

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
МОСКОВСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ИНСТИТУТ
(ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ)

А. К. Платонов, Д. С. Иванов

МЕТОДЫ ОБРАБОТКИ ИЗМЕРЕНИЙ

*Допущено
Учебно-методическим объединением
высших учебных заведений Российской Федерации
по образованию в области прикладных математики и физики
в качестве учебного пособия для студентов вузов
по направлению подготовки «Прикладные математика и физика»*

МОСКВА
МФТИ
2013

УДК 517.983; 519.6; 621.391(075)

ББК 32.811я73

ПЗ7

Рецензенты:

Доктор физико-математических наук, профессор *В. В. Сазонов*

Кандидат физико-математических наук *В. А. Степаньянц*

Платонов, А. К., Иванов, Д. С.

ПЗ7

Методы обработки измерений: учеб. пособие /

А. К. Платонов, Д. С. Иванов. – М. : МФТИ, 2013. – 107 с.

ISBN 978-5-7417-0497-4

Пособие разработано в процессе чтения лекций на кафедре МФТИ "Прикладная математика" по специализации "Управление динамическими системами" и направлено на подготовку студентов-магистров. Цель курса – освоение студентами фундаментальных знаний в области математических свойств навигационных проблем, изучение способов решения задач навигации и методов исследования достижимой навигационной точности и областей их практического применения на поверхности Земли и в космосе. Пособие способствует формированию базовых знаний студентов и аспирантов в области обработки инструментальных данных статистическими и модельно-параметрическими методами на практических примерах определения движения робототехники и космических аппаратов и будет полезно студентам старших курсов и молодым специалистам, интересующимся вопросами определения движения и навигации.

УДК 517.983; 519.6; 621.391(075)

ББК 32.811я73

ISBN 978-5-7417-0497-4

© Платонов А. К., Иванов Д. С. 2013

© Федеральное государственное автономное образовательное учреждение высшего профессионального образования «Московский физико-технический институт (государственный университет)», 2013

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	5
А. СВОЙСТВА НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ	
ВЕРОЯТНОСТЕЙ	7
1. Одномерное нормальное распределение вероятностей	8
2. Двумерное нормальное распределение вероятностей.....	10
Ковариационная матрица	11
Эллипс рассеивания	13
Корреляция	15
3. Геометрия нормального распределения вероятностей.....	18
Преобразования переменных многомерного случайного	
рассеивания	20
Геометрия многомерного нормального распределения	
вероятностей	21
Геометрия процедур модификации случайных векторов.....	23
4. Вероятность присутствия точки внутри области рассеивания	27
Круговое рассеивание.....	28
Вероятность попадания случайной точки внутрь эллипса	
рассеивания	29
Б. МЕТОДЫ СРЕДНЕКВАДРАТИЧНОЙ ФИЛЬТРАЦИИ	33
1. Основные понятия теории оценок.....	33
Максимум правдоподобия оценки	34
2. Простейшие способы осреднения измерений	35
3. Среднеквадратичная фильтрация параметров модели	
измерений.....	38
Способ наименьших квадратов	
(метод максимума правдоподобия).....	40
Учёт априорной информации	42
Метод нормальных мест.....	44
4. Простые среднеквадратичные фильтры измерений.....	47
Аккумулятивный фильтр Гаусса	48
Учёт фазовых шумов	48
Определение состояний процесса в момент наблюдения	49
Аккумулятивная фильтрация состояния процесса с фазовыми	
шумами	52
5. Рекурсивные фильтры Гаусса и Калмана–Бьюси.....	54
Связь рекурсивной процедуры фильтрации с методом	
релаксации	54
Специфика двумерной случайной величины	
"Прогноз–измерение"	56
Рекурсивный фильтр Гаусса	59

Фильтр Калмана–Бьюси	61
Связь рекурсивного фильтра и способа наименьших квадратов	62
Построение линейного приближения нелинейной системы	64
В. ОЦЕНКА ТОЧНОСТИ ЗНАНИЯ ПАРАМЕТРОВ МОДЕЛИ	
ПОСЛЕ ОСРЕДНЕНИЯ ИЗМЕРЕНИЙ	66
1. Формальная оценка	66
2. Расширенная оценка	68
Учёт систематических ошибок погрешностей измерений.	68
Модель наилучшей корреляции погрешностей измерений	75
"Схема бортика"	76
"Мешающие параметры"	77
Параметры μ вне обработки	77
Параметры μ , учитываемые в обработке измерений в качестве	
определяемых параметров	78
Параметры μ в ошибках прогноза изменений параметров модели	79
Г. ВЫЧИСЛИТЕЛЬНЫЕ ПРОБЛЕМЫ ОБРАБОТКИ	
ИЗМЕРЕНИЙ	80
1. Метод квадратных корней	80
2. Метод параболического спуска	82
Д. ПРИМЕРЫ ПРИМЕНЕНИЯ МЕТОДОВ	
СРЕДНЕКВАДРАТИЧНОЙ ФИЛЬТРАЦИИ	90
1. Задача определения ориентации космического аппарата.....	90
2. Оценка смещения ноля магнитометра по полетным данным ..	96
ЛИТЕРАТУРА	100
ПРИЛОЖЕНИЯ	102

ВВЕДЕНИЕ

В учебном пособии рассказывается о способах получения уверенного знания численных значений фазовых параметров экспериментально наблюдаемых движений выбранного объекта слежения или изменений состояний некоторого процесса.

Основным способом получения требуемых знаний является выполнение измерений некоторых сигнальных величин, доступных для инструментального наблюдения параметров объекта или процесса. При этом выполняемые измерения очень часто не дают прямой информации об искомым параметрах наблюдаемого "объекта интереса" и к тому же они всегда содержат инструментальные ошибки.

Отсутствие возможностей прямых измерений нужных параметров вынуждает рассматривать модели их взаимозависимости с получаемыми значениями сигнальных величин. В практике экспериментов это требует использования алгоритмических и программных средств для моделирования устройства измерительной аппаратуры и отображения физики наблюдаемого процесса на измеряемую величину сигнала. При малых уровнях относительных погрешностей используемого оборудования модельная взаимосвязь искомым фазовых и наблюдаемых параметров хорошо выражается в виде системы линейных "условных уравнений", определяющих требуемые вариации фазовых переменных для объяснения наблюдаемых вариаций выполненных измерений.

Там, где недопустимы ошибки принятия решений – в связи с опасностью потери человеческой жизни или дорогого оборудования, – требуется иметь уверенные оценки возможных погрешностей знания искомым параметров. Обстоятельства случайности ошибок измерений диктуют вынужденное применение вероятностных подходов для получения оценки ошибок знания искомым величин. В предположении отсутствия связей между случайными погрешностями выполненных измерений или в случаях, когда такие корреляционные связи хорошо известны, требуемые оценки ошибок определяемых параметров достигаются методами осреднения инструментальных погрешностей измерений.

В более общем случае хорошо известны два типа получаемых вероятностных оценок: "в среднем" и "почти навверное". Ниже большое внимание уделяется рассмотрению методов оценивания именно "почти навверное", в которых "оценка погрешности полученной оценки параметра" (сокращёно – "ОП²") имеет заведомо более надёжное значение, чем дисперсия результата осреднения независимых случайностей.

Необходимость обеспечения надёжного "почти навверного" знания величины "ОП²" привело в своё время к развитию методов оценивания по-

грешностей результатов обработки получаемых измерений с учётом полного числа источников погрешностей оценивания, с необходимостью наряду с учётом взаимосвязи инструментальных погрешностей и имеющихся нелинейностей условных уравнений использовать и методы учёта ошибок моделей движения (так называемых "мешающих параметров").

Достаточно очевидно, что надёжность подобного "почти верного" оценивания зависит от принятых предположений и тоже не абсолютна (шутя, можно поставить вопросы определения и ОПⁿ с $n = 3, 4, \dots$).

В пособии приводится подробное описание имеющихся постановок и способов оценивания искомых величин ОП и ОП² по полученным экспериментальным данным с примерами подобных задач. Значительное внимание уделено описанию содержательной и вычислительной сторон описываемых формализмов оценивания – особенностям обращения матриц коэффициентов нормальных уравнений, геометрическим свойствам линейных фильтров Гаусса и Калмана–Бьюси и проблемам сходимости численных методов решения нелинейных условных уравнений. Следует заметить, что многие из приведенных здесь геометрических и вычислительных свойств инструментального оценивания в подобных курсах обычно не рассматриваются.

Приведённые ниже материалы, направленные на расширение кругозора студента в прикладной области задач обработки экспериментальных данных, должны также на примере этих задач способствовать углублённому пониманию свойств преобразований многомерных случайных величин и вычислительных аспектов их нелинейных преобразований.

Содержание пособия делится на четыре раздела со сквозной нумерацией отдельных фрагментов излагаемого материала.

В первом разделе применительно к задаче оценивания подробно рассматриваются геометрические свойства законов нормального распределения вероятностей случайных величин разной размерности и их линейных преобразований.

Второй раздел посвящён различным свойствам и обстоятельствам применения методов среднеквадратического осреднения совокупности измерений: способа наименьших квадратов, метода максимума правдоподобия, метода нормальных мест и методов линейной фильтрации.

В третьем разделе подробно рассматриваются различные известные методы решения упомянутых ОП²-задач неформального оценивания точности полученного знания.

Четвёртый раздел посвящён вычислительным аспектам решения линейных и нелинейных задач обработки экспериментальных данных.

В приложении приводятся примеры полезных преобразований.

А. СВОЙСТВА НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ

1. Для решения задач навигации и управления нужно выполнять измерения параметров движения, знание которых позволяет определить текущий вектор состояния наблюдаемого объекта. Однако любые измерения обладают погрешностями, что требует от навигатора знаний и умений применения специальных средств обработки измерений. Важным классом таких средств являются рассматриваемые ниже эффективные методы фильтрации погрешностей измерений, основанные на использовании модельного описания параметров наблюдаемости объекта.

Дадим некоторые определения используемых ниже терминов.

Погрешность измерения – отклонение результата измерения от истинного (действительного) значения измеряемой величины.

Инструментальная погрешность обусловлена несовершенством применяемых средств измерений. Причинами её возникновения являются неточности, допущенные при изготовлении и регулировке приборов, изменение параметров элементов конструкции и схемы вследствие старения. В высокочувствительных приборах могут сильно проявляться их внутренние шумы, прежде всего – тепловой и дробовой шумы. Тепловой шум – это равновесный шум, обусловленный тепловым движением носителей заряда в проводнике, в результате чего на концах проводника возникает флуктуирующая разность потенциалов. Дробовой шум – беспорядочные флуктуации напряжений и токов относительно их среднего значения в цепях радиоэлектронных устройств, обусловленные дискретностью носителей электрического заряда – электронов.

Динамическая погрешность измерений – погрешность результата измерений, свойственная измерениям быстро меняющихся величин. Динамическая погрешность появляется при измерении переменных величин и обусловлена инерционными свойствами средств измерений.

Статическая погрешность измерений – погрешность результата измерений, свойственная условиям статического измерения, то есть при измерении постоянных величин после завершения переходных процессов в элементах приборов и преобразователей.

Систематическая погрешность измерения – составляющая общей погрешности, остающаяся постоянной или закономерно изменяющаяся при повторных измерениях одной и той же физической величины. Систематические погрешности являются в общем случае функцией измеряемой величины, влияющих параметров (температуры, влажности, напряжения питания и пр.) и времени.

Случайные погрешности определяются совместным действием ряда причин: внутренними шумами элементов электронных схем, наводками на входные цепи средств измерений, пульсацией постоянного питающего напряжения, а иногда и дискретностью счета.

Ниже будут подробно рассмотрены способы исключения влияния случайных и систематических погрешностей получаемых измерений на результат оценки определяемых параметров.

2. Погрешность измерений представляет собой сумму её составляющих, каждая из которых имеет свою причину. Можно выделить следующие группы причин возникновения погрешностей:

- неверная настройка средства измерений: ошибки калибровки прибора;
- неверная установка объекта измерения на измерительную позицию;
- ошибки в процессе получения, преобразования и выдачи информации в измерительной цепи средства измерений;
- внешние воздействия на средство и/или объект измерений (влияния температуры, давления, электрического и магнитного полей, вибрация и др.);
- свойства измеряемого объекта;
- квалификация и состояние оператора.

Погрешности измерений обусловлены, как правило, неизвестным набором причин, размеры суммарного влияния которых на результат измерения могут быть охарактеризованы неким законом распределения вероятностей его значения. При анализе случайных погрешностей, образуемых большим числом причин, в силу предельной теоремы вероятностей правомерно для анализа принять нормальный закон распределения вероятностей [1–7].

Рассмотрим некоторые важные особенности этого распределения.

1. Одномерное нормальное распределение вероятностей

3. Функция плотности распределения вероятности для одномерного закона Гаусса распределения случайных величин записывается в следующем виде:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}},$$

где x – одномерная случайная величина, m – её математическое ожидание, σ – среднее квадратичное отклонение, $\sigma^2 = D$ – дисперсия. Вероятность принадлежности случайной величины диапазону $\alpha < x < \beta$ вычисляется следующим образом:

$$p(\alpha < x < \beta) = \int_{\alpha}^{\beta} f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{(x-m)^2}{2\sigma^2}} dx.$$

Если нормализовать и центрировать значение рассеивания x , т.е. принять, что $\frac{x-m}{\sigma\sqrt{2}} = t$, то получим

$$p(\alpha < x < \beta) = \frac{1}{\sqrt{\pi}} \int_{\frac{\alpha-m}{\sigma\sqrt{2}}}^{\frac{\beta-m}{\sigma\sqrt{2}}} e^{-t^2} dt.$$

В случае если положить $m = \alpha$, а $z = \frac{\beta-m}{\sigma\sqrt{2}}$, то получим табличный интеграл вероятности (этот интеграл не приводим к квадратурам аналитически):

$$\Phi(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt.$$

Существуют таблицы, в которых $X = (x-m)/\sigma$ измеряется в единицах σ :

$$F(X) = \frac{1}{\sqrt{2\pi}} \int_0^X e^{-\frac{t^2}{2}} dt.$$

4. Величина σ в науке и в технике является единицей ожидаемого или измеренного рассеивания величин (рис. А.1).

Вероятность рассеивания внутри диапазона $\pm\sigma$ равна $p(-\sigma \leq x \leq \sigma) = 0.63$.

Вместе с этим военные специалисты для удобства измеряют рассеивание x в единицах E , где E – такая величина, что вероятность быть внутри $\pm E$ $p(-E \leq x \leq E) = 0.5$. В этом случае всё возможное рассеивание делится на две равные части, что соответствует равным вероятностям "попадания" и "промаха".

Величина E соотносится с величиной σ , как $E = 0.675\sigma$.

Соответственно "доверительные" интервалы отклонений x (при которых вероятность $p(x) > 0.99$) равны или 3σ , или $4E$:

$$p(-3\sigma \leq x \leq 3\sigma) = 0.997 \text{ и } p(-4E \leq x \leq 4E) = 0.993.$$

На этих интервалах построены все теории принятия решений в условиях ситуационной неопределённости.

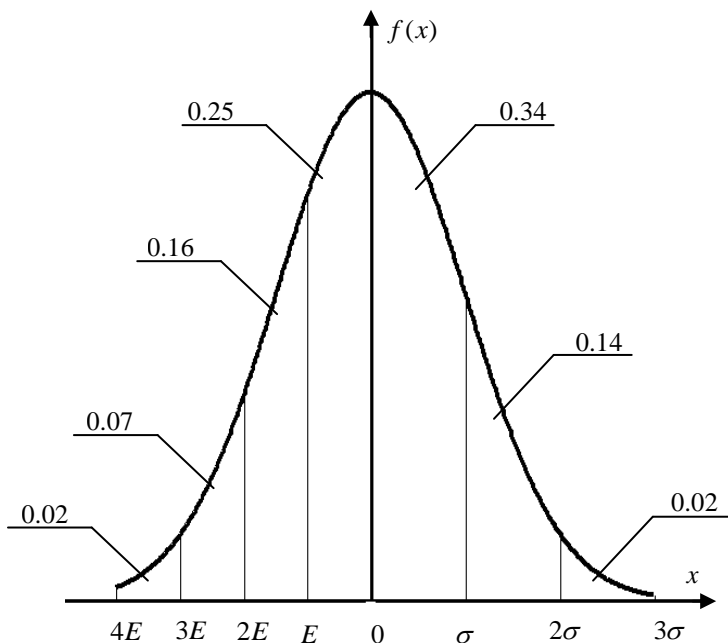


Рис. А.1. Нормальный закон плотности распределения вероятностей случайного одномерного рассеивания и соответствующие вероятности в стандартных интервалах – в единицах E (левая половина графика) и в единицах σ (его правая половина)

2. Двумерное нормальное распределение вероятностей

5. Функция плотности для нормального распределения вероятностей значений X и Y в облаке двумерных случайностей по определению равна (рис. А.2)

$$f(x, y) = \frac{d^2 F(x, y)}{dxdy},$$

где $F(x, y)$ – распределение вероятностей одновременного появления некоторых $x > X$ и $y > Y$.

Если в качестве параметров рассеивания принять *безусловные математические ожидания* m_x, m_y и *безусловные среднеквадратичные отклонения* σ_x, σ_y (т.е. определяемые для каждой из величин x или y при всех возможных значениях другой величины соответственно у или x без

учёта дополнительных условий знания о величинах последних) и дополнительно ввести в рассмотрение некую величину *коэффициента корреляции* $-1 \leq r \leq 1$ (о нём см. ниже), то закон нормального распределения плотности вероятности двумерной случайности нужно записать в следующем виде:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)} \left[\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2r(x-m_x)(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2} \right]},$$

или в более удобном виде векторно-матричного выражения тех же величин:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)} \left\{ \begin{pmatrix} x-m_x \\ y-m_y \end{pmatrix}^T \begin{pmatrix} \frac{1}{\sigma_x^2} & -\frac{r}{\sigma_x\sigma_y} \\ -\frac{r}{\sigma_x\sigma_y} & \frac{1}{\sigma_y^2} \end{pmatrix} \begin{pmatrix} x-m_x \\ y-m_y \end{pmatrix} \right\}},$$

или так:

$$f(x, y) = \frac{1}{2\pi\sqrt{\Delta}} e^{-\frac{1}{2}\mathbf{x}^T K^{-1}\mathbf{x}}. \quad (\text{A.1})$$

Здесь через \mathbf{X} обозначен двумерный вектор центрированных отклонений величин x, y , через K обозначена *ковариационная* матрица с её обратной *весовой* матрицей K^{-1} , входящей в приведённое выше выражение, а $\Delta = \sigma_x^2\sigma_y^2(1-r^2)$.

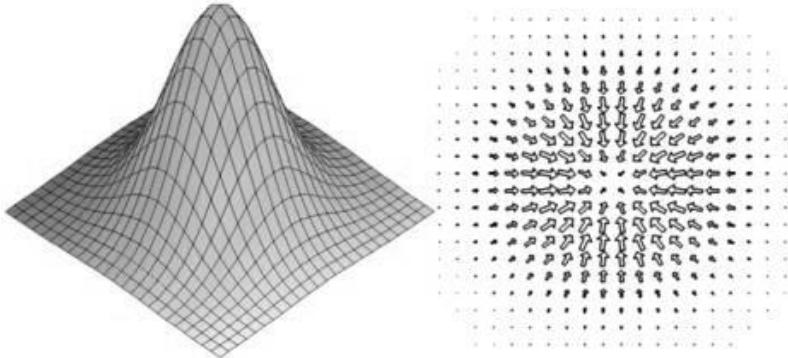


Рис. А.2. Плотность двумерного нормального распределения. Слева – двумерная поверхность "колокола" Гауссиана, справа – её вид сверху

Ковариационная матрица

6. Рассмотрим важные свойства ковариационной матрицы K , тесно связанной с законом нормального распределения вероятностей.

Оказывается, что у функции плотности (А.1) нормального распределения случайных величин (причём любых размерностей!) есть замечательное свойство: эта *функция полностью определена своими первыми двумя моментами*, и это определение содержится именно в матрице ковариаций K .

По аналогии с тензором инерции твёрдого тела ковариационная матрица K представляет собой симметрическую матрицу вторых моментов функции $f(x, y)$ (попробуйте это проверить для двумерного случая центрированной нормальной случайной величины, вычислив первый и второй моменты гауссиана $f(x, y)$). При этом главные моменты функции $f(x, y)$ равны диагональным дисперсиям компонент случайного вектора, а недиагональные элементы матрицы, равные перекрёстному второму моменту k_{xy} , в их выражении через диагональные элементы матрицы однозначно определяют величину и знак получаемого множителя согласования – упомянутого коэффициента корреляции r :

$$K = \begin{pmatrix} k_{xx} & k_{xy} \\ k_{xy} & k_{yy} \end{pmatrix} = \begin{pmatrix} k_{xx} & r \cdot \sqrt{k_{xx}} \sqrt{k_{yy}} \\ r \cdot \sqrt{k_{xx}} \sqrt{k_{yy}} & k_{yy} \end{pmatrix} = \begin{pmatrix} \sigma_x^2 & \sigma_x \sigma_y r \\ \sigma_x \sigma_y r & \sigma_y^2 \end{pmatrix}.$$

7. Определитель ковариационной матрицы равен $\Delta = \sigma_x^2 \sigma_y^2 (1 - r^2)$, отсюда легко получить выражение для обратной матрицы квадратичной формы, стоящей в показателе экспоненты нормального закона плотности вероятности центрированных случайных величин x и y (почему эту матрицу называют *матрицей весов ошибок*, станет ясно ниже):

$$K^{-1} = \frac{1}{\Delta} \begin{pmatrix} \sigma_y^2 & -r \sigma_x \sigma_y \\ -r \sigma_x \sigma_y & \sigma_x^2 \end{pmatrix} = \frac{1}{1 - r^2} \begin{pmatrix} \frac{1}{\sigma_x^2} & -\frac{r}{\sigma_x \sigma_y} \\ -\frac{r}{\sigma_x \sigma_y} & \frac{1}{\sigma_y^2} \end{pmatrix}.$$

8. При $0 \leq r \leq 1$ выражение $\mathbf{X}^T K^{-1} \mathbf{X} = \lambda^2 = \text{const} > 0$ есть положительно определённая квадратичная форма для векторов \mathbf{X} . Поэтому в соответствии с теорией кривых второго порядка уравнение постоянства функции $f(x, y)$

$$\frac{(x - m_x)^2}{\sigma_x^2 (1 - r^2)} - \frac{2r(x - m_x)(y - m_y)}{\sigma_x \sigma_y (1 - r^2)} + \frac{(y - m_y)^2}{\sigma_y^2 (1 - r^2)} = \lambda^2, \quad (\text{A.2})$$

относительно компонент вектора $\mathbf{X}(x, y)$ представляет собой уравнение эллипса (эллипса равной плотности вероятности* или эллипса рассеивания, определяющего минимальную площадь рассеивания \mathbf{X} с данной вероятностью $p(\lambda)$ – см. рис. А.3).

Эллипс рассеивания

9. В общем случае (при $r \neq 0$) главные оси этого эллипса не совпадают с осями системы координат Oxy . Центр эллипса лежит в точке (m_x, m_y) , а направление его осей симметрии в системе координат Oxy определяется из выражения:

$$\operatorname{tg} 2\alpha = \frac{2r\sigma_x\sigma_y}{\sigma_x^2 - \sigma_y^2}. \quad (\text{A.3})$$

Получить эту зависимость можно путём поворота осей системы координат на некоторый угол α по формулам

$$\begin{cases} x' = x \cos \alpha + y \sin \alpha, \\ y' = -x \sin \alpha + y \cos \alpha \end{cases} \quad (\text{A.4})$$

с условием равенства нулю второго слагаемого в уравнении (А.2).

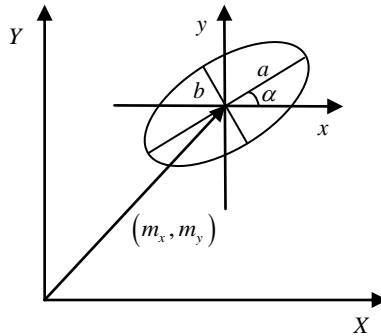


Рис. А.3. Эллипс рассеивания

Выражение (А.3) определяет два значения угла α , отличающихся на угол $\pi/2$. Можно заметить, что если $\sigma_x = \sigma_y$, то при любом r главные оси соответствующей окружности наклонены к оси x на 45° и 135° .

Если элементы весовой матрицы K^{-1} обозначить в виде

* Ниже будет показана роль нормирующего параметра λ , влияющего на обобщённые размеры эллипса.

$$K^{-1} = \begin{bmatrix} A & -C \\ -C & B \end{bmatrix}$$

(с уравнением эллипса вида $Ax^2 - 2Cxy + By^2 = \lambda^2$), то размеры его полуосей определяются из выражений

$$a^{-2} = \frac{A+B - \sqrt{(A-B)^2 + 4C^2}}{2}, \quad b^{-2} = \frac{A+B + \sqrt{(A-B)^2 + 4C^2}}{2}.$$

Уравнение эллипса при постоянстве функции $f(x, y)$ можно записать и иначе:

$$\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2r(x-m_x)(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2} = \lambda^2(1-r^2) = \chi^2.$$

При этом размеры эллипса на плоскости Oxy не изменятся, если в качестве *параметра размера эллипса* использовать не λ , а соответствующее $\chi(\lambda, r)$.

10. Приведём эллипс к главным осям, повернув систему координат по формулам (A.4) на угол α из (A.3) и сдвинув её на m_x и m_y . Тогда *главные средние квадратические отклонения* σ_1 и σ_2 выразятся через *средние квадратические отклонения* σ_x и σ_y следующим образом (проверьте это):

$$\begin{cases} \sigma_1^2 = \sigma_x^2 \cos^2 \alpha + r\sigma_x\sigma_y \sin 2\alpha + \sigma_y^2 \sin^2 \alpha, \\ \sigma_2^2 = \sigma_x^2 \sin^2 \alpha - r\sigma_x\sigma_y \sin 2\alpha + \sigma_y^2 \cos^2 \alpha. \end{cases}$$

Напомним, что след ковариационной матрицы является инвариантом преобразования поворота системы координат. Это обстоятельство определяет связь дисперсий одного и того же эллипса рассеивания в разных системах координат с собственными значениями σ_1^2 , σ_2^2 ковариационной матрицы:

$$\sigma_1^2 + \sigma_2^2 = \sigma_x^2 + \sigma_y^2.$$

В главных осях эллипса коэффициент корреляции $r = 0$ и плотность нормального закона распределения вероятности в этих осях имеет вид

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2} \cdot e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2}\right)},$$

а уравнения эллипсов вероятности имеют вид ($\lambda = \chi$)

$$\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} = \lambda^2,$$

или

$$\frac{x^2}{(\lambda\sigma_1)^2} + \frac{y^2}{(\lambda\sigma_2)^2} = 1.$$

Последнее выражение показывает, что размеры главных среднеквадратичных отклонений эллипсов равных значений плотности вероятности для двумерного закона распределения вероятностей пропорциональны размерам их главных полуосей с коэффициентом пропорциональности, равным параметру размера λ . Это значение определяет высоту соответствующего эллипса, получаемого горизонтальным сечением "колокола" гауссиана. В качестве стандартного размера гауссиана рассеивания двумерной случайной величины ниже будет рассматриваться эллипс при $\lambda = 1$ с его главными среднеквадратичными отклонениями σ_1^2 , σ_2^2 и соответствующими им безусловными среднеквадратичными отклонениями σ_x^2 , σ_y^2 .

Корреляция

11. Рассмотрим свойства корреляции рассеивания координат точек при нормальном двумерном распределении вероятностей. Для этого напишем выражение для закона плотности условного распределения вероятности появления величины y при известном x и аналогично – для условной плотности закона распределения вероятностей x при известном y .

Так как для функции плотности распределения полной вероятности справедливы её выражения через плотности условных вероятностей и плотности вероятностей условий в виде

$$f(x, y) = f(y|x) \cdot f_1(x) = f(x|y) \cdot f_2(y),$$

где $f(y|x)$ – плотность распределения вероятности величины y при известном x , $f(x|y)$ – плотность распределения вероятности величины x при известном y , $f_1(x)$ и $f_2(y)$ – одномерные функции плотности распределения вероятности реализации величины условия, то

$$f(y|x) = \frac{f(y, x)}{f_1(x)} = \frac{\sigma_x \sqrt{2\pi}}{\sigma_x \sigma_y 2\pi \sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)} \left[\frac{(x-m_x)^2}{\sigma_x^2} - \frac{r(x-m_x)(y-m_y)}{\sigma_x \sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2} \right] + \frac{(x-m_x)^2}{2\sigma_x^2}}.$$

Отсюда получаем

$$f(y|x) = \frac{1}{\sigma_y \sqrt{2\pi} \sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)} \left[-r \frac{(x-m_x)}{\sigma_x} + \frac{(y-m_y)}{\sigma_y} \right]^2}$$

и

$$f(x | y) = \frac{1}{\sigma_x \sqrt{2\pi} \sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)} \left[\frac{(x-m_x)}{\sigma_x} - r \frac{(y-m_y)}{\sigma_y} \right]^2}.$$

Преобразуем показатель степени экспоненты к следующему виду:

$$f(y | x) = \frac{1}{\sigma_y \sqrt{1-r^2} \cdot \sqrt{2\pi}} e^{-\frac{\left[y - (m_y + r \frac{\sigma_y}{\sigma_x} (x - m_x)) \right]^2}{2\sigma_y^2 (1-r^2)}}.$$

12. Мы получили одномерный нормальный закон с условной дисперсией:

$$\sigma_{y|x}^2 = \sigma_y^2 (1-r^2) \text{ и } \sigma_{y|x} = \sigma_y \sqrt{1-r^2}$$

и выражение для математического ожидания:

$$m_{y|x} = m_y + r \frac{\sigma_y}{\sigma_x} (x - m_x).$$

Аналогично для условного рассеивания величины x при условии, что y известно, имеем

$$\sigma_{x|y}^2 = \sigma_x^2 (1-r^2) \text{ и } \sigma_{x|y} = \sigma_x \sqrt{1-r^2},$$

$$m_{x|y} = m_x + r \frac{\sigma_x}{\sigma_y} (y - m_y).$$

13. Эти выражения показывают очень важное свойство нормального закона распределения вероятностей.

Дисперсия условной вероятности не зависит от условия! Это свойство нормального распределения ошибок позволяет оценивать ожидаемые точности результатов эксперимента, *не дожидаясь выполнения эксперимента.*

14. Однако условные математические ожидания зависят от условия, и эти зависимости – линейные. Им соответствуют прямолинейные геометрические места центров условного рассеивания величин $y|x$ и $x|y$ при изменении значений x и y соответственно. Эти прямые называются *линиями регрессии* (положительными или отрицательными – в зависимости от знака коэффициента корреляции r). Их уравнения выглядят следующим образом:

$$\begin{cases} x = m_x + r \frac{\sigma_x}{\sigma_y} (y - m_y), \\ y = m_y + r \frac{\sigma_y}{\sigma_x} (x - m_x), \end{cases}$$

или после преобразований:

$$\begin{cases} \frac{x-m_x}{\sigma_x} = r \frac{y-m_y}{\sigma_y}, \\ \frac{y-m_y}{\sigma_y} = r \frac{x-m_x}{\sigma_x}. \end{cases}$$

Как найти эти линии регрессии на эллипсе?

Не трудно увидеть, что геометрически *линии регрессии должны быть диаметрами эллипса, сопряжёнными** его "координатным" диаметрам (т.е. диаметрам, параллельным осям координат x и y), поскольку условные математические ожидания делят пополам хорды эллипса, параллельные осям x и y .

15. Здесь пришло время заметить, что в знаменателе уравнения эллипса у показателя экспоненты в функции плотности нормального распределения вероятностей стоят именно *условные дисперсии*:

$$\frac{(x-m_x)^2}{\sigma_x^2(1-r^2)} = \sigma_{x|y}^2 - \frac{2r(x-m_x)(y-m_y)}{\sigma_x\sigma_y(1-r^2)} + \frac{(y-m_y)^2}{\sigma_y^2(1-r^2)} = \lambda^2,$$

и поэтому плотность нормального закона распределения вероятностей компонент x , y двумерного вектора проще записывать (и запоминать) в следующем виде:

$$f(x, y) = \frac{1}{2\pi\sigma_{x|y}\sigma_y} e^{-\frac{1}{2} \left[\frac{(x-m_x)^2}{\sigma_{x|y}^2} - \frac{2r(x-m_x)(y-m_y)}{\sigma_{x|y}\sigma_{y|x}} + \frac{(y-m_y)^2}{\sigma_{y|x}^2} \right]}. \quad (\text{A.5})$$

16. При этом по аналогии с полной вероятностью для двумерного распределения вероятностей можно определить *полную дисперсию*, равную определителю ковариационной матрицы:

$$\sigma_{x|y}^2\sigma_y^2 = \sigma_x^2\sigma_{y|x}^2 = \Delta = \sigma_x^2\sigma_y^2(1-r^2)$$

и аналогично *полное среднеквадратичное отклонение*:

$$\sigma_{x|y}\sigma_y = \sigma_x\sigma_{y|x} = \sigma_x\sigma_y\sqrt{1-r^2}.$$

Заметим, что наряду с ковариационной матрицей K часто используется *корреляционная матрица* $R = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}$.

Вопрос. Как построить преобразования матриц $K \rightarrow R$ и $R \rightarrow K$?

* Напомним, что в эллипсе диаметр, сопряжённый данному диаметру, делит пополам хорды эллипса, параллельные исходному диаметру.

3. Геометрия нормального распределения вероятностей

17. Покажем, что эллипсы вероятности при любом значении корреляции r подобны эллипсу, вписанному в прямоугольник со сторонами, параллельными осям выбранной системы координат, а их размеры равны σ_x и σ_y . Покажем также, что точки пересечения линий регрессии со сторонами такого прямоугольника являются именно теми точками, где эллипс касается прямоугольника. Геометрически эти факты становятся достаточно очевидными, если рассмотреть хорды из определения сопряжённых диаметров эллипса и их предел при движении точки вдоль линии регрессии к граничной точке на эллипсе.

Для аналитического доказательства рассмотрим точки (x, y) , которые удовлетворяют следующим условиям:

$$\begin{cases} x - m_x = \pm \sigma_x, \\ y - m_y = \pm r \sigma_y, \end{cases}$$

и

$$\begin{cases} x - m_x = \pm r \sigma_x, \\ y - m_y = \pm \sigma_y. \end{cases}$$

Эти точки принадлежат как линиям регрессии, так и эллипсу равной плотности вероятности с нормированным параметром размера $\lambda = 1$:

$$\frac{(x - m_x)^2}{\sigma_x^2(1 - r^2)} - \frac{2r(x - m_x)(y - m_y)}{\sigma_x \sigma_y(1 - r^2)} + \frac{(y - m_y)^2}{\sigma_y^2(1 - r^2)} = 1.$$

Теперь покажем, что x_{\max} лежит на линии регрессии $m_{y/x}$, то есть на следующей прямой:

$$\frac{y}{x} = r \frac{\sigma_y}{\sigma_x}.$$

Действительно, из системы уравнений, определяющих экстремумы эллипса

$$\begin{cases} \frac{\partial}{\partial x} (Ax^2 + By^2 - 2Cxy - 1) = 0, \\ \frac{\partial}{\partial y} (Ax^2 + By^2 - 2Cxy - 1) = 0 \end{cases}$$

следует, что

$$\frac{x}{y} = \frac{C}{B} = r \frac{\sigma_y}{\sigma_x}.$$

Иными словами, выполняются необходимые условия того, что линия регрессии $m_{y/x}$ пересекает эллипс в точках, наиболее удалённых от оси y (при значениях $x = \pm\sigma_x$ и $y = \pm r\sigma_y$ соответственно).

Аналогично, для определения координат точки y_{\max} имеем условие

$$\frac{y}{x} = \frac{A}{C} = \frac{\sigma_y}{r\sigma_x},$$

то есть условие того, что линия регрессии $m_{x/y}$ пересекает эллипс в точках, наиболее удалённых от оси x (при значениях $y = \pm\sigma_y$ и $x = \pm r\sigma_x$ соответственно). И отсюда следует (докажите это очевидное утверждение сами), что касательные к эллипсу в таких точках параллельны координатным осям, и поэтому эллипс "σ-рассеивания" оказывается всегда вписанным в прямоугольник со сторонами, параллельными осям координат, с их размерами, равными $\pm\sigma_x$ и $\pm\sigma_y$ ("координатный прямоугольник эллипса рассеивания", рис. А.4).

18. Важно обратить внимание на геометрический смысл коэффициента корреляции r : в координатном прямоугольнике *точки касания вписанного в него эллипса находятся на одинаковых относительных расстояниях от осей координат, равных именно r .*

Становится понятным, почему (как было сказано) при $\sigma_x = \sigma_y$ эллипс всегда одинаково наклонен: точки касания квадрата $r\sigma_x = r\sigma_y$ симметричны. Параметр корреляции r здесь отражает лишь ширину вписанного в квадрат эллипса с ростом r , линии регрессии всё дальше отстоят от полуосей эллипса, и менее полным становится эллипс.

При $r=1$ эллипс превращается в диагональ координатного прямоугольника.

19. Размеры хорд эллипса, параллельные осям y и x (линии регрессии делят их пополам), показывают относительно σ_y и σ_x размеры одномерного условного рассеивания y/x и x/y соответственно. Вертикальные сечения колокола гауссиана, проходящие через эти хорды эллипса, имеют формы одномерного закона плотности распределения вероятностей, но их уменьшающиеся

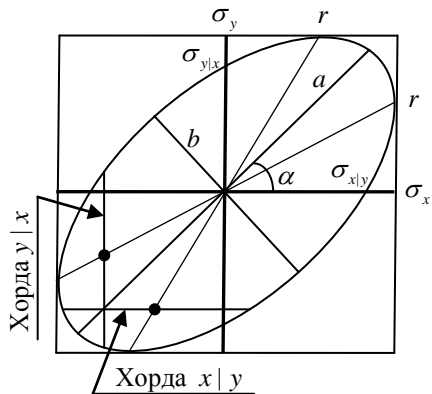


Рис. А.4. Координатный прямоугольник

размеры не удовлетворяют интегральным условиям нормирования этих кривых

$$P(x < t) = \int_{-\infty}^{\infty} f(t) dt = 1.$$

Не трудно доказать (попробуйте), что *центральные хорды* эллипса с параметром $\lambda = 1$ при любом r этим условиям удовлетворяют. И теперь $\sigma_{y|x}$ и $\sigma_{x|y}$ приобретают ясный геометрический смысл: им соответствуют точки пересечения эллипса с центральными координатными осями (рис. А.4).

Упражнение 1. Найдите на рис. А.4 прямоугольные треугольники из следующих очевидных утверждений:

$$\begin{cases} (r\sigma_y)^2 + \sigma_{y|x}^2 = \sigma_y^2, \\ (r\sigma_x)^2 + \sigma_{x|y}^2 = \sigma_x^2. \end{cases}$$

20. Это упражнение отражает одно из полезных свойств геометрии эллипса, позволяющее, зная σ_x и σ_y и r , без вычислений (с помощью циркуля и линейки) построить 8 точек эллипса в его координатном прямоугольнике. Если при этом известны элементы ковариационной матрицы, то можно вычислить и определить дополнительно ещё 4 важных точки на эллипсе, соответствующие большой и малой его полуосям, после чего опять с помощью циркуля и линейки легко построить и точки фокусов эллипса (опишите эту процедуру).

21. Отметим важное свойство геометрии нормального распределения вероятностей: из того, что эллипс вероятности с параметром $\lambda = 1$ вписан в x, y прямоугольник, вытекает следующее утверждение: *безусловные среднеквадратичные отклонения $\pm\sigma_x$ и $\pm\sigma_y$ являются проекциями эллипса на оси системы координат.*

Это обстоятельство нам понадобится ниже: оно по индукции упрощает понимание того, что происходит с распределением вероятностей многомерных векторных случайных величин при изменениях координатного базиса, определяемым линейными преобразованиями рассматриваемой случайной величины.

Преобразования переменных многомерного случайного рассеивания

22. В общем случае нормальный закон плотности распределения вероятности многомерного случайного вектора \mathbf{X} с ковариационной матрицей K выводится по индукции из свойств одномерного и двумерного рассеиваний ($\Delta = |K|$):

$$f(\mathbf{X}) = \frac{1}{\sqrt{(2\pi)^n \Delta}} e^{-\frac{1}{2} \mathbf{x}^T \mathbf{K}^{-1} \mathbf{x}}.$$

23. Линейное преобразование случайного вектора \mathbf{X} с невырожденной квадратной матрицей преобразования W образует новый случайный вектор $\mathbf{Y} = W\mathbf{X}$ с другим законом распределения вероятностей. Аналитически рассеивание вектора \mathbf{Y} описывается преобразованием квадратичной формы в показателе экспоненты для функции плотности распределения вероятностей рассеивания \mathbf{X} .

Поскольку матрица W имеет обратную матрицу U , то можно подставить в выражение квадратичной формы значение \mathbf{X} в виде $\mathbf{X} = U\mathbf{Y}$:

$$\mathbf{X}^T \mathbf{K}_X^{-1} \mathbf{X} = \mathbf{Y}^T U^T \mathbf{K}_X^{-1} U \mathbf{Y} = \mathbf{Y}^T \mathbf{K}_Y^{-1} \mathbf{Y}.$$

Таким образом, *весовая* матрица для случайного вектора \mathbf{Y} определяется из*

$$\mathbf{K}_Y^{-1} = U^T \mathbf{K}_X^{-1} U.$$

И тогда ковариационная матрица вектора \mathbf{Y} получается из (случай использования вырожденного преобразования $\mathbf{Y} = W\mathbf{X}$ будет подробно рассмотрен ниже)

$$\mathbf{K}_Y = U^{-1} \mathbf{K}_X U^{-1T} = W \mathbf{K}_X W^T.$$

Здесь следует запомнить, где именно стоят знаки транспонирования для каждой из преобразуемых матриц. И постараться (например, исходя из размерности величин) ответить на следующий вопрос.

Вопрос. Почему в преобразованиях из пространства \mathbf{X} в пространство \mathbf{Y} для весовых матриц $\mathbf{K}_X^{-1} \rightarrow \mathbf{K}_Y^{-1}$ используется обратная зависимость старой переменной \mathbf{X} от новой \mathbf{Y} ($U: \mathbf{Y} \rightarrow \mathbf{X}$), а для преобразования ковариационных матриц $\mathbf{K}_X \rightarrow \mathbf{K}_Y$ нужна прямая связь новой \mathbf{Y} от старой \mathbf{X} ($W: \mathbf{X} \rightarrow \mathbf{Y}$)?

Геометрия многомерного нормального распределения вероятностей

24. Важным свойством линейного преобразования многомерной случайной величины является *сохранение после преобразования положительной определённости и симметрии новой матрицы квадратичной формы с присущим ей свойством ортогональности собственных векторов, определяющих положения главных осей эллипсов вероятности.*

* Для студента, вникающего в связи наук, заметим, что все матрицы квадратичных форм принадлежат к подклассу *конгруэнтности* $A = C^T B C$ из общего класса эквивалентности *подобных* матриц, определяемых, в свою очередь, соотношением $A = C^{-1} B C$ – с невырожденной матрицей C .

Именно эти факты и определяют при линейных преобразованиях случайных переменных с нормальным распределением вероятностей факт сохранения их свойств. При этом сохраняются и геометрические свойства функции Гаусса плотности распределения вероятностей.

Иными словами, при линейных преобразованиях многомерных случайных векторов эллипсоиды их рассеивания превращаются в другие эллипсоиды. Использование геометрически свойств эллипсоидов позволяет довольно просто определять нужные способы требуемых преобразований вероятностей, и обратно – известные свойства нормального распределения вероятностей позволяют уточнять геометрические свойства преобразуемых эллипсоидов.

25. По аналогии с двумерным случаем легко представить, что в многомерном декартово-координатном пространстве эллипсоиды равной плотности нормального закона распределения вероятностей вписаны в многомерный координатный параллелепипед, образуемый гиперплоскостями, ортогональными осям координат. Размеры этого прямоугольного координатного "ящика" вдоль каждой из осей координат пропорциональны размерам одномерных среднеквадратичных отклонений $\pm\sigma_i$ ($i = 1, \dots, n$, где n – размерность "облака" случайного вектора) наибольшего рассеивания вдоль каждой из координатных осей. Этот размер безусловного σ -отклонения от центра эллипсоида вдоль выбранной координатной оси соответствует предположению о реализации любых возможных значений отклонений остальных компонент случайного вектора, ограниченных вероятностью быть внутри эллипсоида с размерами, соответствующими выбранному значению плотности вероятности.

26. В свою очередь соответствующие значения величин безусловных дисперсий одномерных вариаций координатных компонент многомерного случайного вектора равны квадратам величин *проекций* исходного многомерного эллипсоида с параметром $\lambda = 1$ на орты этих координатных направлений (аналогично двумерному рассеиванию см. выше рис. А.4). Но поскольку выбор направления осей системы координат, в которой описывается многомерное случайное "облако", произволен, то *правило "безусловная дисперсия = квадрат проекции эллипсоида $\lambda = 1$ " определяет σ -рассеивание многомерной случайной величины в любом выбранном направлении этого многомерного пространства.* Отсюда следует простое геометрическое правило:

Для определения величины дисперсии выбранной скалярной функции $h(r)$ многомерного случайного вектора L с ковариационной матрицей K_L нужно спроектировать эллипсоид K_L на направление вектора-строки $w = \text{grad}(h(r))$ и умножить полученную величину на $|w|$.

Аналитически это означает, что $\sigma_h^2 = wK_r w^T$.

27. В свою очередь геометрически очевидно следующее.

- Проекция эллипсоида рассеивания многомерной случайной величины на координатные плоскости образуют двумерные прямоугольные ящики с их границами, определяющими значения безусловного σ -рассеивания двумерного распределения вероятностей соответствующих координат.

- Проекция эллипсоида рассеивания многомерной случайной величины в трёхмерном координатном пространстве образуют трёхмерные эллипсоиды нормального распределения вероятностей рассеивания выбранных троек координат. И так далее по индукции.

- После любого линейного преобразования в виде $\mathbf{X} = \mathbf{U}\mathbf{Y}$ в новом координатном базисе, зависящем от градиентов преобразования (векторов строк матрицы \mathbf{U}), сохраняется свойство ортогональности нового многомерного координатного ящика для нового эллипсоида рассеивания.

- При этом в общем случае

- эллипсоид $\lambda = 1$ касается гиперплоскостей координатного ящика точками линий регрессии $m_{k|(i=1, \dots, n, i \neq k)}$.

- эллипсоид $\lambda = 1$ отсекает на осях координат отрезки, равные условным среднеквадратичным отклонениям $\sigma_{k|(i=1, \dots, n, i \neq k)}$ при условии, что известны остальные координаты случайного вектора (напомним, что это следует из аналога формулы (А.5), где в знаменателях уравнения эллипсоида стоят именно условные дисперсии, величина которых для каждой координаты определена равенством нулю всех остальных координат).

28. Отсюда аналитически следует, что любое "менее мерное" распределение вероятностей компонент случайного вектора описывается выбираемым подмножеством строк и столбцов исходной многомерной весовой матрицы \mathbf{K}^{-1} и соответственных строк и столбцов ковариационной матрицы \mathbf{K} .

Задания.

1. Напишите уравнение линии регрессии для трёхмерного случая.

2. Определите геометрический образ и аналитическое выражение для коэффициента корреляции r_{ij} (например, r_{11}).

Геометрия процедур модификации случайных векторов

29. Из вышесказанного ясно, что в простейшем случае поворота декартовых осей многомерной системы координат параметры нового закона распределения вероятностей соответствуют простому проектированию

старого эллипсоида вероятности на новые (повёрнутые) оси и/или их координатные гиперплоскости и пространства меньшей размерности.

Этот случай чаще всего реализуется при приведении параметров ковариационной матрицы к диагональному виду, а размеров координатного ящика – к главным осям эллипсоида. Тогда процедура проектирования совпадает с определением размеров главных осей эллипсоида, вектор математического ожидания не изменяется, коэффициенты корреляции обнуляются, величины безусловных и условных дисперсий совпадают.

30. Важен и другой случай – более экзотический, но возможный. Пусть все коэффициенты корреляции многомерного закона рассеивания $r_{ij} = 1$, в то время как диагональ ковариационной матрицы имеет не нулевые элементы. Что это означает? Геометрически это соответствует вырождению многомерного эллипсоида в отрезок диагонали его "ящика", над которой располагается вырожденная "верхушка колокола" в виде плоского одномерного гауссиана (рис. А.5).

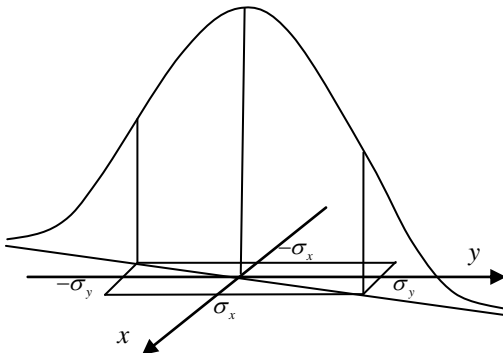


Рис. А.5. Вырожденное двухмерное рассеивание

Иными словами, в таком случае с самого начала в вероятностном пространстве существовала только одномерная случайная величина (в виде её плоского гауссиана), описанная многомерным случайным вектором с его проекциями на оси неудачно выбранной системы координат.

Выполнение поворота системы координат до совпадения оси x с направлением рассеивания превращает в нули все элементы ковариационной и весовой матриц кроме верхнего диагонального члена.

31. Пример. Рассмотрим ковариационную матрицу трёхмерной случайной величины с единичными безусловными дисперсиями и полной корреляцией координатных отклонений:

$$K = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Для определения величин и направлений главных дисперсий такого рассеивания требуется определение корней характеристического полино-

ма матрицы K , собственных чисел этой матрицы и собственных векторов в исходном координатном базисе (сделайте это!).

Геометрически, однако, очевидно, что такой матрице соответствует одномерное распределение, эллипс рассеивания которого представляет собой отрезок диагонали кубического ящика со стороны, равной единице. Геометрически очевиден также и требуемый поворот для приведения ковариационной матрицы к диагональному виду:

$$K = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Этот пример показывает, как в ряде случаев геометрические соображения позволяют достаточно просто определять не только параметры рассеивания, но и собственные числа ковариационной матрицы (в данном случае они равны 3 и 0) и единственный собственный вектор матрицы K (очевидно, что в данном случае он равен $[1 \ 1 \ 1]^T$).

Подтвердите это аналитически и докажите существование здесь лишь одного собственного вектора (подсказка: следует вспомнить свойства собственных векторов матрицы и найти здесь их нарушение).

32. В общем случае при линейном преобразовании центрированной функции плотности нормального распределения вероятности одновременно изменяются как *ориентация*, так и *размеры отклонений* вдоль осей базиса многомерного случайного вектора. Известно [3, с. 132], что матрица W такого ("центроаффинного") преобразования $\mathbf{Y} = W\mathbf{X}$ всегда может быть разложена на композицию ортогонального (поворот) и положительного самосопряжённого (растяжение) операторов.

Вопрос. Коммутативны ли операции координатного поворота и масштабирования закона распределения вероятностей случайного вектора?

В результате происходят поворот и растяжения пространства случайных векторов и соответственно ящика многомерного эллипсоида рассеивания.

33. Очень часто процедурой линейного преобразования приходится пользоваться для определения величины рассеивания *одномерной функции многомерного случайного вектора*. Но в таком одномерном случае существует проблема: нужно определить рассеивание вектора $\mathbf{Y} = W\mathbf{X}$ с использованием матрицы W , которая в случае скалярной искомой величины не обратима, поскольку представляет собой лишь вектор-строку. Иными словами, возникает вопрос: *как применять правила преобразования квадратичной формы, если ищутся параметры случайного вектора или скаляра \mathbf{Y} меньшей размерности, чем размерность порожд-*

дающего его вектора \mathbf{X}^* (т.е. матрица W имеет число строк меньше, чем число столбцов)?

34. Ответ заключается в некотором "трюке", который обычно не оговаривается. По умолчанию якобы выполняются следующие операции.

1. Строится ортогональное дополнение прямоугольной матрицы преобразования до квадратной матрицы полной размерности исходного пространства:

$$\tilde{W} = \begin{bmatrix} W & 0 \\ 0 & W_{\perp} \end{bmatrix},$$

здесь \tilde{W} – квадратная матрица и существует $U = \tilde{W}^{-1}$.

2. Выполнение преобразования квадратичной формы для весовой матрицы K_X^{-1} (матрицы коэффициентов уравнения исходного эллипсоида):

$$\tilde{K}_Y^{-1} = U^T K_X^{-1} U.$$

3. Вычисляется ковариационная матрица фиктивного вектора \mathbf{Y} путём обращения весовой матрицы

$$\left[\tilde{K}_Y^{-1} \right]^{-1} = \tilde{K}_Y = \begin{bmatrix} K_Y & C \\ C & K_{\perp} \end{bmatrix}.$$

4. Из матрицы \tilde{K}_Y вычёркиваются все строки и столбцы, кроме клетки K_Y , являющейся искомой проекцией преобразованного эллипсоида \tilde{K}_Y на новые оси координат пространства \mathbf{Y} , повернутые в соответствии с матрицей W :

$$K_Y = \begin{bmatrix} K_Y & \text{---} \\ \text{---} & \text{---} \end{bmatrix}.$$

Часто формальной записью этой пошаговой процедуры, которая только предполагается, но не упоминается, является такое выражение:

$$K_Y = [U^T K_X^{-1} U]^{-1} = W K_X W^T.$$

35. Пример. Определим вариацию величины интеграла энергии h в кеплеровом приближении, получаемую в начале траектории полёта к Луне в результате ошибок управления полётом ракеты на активном участке траектории. Ошибки вектора $\mathbf{L} = [x \ y \ z \ V_x \ V_y \ V_z]^T$ декартовых координат и компонент скорости полёта в момент выключения двигателя по-

* Не менее частый случай, когда размерность вектора Y много больше размерности вектора X подробно рассматривается ниже.

следней ступени ракеты описываются ковариационной матрицей шестимерного эллипсоида рассеивания:

$$K_0 = \begin{pmatrix} \sigma_x^2 & r_{xy}\sigma_x\sigma_y & r_{xz}\sigma_x\sigma_z & r_{xV_x}\sigma_x\sigma_{V_x} & r_{xV_y}\sigma_x\sigma_{V_y} & r_{xV_z}\sigma_x\sigma_{V_z} \\ r_{xy}\sigma_x\sigma_y & \sigma_y^2 & r_{yz}\sigma_z\sigma_y & r_{yV_x}\sigma_{V_x}\sigma_y & r_{yV_y}\sigma_{V_y}\sigma_y & r_{yV_z}\sigma_{V_z}\sigma_y \\ r_{xz}\sigma_x\sigma_z & r_{yz}\sigma_z\sigma_y & \sigma_z^2 & r_{zV_x}\sigma_{V_x}\sigma_z & r_{zV_y}\sigma_{V_y}\sigma_z & r_{zV_z}\sigma_{V_z}\sigma_z \\ r_{xV_x}\sigma_x\sigma_{V_x} & r_{yV_x}\sigma_{V_x}\sigma_y & r_{zV_x}\sigma_{V_x}\sigma_z & \sigma_{V_x}^2 & r_{V_xV_y}\sigma_{V_x}\sigma_{V_y} & r_{V_xV_z}\sigma_{V_x}\sigma_{V_z} \\ r_{xV_y}\sigma_x\sigma_{V_y} & r_{yV_y}\sigma_{V_y}\sigma_y & r_{zV_y}\sigma_{V_y}\sigma_z & r_{V_xV_y}\sigma_{V_x}\sigma_{V_y} & \sigma_{V_y}^2 & r_{V_zV_y}\sigma_{V_z}\sigma_{V_y} \\ r_{xV_z}\sigma_x\sigma_{V_z} & r_{yV_z}\sigma_{V_z}\sigma_y & r_{zV_z}\sigma_{V_z}\sigma_z & r_{V_xV_z}\sigma_{V_x}\sigma_{V_z} & r_{V_zV_y}\sigma_{V_z}\sigma_{V_y} & \sigma_{V_z}^2 \end{pmatrix}.$$

Случайность величины искомой скалярной величины интеграла энергии h определена её зависимостью от случайной величины шестимерного вектора \mathbf{L} .

Геометрически процедура ответа заключается в проектировании эллипсоида рассеивания $\mathbf{L}^T K_0^{-1} \mathbf{L} = 1$ исходного случайного вектора параметров траектории полёта \mathbf{L} на направление вектора градиента функции интеграла энергии этих параметров $h = V^2 / 2 - \mu / r$ в исходном многомерном пространстве с последующим выполнением требуемых растяжений (умножений на модуль величины градиента). Аналитически эта операция заключается в определении σ_h^2 .

Вектор \mathbf{u} градиента интеграла энергии h в пространстве \mathbf{L} имеет вид

$$\mathbf{u} = \text{grad } h = \begin{bmatrix} \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} & \frac{\partial h}{\partial V_x} & \frac{\partial h}{\partial V_y} & \frac{\partial h}{\partial V_z} \end{bmatrix}^T = \begin{bmatrix} \frac{\mu x}{r^3} & \frac{\mu y}{r^3} & \frac{\mu z}{r^3} & V_x & V_y & V_z \end{bmatrix}^T$$

и

$$h = (\mathbf{u}, \mathbf{L}) = \mathbf{u}^T \cdot \mathbf{L}.$$

Искомая проекция (величина рассеивания интеграла энергии) определяется из правила преобразования ковариационной матрицы:

$$\sigma_h^2 = \mathbf{u}^T K_0 \mathbf{u}.$$

Вопрос. Почему транспонирование оказалось слева, а не справа?

4. Вероятность присутствия точки внутри области рассеивания

36. Отметим важную особенность вероятностного пространства Колмогорова – множества элементов с подмножествами классов случайных событий ("экспериментов") и с вероятностями выборки из каждого класса лишь единственного элемента ("случайного результата эксперимента").

Важно понять, что в формализме математической теории вероятности ответ не однозначен до тех пор, пока не оговорены подробности наблюдаемого случая. Иными словами, ответ зависит от типа объектов и, главное, от способа их выборки в некотором мысленном эксперименте.

37. Хорошим примером этому служит определение вероятности точки (x, y) быть внутри заданной области. Эта вероятность определяется двойным интегрированием функции плотности распределения вероятностей значений величин координат x, y в пределах границ выбранной области:

$$P(\text{Об}) = \iint_{\text{Об}(x,y)} f(x, y) dx dy .$$

Для нормального закона функции $f(x, y)$ соответствующий мысленный вероятностный эксперимент заключается в том, что "бросаются кости" для выборки случайной пары координат и суммируются вероятности только тех получаемых точек, которые оказываются внутри эллипса заданной вероятности.

Круговое рассеивание

38. Рассмотрим частный случай для двумерного нормального рассеивания. При $\sigma_1 = \sigma_2 = \sigma$ и $r = 0$ двумерное круговое рассеивание со случайным радиусом R и с его случайным углом направления α .

Постановка задачи определения вероятности ситуации "внутри-снаружи" для эллипса рассеивания и для окружности мишени совпадают, но вероятностные характеристики у них разные. Это легко понять: случайные координаты x, y могут иметь разные знаки, но размеры случайных радиусов окружности только положительны. Поэтому распределение вероятностей рассеивания величин модулей двумерных радиусов-векторов точек x, y – не гауссово.

39. Это распределение хорошо известно в физике как распределение Рэлея*.

Если измерять радиус отклонения от центра в единицах рассеивания $\lambda = R / \sigma$, то плотность распределения Рэлея имеет вид (рис. А.6)

$$f(\lambda) = \frac{\lambda}{\sigma} e^{-\frac{\lambda^2}{2}} .$$

* Распределение длины вектора с независимыми нормально распределёнными компонентами в 2D-векторном пространстве получено в 1880 г. Д.У.С. Рэлеем, а в 3D-векторном пространстве в 1860 г. Д.К. Максвеллом. Их графики имеют разные производные в нуле (соответственно конечное и нулевое), а их отличие определяется разным числом степеней свободы рассеивания в распределении χ^2 , частными случаями которого эти распределения являются.

Это выражение интегрируется, и вероятность попадания в круг радиуса λ есть

$$P(\lambda) = 1 - e^{-\frac{\lambda^2}{2}}.$$

При этом математическое ожидание попадания в круг радиуса λ равно

$$m_\lambda = \int_0^\infty \lambda f(\lambda) d\lambda = \sqrt{\frac{\pi}{2}} \approx 1.25.$$

Иначе: радиус круга, в котором в среднем будет находиться точка, равен

$$R_{\text{среднее}} = 1.25\sigma.$$

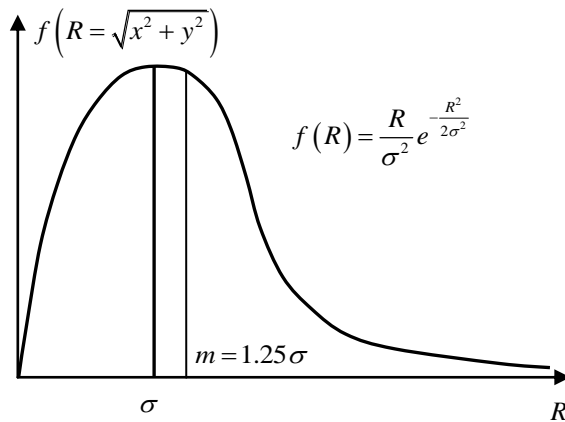


Рис. А.6. Плотность распределения Рэлея для длин радиусов кругового рассеивания

Упражнение 2. Используя распределение Рэлея, рассчитайте 10 радиусов окружностей мишени для стрельбы так, чтобы вероятности попадания в полосы между окружностями были равны.

Вероятность попадания случайной точки внутрь эллипса рассеивания

40. Посмотрим теперь, как определить вероятность принадлежности точки (x, y) внутренности эллипса равной плотности вероятности $\varepsilon(\lambda)$, где λ — параметр размера эллипса.

Искомая вероятность точки (x, y) быть внутри эллипса определяется двойным интегрированием функции плотности в пределах его границ:

$$P(\lambda) = \iint_{\varepsilon(\lambda)} f(x, y) dx dy .$$

Поскольку функция плотности $f(x, y)$ Гаусса в декартовых координатах не интегрируема, то описываемую вероятностную задачу удобнее сформулировать не в виде задачи определения вероятности положения компонент случайного вектора (x, y) внутри эллипса, а в виде задачи о рассеивании в полярной системе координат случайного вектора (ρ, α) – пары значений случайной длины и случайного направления радиусов-векторов точек (x, y) :

$$\begin{cases} x - m_x = \rho \cos \alpha, \\ y - m_y = \rho \sin \alpha. \end{cases}$$

41. Здесь мысленный вероятностный эксперимент другой: бросаются кости для определения случайной длины радиуса точки под гауссианом от центра и для всех возможных направлений суммируются вероятности выпадения тех получаемых векторов, которые оказываются внутри эллипса. При таком подходе *задача практически сводится к одномерной*, поскольку вероятность принадлежности случайных направлений области 2π очевидна: она равна 1.

42. Замена декартовых координат (x, y) на полярные координаты (ρ, α) в функции плотности вероятности рассеивания точек, подчиняющихся двумерному нормальному закону распределения значений их координат в декартовой системе Oxy , разделяет переменные в показателе степени у экспоненты:

$$\rho^2 \left(\frac{\cos^2 \alpha}{\sigma_x^2 (1-r^2)} - \frac{2r \cos \alpha \cdot \sin \alpha}{\sigma_x \sigma_y (1-r^2)} + \frac{\sin^2 \alpha}{\sigma_y^2 (1-r^2)} \right) = \rho^2 s^2 = \lambda^2 .$$

Это позволяет, интегрируя по α в пределах $0 \dots 2\pi$ и по ρ в пределах эллипса ($\rho \leq \lambda/s$), получить новый вид для функции полной вероятности быть внутри эллипса $\varepsilon(\rho, \alpha, \lambda)$ с параметром λ :

$$P(\lambda) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} \int_0^{2\pi} \int_0^{\lambda/s} e^{-\frac{\rho^2 s^2}{2}} \rho d\rho d\alpha .$$

Интегрируя это выражение по ρ , имеем

$$P(\lambda) = \frac{1 - e^{-\frac{\lambda^2}{2}}}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} \int_0^{2\pi} \frac{d\alpha}{s^2} . \tag{A.6}$$

43. Далее (*обратите внимание!*) используется известный (и полезный для применения и в других вероятностных задачах) "трюк" [7], позво-

ляющий заменить оставшееся интегрирование тригонометрических функций вероятностным соображением, делающим заранее известным результат этого интегрирования. Как было упомянуто, вероятность принадлежности подынтегральной функции в (А.6) пределам интегрирования равна 1:

$$P(0 \leq \alpha \leq 2\pi) = 1 = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} \int_0^{2\pi} \frac{d\alpha}{s^2}.$$

Отсюда $\int_0^{2\pi} \frac{d\alpha}{s^2} = 2\pi\sigma_x\sigma_y\sqrt{1-r^2}$, и тогда $P(\lambda) = 1 - e^{-\frac{\lambda^2}{2}}$.

Таким образом, вероятность попадания случайного вектора во внутрь эллипса равной плотности вероятности интегрируема и зависит только от его параметра λ .

44. Как и следовало ожидать, эта зависимость совпадает с выражением для вероятности, получаемой из распределения Релея при $\lambda = \rho s = R / \sigma$, поскольку окружность является частным случаем эллипса.

Вместе с тем заметим, что параметр λ зависит от r , σ_x и σ_y . Поэтому для круга и площади внутри сжатого и наклонённого эллипса значения λ не совпадают.

45. Из сказанного вытекает удобный способ отображения λ -эллипса на круг кругового рассеивания с радиусом, равным λ (рис. А.7).

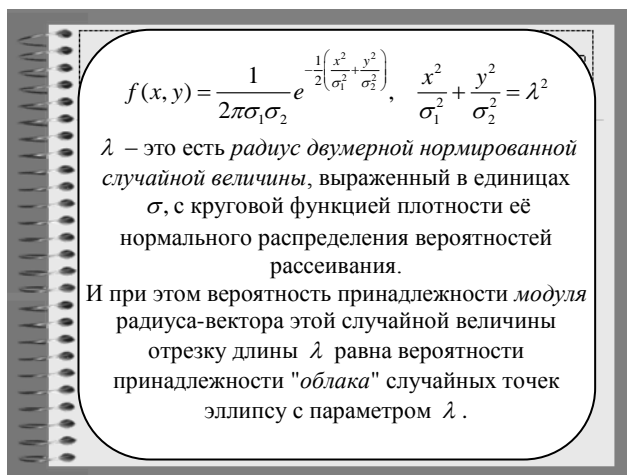


Рис. А.7. Отображение эллипса на равновероятный круг

Упражнение 3. Вычислите вероятность попадания в отрезок $\pm 3\sigma_x$ с центром в m_x при $y = \sigma_y = m_y = 0$. Можете предсказать ожидаемый результат?

46. Важно отметить, что вероятности попадания в круг или эллипс с параметрами $\lambda = \sigma$, $\lambda = 3\sigma$ и $\lambda = E$, $\lambda = 4E$, вычисленные по формуле Рэля, *заметно меньше*, чем в случае одномерного закона: $P(\lambda = 1) = 0.393$, а не $p_\sigma = 0.68$, как в случае одномерного распределения; $P(\lambda = 3) = 0.989$, а не $p_{3\sigma} = 0.993$ для рассеивания одномерного распределения; $P(\lambda = 0.675) = 0.204$, а не $p_E = 0.5$, как в случае одномерного распределения; $P(\lambda = 4 \cdot 0.675) = 0.974$, а не $p_{4E} = 0.993$ для $\pm 4E$ одномерного распределения.

47. И снова видно, что *вероятности попадания в эллипс размером $4E$ или 3σ уменьшаются с увеличением размерности случайной величины!*

Это обстоятельство, как и раньше, отражает факт увеличения числа степеней свободы при увеличении размерности пространства рассеивания случайной величины.

48. Приведём здесь доверительные значения вероятностей попадания в двумерный эллипс рассеивания:

$$p(\lambda) = 0.5 \text{ при } \lambda = \sqrt{-2 \ln 0.5} = 1.177,$$

то есть $\Delta_{0.5} = 1.177\sigma = 1.75E$;

$$p(\lambda) = 0.993 \text{ при } \lambda = \sqrt{-2 \ln 0.007} = 3.151,$$

то есть $\Delta_{\max} = 3.151\sigma = 4.67E$;

$$p(\lambda) = 0.997 \text{ при } \lambda = \sqrt{-2 \ln 0.003} = 3.42,$$

то есть $\Delta_{\max} = 3.42\sigma$, а не 3σ , как в случае одномерного нормального закона распределения вероятностей.

Б. МЕТОДЫ СРЕДНЕКВАДРАТИЧНОЙ ФИЛЬТРАЦИИ

49. Представляется крайне удивительным тот факт, что способ наименьших квадратов и свойства нормального распределения вероятностей были открыты Ф. Гауссом в 1794 г. в возрасте 17 лет [1]. Но ещё более удивительно, что за прошедшие два века не был найден метод обработки, который так или иначе не был бы сводим к среднеквадратичному осреднению.

Способ наименьших квадратов, метод максимума правдоподобия, метод нормальных мест, итерационный и рекурсивный алгоритмы Гаусса, фильтр Калмана–Бьюси, фурье-сглаживание – все эти методы являются методами среднеквадратичного осреднения при различных предположениях. И для их использования, прежде всего, следует проверять выполнимость принятых предположений.

Ниже приведен обзор содержания и особенностей наиболее распространённых методов квадратичного осреднения для задачи определения параметров модели измеряемых величин.

1. Основные понятия теории оценок

50. Рассмотрим основные понятия теории оценок.

Оценка – всякая функция результатов опыта, которая может быть принята за искомую *вероятностную* характеристику $\hat{\alpha}$ некоторого параметра α .

Искомыми вероятностными характеристиками параметра α являются статистические параметры: М.О. $\alpha = m_\alpha$ и среднеквадратичное отклонение σ_α .

Несмещённая оценка – математическое ожидание m_α *совпадает* с искомым истинным значением $\hat{\alpha}$ математического ожидания параметра α .

Смещённая оценка – математическое ожидание m_α оценки *не совпадает* с искомой характеристикой $\hat{\alpha}$, отличаясь от неё на некоторое значение.

Оптимальная оценка – оценка математического ожидания $\hat{m}_\alpha(t)$ искомой величины α , такая, что

$$\hat{m}_\alpha(t) = \sum_{v=1}^n \int_0^T P_v(t, S) \alpha_v(S) dS \text{ и}$$

$$\|\hat{m}_\alpha(t) - m_\alpha(t)\| = \min_{P_v},$$

где $P_v(t, S)$ – функция веса измерений $\alpha_v(t)$, n – число измерений, а $\|\cdot\|$ – норма в некотором смысле.

Состоятельная оценка параметра α – оценка $\hat{\alpha}(\psi_n)$, которая сходится по вероятности к α при $n \rightarrow \infty$ (где ψ_n – вектор из n измерений величины ψ),

$$m(\hat{\alpha}(\psi)) = \int_{-\infty}^{\infty} \hat{\alpha}(\psi) \cdot f(\psi, \alpha) d\psi = \varphi(\alpha),$$

где $f(\psi, \alpha)$ – плотность распределения вероятности.

Математическое ожидание состоятельной оценки зависит от истинного значения α . Для несмещённой оценки верно равенство $\varphi(\alpha) = \alpha$.

Эффективная оценка параметра α – такая оценка $\hat{\alpha}$, которая обладает минимальной дисперсией. Для любой другой оценки $\hat{\alpha}_1$ выполняется неравенство

$$m(\hat{\alpha} - \alpha)^2 \leq m(\hat{\alpha}_1 - \alpha)^2 \text{ для любого } \alpha.$$

Эффективность оценки определяется следующим соотношением:

$$r = \frac{\varphi(\alpha)}{m \left[\frac{d^2 \ln f(\psi, \alpha)}{d\alpha^2} \right] \cdot m(\hat{\alpha} - \alpha)^2} \leq 1,$$

где r – коэффициент корреляции между несмещённой эффективной оценкой и какой-либо другой несмещённой оценкой.

Если $r = 1$, то оценка является эффективной.

Коэффициенты корреляции двух несмещённых эффективных оценок равны. Отсюда следует единственность несмещённой эффективной оценки.

51. Следует помнить, что *любая оценка не может содержать больше информации об определяемом параметре α , чем содержится информации в результатах ψ оценочного опыта.*

Максимум правдоподобия оценки

52. Оценка по максимуму правдоподобия в статистике измерений ψ формулируется следующим образом.

Оценка $\hat{\alpha}$ – это есть то искомое α , при котором условная плотность вероятности определяемых параметров $\hat{\alpha} | \psi$ (получение оценки $\hat{\alpha}$ из опыта) максимальна. Это означает, что параметр $\hat{\alpha}$ лучше всего "согла-

сует" наблюдаемую выборку ψ между собой, и показатель экспоненты нормального закона параметрического распределения вероятностей $\psi | \hat{\alpha}$ максимален. В этом случае

$$\frac{d \ln f(\psi | \alpha)}{d\alpha} = 0. \quad (\text{Б.1})$$

Иными словами, ищется такой параметр $\hat{\alpha}$, который максимальным образом сжимает гистограмму выборки полученных измерений вокруг значения $\psi(\hat{\alpha})$.

53. Если существует наилучшая эффективная оценка, то $\hat{\alpha} = \hat{\alpha}(\psi)$ – единственное решение. Если не существует наилучшей эффективной оценки, то из (Б.1) следует, что $\hat{\alpha}_k = \phi(\psi)$, и решений много. *Это означает плохую наблюдаемость параметра α по наблюдениям ψ .*

54. Выше было отмечено, что центрированные эллипсоиды сечений многомерного гауссиана определяют минимальный размер области рассеивания случайной величины. Таким образом, из геометрии следует утверждение ниже.

Метод максимума правдоподобия в случае нормального распределения сводится к поиску расположения и ориентации минимального эллипсоида, охватывающего полученное облако измерений. Центр такого эллипсоида и определяет искомую оценку $\hat{\alpha}$.

2. Простейшие способы осреднения измерений

55. Рассмотрим простейший случай, когда производятся прямые измерения искомой случайной величины ψ с вычислением её оцениваемой величины $\hat{\psi}$ (оценка $\hat{\psi}$ является математическим ожиданием измерений ψ_i с их ошибками $\delta\psi_i$). Приравняем ошибки $\delta\psi_i$ отклонениям измерений от их среднего значения, предполагаемого с наибольшей вероятностью равным искомой оценке $\hat{\psi}$:

$$\delta\psi_i = \psi_i - \hat{\psi}.$$

Тогда центрированная плотность вероятности величины получаемого измерения ψ_i равна

$$f(\psi_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\delta\psi_i^2}{2\sigma^2}}.$$

56. Если имеется t независимых однородных измерений одной и той же величины ψ , то вероятности их появления умножаются:

$$f(\psi) = \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^m e^{-\frac{1}{2\sigma^2}(\delta\psi_1^2 + \delta\psi_2^2 + \dots + \delta\psi_m^2)}.$$

Найдём такое математическое ожидание величины, при котором для полученного вектора измерений Ψ плотность вероятности будет максимальной. Для этого следует определить необходимые условия существования максимума показателя экспоненты, которые заключаются в условиях минимума квадрата отклонения от принятой оценки этого минимума:

$$\min_{\psi} \sum_{i=1}^m (\psi_i - \hat{\psi})^2.$$

Необходимые условия искомого минимума:

$$\frac{\partial}{\partial \hat{\psi}} \left[\sum_{i=1}^m (\psi_i - \hat{\psi})^2 \right] = 0.$$

Из этого равенства следует формула для вычисления оценки $\hat{\psi}$:

$$\hat{\psi} = \frac{\sum_{i=1}^m \psi_i}{m},$$

представляющая собой обычное выражение для величины среднего значения ряда величин.

57. Если имеется m зависимых разнородных измерений одной величины ψ , то центрированная плотность вероятности вектора измерения Ψ равна (здесь учитываются корреляционные связи погрешностей измерений)

$$\phi(\Psi) = \left(\frac{1}{\sqrt{2\pi}} \right)^m \frac{1}{\sqrt{|K_{\Psi}|}} e^{-\frac{1}{2} \delta \Psi^T K_{\Psi}^{-1} \delta \Psi}, \quad (\text{Б.2})$$

где

$$K_{\Psi} = \begin{vmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_m \end{vmatrix} \cdot \begin{vmatrix} 1 & r_{12} & \dots & r_{1m} \\ r_{12} & 1 & \dots & r_{2m} \\ \dots & \dots & \dots & \dots \\ r_{1m} & r_{2m} & \dots & 1 \end{vmatrix} \cdot \begin{vmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_m \end{vmatrix} = \Sigma \cdot R \cdot \Sigma.$$

Средняя матрица с её диагональю, нормированной среднеквадратичными отклонениями компонент случайного вектора, как уже было сказано (см. выше раздел А.1), носит имя *корреляционной матрицы*. Её элементами являются коэффициенты взаимных корреляций вариаций компонент случайного вектора.

58. Наиболее вероятное значение вектора $\hat{\psi}$ снова получается из условия $\min_{\psi}(\delta\psi^T K_{\psi}^{-1} \delta\psi)$. В случае, когда в измерениях есть корреляция, то есть $R \neq E$, обозначим матрицу K_{ψ}^{-1} следующим образом:

$$P = K_{\psi}^{-1} = \begin{vmatrix} P_{11} & P_{12} & \dots & P_{1m} \\ P_{12} & P_{22} & \dots & P_{2m} \\ \dots & \dots & \dots & \dots \\ P_{1m} & P_{2m} & \dots & P_{mm} \end{vmatrix}.$$

Тогда оценка величины ψ будет равна

$$\hat{\psi} = \frac{\sum_i \sum_j P_{ij} \psi_i \psi_j}{\sum_i \sum_j P_{ij}}.$$

59. В частном случае, если имеется m независимых разнородных измерений ψ_i , то $R = E$, а матрица K_{ψ}^{-1} имеет следующий вид:

$$P = K_{\psi}^{-1} = \begin{vmatrix} \frac{1}{\sigma_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{1}{\sigma_m^2} \end{vmatrix} = \begin{vmatrix} p_1 & 0 & \dots & 0 \\ 0 & p_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p_m \end{vmatrix}.$$

В этом случае наиболее вероятная оценка величины $\hat{\psi}$ будет являться средневзвешенным значением суммы измерений, взятых с их весами p_i :

$$\hat{\psi} = \frac{\sum_{i=1}^m p_i \psi_i}{\sum_{i=1}^m p_i}.$$

Очевидно, что вес измерения p_i тем больше, чем меньше его погрешность.

3. Среднеквадратичная фильтрация параметров модели измерений

60. Наиболее общий случай измерений отличается от описанных выше тем, что доступные измерения ψ несут лишь косвенную информацию о некотором искомом случайном векторе α . В этом случае, когда прямые измерения величин компонент вектора α недоступны оценка последних выполняется с использованием модели связи $\psi(\alpha)$ доступных измерений ψ с компонентами оцениваемого вектора α . Такой метод оценки предполагает, что доступные измерения косвенно связанных с α величин ψ обеспечивают определение искомых величин α . Это означает, что число измерений (размерность вектора ψ) должно быть не меньше числа определяемых параметров (размерности вектора α) и между этими векторами не должно быть заметной линейной зависимости. Иными словами, должна быть *наблюдаемость* параметров α измерениями ψ .

61. Обычно предполагается, что ошибки измерений не велики и лежат в линейной окрестности истинных значений измерений. Поэтому наиболее распространённые методы осреднения сводятся к простым линейным соотношениям. Однако связь определяемых параметров и измеренных величин не обязательно линейная. Она нелинейная, если действительные их значения далеки от принятого первого приближения. Поэтому в общем случае задача фильтрации ошибок измерений не является удобной линейной задачей. В дальнейшем мы рассмотрим связанные с этим обстоятельства и методы, но здесь пока ограничимся рассмотрением классического случая линейной фильтрации.

62. При линейной связи параметров α и ψ в качестве параметров будем рассматривать их значения α и ψ в линейной окрестности некоторых α_0 и ψ_0 (т.е. ниже под обозначениями α и ψ понимаются векторы $\alpha := \alpha - \alpha_0$ и $\psi := \psi - \psi_0$). Если

$$\psi(\alpha) = U\alpha,$$

то в рамках линейности такой же зависимостью связаны и их вариации $\delta\psi$ и $\delta\alpha$ (ошибки измерений ψ и связанные с ними погрешности знания компонент α):

$$\delta\psi(\delta\alpha) = U\delta\alpha.$$

Поэтому в дальнейшем мы не будем отличать разными символами величины компонент ψ_m и $\delta\psi_m$, α_n и $\delta\alpha_n$ рассматриваемых векторов и их ошибок истинности их знания, отмечая их отличие лишь в случае

необходимости. Иными словами, далее везде смысл используемых обозначений ψ и α для линеаризованных векторов или их инструментальных погрешностей зависит от контекста.

63. Для определённости задача модельной фильтрации измерений ниже рассматривается на примере задачи определения элементов орбиты космического аппарата по косвенным данным измерений параметров его движения.

Легко понять, что нет способа прямых измерений элементов орбиты или гелиоцентрических координат и компонент скорости полёта. Поэтому в космической технике (как и во многих других случаях) приходится довольствоваться лишь косвенными измерениями. Обычно это – "наклонные дальности" от пункта наблюдения (вращающегося вместе с Землёй) до космического аппарата, его радиальные ("доплеровские") скорости, наблюдаемые из этого пункта, и различные угловые параметры видимости космического аппарата.

64. Будем предполагать, что состав измерений достаточно широк, измерения выполняются в большом количестве и есть априорные предположения об ожидаемых значениях определяемых параметров. Будем предполагать также, что модель связи измерений с определяемыми параметрами не точна и что возможные ошибки измерений $\delta\psi$ содержат не только случайные, но и систематические составляющие. *В этих предположениях следует определить не только наиболее вероятные оценки величин определяемых параметров $\hat{\alpha}$, но и остаточные погрешности $\Delta\hat{\alpha}$ знания этих величин.*

65. Каждое измерение ψ_i ограничивает в линейном приближении значения параметров α модели наблюдаемого процесса своим *условным уравнением* [5]

$$\mathbf{u}_i \cdot \alpha = \psi_i, \quad i = 1, \dots, m.$$

Здесь \mathbf{u}_i есть *вектор-строка* длины n , представляющая собой компоненты вектора градиента величины измерения ψ_i , в n -мерном α -пространстве, α – вектор-столбец (длины $n \leq m$) значений параметров модели наблюдаемого процесса.

Число m таких условных уравнений должно быть не менее числа n искоемых α -параметров модели процесса, но при большом числе измерений их будет много больше. Соответствующая система условных уравнений имеет вид

$$U \cdot \alpha = \psi, \tag{Б.3}$$

где прямоугольная матрица U $m \times n$ содержит m *векторов-строк* градиентов \mathbf{u}_i , α – n -мерный вектор параметров модели движения, а ψ – слу-

чайный m -вектор измеренных значений ψ_i . Из-за ошибок измерений и погрешностей модели (*мешающих параметров* – см. ниже) эта система оказывается несовместной.

Способ наименьших квадратов (метод максимума правдоподобия)

Найдём оценку вектора параметров $\hat{\mathbf{a}}$ при условии известных значений измерений Ψ с их ковариационной матрицей K_Ψ .

66. В соответствии с правилом Байеса априори случайные значения компонент "длинного" m -мерного вектора измерений Ψ группируются около их модельных значений $\hat{\Psi} = U\hat{\mathbf{a}}$, в то время как апостериорные (после получения измерений) параметры "короткого" n -мерного вектора \mathbf{a} , соответствующие полученным измерениям, рассеиваются вокруг $\hat{\mathbf{a}} = U^{-1}\hat{\Psi}$ (по поводу обращения прямоугольной матрицы $U^{-1} = W$ вспомним сказанное в фрагменте лекций **34***).

67. В то же время максимум вероятности совместного появления совокупности измерений Ψ соответствует вектору $\hat{\Psi}$, дающему минимум квадратичной формы (Б.2) для вектора Ψ случаев измерений:

$$\hat{\Psi} : \min_{\Psi} [(\Psi - \hat{\Psi})^T K_\Psi^{-1} (\Psi - \hat{\Psi})].$$

Преобразуем это условие для случайного вектора \mathbf{a} с учётом условия наибольшего правдоподобия рассеивания $\Psi - \hat{\Psi} = \Psi - U\hat{\mathbf{a}}$. Тогда имеем

$$\hat{\mathbf{a}} : \min_{\mathbf{a}} [(\mathbf{a} - \hat{\mathbf{a}})^T U^T K_\Psi^{-1} U (\mathbf{a} - \hat{\mathbf{a}})].$$

68. Для определения наиболее вероятной оценки $\hat{\mathbf{a}}$ должно выполняться равенство нулю производных элементов этой квадратичной формы по элементам вектора $\hat{\mathbf{a}}$:

$$\frac{\partial}{\partial \hat{a}_i} [(\mathbf{a} - \hat{\mathbf{a}})^T U^T K_\Psi^{-1} U (\mathbf{a} - \hat{\mathbf{a}})] = 0.$$

Подобное дифференцирование заключается в частном (по элементам вектора $\hat{\mathbf{a}}$) дифференцировании скалярных произведений векторов $(\mathbf{a} - \hat{\mathbf{a}})$ с векторами градиентных строк матрицы U – при условии, что отображения измерений в \mathbf{a} – постоянные числа. При дифференцировании этих скалярных произведений исчезают только векторные скобки, а компоненты строк-градиентов в U остаются (*вспомните или проверьте*

* Далее такие ссылки будут приводиться в виде – п. **34**.

это утверждение векторного анализа). В результате получаем следующие уравнения:

$$U^T K_{\Psi}^{-1} U (\alpha - \hat{\alpha}) + (\alpha - \hat{\alpha})^T U^T K_{\Psi}^{-1} U = 0.$$

В силу симметричности матрицы $U^T K_{\Psi}^{-1} U$ левое и правое слагаемые этого выражения приводят к одному и тому же n -мерному вектору, записанному в столбец или в строку. Необходимые условия искомого оптимума говорят, что этот вектор должен быть нуль-вектором.

Таким образом, оба слагаемые выражения для производной приводят к одной и той же системе уравнений:

$$U^T K_{\Psi}^{-1} U \hat{\alpha} = U^T K_{\Psi}^{-1} U \alpha.$$

Заменяя в правой части случайный вектор α его прообразом – вектором Ψ (т.е. используя систему условных уравнений (Б.3): $U\alpha = \Psi$), получаем систему уравнений для решения методом квадратичного осреднения задачи определения искоемых компонент n -мерного вектора по косвенным измерениям:

$$U^T K_{\Psi}^{-1} U \hat{\alpha} = U^T K_{\Psi}^{-1} \Psi. \quad (\text{Б.4})$$

Вопрос. Как будет выглядеть аналогичная система для случая *прямых* измерений n -мерного вектора определяемых параметров?

69. Полученная система (Б.4) представляет собой основное выражение для двух известных методов модельной обработки экспериментальных данных:

- способа наименьших квадратов ("СНК") (когда матрица K_{Ψ} диагональная),
- метода максимума правдоподобия ("ММП") (матрица K_{Ψ} недиагональная).

Эта система обычно называется *системой нормальных уравнений* или просто *нормальной системой*. По-видимому, это название связано с тем, что в отличие от несколько "ненормального" вида несовместной системы условных уравнений с её прямоугольной матрицей U , выражение (Б.4) слева содержит уже $n \times n$ -квадратную и обращаемую ("нормальную") матрицу и, соответственно, n -мерный столбец правых частей системы.

Из решения нормальной системы уравнений следует выражение для оценки вектора параметров $\hat{\alpha}$ через вектор случайных величин Ψ :

$$\hat{\alpha} = [U^T K_{\Psi}^{-1} U]^{-1} U^T K_{\Psi}^{-1} \Psi. \quad (\text{Б.5})$$

70. Важное обстоятельство: для выполнения этой оценки нужно знать ковариационную матрицу K_{Ψ} нормально распределённых ошибок измерений. Это самое "тонкое" место способа наименьших квадратов.

Известный космонавт Г.М. Гречко (который до того, как он стал космонавтом, был прекрасным баллистиком) как-то очень метко заметил: "Данные должны быть либо верные, либо официальные!"

В свете этих слов используемое в оценках описываемого рода предположение о нормальном законе распределения рассматриваемых погрешностей – достаточно *верное*, поскольку оно базируется на предельной теореме вероятностей (*при росте числа независимых случайностей с разными распределениями их вероятностей распределение вероятностей суммарной случайности стремится к нормальному распределению вероятностей*).

Но если говорить о *надёжности* используемых значений элементов ковариационных матриц, то эти данные скорее всегда более "официальные", чем верные... Заметим, полушутя, что богатый русский язык неоднозначно трактует слово "данные".

Приведём пример из практики. В формуляре нужного прибора (например, интегратора кажущихся ускорений) приведено официальное значение его погрешности (например, 6 см/с). Естественно предположить, что имеется в виду величина 3σ погрешности прибора с её вероятностью 0.997. Однако, как в формуляр попало это число? Были проведены испытания и их результаты внесены в формуляр? Нет, как правило, это число взято из технического задания (ТЗ) на разработку прибора.

Ведь достаточно очевидно, что на всех стадиях конструирования, разработки технологии и последующего изготовления прибора каждый из разработчиков, глядя на предъявленные ему требования точности, выполнял эти требования с некоторым своим запасом. В результате статистика цеховых, заводских и приёмочных испытаний могла показать намного меньшую величину истинных 3σ погрешности (скажем, 2 см/с). Но в формуляре главный конструктор меньше, чем требование ТЗ на его прибор, просто на всякий случай подписывать не станет. И после этого никто не посмеет использовать иное значение ошибки прибора, чем то, что написано в его формуляре.

Упражнение 4. Определите истинный уровень вероятности "официального" значения погрешности в 6 см/с (которая и будет заложена в следующие расчёты).

Учёт априорной информации

71. Как правило, заранее известны ожидаемые ("номинальные") значения определяемых параметров α_0 и их предполагаемый разброс, описываемый ковариационной матрицей K_{α_0} . Более того, могут быть известны и некоторые другие параметры \mathbf{a} с их ковариационной матрицей $K_{\mathbf{a}}$, величина которых как-то связана с величиной определяемых параметров

α . Было бы разумно учесть в получаемой оценке параметров α эти, пусть не очень точные, априорные знания α_0 и \mathbf{a} , хотя бы таким образом, чтобы α и \mathbf{a} не противоречили друг другу.

Как это сделать?

С точки зрения логики способа наименьших квадратов любые априорные сведения могут расцениваться как некоторые *дополнительные косвенные измерения*, которым, однако, опасно приписывать высокий вес!

Вопрос. Почему?

Для этого выразим в линейном приближении связь вектора априорных знаний с параметрами α в виде соотношений типа $V\alpha = \mathbf{a}$ (очевидно, что если $\mathbf{a} = \alpha_0$, где α_0 – номинальное (исходное) значение параметра α , то матрица $V = E$). Тогда, предполагая очевидную независимость погрешностей измерений и диапазонов незнания априорных данных (то есть матриц K_ψ и K_a), получим

- новую расширенную систему условных уравнений с её новыми дополнительными измерениями \mathbf{a} ;
- новую расширенную клеточно-диагональную ковариационную матрицу;
- соответствующую систему *нормальных уравнений*, отвечающую условию максимальной вероятности одновременного появления измеренных значений ψ_i и реализации априорных предположений относительно значений компонент вектора априорных параметров \mathbf{a} :

$$\tilde{U} = \begin{bmatrix} U \\ V \end{bmatrix}, \quad \tilde{K}_\psi = \begin{bmatrix} K_\psi & 0 \\ 0 & K_a \end{bmatrix}.$$

Раскрывая клеточную диагональность расширенной ковариационной матрицы, получим более удобное выражение для оценки вектора параметров α :

$$\left[U^T K_\psi^{-1} U + V^T K_a^{-1} V \right] \hat{\alpha} = U^T K_\psi^{-1} \psi + U^T K_a^{-1} \mathbf{a}. \quad (\text{Б.6})$$

В случае $\mathbf{a} = \alpha_0$ – априорной информации о номинальном значении вектора параметров α , с $V_0 = E$, имеем

$$\begin{aligned} \left[U^T K_\psi^{-1} U + K_{\alpha_0}^{-1} \right] \hat{\alpha} &= U^T K_\psi^{-1} \psi + K_{\alpha_0}^{-1} \alpha_0, \\ \hat{\alpha} &= \left(U^T K_\psi^{-1} U + K_{\alpha_0}^{-1} \right)^{-1} \cdot \left(U^T K_\psi^{-1} \cdot \psi + K_{\alpha_0}^{-1} \alpha_0 \right). \end{aligned} \quad (\text{Б.7})$$

72. Обозначим новую *матрицу нормальной системы* СНК или ММП как

$$N = \left[U^T K_\psi^{-1} U + K_{\alpha_0}^{-1} \right]. \quad (\text{Б.8})$$

С этой матрицей N ниже будет связано много обстоятельств:

- запомним пока, что обращённая матрица $N^{-1} = K_a$, т.е. она является оценкой точности знания параметров \mathbf{a} (оценкой формальной и, увы, неполной);

- для её обращения с преодолением, как правило, её плохой обусловленности и для сохранения симметрии обратной матрицы придётся изучить специальный *метод квадратных корней*;

- с этой матрицей связано смысловое понимание

- *нормальных мест* измерений,
- устройства фильтров Гаусса и Калмана,
- *мешающих параметров*
- и решения методом *параболического спуска* нелинейной задачи обработки измерений.

Метод нормальных мест

73. В случае клеточной структуры матрицы K_ψ (*кусочной* корреляции погрешностей измерений) объединение и суммирование по i -м клеткам матрицы коэффициентов нормальной системы и соответствующих элементов столбца правых частей позволяет построить выражение для *нормальных мест* \mathbf{a}_i ($i < j$) – осреднённых в прошлом искомым параметров \mathbf{a} , сгруппированных и принимаемых к осреднению с большим весом (пропорциональным сумме i -х клеток):

$$\begin{aligned} & \left[U_j^T K_{\psi_j}^{-1} U_j + U_{j-1}^T K_{\psi_{j-1}}^{-1} U_{j-1} + \dots + U_1^T K_{\psi_1}^{-1} U_1 + K_{\mathbf{a}_0}^{-1} \right] \mathbf{a}_j = \\ & = U_j^T K_{\psi_j}^{-1} \boldsymbol{\psi}_j + U_{j-1}^T K_{\psi_{j-1}}^{-1} \boldsymbol{\psi}_{j-1} + \dots + U_1^T K_{\psi_1}^{-1} \boldsymbol{\psi}_1 + K_{\mathbf{a}_0}^{-1} \mathbf{a}_0. \end{aligned}$$

Отсюда видна возможность замены всей $0 \dots j-1$ группы старых измерений столбца правых частей и суммы их клеток в матрице коэффициентов нормальной системы одним *нормальным местом* \mathbf{a}_{j-1} с его ковариационной матрицей

$$K_{\mathbf{a}_{j-1}} = \left[U_{j-1}^T K_{\psi_{j-1}}^{-1} U_{j-1} + K_{\mathbf{a}_{j-2}}^{-1} \right]^{-1}$$

(равенство $N^{-1} = K_a$ будет доказано ниже),

$$\hat{\mathbf{a}}_{j-1} = \left[U_{j-1}^T K_{\psi_{j-1}}^{-1} U_{j-1} + K_{\mathbf{a}_{j-2}}^{-1} \right]^{-1} \cdot \left(U_{j-1}^T K_{\psi_{j-1}}^{-1} \boldsymbol{\psi}_{j-1} + K_{\mathbf{a}_{j-2}}^{-1} \hat{\mathbf{a}}_{j-2} \right) \quad (\text{Б.9})$$

и суммирования с этим нормальным местом новой j -й группы измерений.

Нетрудно заметить, что априорная информация и формульно, и по существу также представляет собой некое априорное нормальное место (сравните с (Б.7)).

74. Важно иметь в виду, что при последовательной замене групп измерений их локально-осреднёнными нормальными местами есть угроза

потери информации (см. рис. В.1) о неизвестных корреляционных связях ошибок измерений внутри групп измерений. Это отражается в противоречии наблюдаемой картины рассеивания прошлых измерений с принятой моделью их "облака" чисто случайных ошибок.

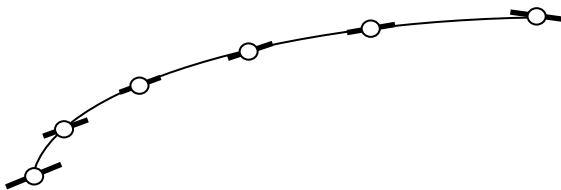


Рис. В.1. Пример неудачных нормальных мест

Поэтому забывать значения прошлых измерений, заменяя их последним осреднённым нормальным местом, нужно очень осторожно. Правильнее – хранить все измерения и анализировать апостериорную картину их рассеивания относительно построенной траектории движения в фазовом пространстве наблюдаемого объекта.

75. Однако в ряде случаев замена группы измерений их нормальным местом очень правомерна. Это, прежде всего, полезно тогда, когда новая группа измерений не может быть усредняемой с последующими измерениями.

Примером такой ситуации служат нарушающие возможность осреднения изломы траектории в точках её коррекции импульсным изменением скорости полёта. В этом случае в её точке коррекции вся предыдущая траектория заменяется двумя нормальными местами: нормальным местом вычисленных с хорошей точностью координат этой точки и нормальным местом с уменьшенным весом компонент скорости, которые для процесса построения послекоррекционной траектории играют роль априорной информации.

Почему вес нормального места скорости полёта в точке коррекции нужно было уменьшить, несмотря на хорошее знание скорости полёта, полученное по предыдущим измерениям? Ответ совпадает с ответом на вопрос, заданным выше при описании способа учёта априорной информации.

Если задать высокую точность априорной информации, то эта информация и станет ответом осреднения (сравните величины обоих слагаемых матрицы N при разных дисперсиях ошибок измерений и априорных знаний).

В упомянутом примере траектория после её коррекции должна быть снова определена уже на новой измерительной базе. Но может случиться, что коррекция была произведена с ошибкой. Поэтому задаваемая априори предполагаемая точность знания вектора скорости полёта в точке коррекции после окончания работы корректирующего двигателя не может быть

выше максимального размера ошибок реализации величины и направления требуемого изменения скорости полёта.

Похожим примером разумного использования технологий нормальных мест является стыковка измерений первичной "приземной" и последующей "дальней" радиосистем траекторных измерений.

Траекторные измерения пунктов радионаблюдения, расположенных на территории России, к концу видимости космического аппарата, улетающего от Земли, хорошо определяют его координаты и хуже определяют вектор его скорости по доплеровскому сдвигу радиочастоты (ввиду почти ортогональности вектора скорости полёта вектору наблюдаемой пунктом наклонной дальности). Этот вектор скорости наблюдаемого полёта будет хорошо определён дальней системой траекторных измерений, но это произойдёт позже – лишь через 12 часов полёта, когда в суточном вращении Земля повернётся на $180,0^\circ$ и с территории России космический аппарат снова станет виден.

Вопрос. На каком расстоянии это произойдёт, если скорость отлёта от Земли была равна 11 км/с?

В этом случае при обработке измерений дальней космической связи удобно все многочисленные и проверенные на правильность измерения приземной системы заменить нормальными местами координат и компонент вектора скорости в конце первого участка видимости космического аппарата.

76. Другим мотивом для использования нормальных мест являются случаи, когда большая группа измерений заметно уточняется последующими измерениями. Хорошим примером этого является обработка измерений при полёте к Луне для фотографирования неизвестной до этого обратной стороны Луны.

Траектория этого полёта (см. рис. Б.2) была необычной. В ней впервые был применён пертурбационный манёвр с "подныриванием" под Луну, поскольку требовалось обеспечить хорошую видимость с нашей территории и при полёте к Луне, и при полёте обратно к Земле. Иными словами, и туда и обратно требовалось лететь в северном полушарии. Для этого надо было очень точно прилететь к Луне, так, чтобы её притяжение забросило космический аппарат на нужную траекторию обратного полёта к Земле после выполнения фотографирования Луны.

Это означало необходимость тщательного наблюдения за полётом для хорошего знания траектории подлёта к Луне ввиду опасности последующей потери космического аппарата антеннами дальней космической связи. Эта опасность проистекала из-за того, что при близком пролёте Луны её гравитационное поле является мощным усилителем траекторных ошибок. В результате узкая трубка возможного рассеивания траекторий подлёта к Луне после её пролёта расширялась в громадный веер траекторий

(что в свою очередь требовало определения верной траектории, чтобы позже – уже у Земли – суметь передать на Землю впервые полученную фотографию обратной стороны Луны).

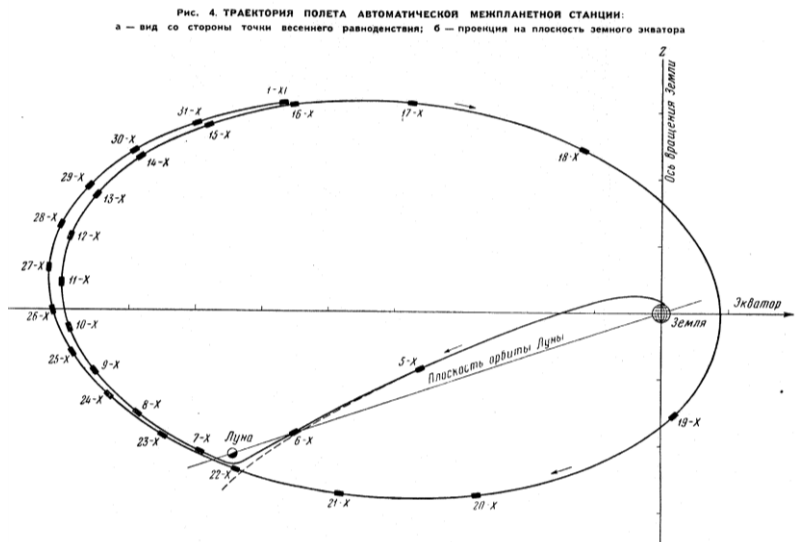


Рис. Б.2. Из книги: Первые фотографии обратной стороны Луны.
М. : АН СССР, 1959

И после этого объяснения обстоятельств того полёта должно стать понятным, что после пролёта Луны все траекторные наблюдения до сближения с Луной было разумным заменить одним нормальным местом. Это было связано с тем, что пространственное разрешение на широкой трубке траекторий после Луны, а значит, и точность траекторного определения, была выше, чем на узкой трубке траекторий полёта к Луне. Иными словами, измерения до сближения с Луной при возвращении к Земле уже мало влияли на результаты траекторного слежения. И их вполне можно было заменить нормальным местом. Это не только сэкономило очень небольшую память (2К слов!) машин того времени, но и, что было важнее, резко сокращало время очередного уточнения параметров орбиты космического аппарата.

4. Простые среднеквадратичные фильтры измерений

77. На примере техники нормальных мест мы встречаемся с широко применяемыми процедурами *фильтрации* измерений [20], когда каждое очередное независимое случайное измерение (или их локальные группы) обрабатывается, учитывается и забывается. Это требует осторожности,

поскольку даже в случае простой линейной связи измерений с определяемыми параметрами с её свойствами локальности и марковости простоте процедуры фильтрации может противоречить необходимость хранения всех полученных данных для учёта их семантической непротиворечивости в апостериорном анализе.

Фильтры делятся на простые и рекурсивные – по типу применяемых весовых или ковариационных матриц. Сначала рассмотрим простые фильтры.

Аккумулятивный фильтр Гаусса

78. В линейном случае некоррелированных погрешностей измерений (с их диагональной ковариационной матрицей) метод нормальных мест приводится к предложенному Гауссом "*аккумулятивному*" фильтру [13, 1] с его свойством минимального рассогласования. Гаусс предложил по мере прихода каждого нового измерения Ψ_i лишь накапливать очередные значения компонент вектора \mathbf{Y}_i (т.е. вектора столбца правых частей нормальной системы уравнений) и элементов её матрицы N_i , а обращение этой матрицы и построение нормального места выполнять по мере необходимости – лишь в редких избранных точках $i = k$.

На стадии накопления данных увеличение \mathbf{Y}_i и $N_i = K_{\mathbf{a}_0}^{-1}$ происходит так:

$$\mathbf{Y}_i = \mathbf{Y}_{i-1} + U_i^T K_{\Psi_i}^{-1} \Psi_i,$$

$$N_i = N_{i-1} + U_i^T K_{\Psi_i}^{-1} U_i.$$

На стадии определения искомого вектора \mathbf{a}_k при $i = k$ вычисляются оценка вектора определяемых параметров и матрица его ошибок:

$$\hat{\mathbf{a}}_k = N_k^{-1} \cdot \mathbf{Y}_k \text{ и } K_{\mathbf{a}_k} = N_k^{-1}.$$

Важно подчеркнуть, что при подобном подходе ошибки обращения матрицы не влияют на погрешности операций накопления. Причём в простейшем случае определения *единственного* параметра α путём фильтрации одиночных измерений ψ_j операция обращения матрицы заменяется операцией деления.

Учёт фазовых шумов

79. В задачах фильтрации текущего состояния наблюдаемого процесса очень часто существует необходимость учёта влияния действующих возмущений фазового состояния процесса внешними факторами.

Примерами задач с фазовыми шумами могут служить задачи самолётной или морской навигации, в которых необходимо учитывать влияние движений внешней среды – ветра и морских течений.

В мобильной и манипуляционной робототехнике фазовые шумы связаны с неопределённостями параметров текущего состояния внешней среды, влияющих, например, на факторы опорной и профильной проходимости машины, на процессы формирования или сопряжения деталей или на свойства зрительного распознавания операционной обстановки.

В космической технике к фазовым шумам сводятся неопределённости орбитального влияния солнечного давления и работы реактивных двигателей системы ориентации при отсутствии точных телеметрических данных об угловом положении космического аппарата. Другой тип фазовых шумов связан с неопределённостями верхней атмосферы и магнитного поля Земли при их влиянии на движение космического аппарата около его центра масс.

80. В каждом из перечисленных случаев в результат фильтрации добавляется дополнительная к погрешностям измерений фазовая неопределённость воздействия внешней среды. Эта неопределённость существенно отличается от неопределённости погрешностей измерения характером её суммирования в результатах обработки данных.

Образно говоря, это отличие связано с отличием суммирования ошибок и суммирования знаний. При суммировании погрешностей главное влияние оказывает наибольшая погрешность, в то время как при суммировании знаний важнее всех наиболее точное знание. Иными словами, при суммировании ошибок мы складываем их дисперсии, в то время как при суммировании сведений мы складываем их веса, т.е. величины, обратные дисперсиям погрешности данных. Здесь прямая аналогия с законом суммарного сопротивления в последовательной или параллельной цепи резисторов.

81. Сказанное делает очевидным присутствие именно *обратной* ковариационной матрицы в законе плотности нормального распределения вероятностей: если в совокупности реализаций опыта присутствует точный результат, то все остальные не так важны. *Суммируя измерения, мы суммируем знания.*

В то же время если разные неопределённости (погрешности знания) рассматриваются вне метода обработки и осреднения данных, то должны суммироваться их *ковариационные* матрицы.

Таким образом, *поскольку фазовые шумы не измеряются* и их присутствие лишь увеличивает погрешность результата, то для учёта их влияния на результат *должны быть просуммированы ковариационные матрицы* результата фильтрации данных и модели фазовых шумов.

Определение состояний процесса в момент наблюдения

82. Выше везде рассмотренными методами осреднения ошибок измерений были те, в которых неизвестными были *постоянные* значения век-

тора определяемых параметров α . Например, по траекторным радиоизмерениям определяются элементы орбиты космического аппарата или формирующие будущую его орбиту координаты и компоненты скорости в момент выключения последней ступени ракеты-носителя.

Вместе с этим существуют задачи, в которых определяемыми параметрами являются значения фазовых переменных наблюдаемого объекта именно в момент его наблюдения. Например, определяются высота, курс и скорость полёта самолёта в момент его лоцирования или координаты и компоненты скорости спутника в момент измерения значений его наклонной дальности и/или доплеровской частоты. В таких случаях вектор определяемых параметров является функцией времени движения $\beta(t)$, и при фильтрации измерений наблюдаемых параметров необходимо использовать модель динамики объекта наблюдения для переноса в текущую точку сведений о прошлом его состоянии.

83. Важно подчеркнуть, что смысл и содержание процедур осреднения определяемого *переменного* параметра существенно отличается от описанных выше. Возникающие изменения алгоритмов способа наименьших квадратов можно показать на примере простейшего случая определения последовательности $\beta(t_i)$ векторов положений $\mathbf{R} = [x_i \ y_i \ z_i]^T$ космического аппарата в моменты выполнения измерений t_i ($i = 1, \dots, m$) в пространстве с выбранной системой координат.

Система условных уравнений в этом случае имеет тот же вид:

$$V\beta = \psi.$$

Но здесь вектор определяемых параметров β , имеющий размерность $3m$, содержит *все прошлые векторы орбитальных позиций* в моменты измерений.

Матрица системы условных уравнений V представляет собой клеточно-диагональную матрицу с m клетками-строками векторов градиентов каждого измерения в пространстве $\mathbf{R} : [x_i \ y_i \ z_i]^T$:

$$V = \begin{vmatrix} v_{11} & v_{12} & v_{13} & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & v_{21} & v_{22} & v_{23} & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & v_{31} & v_{32} & v_{33} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{vmatrix}.$$

84. Заметим, что текущие компоненты вектора β_i связаны с определяемыми ранее постоянными параметрами орбиты α соотношением $\beta_i = \beta(\alpha, t_1, t_2, \dots, t_i)$. Заметим, что параметры орбиты теперь неизвестны, но принципиально такая связь существует. Для достаточно узкой трубки

траекторий в линейном приближении в каждый момент прихода очередного i -го измерения эта связь имеет вид

$$\beta_i = D_i \alpha,$$

где матрица D_i имеет размерность $n \times 3m$, определяемую размерностью n текущего вектора ψ .

В то же время матрица условных уравнений U связи измерений ψ с постоянным вектором α связана с матрицей условных уравнений вектора β соотношением: $U = VD$.

85. Эти зависимости позволяют (либо их формальной подстановкой, либо преобразованием в пространство векторов β условий максимальной вероятности одновременного появления измерений вектора ψ) получить выражение для нормальной системы способа наименьших квадратов (заметим, что размерность её матрицы N теперь равна $m \times m$):

$$\left[V^T K_{\psi}^{-1} V + K_{\beta^+}^{-1} \right] \hat{\beta} = V^T K_{\psi}^{-1} \psi + K_{\beta^+}^{-1} \beta^+. \quad (\text{Б.10})$$

Внешний вид этой системы СНК аналогичен предыдущему, но вероятностный смысл переменных существенно иной. α_{i-1} – это было значение нормального места *постоянного* вектора α по данным вектора измерений ψ_{i-1} , β_{i-1}^+ – теперь это значение в текущей точке t_i нормального места *переменного* вектора β , полученного в момент t_{i-1} измерения по данным вектора ψ_{i-1} и *перенесённого* с учётом траекторных изменений в текущую точку t_i . $K_{\beta_{i-1}^+}^{-1}$ – весовая матрица нормального места β_{i-1} в точке t_{i-1} , полученная в результате обработки всех измерений вектора ψ_{i-1} к моменту t_{i-1} , но также затем *перенесённая* вдоль траектории полёта в точку t_i .

Именно необходимость переносить параметры β_{i-1}^+ и $K_{\beta_{i-1}^+}^{-1}$ нормального места предыдущих данных в точку нового измерения отличает алгоритмы СНК для постоянных оцениваемых параметров от алгоритмов осреднения переменных фазовых параметров текущего состояния подвижного наблюдаемого объекта в каждой новой точке измерения.

86. Для переноса в точку момента t_i нового измерения ψ_i фазового состояния наблюдаемого процесса параметров предыдущей точки в момент t_{i-1} используется линейное прогнозирование параметров динамической системы

$$\beta_{i-1}^+ = \Phi_{i-1,i} \beta_{i-1},$$

$$K_{\beta_{i-1}^+}^{-1} = \Phi_{i,i-1}^T K_{\beta_{i-1}}^{-1} \Phi_{i,i-1} = \left[\Phi_{i-1,i} K_{\beta_{i-1}} \Phi_{i-1,i}^T \right]^{-1}.$$

Здесь $\Phi_{i-1,i}$ – переходная матрица из точки β_{i-1} в точку β_i , определяемая динамикой процесса и интервалом времени его развития между двумя рассматриваемыми дискретными моментами t_{i-1} и t_i . Её обратная матрица – $\Phi_{i,i-1}$. Очевидно, что условия линейности переходной матрицы $\Phi_{i-1,i}$ накладывают определённые ограничения на допустимый нижний уровень частоты наблюдений.

Отметим, что требуемый перенос предыдущего нормального места в точку нового измерения приводит к необходимости на каждом шаге накопления использовать как *прямую*, так и *обратную переходные* матрицы $\Phi_{i-1,i}$ и $\Phi_{i,i-1}$ между точками соседних измерений. При движении в потенциальном поле сил их легко обращать транспонированием клеток. В других случаях можно использовать свойства сопряжённых динамических систем.

87. Нормальные места $\hat{\beta}_i$ фильтра координат точек траектории, получаемых в результате поочерёдного осреднения векторов ψ_i вариантом (Б.10) СНК равны

$$\hat{\beta}_i = \left[V_i^T K_{\psi_i}^{-1} V_i + K_{\beta_{i-1}^+}^{-1} \right]^{-1} \left(V_i^T K_{\psi_i}^{-1} \psi_i + K_{\beta_{i-1}^+}^{-1} \beta_{i-1}^+ \right). \quad (\text{Б.11})$$

Упражнение 5. Постройте СНК для определения переменных значений фазовых параметров в моменты времени, не совпадающие с моментами выполнения измерений.

Аккумулятивная фильтрация состояния процесса с фазовыми шумами

88. Казалось бы, что систему (Б.11) с вектором β_i ($i=1, \dots, m$) можно попытаться записать в виде аккумулятивного фильтра Гаусса (см. п. 78):

Стадия накопления ряда измерений:

$$K_{\beta_{i-1}^+}^{-1} = \Phi_{i,i-1}^T K_{\beta_{i-1}}^{-1} \Phi_{i,i-1},$$

$$N_i = N_{i-1} + V_i^T K_{\psi_i}^{-1} V_i + K_{\beta_{i-1}^+}^{-1},$$

$$\beta_{i-1}^+ = \Phi_{i-1,i} \beta_{i-1},$$

$$Y_i = Y_{i-1} + V_i^T K_{\psi_i}^{-1} \psi_i + K_{\beta_{i-1}^+}^{-1} \beta_{i-1}^+.$$

Акт оценки нужного состояния:

$$\hat{\beta}_k = N_k^{-1} Y_k \quad \text{и} \quad K_{\beta_k} = N_k^{-1}.$$

Однако это не так, поскольку идея аккумуляции без обращения матриц здесь наталкивается на *необходимость на каждом шаге фильтрации определять значение вектора β_{i-1}* (что и видно из 3-й формулы фильтра).

Это нужно для последующего определения нормального места β_{i-1}^+ в следующей i -й точке. И это означает, что в конце каждого предыдущего шага фильтрации требуется выполнять обращение матрицы N нормальной системы, что обесценивает достоинства *аккумулятивного* алгоритма, превращая его в алгоритм фильтрации методом нормальных мест.

89. Другое обстоятельство, связанное с присутствием фазовых шумов, тоже требует промежуточных обращений матриц. Здесь это связано с необходимостью *одновременного накопления как прямых, так и обратных ковариационных матриц*. Действительно, при учёте фазовых возмущений, как было сказано, нужна ковариационная матрица результирующей оценки, а при осреднении ошибок измерений нужна её весовая матрица. Таким образом, на каждом шаге накопления данных требуется обращение ковариационной матрицы.

Обобщение аккумулятивного фильтра Гаусса на случай линейной модели процесса с переменными параметрами $\beta_i(t)$ и с фазовыми шумами в несколько сложном виде нашёл Y. Genin [13]. Преобразование его алгоритма позволяет привести упрощённую запись этого фильтра в принятых выше обозначениях.

90. В отличие от аккумулятивного фильтра Гаусса, приведённого в предыдущем фрагменте, теперь в линейную модель связи прошлого и текущего фазовых состояний наблюдаемого процесса добавляется действующий на объект наблюдения фазовый шум q :

$$\beta_i = \Phi_{i-1,i} \beta_{i-1} + q_i.$$

Предполагается, что в каждый i -й момент наблюдения вектор фазового шума q_i неизвестен, но априори известна его ковариационная матрица Q . Предполагается также, что известны обе переходные матрицы $\Phi_{i,i-1}$ и $\Phi_{i-1,i}$. В этих предположениях аккумулятивная процедура Гаусса имеет следующий вид.

Сначала производится перенос матрицы $K_{\beta_{i-1}} = N_{i-1}^{-1}$ из точки $i-1$ в точку i (обозначим матрицу как $\hat{K}_{\beta_{i-1}}$):

$$\hat{K}_{\beta_{i-1}} = \Phi_{i-1,i} N_{i-1}^{-1} \Phi_{i-1,i}^T.$$

Затем происходит Q -расширение ковариационной матрицы и её обращение:

$$K_{\beta_{i-1}^+}^{-1} = \left[\hat{K}_{\beta_{i-1}} + Q_{i-1} \right]^{-1}.$$

Далее производится усечение весовой матрицы новым измерением:

$$K_{\beta_i}^{-1} = K_{\beta_{i-1}^+}^{-1} + V_i^T K_{\psi_i}^{-1} V_i = N_i.$$

Вектор правых частей накапливается без использования величины β_{i-1}^+ (!):

$$Y_i = \hat{Y}_i + V_i^T K_{\psi_i}^{-1} \psi_i = K_{\beta_{i-1}^+}^{-1} \hat{K}_{\beta_{i-1}}^{-1} \Phi_{i,i-1}^T Y_{i-1} + V_i^T K_{\psi_i}^{-1} \psi_i.$$

Процедура определения оценки $\hat{\beta}_{i=k}$ совпадает с процедурой Гаусса:

$$\hat{\beta}_{i=k} = N_k^{-1} Y_{i=k}.$$

Упражнение 6. Докажите правильность выражения Y_i .

5. Рекурсивные фильтры Гаусса и Калмана–Бьюси

91. Гауссу принадлежит также и первое описание *рекурсивного* (итерационного) фильтра для линейной оценки с минимальным рассогласованием *постоянного* параметра α_i . Впоследствии аналогичный способ оценки *переменного* параметра $\alpha_i(t)$ с фазовыми шумами получил имя фильтра Калмана–Бьюси [10]. Он имеет вид

$$\alpha_k = \alpha_{k-1} + H(\psi_k - U\alpha_{k-1}). \quad (\text{Б.12})$$

Приведём ряд соображений, показывающих сущность и пределы применимости этого популярного метода обработки экспериментальных данных.

Связь рекурсивной процедуры фильтрации с методом релаксации

Интересна связь процедуры такой фильтрации с известным *релаксационным методом* [8, с. 206] для итерационного решения системы алгебраических линейных уравнений (в данном случае – последовательности условных уравнений).

92. Напомним, что для линейной системы типа

$$U\alpha = \psi, \quad (\text{Б.13})$$

где U – квадратная матрица, можно построить итерационную процедуру (*метод последовательных приближений* [8, с. 208]):

$$\alpha_k = \alpha_{k-1} + \delta\alpha_{k-1} = \alpha_{k-1} + \psi - U\alpha_{k-1} = B\alpha_{k-1} + \psi, \quad (\text{Б.14})$$

где $B = E - U$, E – единичная матрица, $\delta\alpha_k$ – невязка линейного уравнения при подстановке $\alpha = \psi - U\alpha_{k-1}$. Выражение (Б.14) получается из (Б.13) прибавлением и отниманием от левой части единичной матрицы:

$$(U + E - E)\alpha = \psi.$$

Отсюда следует

$$\alpha = \alpha - U\alpha + \psi = B\alpha + \psi.$$

Оценка погрешности последовательности итераций будет иметь вид

$$|\alpha_k - \hat{\alpha}| = |B|^n |\alpha_0 - \hat{\alpha}|.$$

Поэтому итерационная процедура будет сходиться, если определитель матрицы $|B| < 1$ (*сжимающее преобразование*).

Но если определитель $|B| \geq 1$, то для выполнения условий сходимости можно попытаться сделать преобразование сжимающим, применив умножение исходного уравнения слева и справа на *матрицу релаксации H*:

$$HU\alpha = H\psi.$$

Для получения сходящейся последовательности необходимо, чтобы было выполнено условие

$$|E - HU| < 1.$$

Тогда

$$\alpha_k = \alpha_{k-1} + H(\psi_k - U\alpha_{k-1}). \quad (\text{B.15})$$

И здесь нужно обратить внимание на то, что рекуррентное выражение (B.15) совпадает с формулой (B.12) фильтра Калмана.

93. А с точки зрения способа управления процессом сходимости матрица релаксации H играет роль размерного *коэффициента обратной связи* при рассогласовании измерения с предвычисленным значением его модели.

94. В методе релаксации для итерационного решения систем линейных уравнений матрица релаксации H выбирается так, чтобы $|E - HU| < 1$. При этом если α и ψ – векторы, то H – матрица. Если α – вектор, а ψ – скаляр, то H – вектор. И, наконец, если α и ψ – скаляры, то H – число. Рассмотрим для простоты последний случай.

95. Применяя этот метод к последовательности условных уравнений

$$u_i \alpha = \psi_i,$$

где α – оцениваемый постоянный (для простоты – скалярный) параметр, можно некоторым выбором *сжимающей последовательности чисел* H_k получить при $k = i$ упомянутую последовательность (B.12) оценок величины α :

$$\alpha_k = \alpha_{k-1} + H_k (\psi_k - u_k \alpha_{k-1}), \quad (\text{B.16})$$

таких, что при поступлении каждого очередного измерения ψ_k происходило бы смещение α_{k-1} к наиболее вероятному значению α_k .

Иными словами, фильтр Гаусса–Калмана представляет собой итерационный способ решения линейной системы последовательности несовместных условных уравнений с выбором матрицы релаксации H так, чтобы на каждой итерации при использовании условного уравнения $u_k \alpha = \psi_k$ для очередного измерения ψ_k получать наиболее вероятное смещение в пространстве определяемых параметров α_k .

Для иллюстрации особенностей этого процесса фильтрации измерений рассмотрим свойства некоторой специальной двумерной случайной величины, связанной с ошибками измерения и с ошибками оценки значений определяемых параметров (их "прогноза") при использовании линейной модели зависимости измерений ψ_k от параметров α_{k-1} .

Специфика двумерной случайной величины "прогноз–измерение"

96. Для требуемого выбора релаксационных чисел H_k , обеспечивающих наиболее вероятное смещение при поступлении нового измерения, рассмотрим двумерную нормально-распределённую случайную величину "прогноз–измерение", которую запишем в следующем виде:

$$[\psi_k(\alpha_{k-1}), \psi_k] = [X, Y],$$

где $\psi_k(\alpha_{k-1})$ – прогноз измерения в следующий момент k на основе значения α_{k-1} , ψ_k – реальное выполненное измерение. Обозначим их как X и Y соответственно. Пусть случайная величина $[X, Y]$ является центрированной на текущем k -м шаге процесса относительно своего математического ожидания $m([X, Y])$.

Рассмотрим, что известно о параметрах распределения этой случайной величины. Здесь наиболее уместна и интересна описанная выше геометрическая картина параметров рассеивания этой двумерной случайной величины, описываемой ковариационной матрицей

$$K = \begin{bmatrix} \sigma_X^2 & r\sigma_X\sigma_Y \\ r\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix},$$

где r – коэффициент корреляции между рассеиванием двух компонент двумерной случайной величины с их безусловными дисперсиями σ_X^2 и σ_Y^2 (т.е. дисперсиями рассеивания одной величины при всех значениях другой величины).

Определим параметры матрицы ковариаций "прогноз–измерение".

97. 1) Пусть в результате последовательной фильтрации измерений в момент t_{k-1} получена текущая оценка $K_{\alpha_{k-1}} = \sigma_{\alpha_{k-1}}^2$, то есть оценка диспер-

сии прогноза безусловного рассеивания параметра α_{k-1} (на первом шаге дана априорная оценка K_{α_0}). Тогда известна также и дисперсия $\sigma_Y^2 = \sigma_{\text{пр}}^2$ безусловного рассеивания прогноза измерения $X = \psi_k(\alpha_{k-1})$ для будущего момента t_k , равная

$$\sigma_{\text{пр}}^2 = u_k^2 K_{\alpha_{k-1}}. \quad (\text{Б.17})$$

Геометрический смысл этой величины, как было показано, заключается в том, что она является проекцией эллипса рассеивания двумерной случайной величины $[X, Y]$ на ось X (см. ниже рис. Б.3).

Итак, мы знаем величину проекции искомого эллипса на ось X .

98. 2) Возможное рассеивание самой наблюдаемой величины ψ_k в момент t_k описывается безусловной дисперсией $\sigma_{\text{изм}}^2$, которая получается при суммировании рассеивания величины $\psi_k(\alpha_{k-1})$ (то есть дисперсии $\sigma_{\text{пр}}^2$) и ожидаемой дисперсией погрешности измерений $\delta\psi_k$, априорно предполагаемой равной заданной постоянной величине K_ψ :

$$\sigma_Y^2 = \sigma_{\text{изм}}^2 = \sigma_{\text{пр}}^2 + K_\psi.$$

Заметим, что *ошибки измерения по правилу Байеса рассеиваются вокруг своего прогноза*, и, следовательно, априорно предполагаемое значение величины K_ψ по существу является значением величины условной дисперсии ошибок измерений при условии прогноза:

$$K_\psi = \sigma_{\text{изм|пр}}^2. \quad (\text{Б.18})$$

99. С другой стороны, ранее было сказано, что дисперсии условных рассеиваний компонент многомерной случайной величины геометрически изображаются координатными диаметрами эллипса, вписанного в прямоугольник со сторонами, равными безусловным дисперсиям этой случайной величины. Диаметр на оси X геометрически показывает условное рассеивание X при условии Y , а диаметр на оси Y показывает рассеивание Y при условии X .

Заметим, что именно здесь опять проявляется упомянутое свойство нормального распределения: *дисперсия условной вероятности отклонений случайной величины не зависит от условия*.

Именно поэтому и существует возможность априорной оценки ожидаемой точности определения искомой величины прогноза – до получения нового измерения.

Из сказанного следует, что рассеивание выполняемого измерения ψ_k вдоль оси Y определено величиной погрешности измерений K_ψ – размером центрального диаметра эллипса. Иными словами, нам *известен ещё*

один параметр искомого эллипса рассеивания: положение точки его пересечения с осью Y .

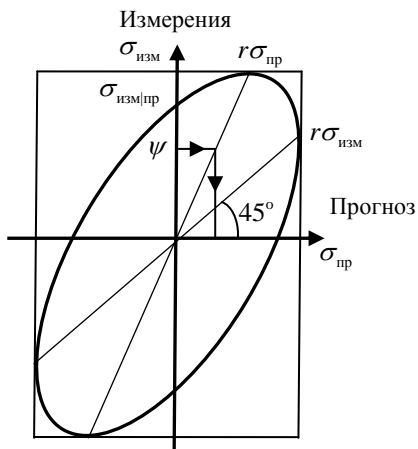


Рис. Б.3. Геометрическая интерпретация фильтра Калмана

Эллипс вероятности определен по трём параметрам:

$X = u\sqrt{K_\alpha} = \sigma_{\text{пр}}$, $\sqrt{K_\psi} = \sigma_{\text{изм|пр}}$ и по углу наклона линии регрессии на 45° . Это позволяет определить размер $Y = \sigma_{\text{изм}}$ и положение другой линии регрессии. После очередного измерения ψ наиболее вероятное значение параметра $\alpha(\psi)$ определяется по его положению в эллипсе на другой линии регрессии $X = \sigma_{\text{пр|изм}}$. Новое значение $\sigma_{\text{пр}}$ (сокращение ширины эллипса по оси X) связано лишь с фактом получения очередного измерения ψ и не зависит от значения величины ψ .

100. 3) Далее, используя правило Байеса, вспомним, что прогнозируемое значение $\psi_k(\alpha_{k-1})$ является математическим ожиданием будущего измерения ψ_k . Этим предположением задаются:

- наклон под углом 45° к оси $X = \psi_k(\alpha_{k-1})$ геометрического места точек условных математических ожиданий измерений (линии регрессии измерений – диаметра эллипса, сопряжённого оси $Y = \psi_k$), так как в этом случае

$$m_{\text{изм|пр}} = m_{\psi_k|\psi_k(\alpha_{k-1})} = \psi_k(\alpha_{k-1}); \quad (\text{Б.19})$$

- положение точек касания эллипсом вертикальных линий $Y = \pm\sigma_{\text{пр}}$ (см. рис. Б.3), определяемое в общем случае расстоянием по

вертикали от оси X , равным $r\sigma_{\text{изм}}$, что в данном случае приводит к соотношению

$$r\sigma_{\text{изм}} = \sigma_{\text{пр}}. \quad (\text{Б.20})$$

101. Таким образом, нам известны 3 параметра, полностью определяющие искомый эллипс (у эллипса 3 свободных параметра):

- его проекция на ось X ,
- точка пересечения с осью Y
- и точки его касания вертикальной стороны описанного прямоугольника.

Итак, с помощью выражений (Б.17), (Б.18), (Б.19) и (Б.20) полностью определяются элементы искомой ковариационной матрицы двумерной случайной величины "прогноз–измерение" $[\psi_k(\alpha_{k-1}), \psi_k]$:

$$\begin{cases} \sigma_X^2 = \sigma_{\text{пр}}^2 = u_k K_{\alpha_{k-1}} u_k, \\ \sigma_Y^2 = \sigma_{\text{изм}}^2 = \sigma_{\text{пр}}^2 + K_{\psi_{k-1}}, \\ r^2 = \sigma_{\text{пр}}^2 / \sigma_{\text{изм}}^2. \end{cases} \quad (\text{Б.21})$$

Рекурсивный фильтр Гаусса

102. В силу условия 1) получаем, что

$$m(\psi_k - \psi_k(\alpha_{k-1})) = m(\text{изм} - \text{пр}) = 0,$$

что отражает независимость ошибок измерений от значений параметра α_{k-1} .

103. В силу условия 2) проекция эллипса на ось $Y = \psi_k$ равна гипотенузе прямоугольного треугольника, образуемого диаметром эллипса по оси Y (ошибками измерений ψ_k) и горизонтальной проекцией эллипса (ошибками прогноза $\psi_k(\alpha_{k-1})$).

По мере уточнения величины $m(\alpha_k)$ горизонтальные размеры прямоугольника, в который вписывается эллипс рассеивания двумерной случайной величины $[\psi_k(\alpha_{k-1}), \psi_k]$, стремятся к нулю как корень квадратный из числа измерений (для принятой модели осреднения чисто случайного шума), и эллипс стремится к вертикальному отрезку с длиной:

$$\sigma_X^2 = \sigma_{\text{изм}}^2 = \sigma_{\text{пр}}^2 + K_{\psi}.$$

104. Наиболее вероятное значение α_{k-1} после реализации очередного измерения ψ_k определяется условным математическим ожиданием $m_{\text{пр|изм}}$, то есть принадлежит диаметру эллипса, сопряжённому горизонтальному осевому диаметру и соединяющему центр эллипса с его точками:

$$[X, Y] = \pm[X = r\sigma_{\text{пр}}, Y = \sigma_{\text{изм}}].$$

Эти точки являются точками касания горизонтальных сторон описанного прямоугольника (см. рис. Б.3). Уравнение в центрированных относительно $m([\psi_k(\alpha_{k-1}), \psi_k])$ осях такой линии регрессии прогноза $\psi_k(\alpha_{k-1})$ при условии измерения ψ_k связывает наиболее вероятное значение α_k с искомой величиной параметра релаксации H_k :

$$m_{\text{пр|изм}} - m(\psi_k(\alpha_{k-1})) = r \cdot (\sigma_{\text{пр}} / \sigma_{\text{изм}}) \cdot (\psi_k - m(\psi_k)). \quad (\text{Б.22})$$

Но, используя (Б.20), (Б.21), имеем

$$r(\sigma_{\text{пр}} / \sigma_{\text{изм}}) = \sigma_{\text{пр}}^2 / \sigma_{\text{изм}}^2,$$

$$\psi_k - m(\psi_k) = \psi_k - u_k \alpha_{k-1},$$

$$m_{\text{пр|изм}} - m(\psi_k(\alpha_{k-1})) = u_k (\alpha_k - \alpha_{k-1}).$$

И тогда, используя (Б.22), получаем *наиболее вероятное* соотношение между α_k и ψ_k (или в более общем случае – α_k и Ψ_k):

$$\frac{u_k K_{\alpha_{k-1}} u_k^T}{u_k K_{\alpha_{k-1}} u_k^T + K_{\psi_k}} \cdot (\psi_k - u_k \alpha_{k-1}) = u_k (\alpha_k - \alpha_{k-1}),$$

которое удовлетворяется, если положить

$$\alpha_k = \alpha_{k-1} + \frac{K_{\alpha_{k-1}} u_k^T}{u_k K_{\alpha_{k-1}} u_k^T + K_{\psi_k}} \cdot (\psi_k - u_k \alpha_{k-1}).$$

105. Отсюда получаем, что искомое выражение для матрицы релаксации выглядит следующим образом:

$$H_k = \frac{K_{\alpha_{k-1}} u_k^T}{u_k K_{\alpha_{k-1}} u_k^T + K_{\psi_k}}.$$

При этом новая оценка K_{α_k} из условия

$$\sigma_{\text{пр|изм}}^2 = (1-r)^2 \sigma_{\text{пр}}^2$$

в общем случае равна

$$K_{\alpha_k} = \left[E - K_{\alpha_{k-1}} u_k^T (u_k K_{\alpha_{k-1}} u_k^T + K_{\psi_k})^{-1} u_k \right] K_{\alpha_{k-1}} = [E - H_k u_k] K_{\alpha_{k-1}}.$$

106. Таким образом, степенью формального улучшения (*усечения измерения*) оценки погрешности определения параметра α является величина $[E - H_k u_k]$, которая при постоянных значениях u_k и K_{ψ_k} сводится к значению \sqrt{k} и не зависит от значений получаемых измерений.

107. Из сказанного следует, что в векторном случае определения постоянных параметров вектора \mathbf{a} рекурсивный фильтр Гаусса выражается следующими соотношениями:

$$\begin{aligned}\mathbf{a}_k &= \mathbf{a}_{k-1} + H_k (\boldsymbol{\Psi}_k - U_k \mathbf{a}_{k-1}), \\ H_k &= K_{\mathbf{a}_{k-1}} U_k^T (U_k K_{\mathbf{a}_{k-1}} U_k^T + K_{\boldsymbol{\Psi}_k})^{-1}, \\ K_{\mathbf{a}_k} &= [E - H_k U_k] K_{\mathbf{a}_{k-1}}.\end{aligned}\tag{Б.21}$$

Именно в такой форме его и предложил Гаусс. Здесь предполагается последовательная обработка элементов выборки измерений $\boldsymbol{\Psi}_k$ (что оправдывает название "фильтр") и обращение матрицы на каждом шаге (в одномерном случае это заменено операцией деления).

Фильтр Калмана–Бьюси

108. Для решения задач с переменными определяемыми параметрами наиболее известным фильтром измерений является фильтр Калмана–Бьюси [10], предложенный в 1961 г. (обычно называемый просто "Фильтр Калмана").

109. Если предположить, что параметры \mathbf{a} являются переменными величинами, определяемыми в каждый момент измерения t_k и подчиняющимися линейной модели процесса с фазовыми и измерительными шумами:

$$\begin{aligned}\mathbf{a}_k &= \Phi_{k-1,k} \mathbf{a}_{k-1} + \mathbf{q}_k, \\ \boldsymbol{\Psi}_k &= U_k \mathbf{a}_k + \boldsymbol{\delta}_k,\end{aligned}$$

то наиболее полный рекурсивный фильтр в виде фильтра Калмана–Бьюси описывается в виде следующего алгоритма ("+" обозначает прогноз до измерения):

1. Линейный $\Phi_{k-1,k}$ перенос предыдущей оценки \mathbf{a}_{k-1} в точку k :

$$\mathbf{a}_k^+ = \Phi_{k-1,k} \mathbf{a}_{k-1}.$$

2. Линейный $\Phi_{k-1,k}$ перенос и Q -расширение ковариационной матрицы

$$K_{\mathbf{a}_k}^+ = \Phi_{k-1,k} K_{\mathbf{a}_{k-1}} \Phi_{k-1,k}^T + Q_k.$$

3. Вычисление матрицы релаксации:

$$H_k = K_{\mathbf{a}_k}^+ U_k^T (U_k K_{\mathbf{a}_k}^+ U_k^T + K_{\boldsymbol{\Psi}_k})^{-1}.$$

4. Вычисление новой оценки:

$$\mathbf{a}_k = \mathbf{a}_k^+ + H_k (\boldsymbol{\Psi}_k - U_k \mathbf{a}_k^+).$$

5. Вычисление $\boldsymbol{\Psi}$ -усечения ковариационной матрицы в новой точке:

$$K_{\alpha_k} = [E - H_k U_k] K_{\alpha_k}^+ .$$

Таким образом, фильтр Калмана отличается от фильтра Гаусса лишь процедурами линейного переноса в следующую точку с её номером k результатов в предыдущей точки $k-1$ оценки значений α_{k-1}^+ и ковариационной матрицы $K_{\alpha_{k-1}}^+$.

Для простого случая фильтрации единственного постоянного параметра α при получении очередного измерения ψ_k эти соотношения имеют вид:

$$\begin{aligned} \alpha_k^+ &= \Phi_{k-1,k} \alpha_{k-1}^+, \\ \sigma_{\alpha_k}^{2+} &= \Phi_{k-1,k}^2 \cdot \sigma_{\alpha_{k-1}}^2 + Q_k, \\ H_k &= \sigma_{\alpha_k}^{2+} u_k / (u_k^2 \sigma_{\alpha_k}^{2+} + K_{\psi_k}), \\ \alpha_k &= \alpha_k^+ + H_k (\psi_k - u_k \alpha_k^+), \\ K_{\alpha_k} \sigma_{\alpha_k}^2 &= [E - H_k u_k] K_{\alpha_k}^+, \end{aligned}$$

110. Заметим, что важными ограничениями для использования описанных методов фильтрации являются предположения о независимости (кусочной независимости) погрешностей измерений и линейности связи фазового пространства определяемых параметров с пространством измерений. Иными словами:

Ковариационные матрицы K_{ψ_k} должны быть диагональными, а матрицы градиентов U в условных уравнениях не должны зависеть от величин α и ψ .

Связь рекурсивного фильтра и способа наименьших квадратов

111. Важное утверждение.

Рекурсивные фильтры Гаусса и Калмана представляют собой тот же СНК, но записанный через ковариационные (а не весовые) матрицы.

112. Свойство сводимости друг к другу формул осреднения фильтрами Гаусса–Калмана и способа наименьших квадратов (или метода максимума правдоподобия – при недиагональных матрицах погрешностей) логически вытекает из следующей посылки.

Оба метода (каждый по своему) обеспечивают *наиболее вероятную оценку* компонент вектора определяемых параметров α_j в рамках одного и того же закона распределения вероятностей и при одних и тех же исходных предпосылках. Свойство единственности максимума вероятности

нормального закона не позволяет ожидать, что эти наиболее вероятные значения определяемого вектора окажутся разными.

113. Как это доказать аналитически?

Решение нормальной системы имеет вид

$$\mathbf{a}_j = \left[U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1} \right]^{-1} (U_j^T K_{\Psi_j}^{-1} \Psi_j + K_{\mathbf{a}_{j-1}}^{-1} \mathbf{a}_{j-1})$$

и

$$K_{\mathbf{a}_j}^{-1} = \left[U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1} \right].$$

Аналогично этому, фильтрация по Гауссу или Калману выражается в виде

$$\mathbf{a}_j = \mathbf{a}_{j-1} + H (\Psi_j - U_j \mathbf{a}_{j-1})$$

и

$$K_{\mathbf{a}_j} = \left[E - H_j U_j \right] K_{\mathbf{a}_{j-1}},$$

где

$$H_j = K_{\mathbf{a}_{j-1}} U_j^T \left(U_j K_{\mathbf{a}_{j-1}} U_j^T + K_{\Psi_j} \right)^{-1}.$$

Для того чтобы убедиться в идентичности этих методов осреднения, нужно доказать, что

$$\mathbf{a}_j = \left[U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1} \right]^{-1} (U_j^T K_{\Psi_j}^{-1} \Psi_j + K_{\mathbf{a}_{j-1}}^{-1} \mathbf{a}_{j-1}) = \mathbf{a}_{j-1} + H (\Psi_j - U_j \mathbf{a}_{j-1})$$

и что

$$K_{\mathbf{a}_j} = \left[E - H_j U_j \right] K_{\mathbf{a}_{j-1}} = \left[U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1} \right]^{-1}.$$

114. Эти выражения приводятся друг к другу с помощью полезного тождества. Для любых матриц A, B, C

$$\left(A + CBC^T \right)^{-1} \equiv A^{-1} - A^{-1} C \left(B^{-1} + C^T A^{-1} C \right)^{-1} C^T A^{-1},$$

и при этом для симметричных матриц A и C при $A = C$

$$\left(A + ABA \right)^{-1} \equiv A^{-1} - \left(B^{-1} + A \right)^{-1}.$$

115. Эти преобразования позволяют получить выражение матрицы релаксации

$$H_j = K_{\mathbf{a}_{j-1}} U_j^T \left(U_j K_{\mathbf{a}_{j-1}} U_j^T + K_{\Psi_j} \right)^{-1},$$

обеспечивающее требуемое доказательство обоих утверждений.

Задача. Для полного овладения предлагаемым материалом настоятельно рекомендуем читателю самому попытаться доказать как упомянутое тождество, так и факт взаимной сводимости формул осреднения методом нормальных мест и фильтров Гаусса–Калмана. Если же не получится, см. Приложение.

Построение линейного приближения нелинейной системы

116. Заметим, что большинство динамических систем и сенсорных измерений не являются линейными системами. Но тем не менее методы фильтрации, разработанные для линейных систем, в ряде случаев могут быть применены и для нелинейных систем. Это связано с тем, что на каждом шаге фильтрации (аналогично шагам любого метода итерации) отклонение величины фильтруемого (итерируемого) параметра \mathbf{a} от принятого значения первого приближения \mathbf{a}_0 достаточно хорошо известно из значений на прошлом шаге.

Метод, позволяющий в этом предположении использовать фильтрацию Калмана, сводится к *хордовой линеаризации* нелинейной зависимости. Заметим, что на границах линейной окрестности хордовые производные правильнее описывают ожидаемое рассеивание. *При численном вычислении производных отклонения аргумента всегда надо брать порядка ожидаемых возмущений.*

117. Итак, пусть область рассеивания фазовых переменных, определяющая истинные значения измерений, выходит за границы области линейности градиентов измерений в фазовом пространстве, *но это отклонение примерно известно* из предыдущего шага фильтра.

Предположим, что рассматриваемый стохастический процесс может быть представлен нелинейными динамическими уравнениями и соответственно нелинейными условными уравнениями, описывающими модельные зависимости полученных значений измерений $\boldsymbol{\psi}$ от фазовых параметров \mathbf{a} (см. Таблицу 1).

Таблица 1

Модель	Непрерывное время	Дискретное время
Система	$\dot{\mathbf{a}}(t) = \mathbf{f}(\mathbf{a}(t), t) + \mathbf{q}(t)$	$\mathbf{a}_k = \mathbf{f}(\mathbf{a}_{k-1}, k-1) + \mathbf{q}_{k-1}$
Измерения	$\boldsymbol{\psi}(t) = \mathbf{u}(\mathbf{a}(t), t) + \mathbf{v}(t)$	$\boldsymbol{\psi}_k = \mathbf{u}(\mathbf{a}_k, k) + \mathbf{v}_k$

118. Обозначим символом δ ожидаемое отклонение вектора \mathbf{a} от оценки:

$$\delta \mathbf{a}_k = \mathbf{a}_k - \hat{\mathbf{a}}_k \text{ и } \delta \boldsymbol{\psi}_k = \boldsymbol{\psi}_k - \mathbf{u}(\hat{\mathbf{a}}_k, k).$$

Тогда $\mathbf{f}(\mathbf{a}, k-1)$ в окрестности $\mathbf{a} = \mathbf{a}_{k-1}^+$ может быть представлена в виде

$$\mathbf{a}_k = \mathbf{f}(\mathbf{a}_{k-1}, k-1) = \mathbf{a}_k^+ + \left. \frac{\partial \mathbf{f}(\mathbf{a}, k-1)}{\partial \mathbf{a}} \right|_{\mathbf{a}=\hat{\mathbf{a}}_{k-1}} \delta \mathbf{a}_{k-1} + \frac{1}{2} \cdot \left. \frac{\partial^2 \mathbf{f}(\mathbf{a}, k-1)}{\partial \mathbf{a}^2} \right|_{\mathbf{a}=\hat{\mathbf{a}}_{k-1}} (\delta \mathbf{a}_{k-1})^2.$$

Выразим величину *хордовой производной* как

$$\Phi_{k-1,k} = \left[\left. \frac{\partial \mathbf{f}(\mathbf{a}, k-1)}{\partial \mathbf{a}} \right|_{\mathbf{a}=\hat{\mathbf{a}}_{k-1}} \delta \mathbf{a}_{k-1} + \frac{1}{2} \cdot \left. \frac{\partial^2 \mathbf{f}(\mathbf{a}, k-1)}{\partial \mathbf{a}^2} \right|_{\mathbf{a}=\hat{\mathbf{a}}_{k-1}} (\delta \mathbf{a}_{k-1})^2 \right] / \delta \mathbf{a}_{k-1}.$$

Тогда имеем линейное приближение требуемой переходной функции для прогнозирования изменения фазового вектора на шаге процесса измерений:

$$\delta \mathbf{a}_k \approx \Phi_{k-1,k} \delta \mathbf{a}_{k-1} + \mathbf{q}_{k-1}.$$

119. Аналогично для построения условных уравнений измерений положим

$$U_k = \left[\frac{\partial \mathbf{h}(\mathbf{a}, k)}{\partial \mathbf{a}} \Big|_{\mathbf{a}=\hat{\mathbf{a}}_k} \delta \mathbf{a}_k + \frac{1}{2} \cdot \frac{\partial^2 \mathbf{h}(\mathbf{a}, k)}{\partial \mathbf{a}^2} \Big|_{\mathbf{a}=\hat{\mathbf{a}}_k} (\delta \mathbf{a}_k)^2 \right] / \delta \mathbf{a}_k.$$

Возмущения $\delta \boldsymbol{\psi}_k$ могут быть представлены в линейном виде:

$$\delta \boldsymbol{\psi}_k = \mathbf{U}_k \delta \mathbf{a}_k.$$

В. ОЦЕНКА ТОЧНОСТИ ЗНАНИЯ ПАРАМЕТРОВ МОДЕЛИ ПОСЛЕ ОСРЕДНЕНИЯ ИЗМЕРЕНИЙ

120. После вычисления значения искомого вектора параметров α каким-либо методом осреднения полученных измерений важно получить также и оценку уровня доверия к полученному значению искомого вектора.

121. Следует отметить, что рассматриваемая задача относится к классу трудно формализуемых задач. Причина этого обстоятельства в том, что доступный формализм её решения не учитывает главных причин возникновения ошибок, количество и вероятностный смысл которых могут быть определены лишь в формализме игровых задач.

Иными словами, определение точности полученного знания предполагает минимаксную постановку задачи игры с природой возникновения возможных погрешностей. Это именно класс задач "игр с природой", определяемый известным тезисом, что "Бог изощрён, но не злонамерен" – с отсутствием формализма полного описания этой "изощрённости".

122. К настоящему времени разработан ряд алгоритмов оценивания точности результата осреднения измерений. *Для полной уверенности нужно выполнить апостериорный анализ характера рассеивания измерений (см. рис. Б.1) и анализ результатов применения всех доступных методов оценивания!*

123. В этом разделе будут описаны следующие методы получения надёжной оценки.

1. Формальная оценка.
2. Расширенная оценка.
 - i. Учёт систематических ошибок погрешностей измерений.
 - ii. Модель наихудшей корреляции погрешностей измерений.
 - iii. "Схема бортика".
 - iv. Учёт мешающих параметров.
3. Метод разделения определяемых и оцениваемых параметров.

1. Формальная оценка

124. Решение нормальной системы

$$N\alpha = [U^T K_{\psi}^{-1} U + K_{\alpha_0}^{-1}] \alpha = U^T K_{\psi}^{-1} \psi + K_{\alpha_0}^{-1} \alpha_0$$

после обращения матрицы её коэффициентов позволяет получить не только оценку вектора определяемых параметров

$$\hat{\alpha} = N^{-1} (U^T K_{\psi}^{-1} \psi + K_{\alpha_0}^{-1} \alpha_0),$$

но и *оценить погрешность* полученного значения компонент вектора $\hat{\alpha}$.

Ранее уже упоминалось, что используемая для получения оценки средних значений определяемых параметров обратная матрица нормальных уравнений

$$N^{-1} = [U^T K_{\psi}^{-1} U + K_{\alpha_0}^{-1}]^{-1}$$

формально является и оценкой точности определения параметров α при осреднении измерений ψ , поскольку она равна искомой ковариационной матрице K_{α} :

$$K_{\alpha} = N^{-1}.$$

125. Заметим, что слово *оценить* было выделено выше специально. Дело в том, что эта популярная часто применяемая оценка полученной точности знания параметров α не учитывает многих факторов, а при большом числе усредняемых величин может привести к заведомо неверным выводам.

И причина этого в том, что *формально* получаемая оценка является следствием преобразований, выполненных в предположении о присутствии только *чисто случайных* погрешностях полученных измерений. Поэтому она определяет лишь величину остаточной погрешности осреднения полученного количества m случайных ошибок, которое иногда достаточно большое.

Иными словами, получаемая ковариационная матрица K_{α} определяет лишь формальное уменьшение в \sqrt{m} раз исходных размеров только случайных погрешностей выполненных измерений.

126. Формальное равенство $K_{\alpha} = N^{-1}$ легко доказывается простым преобразованием квадратичных форм способа наименьших квадратов.

Выразим для этого связь переменных α , ψ и α_0 , получаемую в результате осреднения, в виде линейного выражения с некоторыми матричными коэффициентами A и B :

$$\hat{\alpha} = N^{-1} [U^T K_{\psi}^{-1} \psi + K_{\alpha_0}^{-1} \alpha_0] = A\psi + B\alpha_0,$$

где

$$A = N^{-1} U^T K_{\psi}^{-1}, \quad B = N^{-1} K_{\alpha_0}^{-1}. \quad (B.1)$$

Тогда ковариационная матрица параметра α записывается в виде

$$K_{\alpha} = AK_{\psi}A^T + BK_{\alpha_0}B^T.$$

Подставляя соответствующие выражения для A и B , имеем

$$\begin{aligned} K_{\alpha} &= N^{-1} U^T K_{\psi}^{-1} K_{\psi} (N^{-1} U^T K_{\psi}^{-1})^T + N^{-1} K_{\alpha_0}^{-1} K_{\alpha_0} (N^{-1} K_{\alpha_0}^{-1})^T = \\ &= N^{-1} [U^T K_{\psi}^{-1} U + K_{\alpha_0}^{-1}] N^{-1} = N^{-1} N N^{-1} = N^{-1}. \end{aligned}$$

127. Таким образом, очень удобная (и крайне опасная) эта формальная оценка получается непосредственно в процессе решения нормальной системы уравнений. Другими словами, в процессе осреднения данной совокупности условных уравнений ковариационная матрица ошибок осреднения оказывается равной получаемой обратной матрице нормальной системы уравнений способа наименьших квадратов или метода максимума правдоподобия.

При использовании фильтров ковариационная матрица остаточных ошибок получается непосредственно в процессе фильтрации и совпадает с формальной оценкой способа наименьших квадратов.

Получаемая величина такой формальной оценки обратно пропорциональна корню квадратному из числа измерений.

Например, формальная оценка точности знания координат посадки на Марс по полному числу измерений наклонных дальностей при полёте Земля–Марс не превышает 1 мм(!), что, конечно, не соответствует реальной точности оценки.

2. Расширенная оценка

128. Недостатком формальной оценки точности является использование при оценивании погрешностей обработки измерений *той же самой гипотезы* о погрешностях измерений, что и гипотеза, принимаемая в методе их обработки.

129. Поэтому более правильной является *расширенная* оценка вида

$$K_a = [U^T K_\psi^{-1} U + K_{a_0}^{-1}]^{-1} \{U^T K_\psi^{-1} (K_\psi + L K_\mu L^T) K_\psi^{-1} U + K_{a_0}^{-1}\} \left([U^T K_\psi^{-1} U + K_{a_0}^{-1}]^{-1} \right)^T,$$

или в более простом способе записи с использованием (В.1):

$$K_a = A (K_\psi + L K_\mu L^T) A^T + B K_{a_0} B^T.$$

Здесь K_ψ , K_{a_0} и K_μ – ковариационные матрицы погрешностей измерений ψ , априорных значений параметров a_0 и погрешностей модели (*мешающих параметров*) μ . Эти апостериорные матрицы отвечают более широкой статистической гипотезе, чем та, которая была использована в методе обработки.

Учёт систематических ошибок погрешностей измерений

130. Наиболее очевидным расширением статистической гипотезы о погрешностях измерений ψ является отказ от модели только случайных ошибок с предположением наличия систематических ошибок – постоянных сдвигов значений измерений от их верной величины. Такие ошибки

не усредняются и поэтому формальная оценка при систематических погрешностях измерений неверна.

Известен ряд методов оценки влияния систематических ошибок на точность прогнозирования параметров траектории в космосе.

131. Первая неформальная оценка достижимой точности определения искоемых параметров была получена в 1959 г. Т.М. Энеевым совместно с Н.И. Золотухиной на примере задачи о точности прогнозирования движения космического аппарата.

Если модель движения известна (например, рассматривается возможное множество эллиптических гелиоцентрических траекторий), то постоянные сдвиги измеряемых величин в общем случае противоречат закону движения. В связи с этим влияние постоянных систематических ошибок δ_i может быть оценено путём выбора моментов измерения и выбора эллиптической траектории, некоторым образом (например, в смысле наименьших квадратов) приближенной к полученным значениям измерений.

Для получения оценки достижимой точности знания орбиты надо в предположении наихудшего влияния погрешности измерения в каждой выбранной точке траектории наилучшим образом расположить предполагаемые ошибочные измерения на определяемой траектории движения.

132. Тогда суммарная погрешность прогнозирования движения $\Delta\rho$ определяется выражением

$$\Delta\rho = \min_{\text{изм}} \max_{\delta_i} \sum_{i=1}^6 \frac{\partial\rho}{\partial\delta_i},$$

где $\frac{\partial\rho}{\partial\delta_i}$ определяются в предположениях выбранного способа определения траектории по измерениям $\psi_i + \delta_i$.

133. Задача учёта не случайных погрешностей измерений в описанной минимаксной постановке была решена численно на примере траектории полёта к планете Марс в 1960 г. В процессе её решения неожиданно стало понятным, что наиболее опасными являются отнюдь не постоянные систематические ошибки измерений, а такие *плавающие* их погрешности, которые могут оказаться не противоречащими используемой модели движения.

Например, в кеплеровом приближении одной из моделей геоцентрической траектории межпланетного полёта является гипербола, и пусть так случилось, что ошибки траекторных измерений приводят к оценке расстояний космического аппарата от Земли, имеющей тоже некий гиперболический характер. В этом случае вместо искомого *определения правильных значений* параметров движения α в качестве ответа будет неявно выполнена лишь мало интересная *аппроксимация плавающих* ошибок изме-

рений подходящей моделью некоторой фиктивной траектории движения. И при этом будет получено (!) наилучшее значение формальной оценки точности определения вектора параметров α .

134. Интересно, что когда это соображение было обнаружено, то разработчики системы траекторных радиоизмерений, которые до этого утверждали, что возможны лишь постоянные систематические сдвиги значений измерений, сразу нашли возможный источник любых плавающих ошибок: переменная (из-за солнечных вспышек) концентрация ионов в столбе прохождения радиолуча. После этого заметно возросла практическая значимость астрономической службы Солнца и исследований солнечного ветра.

135. Простейший способ оценки величины погрешности прогноза в модели систематических ошибок применительно к задаче полёта к Луне был в 1959 г. предложен М.Л. Лидовым во время обсуждения рассматриваемой проблемы с автором данных лекций. Этот способ заключается в следующем.

Пусть число измерений $m \rightarrow \infty$ ($1/\sqrt{m} \rightarrow 0$). Тогда искомая погрешность оценки определяемых параметров зависит только от неизвестных систематических ошибок. Поскольку любые искомые параметры движения (например, координаты в картинной плоскости Луны) зависят от параметров орбиты, то размерность вектора определяемых параметров не может быть меньше шести (6 элементов орбиты или 6 параметров траектории полёта).

Поскольку минимальное требуемое число наблюдений не может быть меньше числа определяемых параметров, то следует выбрать 6 предельно возможных значений систематических погрешностей измерений. Затем, предполагая нормальное распределение вероятностей знания истинной величины систематических ошибок, следует построить эллипс рассеивания параметров орбиты и следующим линейным преобразованием определить большую полуось соответствующего эллипса в картинной плоскости Луны.

136. В описываемом случае для приближённой оценки влияния систематической ошибки был выбран простейший вариант: шесть измерений дальности от центра Земли до космического аппарата были равномерно расположены на выбранном интервале наблюдения вдоль траектории полёта к Луне с последним измерением вблизи Луны.

Задача. Интересной курсовой НИР является задача определения наилучших расположений шести систематических ошибок для измерений каждого состава (наклонные дальности, углы, доплеровские скорости от каждого из пунктов наблюдения) с последующим суммированием получаемых величин их влияния на прогнозируемые параметры траектории полёта космического аппарата.

137. Аналитически процедура расчёта выглядит следующим образом. Пусть априорная информация об ошибках траектории описывается в удобной системе координат шестимерной ковариационной матрицей K_{r_0} . Для простоты будем считать, что прямым образом измеряется сама очередная координата.

Тогда *усечение* этого знания более точным измерением с погрешностью σ_k какой-либо k -й компоненты этой системы координат описывается выражением $K_r = [K_{r_0}^{-1} + G_k^{-1}]^{-1}$.

Здесь матрица G_k^{-1} является весовой матрицей измерения k -й координаты с ошибкой σ_k , состоящей из нулевых клеток везде, кроме единственного k -го диагонального элемента, равного $1/\sigma_k^2$:

$$G_k^{-1} = \begin{bmatrix} 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_k^{-2} & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix}.$$

Если последовательно или одновременно измеряются несколько координат, то полная матрица G_k^{-1} содержит несколько ненулевых диагональных элементов.

Итак, пусть в результате преобразований исходной весовой матрицы ошибок измерений K_ψ^{-1} (она обычно диагональная) для шести выбранных моментов на траектории полёта космического аппарата с использованием соответствующей этим моментам квадратной матрицы U^{-1} получена *матрица измерений* определяемых параметров $K_{\psi_r}^{-1}$. Поскольку её диагональные элементы представляют собой обратные величины безусловных погрешностей измерений каждой отдельной координаты и поскольку в *официальных* исходных данных систематические и случайные погрешности обычно просуммированы, то, полагая их только систематическими, для получения матрицы G_k^{-1} можно использовать преобразование с матрицей Δ_k , заполненной нолями везде, кроме диагональной единицы на k -м месте:

$$G_k^{-1} = \Delta_k K_{\psi_r}^{-1} \Delta_k, \Delta_k = \begin{bmatrix} 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix}.$$

В результате на k -м месте диагонали мы получим вес $1/\sigma_k^2$ безусловного рассеивания измерения k -й координаты при любых значениях измерений других координат и при всех остальных элементах матриц, равных нулю.

Полная шестимерная матрица G^{-1} получается суммированием по k матриц G_k^{-1} , что даёт диагональную матрицу безусловного рассеивания.

Практически здесь описан аналитический приём выделения диагональной матрицы из недиагональной исходной матрицы.

138. Таким образом, оценка остаточной погрешности от совместного влияния случайных и неслучайных ошибок (типа систематических) выполняется достаточно простым отображением шестимерной матрицы $K_r = [K_{r_0}^{-1} + G_k^{-1}]^{-1}$ в пространство прогнозируемых параметров \mathbf{p} (на пример, координат в картинной плоскости Луны) преобразованием вида

$$K_p = DK_r D^T,$$

где $D = \frac{\partial \mathbf{p}}{\partial \mathbf{r}}$ – матрица преобразования координат точки последнего измерения в координаты картинной плоскости траектории полёта к Луне.

139. Наиболее строгая модель учёта систематических ошибок – в форме линейного фильтра – была описана в 1964 г. сотрудником Bell Telephone Laboratory А.Дж. Клаусом [12].

Положим, что m -размерный вектор ошибок траекторных наблюдений $\delta\psi$ содержит две составляющие: случайную статистически независимую ошибку (шум измерений) $\boldsymbol{\varepsilon}$ и результат смещения этих m измерений из-за влияния на измерение r неизвестных систематических отклонений $\delta\mathbf{v}$ параметров \mathbf{v} от их номинальных значений:

$$\delta\psi = \boldsymbol{\varepsilon} + S\delta\mathbf{v}.$$

Здесь строки матрицы чувствительности S (ее размерность – $r \times m$) содержат компоненты градиентов влияния параметров \mathbf{v} на измерения ψ :

$$S = \left[\frac{\partial \psi}{\partial \mathbf{v}} \right].$$

Например, если r соответствует числу типов измерений, то матрица S – единичная. Но если присутствуют ошибки модели измерений (мешающие параметры – см. ниже), то матрица S может отличаться от единичной.

Тогда новая ковариационная матрица ошибок измерений K_ψ содержит сумму двух ковариационных матриц – ковариационной матрицы случайных ошибок ε (обозначим ее как K_ε) и ковариационной матрицы результирующего влияния на измерения ψ неизвестных (и только поэтому предполагаемых случайными) систематических ошибок \mathbf{v} с их ковариационной матрицей K_v :

$$K_\psi = K_\varepsilon + SK_v S^T.$$

140. У такой матрицы K_ψ есть одно крайне неприятное свойство: ввиду наличия члена $SK_v S^T$ может оказаться практически невозможным вычислить нужную для осреднения весовую обратную матрицу K_ψ^{-1} .

Действительно, если $r \ll m$ и $\varepsilon \ll \mathbf{v}$, то матрица $SK_v S^T$ – почти вырожденная, поскольку влияние малого числа систематических ошибок на большое число измерений приводит к жёсткой r -корреляционной связи среди m -многих значений измерений. Если систематические погрешности превалируют над случайными погрешностями (что рано или поздно случается в процессе осреднения), то матрица K_ψ^{-1} становится трудно определимой.

Например, в случае m измерений со средней квадратической ошибкой σ и одинаковым для всех измерений средним квадратическим систематическим сдвигом γ новая ковариационная матрица ошибок может быть выражена в виде

$$K_\psi = \sigma^2 E + \gamma I.$$

Здесь E – единичная матрица, а I – не обратимая матрица, состоящая из одних единиц: все коэффициенты корреляции ошибок измерений в случае единственной систематической ошибки равны единице. При $\sigma = 0$ матрица K_ψ^{-1} не существует.

141. Если K_ψ^{-1} существует, то находим, как и прежде, оценку

$$\alpha = N^{-1} \left(U^T K_\psi^{-1} \psi + K_{\alpha_0}^{-1} \alpha_0 \right),$$

где $N = [U^T K_{\psi}^{-1} U + K_{\alpha_0}^{-1}]$ – матрица коэффициентов нормальной системы. В результате осреднения измерений $\delta\psi = \varepsilon + S\delta\mathbf{v}$ мы получим формальную оценку осреднённых по $\delta\psi$ параметров α в виде

$$K_{\alpha} = [U^T K_{\psi}^{-1} U + K_{\alpha_0}^{-1}]^{-1}.$$

142. Эта оценка при преобладании систематических ошибок ($\varepsilon \ll S\delta\mathbf{v}$) представляет собой попытку оценить рассеивание оценки вектора α из-за присутствия в $\delta\psi$ неизвестных систематических ошибок с известной ковариационной матрицей. При этом очевидно, что сама оценка *не будет смещенной*, только в случае, если *сдвиги измерений от систематических ошибок противоречат* модели зависимости ψ от α , выражаемой матрицей условных уравнений U . В противном случае будет выполнен подбор параметров траектории, объясняющий в среднем сдвиг измерений из-за присутствия систематических ошибок (этот случай подробно исследуется в работе А. Дж. Клауса).

143. В любом случае здесь основной проблемой, как было сказано, является вычисление обратной матрицы K_{ψ}^{-1} . А. Дж. Клаус нашёл способ выполнить это обращение аналитически, используя матричное тождество:

$$(A + CBC^T)^{-1} \equiv A^{-1} - A^{-1}C(B^{-1} + C^T A^{-1}C)^{-1} C^T A^{-1} \quad \forall A, B, C.$$

Его удобство заключается в отсутствии необходимости обращения матрицы C .

Использование этого тождества для матриц

$$U^T K_{\psi}^{-1} U = U^T [K_{\varepsilon} + SK_{\mathbf{v}} S^T] U$$

и

$$U^T K_{\psi}^{-1} \psi = U^T [K_{\varepsilon} + SK_{\mathbf{v}} S^T] \psi$$

позволяет получить следующие выражения, требуемые для вычисления оценки вектора α и матрицы K_{α} :

$$U^T K_{\psi}^{-1} U = U^T K_{\varepsilon}^{-1} U - U^T K_{\varepsilon}^{-1} S [K_{\mathbf{v}}^{-1} + S^T K_{\varepsilon}^{-1} S]^{-1} S^T K_{\varepsilon}^{-1} U,$$

$$U^T K_{\psi}^{-1} \delta\psi = U^T K_{\varepsilon}^{-1} \psi - U^T K_{\varepsilon}^{-1} S [K_{\mathbf{v}}^{-1} + S^T K_{\varepsilon}^{-1} S]^{-1} S^T K_{\varepsilon}^{-1} \psi.$$

Отсюда следует фильтр Клауса:

$$K_{\alpha} = \left(U^T K_{\varepsilon}^{-1} U - U^T K_{\varepsilon}^{-1} S [K_{\mathbf{v}}^{-1} + S^T K_{\varepsilon}^{-1} S]^{-1} S^T K_{\varepsilon}^{-1} U + K_{\alpha_0}^{-1} \right)^{-1},$$

$$\alpha = K_{\alpha} \left(U^T K_{\varepsilon}^{-1} \psi - U^T K_{\varepsilon}^{-1} S [K_{\mathbf{v}}^{-1} + S^T K_{\varepsilon}^{-1} S]^{-1} S^T K_{\varepsilon}^{-1} \psi + K_{\alpha_0}^{-1} \alpha_0 \right).$$

144. Этот фильтр работает лучше всего, когда систематические погрешности и случайный шум соизмеримы и матрица

$$M = \left[U^T U - U^T S (S^T S)^{-1} S^T U \right]$$

не вырождена. Вырождение матрицы M означает наличие существования равенства двух векторов в пространстве измерений:

$$S \delta \mathbf{v} = U \delta \mathbf{a},$$

что означает неоднозначность интерпретации измеренного значения Ψ (это как раз тот, упомянутый выше, случай, когда ошибки не противоречат модели измерений).

Несоизмеримость априорных представлений о величине шума и систематических ошибок измерений приводит в пределе к вырождению фильтра в тривиальный вид осреднения методом максимума правдоподобия либо с матрицей шума $U^T K_{\varepsilon}^{-1} U$, либо с матрицей систематических погрешностей $U^T S^T K_{\nu}^{-1} S U$ (это легко проверить в качестве упражнения на приведённом выше примере $K_{\Psi} = \sigma^2 E + \gamma I$).

145. Когда систематические ошибки противоречат модели ($\det M \neq 0$), систематические ошибки практически не влияют на результат оценки вектора определяемых параметров \mathbf{a} , несколько ухудшая, однако, формальную оценку $K_{\mathbf{a}}$. Самый правильный способ почувствовать влияние систематических ошибок на \mathbf{a} и $K_{\mathbf{a}}$ – это проанализировать в расчётах фильтра Клауса результаты изменения априорных данных K_{ε}^{-1} и K_{ν}^{-1} .

Модель наилучшей корреляции погрешностей измерений

146. Э.Л. Аким предложил [11] использовать модель плавающих корреляционных связей погрешностей измерений для неформальной расширенной оценки точности определения искомых параметров.

По этой модели все измерения имеют одинаковую дисперсию безусловного рассеивания σ^2 , а каждые два измерения в моменты t_{α} и t_{β} связаны в ковариационной матрице недиагональным корреляционным элементом вида

$$\left[K_{\Psi} \right]_{\alpha, \beta} = \sigma^2 \left[e^{-k\tau} \frac{\cos(\omega |\tau| + \varphi)}{\cos \varphi} \right]_{\tau=t_{\alpha}-t_{\beta}}.$$

Для вычисления расширенной оценки используются полученные в процессе обработки измерений коэффициенты A и B (см. формулу (B.1)) и подбираются наиболее неблагоприятные на множестве моментов

измерений общие для всей матрицы K_ψ значения радиуса корреляции k и параметров плавающей ошибки ω и φ .

147. Аналогичный подход к оценке погрешностей обработки измерений был использован в кандидатской диссертации В.Б. Бриткова [18], в которой выбирались наихудшие знаки единичных коэффициентов корреляции путём вычисления их функций влияния на отклонения определяемых параметров.

"Схема бортика"

148. Метод наиболее надёжной оценки погрешности определения параметров модели предложен автором этих лекций в упомянутой дискуссии с М.Л. Лидовым, но был тогда отвергнут нами ввиду сложности его программной реализации на машине "Стрела". Позднее он был независимо предложен в ИКИ РАН и реализован в кандидатской диссертации А.А. Суханова. В ИПМ им. М.В. Келдыша РАН этот метод получил наименование "Схема бортика". Рассматриваемый метод опирается на следующую модель влияния погрешностей измерений.

149. Пусть каждое измерение ψ_i характеризуется своей фиксированной максимальной погрешностью σ_i неизвестной природы и неизвестного распределения вероятностей. Рассмотрим, что в этом случае определяет такое измерение в пространстве определяемых параметров \mathbf{a} . В силу условного уравнения

$$\mathbf{u}_i \mathbf{a} = \psi_i$$

измерению $\psi_i \pm \sigma_i$ соответствует множество векторов \mathbf{a} , удовлетворяющих равенству скалярного произведения

$$\mathbf{u}_i \mathbf{a} = \psi_i \pm \sigma_i.$$

Множество концов этих векторов (их годограф) в пространстве определяемых параметров принадлежит гиперслою, который ортогонален градиенту \mathbf{u}_i измерения ψ_i в пространстве неизвестных векторов \mathbf{a} . То есть измерение $\psi_i \pm \sigma_i$ существенным образом ограничивает возможное множество значений искомых параметров, хотя и оставляет его бесконечным.

150. Если получить n измерений с линейно независимыми векторами градиентов \mathbf{u}_i для n неизвестных параметров \mathbf{a} , то соответствующее множество гиперслоёв концов векторов \mathbf{a} станет замкнутым, определив тем самым множество возможных значений оцениваемых параметров.

Для определения случая наихудшего знания параметров \mathbf{a} следует решить задачу линейного программирования на симплексе пересечения гиперслоёв относительно или наиболее критического направления вектора отклонения $\delta \mathbf{a}$ или его наибольшей величины.

151. Заметим, что толщина и ориентация каждого гиперслоя в пространстве определяемых параметров (α значит, и их пересечения) зависит лишь от принятой величины σ_i и ориентации вектора градиента измерения $\mathbf{u}_i = \text{grad}\psi_i$, но не зависит от значения измерения ψ_i .

Иными словами, как и раньше, размеры и ориентация области неопределённости параметров α , при условии полученных значений измерений, не зависит от этих значений (это – следствие линейной постановки задачи). Поэтому погрешности определения параметров α при заданных моментах и погрешностях выполнения измерений можно оценивать, не дожидаясь самих измерений.

Внешнее удобство этого обстоятельства обладает значительным внутренним недостатком – все эти оценки делаются по априорно назначенным значениям погрешностей σ_i . Отсюда следует необходимость обязательного апостериорного анализа надёжности полученного результата.

Такой анализ выполняется по характеру расположения и рассеивания выполненных измерений вдоль определённой траектории полёта.

"Мешающие параметры"

152. В общем случае измерения ψ (например, наклонные дальности) и их линейные отклонения $\delta\psi$ зависят не только от определяемых параметров α (например, элементов орбиты) и их линейных отклонений $\delta\alpha$, но и от некоторых плохо известных параметров μ принятой модели движения (например, массы Земли, координат пункта наблюдения и т.п.).

Такие параметры называют *мешающими* [14, 17, 19] (имеются в виду их помехи применению методов обработки измерений и/или оценки точности прогноза развития наблюдаемого процесса).

153. В линейном приближении можно ввести в условное уравнение дополнительную связь:

$$\psi = \psi(\alpha, \mu),$$

отсюда

$$\delta\psi = \frac{d\psi}{d\alpha} \delta\alpha + \frac{d\psi}{d\mu} \delta\mu.$$

Рассмотрим ряд следующих способов учёта влияния мешающих параметров на результаты обработки получаемых измерений.

Параметры μ вне обработки

154. Если мешающие параметры не учитываются в способе обработки измерений, то их следует учитывать в расширенной оценке погрешностей

определения α . Способ обработки в этом случае выглядит следующим образом:

$$\alpha = [U^T \tilde{K}_\psi^{-1} U + K_\alpha^{-1}]^{-1} U^T \tilde{K}_\psi^{-1} \tilde{\psi} = A \tilde{\psi},$$

где

$$\tilde{K}_\psi = K_\psi + K_\mu,$$

$$\tilde{\psi} = U\alpha + L\mu.$$

Здесь матрица L определяет преобразование в пространство измерений модели наблюдаемого процесса.

Тогда оценка погрешности α равна

$$K_\alpha = A(K_\psi + K_\mu)A^T = AK_\psi A^T + ALK_\mu L^T A^T.$$

Параметры μ , учитываемые в обработке измерений в качестве определяемых параметров

155. Можно попытаться определить параметры μ , введя их в качестве дополнительных определяемых параметров в процесс обработки имеющихся измерений.

Например, такая попытка оказалась необходимой при первых полётах к Венере и Марсу, когда линейные размеры Солнечной системы ещё были плохо известны. Заметным мешающим параметром в то время была величина астрономической единицы a_e . Поскольку величина её незнания приводила к погрешностям траектории, превышающим размеры и Марса, и Венеры, то пришлось попытаться определить величину a_e по наблюдениям траектории движения в поле сил гравитации Солнца и планет (зависящем от модели размеров Солнечной системы).

156. Для определения мешающего параметра надо следующим образом расширить используемые векторы и матрицы принятого метода обработки измерений:

$$\tilde{\alpha} = \begin{bmatrix} \alpha \\ \mu \end{bmatrix},$$

$$\tilde{U} = [U \quad L],$$

$$\tilde{K}_\alpha^{-1} = \begin{bmatrix} K_\alpha^{-1} & 0 \\ 0 & K_\mu^{-1} \end{bmatrix}.$$

В результате получаются новые выражения для оценки вектора $\tilde{\alpha}$ и ковариационной матрицы \tilde{K}_α (см. (B.1)):

$$\tilde{\alpha} = [\tilde{U}^T K_\psi^{-1} \tilde{U} + \tilde{K}_{\alpha_0}^{-1}]^{-1} (\tilde{U}^T K_\psi^{-1} \tilde{\psi} + \tilde{K}_{\alpha_0}^{-1} \tilde{\alpha}_0) \quad \text{и} \quad K_{\tilde{\alpha}} = \tilde{A} K_\psi \tilde{A}^T.$$

Важно отметить, что пользоваться подобным приёмом нужно крайне осторожно, поскольку он увеличивает число степеней свободы в задаче согласования ошибочных измерений с искомыми величинами определяемых параметров.

Параметры μ в ошибках прогноза изменений параметров модели

157. Более приятной является ситуация, когда параметры μ прямо не влияют на модельную зависимость $\psi(\alpha)$, но влияют в линейном приближении на прогноз некоторых параметров ρ , зависящих от определяемых параметров.

Примером подобной ситуации является плохое знание параметров небесного тела (характеристик вращения астероида, высот гор на Луне и на Венере, параметров атмосферы Марса и т.п.), влияющих на точность посадки на его поверхность.

В этом случае если положить

$$\rho = B\alpha + C\mu,$$

то ожидаемое рассеивание прогнозируемых величин ρ определяется ковариационной матрицей

$$K_\rho = BK_\alpha B^T + CK_\mu C^T.$$

Г. ВЫЧИСЛИТЕЛЬНЫЕ ПРОБЛЕМЫ ОБРАБОТКИ ИЗМЕРЕНИЙ

158. Приведём цитату Р. Калмана из его книги "Элементарная теория управления с современной точки зрения": *"Но независимо от того, о каком методе или алгоритме идёт речь, так или иначе приходится выяснять основные проблемы сходимости, точности и влияния вычислительной машины"*.

Рассмотрим две проблемы подобного рода.

1. Метод квадратных корней

159. Опыт показывает, что при использовании метода наименьших квадратов или максимума правдоподобия с ростом числа обрабатываемых измерений заметно ухудшаются параметры обусловленности матрицы N нормальной системы уравнений. Это приводит к численным ошибкам её обращения, вызывающим, прежде всего, потерю свойства симметричности полученной матрицы N^{-1} . Потеря симметрии в элементах левой и правой частей решения нормальной системы подчас приводит к катастрофически неверным результатам определения искомых величин.

Известен случай, когда единственный неверный последний 36-двоичный разряд мантиссы чисел на машине "Стрела" в ВЦ МГУ привёл к большому отличию тестовых результатов определения элементов орбиты – не только в порядках величин, но и их знаков!

В связи с этими обстоятельствами давно понятно, что для обращения матриц СНК и ММП необходимо использовать описываемый ниже метод квадратных корней. Устройство этого метода заведомо обеспечивает сохранение симметрии матрицы N^{-1} .

160. Метод квадратных корней предназначен для прямого решения линейных систем вида $Ax = f$ (без явного обращения матрицы A), в случае, когда матрица A – симметричная.

В случае симметрии обращаемая матрица всегда может быть предварительно разложена на произведение двух транспонированных треугольных матриц [8, с. 165], каждая из которых легко обратима, а произведение с транспонированием полученных матриц определяет симметрию результата. Этот метод хорошо работает даже для плохо обусловленных матриц. Преимущество метода заключается и в том, что он сокращает количество операций вычисления, выполняя неявное обращение матрицы A .

161. Рассмотрим симметричную матрицу, разложенную на транспонированное произведение двух одинаковых треугольных матриц:

$$A = S^T S,$$

где матрица S имеет вид

$$S = \begin{bmatrix} s_{11} & & & s_{ij} \\ 0 & \ddots & & \\ 0 & 0 & s_{ii} & \\ 0 & 0 & & s_{mm} \end{bmatrix}.$$

Упражнение 7. Докажите единственность такого разложения.

162. Определим элементы s_{ij} матрицы S для элементов матрицы $A = [a_{ij}]$ с помощью условий разложения матрицы на транспонированное произведение двух одинаковых матриц:

$$a_{ij} = s_{1i}s_{1j} + s_{2i}s_{2j} + \dots + s_{ii}s_{ij} \quad \text{при } i < j,$$

$$a_{ii} = s_{ii}^2 + s_{2i}^2 + \dots + s_{ii}^2 \quad \text{при } i = j.$$

Аналогия с извлечением корня из квадрата числа и породила название метода.

Упражнение 8. Докажите, что транспонированное произведение любой матрицы на себя образует симметричную матрицу. Что из этого следует для проблемы определения свойств разложения любой матрицы на транспонированное произведение двух одинаковых матриц? Единственно ли такое разложение, если оно существует?

163. Для элементов s_{ij} треугольной матрицы имеем

$$s_{11} = \sqrt{a_{11}}, \quad s_{1j} = \frac{a_{1j}}{s_{11}} \quad \text{при } i = 1,$$

$$s_{ii} = \sqrt{a_{ii} - \sum_{l=1}^{i-1} s_{il}^2} \quad \text{при } i > 1,$$

$$s_{ij} = \frac{a_{ij} - \sum_{l=1}^{i-1} s_{il}s_{lj}}{s_{ii}} \quad \text{при } i < j,$$

$$s_{ij} = 0 \quad \text{при } i > j.$$

Тогда линейная система $Ax = f$ разбивается на две следующие системы:

$$S^T k = f \quad \text{и} \quad Sx = k.$$

164. Элементы вектора k определяются аналогично по рекуррентной формуле:

$$k_1 = \frac{f_1}{S_{11}},$$

$$k_i = \frac{f_i - \sum_{l=1}^{i-1} S_{li} \cdot k_l}{S_{ii}} \text{ при } i > 1.$$

Таким способом решение системы $Ax = f$ с симметричной матрицей A приводится к решению двух треугольных систем.

165. Окончательное решение для вектора x (в результате "обратного хода" метода) выглядит так:

$$x_n = \frac{k_n}{S_{nn}},$$

$$x_i = \frac{k_i - \sum_{l=i+1}^n s_{il} \cdot x_l}{s_{ii}} \text{ при } i < n.$$

Упражнение 9. По существу в обратном ходе метода выполняется неявное обращение матрицы A . Постройте алгоритм обратного хода для явного обращения симметричной матрицы методом квадратных корней.

2. Метод параболического спуска

(Обобщение метода Ньютона: градиентный спуск в среде с оврагами)

166. Как быть, если искомые значения $\hat{\alpha}$ определяемых параметров α лежат вне области линейности градиентов $U(\alpha)$ условного уравнения?

Если в фильтрующем алгоритме ответ на этот вопрос важен лишь в начале процесса сходимости (см. п. 107 и п. 116), то в СНК или ММП обработка одновременно всех измерений требует построения итерационного процесса для поиска значений α , лежащих в линейной окрестности искомого значения $\hat{\alpha}$. На каждом шаге итераций в пространстве α решается стандартная линейная задача СНК и вычисляется мера сходимости процесса $\Phi(\alpha)$. В качестве $\Phi(\alpha)$ годится любая функция параметра α , для которой при $\Phi(\alpha) \rightarrow 0$, градиенты $U(\alpha) \rightarrow \text{const}$.

Итак, пусть требуется найти вектор α , доставляющий минимум функции $\Phi(\alpha)$.

167. В случае, когда размерность вектора α не велика, для решения годится метод *наискорейшего спуска* по направлению антиградиента функции $\Phi(\alpha)$:

1. В точке \mathbf{a} вычисляется n -мерный вектор и модуль градиента функции $\Phi(\mathbf{a})$:

$$\text{grad } \Phi(\mathbf{a}) = \left[\frac{\partial \Phi}{\partial \alpha_j} \right] \text{ и } |\text{grad } \Phi(\mathbf{a})| \quad (j = n).$$

2. После этого направление антиградиента и величина текущего изменения координат вектора \mathbf{a} определяются из уравнения

$$\Delta \mathbf{a} = \left[\frac{d\alpha_j}{ds} \right] \Delta s = -|\text{grad } \Phi(\mathbf{a})|^{-1} \left[\frac{\partial \Phi}{\partial \alpha_j} \right] \Delta s,$$

где s – параметр пути, а Δs – это некоторый достаточно малый шаг итераций. Процесс продолжается до тех пор, пока $|\text{grad } \Phi(\mathbf{a})| > \varepsilon$, где ε – заданное число.

168. Здесь необходимы следующие комментарии.

- Практически при $n > 3$ метод наискорейшего спуска не обеспечивает быстрого и надёжного определения $\min(\Phi(\mathbf{a}))$ (и соответствующего искомого вектора \mathbf{a}) из-за резких изменений в процессе спуска как величины, так и направления вектора градиента $\text{grad } \Phi(\mathbf{a})$. Последнее обстоятельство связано с неоднородностью различных компонент вектора \mathbf{a} (разница единиц измерения и/или их масштабов, нелинейное искривление "поверхности" $\Phi(\mathbf{a})$ и др.), что приводит к неизотропности размеров смещения $\Delta \mathbf{a}$ в разных направлениях.

- Геометрически этому соответствует образ (гипер) "оврага" с крутыми склонами и пологим искривлённым дном.

- Существуют многие способы борьбы с овражным эффектом для обеспечения сходимости процесса поиска решения. Все они работают в определённых предположениях относительно свойств функции $\Phi(\mathbf{a})$. Например, метод, имитирующий инерционное накопление скорости движения тяжёлым шариком в сторону осреднённого направления спуска при более или менее монотонном направлении его векового движения; другой пример – при квадратичности, $\Phi(\mathbf{a})$ для реализации простого и красивого метода обеспечения сходимости используется свойство градиентов, касательных к изолиниям функции $\Phi(\mathbf{a})$.

Упражнение 10. Докажите, что для квадратичной функции $\Phi(\mathbf{a})$ линии градиентов $\text{grad } \Phi(\mathbf{a})$, последовательно касательные к изолиниям, образуют две стороны треугольника, третья сторона которого своим продолжением ведёт в точку минимума.

169. Вследствие овражного эффекта возникают две проблемы, связанные с концом процесса сходимости: проблема его достижения и проблема его определения. Классический график процесса сходимости по числу шагов имеет вид, представленный на рис. Г.1. При таком характере кривой сходимости конца процесса можно и не дожидаться.

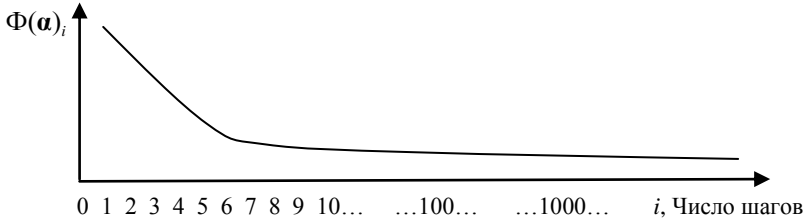


Рис. Г.1. График процесса сходимости по числу шагов

Вопрос. Как геометрически объяснить причину резкого излома кривой сходимости к минимуму функции $\Phi(\mathbf{a})$, показанного на рис. Г.1.

170. Заметим, что даже в случае, если $|\text{grad } \Phi(\mathbf{a})| \leq \varepsilon$, то в ситуации оврага это ещё не означает, что достигнута точка минимума $\Phi(\mathbf{a})$. Правильным критерием конца итераций (который обычно не используется в силу его вычислительной сложности) является положительная определённая вариация $\delta\Phi(\mathbf{a})$ в базисе касательного пространства в точке решения. В противном случае может существовать направление в пространстве \mathbf{a} , в котором функция $\Phi(\mathbf{a})$ убывает.

171. Для решения проблемы оврага уже при самом первом опыте компьютерного определения орбиты [9, 15] Т.М. Энеевым был найден мощный метод, обеспечивающий сходимость итерационного спуска в многомерном пространстве. Идею этого метода лучше всего объяснить на геометрическом языке.

Пусть в качестве примера требуется найти вектор $\mathbf{a}(x, y, z)$, доставляющий минимум некоторой функции $\Phi(\mathbf{a})$ в пространстве \mathbf{a} , достаточно изрезанном оврагами. Тогда надо поступить так: преобразовать пространство переменных \mathbf{a} в другое, более гладкое, пространство переменных $\mathbf{\beta}(\xi, \eta, \zeta)$ той же размерности и реализовать поиск решения в этом новом пространстве без оврагов.

172. Заметим, что наиболее гладкой функцией $\Psi(\mathbf{\beta})$, имеющей единственный минимум в любом пространстве $\mathbf{\beta}(\xi, \eta, \zeta)$, является простейший параболоид вращения $\Psi(\mathbf{\beta}) = \xi^2 + \eta^2 + \zeta^2$. Следовательно, надо вы-

брать такие параметры (ξ, η, ζ) и так построить связь $\beta = \beta(\alpha)$, чтобы минимальные величины функций $\Phi(\alpha)$ и $\Psi(\beta)$ достигались одновременно на векторах α и β , связанных найденным условием их связи $\beta = \beta(\alpha)$.

Если такой параболоид построен, то надо реализовать движение *наискорейшего* спуска по линии тока на поверхности параболоида $\Psi(\beta)$, после чего отобразить полученную траекторию спуска в этом пространстве β на соответствующую траекторию в исходном пространстве α .

Эта траектория не будет связана с уклонами "плохого" пространства α , и будет спускаться, минуя овраги, прямо в точку минимума $\Phi(\alpha)$ так, как это показано сплошной линией рядом с пунктиром на рис. Г.2.

173. Достаточно очевидно, что искомая связь $\beta = \beta(\alpha)$ гладкого и искривлённого пространств в общем случае может быть выписана лишь в локально дифференциальном виде. Поэтому для отображения пути в пространстве параболоида в путь в координатах α заведомо должно потребоваться интегрирование системы дифференциальных уравнений.

174. Аналитически реализация такого метода параболического спуска выглядит так.

Введём в рассмотрение искомую непрерывную связь пространств α и β (в общем случае – с их одинаковой размерностью n):

$$\beta = \beta(\alpha_1, \alpha_2, \dots, \alpha_n)$$

и построим в пространстве β n -мерный параболоид вращения, "уничтожающий" проблему оврагов:

$$\Psi(\beta) = \beta_1^2 + \beta_2^2 + \dots + \beta_n^2.$$

Пусть удалось найти такое соответствие между переменными α и β , что из условия $\min(\Phi(\alpha))$ следует $\min(\Psi(\beta(\alpha))) = 0$.

175. Заметим, что $\text{grad } \Psi = 2\beta$ и $|\text{grad } \Psi| = 2\sqrt{\Psi}$.

Достаточно очевидно, что компоненты единичного вектора антиградиента функции $\Psi(\beta)$ (направления пути s наискорейшего спуска по поверхности параболоида в пространстве β – прямо к $\min(\Psi(\beta))$) определены соотношением

$$\frac{d\beta_i}{ds} = -\frac{\beta_i}{\sqrt{\Psi}}. \quad (\text{Г.1})$$

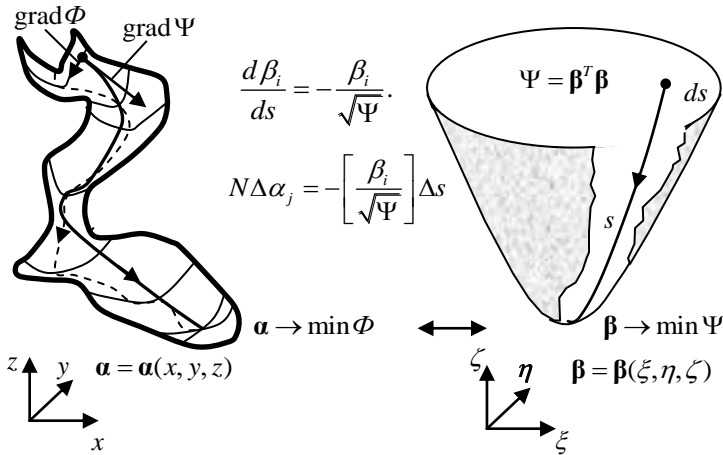


Рис. Г.2. Слева "плохое" пространство оврага, справа "хорошее" – параболоида

176. Итак, для того чтобы найти вектор $\alpha(x, y, z)$, доставляющий минимум некоторой функции $\Phi(\alpha)$, когда пространство α сильно изрезано оврагами, надо:

- построить новое пространство векторов $\beta(\xi, \eta, \zeta)$ и его связь $\beta = \beta(\alpha)$ с пространством векторов α ;
- согласовать эти векторы так, чтобы минимальные величины суммы квадратов $\Psi = \beta^T \beta$ и заданной функции $\Phi(\alpha)$ достигались одновременно на векторах, связанных условием $\beta = \beta(\alpha)$;
- двигаться в "хорошем" пространстве β из начальной точки $\beta_0 = \beta(\alpha_0)$ в направлении антиградиента поверхности параболоида $\Psi = \beta^T \beta$;
- отображать получаемую траекторию спуска в пространстве β на соответствующую траекторию спуска в пространстве α .

Эта траектория будет идти, как это показано на рис. Г.2, минуя овраги, прямо в точку минимума $\Phi(\alpha)$.

Искомое отображение траекторий спуска получается интегрированием дифференциальных уравнений, являющихся решением линейной алгебраической системы, матрица которой, как выяснилось, совпадает с матрицей метода Ньютона для поиска минимума функции $\Phi(\alpha)$.

При поиске минимума методом Ньютона делается очень большой шаг интегрирования в направлении, касательном к требуемой линии правиль-

ного спуска, в то время как в методе параболического спуска (оказавшимся, таким образом, обобщением метода Ньютона) точное интегрирование вдоль этого направления все время ведёт решение вдоль наискорейшей кривой спуска. На практике, однако, точное интегрирование, как правило, не используется, а просто применяется дробление пополам шага Ньютона, до тех пор пока функция $\Phi(\mathbf{a})$ не уменьшится заметно, после чего снова пробуют рассчитать новый шаг Ньютона из полученной новой точки \mathbf{a} .

При этом точное решение этого дифференциального уравнения однозначно определяет конец процесса интегрирования (достижения искомого минимума) длиной пути на поверхности параболоида $\Psi(\boldsymbol{\beta})$, равной

$$s_k = \sqrt{\Psi_0}.$$

Упражнение 11. Докажите это утверждение.

177. Таким образом, спуск в "хорошем" пространстве $\boldsymbol{\beta}$ действительно является "наискорейшим". Но наша задача – спустаться в "плохом" пространстве \mathbf{a} . Для этого необходимо интегрировать дифференциальное отображение "хорошего" пространства в "плохое".

178. Это отображение получается следующим образом. Продифференцируем зависимость $\boldsymbol{\beta} = \boldsymbol{\beta}(\alpha_1, \alpha_2, \dots, \alpha_n)$, получив квадратную матрицу N :

$$N = \|\|a_{ij}\|\| = \|\|\partial\beta_i/\partial\alpha_j\|\|.$$

Тогда, применив соотношение (Г.1), построим систему линейных алгебраических уравнений относительно производных компонент вектора \mathbf{a} в "плохом" пространстве по пути s на поверхности параболоида в "хорошем" пространстве:

$$N \left[\frac{d\alpha_j}{ds} \right] = - \left[\frac{\beta_i}{\sqrt{\Psi}} \right].$$

Из её решения получается явная система n дифференциальных уравнений

$$\left[\frac{d\alpha_j}{ds} \right] = -N^{-1} \left[\frac{\beta_i}{\sqrt{\Psi}} \right]. \quad (\text{Г.2})$$

179. Полученные дифференциальные уравнения можно достаточно точно любым методом высокого порядка численно проинтегрировать по параметру пути s в пределах от 0 до $s_k = \sqrt{\Psi_0}$. В результате этого интегрирования будет получена траектория пути в "плохом" пространстве \mathbf{a} для реализации "прямого" движения в точку $\min(\Phi(\mathbf{a}))$. *Это движение никак не связано с градиентами функции $\Phi(\mathbf{a})$, оно отображает в про-*

странстве α линию тока на поверхности параболоида $\Psi(\beta)$ в пространстве β .

180. Однако проще и более правильно использовать для интегрирования простейший метод Эйлера первого порядка. Это в конечно-разностном приближении даёт пошаговую линейную аппроксимацию образа линии тока:

$$\Delta\alpha_j = -N^{-1} \left[\frac{\beta_i}{\sqrt{\Psi}} \right] \Delta s = -N^{-1} \beta_i \Delta\sigma. \quad (\text{Г.3})$$

Шаг такого интегрирования $\Delta\sigma = \Delta s / \sqrt{\Psi}$ выбирается из условий наилучшей сходимости процесса. При достаточно грубом шаге интегрирования траектория движения на параболоиде не обязательно совпадает с линией тока, и в связи с этим длина пути интегрирования заранее не определена.

181. *Теперь очень важный момент:* при выборе максимальной величины шага $\Delta\sigma = 1$ выражение (Г.3) для вычисления величины $\Delta\alpha$ совпадает с выражением метода Ньютона для решения многомерной системы трансцендентных уравнений, получаемой приравнением нулю зависимости $\beta_i(\alpha_j) = 0$.

Это следует из того факта, что в методе Ньютона используется та же матрица

$$\|\partial\beta_i/\partial\alpha_j\| = \|a_{ij}\|.$$

182. Отсюда следует любопытное свойство метода Ньютона:

Решение методом Ньютона любой системы трансцендентных уравнений $Y(X) = 0$ можно интерпретировать как задачу поиска нуля суммы квадратов аргумента X линейного приближения уравнений решаемой системы $Y = AX$.

183. Таким образом, параболический спуск приводит к итерационному решению нелинейной алгебраической системы, причём матрица линейной системы на каждой итерации совпадает с матрицей метода Ньютона для поиска в текущей точке минимума функции $\Phi(\alpha)$.

Отличие этих двух методов заключается в том, что при поиске минимума методом Ньютона делается сразу очень большой шаг интегрирования в направлении, касательном к требуемой линии правильного спуска в текущей точке движения, в то время как в методе параболического спуска (оказавшимся, таким образом, обобщением метода Ньютона) точное интегрирование все время обеспечивает движение вдоль этой правильной кривой спуска.

184. При *ньютоновом* единичном шаге интегрирования системы параболического спуска ($\Delta\sigma = 1$) движение в пространстве α происходит достаточно далеко по прямой, касательной в её начальной точке к кривой, отображающей линию тока параболоида Ψ (см. рис. Г.2). Соответствующее движение на параболоиде описывается некоторой кривой линией, всё более отклоняющейся от линии тока, которая, в конце концов, может подняться и выше исходной точки.

185. На практике точное интегрирование, как правило, не используется. Вместо этого применяется иной метод, часто называемый *модифицированным методом Ньютона*, – уменьшение размера $\Delta\alpha$ происходит до тех пор, пока минимизируемая функция Ψ не уменьшится заметно. Тогда в полученной новой точке α вычисляется и обращается матрица $\|a_{ij}\|$ (см. Г.3) и делается шаг Ньютона $\Delta\sigma = 1$.

Действительно, если в процессе решения линейное приближение спуска оказывается плохим и мера сходимости Ψ на очередном шаге процесса не уменьшилась заметно (или даже увеличилась), то не следует начинать новый спуск из полученной точки: это потребует очередного вычисления и обращения матрицы производных $\|a_{ij}\|$. Вместо этого следует уменьшить длину шага спуска из старой точки (где все эти действия уже выполнены), т.е. попросту уменьшить полученное приращение $\Delta\alpha$ (например, уполовинить его) и действовать так до тех пор, пока Ψ не станет заметно меньше предыдущего значения.

186. Метод *параболического спуска* впервые был описан Канторовичем применительно к экономическим моделям [4]. Позже при первом решении задачи траекторных определений с помощью ЭВМ он был независимо вторично открыт Т.М. Энеевым и еще позже в США – С. Брайсоном.

187. Остаётся вопрос. Как использовать метод параболического спуска для решения нормальной системы в нелинейной задаче обработки измерений?

Ответ. В качестве квадратичного критерия спуска $\Psi(\beta)$ надо выбрать сумму квадратов n -мерного столбца невязок (рассогласований левых и правых частей) нормальной системы (Б.4) или (Б.7):

$$\beta = U^T K_{\Psi}^{-1} U \hat{\alpha} - U^T K_{\Psi}^{-1} \psi$$

или

$$\beta = [U^T K_{\Psi}^{-1} U + K_{\alpha_0}^{-1}] \hat{\alpha} - (U^T K_{\Psi}^{-1} \psi + K_{\alpha_0}^{-1} \alpha_0) \text{ и } \Psi(\beta) = \beta^T \beta.$$

Очевидно, что при $\beta \rightarrow 0, \Psi(\beta) \rightarrow 0$ градиенты условных уравнений $U(\alpha) \rightarrow \text{const}$, т.е. условия применения параболического спуска выполняются.

Д. ПРИМЕРЫ ПРИМЕНЕНИЯ МЕТОДОВ СРЕДНЕКВАДРАТИЧНОЙ ФИЛЬТРАЦИИ

1. Задача определения ориентации космического аппарата

188. Для того чтобы использовать фильтр Калмана для определения ориентации, необходимо иметь математическую модель вращения космического аппарата (КА). Предполагается, что задача определения ориентации КА будет решаться на бортовом компьютере, обладающем весьма ограниченными вычислительными ресурсами. Поэтому имеет смысл построить достаточно грубую модель движения КА относительно его центра масс, учитывающую только основные действующие моменты [21]. В качестве последних ниже рассматриваются гравитационный момент и момент системы управления ориентацией КА.

Часто на космических аппаратах в качестве исполнительных органов систем ориентации устанавливаются маховичные системы. Фильтр Калмана в таких случаях используется для определения параметров текущей ориентации при выполнении управляющих ориентационных манёвров маховиками в задачах, к примеру, дистанционного зондирования Земли. Рассмотрим такую задачу для иллюстрации обстоятельств практики применения фильтра Калмана. В качестве датчиков информации, используемых для определения ориентации КА, рассмотрим магнитометр и солнечный датчик, которые измеряют вектор локального геомагнитного поля и вектор направления на центр диска Солнца в связанной с КА системе координат (ССК). В качестве вектора оцениваемых параметров выберем кватернион поворота ССК относительно орбитальной системы координат (ОСК) и вектор угловой скорости ССК относительно инерциальной системы координат (ИСК).

189. В векторном виде динамическое уравнение движения КА можно записать в следующем виде:

$$J\dot{\boldsymbol{\omega}} = -\dot{\mathbf{h}} + \frac{3\mu}{R^3}(\boldsymbol{\eta} \times J\boldsymbol{\eta}) - \boldsymbol{\omega} \times (J\boldsymbol{\omega} + \mathbf{h}), \quad (\text{Д.1})$$

где $\boldsymbol{\omega} = [\omega_x \ \omega_y \ \omega_z]^T$ – вектор угловой скорости ССК относительно ИСК,

J – тензор инерции микроспутника, $\boldsymbol{\eta} = [\eta_x \ \eta_y \ \eta_z]^T$ – единичный вектор направления местной вертикали, записанный в ССК, $\mu = GM_3$ – гравитационный параметр Земли, R – расстояние от КА до центра Земли, \mathbf{h} –

кинетический момент маховиков, $\dot{\mathbf{h}}$ – изменение управляющего кинетического момента маховиков, которое задается выражением

$$\dot{\mathbf{h}} = k_a \boldsymbol{\lambda} + k_\omega \tilde{\boldsymbol{\omega}} - \boldsymbol{\omega} \times (J\boldsymbol{\omega} + \mathbf{h}),$$

$\boldsymbol{\lambda}$ – векторная часть кватерниона $\Lambda = [\boldsymbol{\lambda} \quad \lambda_4]$ перехода из ОСК в ССК, $\tilde{\boldsymbol{\omega}}$ – вектор угловой скорости ССК относительно ОСК, k_a , k_ω – параметры пропорционально-дифференциального-регулятора управления маховиками. Последнее слагаемое в (Д.1) описывает гравитационный момент.

Таким образом, динамическое уравнение можно переписать следующим образом:

$$J\dot{\boldsymbol{\omega}} = -k_a \mathbf{q} - k_\omega \tilde{\boldsymbol{\omega}} + \frac{3\mu}{R^3} (\boldsymbol{\eta} \times J\boldsymbol{\eta}). \quad (\text{Д.2})$$

Запишем кинематические уравнения в виде

$$\dot{\Lambda} = \frac{1}{2} \Omega \Lambda. \quad (\text{Д.3})$$

Здесь

$$\Omega = \begin{pmatrix} \tilde{W} & \tilde{\boldsymbol{\omega}} \\ -\tilde{\boldsymbol{\omega}}^T & 0 \end{pmatrix},$$

где \tilde{W} – кососимметрическая матрица угловой скорости $\tilde{\boldsymbol{\omega}}$, которая вычисляется по формуле

$$\tilde{\boldsymbol{\omega}} = \boldsymbol{\omega} - \omega_0 \mathbf{e}_2,$$

ω_0 – орбитальная угловая скорость, \mathbf{e}_2 – орт нормали к плоскости орбиты, записанный в ССК.

190. Для фильтра, основанного на измерениях солнечного датчика и магнитометра, в качестве вектора оцениваемых величин возьмем векторную часть кватерниона перехода из ОСК в ССК и угловую скорость ССК относительно ИСК, $\boldsymbol{\alpha} = [\boldsymbol{\lambda}^T \quad \boldsymbol{\omega}^T]^T$. Теперь линеаризуем динамическое и кинематическое уравнения движения в окрестности текущего положения. Запишем уравнения (Д.2) и (Д.3) в виде

$$\frac{d}{dt} \delta \boldsymbol{\alpha}(t) = F(t) \delta \boldsymbol{\alpha}(t),$$

где $\delta \boldsymbol{\alpha}(t)$ – малое приращение вектора состояния, а $F(t)$ – линеаризованная матрица уравнения движения в окрестности состояния $\boldsymbol{\alpha}(t)$. Тогда получим

$$F(t) = \begin{pmatrix} -W(t) & \frac{1}{2}E \\ J^{-1} \left(\frac{6\mu}{R^3} F_g(t) - Ek_a \right) & -J^{-1}k_\omega \end{pmatrix},$$

где E – единичная матрица, W – кососимметрическая матрица угловой скорости ω , F_g – матрица линеаризованного гравитационного момента, в случае диагонального тензора инерции $J = \text{diag}(I_1, I_2, I_3)$ равная

$$F_g = \begin{pmatrix} (\eta_2^2 - \eta_3^2)(I_2 - I_3) & -\eta_1\eta_2(I_2 - I_3) & \eta_3\eta_1(I_2 - I_3) \\ \eta_1\eta_2(I_3 - I_1) & (\eta_3^2 - \eta_1^2)(I_3 - I_1) & -\eta_3\eta_2(I_3 - I_1) \\ -\eta_1\eta_3(I_1 - I_2) & \eta_2\eta_3(I_1 - I_2) & (\eta_1^2 - \eta_2^2)(I_1 - I_2) \end{pmatrix},$$

где η_1, η_2, η_3 – компоненты вектора вертикали η .

191. Априорная оценка $\hat{\mathbf{a}}_k^+$ на шаге k работы фильтра представляет собой результат совместного интегрирования уравнений движения (Д.2) и кинематических уравнений (Д.1) [21]. Априорная оценка ковариационной матрицы ошибок на шаге k рассчитывается как

$$K_k^+ = \Phi_{k-1,k} K_{k-1} \Phi_{k-1,k}^T + Q,$$

где Q – ковариационная матрица ошибок модели движения, K_{k-1} – апостериорная оценка матрицы ошибки на предыдущем шаге.

192. Апостериорная оценка строится как корректировка априорной оценки после получения измерений. Условное уравнение вектора измерений имеет вид

$$\Psi_k = \mathbf{u}(\hat{\mathbf{a}}_k^+, k) + \mathbf{v}_k.$$

Здесь $\omega(t=t_0) = \omega_m(t=t_0)$ – полученный на k -м шаге вектор измерений, $\mathbf{h}(\hat{\mathbf{a}}_k^+, k)$ – модель измерений, \mathbf{v}_k – вектор шумов измерений с ковариационной матрицей R . Вектор измерений состоит из вектора магнитного поля и вектора направления на Солнце в ССК:

$$\Psi_k = \begin{bmatrix} \mathbf{b}_k^T & \mathbf{s}_k^T \end{bmatrix}^T.$$

Вектор \mathbf{u} можно записать в виде

$$\mathbf{u} = \begin{bmatrix} (A(\hat{\Lambda}_k^+) \mathbf{b}_o)^T & (A(\hat{\Lambda}_k^+) \mathbf{s}_o)^T \end{bmatrix}^T,$$

где A – матрица перехода из ССК в ОСК, записанная через оценку кватерниона $\hat{\Lambda}_k^+$, \mathbf{b}_o и \mathbf{s}_o – векторы магнитного поля и направления на Солнце

це, записанные в ОСК. Матрица A выражается через кватернион $\Lambda = [\lambda \ \lambda_4]$ следующим образом:

$$A(\Lambda) = (\lambda_4^2 - \|\Lambda\|^2)E + 2\lambda\lambda^T - 2\lambda_4 W_\lambda,$$

где

$$W_\lambda = \begin{pmatrix} 0 & -\lambda_3 & \lambda_2 \\ \lambda_3 & 0 & -\lambda_1 \\ -\lambda_2 & \lambda_1 & 0 \end{pmatrix}.$$

193. Линейризованная модель измерений записывается следующим образом:

$$\delta\boldsymbol{\Psi}(t) = U(t)\delta\boldsymbol{\alpha}(t).$$

Здесь $\delta\boldsymbol{\Psi}(t)$ – малое приращение измерений при малом изменении вектора состояния $\delta\boldsymbol{\alpha}(t)$ в момент времени t . Матрица чувствительности U в нашем случае имеет вид

$$U = \begin{bmatrix} U_1 \mathbf{b}_o & U_2 \mathbf{b}_o & U_3 \mathbf{b}_o & \mathbf{0}_{3 \times 3} \\ U_1 \mathbf{s}_o & U_2 \mathbf{s}_o & U_3 \mathbf{s}_o & \mathbf{0}_{3 \times 3} \end{bmatrix},$$

где U_i определяется из уравнения

$$U_i = \left[\frac{\partial A(\Lambda)}{\partial \lambda_i} \right], \quad i = \overline{1, 3}.$$

Таким образом,

$$U_1 = 2 \begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_2 & -\lambda_1 & \lambda_4 \\ \lambda_3 & -\lambda_4 & -\lambda_1 \end{pmatrix}, \quad U_2 = 2 \begin{pmatrix} -\lambda_2 & \lambda_1 & -\lambda_4 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_4 & \lambda_3 & -\lambda_2 \end{pmatrix}, \quad U_3 = 2 \begin{pmatrix} -\lambda_3 & \lambda_4 & \lambda_1 \\ -\lambda_4 & -\lambda_3 & \lambda_2 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix}.$$

194. Матрица релаксации (коэффициентов обратной связи) записывается следующим образом: $H_k = K_k^+ U_k^T [U_k K_k^+ U_k^T + R]^{-1}$.

195. Апостериорная оценка фильтра Калмана выглядит так:

$$\boldsymbol{\alpha}_k = \boldsymbol{\alpha}_k^+ + H_k (\boldsymbol{\Psi}_k - \mathbf{u}(\boldsymbol{\alpha}_k^+)).$$

196. Апостериорная оценка для матрицы ошибки имеет вид

$$K_k = (E - H_k U_k) K_k^+.$$

197. Несмотря на большую популярность рекурсивной фильтрации как метода оценки параметров ориентации в режиме реального времени, существует ряд проблем при его использовании. Выбор матрицы ошибок модели движения Q , который называется часто как "настройка фильтра", является основной проблемой использования фильтра Калмана [22]. Эта

матрица имеет значительное влияние на качество работы фильтра: точность оценок вектора состояния и время сходимости.

Суть проблемы заключается в следующем. Уравнения движения (Д.1) являются приближенными и не учитывают всех возмущений, которые действуют на КА. Модель этих возмущений либо достаточно сложная, чтобы использовать её на бортовом компьютере (который может быть слабомощным), либо требует задания неизвестных или сложно определенных параметров. Заранее величина действующих возмущений, неучтенных в модели движения, как правило, неизвестна и может быть лишь оценена сверху из общих соображений. Согласно предположениям, на которых строится фильтр Калмана, эти неучтенные возмущения должны иметь случайный характер с нормальным распределением, которое задается матрицей ошибок Q . Однако реальные возмущения могут не удовлетворять этому условию, и это приводит к тому, что оценки фильтра Калмана уже не являются минимальными по среднеквадратичному критерию. Поэтому при нарушении исходных предположений относительно нормального распределения возмущений возникает задача выбора такой матрицы Q , которая бы обеспечивала приемлемую точность определения параметров α .

198. На практике настройка фильтра – это некоторый специальный процесс поиска матриц шумов для достижения желаемых характеристик работы фильтра Калмана, часто основанный на методе проб и ошибок. Однако существует ряд автоматизированных методик настройки фильтра. Существует метод настройки фильтра Калмана с помощью техники численной минимизации. В качестве функционала выбирается сумма квадратов разностей оценок от вектора состояния и его реальной величины, известной при математическом моделировании, в качестве параметров рассматривались элементы матриц ошибок. Далее с помощью моделирования работы фильтра Калмана проводится процедура численной минимизации функционала.

Другой подход настройки фильтра заключается в применении метода Монте-Карло, который основан на множественном моделировании работы фильтра при случайно выбираемых значениях матриц ошибок и начальных условий. Существуют также генетические алгоритмы для настройки фильтра Калмана. Суть генетических алгоритмов заключается в случайном изменении вектора параметров ("мутации"), которое может привести либо к ухудшению, либо к улучшению точности оценок фильтра Калмана. После нескольких "мутаций" выбирается вектор параметров, который привел к наилучшей точности, и на следующей стадии "мутации" подвергается уже этот вектор параметров. Так "эволюция" продолжается до тех пор, пока на некотором "поколении" все "мутации" не приведут к улучшению точности оценок фильтра. Главной особенностью генетиче-

ских алгоритмов является то, что с их помощью возможно найти только локальные минимумы функционала, и поэтому они плохо подходят для задачи настройки фильтров.

199. Все вышеперечисленные методы численной настройки фильтра требуют больших вычислительных мощностей, так как основаны на множественном моделировании работы фильтра. Так как при моделировании работы фильтра используются случайные шумы системы и измерений, то и результат методов настройки будет некоторой случайной величиной, математическое ожидание и дисперсию которой также необходимо определить. Для оценки влияния неучтенных в модели движения возмущений на точность фильтра моделирование работы фильтра проводится с учетом этих возмущений на исходное "идеальное" движение.

200. Другой метод исследования точности оценок движения может быть использован для стационарного движения. Этот метод не требует моделирования работы фильтра Калмана и является аналитическим. В работе [23] показано, что для стационарной системы матрица ошибок фильтра после сходимости может быть получена из квадратного матричного уравнения. Для стационарной системы справедливо:

$$\Phi_k = \Phi = \text{const}, U_k = U = \text{const}, Q_k = Q = \text{const}, R_k = R = \text{const}.$$

Для вычисления *формальной* точности оценки вектора состояния $\hat{\mathbf{a}}_k$ после переходных процессов необходимо вычислить матрицу K_∞ из условия равенства ковариационной матрицы на двух соседних шагах $K_\infty = K_k = K_{k-1}$, что можно представить в виде квадратного матричного уравнения относительно неизвестной K_∞ :

$$K_\infty = \left[E - (\Phi K_\infty \Phi^T + Q)U^T [U(\Phi K_\infty \Phi^T + Q)U^T + R]^{-1}U \right] \cdot (\Phi K_\infty \Phi^T + Q).$$

В общем случае это уравнение аналитически решить не удастся. Учитывая, что матрица K_∞ симметричная, рассматриваемое нелинейное матричное уравнение можно записать как систему из n уравнений с n неизвестными, в качестве неизвестных рассматриваются элементы матрицы K_∞ . Эту систему нелинейных уравнений можно решить, например, методом простой итерации или методом Ньютона.

Заметим, что K_∞ вычисляется как предельное значение для заданных и постоянных матриц Φ , U , Q и R . Однако, если рассмотреть систему, значения матрицы динамики и матрицы чувствительности которой изменятся настолько медленно, что в каждый момент времени можно считать, что оценки фильтра Калмана сошлись, то и для такой системы можно оценить *формальную* точность определения вектора состояния с помощью решения вышеприведенного матричного уравнения.

2. Оценка смещения нуля магнитометра по полетным данным

201. Рассмотрим пример использования способа наименьших квадратов для оценки смещения магнитометра по полётным данным, если дополнительно имеются измерения солнечного датчика.

Модель измерений магнитометра может быть записана в виде

$$\mathbf{m}_i = \mathbf{b}_i + \boldsymbol{\mu}. \quad (\text{Д.4})$$

Здесь \mathbf{m}_i – вектор измерений магнитометра; \mathbf{b}_i – вектор напряжённости магнитного поля Земли в связанной системе координат; $\boldsymbol{\mu}$ – искажающий измерения вектор смещения нуля магнитометра.

202. Смещение в измерениях магнитометра может быть вызвано магнитными полями бортовых приборов. В общем случае эти "возмущающие" поля могут быть непостоянными и зависеть, например, от величины токов с солнечных батарей, от степени заряда аккумуляторной батареи, от режима работы конкретного бортового прибора. Несмотря на то, что разработчики космических аппаратов стараются уменьшить бортовые магнитные поля путем экранирования приборов, специальной разводки проводов, полностью их устранить не удается. Поэтому возникает задача оценки бортовых магнитных полей, чтобы исключить их из измерений магнитометра и определить вектор геомагнитного поля в связанной с аппаратом системе координат.

203. Предположим, что смещение нуля измерений магнитометра постоянно. Для его оценки на достаточно продолжительном промежутке времени полёта необходимы следующие телеметрические данные:

- измерения магнитометра,
- измерения солнечных датчиков,
- параметры орбиты участка траектории, на которой были получены телеметрические данные.

Как правило, каждый кадр телеметрии привязан к бортовому времени выполнения измерения и такие кадры следуют с некоторой периодичностью.

204. Определить смещение нуля магнитометра можно следующим образом.

Используя модель магнитного поля Земли и зная местонахождение КА в моменты времени, для которых доступна телеметрия, можно вычислить величину вектора напряжённости магнитного поля Земли в ОСК. Далее, зная начальное приближение вектора состояния КА (кватернион поворота ССК относительно ОСК и вектор угловой скорости), интегрируя уравнения движения (Д.1) и используя модель измерений (Д.4), вычисляется ожидаемый вектор измерений для каждого кадра телеметрии. Затем с

помощью метода наименьших квадратов на полученной совокупности данных определяется смещение нуля магнитометра в виде математического ожидания рассогласования между векторами измерений магнитометра и векторами ожидаемых измерений.

205. Для интегрирования уравнений движения (Д.1) необходимо иметь начальные условия положения КА и оценку действующих моментов. Первое приближение нужных начальных условий можно получить, используя следующий алгоритм определения ориентации по одному измерению.

Составим систему уравнений, позволяющую по имеющимся расчётным данным однозначно определить элементы матрицы направляющих косинусов A . Пусть \mathbf{S} , \mathbf{H} – единичные векторы направления на Солнце и напряженности геомагнитного поля в ОСК соответственно, \mathbf{s} , \mathbf{h} – те же векторы в ССК. В ОСК конструируется следующая ортогональная матрица:

$$G = \begin{pmatrix} \mathbf{H} & \frac{\mathbf{H} \times \mathbf{S}}{|\mathbf{H} \times \mathbf{S}|} & \frac{\mathbf{H} \times (\mathbf{H} \times \mathbf{S})}{|\mathbf{H} \times (\mathbf{H} \times \mathbf{S})|} \end{pmatrix}.$$

Аналогично по измеренным данным одного кадра вычисляется ортогональная матрица Υ в ССК:

$$\Upsilon = \begin{pmatrix} \mathbf{h} & \frac{\mathbf{h} \times \mathbf{s}}{|\mathbf{h} \times \mathbf{s}|} & \frac{\mathbf{h} \times (\mathbf{h} \times \mathbf{s})}{|\mathbf{h} \times (\mathbf{h} \times \mathbf{s})|} \end{pmatrix}.$$

Связь между измеренными и вычисленными значениями векторов даётся выражением

$$\Upsilon = A G.$$

Матрица A определяется как

$$A = \Upsilon G^T \tag{Д.5}$$

в силу ортогональности G .

206. С другой стороны, матрица направляющих косинусов перехода между связанной и орбитальной системами координат A выражается через компоненты кватерниона согласно следующей формуле:

$$A = \begin{bmatrix} \lambda_1^2 - \lambda_2^2 - \lambda_3^2 + \lambda_4^2 & 2(\lambda_1 \lambda_2 + \lambda_3 \lambda_4) & 2(\lambda_1 \lambda_3 - \lambda_2 \lambda_4) \\ 2(\lambda_1 \lambda_2 - \lambda_3 \lambda_4) & -\lambda_1^2 + \lambda_2^2 - \lambda_3^2 + \lambda_4^2 & 2(\lambda_2 \lambda_3 + \lambda_1 \lambda_4) \\ 2(\lambda_1 \lambda_3 + \lambda_2 \lambda_4) & 2(\lambda_2 \lambda_3 - \lambda_1 \lambda_4) & -\lambda_1^2 - \lambda_2^2 + \lambda_3^2 + \lambda_4^2 \end{bmatrix}.$$

Из матрицы можно вычислить 4 неизвестных компоненты λ_i , если добавить к трем независимым уравнениям условие нормировки кватерниона $\lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2 = 1$.

207. Таким образом, для нахождения первого приближения начальных условий можно использовать первые измерения магнитометра и солнечных датчиков и получить кватернион $\Lambda_0 = \Lambda(t_0)$ на момент, которому соответствует первый кадр телеметрии.

Для получения вектора угловой скорости в начальный момент ω_0 можно воспользоваться измерениями датчиков угловой скорости ω_m :

$$\omega_0 = \omega(t_0) = \omega_m(t_0).$$

208. Для интегрирования уравнений (Д.1), (Д.3) необходима информация для каждого шага интегрирования k следующих данных:

- текущего кинетического момента маховиков $\mathbf{h}(t_k)$,
- текущего управляющего момента со стороны маховиков $\dot{\mathbf{h}}(t_k)$,
- текущего расстояния КА до центра Земли $R(t_k)$.

209. В результате интегрирования уравнений движения для каждого момента, к которому привязан кадр телеметрии, будет найден кватернион поворота Λ_k . Зная ориентацию КА в пространстве в эти моменты времени, используя модель магнитного поля Земли, определяется вектор магнитного поля в ОСК \mathbf{b}_k^o . Тогда величина этого вектора в ССК вычисляется по формуле

$$\mathbf{b}_{t_k}^s = A_k \mathbf{b}_k^o.$$

210. Чтобы получить прогнозируемый вектор измерений, необходимо воспользоваться моделью измерений магнитометра (Д.4):

$$\mathbf{m}_{t_k} = \mathbf{b}_{t_k}^s + \boldsymbol{\mu}.$$

211. Далее, когда последовательность ожидаемых векторов измерения \mathbf{m}_{t_k} сформирована, можно составить функционал

$$\Phi = \sum_{k=1}^K (\mathbf{m}_{t_k} - \tilde{\mathbf{m}}_{t_k})^2, \quad (\text{Д.6})$$

где $\tilde{\mathbf{m}}_{t_k}$ – вектор измерения, взятый из кадра телеметрии, привязанного ко времени $t = t_k$, K – количество кадров телеметрии. Функционал Φ необходимо минимизировать по компонентам вектора смещения нуля:

$$\mathbf{f} = [\mu_1, \mu_2, \mu_3].$$

Для минимизации функционала (Д.6) необходимо решить систему из 3-х нелинейных уравнений

$$\frac{\partial \Phi}{\partial f_i} = 0, \quad i = 1, 2, 3, \quad (\text{Д.7})$$

где f_i – i -я компонента вектора параметров. Эту систему можно решить, например, итерационным методом Ньютона по формуле

$$\mathbf{f}_{i+1} = \mathbf{f}_i - F^{-1}(\mathbf{f}_i)\boldsymbol{\varphi}(\mathbf{f}_i), \quad (\text{Д.8})$$

где $\boldsymbol{\varphi} = \left[\frac{\partial \Phi}{\partial f_i} \right]$, $F = \left[\frac{\partial \varphi_i}{\partial f_j} \right]$ – якобиан системы (Д.7). Начальное значение

\mathbf{f}_0 можно принять нулевым. Приближения заканчиваются, когда $|\mathbf{f}_{i+1} - \mathbf{f}_i| < \varepsilon$, где ε – некоторая заданная величина, которая выбирается эмпирически.

212. Заметим, что так как начальное приближение для динамических уравнений вычисляется с помощью измерений, которые зависят в том числе и от смещения нуля магнитометра, то имеет смысл в итерационном подсчёте производных функционала Φ варьировать и начальную ориентацию согласно формуле (Д.5), которая строится по векторам \mathbf{s} и \mathbf{h} , записанным в ССК.

213. После каждой итерации нахождения вектора параметров желательно строить зависимость $\mathbf{m}_{i_k} - \tilde{\mathbf{m}}_{i_k}$ от времени, чтобы визуально наблюдать сходимость метода Ньютона. При правильно найденном векторе параметров \mathbf{f} математическое ожидание $\mathbf{m}_{i_k} - \tilde{\mathbf{m}}_{i_k}$ должно приближаться к нулю. Для оценки точности полученной оценки измерений требуется апостериорный анализ лежащих на результирующем решении измеренных отклонений.

Литература

1. Гаусс К.Ф. Сборник статей к столетию со дня смерти. — М.: АН СССР, 1956.
2. Уитеккер Э., Робинсон Г. Математическая обработка результатов наблюдений. — М.: ГТТИ, 1933.
3. Гельфанд И.М. Лекции по линейной алгебре. — М.-Л.: ОГИЗ ГИТТЛ, 1948.
4. Канторович Л.В., Акилов Г.П. Функциональный анализ. — М.: Наука, 1984.
5. Крылов А.Н. Лекции о приближённых вычислениях. — М.: ГИТТЛ, 1950.
6. Линник Ю.В. Метод наименьших квадратов и основы теории обработки наблюдений. — М.: Физ-мат. лит-ра, 1958.
7. Гнеденко Б.В. Теория вероятности. — М.: ГИТЛ, 1954.
8. Фадеев Д.Л., Фадеева В.Н. Вычислительные методы линейной алгебры. — М.: Физматгиз, 1960.
9. Энеев Т.М., Платонов А.К., Казакова Р.К. Определение параметров движения космического летательного аппарата по данным траекторных измерений. — М.: АН СССР, 1960.
10. Kalman R.E., Bucy R.S. New Results in Linear Filtering and Prediction Theory // *Journal of Basic Engineering*. — 1961. — V. 83. — P. 95–108.
11. Аким Э.Л., Энеев Т.М. Определение параметров летательного аппарата по данным траекторных измерений // *Космические исследования*. 1963. — Т. 1. — № 1. — С. 5–50.
12. Клаус А. О систематических ошибках в задачах определения траекторий // *Автоматическое управление космическими летательными аппаратами: сб.* — М.: Наука, 1968. — С. 296-308.
13. Genin Y. // *IEEE Trans. on Automatic and Control*. — 1968. — V. AC13. — № 1. — P. 103.
14. Линник Ю.В. Статистические задачи с мешающими параметрами. — М.: Наука, 1966.
15. Энеев Т.М. Некоторые вопросы применения метода наискорейшего спуска / препринт: ИПМ АН СССР. — 1970. — № 17. — 57 с.
16. Козлов Н.Н. // *Космические исследования*. — 1971. — Т. 9. — № 1. С. 3–22.
17. Мазо В.Л. // *Космические исследования*. — 1973. — Т. 11. — № 4. — С. 530.
18. Бритков В.Б. Определение оптимальной программы траекторных измерений с коррелированными ошибками / препринт: ИПМ АН СССР. — 1975. — № 67. — 28 с.

19. Бритков В.Б. Влияние мешающих параметров на оптимальную программу траекторных измерений / препринт: ИПМ АН СССР. — 1976. — № 108. — 23 с.

20. Колос М.В., Колос И.В. Методы линейной оптимальной фильтрации. — М.: Изд-во Московского университета, 2000.

21. Иванов Д.С., Карпенко С.О., Овчинников М.Ю. Алгоритм оценки параметров ориентации малого космического аппарата с использованием фильтра Калмана / препринт: ИПМ им. М.В. Келдыша РАН. — 2009. — № 48. — 32 с.

22. Иванов Д.С. Метод исследования алгоритмов определения ориентации микроспутников и его верификация / препринт: ИПМ им. М.В. Келдыша РАН. — 2013. — № 80. — 32 с.

23. Голован А.А., Парусников Н.А. Математические основы навигационных систем. Часть II. Приложения методов оптимального оценивания к задачам навигации. — М.: Изд-во Моск. ун-та, 2008.

ПРИЛОЖЕНИЯ

1. Доказательство тождества

$$(A + CBC^T)^{-1} \equiv A^{-1} - A^{-1}C(B^{-1} + C^T A^{-1}C)^{-1} C^T A^{-1}.$$

Умножим прямое левое на обратное правое части тождества:

$$\begin{aligned} \square (A + CBC^T)(A^{-1} - A^{-1}C(B^{-1} + C^T A^{-1}C)^{-1} C^T A^{-1}) &= \\ = E - C(B^{-1} + C^T A^{-1}C)^{-1} C^T A^{-1} + \\ + CBC^T A^{-1} - CBC^T A^{-1}C(B^{-1} + C^T A^{-1}C)^{-1} C^T A^{-1} &= \\ = [E + CBC^T A^{-1}][E - C(B^{-1} + C^T A^{-1}C)^{-1} C^T A^{-1}] &= \\ = [E + CBC^T A^{-1}][E - (A(C^T)^{-1} B^{-1} C^{-1} + E)^{-1}] &= \\ = (CBC^T A^{-1} + E)[E - ((CBC^T A^{-1})^{-1} + E)^{-1}] &= \\ = CBC^T A^{-1}(E + (CBC^T A^{-1})^{-1}) - CBC^T A^{-1} &= \\ = CBC^T A^{-1} + E - CBC^T A^{-1} = E. \quad \blacksquare \end{aligned}$$

2. Матрица H может быть приведена к виду

$$\begin{aligned} H_j = K_{a_{j-1}} U_j^T (U_j K_{a_{j-1}} U_j^T + K_{\psi_j})^{-1} &= (U_j + K_{\psi_j} (U_j^T)^{-1} K_{a_{j-1}}^{-1})^{-1} = \\ = [U_j^T K_{\psi_j}^{-1} U_j + K_{a_{j-1}}^{-1}]^{-1} U_j^T K_{\psi_j}^{-1}. \end{aligned}$$

2.1. Для нас более необходимо правое выражение.

Докажем его, используя то обстоятельство, что из тождества

$$(A + CBC^T)^{-1} \equiv A^{-1} - A^{-1}C(B^{-1} + C^T A^{-1}C)^{-1} C^T A^{-1}$$

при симметричных матрицах $C = A$ следует

$$(A + ABA)^{-1} \equiv A^{-1} - A^{-1}A(B^{-1} + AA^{-1}A)^{-1} AA^{-1} = A^{-1} - (B^{-1} + A)^{-1}.$$

Заменяя скобку обратной матрицы в формуле для H её тождественным выражением, получаем

$$\begin{aligned}
\square \quad H &= K_{\alpha_{j-1}} U_j^T (K_{\psi_j} + U_j K_{\alpha_{j-1}} U_j^T)^{-1} = \\
&= K_{\alpha_{j-1}} U_j^T (K_{\psi_j}^{-1} - K_{\psi_j}^{-1} U_j [K_{\alpha_{j-1}}^{-1} + U_j^T K_{\psi_j}^{-1} U_j]^{-1} U_j^T K_{\psi_j}^{-1}) = \\
&= K_{\alpha_{j-1}} U_j^T K_{\psi_j}^{-1} - K_{\alpha_{j-1}} U_j^T K_{\psi_j}^{-1} U_j [U_j^T K_{\psi_j}^{-1} U_j + K_{\alpha_{j-1}}^{-1}]^{-1} U_j^T K_{\psi_j}^{-1} = \\
&= K_{\alpha_{j-1}} U_j^T K_{\psi_j}^{-1} - (U_j^{-1} K_{\psi_j} (U_j^T)^{-1} K_{\alpha_{j-1}}^{-1})^{-1} [U_j^T K_{\psi_j}^{-1} U_j + K_{\alpha_{j-1}}^{-1}]^{-1} U_j^T K_{\psi_j}^{-1} = \\
&= K_{\alpha_{j-1}} U_j^T K_{\psi_j}^{-1} - [(U_j^T K_{\psi_j}^{-1} U_j + K_{\alpha_{j-1}}^{-1})(U_j^{-1} K_{\psi_j} (U_j^T)^{-1} K_{\alpha_{j-1}}^{-1})]^{-1} U_j^T K_{\psi_j}^{-1} = \\
&= \{K_{\alpha_{j-1}}^{-1} - [K_{\alpha_{j-1}}^{-1} + K_{\alpha_{j-1}}^{-1} U_j^{-1} K_{\psi_j} (U_j^T)^{-1} K_{\alpha_{j-1}}^{-1}]^{-1}\} U_j^T K_{\psi_j}^{-1} = \\
&= [(U_j^{-1} K_{\psi_j} (U_j^T)^{-1})^{-1} + K_{\alpha_{j-1}}^{-1}]^{-1} U_j^T K_{\psi_j}^{-1} = \\
&= [U_j^T K_{\psi_j}^{-1} U_j + K_{\alpha_{j-1}}^{-1}]^{-1} U_j^T K_{\psi_j}^{-1}. \blacksquare
\end{aligned}$$

Среднее равенство доказывается более просто:

$$\begin{aligned}
\square \quad H_j &= K_{\alpha_{j-1}} U_j^T (U_j K_{\alpha_{j-1}} U_j^T + K_{\psi_j})^{-1} = \\
&= ((U_j^T)^{-1} K_{\alpha_{j-1}}^{-1})^{-1} (U_j K_{\alpha_{j-1}} U_j^T + K_{\psi_j})^{-1} = \\
&= (U_j K_{\alpha_{j-1}} U_j^T + K_{\psi_j})(U_j^T)^{-1} K_{\alpha_{j-1}}^{-1})^{-1} = \\
&= (U_j + K_{\psi_j} (U_j^T)^{-1} K_{\alpha_{j-1}}^{-1})^{-1}. \blacksquare
\end{aligned}$$

3. Доказательство сводимости фильтра Калмана к способу наименьших квадратов

Для того чтобы убедиться в идентичности этих методов осреднения, нужно доказать, что

$$\alpha_j = [U_j^T K_{\psi_j}^{-1} U_j + K_{\alpha_{j-1}}^{-1}]^{-1} (U_j^T K_{\psi_j}^{-1} \psi_j + K_{\alpha_{j-1}}^{-1} \alpha_{j-1}) = \alpha_{j-1} + H(\psi_j - U_j \alpha_{j-1})$$

и что

$$K_{\alpha_j} = [E - H_j U_j] K_{\alpha_{j-1}} = [U_j^T K_{\psi_j}^{-1} U_j + K_{\alpha_{j-1}}^{-1}]^{-1}.$$

3.1. Приведём выражение фильтра в правой части верхнего равенства к выражению способа наименьших квадратов в его левой части:

$$\begin{aligned}
\square \quad \mathbf{a}_j &= \mathbf{a}_{j-1} + K_{\mathbf{a}_{j-1}} U_j^T (K_{\Psi_j} + U_j K_{\mathbf{a}_{j-1}} U_j^T)^{-1} (\Psi_j - U_j \mathbf{a}_{j-1}) = \\
&= [(U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1})^{-1} U_j^T K_{\Psi_j}^{-1}] \Psi_j + \\
&+ [E - (U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1})^{-1} U_j^T K_{\Psi_j}^{-1} U_j] \mathbf{a}_{j-1} = \\
&= [(U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1})^{-1} U_j^T K_{\Psi_j}^{-1}] \Psi_j + \\
&+ [(U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1})^{-1} (U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1}) - (U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1})^{-1} U_j^T K_{\Psi_j}^{-1} U_j] \mathbf{a}_{j-1} = \\
&= (U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1})^{-1} (U_j^T K_{\Psi_j}^{-1} \Psi_j + \\
&+ [U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1} - U_j^T K_{\Psi_j}^{-1} U_j] \mathbf{a}_{j-1}) = \\
&= (U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1})^{-1} (U_j^T K_{\Psi_j}^{-1} \Psi_j + K_{\mathbf{a}_{j-1}}^{-1} \mathbf{a}_{j-1}). \blacksquare
\end{aligned}$$

3.2. Наконец докажем, что

$$K_{\mathbf{a}_j} = [E - H_j U_j] K_{\mathbf{a}_{j-1}} = [U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1}]^{-1}.$$

$$\begin{aligned}
\square \quad [E - H_j U_j] K_{\mathbf{a}_{j-1}} &= \\
&= \{E - [U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1}]^{-1} U_j^T K_{\Psi_j}^{-1} U_j\} K_{\mathbf{a}_{j-1}} = \\
&= [U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1}]^{-1} \{U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1} - U_j^T K_{\Psi_j}^{-1} U_j\} K_{\mathbf{a}_{j-1}} = \\
&= [U_j^T K_{\Psi_j}^{-1} U_j + K_{\mathbf{a}_{j-1}}^{-1}]^{-1}. \blacksquare
\end{aligned}$$

4. Определение параметров эллипса рассеивания

На примере двумерной случайной величины $\{x, y\}$ эллипс рассеивания может быть задан следующими способами:

- матрицей коэффициентов уравнения эллипса:

$$K^{-1} = \begin{pmatrix} A & -C \\ -C & B \end{pmatrix},$$

тогда уравнение эллипса записывается как

$$Ax^2 + By^2 - 2Cxy = 1$$

или

$$\mathbf{r}^T K^{-1} \mathbf{r} = (x \quad y) \begin{pmatrix} A & -C \\ -C & B \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1;$$

- ковариационной матрицей K параметров рассеивания случайной величины $\{x, y\}$ с элементами матрицы σ_x^2 , σ_y^2 , $\sigma_{xy} = r\sigma_x\sigma_y$.

Величины недиагональных элементов σ_{xy} , стоящие на пересечении строк и столбцов x и y , однозначно определяют коэффициенты кор-

реляции r ($|r| \leq 1$) по величине корней квадратных из диагональных элементов. Матрица K имеет вид

$$K = \begin{pmatrix} \sigma_x^2 & r\sigma_x\sigma_y \\ r\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix};$$

- полуосями a, b и углом наклона α эллипса к оси x (см. рис. Г.1).

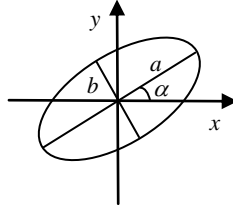


Рис. Г.1. Задание эллипса величинами полуосей и углом наклона диаметра

Определим связи между этими величинами.

- Определитель матрицы K равен: $\Delta_K = \sigma_x^2\sigma_y^2(1-r^2)$.

Тогда матрица K^{-1} равна

$$K^{-1} = \frac{1}{\Delta} \begin{pmatrix} \sigma_y^2 & -r\sigma_x\sigma_y \\ -r\sigma_x\sigma_y & \sigma_x^2 \end{pmatrix} = \frac{1}{1-r^2} \begin{pmatrix} \frac{1}{\sigma_x^2} & -\frac{r}{\sigma_x\sigma_y} \\ -\frac{r}{\sigma_x\sigma_y} & \frac{1}{\sigma_y^2} \end{pmatrix}.$$

Таким образом, получаем

$$A = \frac{1}{\sigma_x^2(1-r^2)}, \quad B = \frac{1}{\sigma_y^2(1-r^2)}, \quad C = \frac{r}{\sigma_x\sigma_y(1-r^2)}$$

и обратно:

$$\sigma_x = \frac{1}{\sqrt{A}\sqrt{1-r^2}}, \quad \sigma_y = \frac{1}{\sqrt{B}\sqrt{1-r^2}}, \quad r = \frac{C}{\sqrt{A}\sqrt{B}}.$$

- Для определения α приведём матрицу K^{-1} к диагональному виду поворотом на угол α . Матрица поворота R равна

$$R = \begin{pmatrix} \cos \alpha & +\sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}.$$

После поворота эллипса на угол α имеем диагональный вид матрицы K^{-1} :

$$\begin{pmatrix} \cos \alpha & +\sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} A & -C \\ -C & B \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ +\sin \alpha & \cos \alpha \end{pmatrix} = \begin{pmatrix} a^{-2} & 0 \\ 0 & b^{-2} \end{pmatrix}.$$

Это уравнение эллипса можно записать в следующем виде:

$$\mathbf{p}^T R R^{-1} R^T \mathbf{p} = \mathbf{p}^T \begin{pmatrix} a^{-2} & 0 \\ 0 & b^{-2} \end{pmatrix} \mathbf{p} = 1, \quad a^{-2} < b^{-2},$$

где $\mathbf{p} = R\mathbf{r}$. Из равенства элементов матриц имеем следующие соотношения:

$$\begin{cases} A \cos^2 \alpha + B \sin^2 \alpha - 2C \sin \alpha \cos \alpha = a^{-2}, \\ A \sin^2 \alpha + B \cos^2 \alpha + 2C \sin \alpha \cos \alpha = b^{-2}, \\ -A \cos \alpha \sin \alpha + B \cos \alpha \sin \alpha - C(\cos^2 \alpha - \sin^2 \alpha) = 0 \end{cases}$$

или после преобразования:

$$\begin{cases} A + B = a^{-2} + b^{-2}, \\ A \cos 2\alpha - B \cos 2\alpha - 2C \sin 2\alpha = a^{-2} - b^{-2}, \\ (A - B) \sin 2\alpha + 2C \cos 2\alpha = 0. \end{cases} \quad (\text{П.1})$$

Отсюда получаем

$$\operatorname{tg} 2\alpha = -\frac{2C}{A - B} \quad (\text{П.2})$$

(при $A = B$ и $C = 0$ получаем $\alpha = 0$). Далее из следующих соотношений

$$\begin{cases} (A - B) \cos 2\alpha - 2C \sin 2\alpha = a^{-2} - b^{-2}, \\ (A - B) \sin 2\alpha + 2C \cos 2\alpha = 0 \end{cases} \quad (\text{П.3})$$

получаем

$$(a^{-2} - b^{-2})^2 = (A - B)^2 + 4C^2. \quad (\text{П.4})$$

Преобразуем систему (П.3) к следующему соотношению:

$$(A - B) \cos 2\alpha - 2C \left(-\frac{2C}{A - B} \cos 2\alpha \right) = a^{-2} - b^{-2}.$$

Отсюда получаем

$$\cos 2\alpha = \frac{a^{-2} - b^{-2}}{A - B + \frac{4C^2}{A - B}} = (A - B) \frac{a^{-2} - b^{-2}}{(A - B)^2 + 4C^2} = (A - B) \frac{-\sqrt{(A - B)^2 + 4C^2}}{(A - B)^2 + 4C^2}.$$

Знак минус здесь возникает вследствие учёта знака в (П.3). Окончательно для угла α получаем выражения:

$$\cos 2\alpha = -\frac{A - B}{\sqrt{(A - B)^2 + 4C^2}}, \quad \sin 2\alpha = \frac{2C}{\sqrt{(A - B)^2 + 4C^2}}.$$

Для определения величин полуосей a и b , используя первое выражение из (П.1) и формулу (П.4), получим значения их обратных квадратов:

$$a^{-2} = \frac{A+B-\sqrt{(A-B)^2+4C^2}}{2},$$

$$b^{-2} = \frac{A+B+\sqrt{(A-B)^2+4C^2}}{2},$$

причем здесь $a^{-2} < b^{-2}$.

Связь коэффициентов уравнения эллипса с его параметрами определяется из уравнений (П.1) и (П.3):

$$A = \frac{a^{-2}+b^{-2}}{2} + \frac{a^{-2}-b^{-2}}{2} \cos 2\alpha,$$

$$B = \frac{a^{-2}+b^{-2}}{2} - \frac{a^{-2}-b^{-2}}{2} \cos 2\alpha,$$

$$2C = -(a^{-2}-b^{-2}) \sin 2\alpha.$$

Упражнение 12. Постройте алгоритмы определения размеров полуосей и углов их ориентации для трёхмерного эллипсоида.

Учебное издание

**Платонов Александр Константинович,
Иванов Данил Сергеевич**

МЕТОДЫ ОБРАБОТКИ ИЗМЕРЕНИЙ

Редакторы: *О. П. Котова, Л. В. Себова.*

Подписано в печать 25.12.2013. Формат 60 × 84 ¹/₁₆. Усл. печ. л. 6,75.
Уч.-изд. л. 6,5. Тираж 50 экз. Заказ № 414.

Федеральное государственное автономное образовательное учреждение
высшего профессионального образования
«Московский физико-технический институт (государственный университет)»
141700, Московская обл., г. Долгопрудный, Институтский пер., 9
Тел. (495) 408-58-22, e-mail: rio@mail.mipt.ru

Отдел оперативной полиграфии «Физтех-полиграф»
141700, Московская обл., г. Долгопрудный, Институтский пер., 9
Тел. (495) 408-84-30, e-mail: polygraph@mail.mipt.ru