

Б. Ц. БАХШИЯН, К. С. ФЕДЯЕВ

**ОСНОВЫ
КОСМИЧЕСКОЙ БАЛЛИСТИКИ
И НАВИГАЦИИ
КУРС ЛЕКЦИЙ**

Серия «**МЕХАНИКА, УПРАВЛЕНИЕ И ИНФОРМАТИКА**»

МОСКВА
2013

УДК 519.7
ББК 22.2
Б30

ISSN 2075-6839

Бахшиян Б. Ц., Федяев К. С. Основы космической баллистики и навигации : Курс лекций. М.: ИКИ РАН, 2013. 119 с. (Серия «Механика, управление и информатика»).

ISBN 978-5-9903101-4-8

Данный курс лекций представляет собой введение в космическую баллистику — теорию полета искусственных небесных тел. Приведены законы невозмущенного движения спутников, излагаются методы оценивания и коррекции их параметров. Приводятся примеры решения некоторых типовых задач.

Для студентов, аспирантов и преподавателей факультетов прикладной математики технических вузов, может быть полезно специалистам по небесной механике и астродинамике.

Ключевые слова: невозмущенное движение, законы Кеплера, элементы орбиты, возмущенное движение, теория оценивания, гарантирующее оценивание, линейное программирование, планирование эксперимента, коррекция движения, линейная импульсная коррекция.

Bakhshiyani B. Ts., Fedyaev K. S. Fundamentals of Space Ballistics and Navigation. M.: IKI RAN, 2013. 119 p. (Series “Mechanics, Control, Informatics”).

ISBN 978-5-9903101-4-8

The book contains lectures on the theory of space ballistics (satellite motion theory). Basic concepts of unperturbed satellite motion as well as its parameters determination and correction methods are discussed.

This book may be of interest to students, graduates and lecturers of applied mathematics departments of technical universities. Also it can be useful to specialists in celestial mechanics and astrodynamics.

Keywords: unperturbed motion, Kepler’s laws, orbital elements, perturbed motion, estimation theory, guaranteed estimation, linear programming, design of experiment, motion correction, linear impulse correction.

Редактор: Корниленко В. С.

Дизайн обложки: Захаров А. Н.

Компьютерная верстка: Комарова Н. Ю.

ОГЛАВЛЕНИЕ

Предисловие	5
1. Предварительные сведения	7
1.1. Что такое космическая баллистика?	7
1.2. Необходимые сведения из линейной алгебры	11
1.2.1. Векторы и матрицы.	11
1.2.2. Операции над векторами и матрицами	14
1.2.3. Скалярное и векторное произведение векторов	15
1.2.4. Вектор-функции и функции вектора.	17
1.3. Основные сведения из теории линейного программирования	18
1.3.1. Постановка задачи и симплекс-метод.	18
1.3.2. Алгоритм симплекс-метода.	20
1.3.3. Оценка близости текущего решения к оптимальному для некоторых типов задач	22
1.3.4. Геометрический способ решения двумерной задачи линейного программирования специального вида	22
2. Основные законы небесной механики	24
2.1. Законы Кеплера	24
2.2. Движение в гравитационном поле. Закон всемирного тяготения.	27
2.3. Основная задача небесной механики.	30
3. Невозмущенное движение.	33
3.1. Задача двух тел	33
3.2. Первые интегралы уравнения движения спутника	35
3.2.1. Интеграл энергии	35
3.2.2. Интеграл площадей.	36
3.2.3. Интеграл Лапласа	39
4. Невозмущенное движение спутника в плоскости орбиты	42
4.1. Уравнение орбиты спутника	42
4.2. Скорость спутника на орбите	45
4.3. Параметры орбиты спутника.	48
4.3.1. Эллиптические орбиты	48
4.3.2. Круговые орбиты.	51
4.3.3. Гиперболические орбиты.	52
4.3.4. Параболические орбиты	53
4.4. Время прохождения спутника через заданную точку орбиты ..	54
4.4.1. Эллиптические орбиты	54
4.4.2. Гиперболические орбиты.	58
4.5. Формула Ламберта.	60

5. Невозмущенное движение спутника в инерциальной системе координат.	64
5.1. Элементы орбиты	64
5.2. Определение положения и скорости спутника по элементам орбиты	66
5.3. Определение элементов орбиты спутника по положению и скорости.	67
6. Возмущенное движение спутника.	70
6.1. Понятие о возмущенном движении. Метод оскулирующих элементов	70
7. Задача определения движения	76
7.1. Модель оценивания.	76
7.2. Несмещенный алгоритм оценивания	77
7.3. Линеаризация модели измерений.	79
7.4. Одномерная линейная модель.	79
7.5. Линейный несмещенный алгоритм и метод наименьших квадратов.	80
7.6. Вычисление точности оценок при известной ковариационной матрице ошибок измерений. Теорема Гаусса— Маркова	82
8. Классические задачи планирования эксперимента	84
8.1. Усреднение модели измерений.	84
8.2. Постановка задачи.	86
8.3. Скалярная задача планирования эксперимента и алгоритм её решения	89
8.4. Оценивание параметров параболической траектории по измерениям дальности	91
9. Задача оптимальной линейной импульсной коррекции и алгоритм её решения	95
9.1. Задача оптимальной коррекции при известном корректируемом векторе	95
9.2. Проектная задача идеальной коррекции	101
9.3. Коррекция параболической траектории летательного аппарата.	101
10. Гарантированные характеристики точности	108
10.1. Критика классического подхода к оцениванию точности	108
10.2. Гарантирующий подход к вычислению точности оценивания	108
10.3. Сравнение решений задач оптимального оценивания в двух простейших случаях при гарантирующем и классическом подходах	111
10.3.1. Оптимизация гарантированной дисперсии D_1	111
10.3.2. Минимаксная задача оценивания при ограниченных по модулю ошибках измерений	112
10.4. Оптимизация гарантированных характеристик точности методом генерации столбцов.	114
Литература.	117

ПРЕДИСЛОВИЕ

В настоящее время, несмотря на значительный интерес к изучению небесной механики и теории полета искусственных спутников, весьма невелик список литературы, в которой теория и методы решения практических задач излагались бы достаточно подробно, в форме, доступной для начинающих изучение этих дисциплин. При этом значительная часть таких изданий была осуществлена много лет назад, и сегодня стала трудно доступной. Настоящее издание имеет целью восполнить этот дефицит. Оно содержит материал для вводных лекций по основам космической баллистики, в котором подробно излагаются основные понятия теории движения искусственных спутников, а также рассматриваются методы оценивания и коррекции параметров этого движения.

В разделах 1–6 излагается терминологическая база и приводятся справочные сведения об используемом математическом аппарате, затем подробно рассматриваются вопросы невозмущенного движения искусственных спутников, дается определение возмущенного движения, рассматривается решение типовых примеров.

В разделах 7–9 рассматриваются постановки и излагаются методы решения задач оценивания и коррекции параметров движения спутников. Сюда вошли основные разделы книги [Бахшиян, 2012] за исключением некоторых разделов специального характера, которые, как правило, труднее воспринимаются студентами и могут быть опущены в рамках вводного курса.

Раздел 10 посвящен проблеме нахождения точности оценок параметров движения на основе гарантирующего подхода. Приводятся методы оценивания точности при различной информации о ковариационной матрице вектора ошибок измерений.

С целью облегчить восприятие материала в некоторых местах в текст включены короткие исторические отступления и биографические сноски, материал для которых брался из источников [БСЭ, 1969–1978; Википедия; Савченко, 2010].

Данный материал предназначен для чтения вводных лекций по курсу космической баллистики и не претендует на полноту описания всех вопросов, традиционно рассматриваемых в рамках этой дисциплины. Для дальнейшего и более глубокого изучения курса могут быть рекомендованы издания [Белецкий, 2009; Бордовицына, Авдюшев, 2007; Иванов, Лысенко, 2004; Охоцимский, Сихарулидзе, 1990; Сихарулидзе, 2011; Суханов, 2010; Эльясберг, 1965]. Большое количество задач по различным разделам космической баллистики можно найти в сборниках [Авдеев и др., 1965; Богачев и др., 2001].

К сожалению, эта книга выходит в свет уже после кончины Б. Ц. Бахшияна, замечательного ученого и педагога. В написанных им п. 1.3 и разделах 7–10 при подготовке текста к печати была сделана попытка в максимальной степени сохранить стиль и последовательность изложения материала, были лишь добавлены сноски и отдельные пояснения.

Хочется выразить особую благодарность В. И. Прохоренко за внимательное прочтение рукописи и ряд существенных замечаний, а также А. А. Савченко за помощь в работе над текстом лекций и полезные советы.

К. С. Федяев

1. ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ

1.1. ЧТО ТАКОЕ КОСМИЧЕСКАЯ БАЛЛИСТИКА?

Звездное небо, наверное, было одним из самых первых объектов изучения человека. С глубокой древности люди всматривались в скопления звезд, пытались разгадать их тайны. Почему одни звезды светятся ярко, а другие едва заметны? Почему одни звезды кажутся неподвижными, а другие перемещаются по небосклону? На эти и многие другие вопросы люди веками пытались найти ответы — наблюдали за звездами, делали самые разные догадки и предположения, создавали как научные объяснения, так и легенды, которые передавались из поколения в поколение. С развитием человечества наши знания о космосе становились все более совершенными: накапливались наблюдения, на их основе появлялись теории, все точнее и точнее объясняющие строение космического пространства. Так развивалась *астрономия* — наука о Вселенной, изучающая расположение, движение, строение, происхождение и развитие небесных тел и образованных ими систем [Кононович, Мороз, 2009].

Со временем из астрономии выделилось особое научное направление, изучающее законы движения небесных тел, — *небесная механика*.

Термин «небесная механика» был впервые предложен французским астрономом и математиком Пьером Симоном Лапласом в 1798 году, а основой небесной механики стал закон всемирного тяготения великого английского физика и математика Исаака Ньютона. Ньютон сформулировал этот закон, основываясь на трудах своих предшественников, среди которых особое место занимал немецкий астроном Иоганн Кеплер, опытным путем открывший в начале XVII века три основных закона движения небесных тел. Ньютон показал, что законы Кеплера являются следствиями закона всемирного тяготения,

и тем самым впервые дал им теоретическое объяснение. Кроме законов Кеплера Ньютону удалось обосновать такие необъяснимые ранее явления как некоторые особенности движения Луны (вариация, попятное движение узлов и т.д.), прецессию и сжатие Юпитера, чередование на Земле приливов и отливов.

Дальнейшее развитие небесной механики связано с именами Ж. Лагранжа, П. Лапласа, У. Леверье и др. В середине XIX века этими учеными были разработаны классические методы теории возмущений, созданы теория движения больших планет, теория движения Луны.

Начало освоения космоса в XX веке привело к новому витку развития науки о космосе. Наряду с появлением новых возможностей и новых научных данных возникла необходимость изучать движение в космическом пространстве искусственных объектов — спутников, межпланетных станций и др., в том числе управляемых. Научное направление, занимающееся этими проблемами, получило название *космической баллистики*.

Определение 1.1. *Космическая баллистика* — это наука о движении в космическом пространстве различных искусственных объектов.

Замечание 1.1. Термин «космическая баллистика» получил распространение в основном только в отечественной литературе и является, вообще говоря, не совсем точным. Его появление связано с тем, что в нашей стране изначально разработкой теории движения искусственных спутников стали заниматься баллистики — военные специалисты в области движения реактивных снарядов и разработки артиллерийского вооружения. Создавая новое научное направление, они назвали его привычным термином «баллистика», прибавив к нему определение «космическая» [Иванов, Лысенко, 2004]. В зарубежной науке это научное направление получило название “orbital mechanics”, “flight mechanics” или “astrodynamics”. Соответственно, русские переводы этих терминов («орбитальная механика», «механика полета», «аэродинамика») стали употребляться одновременно с термином «космическая баллистика» в качестве его синонимов. Также иногда используются названия «наука о движении искусственных спутников Земли», «прикладная небесная механика», «космическая динамика» и др.

Как ясно из названия, в основе космической баллистики наряду с небесной механикой лежит *баллистика* — военно-тех-

ническая наука о движении различных искусственных тел (артиллерийских снарядов, пуль, мин, авиабомб, реактивных снарядов, гарпунов и т. п.) [БСЭ, 1969—1978, т. 2, 1970]. При этом различают *внутреннюю баллистику*, занимающуюся изучением движения снаряда или другого искусственного тела внутри орудийного ствола, и *внешнюю баллистику*, которая рассматривает неуправляемое движение искусственного тела в пространстве после вылета из орудия. Понятно, что это движение является, вообще говоря, одним из случаев движения искусственного тела в гравитационном поле (в данном случае в поле тяготения Земли). Поэтому при описании движения искусственных объектов в космосе наряду с методами классической небесной механики вполне могут быть применимы и методы внешней баллистики.

Когда снаряд приведен в движение, его центр масс очерчивает в пространстве кривую, называемую *траекторией*. Одна из главных задач внешней баллистики (так называемая *прямая задача баллистики*) состоит в том, чтобы по известным начальным условиям движения снаряда рассчитать его траекторию, определив положение центра масс и пространственное положение снаряда как функции времени полета (времени после запуска). Для этого нужно решить систему уравнений, в которых учитывались бы силы и моменты сил, действующие на снаряд. Важнейшее значение имеет и другая задача (*обратная задача баллистики*) — определение начальных условий движения снаряда, необходимых для достижения заданной цели полета.

Замечание 1.2. Одновременно с термином «траектория» в небесной механике часто употребляется термин «*орбита*». Многие авторы вовсе не делают различия между этими терминами, считая их синонимами. Однако в ряде источников указывается, что термин «орбита» более правильно использовать при описании относительного движения нескольких небесных тел: «орбита Сатурна» (подразумевается движение этой планеты вокруг Солнца), «орбита астероида при сближении с Землей», «околосолнечная орбита космического зонда» и т. д. Термин «траектория», который можно применить к любой материальной точке, чаще используется, когда речь идет просто о движении: «траектория кометы в межзвездном пространстве», «траектория перелета спутника между двумя орбитами» и т. д.

Перечислим ряд важнейших задач, решаемых в рамках космической баллистики и обусловленных спецификой движения в космосе искусственных объектов, следуя книге [Балк, 1965]:

- 1 *Расчёт движения космического аппарата вблизи большого небесного тела.* Рассматривается движение космического аппарата на небольшом расстоянии от поверхности большого небесного тела (планеты), когда это небесное тело уже нельзя считать материальной точкой. При расчете траектории в этом случае необходимо учитывать несимметричность формы небесного тела. Так, например, для спутников, движущихся на расстоянии менее 40 тысяч километров от центра Земли, пренебрежение несферичностью формы Земли (эффект сжатия) приводит к ошибке, большей, чем ошибка в случае пренебрежения влиянием Луны и Солнца.
- 2 *Расчёт движения космического аппарата в вязкой среде.* При движении спутника вблизи планеты ее атмосфера может оказывать значительное тормозящее воздействие и приводить к существенному изменению траектории аппарата. Поэтому при построении траектории спутника вблизи планеты необходимо знать свойства ее атмосферы, чтобы избежать ошибок в расчетах.
- 3 *Применение космических маневров.* Основным способом изменения траектории космического аппарата является кратковременное (импульсное) включение ракетного двигателя, приводящее к резкому изменению параметров траектории аппарата. Переход аппарата с первоначальной траектории на иную вследствие импульсного воздействия называется *космическим маневром*.
- 4 *Расчет движения аппарата с переменной массой.* Использование двигателей в процессе движения спутника приводит к уменьшению его массы, что также необходимо учитывать при расчете траектории движения.
- 5 *Использование двигателей с малой тягой.* При движении аппарата могут использоваться не только импульсные воздействия мощных двигателей, но и двигатели небольшой мощности, работающие в течение длительного времени, в том числе с использованием альтернативных видов топлива.
- 6 *Учёт других факторов.* К прочим факторам, оказывающим влияние на движение космического аппарата, относятся: движение атмосферы, собственное вращение аппарата относительно своего барицентра, влияние магнитного поля планеты, солнечной радиации и др.

В рамках космической баллистики выделяют два направления: *фундаментальную* космическую баллистику, которая зани-

мается вопросами, связанными с теорией полета искусственных небесных тел, и *прикладную* космическую баллистику, которая рассматривает вопросы практического построения управляемых космических аппаратов, занимается расчетами траектории их движения, осуществляет подбор требуемых характеристик и др.

Часто в космической баллистике в качестве отдельного направления выделяется *космическая навигация* — *наука об управлении движением космических аппаратов*. Иногда, впрочем, термин «космическая навигация» рассматривается в более узком смысле как задача об определении положения и скорости движения искусственного спутника по результатам обработки измерительной информации.

К основным задачам космической навигации относятся:

1. *Определение движения* — задача оценивания параметров движения космического аппарата в результате обработки измерительной информации.
2. *Коррекция движения* — задача об исправлении траектории космического аппарата с целью достижения заданной цели полета.

В настоящее время космическая баллистика продолжает активно развиваться в рамках различных научных школ, используя различные подходы, пополняясь новыми методами, теоретическими и практическими результатами. Поэтому в литературе по космической баллистике нередко до сих пор отсутствует общепринятая терминология, одно и то же понятие может обозначаться различными терминами, один и тот же термин может иметь в разных источниках различный смысл. В изложении материала мы использовали в основном терминологию, принятую в отечественной литературе, указывая при необходимости альтернативные термины из зарубежных источников.

1.2. НЕОБХОДИМЫЕ СВЕДЕНИЯ ИЗ ЛИНЕЙНОЙ АЛГЕБРЫ

1.2.1. ВЕКТОРЫ И МАТРИЦЫ

Определение 1.2. *Вектором* или *вектором-столбцом* $\mathbf{a}^*$ в m -мерном евклидовом пространстве $\mathbb{R}^m$ (или, более кратко,

* Здесь и всюду далее векторные и матричные величины будут выделяться жирным шрифтом. При этом векторы будем обозначать строчными буквами, матрицы — прописными.

m -мерным вектором) будем называть упорядоченный столбец чисел $a_1, a_2, \dots, a_m$, называемых *компонентами* или *координатами* вектора $\mathbf{a}$:

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}.$$

Длиной, абсолютной величиной или *модулем* вектора $\mathbf{a}$ будем называть число, равное евклидовой норме этого вектора:

$$a = |\mathbf{a}| = \|\mathbf{a}\|_m = \sqrt{a_1^2 + a_2^2 + \dots + a_m^2}.$$

В трехмерном пространстве $\mathbb{R}^3$ под вектором обычно понимают направленный отрезок прямой. Если A и B — точки, являющиеся соответственно началом и концом вектора $\mathbf{a}$, $(A_x, A_y, A_z)'$, $(B_x, B_y, B_z)'$ — координаты этих точек в прямоугольной декартовой системе координат, то координатами вектора $\mathbf{a}$ будет столбец чисел $B_x - A_x, B_y - A_y, B_z - A_z$:

$$\mathbf{a} = \begin{pmatrix} B_x - A_x \\ B_y - A_y \\ B_z - A_z \end{pmatrix} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}, \quad (1.1)$$

a длиной — число

$$a = |\mathbf{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2}. \quad (1.2)$$

Если задано положение в пространстве точки A , т.е. известны ее координаты A_x, A_y и A_z в некоторой системе координат, то вектор (1.1) однозначно определяет пространственное положение точки B и называется *радиус-вектором* этой точки относительно точки A . Саму точку A в этом случае часто называют *полюсом*. Если полюс совпадает с началом координат, то координаты точки B будут совпадать с координатами радиус-вектора этой точки.

Вектор, длина которого равна единице, будем называть *единичным*. Вектор, длина которого равна нулю, будем называть *нулевым*. Очевидно, что вектор является нулевым тогда и только тогда, когда все его компоненты равны нулю.

Будем говорить, что *вектор $\mathbf{b}$ больше вектора $\mathbf{a}$* , если каждая компонента вектора $\mathbf{b}$ больше соответствующей компоненты вектора $\mathbf{a}$:

$$\mathbf{b} > \mathbf{a} \Leftrightarrow b_1 > a_1, b_2 > a_2, \dots, b_n > a_n.$$

Остальные операции сравнения векторов определяются аналогично.

Определение 1.3. Матрицей $\mathbf{A}$ называется прямоугольная таблица чисел, содержащая m строк и n столбцов:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Числа, входящие в состав матрицы, называются ее *элементами*. Каждый элемент матрицы имеет два индекса — номер строки и номер столбца, которым принадлежит этот элемент. Число элементов матрицы, равное произведению числа строк m на число столбцов n , называется *размерностью матрицы*: будем говорить, что матрица $\mathbf{A}$ имеет размерность $m \times n$, записывая это символически в виде

$$\mathbf{A} = (m \times n).$$

Также матрицу $\mathbf{A}$ можно рассматривать как совокупность из n векторов $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \in \mathbb{R}^m$:

$$\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n).$$

Матрица размерности $m \times 1$, очевидно, представляет собой вектор-столбец размерности m . По аналогии матрицу размерности $1 \times n$ будем называть *вектором-строкой* размерности n .

Матрица, все элементы которой равны нулю, называется *нулевой матрицей*.

Определение 1.4. Совокупность элементов матрицы, у которых номер строки совпадает с номером столбца, называется *главной диагональю* этой матрицы. Элементы главной диагонали называются *диагональными*, а все остальные элементы — *внедиагональными*.

Матрица, у которой все внедиагональные элементы равны нулю, называется *диагональной*.

Матрица $\mathbf{A}$ размерности $n \times n$ называется *квадратной матрицей порядка n* .

Квадратная матрица, у которой все диагональные элементы равны единице, а все внедиагональные элементы равны нулю, называется *единичной матрицей*.

Определение 1.5. Пусть $\mathbf{A}$ — квадратная матрица порядка n . *Определителем* матрицы $\mathbf{A}$ называется действительная функция элементов a_{ij} этой матрицы

$$\det \mathbf{A} = \sum \pm a_{i_1} a_{i_2} \dots a_{i_n},$$

где суммирование ведется по всевозможным перестановкам $(i_1, i_2, \dots, i_n)$ чисел $(1, 2, \dots, n)$. Знак каждого сомножителя определяется четностью перестановки [Бахшиян и др., 1980].

1.2.2. ОПЕРАЦИИ НАД ВЕКТОРАМИ И МАТРИЦАМИ

Определение 1.6. Пусть $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \in \mathbb{R}^m$ — некоторые векторы. *Линейной комбинацией* этих векторов называется вектор

$$\mathbf{b} = x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_n \mathbf{a}_n, \quad (1.3)$$

где $x_1, x_2, \dots, x_n$ — некоторые действительные числа. Если выполняются условия

$$x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0, \quad x_1 + x_2 + \dots + x_n = 1,$$

то линейная комбинация (1.3) называется *выпуклой*.

Определение 1.7. Векторы $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \in \mathbb{R}^m$ называются *линейно независимыми*, если равенство $x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_n \mathbf{a}_n = \mathbf{0}$ выполняется тогда и только тогда, когда $x_1 = 0, x_2 = 0, \dots, x_n = 0$.

Определение 1.8. Рангом матрицы $\mathbf{A}$ размерности $m \times n$ называется число $\text{rang } \mathbf{A}$, равное максимальному числу линейно независимых строк (столбцов) этой матрицы.

Очевидно, что $\text{rang } \mathbf{A} \leq \min\{m, n\}$.

Пусть $\mathbf{A}$ — матрица размерности $m \times n$. Транспонированной по отношению к $\mathbf{A}$ называется матрица $\mathbf{A}'$ размерности $n \times m$ с элементами

$$a'_{ij} = a_{ji}, \quad 1 \leq i \leq n, \quad 1 \leq j \leq m,$$

т. е. матрица, строки которой являются столбцами матрицы $\mathbf{A}$.

Очевидно, что $(\mathbf{A}')' = \mathbf{A}$.

Кроме того, если для матриц $\mathbf{A}$ и $\mathbf{B}$ определены соответствующие операции суммы или произведения, то справедливы тождества

$$\begin{aligned}(\mathbf{A} + \mathbf{B})' &= \mathbf{A}' + \mathbf{B}', \\ (\mathbf{AB})' &= \mathbf{B}'\mathbf{A}'.\end{aligned}$$

Определение 1.9. Пусть $\mathbf{A}$ — квадратная матрица порядка n . Квадратная матрица $\mathbf{A}^{-1}$ того же порядка n называется *обратной к матрице $\mathbf{A}$* , если справедливо равенство

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{E},$$

где $\mathbf{E}$ — единичная матрица порядка n .

Квадратная матрица $\mathbf{A}$ порядка n имеет обратную матрицу тогда и только тогда, когда $\text{rang } \mathbf{A} = n$.

Для любой квадратной матрицы $\mathbf{A}$, имеющей обратную матрицу $\mathbf{A}^{-1}$, справедливы равенства

$$\begin{aligned}\det \mathbf{A}^{-1} &= \frac{1}{\det \mathbf{A}}, \\ (\mathbf{A}')^{-1} &= (\mathbf{A}^{-1})'.\end{aligned}$$

Если $\mathbf{A}$ и $\mathbf{B}$ — две квадратные матрицы одинакового порядка, и существуют матрицы $\mathbf{A}^{-1}$ и $\mathbf{B}^{-1}$, то

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

1.2.3. СКАЛЯРНОЕ И ВЕКТОРНОЕ ПРОИЗВЕДЕНИЕ ВЕКТОРОВ

Пусть $\mathbf{a}$ и $\mathbf{b}$ — два вектора в n -мерном евклидовом пространстве $\mathbb{R}^n$, $(a_1, a_2, \dots, a_n)'$, $(b_1, b_2, \dots, b_n)'$ — координаты этих векторов в прямоугольной декартовой системе координат. Тогда *скалярное произведение* векторов $\mathbf{a}$ и $\mathbf{b}$ представляет собой число $\mathbf{a} \cdot \mathbf{b}$, равное сумме попарных произведений соответствующих координат этих векторов:

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \dots + a_n b_n = \mathbf{a}' \mathbf{b}. \quad (1.4)$$

В трехмерном пространстве скалярное произведение векторов $\mathbf{a}$ и $\mathbf{b}$ равно произведению длин этих векторов на косинус угла между ними:

$$\mathbf{a} \cdot \mathbf{b} = ab \cos \alpha, \quad (1.5)$$

где a и b — длины векторов $\mathbf{a}$ и $\mathbf{b}$, α — угол между этими векторами.

Из последней формулы, в частности, следует, что скалярное произведение двух векторов ненулевой длины равно нулю тогда и только тогда, когда эти векторы перпендикулярны.

Очевидно, что скалярное произведение векторов *коммутативно*:

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}.$$

Также отметим, что скалярное произведение вектора на самого себя (скалярный квадрат вектора) равно квадрату длины этого вектора:

$$\mathbf{a} \cdot \mathbf{a} = a_1^2 + a_2^2 + \dots + a_n^2 = a^2.$$

Пусть $\mathbf{a}$ и $\mathbf{b}$ — два вектора в трехмерном евклидовом пространстве, α — угол между ними. Тогда *векторное произведение* векторов $\mathbf{a}$ и $\mathbf{b}$ представляет собой вектор, обозначаемый $\mathbf{a} \times \mathbf{b}$, перпендикулярный векторам $\mathbf{a}$ и $\mathbf{b}$, образующий с ними правую тройку векторов и имеющий длину

$$|\mathbf{a} \times \mathbf{b}| = ab \sin \alpha, \quad (1.6)$$

где a и b — длины векторов $\mathbf{a}$ и $\mathbf{b}$, α — угол между ними.

Если в трехмерном пространстве задана прямоугольная декартова система координат, в которой $\mathbf{a} = (a_x, a_y, a_z)'$, $\mathbf{b} = (b_x, b_y, b_z)$, то векторное произведение этих векторов представляет собой вектор с координатами

$$\mathbf{a} \times \mathbf{b} = \begin{pmatrix} a_y b_z - a_z b_y \\ a_z b_x - a_x b_z \\ a_x b_y - a_y b_x \end{pmatrix}. \quad (1.7)$$

Из последних выражений следует, что векторное произведение двух векторов ненулевой длины равно нулю тогда и только тогда, когда эти векторы коллинеарны, т.е. лежат на параллельных прямых.

Для векторного произведения справедливо равенство

$$\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a},$$

т.е. векторное произведение *антикоммутативно*.

Двойным векторным произведением векторов $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ называется векторное произведение вектора $\mathbf{a}$ на векторное произведение векторов $\mathbf{b}$ и $\mathbf{c}$. Для двойного векторного произведения справедливо тождество Лагранжа

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}). \quad (1.8)$$

Смешанным произведением векторов $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ называется скалярное произведение вектора $\mathbf{a}$ на векторное произведение векторов $\mathbf{b}$ и $\mathbf{c}$:

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}).$$

Смешанное произведение удовлетворяет соотношению

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{b}, \mathbf{c}, \mathbf{a}) = (\mathbf{c}, \mathbf{a}, \mathbf{b}) = -(\mathbf{b}, \mathbf{a}, \mathbf{c}) = -(\mathbf{c}, \mathbf{b}, \mathbf{a}) = -(\mathbf{a}, \mathbf{c}, \mathbf{b}). \quad (1.9)$$

1.2.4. ВЕКТОР-ФУНКЦИИ И ФУНКЦИИ ВЕКТОРА

Определение 1.10. Вектор $\mathbf{f} \in \mathbb{R}^n$, компоненты которого $f_1, f_2, \dots, f_n$ представляют собой функции скалярного аргумента x или векторного аргумента $\mathbf{x} \in \mathbb{R}^m$, будем называть *n -мерной вектор-функцией* аргумента x или $\mathbf{x}$ соответственно.

Определение 1.11. Производной n -мерной вектор-функции $\mathbf{f}(\mathbf{x})$ по вектору $\mathbf{x} \in \mathbb{R}^m$ будем называть матрицу размерности $n \times m$

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_m} \end{pmatrix},$$

если все элементы такой матрицы существуют.

В частности, производная от вектор-функции скалярного аргумента будет представлять собой вектор-столбец

$$\frac{\partial \mathbf{f}}{\partial x} = \begin{pmatrix} \frac{\partial f_1}{\partial x} \\ \frac{\partial f_2}{\partial x} \\ \vdots \\ \frac{\partial f_n}{\partial x} \end{pmatrix},$$

а производная скалярной функции векторного аргумента — вектор-строку

$$\frac{\partial f}{\partial \mathbf{x}} = \left(\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \dots \quad \frac{\partial f}{\partial x_m} \right).$$

1.3. ОСНОВНЫЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

1.3.1. ПОСТАНОВКА ЗАДАЧИ И СИМПЛЕКС-МЕТОД

Определение 1.12. Задачей линейного программирования (ЛП) будем называть задачу нахождения минимума линейной функции $z(\mathbf{x})$ векторного аргумента $\mathbf{x} \in \mathbb{R}^n$

$$z(\mathbf{x}) = \mathbf{c} \cdot \mathbf{x} = c_1 x_1 + c_2 x_2 + \dots + c_n x_n,$$

где $c_1, c_2, \dots, c_n$ — числовые коэффициенты, а компоненты $x_1, x_2, \dots, x_n$ вектора $\mathbf{x}$ связаны линейными соотношениями

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \dots + \mathbf{a}_n x_n = \mathbf{b},$$

где $\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{b} \in \mathbb{R}^m$. Функция $z(\mathbf{x})$ называется *целевой функцией задачи ЛП*, а ее аргументы $x_1, \dots, x_n$ — *переменными* задачи ЛП.

В сокращенной форме задачу линейного программирования будем записывать в виде

$$z^* = \min_{\mathbf{x}} \left\{ z(\mathbf{x}) = \sum c_i x_i : \sum \mathbf{a}_i x_i = \mathbf{b}, x_i \geq 0, i = 1, \dots, n \right\}, \quad (1.10)$$

где z^* — искомое минимальное (оптимальное) значение целевой функции $z(\mathbf{x})$.

В векторно-матричной форме задача (1.10) имеет вид

$$z^* = \min_{\mathbf{x}} \left\{ z(\mathbf{x}) = \mathbf{c}' \mathbf{x} : \mathbf{A} \mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \right\}, \quad (1.11)$$

где $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{A}$ — матрица коэффициентов ограничений размерности $m \times n$, $\mathbf{b} \in \mathbb{R}^m$ — вектор правых частей ограничений.

В дальнейшем будем полагать $n \geq m$ и $\text{rang } \mathbf{A} = m$ (невыполнение условия $n \geq m$ может привести к неразрешимости системы ограничений, а условие $\text{rang } \mathbf{A} < m$ говорит о наличии избыточных ограничений либо об их несовместности).

Выделим в матрице $\mathbf{x}$ невырожденную подматрицу $\mathbf{B}$ размерности $m \times m$, которую будем называть *базисной матрицей* или, короче, *базисом*. Перенумеровав при необходимости переменные, приведем матрицу $\mathbf{A}$ к виду

$$\mathbf{A} = (\mathbf{B}, \mathbf{N}),$$

где $\mathbf{N}$ — матрица размерности $m \times (n - m)$. Соответственно разобьем векторы $\mathbf{x}$ и $\mathbf{c}$:

$$\mathbf{x}' = (\mathbf{x}'_B, \mathbf{x}'_N), \quad \mathbf{c}' = (\mathbf{c}'_B, \mathbf{c}'_N).$$

Вектор $\mathbf{x} = (\mathbf{x}'_B, 0, \dots, 0)'$, такой что $\mathbf{Ax} = \mathbf{b}$ и $\mathbf{x} \geq \mathbf{0}$, будем называть *допустимым базисным вектором* задачи ЛП, соответствующим базису $\mathbf{B}$, который также назовем *допустимым*. Вектор $\mathbf{x}^*$ такой что $\mathbf{c}'\mathbf{x}^* = z^*$, будем называть *оптимальным вектором* или *решением* задачи (1.10). Можно доказать, что среди оптимальных векторов задачи (1.10) всегда есть базисный. Базис, соответствующий оптимальному вектору, будем также называть *оптимальным*.

Таким образом, поиск решения задачи (1.10) может быть сведен к перебору всевозможных базисных векторов и выбору среди них того, для которого значение целевой функции $z(\mathbf{x})$ минимально. Идея *симплекс-метода* состоит в том, чтобы сделать такой перебор направленным, рассматривая на каждом последующем шаге только те базисные векторы, на которых значение целевой функции не больше, чем на текущем.

Представим вектор $\mathbf{x}_B$ в виде

$$\mathbf{x}_B = \mathbf{B}^{-1}(\mathbf{b} - \mathbf{N}\mathbf{x}_N).$$

Полагая $\mathbf{x}_N = \mathbf{0}$, получим выражение для вектора $\mathbf{x}_B$, соответствующего базисному вектору:

$$\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b}. \quad (1.12)$$

Введем вспомогательный вектор $\boldsymbol{\pi}$ из условия

$$\boldsymbol{\pi}'\mathbf{B} = \mathbf{c}'_B.$$

Тогда справедливо равенство $\mathbf{c}'_B\mathbf{x}_B = \mathbf{c}'_B\mathbf{B}^{-1}\mathbf{b} = \boldsymbol{\pi}'\mathbf{b}$.

Пусть $\mathbf{x}$ — текущий допустимый базисный вектор, соответствующий базису $\mathbf{B}$. Рассмотрим произвольный вектор $\tilde{\mathbf{x}} \neq \mathbf{x}$ и найдем $z(\tilde{\mathbf{x}})$:

$$\begin{aligned} z(\tilde{\mathbf{x}}) &= \mathbf{c}'\tilde{\mathbf{x}} = \mathbf{c}'_B\tilde{\mathbf{x}} + (\mathbf{c}'_N\tilde{\mathbf{x}}_N - \boldsymbol{\pi}'\mathbf{b}) = \mathbf{c}'_B\tilde{\mathbf{x}}_B + (\mathbf{c}'_N\tilde{\mathbf{x}}_N - \boldsymbol{\pi}'\mathbf{A}\tilde{\mathbf{x}}) = \\ &= \mathbf{c}'_B\tilde{\mathbf{x}}_B + (\mathbf{c}' - \boldsymbol{\pi}'\mathbf{A})\tilde{\mathbf{x}} = z(\mathbf{x}) + \boldsymbol{\Delta}'\tilde{\mathbf{x}}, \end{aligned} \quad (1.13)$$

где $\boldsymbol{\Delta}' = \mathbf{c}' - \boldsymbol{\pi}'\mathbf{A}$. Отсюда следует, что текущий базис $\mathbf{B}$ оптимален тогда и только тогда, когда

$$\boldsymbol{\Delta}'\tilde{\mathbf{x}} \geq 0 \quad \forall \tilde{\mathbf{x}} \neq \mathbf{x}. \quad (1.14)$$

В частности, *достаточным условием оптимальности* является выполнение неравенства

$$\Delta \geq 0.$$

Предположим, что это достаточное условие не выполняется, т. е. существует номер s такой, что $\Delta_s < 0$. При увеличении от нуля значения компоненты x_s вектора $\mathbf{x}$, как следует из формулы (1.13), значение целевой функции будет уменьшаться. При этом, однако, должно выполняться условие

$$\mathbf{B}\tilde{\mathbf{x}}_B + \mathbf{a}_s \tilde{x}_s = \mathbf{b},$$

откуда

$$\tilde{\mathbf{x}}_B = \mathbf{B}^{-1}(\mathbf{b} - \mathbf{a}_s \tilde{x}_s) \geq 0$$

или

$$\tilde{\mathbf{x}}_B = \mathbf{x}_B - \alpha \tilde{x}_s \geq 0,$$

где $\alpha = \mathbf{B}^{-1} \mathbf{a}_s$. Из последнего неравенства находим

$$\min_i \left\{ \frac{x_i}{\alpha_i} : \alpha_i > 0, i = 1, \dots, m \right\} = v = \frac{x_r}{\alpha_r}$$

— максимально возможное значение $\tilde{x}_s$, при котором вектор $\tilde{\mathbf{x}}$ остается допустимым. Если же $\alpha_i \leq 0$ при всех $i = 1, \dots, m$, то компоненту x_s можно неограниченно увеличивать, не нарушая при этом условий допустимости. Из формулы (1.13) следует, что в этом случае целевая функция $z(\tilde{\mathbf{x}})$ не ограничена. В остальных случаях, полагая $x_s = v$, получим $\tilde{x}_r = 0$ и

$$z(\tilde{\mathbf{x}}) = z(\mathbf{x}) + \Delta' \tilde{\mathbf{x}} = z(\mathbf{x}) + \Delta'_N \tilde{\mathbf{x}} = z(\mathbf{x}) + v \Delta_s.$$

так как $\Delta_B = \mathbf{c}'_B - \boldsymbol{\pi}' \mathbf{B} = 0$.

1.3.2. АЛГОРИТМ СИМПЛЕКС-МЕТОДА

На основе проведенных рассуждений можно построить следующий итерационный алгоритм нахождения решения $\mathbf{x}^*$ задачи ЛП (1.10).

1. Пусть $\mathbf{B}$ — некоторый базис, допустимый для задачи ЛП. Вычисляется вспомогательный вектор $\boldsymbol{\pi}$ из условия

$$\boldsymbol{\pi}' = \mathbf{c}'_B \mathbf{B}^{-1}. \quad (1.15)$$

2. Вычисляется вектор *симплекс-разностей* Δ :

$$\Delta' = \mathbf{c}' - \boldsymbol{\pi}' \mathbf{A}.$$

Заметим, что условие (1.15) эквивалентно условию

$$\Delta_B \doteq \mathbf{c}'_B - \boldsymbol{\pi}'\mathbf{B} = \mathbf{0},$$

поэтому на практике при заданном базисе $\mathbf{B}$ вычисляют только вектор $\Delta_N = (\Delta_{m+1}, \dots, \Delta_n)'$.

3. Вычисляется минимальная симплекс-разность

$$\Delta_{\min} = \Delta_s = \min_i \Delta_i. \quad (1.16)$$

4. Проверяется достаточное условие оптимальности

$$\Delta_{\min} \geq 0. \quad (1.17)$$

Если это условие выполняется, то текущий базис $\mathbf{B}$ оптимален, и вычисления завершаются. В противном случае столбец $\mathbf{a}_s$ вводится в базис.

5. Вычисляется вектор

$$\boldsymbol{\alpha} = \mathbf{B}^{-1}\mathbf{a}_s$$

координат вектора $\mathbf{a}_s$ в базисе $\mathbf{B}$, и ищется величина

$$v = \frac{x_r}{\alpha_r} = \min_i \left\{ \frac{x_i}{\alpha_i} : \alpha_i > 0, i = 1, \dots, m \right\}. \quad (1.18)$$

Если вектор $\boldsymbol{\alpha}$ не содержит положительных компонент, то целевая функция $z(\mathbf{x})$ не ограничена на множестве допустимых векторов, и вычисления завершаются. В противном случае столбец $\mathbf{a}_r$ выводится из базиса.

6. Производится пересчет значений базисных переменных, целевой функции и матрицы $\mathbf{B}^{-1}$ по формулам

$$\tilde{x}_s = v, \quad \tilde{x}_i = x_i - v\alpha_i, \quad (1.19)$$

$$\tilde{z} = z + v\Delta_{\min}, \quad (1.20)$$

$$\tilde{g}_{ij} = \begin{cases} g_{ij} - g_{rj} \frac{\alpha_i}{\alpha_r}, & i \neq r, \\ g_{rj} \frac{1}{\alpha_r}, & i = r, \end{cases}$$

где g_{ij} — элементы матрицы $\mathbf{B}^{-1}$, а $\tilde{g}_{ij}$ — элементы матрицы $\tilde{\mathbf{B}}^{-1}$.

После этого происходит возврат к шагу 1. Новый допустимый базис $\tilde{\mathbf{B}}$ будет состоять из векторов $\mathbf{a}_i$, $i = 1, \dots, m$, $i \neq r$, и вектора $\mathbf{a}_s$.

Легко видно, что если на каждой итерации симплекс-метода величина v положительна, то приведенный алгоритм за конечное число итераций позволяет либо найти решение задачи, либо установить, что целевая функция не ограничена.

1.3.3. ОЦЕНКА БЛИЗОСТИ ТЕКУЩЕГО РЕШЕНИЯ К ОПТИМАЛЬНОМУ ДЛЯ НЕКОТОРЫХ ТИПОВ ЗАДАЧ

1. *Случай, когда $c_i = 1$ для всех i .* Рассмотрим задачу ЛП, в которой целевая функция имеет единичные коэффициенты, т. е. $z(\mathbf{x}) = x_1 + x_2 + \dots + x_n$. Тогда из формулы (1.13), полагая $\tilde{\mathbf{x}} = \mathbf{x}^*$, получим

$$z(\mathbf{x}^*) = z(\mathbf{x}) + \Delta' \tilde{\mathbf{x}}^* \geq z(\mathbf{x}) + \Delta_{\min} z(\mathbf{x}^*), \quad (1.21)$$

где $\mathbf{x}^*$ — решение задачи (1.10). Отсюда получаем оценку для искомой величины $z^* = z(\mathbf{x}^*)$

$$\frac{z(\mathbf{x})}{1 - \Delta_{\min}} \leq z(\mathbf{x}^*) \leq z(\mathbf{x}). \quad (1.22)$$

Пользуясь этими неравенствами, при решении задачи (1.10) в случае, когда $z(\mathbf{x}) = x_1 + \dots + x_n$, на каждом шаге симплекс-метода можно оценить близость текущего решения $z(\mathbf{x})$ к оптимальному. Это бывает полезно, когда число n переменных решаемой задачи весьма велико и сходимость симплекс-метода медленная. В этом случае неравенства (1.22) позволяют получить приближенное решение с заданной точностью.

2. *Случай наличия ограничения вида $\sum_{i=1}^n x_i = 1$.* При этом условии из (1.13) легко получить оценку

$$z(\mathbf{x}) + \Delta_{\min} \leq z(\mathbf{x}^*) \leq z(\mathbf{x}). \quad (1.23)$$

1.3.4. ГЕОМЕТРИЧЕСКИЙ СПОСОБ РЕШЕНИЯ ДВУМЕРНОЙ ЗАДАЧИ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ СПЕЦИАЛЬНОГО ВИДА

Рассмотрим двумерную задачу ЛП вида

$$\min_{\mathbf{x}} \left\{ \sum_{i=1}^n x_i : \sum_{i=1}^n \mathbf{a}_i x_i = \mathbf{b}, x_i \geq 0, i = 1, \dots, n \right\}, \quad (1.24)$$

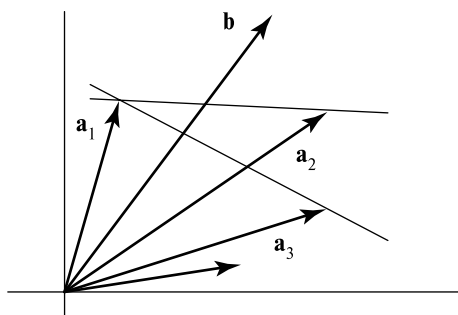


Рис. 1. Геометрический способ решения двумерной задачи (1.24)

в которой целевая функция имеет единичные коэффициенты; $\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{b} \in \mathbb{R}^2$.

Решение такой задачи может быть построено геометрически (рис. 1). Для этого на координатной плоскости из начала координат строятся векторы $\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{b}$. Затем выбирается базис, состоящий из каких-либо двух векторов $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}$ (на рисунке выбран начальный базис, состоящий из векторов $\mathbf{a}_1$ и $\mathbf{a}_3$), и через концы векторов выбранного базиса проводится прямая. Если среди $\mathbf{a}_1, \dots, \mathbf{a}_n$ нет векторов, пересекающих проведенную прямую, то выбранный базис оптимален. В самом деле, как нетрудно убедиться, вектор $\mathbf{b}$ представляется в виде линейной комбинации $\mathbf{b} = x_{i_1} \mathbf{a}_{i_1} + x_{i_2} \mathbf{a}_{i_2}$, причем $x_{i_1} \geq 0, x_{i_2} \geq 0$, а сумма $x_{i_1} + x_{i_2}$ минимальна среди всех возможных значений пар индексов (i_1, i_2) .

Рассмотрим случай, когда существует некоторый вектор $\mathbf{a}_{i_3}$, пересекающий построенную прямую (на рисунке — вектор $\mathbf{a}_2$). Если этот вектор ввести в базис так, чтобы коэффициенты разложения вектора $\mathbf{b}$ по новому базису были неотрицательны (на рисунке таким базисом является базис из векторов $\mathbf{a}_1$ и $\mathbf{a}_2$), то, как легко убедиться, сумма коэффициентов разложения вектора $\mathbf{b}$ по векторам этого базиса уменьшится. Далее снова проводится прямая через концы базисных векторов, и т.д. Такой упорядоченный перебор базисов и составляет идею симплекс-метода.

2. ОСНОВНЫЕ ЗАКОНЫ НЕБЕСНОЙ МЕХАНИКИ

2.1. ЗАКОНЫ КЕПЛЕРА

Основу небесной механики составляют три закона движения планет Солнечной системы, открытые в начале XVII века великим немецким математиком и астрономом Иоганном Кеплером*.

На протяжении почти полутора тысяч лет общепринятой теорией строения мира была геоцентрическая система, сформулированная около 140 г. н.э. греческим астрономом Клавдием Птолемеем (ок. 87 — ок. 165). Согласно этой системе, в центре вселенной находилась Земля, вокруг которой по круговым орбитам вращались Солнце, Луна и пять известных в то время планет: Меркурий, Венера, Марс, Юпитер и Сатурн. За этими планетами находилась сфера неподвижных звезд.

В 1543 г. польский астроном Николай Коперник опубликовал книгу «О вращении небесных сфер», в которой сформулировал свою *гелиоцентрическую* теорию строения солнечной

* Иоганн Кеплер (1571—1630) — немецкий астроном. Помимо открытых им законов движения планет вывел уравнение для определения положения небесных тел, названное его именем. Является автором так называемого *фотометрического парадокса*: если число звёзд бесконечно, то в любом направлении взгляд наткнулся бы на звезду, и на небе не существовало бы тёмных участков.

Внес также большой вклад в развитие математики, механики, физики. Изобрел способ определения объемов различных тел вращения, составил одну из первых таблиц логарифмов. Пытался объяснить гравитацию, изучал различные оптические явления, заложив основы оптики как науки. Создал один из видов телескопа. Кеплеру принадлежат такие термины как *среднее арифметическое*, *инерция*, *оптическая ось*.

Именем Кеплера названы кратеры на Луне и Марсе, астероид, орбитальная обсерватория NASA и др.

системы (такая гипотеза существовала и ранее, еще во времена античности, но ко времени Коперника она оказалась совершенно забытой). Согласно этой теории, Земля вместе с другими планетами совершает круговое вращение вокруг Солнца, а Луна совершает круговое вращение вокруг Земли.

Появление новой теории вызвало, разумеется, бурные споры. Появились ее горячие приверженцы, были и ярые противники. К числу последних относился и выдающийся датский астроном Тихо Браге (1546—1601). Этот ученый, в течение многих лет наблюдая за движением небесных тел, собрал уникальный по своему объему и научной ценности материал. В конце жизни он организовал астрономическую лабораторию в Праге, и там началась его совместная работа с молодым талантливым математиком-немцем Иоганном Кеплером. Кеплеру была поставлена задача описать движение Марса на основании результатов наблюдений, собранных Браге.

Однако в отличие от своего учителя Кеплер был приверженцем гелиоцентрической теории Коперника. После многочисленных попыток подобрать для Марса подходящую круговую орбиту Кеплер предположил, что Марс может двигаться вокруг Солнца по орбите эллиптической формы. Оказалось, что если Солнце поместить в один из фокусов такой орбиты, то полученная кривая будет очень точно соответствовать результатам наблюдений. Затем Кеплер предположил, что каждая из остальных планет, включая Землю, также движется по эллиптической орбите вокруг Солнца, находящегося в одном из фокусов эллипса. А движение Луны Кеплер описал эллипсом, в одном из фокусов которого находится Земля.

Проделав огромную вычислительную работу по проверке выдвинутых гипотез, Кеплер сформулировал следующий закон движения планет Солнечной системы.

Первый закон Кеплера: каждая планета совершает движение по орбите в форме эллипса, в одном из фокусов которого находится Солнце.

Затем Кеплер попытался оценить скорость движения планет по эллиптической орбите. Ученый обратил внимание на то, что в течение года скорость видимого движения Солнца по небесной сфере различна — зимой видимое перемещение Солнца происходит быстрее, чем весной и осенью, а летом движение Солнца замедляется. В результате длительных расчетов Кеплер вывел следующее заключение.

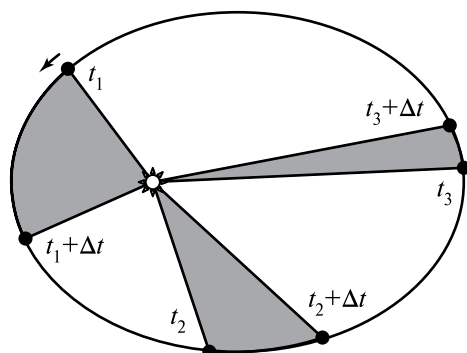


Рис. 2. Иллюстрация второго закона Кеплера. Темным цветом выделены сектора, заметаемые радиус-вектором планеты, обращающейся вокруг Солнца, за одинаковые промежутки времени Δt . Когда планета находится на разных участках орбиты, эти сектора различны, но их площади совпадают

Второй закон Кеплера (закон площадей): за равные промежутки времени радиус-вектор, соединяющий Солнце и планету, заметает сектора равной площади.

Второй закон Кеплера означает, что каждая планета в перигелии (ближайшая к Солнцу точка орбиты) движется с максимальной скоростью, в афелии (наиболее удаленная от Солнца точка орбиты) — с минимальной (рис. 2). Закон площадей указывает также, что сила, управляющая орбитальным движением планет, направлена к Солнцу.

Два своих первых закона Кеплер опубликовал в 1609 г. в книге «Новая астрономия». Третий закон стал результатом его дальнейших исследований и был опубликован спустя десять лет, в 1619 г., в книге «Гармония мира». Этот закон устанавливает связь между периодом обращения планеты вокруг Солнца и размерами ее орбиты.

Определение 2.1. Периодом обращения планеты при ее движении вокруг Солнца называется промежуток времени между двумя последовательными прохождением планеты одной и той же точки своей орбиты.

Третий (гармонический) закон Кеплера: квадраты периодов обращения двух планет вокруг Солнца относятся как кубы больших полуосей орбит этих планет:

$$\frac{T_1^2}{T_2^2} = \frac{a_1^3}{a_2^3}, \quad (2.1)$$

где T_1 и T_2 , a_1 и a_2 — соответственно периоды обращения и величины больших полуосей орбит любых двух планет Солнечной системы.

Пример 2.1. Известно, что среднее расстояние от планеты до Солнца равно большой полуоси орбиты этой планеты. Найти среднее расстояние от Юпитера до Солнца, если период обращения Юпитера составляет 11,86 лет.

Решение. Искомое расстояние можно найти, используя третий закон Кеплера, если известны период обращения и среднее расстояние до Солнца для какой-нибудь другой планеты Солнечной системы. Воспользуемся, например, известной информацией о том, что период обращения Земли вокруг Солнца равен одному году, а среднее расстояние от Земли до Солнца равно 1 а. е. = 149,6 млн км. Тогда по формуле (2.1) находим:

$$\frac{1^2}{11,86^2} = \frac{1^3}{x^3},$$

откуда $x^3 = 11,86^2$, $x = 5,2027$ а. е. = 778,324 млн км.

Замечание 2.1. Позднее Исаак Ньютон на основе закона всемирного тяготения уточнил третий закон Кеплера, сформулировав его в виде

$$\frac{T_1^2 \cdot (M + m_1)}{T_2^2 \cdot (M + m_2)} = \frac{a_1^3}{a_2^3}, \quad (2.2)$$

где M — масса Солнца, а m_1 и m_2 — массы планет. Поскольку движение и масса оказались связаны, эту комбинацию гармонического закона Кеплера и закона тяготения Ньютона используют для определения массы планет и спутников, если известны их расстояния и орбитальные периоды.

2.2. ДВИЖЕНИЕ В ГРАВИТАЦИОННОМ ПОЛЕ. ЗАКОН ВСЕМИРНОГО ТЯГОТЕНИЯ

В XVII веке ряд математиков, в том числе и Кеплер, высказывали соображения о том, что движение планет может быть объяснено

действием силы, которая притягивает каждую планету к Солнцу и убывает пропорционально квадрату расстояния до него. Однако строго доказать это удалось только в 1687 г. английскому математику и физiku Исааку Ньютону*. Ньютон доказал, что движение каждой планеты должно подчиняться первым двум законам Кеплера именно в том случае, если эти планеты движутся под действием силы тяготения Солнца. Далее Ньютон показал, что движение Луны может быть приближённо объяснено с помощью аналогичного действия силового поля Земли, и что сила тяжести на Земле есть результат воздействия этого же силового поля на материальные тела вблизи поверхности Земли. На основании третьего закона механики Ньютон заключил, что притяжение есть взаимное свойство, и пришёл к формулировке следующего закона.

Закон всемирного тяготения: *всякая материальная точка притягивает каждую другую материальную точку с силой, прямо пропорциональной произведению их масс и обратно пропорциональной квадрату расстояния между этими точками.*

Согласно этому закону, величина силы гравитационного взаимодействия двух материальных точек массами m_1 и m_2 , находящихся на расстоянии r друг от друга, определяется формулой

* Исаак Ньютон (1643–1727) — великий английский физик, математик, механик и астроном, один из создателей классической физики. Автор фундаментального труда «Математические начала натуральной философии», в котором изложен закон всемирного тяготения и три закона механики, ставшие основой классической механики.

Фактически открыл новую эпоху в физике и математике. Разработал дифференциальное и интегральное исчисления, теорию цвета, численные методы и многие другие математические и физические теории. Создал теорию физической оптики, открыл дисперсию света, доказал сложный состав белого цвета. Построил первый зеркальный телескоп.

Наряду с занятием научной деятельностью, Ньютон проявил незаурядные административные способности. Он долгое время состоял в должности управляющего английским Монетным двором, способствовал укреплению финансовой системы и успешному проведению в стране денежной реформы. Также занимался философией и алхимией. Вся жизнь был верующим, опубликовал ряд богословских работ.

В честь Ньютона названы единица силы в системе СИ, российский остров в Северном Ледовитом океане, а также множество научных законов, теорем и понятий.

$$F = G \frac{m_1 m_2}{r^2}, \quad (2.3)$$

где G — коэффициент пропорциональности, называемый *универсальной постоянной тяготения*:

$$G = 6,67 \cdot 10^{-20} \text{ км}^3 \cdot \text{кг}^{-1} \cdot \text{с}^{-2}.$$

Если рассмотреть радиус-вектор $\mathbf{r}$, характеризующий положение одной взаимодействующей точки относительно другой, то направление силы гравитационного взаимодействия будет противоположным направлению этого радиус-вектора, т.е. в векторной форме закон всемирного тяготения может быть записан в виде

$$\mathbf{F} = -G \frac{m_1 m_2}{r^2} \frac{\mathbf{r}}{r} = -G \frac{m_1 m_2}{r^3} \mathbf{r}, \quad (2.3)$$

где $\frac{\mathbf{r}}{r}$ — единичный вектор, направленный вдоль вектора $\mathbf{r}$.

Отметим наиболее важные особенности гравитационного взаимодействия [Балк, 1965]:

1. *Гравитационное воздействие имеет место на очень больших расстояниях.* Например, притяжение к Солнцу определяет движение малых тел Солнечной системы, находящихся на расстоянии от Солнца, в десятки тысяч раз большем, чем расстояние от Солнца до Земли.
2. *Гравитационное взаимодействие не зависит от среды, находящейся между взаимодействующими телами,* т.е. сила взаимного притяжения двух тел не зависит от того, расположены ли между этими телами какие-либо другие объекты.
3. *Гравитационное взаимодействие распространяется мгновенно.*

Гравитация была первым взаимодействием, описанным математической теорией. Из четырех видов взаимодействий (сильное, электромагнитное, слабое, гравитационное) гравитационное является слабейшим. Однако из-за перечисленных выше особенностей оно играет во Вселенной огромную роль. Гравитация, в отличие от других взаимодействий, универсальна в действии на всю материю и энергию. Не существует объектов, у которых вообще отсутствует гравитационный заряд. Из-за своего глобального характера гравитация является причиной самых разнообразных явлений природы от строения галактик, черных дыр и расширения Вселенной до смены времен года, чередования

дня и ночи, выпадения дождя или снега и даже обыденного падения какого-нибудь предмета на пол.

2.3. ОСНОВНАЯ ЗАДАЧА НЕБЕСНОЙ МЕХАНИКИ

Как уже говорилось выше, движение небесных тел в гравитационном поле является предметом изучения небесной механики. При этом, как правило, принимаются следующие упрощающие допущения [Балк, 1965]:

1. При изучении движения некоторого небесного тела (планеты, астероида или спутника) рассматривается его гравитационное взаимодействие лишь с небольшим числом других тел (как правило, одним или двумя). Гравитационным взаимодействием с бесконечным множеством прочих небесных тел, имеющих малую массу или расположенных на большом расстоянии от изучаемого тела, пренебрегают.
2. Все небесные тела считаются абсолютно твердыми.
3. Если расстояние между двумя объектами велико по сравнению с их размерами, то эти объекты рассматриваются как материальные точки.

Замечание 2.2. Можно показать, что гравитационное поле, создаваемое сферическим телом массы m со сферическим распределением плотности*, эквивалентно гравитационному полю, создаваемому материальной точкой той же массы m , помещенной в центре сферического тела. Поэтому в тех случаях, когда несферичностью формы и несферичностью распределения плотности небесного тела можно пренебречь, это тело также обычно рассматривают как материальную точку.

Основной задачей небесной механики является так называемая *задача n тел*: *изучить движение n материальных точек в их взаимном гравитационном поле, если известны массы этих точек, а также их положения и скорости в некоторый заданный момент времени.*

Самый простой пример задачи n тел — изучение движения Солнца и планет Солнечной системы.

* Шар имеет сферическое распределение плотности, если во всех точках, равноудалённых от центра шара, плотности равны.

Замечание 2.3. Задача n тел имеет простое аналитическое решение только при $n=2$. При $n=3$ аналитическое решение существует, но не может быть выражено в элементарных функциях, при $n>3$ аналитического решения не найдено. Поэтому для практического решения задачи трех и более тел применяются методы численного интегрирования.

Аналитическое решение задачи n тел для случаев $n=1$ и $n=2$ было найдено Ньютоном. При $n=1$ взаимодействия отсутствуют, и тело остается в состоянии покоя либо движется равномерно и прямолинейно — это утверждение составляет сформулированный Ньютоном первый закон классической механики. При $n=2$ движение взаимодействующих тел подчиняется закону всемирного тяготения. Ньютон показал, что в этом случае тела движутся в фиксированной плоскости, определяемой начальными условиями, а их орбиты друг относительно друга и относительно общего центра масс представляют собой кривые, называемые коническими сечениями (эллипсы, параболы или гиперболы).

С XVIII века многие выдающиеся ученые, среди которых Ж. Лагранж, К. Якоби, К. Вейерштрасс, А. Пуанкаре, Дж. Биркгоф и др., предпринимали попытки найти аналитическое решение при $n=1$. Было получено множество интересных результатов относительно тех или иных частных случаев, однако общего решения получить так и не удалось. Наконец, в конце XIX века Г. Э. Брунс и А. Пуанкаре показали, что общее решение задачи трех тел нельзя выразить через алгебраические или через однозначные трансцендентные функции координат и скоростей взаимодействующих тел. А в 1912 г. финский математик К. Зундман нашел общее решение задачи трех тел в виде сходящихся бесконечных степенных рядов. Однако сходимости этих рядов была чрезвычайно медленной, и уже через двадцать лет французский математик Д. Белорицкий показал, что для определения положения одного из трёх взаимодействующих тел при помощи рядов Зундмана с удовлетворительной точностью тре-

буется просуммировать не менее $10^{8 \cdot 10^6}$ членов этих рядов, что абсолютно невозможно даже для современной вычислительной техники. Так что найденное общее решение оказалось практически совершенно бесполезным [Маркеев, 1999].

Важнейшим частным случаем задачи n тел является изучение движения тела пренебрежимо малой массы в гравитационном поле нескольких массивных тел, которые обычно называют

центральными телами. Если несферичностью формы и несферичностью распределения плотности центральных тел можно пренебречь, то, согласно замечанию 2.2, эта задача может быть сформулирована как *задача изучения движения материальной точки с пренебрежимо малой массой в гравитационном поле $n-1$ материальных точек, масса и движение которых известны.* Такая задача называется **ограниченной задачей n тел.** Массивные материальные точки в задаче n тел далее будем называть *притягивающими центрами*, а тело с бесконечно малой массой — *спутником*. Примером ограниченной задачи n тел является изучение движения космического аппарата в гравитационном поле Солнца, Земли, Луны и, быть может, еще нескольких небесных тел (планет или, например, астероидов).

3. НЕВОЗМУЩЕННОЕ ДВИЖЕНИЕ

3.1. ЗАДАЧА ДВУХ ТЕЛ

Наиболее популярной и наиболее изученной задачей небесной механики является *задача двух тел*: *изучить движение материальной точки в гравитационном поле другой материальной точки, если известны массы этих точек, а также их положения и скорости в некоторый заданный момент времени*. К задаче двух тел относятся, очевидно, задачи расчета движения небесного тела в гравитационном поле другого небесного тела, если несферичностью формы и несферичностью распределения плотности этих двух тел, а также воздействием гравитационных полей всех прочих небесных тел можно пренебречь.

Определение 3.1. Движение материальной точки в гравитационном поле другой материальной точки в условиях задачи двух тел называется *невозмущенным* или *кеплеровым* движением, а траектория этой материальной точки — *орбитой кеплерова движения* или, короче, *кеплеровой орбитой*.

Важным частным случаем задачи двух тел является *задача изучения движения спутника в гравитационном поле некоторого притягивающего центра с известной массой*. Эта задача называется *ограниченной задачей двух тел*.

Рассмотрим задачу двух тел подробнее. Пусть C и S — две материальные точки массами M и m соответственно, $\mathbf{r}_o(C)$ и $\mathbf{r}_o(S)$ — радиус-векторы этих точек в некоторой инерциальной системе координат $Ox_o y_o z_o$ (рис. 3). Положение точки S относительно точки C будем определять вектором $\mathbf{r} = \mathbf{r}_o(S) - \mathbf{r}_o(C)$.

Согласно закону всемирного тяготения (2.3), на точку S со стороны точки C будет действовать сила гравитационного притяжения

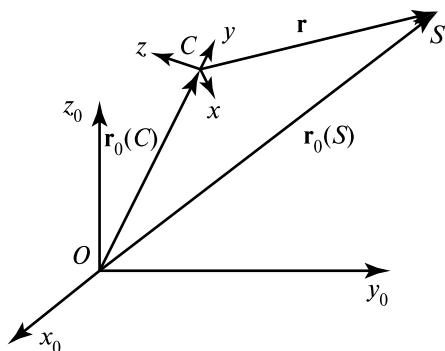


Рис. 3. Иллюстрация к задаче двух тел

$$\mathbf{F}_S = -G \frac{Mm}{r^2} \frac{\mathbf{r}}{r} = -\mu_s \frac{M}{r^3} \mathbf{r},$$

где $r = |\mathbf{r}|$; $\frac{\mathbf{r}}{r}$ — вектор единичной длины, определяющий направление гравитационной силы; $\mu_s = Gm$ — постоянная, называемая *гравитационным параметром* точки S . Соответственно на точку C со стороны точки S будет действовать такая же по величине и обратная по направлению сила

$$\mathbf{F}_C = -\mathbf{F}_S = G \frac{Mm}{r^2} \frac{\mathbf{r}}{r} = \mu_c \frac{m}{r^3} \mathbf{r},$$

где $\mu_c = GM$ — гравитационный параметр точки C . Тогда уравнения движения точек S и C в системе координат $Ox_0y_0z_0$ запишутся в виде

$$\ddot{\mathbf{r}}_o(C) = G \frac{m}{r^2} \frac{\mathbf{r}}{r}, \quad \ddot{\mathbf{r}}_o(S) = -G \frac{M}{r^2} \frac{\mathbf{r}}{r},$$

откуда

$$\ddot{\mathbf{r}} = \ddot{\mathbf{r}}_o(S) - \ddot{\mathbf{r}}_o(C) = -G \frac{M+m}{r^2} \frac{\mathbf{r}}{r}.$$

Рассмотрим некоторую декартову систему координат $Sx_0y_0z_0$, начало которой совпадает с притягивающим центром C . Тогда полученное равенство является выражением для ускорения точки S в системе координат $Sx_0y_0z_0$ и приводит к важному выводу: *движение материальной точки массой m относительно другой материальной точки массой M в задаче двух тел эквивалентно движению спутника в ограниченной задаче двух тел относительно*

притягивающего центра массой $M+m$. Величина гравитационного параметра спутника в ограниченной задаче двух тел будет равна $\mu = G(M+m)$. Так как масса спутника пренебрежимо мала по сравнению с массой притягивающего центра, то можно считать $\mu \approx GM$ и

$$\ddot{\mathbf{r}} = -\frac{\mu}{r^3} \mathbf{r}. \quad (3.1)$$

Это равенство называется *уравнением движения спутника в ограниченной задаче двух тел*.

Обозначим через $\mathbf{v} = \dot{\mathbf{r}}$ вектор скорости спутника и рассмотрим вектор $\mathbf{x} = \begin{pmatrix} \mathbf{r} \\ \mathbf{v} \end{pmatrix}$, называемый *вектором состояния*. Рассмотрим также вектор

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \mathbf{v} \\ -\frac{\mu}{r^3} \mathbf{r} \end{pmatrix}.$$

Тогда уравнение (3.1) в векторной форме будет записываться в виде

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}).$$

3.2. ПЕРВЫЕ ИНТЕГРАЛЫ УРАВНЕНИЯ ДВИЖЕНИЯ СПУТНИКА

3.2.1. ИНТЕГРАЛ ЭНЕРГИИ

Умножим обе части уравнения (3.1) скалярно на величину $2\dot{\mathbf{r}}$:

$$2\dot{\mathbf{r}}\ddot{\mathbf{r}} = -\frac{2\mu}{r^3} \dot{\mathbf{r}}\mathbf{r}.$$

Это равенство можно записать в виде

$$\frac{d(\dot{\mathbf{r}}^2)}{dt} = -\frac{\mu}{r^3} \frac{d(r^2)}{dt},$$

или, с учетом равенств $\dot{\mathbf{r}}^2 = \mathbf{v}^2 = v^2$ и $\mathbf{r}^2 = r^2$,

$$\frac{d(v^2)}{dt} = -\frac{\mu}{r^3} \frac{d(r^2)}{dt} = \frac{d}{dt} \left(\frac{2\mu}{r} \right).$$

Интегрируя левую и правую части последнего равенства по времени, получим выражение

$$v^2 - \frac{2\mu}{r} = h, \quad (3.2)$$

где h — некоторая постоянная величина. Данное выражение называется *интегралом энергии*. Первое слагаемое в левой части представляет собой удвоенную кинетическую энергию единицы массы спутника, а второе слагаемое — удвоенную потенциальную энергию единицы массы спутника. Следовательно, постоянная h равна удвоенной величине полной энергии единицы массы спутника. Эта величина называется *постоянной энергии*.

Из формулы (3.2) следует, что *полная энергия в задаче двух тел является постоянной величиной*. В частности, это означает, что при удалении спутника от притягивающего центра его скорость уменьшается, а при приближении к притягивающему центру — увеличивается.

3.2.2. ИНТЕГРАЛ ПЛОЩАДЕЙ

Предположим вначале, что векторы $\mathbf{r}$ и $\mathbf{v}$ неколлинеарны, т. е. $\mathbf{r} \times \mathbf{v} \neq \mathbf{0}$. Умножим обе части уравнения движения спутника (3.1) векторно на величину $\mathbf{r}$:

$$\mathbf{r} \times \ddot{\mathbf{r}} = -\frac{\mu}{r^3} \mathbf{r} \times \mathbf{r} = \mathbf{0},$$

так как по определению векторного произведения $\mathbf{r} \times \mathbf{r} = \mathbf{0}$.

Продифференцируем теперь по времени выражение $\mathbf{r} \times \mathbf{v}$:

$$\frac{d}{dt}(\mathbf{r} \times \mathbf{v}) = \frac{d}{dt}(\mathbf{r} \times \dot{\mathbf{r}}) = \dot{\mathbf{r}} \times \dot{\mathbf{r}} + \dot{\mathbf{r}} \times \ddot{\mathbf{r}} = \mathbf{r} \times \ddot{\mathbf{r}} = \mathbf{0},$$

откуда

$$\mathbf{r} \times \mathbf{v} = \mathbf{c}, \quad (3.3)$$

где $\mathbf{c}$ — некоторый постоянный вектор. Полученное равенство называется *векторным интегралом площадей*, а вектор $\mathbf{c}$ — *векторной постоянной площадей*. Как видно из последнего равенства, вектор $\mathbf{r}$ во время движения всегда остается ортогональным вектору $\mathbf{c}$, т. е. вектор $\mathbf{r}$ всегда находится в плоскости, проходящей через притягивающий центр и определяемой нормальным к ней постоянным вектором $\mathbf{c}$. Таким образом, интеграл площадей показывает, что *движение спутника в ограниченной*

задаче двух тел происходит в одной плоскости, проходящей через притягивающий центр. Эта плоскость, т.е. плоскость движения спутника, называется *неизменяемой плоскостью Лапласа*. Равенство (3.3) является общим уравнением плоскости Лапласа в векторной форме.

Введем в пространстве декартову систему координат $Cxuz$ с началом в притягивающем центре C , причем оси Cx и Cy расположим в плоскости движения спутника, ось Cz направим вдоль вектора $\mathbf{c}$ (рис. 4).

Пусть радиус-вектор $\mathbf{r}$ и вектор скорости $\mathbf{v}$ спутника в этой системе имеют координаты

$$\mathbf{r} = \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}.$$

Тогда интеграл площадей (3.3) с учетом формулы (1.7) в координатной форме записывается в виде

$$\begin{pmatrix} r_y v_z - r_z v_y \\ r_z v_x - r_x v_z \\ r_x v_y - r_y v_x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}. \quad (3.4)$$

Введем в плоскости движения полярные координаты (r, ϑ) , полагая

$$\begin{aligned} x &= r \cos \vartheta, \\ y &= r \sin \vartheta. \end{aligned}$$

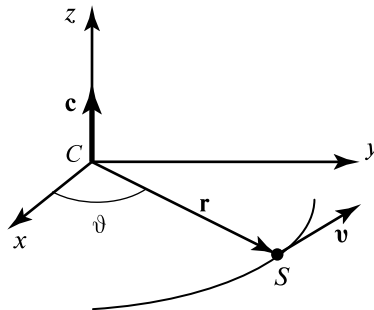


Рис. 4. Иллюстрация к выводу полярной формы интеграла площадей

Тогда

$$\begin{aligned}v_x &= \dot{x} = \dot{r} \cos \vartheta - r \sin \vartheta \cdot \dot{\vartheta}, \\v_y &= \dot{y} = \dot{r} \sin \vartheta + r \cos \vartheta \cdot \dot{\vartheta}.\end{aligned}$$

Подставляя эти равенства в выражение для третьей координаты в равенстве (3.4), получим после преобразований соотношение

$$r^2 \dot{\vartheta} = c, \quad (3.5)$$

которое называется *полярной формой интеграла площадей*. Эта формула имеет несколько важных следствий:

1. Если направление оси Cz совпадает с направлением вектора $\mathbf{c}$, т. е. $c > 0$, то в любой момент времени t справедливо условие $\dot{\vartheta} > 0$. Это означает, что угол наклона радиус-вектора спутника по отношению к оси Cx постоянно возрастает и *движение спутника происходит в положительном направлении* (т. е. в направлении против часовой стрелки, если смотреть на плоскость Cxy со стороны положительного направления оси Cz). Такое движение спутника называется *прямым*. Если же $c < 0$, то *движение спутника все время происходит в отрицательном направлении*, и такое движение называется *обратным*.
2. Из полярной формы интеграла площадей (3.5) следует выражение для угловой скорости спутника

$$\dot{\vartheta} = \frac{c}{r^2}, \quad (3.6)$$

которое показывает, что, чем дальше спутник от притягивающего центра, тем меньше его угловая скорость.

3. Пусть радиус-вектор спутника за время Δt успел описать некоторый угол $\Delta \vartheta$ и «замести» некоторую площадь ΔS . Площадь замеченного сектора ΔS приближенно равна

$$\Delta S = \frac{1}{2} r^2 \Delta \vartheta,$$

откуда

$$\frac{\Delta S}{\Delta t} = \frac{1}{2} r^2 \frac{\Delta \vartheta}{\Delta t}.$$

При $\Delta t \rightarrow 0$ отсюда следует

$$\frac{dS}{dt} = \frac{1}{2} r^2 \frac{d\vartheta}{dt}.$$

Величина $\frac{dS}{dt}$ представляет собой секториальную скорость спутника относительно притягивающего центра. Из формулы (3.5) следует, что

$$\frac{dS}{dt} = \frac{1}{2} c = \text{const},$$

т.е. *секториальная скорость невозмущенного движения спутника относительно притягивающего центра постоянна.*

Полученный результат выражает второй закон Кеплера: *за равные промежутки времени радиус-вектор спутника заметает сектора равной площади.*

Замечание 3.1. При выводе формулы интеграла площадей предполагалось, что векторы $\mathbf{r}$ и $\mathbf{v}$ неколлинеарны. Легко показать, что в случае коллинеарности этих векторов спутник будет совершать прямолинейное движение в направлении своего радиус-вектора, и понятие плоскости движения теряет смысл.

3.2.3. ИНТЕГРАЛ ЛАПЛАСА

Перемножим векторно левые и правые части уравнений (3.1) и (3.3):

$$\ddot{\mathbf{r}} \times \mathbf{c} = -\frac{\mu}{r^3} \mathbf{r} \times (\mathbf{r} \times \mathbf{v}).$$

С учетом равенств $\ddot{\mathbf{r}} = \dot{\mathbf{v}}$ и $\mathbf{c} = \text{const}$ левая часть последнего соотношения может быть записана в виде

$$\ddot{\mathbf{r}} \times \mathbf{c} = \frac{d}{dt}(\mathbf{v} \times \mathbf{c}),$$

а правая — в виде

$$\begin{aligned} -\frac{\mu}{r^3} \mathbf{r} \times (\mathbf{r} \times \mathbf{v}) &= -\frac{\mu}{r^3} [\mathbf{r}(\mathbf{r} \cdot \dot{\mathbf{r}}) - \dot{\mathbf{r}}(\mathbf{r}^2)] = \\ &= -\frac{\mu}{r^3} (\mathbf{r} r \dot{r} - \dot{\mathbf{r}} r^2) = -\mu \left(\frac{\mathbf{r} \dot{r}}{r^2} - \frac{\dot{\mathbf{r}}}{r} \right) = \frac{d}{dt} \left(\mu \frac{\mathbf{r}}{r} \right) \end{aligned}$$

(здесь использованы равенства $\mathbf{r}^2 = r^2$ и $\mathbf{r} \cdot \dot{\mathbf{r}} = r\dot{r}$; последнее равенство представляет собой результат дифференцирования обеих частей первого равенства по времени).

Приравнивая преобразованные левую и правую части, находим:

$$\frac{d}{dt}(\mathbf{v} \times \mathbf{c}) = \frac{d}{dt} \left(\mu \frac{\mathbf{r}}{r} \right)$$

или

$$\frac{d}{dt} \left(\mathbf{v} \times \mathbf{c} - \mu \frac{\mathbf{r}}{r} \right) = \mathbf{0},$$

откуда

$$\mathbf{v} \times \mathbf{c} - \mu \frac{\mathbf{r}}{r} = \mathbf{l}, \quad (3.7)$$

где $\mathbf{l}$ — некоторый постоянный вектор, называемый *вектором Лапласа* или *постоянной Лапласа**. Полученное векторное равенство (3.7) называется *интегралом Лапласа*. Так как оба слагаемых в левой части интеграла Лапласа лежат в плоскости движения спутника, то, следовательно, вектор Лапласа также всегда лежит в этой плоскости.

* Пьер Симон Лаплас (1749–1827) — французский астроном, математик, физик. Автор классических трудов по теории вероятностей и небесной механике, математической физике, акустике, геодезии и др. Был членом Французского Географического общества.

Лаплас доказал устойчивость Солнечной системы, развил теорию возмущений, внес большой вклад в развитие теории движения Луны, уточнил сжатие земного сфероида, впервые построил точную теорию движения галилеевых спутников Юпитера, разработал теорию приливов. Кроме того, Лаплас первым предложил математически обоснованную теорию образования всех тел Солнечной системы (космогоническая гипотеза Лапласа).

Среди математических заслуг Лапласа — разработка методов математической физики, создание математического фундамента теории вероятностей, доказательство предельных теорем и их приложения к обработке наблюдений, формулировка разложения определителя по минорам.

В физике Лапласу принадлежит барометрическая формула, связывающая плотность воздуха, высоту, влажность и ускорение свободного падения. Также он занимался акустикой, вывел формулу для скорости распространения звука в воздушной среде, внёс значительный вклад в развитие гидродинамики. Лаплас состоял в дружеских отношениях с Наполеоном, который наградил его титулом графа и множеством орденов, даже назначил на должность министра внутренних дел, правда на этой должности Лаплас пробыл всего полтора месяца.

В честь учёного названы кратер на Луне, астероид, а также многочисленные математические понятия и теоремы.

Теорема 3.1. *Постоянные $\mathbf{l}$, $\mathbf{c}$, h первых интегралов уравнения движения спутника связаны соотношением*

$$l^2 = \mu^2 + hc^2. \quad (3.8)$$

Доказательство. Рассмотрим скалярное произведение

$$\mathbf{l} \cdot \mathbf{l} = l^2 = \left(\mathbf{v} \times \mathbf{c} - \mu \frac{\mathbf{r}}{r} \right)^2 = (\mathbf{v} \times \mathbf{c})^2 - 2\mu(\mathbf{v} \times \mathbf{c}) \cdot \frac{\mathbf{r}}{r} + \mu^2 \frac{r^2}{r^2}.$$

Так как вектор Лапласа $\mathbf{l}$ лежит в плоскости движения, а постоянная площадей $\mathbf{c}$ ортогональна этой плоскости, т.е. $\mathbf{v} \perp \mathbf{c}$, то

$$(\mathbf{v} \times \mathbf{c})^2 = |\mathbf{v} \times \mathbf{c}|^2 = v^2 c^2.$$

Кроме того, из соотношения (1.9) следует

$$(\mathbf{v} \times \mathbf{c}) \cdot \mathbf{r} = \mathbf{c} \cdot (\mathbf{r} \times \mathbf{v}) = \mathbf{c}^2 = c^2 \quad \text{и} \quad \frac{r^2}{r^2} = 1.$$

Тогда

$$l^2 = v^2 c^2 - \frac{2\mu c^2}{r} + \mu^2 = \mu^2 + c^2 \left(v^2 - \frac{2\mu}{r} \right).$$

В последнем равенстве выражение в скобках представляет собой, по определению, постоянную энергии h , что и доказывает теорему.

4. НЕВОЗМУЩЕННОЕ ДВИЖЕНИЕ СПУТНИКА В ПЛОСКОСТИ ОРБИТЫ

4.1. УРАВНЕНИЕ ОРБИТЫ СПУТНИКА

При помощи интеграла площадей и интеграла Лапласа нетрудно получить уравнение орбиты спутника при его движении относительно притягивающего центра.

Рассмотрим вначале случай $\mathbf{c} = \mathbf{0}$. Тогда из интеграла Лапласа находим:

$$\mathbf{r} = \frac{r}{\mu} \mathbf{l},$$

т. е. спутник движется по прямолинейной орбите.

Пусть теперь $\mathbf{c} \neq \mathbf{0}$. Найдём скалярное произведение векторов $\mathbf{l}$ и $\mathbf{r}$:

$$\mathbf{l} \cdot \mathbf{r} = (\mathbf{v} \times \mathbf{c}) \cdot \mathbf{r} - \frac{\mu \mathbf{r}}{r} \cdot \mathbf{r} = \mathbf{c} \cdot (\mathbf{r} \times \mathbf{v}) - \mu r = c^2 - \mu r.$$

Но, согласно (1.5),

$$\mathbf{l} \cdot \mathbf{r} = l r \cos \vartheta,$$

где ϑ — угол между векторами $\mathbf{l}$ и $\mathbf{r}$. Тогда

$$l r \cos \vartheta = c^2 - \mu r$$

или

$$r(\mu + l \cos \vartheta) = c^2,$$

откуда

$$r = \frac{c^2}{\mu + l \cos \vartheta}.$$

Введем обозначения

$$p = \frac{c^2}{\mu} \quad (4.1)$$

и

$$e = \frac{l}{\mu}. \quad (4.2)$$

Тогда окончательно запишем

$$r = \frac{p}{1 + e \cos \vartheta}. \quad (4.3)$$

Полученное уравнение представляет собой зависимость величины радиус-вектора спутника от угла ϑ при заданных значениях параметров p и e . Так как ϑ и r являются полярными координатами спутника в плоскости его движения, то уравнение (4.3) является *уравнением орбиты спутника в полярных координатах*. Параметр p называется *фокальным параметром* орбиты, параметр e — ее *эксцентриситетом*.

Определение 4.1. Ближайшую к притягивающему центру точку орбиты спутника называют *перигентром*, наиболее удаленную от притягивающего центра точку орбиты (если она существует, т. е. если орбита замкнутая) — *апоцентром*. Полярный угол ϑ между направлениями от притягивающего центра к спутнику и от притягивающего центра к перигентру называется *истинной аномалией*. Главная или фокальная ось орбиты, проходящая от притягивающего центра к перигентру, называется в астрономии *линией апсид*. Точки пересечения этой линии с орбитой называются *апсидальными*, или просто *апсидами*.

Замечание 4.1. При рассмотрении движения вокруг некоторых конкретных небесных тел апоцентр и перигентр орбиты имеют специальные названия. Так, при движении вокруг Солнца апоцентр и перигентр орбиты принято называть соответственно *афелием* и *перигелием*, при движении вокруг Земли — *апогеем* и *перигеем*, вокруг Луны — *апоселением* и *периселением*, вокруг какой-либо звезды — *апоастром* и *периастром*.

Итак, *истинная аномалия характеризует угловое смещение спутника относительно перигентра, эксцентриситет e характеризует форму орбиты, фокальный параметр p — ее линейные размеры. Линия апсид является осью симметрии орбиты. Можно показать, что направление линии апсид совпадает с направлением вектора Лапласа.*

Таким образом, значения *фокального параметра и эксцентриситета полностью определяют форму и размер орбиты невозмущенного движения (кеплеровой орбиты) спутника.*

Из уравнения (4.3) непосредственно следует, что:

- при $\vartheta = 0$ спутник находится на минимальном расстоянии от притягивающего центра

$$r_{\min} = \frac{p}{1+e}, \quad (4.4)$$

которое называют *расстоянием от притягивающего центра до перицентра орбиты*, *фокальным расстоянием до перицентра* или, короче, *расстоянием до перицентра*;

- при $\vartheta = \pm(\pi/2)$ расстояние от притягивающего центра до спутника равно $r=p$, т. е. *фокальный параметр орбиты равен расстоянию от притягивающего центра до точки орбиты, лежащей на прямой, проходящей через притягивающий центр и перпендикулярной линии апсид*;
- если $e < 1$, то при $\vartheta = \pi$ спутник находится на максимальном расстоянии от притягивающего центра

$$r_{\max} = r_a = \frac{p}{1-e}, \quad (4.5)$$

которое называют *расстоянием от притягивающего центра до апоцентра орбиты*, *фокальным расстоянием до апоцентра* или, короче, *расстоянием до апоцентра*;

- если $e \geq 1$, то при возрастании значений угла ϑ от нуля спутник удаляется от притягивающего центра на бесконечно большое расстояние.

Из курса аналитической геометрии (см., например, [Ильин, Позняк, 1968]) известно, что полученное уравнение орбиты спутника представляет собой уравнение конического сечения в полярных координатах с полюсом в фокусе.

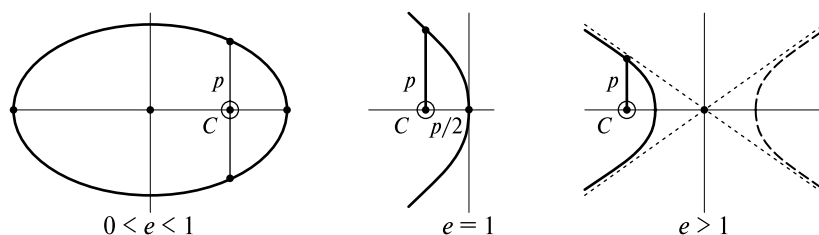


Рис. 5. Вид орбиты спутника при различных значениях e . Для каждого случая показаны положение притягивающего центра C и величина фокального параметра p

В зависимости от величины параметра e плоская кривая, описываемая этим уравнением, будет иметь различный вид. При $e=0$ она будет являться окружностью радиуса p , при $0 < e < 1$ — эллипсом, при $e=1$ — параболой, при $e > 1$ — гиперболой (рис. 5).

Таким образом, уравнение орбиты спутника (4.3) выражает первый закон Кеплера и означает, что *невозмущенное движение спутника относительно притягивающего центра всегда совершается по коническому сечению (окружности, эллипсу, параболе или гиперболе), в одном из фокусов которого находится притягивающий центр.*

4.2. СКОРОСТЬ СПУТНИКА НА ОРБИТЕ

Рассмотрим спутник S , движущийся вокруг притягивающего центра C по орбите с известными параметрами p , e , ϑ и найдем вектор его скорости $\mathbf{v}$ в произвольный момент времени (рис. 6).

Пусть $\mathbf{r} = \overline{CS}$ — радиус-вектор спутника. Вектор $\mathbf{v}$ можно представить в виде суммы двух компонент: вектора $\mathbf{v}_r$, направление которого совпадает с направлением радиус-вектора $\mathbf{r}$, и вектора $\mathbf{v}_n$, перпендикулярного вектору $\mathbf{r}$. Эти компоненты вектора скорости называют *радиальной* и *трансверсальной* скоростями соответственно. При этом величины радиальной и трансверсальной скоростей равны

$$v_r = \dot{r}, \quad v_n = r \dot{\vartheta}. \quad (4.6)$$

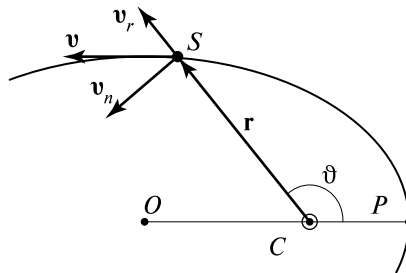


Рис. 6. Компоненты вектора скорости спутника S при движении вокруг притягивающего центра C

С учетом (4.3), (3.5), (4.1) и (4.2) находим:

$$\begin{aligned} v_r &= \frac{dr}{dt} = \frac{dr}{d\vartheta} \frac{d\vartheta}{dt} = \frac{pe \sin \vartheta}{(1 + e \cos \vartheta)^2} \frac{c}{r^2} = \frac{p^2}{(1 + e \cos \vartheta)^2} \frac{1}{r^2} \frac{e}{p} c \sin \vartheta = \\ &= \sqrt{\frac{\mu}{p}} e \sin \vartheta, \\ v_n &= \frac{c}{r} = \sqrt{\frac{\mu}{p}} (1 + e \cos \vartheta). \end{aligned}$$

Тогда величина скорости спутника равна

$$v = \sqrt{v_r^2 + v_n^2} = \sqrt{\frac{\mu}{p}} \sqrt{1 + 2e \cos \vartheta + e^2}. \quad (4.7)$$

Из этой формулы следует, что максимальная по величине скорость достигается в перигеуме орбиты (при $\vartheta = 0$):

$$v_{\max} = v_p = \sqrt{\frac{\mu}{p}} (1 + e), \quad (4.8)$$

минимальная — в апоцентре (т.е. при $\vartheta = \pi$), если апоцентр существует:

$$v_{\min} = v_a = \sqrt{\frac{\mu}{p}} (1 - e). \quad (4.9)$$

Из формул (4.2) и (3.8) следует, что эксцентриситет орбиты спутника может быть вычислен по формуле

$$e = \sqrt{1 + h \frac{c^2}{\mu^2}}, \quad (4.10)$$

где $h = v^2 - \frac{2\mu}{r}$ — постоянная энергии. Найдем выражения для скорости спутника при различных значениях эксцентриситета.

1. $e=0$. В этом случае, как уже отмечалось выше, орбита спутника будет иметь форму окружности, т.е. $r=p=\text{const}$. Тогда радиальная и трансверсальная скорости будут равны

$$v_r = 0, \quad v_n = \sqrt{\frac{\mu}{r}}.$$

Отсюда следует, что *скорость движения спутника по круговой орбите радиуса r равна*

$$v = v_{\text{кр}}(r) = \sqrt{\frac{\mu}{r}}, \quad (4.11)$$

причём вектор скорости в любой момент времени ортогонален радиус-вектору спутника.

2. $0 < e < 1$. В этом случае орбита спутника будет эллиптической. При этом $h < 0$ и, следовательно,

$$v_{\text{кр}}(r) < v < \sqrt{\frac{2\mu}{r}}.$$

3. $e = 1$. В этом случае орбита представляет собой параболу, $h = 0$ и

$$v = v_{\text{пар}}(r) = \sqrt{\frac{2\mu}{r}}. \quad (4.12)$$

4. $e > 1$. В этом случае спутник будет двигаться по гиперболической орбите, для которой $h > 0$ и

$$v > \sqrt{\frac{2\mu}{r}}.$$

Определение 4.2. *Первой космической скоростью* для планеты радиуса R называется скорость кругового движения на поверхности этой планеты:

$$v_1 = v_{\text{кр}}(R) = \sqrt{\frac{\mu}{R}}. \quad (4.13)$$

Иначе говоря, первая космическая скорость — это скорость, которую необходимо сообщить спутнику на поверхности планеты для того, чтобы он вращался вокруг планеты по круговой орбите на нулевой высоте над поверхностью.

Пример 4.1. Найти величину первой космической скорости на поверхности Земли.

Решение. Принимая радиус Земли равным $R_E = 6371$ км, а гравитационный параметр Земли равным $\mu_E = 3,986 \cdot 10^5$ км³/с², найдём:

$$v_{1E} = v_{\text{кр}}(R_E) = \sqrt{\frac{3,986 \cdot 10^5}{6371}} \approx 7,91 \text{ км} \cdot \text{с}^{-1}.$$

Определение 4.3. *Второй космической скоростью* для планеты радиуса R называется скорость параболического движения на поверхности этой планеты:

$$v_2 = v_{\text{пар}}(R) = \sqrt{\frac{2\mu}{R}}. \quad (4.14)$$

Другими словами, вторая космическая скорость — это скорость, которую необходимо сообщить спутнику, находящемуся на поверхности планеты, для того, чтобы он мог преодолеть гравитационное притяжение этой планеты и уйти от нее на бесконечно большое расстояние.

Пример 4.2. Найти величину второй космической скорости на поверхности Земли.

Решение. Используя результаты решения примера 4.1, найдём:

$$v_{2E} = \sqrt{2} v_{1E} \approx 11,19 \text{ км} \cdot \text{с}^{-1}.$$

Замечание 4.2. Иногда рассматривают третью и четвертую космические скорости. *Третья космическая скорость* — это скорость, которую необходимо сообщить спутнику у поверхности планеты для того, чтобы он смог преодолеть гравитационное притяжение не только данной планеты, но и звезды, вокруг которой вращается планета, и покинуть пределы своей звездной системы. *Четвертую космическую скорость* необходимо сообщить спутнику для того, чтобы он смог преодолеть гравитационное взаимодействие со своей галактикой и уйти за ее пределы.

4.3. ПАРАМЕТРЫ ОРБИТЫ СПУТНИКА

4.3.1. ЭЛЛИПТИЧЕСКИЕ ОРБИТЫ

Рассмотрим невозмущенное движение спутника вокруг притягивающего центра по эллиптической орбите. Как уже было установлено выше, при движении по эллиптическим орбитам выполняются неравенства

$$0 < e < 1, \quad h < 0, \quad \sqrt{\frac{\mu}{r}} < v < \sqrt{\frac{2\mu}{r}}.$$

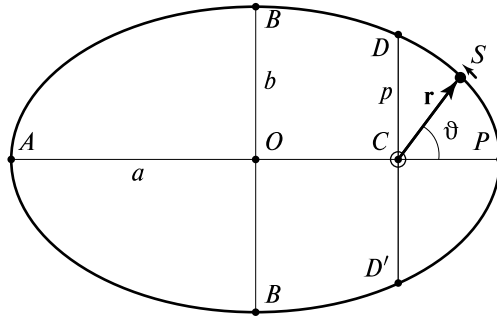


Рис. 7. Параметры эллиптической орбиты

Пусть S — текущее положение спутника, движущегося вокруг притягивающего центра C (рис. 7). Введем обозначения: O — центр орбитального эллипса, P — перигецентр орбиты, A — апоцентр, AP — линия апсид. Истинная аномалия ϑ представляет собой угол между текущим радиус-вектором спутника $\mathbf{r} = \overrightarrow{CS}$ и направлением $\overrightarrow{CP}$ из притягивающего центра на перигецентр. Проведем через точку C хорду DD' , перпендикулярную линии апсид (такая хорда называется *фокальной*). Тогда фокальный параметр p будет равен длине отрезка CD :

$$p = |CD|.$$

Эксцентриситет орбиты равен отношению расстояния OC от центра орбитального эллипса до фокуса к расстоянию OP от центра эллипса до перигецентра:

$$e = \frac{OC}{OP}.$$

Определение 4.4. *Большой полуось a эллиптической орбиты спутника называется длина отрезка, соединяющего центр орбитального эллипса с перигецентром (или апоцентром):*

$$a = |OP| = |OA|.$$

Малой полуось b эллиптической орбиты будем называть длину отрезка, соединяющего центр орбитального эллипса с одной из двух ближайших к нему точек орбиты (см. рис. 7):

$$b = |OB| = |OB'|.$$

Из свойств эллипса следует, что

$$a = \frac{p}{1-e^2}, \quad (4.15)$$

$$b^2 = a^2(1-e^2), \quad (4.16)$$

$$p = a(1-e^2) = \frac{b^2}{a}. \quad (4.17)$$

Расстояния от притягивающего центра до перицентра и апоцентра орбиты равны

$$r_p = r_{\min} = \frac{p}{1+e}, \quad r_a = r_{\max} = \frac{p}{1-e}. \quad (4.18)$$

Приведенные соотношения показывают, что форма и размер эллиптической орбиты кеплеровского движения полностью определяются любыми двумя из следующих линейных параметров:

- фокальный параметр p ;
- эксцентриситет e ;
- большая полуось a ;
- малая полуось b ;
- расстояние до перицентра r_p ;
- расстояние до апоцентра r_a .

Пример 4.3. Найти наибольшее и наименьшее расстояния от Земли до Солнца, если большая полуось орбиты Земли равна $149,6 \cdot 10^6$ км, а эксцентриситет орбиты составляет $1/60$.

Решение. Используя приведенные выше соотношения, найдем:

$$r_p = \frac{p}{1+e} = \frac{a(1-e^2)}{1+e} = a(1-e) \approx 147,1 \cdot 10^6 \text{ км}$$

— расстояние до перигелия Земли. Аналогично находим, что расстояние до афелия Земли равно

$$r_a = a(1+e) \approx 152,1 \cdot 10^6 \text{ км}.$$

Теорема 4.1 (правило рычага). При невозмущенном движении спутника по эллиптической орбите отношение величин его скорости в перицентре и апоцентре равно отношению расстояний до апоцентра и перицентра орбиты, т. е.

$$\frac{v_p}{v_a} = \frac{r_a}{r_p}. \quad (4.19)$$

Доказательство. Отношение расстояния до апоцентра к расстоянию до перигея равно

$$\frac{r_a}{r_p} = \frac{p/(1-e)}{p/(1+e)} = \frac{(1+e)}{(1-e)}.$$

С другой стороны,

$$\frac{v_p}{v_a} = \frac{\sqrt{\frac{\mu}{p}}(1+e)}{\sqrt{\frac{\mu}{p}}(1-e)} = \frac{(1+e)}{(1-e)},$$

что и доказывает теорему.

Пример 4.4. Космический аппарат на высоте 300 км имел скорость 10 км/с, причем вектор скорости был направлен параллельно поверхности Земли. Рассчитать скорость аппарата в апогее, если известно, что расстояние от центра Земли до апогея составляет около 370 тыс. км.

Решение. Тот факт, что на высоте 300 км вектор скорости спутника был направлен параллельно поверхности Земли, означает, что в этой точке орбита спутника имеет перигей, т.е. расстояние до перигея равно 6671 км. Тогда по правилу рычага величина скорости спутника в апогее составляет

$$v_a = \frac{r_a}{r_p} v_p = 554,6 \text{ км} \cdot \text{с}^{-1}.$$

4.3.2. КРУГОВЫЕ ОРБИТЫ

Круговые орбиты являются частным случаем эллиптических. Выше было показано, что для круговых орбит выполняются условия

$$e = 0, \quad h < 0, \quad v = v_{кр}(r) = \sqrt{\frac{\mu}{r}}.$$

В отличие от эллиптической, круговая орбита полностью определяется одним параметром — чаще всего используется

радиус круговой орбиты или высота спутника над поверхностью притягивающего тела.

4.3.3. ГИПЕРБОЛИЧЕСКИЕ ОРБИТЫ

При движении спутника вокруг притягивающего центра по гиперболической орбите выполняются соотношения

$$e > 1, \quad h > 0, \quad v > \sqrt{\frac{2\mu}{r}}.$$

Рассмотрим такое движение подробнее (рис. 8).

Пусть, как и ранее, S — текущее положение спутника, движущегося вокруг притягивающего центра C ; O — центр гиперболы; P — перицентр орбиты. Истинная аномалия ϑ , так же как и для эллиптических орбит, представляет собой угол между текущим радиус-вектором спутника $\vec{r} = \overrightarrow{CS}$ и направлением $\overrightarrow{CP}$ из притягивающего центра на перицентр. Однако для гиперболических орбит истинная аномалия изменяется в пределах

$$-\vartheta_{\text{lim}} < \vartheta < \vartheta_{\text{lim}},$$

где ϑ_{lim} — предельное значение угла ϑ , при котором знаменатель правой части уравнения орбиты спутника (4.3) обращается в нуль, т. е. расстояние от спутника до притягивающего центра неограниченно увеличивается:

$$1 + e \cos \vartheta_{\text{lim}} = 0 \Rightarrow \vartheta_{\text{lim}} = \arccos\left(-\frac{1}{e}\right) = \pi - \arccos\left(\frac{1}{e}\right).$$

Фокальный параметр p и эксцентриситет e гиперболической орбиты определяются так же, как и для эллиптических орбит:

$$p = |CD|, \quad e = \frac{OC}{OP}.$$

Для гиперболических орбит определяют *действительную полуось* a и *мнимую полуось* b :

$$a = |OA| = |OP|, \quad b = |OB|.$$

Расстояние до перицентра равно

$$r_p = \frac{p}{1+e} = a(e-1).$$

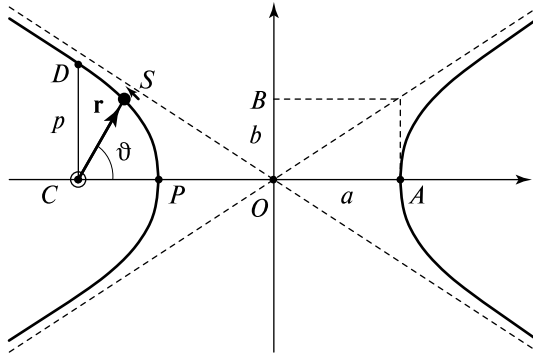


Рис. 8. Параметры гиперболической орбиты

Можно доказать справедливость следующих соотношений, аналогичных соотношениям (4.15)–(4.17):

$$a = \frac{p}{e^2 - 1}, \quad (4.20)$$

$$b^2 = a^2(e^2 - 1), \quad (4.21)$$

$$p = a(e^2 - 1) = \frac{b^2}{a}. \quad (4.22)$$

Эти соотношения позволяют полностью определить гиперболическую орбиту по любым двум из следующих параметров:

- фокальный параметр p ;
- эксцентриситет e ;
- действительная полуось a ;
- мнимая полуось b ;
- расстояние до перицентра r_p .

4.3.4. ПАРАБОЛИЧЕСКИЕ ОРБИТЫ

Параболические орбиты являются предельным случаем эллиптических и гиперболических орбит, соответствующим выполнению равенств

$$e = 1, \quad h = 0, \quad v = \sqrt{\frac{\mu}{r}}.$$

Уравнение параболической орбиты имеет вид

$$r = \frac{p}{1 + \cos \vartheta},$$

откуда непосредственно следует, что

$$r_p = \frac{p}{2}.$$

Из уравнения (4.15) или (4.20) следует, что для параболических орбит можно считать

$$a = \infty.$$

Истинная аномалия для параболических орбит меняется в пределах

$$-\pi < \vartheta < \pi,$$

форма и размер параболической орбиты, так же как и круговой, определяются одним параметром — например, расстоянием до перицентра.

4.4. ВРЕМЯ ПРОХОЖДЕНИЯ СПУТНИКА ЧЕРЕЗ ЗАДАННУЮ ТОЧКУ ОРБИТЫ

4.4.1. ЭЛЛИПТИЧЕСКИЕ ОРБИТЫ

Рассмотрим движение спутника S вокруг притягивающего центра C по эллиптической орбите с центром в точке O , большой полуосью a и эксцентриситетом e (рис. 9). Проведем окружность радиуса a с центром в точке O . Проведем через точку S перпендикулярно линии апсид отрезок прямой от линии апсид до пересечения с окружностью. Точку пересечения обозначим через S' , а величину угла COS' — через E .

Угловая величина E называется *эксцентрической аномалией*, она наряду с истинной аномалией ϑ характеризует смещение спутника относительно перицентра. Отличие заключается в том, что для эксцентрической аномалии угловое смещение спутника отсчитывается из центра орбитального эллипса O , а для истинной аномалии — из притягивающего центра C . Можно показать [Сихарулидзе, 2011], что истинная аномалия ϑ связана с эксцентрической аномалией E соотношением

$$\operatorname{tg} \frac{\vartheta}{2} = \sqrt{\frac{1+e}{1-e}} \operatorname{tg} \frac{E}{2}.$$

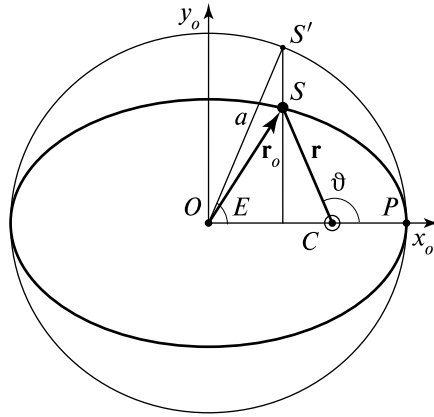


Рис. 9. Иллюстрация к выводу уравнения Кеплера для эллиптических орбит

Рассмотрим в плоскости орбиты декартову систему координат $Ox_o y_o$ с центром, совпадающим с центром орбитального эллипса O , и осью абсцисс, направленной на перигей P (см. рис. 9). Пусть текущее положение спутника определяется вектором $\mathbf{r}_o = \overline{OS}$, имеющим в этой системе координаты

$$\mathbf{r}_o = \begin{pmatrix} r_{x_o} \\ r_{y_o} \end{pmatrix}.$$

Тогда нетрудно убедиться, что

$$r_{x_o} = a \cos E,$$

$$r_{y_o} = b \sin E.$$

Пусть в некоторый момент времени t спутник находится в точке орбиты, эксцентрическая аномалия которой равна E . Обозначим через t_p время прохождения спутником перигея P . Тогда можно показать [Охочимский, Сихарулидзе, 1990], что моменты времени t и t_p связаны соотношением

$$t - t_p = \sqrt{\frac{a^3}{\mu}} (E - e \sin E), \quad (4.23)$$

которое называется *уравнением Кеплера для эллиптических орбит*. В этом уравнении μ — гравитационный параметр притягивающего центра.

Рассмотрим также величину

$$n = \sqrt{\frac{\mu}{a^3}},$$

называемую *средним движением*, и величину $M = n(t - t_p)$, называемую *средней аномалией*. Тогда уравнение (4.23) примет вид

$$E - e \sin E = M. \quad (4.24)$$

Теорема 4.2. При известных значениях средней аномалии M и эксцентриситета e уравнение Кеплера (4.24) всегда имеет единственное действительное решение относительно эксцентрической аномалии E .

Доказательство. Рассмотрим функцию $F(E) = E - \sin E - M$. Эта функция, очевидно, является непрерывной. Кроме того,

$$F(-\infty) = -\infty, \quad F(+\infty) = +\infty, \quad \frac{dF}{dE} = 1 - e \cos E > 0,$$

т. е. функция $F(E)$ — монотонно возрастающая, и, следовательно, уравнение $F(E) = 0$ всегда имеет единственное решение.

Решение уравнения (4.24) может быть найдено приближенно — например, с использованием схемы метода последовательных приближений [Эльясберг, 1965]:

$$E_n = M + e \sin E_{n-1},$$

где E_{n-1} и E_n — предыдущее и последующее приближения искомой величины E . В качестве первого приближения можно взять значение $E_1 = M$.

Из уравнения (4.23) следует, что время перелета спутника между двумя точками орбиты с эксцентрическими аномалиями E_1 и E_2 равно

$$t_2 - t_1 = \sqrt{\frac{a^3}{\mu}} (E_2 - E_1 - e(\sin E_2 - \sin E_1)).$$

Определение 4.5. *Периодом обращения спутника* при невозмущенном движении вокруг притягивающего центра называется промежуток времени между двумя последовательными прохождениями спутником одной и той же точки своей орбиты.

Пусть t_p — момент прохождения спутника через перигеицентр, t'_p — момент следующего прохождения через перигеицентр,

соответствующий значению эксцентрической аномалии $E=2\pi$. Тогда период обращения спутника T , согласно уравнению (4.23), будет равен

$$T = t'_p - t_p = 2\pi \sqrt{\frac{a^3}{\mu}}. \quad (4.25)$$

Как следует из последнего соотношения, период обращения спутника зависит только от величины большой полуоси орбиты и от гравитационного параметра притягивающего центра.

Пример 4.5. Космический аппарат «Восток-2» имел на первых витках высоту перигея 244 км и высоту апогея 183 км. Найти период обращения этого спутника.

Решение. Принимая радиус Земли равным $R_E=6371$ км, найдем расстояние до перигея и расстояние до апогея орбиты спутника:

$$r_p = R_E + h_p = 6371 + 183 = 6554 \text{ км},$$

$$r_a = R_E + h_a = 6371 + 244 = 6615 \text{ км}.$$

Тогда величина большой полуоси орбиты данного спутника равна

$$a = \frac{r_a + r_p}{2} = 6584,5 \text{ км}.$$

Теперь, принимая значение гравитационного параметра Земли равным $\mu = 3,986 \cdot 10^5 \text{ км}^3 \cdot \text{с}^{-2}$, найдем искомый период обращения:

$$T = 2\pi \sqrt{\frac{a^3}{\mu_E}} = 5317 \text{ с} = 1 \text{ ч } 28 \text{ мин } 37 \text{ с}.$$

Пример 4.6. Рассчитать радиус орбиты и скорость, которые должен иметь спутник, вращающийся по круговой орбите вокруг Земли в плоскости экватора, если требуется, чтобы этот спутник все время находился над одной и той же точкой поверхности Земли.

Решение. Для того чтобы спутник находился постоянно над одной и той же точкой поверхности Земли, необходимо, чтобы период T его обращения по орбите совпадал с периодом вращения Земли вокруг своей оси. Примем вначале для простоты значение этого периода равным 24 ч или 86 400 с. Тогда

величина большой полуоси орбиты спутника, которая для круговой орбиты совпадает, очевидно, с ее радиусом, должна составлять

$$a = \sqrt[3]{\left(\frac{T}{2\pi}\right)^2 \mu_E} = 42\,241,1 \text{ км},$$

т.е. высота такой орбиты будет равна 35 870 км. Скорость движения спутника по этой орбите, согласно формуле (4.13), составит

$$v = v_{\text{кр}}(a) = \sqrt{\frac{\mu_E}{a}} \approx 3,072 \text{ км} \cdot \text{с}^{-1}.$$

Заметим, что если рассмотреть более точное значение периода обращения Земли, равное 23 ч 56 мин 4 с или 86 164 с, то радиус круговой орбиты спутника получится равным 42 164 км, т.е. высота орбиты будет равной 35 793 км. Скорость движения спутника по круговой орбите такого радиуса составит 3,075 км·с⁻¹. Как видим, изменение требуемого значения периода обращения на 236 с привело к изменению высоты орбиты спутника на 77 км.

Замечание 4.3. Если рассмотреть два спутника, движущиеся вокруг общего притягивающего центра по эллиптическим орбитам со значениями большой полуоси a_1 и a_2 и имеющие периоды обращения T_1 и T_2 соответственно, то из формулы (4.25) следует соотношение

$$\frac{T_1^2}{T_2^2} = \frac{a_1^3}{a_2^3},$$

выражающее третий закон Кеплера.

4.4.2. ГИПЕРБОЛИЧЕСКИЕ ОРБИТЫ

Теперь рассмотрим случай, когда спутник S движется вокруг притягивающего центра C по гиперболической орбите с центром в точке O , действительной полуосью a и эксцентриситетом e (рис. 10). Рассмотрим новую переменную H такую, что

$$E = iH, \quad H = -iE,$$

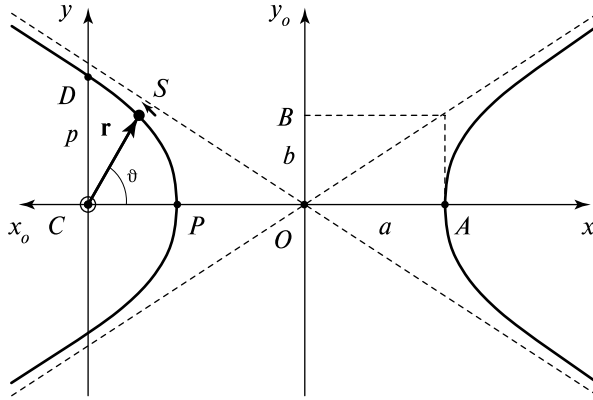


Рис. 10. Иллюстрация к выводу уравнения Кеплера для гиперболических орбит

где i — мнимая единица, определяемая как решение уравнения $i^2 = -1$. Величина H является аналогом эксцентрической аномалии E для случая гиперболического движения. Истинная аномалия ϑ связана с H соотношением [Сихарулидзе, 2011]

$$\operatorname{tg} \frac{\vartheta}{2} = \sqrt{\frac{e+1}{e-1}} \operatorname{th} \frac{H}{2}.$$

Рассмотрим, как и в предыдущем случае, в плоскости орбиты декартову систему координат $Ox_o y_o$ с центром, совпадающим с центром орбитального эллипса O , и осью абсцисс, направленной из точки O на перицентр P (см. рис. 10). Обозначим через $\mathbf{r}_o$ радиус-вектор, определяющий текущее положение спутника в этой системе координат. Тогда компоненты r_{x_o} , r_{y_o} этого радиус-вектора будут определяться по формулам

$$r_{x_o} = -a \operatorname{ch} H, \quad r_{y_o} = b \operatorname{sh} H$$

(эти формулы справедливы для случая, когда рассматривается движение по левой ветви гиперболы, как показано на рис. 10).

Если в некоторый момент времени t спутник находится в точке орбиты, эксцентрическая аномалия которой равна H , и t_p — время прохождения спутника через перицентр P , то можно показать [Охоцимский, Сихарулидзе, 1990], что моменты времени t и t_p связаны соотношением

$$t - t_p = \sqrt{\frac{a^3}{\mu}} (e \operatorname{sh} H - H), \quad (4.26)$$

которое называется *уравнением Кеплера для гиперболических орбит*.

Последнее соотношение приводимо к виду

$$e \operatorname{sh} H - H = M, \quad (4.27)$$

где $M = n(t - t_p)$ — средняя аномалия; n — среднее движение.

4.5. ФОРМУЛА ЛАМБЕРТА

Рассмотрим движение спутника S по эллиптической орбите с большой полуосью a вокруг притягивающего центра C с известным гравитационным параметром μ (рис. 11).

Пусть S_1 и S_2 — две различные точки орбиты, s — расстояние между ними, r_1 и r_2 — расстояния от этих точек до притягивающего центра. Тогда время перелета спутника от точки S_1 до точки S_2 может быть найдено по *формуле Ламберта** [Балк, 1965; Иванов, Лысенко, 2004]

* Иоганн Генрих Ламберт (1728–1777) — выдающийся немецкий астроном, физик, математик, философ. Разработал принципы фотометрии — основного метода наблюдательной астрономии. С помощью своих методов Ламберт смог точно оценить относительную яркость Луны, ослабление света в земной атмосфере, также на основе фотометрии им проводилась оценка межзвёздных расстояний. Занимался исследованиями орбит комет. Ламберту принадлежит и космологическая концепция развития Вселенной, во многом предвосхитившая научные открытия будущего.

Ламберт внес также огромный вклад в развитие математики, особенно тригонометрии, теории конических сечений и гиперболических функций, стал одним из основателей неевклидовой геометрии. Он впервые доказал иррациональность чисел π и e , составил таблицу простых чисел до 102 000. Значительный вклад Ламберт внес и в развитие философии, он состоял в переписке с И. Кантом, высоко ценившим его ум и способности.

Ламберт был членом нескольких научных обществ — литературного и физического в Куре, членом-корреспондентом научного общества в Гёттингене, членом Баварского научного общества и Берлинской Академии наук. Скончался в возрасте 49 лет.

В честь Ламберта назван астероид, его имя присвоено целому ряду математических понятий.

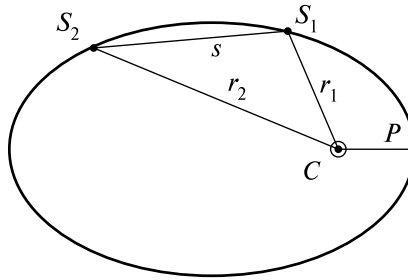


Рис. 11. Иллюстрация к формуле Ламберта

$$t_2 - t_1 = \sqrt{\frac{a^3}{\mu}} [(\lambda_1 - \sin \lambda_1) - (\lambda_2 - \sin \lambda_2)],$$

где λ_1 и λ_2 — корни уравнений

$$\cos \lambda_1 = 1 - \frac{r_1 + r_2 + s}{2a},$$

$$\cos \lambda_2 = 1 - \frac{r_1 + r_2 - s}{2a}.$$

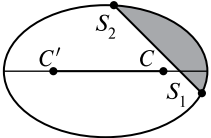
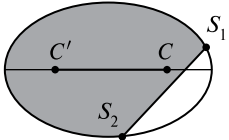
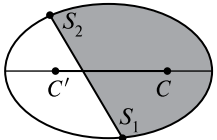
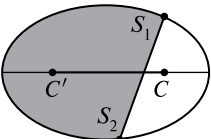
Очевидно, что эти уравнения имеют бесконечное множество решений. Приведем правила, позволяющие однозначно выбрать числа λ_1 и λ_2 для подстановки в формулу Ламберта [Балк, 1965].

1. Если спутник, двигаясь по эллиптической орбите, проходит дугу $\overline{S_1 S_2}$ после прохождения перицентра и до прохождения следующего апоцентра (т.е. так, как показано на рис. 11), то следует выбирать числа λ_1 и λ_2 из условия

$$0 < \lambda_1 < \pi, \quad 0 < \lambda_2 < \pi.$$

2. В других случаях числа λ_1 и λ_2 следует выбирать в зависимости от того, как расположены фокусы C и C' эллипса орбиты по отношению к сегменту, ограниченному дугой $\overline{S_1 S_2}$ и хордой $\overline{S_1 S_2}$. Возможные варианты выбора приводятся в табл. 1. В двух последних колонках символами $\overline{\lambda_1}$ и $\overline{\lambda_2}$ обозначены наименьшие положительные углы, удовлетворяющие двум приведенным выше уравнениям для определения λ_1 и λ_2 .

Таблица 1. Выбор чисел λ_1 и λ_2 при различных случаях расположения точек S_1 и S_2 на орбитальном эллипсе

Возможные случаи		λ_1	λ_2
1. Сегмент не содержит ни C , ни C'		$\overline{\lambda_1}$	$\overline{\lambda_2}$
2. Сегмент содержит C и C'		$2\pi - \overline{\lambda_1}$	$-\overline{\lambda_2}$
3. Сегмент содержит C и не содержит C'		$\overline{\lambda_1}$	$-\overline{\lambda_2}$
4. Сегмент не содержит C , но содержит C'		$2\pi - \overline{\lambda_1}$	$\overline{\lambda_2}$

Пример 4.7. Космический корабль совершает перелет между орбитами Земли и Марса. Можно принять, что орбита перелета является эллиптической, причем в одном из ее фокусов расположено Солнце, а расстояния перигелия и афелия данной орбиты составляют 120 и 240 млн км соответственно. Точка начала перелета расположена на расстоянии 150 млн км от Солнца, точка окончания перелета — на расстоянии 228 млн км от Солнца. Сколько времени должен занять такой перелет?

Решение. Для нахождения времени перелета между двумя точками орбиты удобно воспользоваться формулой Ламберта. Для этого по известным значениям расстояний до афелия и перигелия сначала найдем значения a , e и p :

$$a = \frac{r_a + r_p}{2} = 180 \text{ млн км}, \quad e = \frac{r_a - r_p}{r_a + r_p} = \frac{1}{3},$$

$$p = a(1 - e^2) = 160 \text{ млн км}.$$

Теперь найдем истинные аномалии ϑ_1 и ϑ_2 точек S_1 и S_2 . Из уравнения орбиты (4.3) находим:

$$\vartheta = \arccos \frac{p-r}{er},$$

откуда $\vartheta_1 = 1,369$ рад, $\vartheta_2 = 2,679$ рад. Тогда расстояние s между точками S_1 и S_2 равно

$$s = \sqrt{r_1^2 + r_2^2 - 2r_1r_2 \cos(\vartheta_2 - \vartheta_1)} \approx 238,3 \text{ млн км.}$$

Найдем теперь множители λ_1 и λ_2 :

$$\cos \lambda_1 = 1 - \frac{r_1 + r_2 + s}{2a} \approx -0,71, \quad \cos \lambda_2 = 1 - \frac{r_1 + r_2 - s}{2a} \approx 0,61,$$

откуда

$$\lambda_1 = 2,363 \text{ рад,} \quad \sin \lambda_1 = 0,702;$$

$$\lambda_2 = 0,912 \text{ рад,} \quad \sin \lambda_2 = 0,791.$$

Подставляя найденные величины в формулу Ламберта, найдём

$$\begin{aligned} t_2 - t_1 &= \sqrt{\frac{(180 \cdot 10^6)^3}{132,5 \cdot 10^9}} \cdot [(2,363 - 0,702) - (0,912 - 0,791)] = \\ &= 10,21406 \cdot 10^6 \text{ с} = 118,22 \text{ сут.} \end{aligned}$$

5. НЕВОЗМУЩЕННОЕ ДВИЖЕНИЕ СПУТНИКА В ИНЕРЦИАЛЬНОЙ СИСТЕМЕ КООРДИНАТ

5.1. ЭЛЕМЕНТЫ ОРБИТЫ

До сих пор нами рассматривалось движение спутника в плоскости орбиты и ставилась задача определения положения спутника на орбите по тем или иным известным величинам. Перейдем теперь к рассмотрению задачи определения пространственного положения спутника в какой-либо заданный момент времени относительно некоторой инерциальной системы координат.

Прежде всего отметим, что положение в пространстве любой плоскости (в частности, плоскости орбиты) можно определить значениями трех независимых параметров. Далее, как уже было ранее показано, форма и размеры невозмущенной орбиты спутника в плоскости орбиты полностью определяются значениями двух независимых параметров. Наконец, если форма и размер орбиты известны, то для определения положения спутника на орбите в произвольный момент времени требуется определить еще один параметр — например, истинную аномалию спутника в текущий момент времени или время прохождения спутником некоторой точки орбиты. Таким образом, орбита невозмущенного движения спутника в пространстве может быть определена с помощью пяти постоянных величин. Шестой параметр позволит определить, в какой точке своей орбиты находится спутник в заданный момент времени (т.е. определить пространственное положение спутника в этот момент времени).

Определение 5.1. *Элементами орбиты* будем называть величины, характеризующие орбиту спутника и его положение на орбите. *Полной системой* элементов орбиты будем называть систему из шести независимых элементов орбиты, однозначно определяющих положение спутника в пространстве в любой заданный момент времени.

Рассмотрим движение спутника S по эллиптической орбите вокруг притягивающего центра C (рис. 12). Зададим в пространстве некоторую инерциальную систему отсчета $Cxuz$, начало которой примем совпадающим с притягивающим центром C . Орбита спутника будет пересекать плоскость Cxu в двух точках. Точку, в которой при движении по орбите аппликата z спутника будет менять знак с отрицательного на положительный, будем называть *восходящим узлом* и обозначать символом ϱ . Точку, в которой при движении по орбите аппликата спутника будет менять знак с положительного на отрицательный, будем называть *нисходящим узлом* и обозначать символом $\wp$. Линию, соединяющую точки ϱ и $\wp$, будем называть *линией узлов*.

Определение 5.2. *Долготой восходящего узла орбиты* называется угол Ω между положительным направлением оси абсцисс Cx и направлением из притягивающего центра на восходящий узел (см. рис. 12). Долгота восходящего узла меняется в пределах от 0 до 2π .

Определение 5.3. *Наклоением орбиты* называется угол i между плоскостью орбиты и координатной плоскостью Cxu . Наклонение орбиты меняется в пределах от 0 до π . Если для заданной орбиты $i < \frac{\pi}{2}$, то движение спутника называется *прямым*, если же $\frac{\pi}{2} < i < \pi$, то движение называется *обратным*. При $i = 0$ орбита спутника называется *экваториальной*, при $i = \frac{\pi}{2}$ — *полярной*.

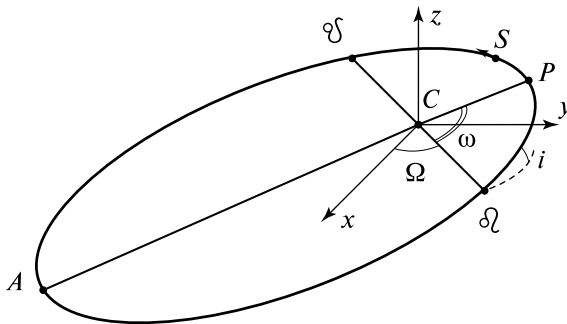


Рис. 12. Элементы орбиты спутника: наклонение i , долгота восходящего узла Ω и аргумент перигея ω

Определение 5.4. *Аргументом перицентра орбиты* называется угол ω между направлениями из притягивающего центра на восходящий узел и на перицентр орбиты. Аргумент перицентра отсчитывается в направлении движения спутника и меняется в пределах от 0 до 2π .

Плоскость S_{xy} часто называют *базовой плоскостью*. При рассмотрении движения вокруг Земли в качестве базовой плоскости как правило выбирается экваториальная плоскость, при рассмотрении движения вокруг Солнца — плоскость эклиптики и т. д.

В качестве полной системы элементов орбиты для определения пространственного положения спутника выберем следующие величины:

- наклонение орбиты i , долготу восходящего узла Ω и аргумент перицентра ω , определяющие пространственное положение плоскости орбиты;
- большую полуось a и эксцентриситет e , характеризующие форму и размер орбиты;
- время $\tau = t_p$ прохождения спутника через перицентр.

Определение 5.5. Элементы орбиты a , e , i , Ω , ω , τ образуют полную систему и называются *кеплеровыми элементами невозмущенного движения*.

Замечание 5.1. Полную систему элементов орбиты, очевидно, можно выбрать множеством способов. В частности, вместо большой полуоси орбиты a можно рассмотреть фокальный параметр p , вместо аргумента перицентра ω или вместо времени прохождения перицентра τ — угол $u = \vartheta + \omega$, называемый *аргументом широты*, вместо параметра τ также часто рассматривают среднюю аномалию M и т. д. Особенностью кеплеровой системы элементов является то, что с ее помощью движение спутника определяется шестью постоянными величинами, не зависящими от времени.

5.2. ОПРЕДЕЛЕНИЕ ПОЛОЖЕНИЯ И СКОРОСТИ СПУТНИКА ПО ЭЛЕМЕНТАМ ОРБИТЫ

Покажем, как по известным значениям кеплеровых элементов орбиты определить компоненты r_x , r_y и r_z радиус-вектора $\mathbf{r}$

и компоненты v_x, v_y, v_z вектора скорости $\mathbf{v}$ спутника в заданный момент времени t [Эльясберг, 1965].

1. Находим фокальный параметр $p = a(1 - e^2)$ и среднюю аномалию $M = \sqrt{\frac{\mu}{a^3}} (t - \tau)$.

2. Из уравнения Кеплера $E - e \sin E = M$ какой-либо численной процедурой находим эксцентрическую аномалию E .

3. По формуле $\vartheta = \sqrt{\frac{1+e}{1-e}} \operatorname{tg} \frac{E}{2}$ находим значение истинной аномалии ϑ , а затем — аргумент широты $u = \vartheta + \omega$.

4. Из уравнения орбиты $r = \frac{p}{1 + e \cos \vartheta}$ находим расстояние до спутника r .

5. Координаты спутника определяем из следующих соотношений:

$$r_x = r(\cos \Omega \cdot \cos u - \sin \Omega \cdot \sin u \cdot \cos i),$$

$$r_y = r(\sin \Omega \cdot \cos u + \cos \Omega \cdot \sin u \cdot \cos i),$$

$$r_z = r \sin u \cdot \sin i.$$

6. Определяем величину радиальной и трансверсальной компонент вектора скорости:

$$v_r = \sqrt{\frac{\mu}{p}} e \cos \vartheta, \quad v_n = \sqrt{\frac{\mu}{p}} (1 + e \cos \vartheta).$$

7. Компоненты вектора скорости определяем по формулам

$$v_x = v_r (\cos \Omega \cdot \cos u - \sin \Omega \cdot \sin u \cdot \cos i) -$$

$$- v_n (\cos \Omega \cdot \sin u + \sin \Omega \cdot \cos u \cdot \cos i),$$

$$v_y = v_r (\sin \Omega \cdot \cos u + \cos \Omega \cdot \sin u \cdot \cos i) -$$

$$- v_n (\sin \Omega \cdot \sin u - \cos \Omega \cdot \cos u \cdot \cos i),$$

$$v_z = v_r \sin u \cdot \sin i + v_n \cos u \cdot \sin i.$$

5.3. ОПРЕДЕЛЕНИЕ ЭЛЕМЕНТОВ ОРБИТЫ СПУТНИКА ПО ПОЛОЖЕНИЮ И СКОРОСТИ

Теперь покажем, как вычислить кеплеровы элементы орбиты, если в некоторый заданный момент времени t известны поло-

жение спутника, определяемое радиус-вектором $\mathbf{r} = (r_x, r_y, r_z)'$, и вектор его скорости $\mathbf{v} = (v_x, v_y, v_z)'$ [Эльясберг, 1965].

1. Вычислим вспомогательные величины (компоненты векторной постоянной площадей $\mathbf{c}$)

$$c_x = r_y v_z - r_z v_y,$$

$$c_y = r_z v_x - r_x v_z,$$

$$c_z = r_x v_y - r_y v_x,$$

$$c = |\mathbf{c}| = \sqrt{c_x^2 + c_y^2 + c_z^2}$$

и найдем долготу восходящего узла Ω , а также наклонение орбиты i по формулам

$$\Omega = -\frac{c_x}{c_y}, \quad \cos i = \frac{c_z}{c}, \quad 0 \leq i < \pi.$$

2. Введем вспомогательный угол γ (угол между вектором скорости и нормалью к радиус-вектору спутника) и найдем его из условия

$$\sin \gamma = \frac{\mathbf{r} \cdot \mathbf{v}}{rv}, \quad -\frac{\pi}{2} \leq \gamma \leq \frac{\pi}{2},$$

$$\text{где } r = |\mathbf{r}| = \sqrt{r_x^2 + r_y^2 + r_z^2}, \quad v = |\mathbf{v}| = \sqrt{v_x^2 + v_y^2 + v_z^2}.$$

Тогда, очевидно,

$$\cos \gamma = \frac{\sqrt{r^2 v^2 - (\mathbf{r} \cdot \mathbf{v})^2}}{rv}.$$

Вычислим также величину $k = \frac{rv^2}{\mu}$.

3. Большую полуось a и эксцентриситет орбиты e найдем по формулам

$$a = \frac{r}{2-k}, \quad e = \sqrt{1 - k(2-k)\cos^2 \gamma}.$$

4. Истинную аномалию ϑ определим из условий

$$\sin \vartheta = \frac{k \sin \theta \cdot \cos \gamma}{e}, \quad \cos \vartheta = \frac{k \cos^2 \gamma - 1}{e},$$

аргумент широты u — из условий

$$\cos u = \frac{r_x \cos \Omega + r_y \sin \Omega}{r}, \quad \sin u = \frac{r_z}{r \sin i},$$

и затем аргумент перицентра ω — из равенства $\omega = u - \vartheta$.

5. Время прохождения перицентра τ найдем по формуле

$$\tau = t - \sqrt{\frac{a^3}{\mu}} (E - e \sin E),$$

где эксцентрическая аномалия E определяется по формуле

$$\operatorname{tg} \frac{E}{2} = \sqrt{\frac{1-e}{1+e}} \operatorname{tg} \frac{\vartheta}{2}.$$

6. ВОЗМУЩЕННОЕ ДВИЖЕНИЕ СПУТНИКА

6.1. ПОНЯТИЕ О ВОЗМУЩЕННОМ ДВИЖЕНИИ. МЕТОД ОСКУЛИРУЮЩИХ ЭЛЕМЕНТОВ

Рассмотренная выше ограниченная задача двух тел описывает, вообще говоря, идеальный случай, когда спутник движется при наличии только гравитационного взаимодействия с притягивающим телом. При описании реального движения спутника приходится учитывать целый ряд других факторов, оказывающих влияние на характер его орбиты. Такие факторы носят название *возмущающих факторов орбиты*, *возмущающих воздействий* или, короче, *возмущений*. К основным типам возмущений относятся [Суханов, 2010]:

- гравитационные возмущения: влияние несферичности притягивающего тела, влияние притяжения других небесных тел;
- негравитационные возмущения: сопротивление атмосферы, световое давление, реактивные возмущения, вызываемые воздействием системы ориентации или утечкой газа с космического аппарата, уменьшение массы спутника за счет расхода топлива, воздействия различных электромагнитных полей и т. д.

Учёт возмущающих факторов приводит к гораздо более сложным математическим моделям, но позволяет получить решения, описывающие реальное движение спутника с большей точностью.

Определение 6.1. Движение спутника с учетом тех или иных возмущающих воздействий называется *возмущенным движением*, а математическая модель, описывающая возмущенное движение, — *моделью возмущенного движения*.

Уравнение возмущенного движения спутника вокруг притягивающего тела в самом общем виде записывается следующим образом:

$$\ddot{\mathbf{r}} = -\frac{\mu}{r^3} \mathbf{r} + \boldsymbol{\psi}, \quad (6.1)$$

где $\boldsymbol{\psi}$ — суммарное ускорение, получаемое спутником в результате действия всех возмущающих факторов и называемое *возмущающим ускорением*. Если возмущающее ускорение отлично от нуля, то движение, очевидно, уже не будет совершаться по одному из конических сечений. Однако в каждый момент времени t по известным значениям $\mathbf{r}(t)$ и $\mathbf{v}(t)$ при помощи описанной выше процедуры можно подобрать значения кеплеровых элементов орбиты, соответствующих невозмущенному движению в данной точке пространства с данной скоростью. Иначе говоря, в каждый момент времени реальная орбита движения спутника будет касаться некоторой орбиты невозмущенного движения, соответствующей найденным значениям кеплеровых элементов орбиты. Таким образом мы можем получить зависимости кеплеровых элементов от времени, т.е. рассматривать функции $a(t)$, $e(t)$, $i(t)$, $\Omega(t)$, $\omega(t)$, $\tau(t)$, описывающие возмущенное движение спутника. Если эти функции известны, то положение и скорость спутника в любой момент времени t могут быть определены аналогично случаю невозмущенного движения.

Определение 6.2. Рассмотренные функции $a(t)$, $e(t)$, $i(t)$, $\Omega(t)$, $\omega(t)$, $\tau(t)$, представляющие собой элементы непрерывно меняющейся со временем орбиты невозмущенного движения, называются *оскулирующими элементами*, а сама такая орбита — *оскулирующей орбитой*. Описание движения спутника при помощи оскулирующей орбиты называется *методом оскулирующих элементов**.

Введем обозначения:

$$\mathbf{x} = \begin{pmatrix} \mathbf{r} \\ \mathbf{v} \end{pmatrix}, \quad \mathbf{f}(\mathbf{x}) = \begin{pmatrix} \mathbf{v} \\ -\frac{\mu}{r^3} \mathbf{r} \end{pmatrix}, \quad \mathbf{g}(\mathbf{x}, t) = \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\psi} \end{pmatrix},$$

где $\mathbf{x}$ — вектор состояния, $\mathbf{f}(\mathbf{x})$ — правая часть векторной формы уравнения невозмущенного движения спутника (3.1).

* Идея и разработка метода оскулирующих элементов принадлежат Ж. Лагранжу.

Тогда уравнение возмущенного движения в векторной форме запишется в виде

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}, t).$$

Обозначим через $\mathbf{q}$ вектор оскулирующих элементов орбиты

$$\mathbf{q} = \mathbf{q}(\mathbf{x}, t) = (a(t), e(t), i(t), \Omega(t), \omega(t), \tau(t))'.$$

Тогда

$$\dot{\mathbf{q}} = \frac{\partial \mathbf{q}}{\partial \mathbf{x}} \dot{\mathbf{x}} = \frac{\partial \mathbf{q}}{\partial \mathbf{x}} [\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})].$$

Если $\mathbf{g}(\mathbf{x}, t) = \mathbf{0}$ т. е. движение невозмущенное, то $\mathbf{q} = \text{const}$ и, следовательно,

$$\frac{\partial \mathbf{q}}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}) = \mathbf{0}.$$

Тогда

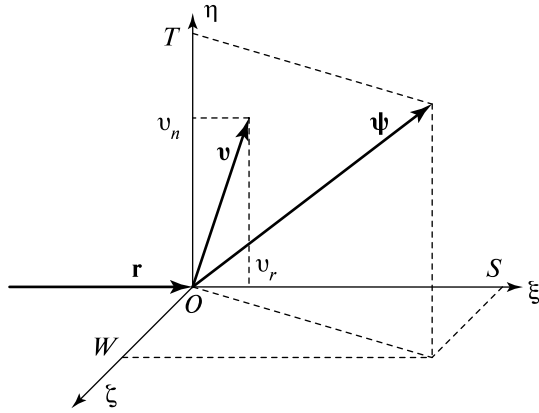
$$\dot{\mathbf{q}} = \frac{\partial \mathbf{q}}{\partial \mathbf{x}} \mathbf{g}(\mathbf{x}, t) = \frac{\partial \mathbf{q}}{\partial \mathbf{v}} \boldsymbol{\psi}.$$

Введем орбитальную систему координат $O\xi\eta\zeta$ (рис. 13). Начало O этой системы координат совместим с центром масс спутника, ось $O\xi$ направим вдоль вектора $\mathbf{r}$, ось $O\eta$ расположим в плоскости векторов $\mathbf{r}$ и $\mathbf{v}$ перпендикулярно оси $O\xi$ и направим так, чтобы ось $O\eta$ составляла с вектором скорости $\mathbf{v}$ острый угол. Ось $O\xi$ расположим перпендикулярно осям $O\xi$ и $O\eta$ так, чтобы система координат была правой.

Пусть векторы $\mathbf{v}$ и $\boldsymbol{\psi}$ в орбитальной системе имеют координаты

$$\mathbf{v} = \begin{pmatrix} v_\xi \\ v_\eta \\ v_\zeta \end{pmatrix}, \quad \boldsymbol{\psi} = \begin{pmatrix} S \\ T \\ W \end{pmatrix}.$$

Очевидно, что $v_\xi = v_r$ — радиальная скорость, $v_\eta = v_n$ — трансверсальная скорость, $v_\zeta = 0$. Также нетрудно заметить, что $\dot{v}_\zeta = W$ (в правой части уравнения возмущенного движения (6.1) первое слагаемое представляет собой правую часть уравнения невозмущенного движения (3.1), т. е. вектор, лежащий в плоскости орбиты; следовательно, третья компонента этого вектора в орбитальной системе координат равна нулю).

Рис. 13. Векторы $\mathbf{r}$ и $\mathbf{v}$ в орбитальной системе координат

Тогда

$$\dot{\mathbf{q}} = \frac{\partial \mathbf{q}}{\partial v_r} S + \frac{\partial \mathbf{q}}{\partial v_n} T + \frac{\partial \mathbf{q}}{\partial v_\zeta} W.$$

Подставляя в это соотношение выражения для кеплеровых элементов орбиты, в результате можно получить следующую систему уравнений для оскулирующих элементов, которая называется *уравнениями Ньютона – Лагранжа*:

$$\begin{aligned} \dot{a} &= \sqrt{\frac{p}{\mu}} \frac{2a}{1-e^2} \left(e \sin \vartheta S + \frac{p}{r} T \right), \\ \dot{e} &= \sqrt{\frac{p}{\mu}} \left(\sin \vartheta S + \left(1 + \frac{r}{p} \right) \cos \vartheta T + \frac{r}{p} e T \right), \\ \dot{i} &= \frac{r}{\sqrt{\mu p}} \cos u W, \\ \dot{\Omega} &= \frac{r}{\sqrt{\mu p}} \frac{\sin u}{\sin i} W, \\ \dot{\omega} &= \sqrt{\frac{p}{\mu}} \left(-\frac{\cos \vartheta}{e} S + \frac{1}{e} \left(1 + \frac{r}{p} \right) \sin \vartheta T - \frac{r}{p} \frac{\sin u}{\operatorname{tg} i} W \right), \\ \dot{\tau} &= \frac{r^2}{\mu e} \left(e N \sin \vartheta S - \cos \vartheta S + \frac{p}{r} N T \right), \end{aligned} \tag{6.2}$$

где $u = \vartheta + \omega$ — аргумент широты, $N = \frac{2p^2}{r^2} \int_0^{\vartheta} \frac{\cos \vartheta d\vartheta}{(1 + e \cos \vartheta)^2}$.

Часто на практике вместо элемента орбиты τ рассматривают истинную аномалию ϑ . Тогда последнее из уравнений Ньютона—Лагранжа* (6.2) заменяется более простым уравнением

$$\dot{\vartheta} = \sqrt{\frac{p}{\mu}} \left(\frac{\mu}{r^2} + \frac{\cos \vartheta}{e} S - \frac{1}{e} \left(1 + \frac{r}{p} \right) \sin \vartheta T \right).$$

Если возмущающее ускорение Ψ не зависит явно от времени, то правые части уравнений Ньютона—Лагранжа также не зависят явно от t . Поэтому в качестве независимой переменной в этом случае как правило выбирают аргумент широты $u = \omega + \vartheta$ учитывая, что

$$\dot{u} = \frac{\sqrt{\mu p}}{r^2} \left(1 - \frac{r^3}{\mu p} \frac{\sin u}{\operatorname{tg} i} W \right).$$

* Жозеф Луи Лагранж (1736–1813) — великий французский математик, астроном и механик итальянского происхождения. Наряду с Эйлером признан лучшим математиком XVIII века. Особенно прославился исключительным мастерством в области обобщения и синтеза накопленного научного материала.

Автор классического трактата «Аналитическая механика», в котором установил фундаментальный «принцип возможных перемещений» и завершил математизацию механики. Внёс грандиозный вклад в развитие математического анализа, теории дифференциальных уравнений, теории чисел, теории вероятностей и численных методов, создал вариационное исчисление.

Среди достижений Лагранжа — открытие точек либрации в ограниченной задаче трех тел, ряд работ по теории движения Луны и спутников Юпитера. Также Лагранжу принадлежат знаменитая интерполяционная формула для приближения функции многочленом, формула остаточного члена ряда Тейлора, метод множителей Лагранжа для решения задач на условный экстремум и многие другие выдающиеся результаты.

Лагранж состоял членом нескольких Академий наук, причем в Берлинскую был избран в возрасте 20 лет (!) по рекомендации Л. Эйлера и Д'Аламбера.

Имя Лагранжа внесено в список 72 величайших учёных Франции, помещённый на первом этаже Эйфелевой башни.

Обозначим

$$\kappa = \frac{1}{1 - \frac{r^3 \sin u}{\mu p} W}. \quad (6.3)$$

Тогда $\dot{u} = \frac{\sqrt{\mu p}}{r^2 \kappa}$, и система уравнений (6.2) преобразуется к виду

$$\begin{aligned} \frac{da}{du} &= \frac{\dot{a}}{\dot{u}} = \frac{r^2 \kappa}{\mu} \frac{2a}{1-e^2} \left(e \sin \vartheta S + \frac{p}{r} T \right), \\ \frac{de}{du} &= \frac{\dot{e}}{\dot{u}} = \frac{r^2 \kappa}{\mu} \left(\sin \vartheta S + \left(1 + \frac{r}{p} \right) \cos \vartheta T + \frac{r}{p} e T \right), \\ \frac{di}{du} &= \frac{\dot{i}}{\dot{u}} = \frac{r^3 \kappa}{\mu p} \cos u W, \\ \frac{d\Omega}{du} &= \frac{\dot{\Omega}}{\dot{u}} = \frac{r^3 \kappa \sin u}{\mu p \sin i} W, \\ \frac{d\omega}{du} &= \frac{\dot{\omega}}{\dot{u}} = \frac{r^2 \kappa}{\mu e} \left(-\cos \vartheta S + e \left(1 + \frac{r}{p} \right) \sin \vartheta T - e \frac{r \sin u}{p} \frac{1}{\operatorname{tg} i} W \right), \\ \frac{d\tau}{du} &= \frac{\dot{\tau}}{\dot{u}} = \frac{r^4 \kappa}{\mu e} \frac{1}{\sqrt{\mu p}} \left(e N \sin \vartheta S - \cos \vartheta S + \frac{p}{r} N T \right), \end{aligned} \quad (6.4)$$

где κ определяется формулой (6.3).

Начальные условия для уравнений Ньютона—Лагранжа определяются по начальному положению спутника согласно формулам, приведенным в п. 5.3.

7. ЗАДАЧА ОПРЕДЕЛЕНИЯ ДВИЖЕНИЯ

7.1. МОДЕЛЬ ОЦЕНИВАНИЯ

Каждый человек в своей жизни и общество в целом ежеминутно решают задачи нахождения оценок (или, короче, *оценивания*) параметров тех или иных объектов (систем). При этом исследуемая система описывается с помощью совокупности некоторых *параметров*, характеризующих ее и изменяющихся по каким-либо законам. Бытовые оценки параметров носят обычно описательный характер: богатый — бедный, высокий — невысокий, умный — неумный и т.п. Способ нахождения оценок параметров называют *методом оценивания*. При использовании статистических методов оценивания рассматривается *математическая модель*, описывающая поведение системы с некоторой точностью, и определяются количественные оценки одного или нескольких параметров этой модели.

Оценивание параметров производится обычно на основе наблюдения за поведением исследуемой системы на некотором интервале времени, при этом измеряются какие-либо величины, связанные с системой. Например, при изучении движения искусственного спутника Земли может измеряться дальность до него в различные моменты времени, и на основе этих наблюдений может быть получена оценка элементов орбиты спутника либо, например, таких параметров его движения как скорость или ускорение. Используемая математическая модель в этом случае будет представлять собой систему уравнений, связывающих измеренные значения дальности до спутника с оцениваемыми параметрами.

Определение 7.1. *Определением движения* будем называть задачу оценивания параметров математической модели, описывающей рассматриваемую движущуюся систему, по результатам обработки измерительной информации. Саму математическую

модель, связывающую измерения с оцениваемыми параметрами, будем называть далее *моделью измерений*.

Будем считать, что модель измерений включает m неизвестных параметров. Обозначим через $\boldsymbol{\theta} \in \mathbb{R}^m$ вектор из этих параметров. Обычно вектор $\boldsymbol{\theta}$ включает начальные условия движения системы и ряд постоянных, определяющих модель движения (например, для движения спутника — параметры атмосферы, аэродинамические характеристики спутника и т. п.). В соответствии с этим будем полагать, что модель измерений представляется в виде

$$\mathbf{y} = \mathbf{f}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}. \quad (7.1)$$

Здесь $\mathbf{y} \in \mathbb{R}^n$ — вектор измерений, $\mathbf{f}: \mathbb{R}^m \rightarrow \mathbb{R}^n$ — вектор-функция, определяющая модель, $\boldsymbol{\varepsilon}$ — суммарный вектор ошибок модели и измерений.

7.2. НЕСМЕЩЕННЫЙ АЛГОРИТМ ОЦЕНИВАНИЯ

Если $n = m$, то естественно искать оценку $\hat{\boldsymbol{\theta}}$ вектора $\boldsymbol{\theta}$ как решение системы уравнений

$$\mathbf{y} = \mathbf{f}(\hat{\boldsymbol{\theta}}).$$

В случае, когда число измерений равно числу оцениваемых параметров ($n = m$) и существует однозначная обратная функция $\mathbf{f}^{-1}(\mathbf{y})$, решение последних уравнений имеет вид

$$\hat{\boldsymbol{\theta}} = \mathbf{f}^{-1}(\mathbf{y}).$$

Если положить $\boldsymbol{\varepsilon} = \mathbf{0}$, то определенный таким образом *оценитель* при $n = m$ обладает свойством $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}$, т. е. оценка параметра совпадает с его истинным значением при отсутствии ошибок исходных данных. Такое свойство будем называть *несмещённостью оценителя*.

Рассмотрим теперь понятие несмещенного оценителя при $n > m$.

Теорема 7.1 (о несмещенном оценителе). Пусть $\|\cdot\|$ — некоторая векторная норма. Оценитель

$$\hat{\boldsymbol{\theta}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{f}(\mathbf{x})\| \quad (7.2)$$

является несмещённым.

Доказательство. Если в (7.1) положить $\boldsymbol{\varepsilon} = \mathbf{0}$, то минимум в (7.2) достигается при $\mathbf{x} = \boldsymbol{\theta}$ — в этом случае норма равна нулю, а при $\mathbf{x} \neq \boldsymbol{\theta}$ норма положительна. Поэтому для рассмотренного оценителя при $\boldsymbol{\varepsilon} = \mathbf{0}$ имеем $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}$, что и доказывает его несмещённость.

Рассмотрим два вида нормы.

1. Пусть норма евклидова, т. е.

$$\|\mathbf{x}\| = \|\mathbf{x}\|_2 \doteq \sqrt{\mathbf{x}'\mathbf{x}} = \sqrt{\sum_{i=1}^m x_i^2},$$

где x_i — компоненты вектора $\mathbf{x} \in \mathbb{R}^m$. Тогда оценитель (7.2) имеет вид

$$\hat{\boldsymbol{\theta}}_{\text{МНК}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{f}(\mathbf{x})\|_2 \quad (7.3)$$

и называется *методом наименьших квадратов (МНК)*. Оценку $\hat{\boldsymbol{\theta}}_{\text{МНК}}$ вектора $\boldsymbol{\theta}$, полученную с использованием оценителя (7.3), будем называть оценкой метода наименьших квадратов или, короче, *оценкой наименьших квадратов*. В более общем случае, когда норма определяется выражением

$$\|\mathbf{x}\| = \|\mathbf{x}\|_w \doteq \sqrt{\mathbf{x}'\mathbf{W}\mathbf{x}}$$

для некоторой положительно определенной матрицы $\mathbf{W}$, которую далее будем называть *весовой*, оценитель имеет вид

$$\hat{\boldsymbol{\theta}}_{\text{МНК-}w} = \arg \min_{\mathbf{x}} (\mathbf{y} - \mathbf{f}(\mathbf{x}))' \mathbf{W} (\mathbf{y} - \mathbf{f}(\mathbf{x}))$$

и называется *обобщенным методом наименьших квадратов (ОМНК)*. Соответствующую оценку $\hat{\boldsymbol{\theta}}_{\text{МНК-}w}$ будем называть *оценкой обобщенного метода наименьших квадратов* или, короче, *оценкой ОМНК*.

2. Пусть норма имеет вид

$$\|\mathbf{x}\| = \|\mathbf{x}\|_1 \doteq \sum_{i=1}^m |x_i|.$$

Тогда оценитель (7.2) имеет вид

$$\hat{\boldsymbol{\theta}}_{\text{МНК}} = \arg \min_{\mathbf{x}} \sum_{i=1}^n |y_i - f_i(\mathbf{x})|$$

и называется *методом наименьших модулей (МНМ)*, а оценка $\hat{\boldsymbol{\theta}}_{\text{МНК}}$ — *оценкой наименьших модулей*.

Качество оценки характеризуется величиной ее ошибки $\|\hat{\theta} - \theta\|$, а точнее — статистическими характеристиками этой ошибки. Обычно эти характеристики удается найти лишь для линейной модели измерений.

7.3. ЛИНЕАРИЗАЦИЯ МОДЕЛИ ИЗМЕРЕНИЙ

Пусть θ^0 — некоторое номинальное значение вектора θ . Предполагая разность $\theta - \theta^0$ достаточно малой, проведем линеаризацию модели измерений, оставляя только слагаемые первого порядка. Тогда она запишется в виде

$$y - f(\theta^0) = H'(\theta - \theta^0) + \varepsilon,$$

где

$$H' = \left. \frac{\partial f(\theta)}{\partial \theta} \right|_{\theta^0}$$

— матрица размерности $n \times m$. Окончательно линеаризованная модель записывается в виде

$$\tilde{y} = H'\theta + \varepsilon,$$

где $\tilde{y} = y - f(\theta^0) + H'\theta^0$ — приведенный вектор измерений. Эта приближенная модель обычно используется для вычисления оценки $\hat{\theta} = \theta^1$, относительно которой модель снова линеаризуется. Процесс последовательных приближений продолжается до тех пор, пока оценки θ^s и θ^{s+1} не окажутся достаточно близкими друг к другу. Тогда оценка $\hat{\theta} = \theta^{s+1}$ принимается за оценку в первоначальной нелинейной модели.

7.4. ОДНОМЕРНАЯ ЛИНЕЙНАЯ МОДЕЛЬ

Одномерная линейная модель имеет вид

$$y_i = \theta + \varepsilon_i, \quad i = 1, \dots, n,$$

где θ — неизвестный скалярный параметр.

Можно показать, что в этом случае оценка наименьших квадратов представляет собой выборочное среднее:

$$\hat{\theta} = \bar{y} \doteq \frac{1}{n} \sum_{i=1}^n y_i,$$

а оценка наименьших модулей — выборочную медиану вектора $\mathbf{y} = (y_1, y_2, \dots, y_n)'$.

Из приведенных формул следует, что оценка наименьших квадратов не является устойчивой к аномальным измерениям в отличие от оценки наименьших модулей.

7.5. ЛИНЕЙНЫЙ НЕСМЕЩЕННЫЙ АЛГОРИТМ И МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

Будем рассматривать только линейную модель оценивания. Нас будет интересовать значение некоторого скалярного параметра l , являющегося линейной функцией вектора $\boldsymbol{\theta}$:

$$l = \mathbf{b}'\boldsymbol{\theta}, \quad \mathbf{b} \in \mathbb{R}^m$$

(далее параметр l будем называть *контролируемым*), а также совокупности подобных параметров

$$\mathbf{l} = \mathbf{B}'\boldsymbol{\theta},$$

где $\mathbf{l} \in \mathbb{R}^s$ — *вектор контролируемых параметров*; $\mathbf{B}$ — матрица размерности $m \times s$; $\mathbf{b}$ и $\mathbf{B}$ известны. Если по результатам измерений найдена оценка $\hat{\boldsymbol{\theta}}$ вектора $\boldsymbol{\theta}$, то оценки указанных параметров можно найти по формулам

$$\hat{l} = \mathbf{b}'\hat{\boldsymbol{\theta}}, \quad \hat{\mathbf{l}} = \mathbf{B}'\hat{\boldsymbol{\theta}}.$$

Для линейной модели оценка ОМНК, соответствующая весовой матрице $\mathbf{W}$, имеет вид

$$\hat{\boldsymbol{\theta}} = \arg \min_{\mathbf{x}} [(\mathbf{y} - \mathbf{H}'\mathbf{x})'\mathbf{W}(\mathbf{y} - \mathbf{H}'\mathbf{x})] = \mathbf{Y}'\mathbf{y},$$

где

$$\mathbf{Y}' \doteq (\mathbf{H}\mathbf{W}\mathbf{H}')^{-1}\mathbf{H}\mathbf{W}.$$

Как видим, если выполняется условие $\det(\mathbf{H}\mathbf{W}\mathbf{H}') \neq 0$, то ОМНК представляет собой линейный относительно вектора $\mathbf{y}$ несмещенный алгоритм (доказательство несмещенности предоставляется читателю). При этом

$$\hat{\mathbf{l}} = \mathbf{B}'\hat{\boldsymbol{\theta}} = \mathbf{X}'\mathbf{y}, \quad \mathbf{X}' = \mathbf{B}'\mathbf{Y}' = \mathbf{B}'(\mathbf{H}\mathbf{W}\mathbf{H}')^{-1}\mathbf{H}\mathbf{W}.$$

Отвлечемся от метода наименьших квадратов и рассмотрим сразу линейную оценку вектора $\mathbf{l}$:

$$\hat{\mathbf{l}} = \mathbf{X}'\mathbf{y},$$

где $\mathbf{X}$ — матрица размерности $n \times s$ определяющая эту оценку. Условие несмещённости линейного алгоритма имеет вид

$$\hat{\mathbf{l}} = \mathbf{l} \Leftrightarrow \mathbf{HX} = \mathbf{B}.$$

Тогда ошибка оценки будет равна

$$\delta \mathbf{l} \doteq \hat{\mathbf{l}} - \mathbf{l} = \mathbf{X}'\boldsymbol{\varepsilon}.$$

Таким образом, ОМНК дает линейную несмещенную оценку любого векторного параметра $\mathbf{l} = \mathbf{B}'\boldsymbol{\theta}$ (а значит, и любого скалярного параметра $l = \mathbf{b}'\boldsymbol{\theta}$). Точнее, каждой весовой матрице $\mathbf{W}$ однозначно соответствует одна матрица $\mathbf{X}$ для линейного оценивания вектора $\mathbf{l} = \mathbf{B}'\boldsymbol{\theta}$ (или один вектор $\mathbf{x}$ для оценивания параметра $l = \mathbf{b}'\boldsymbol{\theta}$).

Рассмотрим соотношение между оценкой ОМНК и линейной несмещенной оценкой с другой точки зрения. Пусть известна матрица $\mathbf{X}$ размерности $n \times s$, удовлетворяющая условию несмещённости

$$\mathbf{HX} = \mathbf{B}$$

(в скалярном виде это условие имеет упомянутый выше вид $\mathbf{Hx} = \mathbf{b}$). Как найти весовую матрицу $\mathbf{W}$ такую, чтобы оценка ОМНК вектора $\mathbf{l} = \mathbf{B}'\boldsymbol{\theta}$, соответствующая этой матрице, имела вид $\hat{\mathbf{l}} = \mathbf{X}'\mathbf{y}$? Эта задача эквивалентна задаче нахождения неотрицательно определенной матрицы $\mathbf{W}$, являющейся решением матричного уравнения

$$\mathbf{X}' = \mathbf{B}'(\mathbf{H}\mathbf{W}\mathbf{H}')^{-1}\mathbf{H}\mathbf{W}.$$

Такая задача рассматривалась различными авторами (см., например, [Рао, 1975], [Бахшиян, 1983]). Было показано, что существует целое множество весовых матриц, для каждой из которых оценка наименьших квадратов совпадает с заданной несмещенной оценкой. В явном виде это множество в настоящем курсе мы приводить не будем.

Таким образом, на основе предыдущих рассуждений можно сделать вывод о том, что множество всех линейных несмещенных оценок любого векторного параметра $\mathbf{l} = \mathbf{B}'\boldsymbol{\theta}$ совпадает с множеством всех оценок ОМНК, получающихся при всевозможных весовых матрицах. Поэтому в дальнейшем будем анализировать только линейные несмещенные алгоритмы.

Замечание 7.1. Рассмотренный выше метод наименьших модулей не является линейным несмещенным оценителем. Поэтому он, вообще говоря, не эквивалентен никакой оценке ОМНК.

7.6. ВЫЧИСЛЕНИЕ ТОЧНОСТИ ОЦЕНОК ПРИ ИЗВЕСТНОЙ КОВАРИАЦИОННОЙ МАТРИЦЕ ОШИБОК ИЗМЕРЕНИЙ. ТЕОРЕМА ГАУССА – МАРКОВА

Пусть вектор ошибок ϵ представляет собой случайный вектор с математическим ожиданием $E(\epsilon)$ равным нулевому вектору, и известной матрицей ковариаций $K \doteq E(\epsilon\epsilon')$. Тогда дисперсия ошибки оценки $\hat{l} = x'y$ скалярного параметра $l = b'\theta$ равна

$$D(\delta l) = x'Kx,$$

где $\delta l = l - \hat{l}$. Данная формула является универсальной, так как она справедлива для любого оценителя x и используется для вычисления не только дисперсии оценок скалярных параметров, но и для определения ковариационной матрицы оценок векторных параметров. В этом случае возникает вопрос о нахождении оптимального оценителя, для которого дисперсия оценки будет минимальной. Можно показать, что минимум дисперсии по оценителю x при условии несмещённости $Hx = b$ достигается при

$$x' = b'(HK^{-1}H')^{-1}HK^{-1}.$$

Это устанавливается с помощью метода множителей Лагранжа. Можно заметить, что оценитель получился такой же, как если бы использовался ОМНК с весовой матрицей $W = K^{-1}$. Таким образом, с точки зрения минимизации дисперсии оценки любого скалярного параметра оптимальным является ОМНК с весовой матрицей, обратной матрице ковариаций ошибок измерений. Оптимальная дисперсия при этом равна

$$D^* = b'K_{\theta W}b,$$

где

$$K_{\theta W} \doteq (HK^{-1}H')^{-1}$$

— ковариационная матрица оценки вектора θ .

Этот известный результат называется *теоремой Гаусса**–Маркова**.

Замечание 7.2. Можно показать, что матрица

$$\mathbf{K}_{\hat{\theta}} - \mathbf{K}_{\hat{\theta}|W}$$

является неотрицательно определенной, если $\mathbf{K}_{\hat{\theta}}$ — ковариационная матрица оценки $\hat{\theta}$, полученной любым линейным несмещенным алгоритмом.

* Карл Фридрих Гаусс (1777–1855) — немецкий математик, механик, физик и астроном. Считается одним из величайших математиков всех времён, «королём математиков». Число научных открытий Гаусса настолько велико, что ему приходилось вести для их записи специальный дневник.

Гаусс впервые доказал основную теорему алгебры, явился создателем метода наименьших квадратов, исследовал нормальное распределение ошибок, положил начало дифференциальной геометрии, создал основы теории потенциала, доказал основную теорему электростатики. Им же был изобретён электрический телеграф. В трактате Гаусса «Теория движения небесных тел» изложена каноническая теория учёта возмущений орбит.

Гаусс знал в совершенстве множество языков, в 62 года стал изучать русский, чтобы прочитать труды Лобачевского. До старости привык большую часть вычислений производить в уме.

Именем Гаусса названо множество теорем и научных терминов, в том числе единица измерения магнитной индукции и одна из фундаментальных астрономических постоянных. Имя Гаусса носят малая планета и вулкан в Антарктиде.

** Андрей Андреевич Марков (1856–1922) — русский математик, академик, внёсший большой вклад в развитие теории вероятностей, математического анализа и теории чисел. С его именем связаны, в частности, такие понятия теории случайных процессов как *марковский процесс*, *цепь Маркова*, *неравенство Маркова*.

8. КЛАССИЧЕСКИЕ ЗАДАЧИ ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА

8.1. УСРЕДНЕНИЕ МОДЕЛИ ИЗМЕРЕНИЙ

Пусть $\theta \in \mathbb{R}^m$ — вектор неизвестных параметров системы, и при каждом значении $i=1, \dots, n$ проводится r_i измерений y_{ij} , $j=1, \dots, r_i$ некоторой скалярной функции $\mathbf{h}_i'\theta$, где $\mathbf{h}_i \in \mathbb{R}^m$ — известный вектор. Обычно считают, что измерения y_{ij} при заданном i проводятся в заданный момент времени t_i (возможен и другой параметр привязки всех измерений y_{ij} при заданном i , например дальность полета при движении точки по траектории). Тогда можно измерения в этот момент считать сеансом измерения. Например, космические измерения проводятся сеансами, соответствующими небольшим интервалам времени, в течение которых есть радиовидимость космического объекта. Эти интервалы часто можно при планировании космического эксперимента заменять отдельными моментами t_i .

Предполагается, что ошибки измерений являются некоррелированными между собой случайными величинами с нулевыми математическими ожиданиями и единичными дисперсиями. Согласно теореме Гаусса — Маркова, в классе линейных несмещенных алгоритмов оценивания минимальную дисперсию оценки любого параметра $l = \mathbf{b}'\theta$ дает МНК, соответствующий весовой матрице $\mathbf{W} = \mathbf{I}$. Учитывая это, осредним модель оценивания.

Лемма 8.1. *При использовании МНК исходная модель измерений*

$$y_{ij} = \mathbf{h}_i'\theta + \varepsilon_{ij}, \quad i=1, \dots, n, \quad j=1, \dots, r_i$$

эквивалентна усредненной модели

$$y_i = \mathbf{h}_i'\theta + \varepsilon_i, \quad i=1, \dots, n, \quad (8.1)$$

где

$$y_i = \frac{1}{r_i} \sum_{j=1}^{r_i} y_{ij}, \quad \varepsilon_i = \frac{1}{r_i} \sum_{j=1}^{r_i} \varepsilon_{ij}$$

есть средние арифметические указанных выше r_i измеренных значений y_{ij} и ошибок ε_{ij} . Эквивалентность данных моделей означает, что оценки параметра θ , получаемые при использовании этих моделей, совпадают. При этом усредненные ошибки ε_i являются некоррелированными между собой случайными величинами с нулевыми математическими ожиданиями и дисперсиями $D\varepsilon_i = \frac{1}{r_i}$.

Для доказательства леммы, согласно теореме Гаусса – Маркова, запишем дисперсию оценки МНК некоторого скалярного параметра $l = \mathbf{b}'\theta$, которая соответствует весовой матрице $\mathbf{W} = \mathbf{I}$, как минимальную дисперсию в классе всех линейных несмещенных алгоритмов. Получаем

$$\hat{l}_{\text{МНК}} = \sum_{i=1}^n \sum_{j=1}^{r_i} x_{ij} y_{ij},$$

где коэффициенты x_{ij} оценщика находятся из условия минимума дисперсии оценки при условиях несмещённости:

$$D\hat{l}_{\text{МНК}} = \min_{x_i} \left\{ \sum_{i=1}^n \sum_{j=1}^{r_i} x_{ij}^2 : \sum_{i=1}^n \sum_{j=1}^{r_i} x_{ij} \mathbf{h}_i = \mathbf{b} \right\}. \quad (8.2)$$

Обозначим

$$x_i = \sum_{j=1}^{r_i} x_{ij}. \quad (8.3)$$

Минимум по x_{ij} суммы $\sum_{j=1}^{r_i} x_{ij}^2$ в (8.2) при условии (8.3) и фиксированных x_i достигается при $x_{ij} = \frac{x_i}{r_i}$ для всех j , т.е. все повторяющиеся измерения в оптимальном случае оцениваются с одним оценщиком (это ясно и из соображений симметрии). Тогда (8.2) переписывается в виде

$$D\hat{l}_{\text{МНК}} = \min_{x_i} \left\{ \sum_{i=1}^n \frac{x_i^2}{r_i} : \sum_{i=1}^n x_i \mathbf{h}_i = \mathbf{b} \right\}. \quad (8.4)$$

Согласно теореме Гаусса – Маркова равенство (8.4) доказывает лемму 8.1, так как линейный оценщик с коэффициентами x_i соответствует модели измерений (8.1).

Замечание 8.1. Так как в (8.4) суммирование записано по всем i , то, чтобы исключить суммирование по тем элементам, где $r_i = 0$, целевая функция должна быть доопределена следующим образом:

$$\frac{x_i^2}{r_i} = \begin{cases} 0, & \text{если } x_i = 0 \text{ и } r_i = 0, \\ +\infty, & \text{если } x_i \neq 0 \text{ и } r_i = 0. \end{cases}$$

8.2. ПОСТАНОВКА ЗАДАЧИ

Пусть задано общее число измерений:

$$\sum_{i=1}^n r_i = N.$$

Вместо чисел r_i будем использовать вектор $\mathbf{p}$ с координатами

$$p_i = \frac{r_i}{N}, \quad i = 1, \dots, n.$$

Далее будем, как это обычно делают, пренебрегать целочисленностью r_i и считать, что $\mathbf{p}$ — непрерывный план эксперимента, т. е. любой вектор из симплекса

$$\Sigma_n = \left\{ \mathbf{p} \in \mathbb{R}^n : p_i \geq 0, \sum_{i=1}^n p_i = 1 \right\}.$$

Пусть $\mathbf{b}_j \in \mathbb{R}^m, j = 1, \dots, s$ — заданные векторы, а $l_j = \mathbf{b}_j' \boldsymbol{\theta}$ — контролируемые параметры (обычно $s \leq m$). При заданном плане $\mathbf{p}$ рассмотрим линейные оценки каждого из этих параметров:

$$\hat{l}_j = \sum_{i \in \mathcal{I}_p} \Phi_{ij} y_i, \quad \mathcal{I}_p \doteq \{i : p_i > 0\}, \quad (8.5)$$

где числа $\Phi_{ij}, i = 1, \dots, n$ определяют линейный оценщик параметра l_j . Этот оценщик дает несмещенные оценки параметров $\mathbf{b}_j' \boldsymbol{\theta}$ тогда и только тогда, когда он удовлетворяет условиям несмещенности

$$\sum_{i \in \mathcal{I}_p} \mathbf{h}_i \Phi_{ij} = \mathbf{b}_j, \quad j=1, \dots, s, \quad (8.6)$$

т.е. $\mathbf{b}_j$ есть линейная комбинация тех векторов $\mathbf{h}_i$, для которых $p_i > 0$. Далее будем обозначать символом $\mathcal{P}_n$ множество тех $\mathbf{p}$ из Σ_n , для которых уравнения (8.6) разрешимы относительно Φ_{ij} .

Для каждого $\mathbf{p} \in \mathcal{P}_n$ рассмотрим наилучшие линейные несмещенные оценки, для которых оценитель определяется из условия минимальных дисперсий при линейном несмещенном оценивании:

$$ND\hat{l}_j = \min_{\Phi_{ij}} \left\{ \sum_{i \in \mathcal{I}_p} \frac{\Phi_{ij}^2}{p_i} : \sum_{i \in \mathcal{I}_p} \mathbf{h}_i \Phi_{ij} = \mathbf{b}_j \right\}, \quad j=1, \dots, s. \quad (8.7)$$

Задача оптимального планирования эксперимента состоит в нахождении плана $\mathbf{p} \in \mathcal{P}_n$, доставляющего нижнюю грань заданному критерию оптимальности $L(\mathbf{p})$ [Бахшиян, Соловьев, 1998; Бахшиян, 2012]:

$$L = \inf \{ L(\mathbf{p}) : \mathbf{p} \in \mathcal{P}_n \}. \quad (8.8)$$

Будем рассматривать два критерия оптимальности, предполагая далее, что оценки (8.5) являются наилучшими нелинейными несмещенными. Первый критерий

$$L(\mathbf{p}) = \left(N \sum_{j=1}^s D\hat{l}_j \right)^{1/2} \quad (8.9)$$

называется критерием *L-оптимальности*. Второй критерий

$$L(\mathbf{p}) = \left(N \max \{ D\hat{l}_1, \dots, D\hat{l}_s \} \right)^{1/2} \quad (8.10)$$

будем называть *MV_s-критерием* (при $s=m$ и $l_j = \theta_j$, $j=1, \dots, m$, критерий (8.10) называют просто *MV-критерием*). В зависимости от критерия (8.9) или (8.10) задачу (8.8) будем соответственно называть *L-задачей* или *MV_s-задачей*.

При решении задачи (8.8) будем предполагать, что

$$\text{rang}(\mathbf{h}'_1, \dots, \mathbf{h}'_n) = m \quad (8.11)$$

(т.е. ранг составной матрицы максимален), что равносильно допущению о возможности линейного несмещенного оценивания всех компонент вектора $\boldsymbol{\theta}$ в случае использования всех n групп измерений.

Замечание 8.2. Нетрудно показать, что суммирование по $i \in \mathcal{I}_p$ в соотношениях (8.5)–(8.7) можно заменить суммированием по $i = 1, \dots, n$, если доопределить слагаемые в (8.7) при $p_i = 0$ следующим образом:

$$\frac{\Phi_{ij}^2}{p_i} = \begin{cases} 0, & \text{если } \Phi_{ij} = 0 \text{ и } p_i = 0, \\ +\infty, & \text{если } \Phi_{ij} \neq 0 \text{ и } p_i = 0. \end{cases}$$

В дальнейшем будем считать указанные замены при суммировании и оптимизации произведенными.

Замечание 8.3. Согласно теореме Гаусса – Маркова, при выполнении условия (8.6) наилучшая нелинейная несмещенная оценка совпадает с оценкой взвешенного метода наименьших квадратов $\hat{l}_j = \mathbf{b}'_j \hat{\boldsymbol{\theta}}$, где $\hat{\boldsymbol{\theta}}$ — любое решение нормальных уравнений

$$\mathbf{M}(\mathbf{p})\hat{\boldsymbol{\theta}} = \sum_{i=1}^n p_i \mathbf{h}_i y_i, \quad (8.12)$$

а

$$\mathbf{M}(\mathbf{p}) = \sum_{i=1}^n p_i \mathbf{h}_i \mathbf{h}'_i$$

— нормированная на число N информационная матрица. Эти уравнения всегда совместны (доказательство этого факта предоставляется читателю), но решение $\hat{\boldsymbol{\theta}}$ будет единственно только в случае невырожденной матрицы $\mathbf{M}(\mathbf{p})$. План $\mathbf{p}$ в этом случае называется *невырожденным*. В противном случае план $\mathbf{p}$ является *вырожденным*, и решение уравнений (8.12) неоднозначно.

Однако условие (8.6) эквивалентно тому, что оценка $\mathbf{b}'_j \hat{\boldsymbol{\theta}}$ каждой функции $\mathbf{b}'_j \boldsymbol{\theta}$, $j = 1, \dots, s$, одна и та же для всех решений уравнения (8.12). При этом параметры $\mathbf{b}'_j \boldsymbol{\theta}$ называются *оцениваемыми*. Таким образом, условия несмещенности и оцениваемости эквивалентны. Далее, условия несмещенности (8.6) можно трактовать как принадлежность векторов $\mathbf{b}_j$ линейному многообразию, порожденному векторами $\sqrt{p_1} \cdot \mathbf{h}_1, \dots, \sqrt{p_n} \cdot \mathbf{h}_n$ или, эквивалентно, столбцами информационной матрицы. Это означает, что найдутся векторы $\boldsymbol{\lambda}_j \in \mathbb{R}^m$ такие, что выполняется условие оцениваемости параметров $\mathbf{b}'_j \boldsymbol{\theta}$:

$$\mathbf{M}(\mathbf{p})\boldsymbol{\lambda}_j = \mathbf{b}_j, \quad j = 1, \dots, s. \quad (8.13)$$

8.3. СКАЛЯРНАЯ ЗАДАЧА ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА И АЛГОРИТМ ЕЁ РЕШЕНИЯ

При $s=1$ L -задача (а также MV_s -задача) сводится к минимизации дисперсии $D\hat{l}$, где $l=l_1$ и называется *C-задачей* планирования эксперимента. Согласно (8.4), C -задача записывается в виде

$$ND_0^* = \min_{p_i, x_i} \left\{ \sum_{i=1}^n \frac{x_i^2}{p_i} : \sum_{i=1}^n x_i \mathbf{h}_i = \mathbf{b}, \mathbf{x} \in \mathbb{R}^n, \mathbf{p} \in \Sigma_n \right\}. \quad (8.14)$$

Проведем далее минимизацию дисперсии по p_i , пренебрегая тем, что r_i должно быть целым. Это оправдано при достаточно большом N . Получим, используя метод множителей Лагранжа, что оптимальный план измерений есть

$$p_i = \frac{|x_i|}{\sum_{k=1}^n |x_k|}.$$

При этом для заданного оценителя с коэффициентами x_i

$$ND_0 = \frac{1}{N} \left(\sum_{i=1}^n |x_i| \right)^2. \quad (8.15)$$

Замечание 8.4. Указанная оптимизация по r_i здесь проведена не совсем строго (так же как и в основополагающей работе [Elfving, 1952]). Действительно, мы проводили оптимизацию по p_i только при условии $p_1 + \dots + p_n = 1$. Однако, условие несмещённости может выполняться не при всех таких p_i , т.е. оценка параметра l может быть получена не всегда. В связи с этим условие несмещённости также называют *условием оцениваемости*. Поэтому множество допустимых p_i может быть несколько меньше, а искомый минимум — не меньше. Однако, ввиду того, что получившееся решение приводит к допустимым оценкам, оно является верным.

Оптимизация величины D_0 , полученной при оптимальных r_i , по оценителю $\mathbf{x}$ приводит к задаче линейного программирования

$$\sqrt{N} \cdot \sigma_0^* \doteq \sqrt{D_1^*} = \min_{x_i} \left\{ \sum_{i=1}^n |x_i| : \sum_{i=1}^n x_i \mathbf{h}_i = \mathbf{b} \right\}.$$

Введем новые переменные $z_i = |x_i|$, $z_i \geq 0$. Тогда $x_i \mathbf{h}_i = z_i \mathbf{a}_i$, где

$$\mathbf{a}_i = \begin{cases} \mathbf{h}_i, & x_i \geq 0, \\ -\mathbf{h}_i, & x_i < 0. \end{cases}$$

При этом оптимальную задачу можно записать в виде

$$\sigma_1^* = \min_{z_i, \mathbf{a}_i} \left\{ \sum_{i=1}^n z_i : \sum_{i=1}^n z_i \mathbf{a}_i = \mathbf{b}, z_i \geq 0, \mathbf{a}_i \in \mathcal{H}_i \right\}, \quad (8.16)$$

где $\mathcal{H}_i = \{\mathbf{h}_i, -\mathbf{h}_i\}$ — множества, состоящие из двух векторов (рис. 14).

Сформулированную задачу можно интерпретировать как задачу линейного программирования, состоящую в минимизации суммы коэффициентов $z_i > 0$ разложения вектора $\mathbf{b}$ по векторам $\pm \mathbf{h}_i$. Геометрический способ ее решения (в двумерном случае) обсуждался в пп. 1.3.4. У этой задачи существует решение, содержащее не более m отличных от нуля коэффициентов, т.е. вектор $\mathbf{b}$ лежит внутри многогранного угла, образованного некоторыми m векторами $\pm \mathbf{h}_i$ (базисом), причем гиперплоскость, проведенная через их концы, отсекает от вектора $\mathbf{b}$ наименьшую (считая от конца вектора) часть по сравнению с другими комбинациями из m векторов. Легко заметить, что указанная гиперплоскость, проходящая через концы векторов оптимального базиса, принадлежит выпуклой оболочке всех $2n$ векторов $\pm \mathbf{h}_i$, $i = 1, \dots, n$ (см. рис. 2).

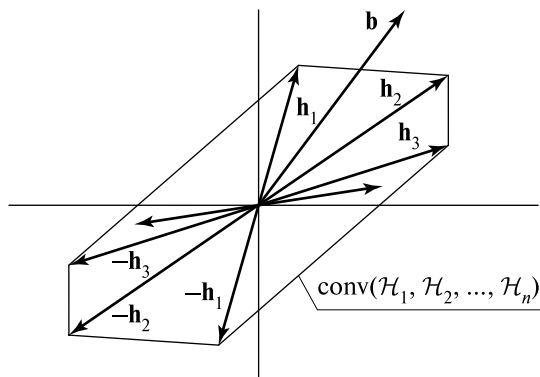


Рис. 14. Построение выпуклой оболочки $\text{conv}(\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_n)$ множеств $\mathcal{H}_i$ при решении задачи (8.16)

Таким образом, можно сделать следующий важный вывод. При решении скалярной задачи планирования эксперимента следует выбрать оптимальные m измерений и найти оценщик для усредненной модели из соотношения

$$\mathbf{H}^0 \mathbf{x}^0 = \mathbf{b} \Rightarrow \mathbf{x}^0 = (\mathbf{H}^0)^{-1} \mathbf{b},$$

где $\mathbf{H}^0$ — матрица размерности $m \times m$, составленная из векторов $\mathbf{h}_i$, входящих в оптимальный базис. Это соответствует тому, что вектор $\boldsymbol{\theta}$ находится из m уравнений

$$y_i = \mathbf{h}_i' \boldsymbol{\theta}, \quad i = 1, \dots, m,$$

где нумерация по i условна. После этого оптимальные значения величин r_i находятся из условия их пропорциональности величинам x_i .

8.4. ОЦЕНИВАНИЕ ПАРАМЕТРОВ ПАРАБОЛИЧЕСКОЙ ТРАЕКТОРИИ ПО ИЗМЕРЕНИЯМ ДАЛЬНОСТИ

Рассмотрим в качестве примера плоское движение материальной точки, брошенной под углом α к горизонту с начальной скоростью v_0 . Введем следующие обозначения: $\mathbf{r}$, $\mathbf{v}$ — векторы положения и скорости точки в момент времени t ; $\mathbf{r}_0$, $\mathbf{v}_0$ — векторы положения и скорости в начальный момент времени $t_0 = 0$;

$\mathbf{x}(t) = \begin{pmatrix} \mathbf{r}(t) \\ \mathbf{v}(t) \end{pmatrix}$ — вектор состояния.

Движение точки описывается радиус-вектором

$$\mathbf{r}(t) = \begin{pmatrix} r_x(t) \\ r_y(t) \end{pmatrix},$$

где

$$r_x(t) = v_{x_0} t,$$

$$r_y(t) = v_{y_0} t - \frac{1}{2} g t^2,$$

$v_{x_0} = |v_0| \cos \alpha$, $v_{y_0} = |v_0| \sin \alpha$, $t \in [0, T]$, $\frac{T}{2} = \frac{v_{y_0}}{g}$ — время подъёма на максимальную высоту. Пусть на интервале $[0, t_k]$ могут производиться измерения дальности $l = |\mathbf{r}(t)|$ до движущейся

точки. Ошибки дальности будем считать случайными величинами с нулевым математическим ожиданием и стандартным отклонением $\sigma_l = 1$ м.

В качестве вектора оцениваемых параметров рассмотрим начальный вектор состояния $\boldsymbol{\theta} = \mathbf{x}(0)$. При этом будем считать, что положение точки в начальный момент известно, а именно $r_x(0) = 0$, $r_y(0) = 0$, т.е. число оцениваемых параметров равно двум (оцениваются компоненты вектора $\mathbf{v}_0$). В качестве контролируемого параметра выберем дальность до точки в момент падения.

Требуется:

- решить задачу оптимального оценивания дальности на интервале $[0, t_k]$, т.е. найти оптимальные значения моментов измерений t_1, t_2 и соответствующую среднеквадратичную ошибку оценки $\sigma_1 = \sqrt{D_1}$;
- доказать, что при любом t_k справедливо $t_2 = t_k$;
- интерпретировать результат как C -задачу планирования эксперимента, найти соответствующие доли p_1, p_2 общего числа измерений в моменты времени t_1 и t_2 .

Запишем линеаризованную модель измерений в виде

$$l(t) = \mathbf{h}'_r \mathbf{r}_0 + \mathbf{h}'_v \mathbf{v}_0 + \varepsilon(t) = \mathbf{h}' \boldsymbol{\theta} + \varepsilon(t),$$

где $\mathbf{h}' \doteq (\mathbf{h}'_r, \mathbf{h}'_v)$ — составной вектор-строка.

Нахождение оптимальных моментов измерений, согласно изложенной выше теории, сводится в данном случае к решению двумерной задачи линейного программирования (используется только вектор $\mathbf{h}_v(t)$)

$$\min_{x_i} \left\{ \sum_{i=0}^k |x_i| : \sum_{i=0}^k x_i \mathbf{h}_v(t_i) = \mathbf{b} \right\}, \quad (8.17)$$

где $\mathbf{b} = (T, 0)'$. Будем искать решение этой задачи геометрически в соответствии с подходом, изложенным в пп. 1.3.4. При этом для простоты и наглядности будем пренебрегать дискретностью моментов измерений, рассматривая вместо моментов t_i непрерывную переменную t .

Компоненты вектора $\mathbf{h}_v(t)$ вычисляются из соотношений

$$\mathbf{h}'_v(t) = \frac{\partial l(t)}{\partial \boldsymbol{\theta}} \bigg|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0} = \frac{\partial l(t)}{\partial \mathbf{r}(t)} \frac{\partial \mathbf{r}(t)}{\partial \boldsymbol{\theta}} \bigg|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0} + \frac{\partial l(t)}{\partial \mathbf{v}(t)} \frac{\partial \mathbf{v}(t)}{\partial \boldsymbol{\theta}} \bigg|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0}.$$

Производные измеряемой функции $l(t)$ вычисляются по формулам

$$\left(\frac{\partial l(t)}{\partial \mathbf{r}(t)} \right)' = \frac{\mathbf{r}'}{|\mathbf{r}|}, \quad \left(\frac{\partial l(t)}{\partial \mathbf{v}(t)} \right) = \mathbf{0}.$$

Окончательно векторы $\mathbf{h}_v(t)$ для каждого t будут иметь вид

$$\mathbf{h}_v(t) = \left[\frac{r_x(t)t}{\sqrt{r_x^2(t) + r_y^2(t)}}, \frac{r_y(t)t}{\sqrt{r_x^2(t) + r_y^2(t)}} \right]', \quad (8.18)$$

где $r_x(t) = v_{x_0} t$, $r_y(t) = v_{y_0} t - \frac{1}{2} g t^2$.

Годограф функции $\mathbf{h}_v(t)$ имеет вид, показанный на рис. 15.

Для нахождения решения задачи (8.17) необходимо построить линейную оболочку векторов $\mathbf{h}_v(t)$ и получить разложение вектора $\mathbf{b}$ по базису из этих векторов с минимальной суммой коэффициентов. Моменты времени, соответствующие векторам оптимального базиса, будут искомыми моментами измерений. Построение линейной оболочки и оптимальный базис показаны на рис. 16. Нетрудно заметить, что при любом значении $t_k \in [0, T]$ вектор $\mathbf{h}_v(t_k)$ входит в состав оптимального базиса, т. е. выполняется условие $t_2 = t_k$.

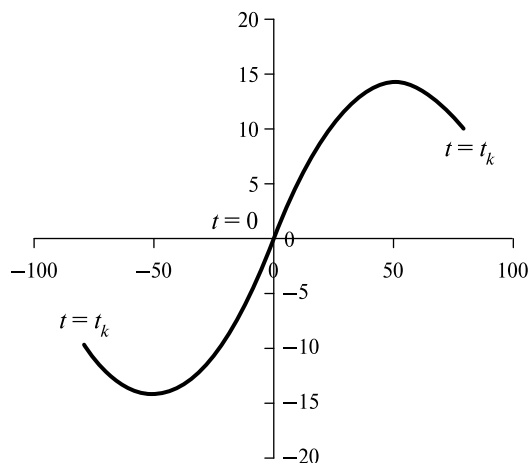


Рис. 15. Годограф функции $\mathbf{h}_v(t)$ (см. формулу (8.18))

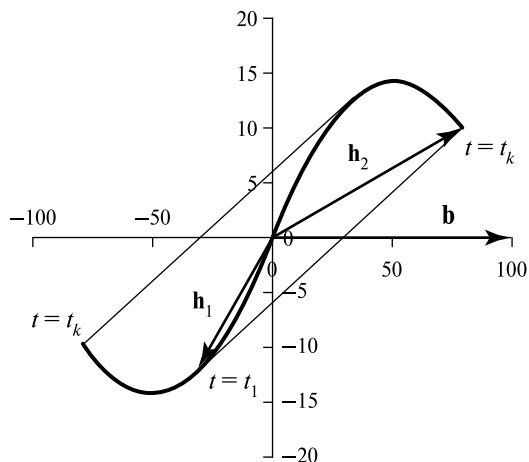


Рис. 16. Построение оптимального базиса в задаче (8.17)

Момент времени t_1 может быть найден из условия касания годографа и прямой, проходящей через его крайнюю точку, соответствующую моменту времени t_k . После этого из (1.12) определяются x_1 и x_2 , а по ним — p_1 и p_2 . Значение среднеквадратической ошибки, согласно (8.16), равно оптимальному значению целевой функции задачи (8.17).

Возьмем, например, $t_k = 80$ с. Для этого случая найдем: $t_2 = t_k = 80$ с, $t_1 = 38$ с, $x_1 = -1,49$, $x_2 = 1,96$, $p_1 = 0,43$, $p_2 = 0,57$, $\sigma_1 = 3,46$.

9. ЗАДАЧА ОПТИМАЛЬНОЙ ЛИНЕЙНОЙ ИМПУЛЬСНОЙ КОРРЕКЦИИ И АЛГОРИТМ ЕЁ РЕШЕНИЯ

Определение 9.1. *Коррекцией движения* будем называть задачу исправления траектории движения космического аппарата по результатам измерений и составления на их основе прогноза движения.

Будем предполагать, что коррекция траектории может производиться при помощи импульсных воздействий бортовых двигателей, позволяющих мгновенно изменять вектор скорости аппарата. Такой вид коррекции принято называть *импульсной коррекцией*.

Нас будет интересовать задача определения оптимальной импульсной коррекции, т.е. ответ на вопрос, в какие моменты времени необходимо включить бортовые двигатели и какой должна быть величина корректирующих импульсов для того, чтобы цель коррекции (т.е. исправление траектории) была достигнута, и при этом затраты на проведение коррекции были минимальными.

9.1. ЗАДАЧА ОПТИМАЛЬНОЙ КОРРЕКЦИИ ПРИ ИЗВЕСТНОМ КОРРЕКТИРУЕМОМ ВЕКТОРЕ

Задача оптимальной коррекции при известном векторе корректируемых параметров имеет большой самостоятельный интерес и разнообразные практические приложения. Пусть $\mathbf{l} \in \mathbb{R}^s$ — некоторый вектор параметров системы, который может быть изменен путем импульсной коррекции ее траектории. Такая коррекция заключается в мгновенном изменении компонент вектора скорости в моменты времени $t_1, \dots, t_n$. Изменение в каждый момент t_i будем характеризовать корректирующим вектором (импульсом) $\mathbf{u}_i$ размерности s_i , принадлежащим евклидовому пространству $\mathbb{R}^{s_i}$. Например, в случае коррекции траектории

космического аппарата, если импульс $\mathbf{u}_i$ может производиться в произвольном направлении, то $s_i=3$, $\mathbb{R}^{s_i} = \mathbb{R}^3$ — трехмерное евклидово пространство; если же заранее задана плоскость или прямая, которым должен принадлежать импульс, то s_i равно двум или единице, а $\mathbb{R}^{s_i}$ есть плоскость или прямая. Будем рассматривать идеальную линейную коррекцию, т.е. предполагать, что импульсы производятся без ошибок и изменение вектора $\mathbf{l}$ в момент t_i под действием импульса $\mathbf{u}_i$ равно $\mathbf{U}_i \mathbf{u}_i$, где $\mathbf{U}_i$ — матрица размерности $s \times s_i$, характеризующая влияние импульса и называемая *матрицей влияния*. Тогда для изменения вектора $\mathbf{l}$ на векторную величину $\mathbf{b} \in \mathbb{R}^s$ нужно импульсы $\mathbf{u}_i$ произвольно выбирать из условия

$$\sum_{i=1}^n \mathbf{U}_i \mathbf{u}_i = \mathbf{b}.$$

Пусть затраты на коррекцию в момент t_i характеризуются величиной $p_i(\mathbf{u}_i)$, где $p_i(\cdot)$ — некоторая норма в $\mathbb{R}^{s_i}$. Например, в случае коррекции траектории космического аппарата путем ориентации в пространстве с использованием одного двигателя, жестко связанного с аппаратом, затраты на коррекцию малого импульса пропорциональны величине

$$p_i(\mathbf{u}_i) = \|\mathbf{u}_i\|_2 \doteq |\mathbf{u}_i| \doteq \sqrt{\mathbf{u}_i' \mathbf{u}_i} = \left(\sum_{j=1}^3 \mathbf{u}_{ij}^2 \right)^{1/2}.$$

Если же для коррекции используются двигатели, действующие вдоль трех осей координат, то

$$p_i(\mathbf{u}_i) = \|\mathbf{u}_i\|_1 = \sum_{j=1}^3 |\mathbf{u}_{ij}|,$$

где $\mathbf{u}_{ij}$ — составляющие вектора $\mathbf{u}_i$ вдоль направления действия двигателей.

При сделанных допущениях поставим задачу оптимизации суммарных затрат на коррекцию. Эта задача имеет вид

$$L = \min_{\mathbf{u}_i} \left\{ \sum_{i=1}^n p_i(\mathbf{u}_i) : \sum_{i=1}^n \mathbf{U}_i \mathbf{u}_i = \mathbf{b} \right\}.$$

Решение данной задачи определяет и оптимальные моменты приложения импульса, так как, если $\mathbf{u}_i = 0$, то в момент t_i импульс не проводится. Представим каждый вектор $\mathbf{u}_i$ в виде

$$\mathbf{u}_i = x_i \gamma_i, \quad x_i \doteq p_i(\mathbf{u}_i) \geq 0,$$

где $\gamma_i \in \mathbb{R}^{s_i}$ принадлежит множеству $\{\gamma_i : p_i(\gamma_i) = 1\}$. Тем самым, x_i есть длина вектора $\mathbf{u}_i$ в метрике, определяемой нормой $p_i(\cdot)$, а γ_i есть единичный вектор в этой метрике, направленный вдоль импульса.

Используя введенные новые переменные, запишем оптимальную задачу коррекции в виде

$$L = \min_{x_i, \mathbf{a}_i} \left\{ \sum_{i=1}^n x_i : \sum_{i=1}^n x_i \mathbf{a}_i = \mathbf{b}, x_i \geq 0, \mathbf{a}_i \in \mathcal{A}_i, i = 1, \dots, n \right\},$$

где $\mathcal{A}_i = \{\mathbf{a}_i : \mathbf{a}_i = \mathbf{U}_i \gamma_i, p_i(\gamma_i) = 1, \gamma_i \in \mathbb{R}^{s_i}\}$. Эта задача по виду напоминает задачу линейного программирования, однако здесь переменными оптимизации являются не только числа x_i , но и векторы $\mathbf{a}_i$ в линейных условиях-равенствах, выбираемые из множеств $\mathcal{A}_i$. Такая задача называется *обобщенной задачей линейного программирования* [Данциг, 1966, Бахшиян и др., 1980]. Ее можно также назвать *многопараметрической задачей линейного программирования*. Приведение задачи коррекции к обобщенной задаче линейного программирования и алгоритм ее решения были предложены М.Л. Лидовым* [Лидов, 1971]. Фактически этот алгоритм представляет собой реализацию так называемого *метода генерации столбцов* в обобщенном линейном программировании. В этой же статье, а также в книге [Бахшиян и др., 1980] рассмотрены более общие постановки задач коррекции. Далее изложение следует этим работам.

Геометрический способ решения обобщенной задачи линейного программирования аналогичен рассмотренному выше геометрическому способу решения обычной задачи линейного программирования. Для случая, когда нормы p_i являются евклидовыми, множества $\mathcal{A}_i$ представляют собой линейное преобразование сфер $\|\gamma_i\| = 1$, т.е. являются эллипсоидами. Такой случай изображен на рис. 17. Если вектор $\mathbf{b}$ находится внутри многогранного угла, образованного некоторыми k векторами $\mathbf{a}_i \in \mathcal{A}_i, i = 1, \dots, k$, то векторы $\mathbf{a}_i, i = 1, \dots, k$ дают решение задачи в следующем смысле (см. рис. 17).

* Михаил Львович Лидов (1926–1993) — советский и российский учёный в области прикладной небесной механики, известен результатами в теории эволюции орбит и в космической баллистике.

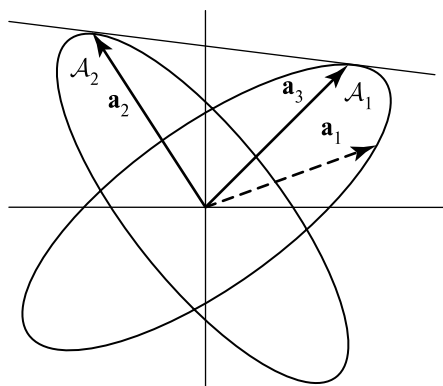


Рис. 17. Геометрический способ решения оптимальной задачи коррекции

Нужно провести $k \leq s$ корректирующих импульсов, величина каждого импульса равна x_j , направление определяется вектором γ_j . При этом случай $k < s$ уже не является исключительным. Например, если вектор $\mathbf{b}$ совпадает по направлению с вектором $\mathbf{a}_i$, принадлежащим выпуклой оболочке $\mathcal{A}_i$, то $k = 1$. Симплексный алгоритм также сводится к вводу в базис вектора $\mathbf{a}_s$, который пересекает гиперплоскость, проходящую через концы векторов базиса. Проверка достаточного условия оптимальности, таким образом, сводится к проверке условия непересечения этой гиперплоскости любым из бесконечного множества векторов $\mathbf{a}_i \in \mathcal{A}_i$, $i = 1, \dots, n$. Если это условие проверяется эффективно, то алгоритм симплекс-метода также эффективен. Для того чтобы понять, насколько эффективен этот измененный для обобщенной задачи линейного программирования симплексный алгоритм, запишем достаточное условие оптимальности базиса в формульном виде. При этом рассмотрим сначала случай, когда нормы p_i являются евклидовыми. Именно к такому случаю сводится L -задача оптимального планирования эксперимента.

Пусть $\mathbf{B} = (\mathbf{a}_1, \dots, \mathbf{a}_s)$ — матрица базиса на текущей итерации, а вектор $\boldsymbol{\pi}$ однозначно находится из системы уравнений

$$\boldsymbol{\pi}'\mathbf{B} = (1, \dots, 1) \Leftrightarrow \Delta_j = 1 - \boldsymbol{\pi}'\mathbf{a}_j = 0, \quad j = 1, \dots, s. \quad (9.1)$$

Таким образом, относительные разности Δ_j для векторов базиса равны нулю (см. шаг 2 алгоритма, приведенного в пп. 1.3.2).

Равны они нулю и для любого вектора, конец которого находится на гиперплоскости, проходящей через концы векторов базиса. Выше этой гиперплоскости (высота измеряется от начала координат) относительные разности отрицательны, а ниже — положительные.

Достаточное условие оптимальности базиса состоит в том, что все векторы из множеств $\mathcal{A}_i$ лежат ниже указанной гиперплоскости. Математически это записывается в виде равенства нулю минимума относительной разности на множестве всех векторов $\mathbf{a} \in \mathcal{A}_i, i = 1, \dots, n$,

$$\Delta_{\min} \doteq \min_i \min_{\mathbf{a} \in \mathcal{A}_i} (1 - \boldsymbol{\pi}' \mathbf{a}) = 0. \quad (9.2)$$

Из (9.2) получаем

$$\Delta_{\min} = 1 - \max_i \max_{\mathbf{a} \in \mathcal{A}_i} \boldsymbol{\pi}' \mathbf{a} = 1 - \max_i \beta_i. \quad (9.3)$$

Здесь

$$\beta_i = \max_{\mathbf{a} \in \mathcal{A}_i} \boldsymbol{\pi}' \mathbf{a} = \max_{\|\boldsymbol{\gamma}\|=1} \boldsymbol{\pi}' \mathbf{U}_i' \boldsymbol{\gamma} = \|\mathbf{U}_i' \boldsymbol{\pi}\|, \quad (9.4)$$

и максимум реализуется при $\boldsymbol{\gamma} = \frac{\mathbf{U}_i' \boldsymbol{\pi}}{\|\mathbf{U}_i' \boldsymbol{\pi}\|}$.

Если условие (9.2) не выполняется, т. е. существует номер k такой, что $\beta_k = \max_i \beta_i > 1$, то в базис вводится любой вектор $\mathbf{a}_k \in \mathcal{A}_k$, для которого не выполняется условие (9.2). Вывод вектора из базиса осуществляется по обычным правилам симплекс-метода.

Замечание 9.1. Если в рассмотренной задаче коррекции все импульсы одномерные, т. е. $s_i = 1$ для всех i , то эта задача сводится к задаче линейного программирования, которая возникает при рассмотрении C -задачи планирования эксперимента.

Замечание 9.2. Если полученный указанным алгоритмом оптимальный базис содержит не более одного вектора с ненулевым весом x_i из каждого множества $\mathcal{A}_i$, то этот базис дает решение задачи оптимальной коррекции. Данное условие может не выполняться только в случае, когда некоторое множество $\mathcal{A}_i$ содержит плоскость, проходящую через совокупность $\mathbf{a}_{ij}, j = 1, \dots, r_i$ векторов из $\mathcal{A}_i$, имеющих ненулевые веса x_{ij} (подробнее об этом см. также в статье [Бахшиян, 1989], где исследуется

обобщенная задача линейного программирования). Но в этом случае один из векторов указанной совокупности векторов заменяется некоторой их выпуклой комбинацией — вектором

$$\mathbf{a}_i = \sum_{j=1}^{r_i} \lambda_j \mathbf{a}_{ij},$$

где

$$\lambda_j = \frac{x_j}{\sum_{l=1}^{r_i} x_{il}}.$$

После этого множество $\mathcal{A}_i$ будет содержать лишь один вектор $\mathbf{a}_i$ с ненулевым весом, т. е. новый базис дает решение задачи оптимальной коррекции.

Таким образом, для представляющих практический интерес случаев решение оптимальной задачи коррекции осуществляется эффективным алгоритмом, который представляет собой модификацию симплекс-метода, состоящую в особом правиле нахождения вводимого в базис вектора, выбираемого из совокупности множеств $\mathcal{A}_i$. Поэтому данный метод и получил название «метод генерации столбцов». Этот метод эквивалентен по сложности алгоритму решения обычной задачи линейного программирования, размерность которой (т. е. число линейных уравнений в ограничениях) равна s . При этом приходится оперировать с базисной матрицей размерности $s \times s$.

Замечание 9.3. В отличие от обычной задачи линейного программирования, в обобщенной задаче линейного программирования изложенный алгоритм требует, вообще говоря, бесконечного числа итераций (см. рис. 17). Однако за конечное число итераций удастся найти приближенное решение с заданной точностью с использованием на каждой итерации оценки для оптимума

$$\frac{\sum_{i=1}^n x_i}{\beta} \leq L \leq \sum_{i=1}^n x_i,$$

где $\beta \doteq \max\{\beta_1, \dots, \beta_n\}$. Данная оценка является следствием соотношения (1.22).

9.2. ПРОЕКТНАЯ ЗАДАЧА ИДЕАЛЬНОЙ КОРРЕКЦИИ

Пусть вектор требуемой величины коррекции не задан, а известно лишь, что он лежит в заданной области:

$$\mathbf{b} \in \mathbb{B}.$$

Такой случай имеет место, например, когда при расчетах траектории движения летательного аппарата заранее, еще до запуска, необходимо рассчитать возможный расход топлива. Естественно поставить задачу нахождения максимального значения суммарного импульса на оптимальную коррекцию при условии, что корректируемый вектор может принимать любые значения из $\mathbb{B}$. Таким образом отыскивается

$$L_{\mathbb{B}} = \max_{\mathbf{b} \in \mathbb{B}} \min_{\mathbf{u}_i} \left\{ \sum_{i=1}^n p_i(\mathbf{u}_i) : \sum_{i=1}^n \mathbf{U}_i \mathbf{u}_i = \mathbf{b} \right\}. \quad (9.5)$$

Можно показать, что если множество $\mathbb{B}$ представляет собой выпуклый многогранник, то максимум в (9.5) достигается в одной из вершин этого многогранника.

9.3. КОРРЕКЦИЯ ПАРАБОЛИЧЕСКОЙ ТРАЕКТОРИИ ЛЕТАТЕЛЬНОГО АППАРАТА

Рассмотрим задачу коррекции траектории летательного аппарата (ЛА), движущегося в гравитационном поле по параболической орбите в некоторой плоскости с известной начальной скоростью. Такое движение может рассматриваться как частный случай движения материальной точки, брошенной под углом α к горизонту (см. п. 8.4). Рассмотрим для данного случая задачу идеальной коррекции. Предположим, что коррекция может производиться либо в произвольном, либо в заданном направлении с помощью одного двигателя, жестко связанного с ЛА.

Пусть известна начальная скорость ЛА $\mathbf{v}_0 = (v_{x_0}, v_{y_0})'$, и проводятся измерения двух параметров движения — дальности и высоты полета (т.е. вектор контролируемых параметров $\mathbf{l}$ двумерный). Допустим, что в момент времени T ЛА должен прийти в требуемую точку, которая является целью полета, однако из-за ошибок исполнения начального импульса реальная траектория ЛА отличается от требуемой и заканчивается в какой-то другой точке. Рассмотрим вектор $\mathbf{b}$, компоненты

которого равны разности контролируемых параметров ЛА для требуемой и реальной траекторий в конечный момент времени (т.е. представляют собой требуемую величину коррекции по каждому из параметров).

Необходимо решить аналитически задачу коррекции параболической траектории ЛА, обеспечивающей достижение заданной цели полета, т.е. попадания ЛА в нужную точку.

Пусть $t_i, i=1, \dots, n$ — моменты времени, в которые может быть осуществлена коррекция, $\mathbf{u}_i$ — корректирующий импульс в момент t_i , $p(\mathbf{u}_i) = \|\mathbf{u}_i\|$ — затраты на проведение импульса в момент t_i . Тогда задача оптимизации суммарных затрат на коррекцию имеет вид

$$L = \min_{\mathbf{u}_i} \left\{ \sum_{i=1}^n \|\mathbf{u}_i\| : \sum_{i=1}^n \mathbf{U}_i \mathbf{u}_i = \mathbf{b} \right\},$$

где $\mathbf{U}_i$ — матрица влияния импульса, произведенного в момент t_i , на вектор контролируемых параметров $\mathbf{l}$.

Представим $\mathbf{u}_i$ в виде $\mathbf{u}_i = x_i \gamma_i$, где $x_i \doteq \|\mathbf{u}_i\|$ — длина вектора импульса, определяемая нормой $\|\cdot\|$; γ_i — единичный вектор, определяющий направление импульса.

В результате приходим к следующей задаче:

$$L = \min_{x_i, \gamma_i} \left\{ \sum_{i=1}^n x_i : \sum_{i=1}^n x_i \mathbf{U}_i \gamma_i = \mathbf{b}, x_i \geq 0, i=1, \dots, n \right\}.$$

Введем векторы $\mathbf{a}_i = \mathbf{U}_i \gamma_i$, $\mathbf{a}_i \in \mathcal{A}_i = \{\mathbf{U}_i \gamma_i : \gamma_i \in \mathbb{R}^2, \|\gamma_i\| = 1\}$.

Тогда задача оптимизации принимает вид

$$L = \min_{x_i, \mathbf{a}_i} \left\{ \sum_{i=1}^n x_i : \sum_{i=1}^n x_i \mathbf{a}_i = \mathbf{b}, x_i \geq 0, \mathbf{a}_i \in \mathcal{A}_i, i=1, \dots, n \right\}.$$

Уравнения движения ЛА в произвольный момент времени t записываются в виде

$$\begin{cases} r_x(t) = v_{x_0} t, \\ r_y(t) = v_{y_0} t - \frac{gt^2}{2}, \\ v_x(t) = v_{x_0}, \\ v_y(t) = v_{y_0} - gt. \end{cases}$$

Вектор $\mathbf{b}$ определяется как

$$\mathbf{b} = \begin{bmatrix} \overset{\circ}{r}_x(T) - r_x(T) \\ \overset{\circ}{r}_y(T) - r_y(T) \end{bmatrix},$$

где T — время полета ЛА, $\overset{\circ}{r}_x(T)$, $\overset{\circ}{r}_y(T)$ — требуемые значения координат в момент времени T .

Запишем уравнения движения ЛА в момент времени T через значения координат и скорости в текущий момент времени t :

$$\begin{cases} r_x(T) = r_x(t) + v_x(t)(T-t), \\ r_y(T) = r_y(t) + v_y(t)(T-t) - \frac{g(T-t)^2}{2}, \\ v_x(T) = v_x(t), \\ v_y(T) = v_y(t) - g(T-t). \end{cases} \quad (9.6)$$

Рассмотрим два случая, когда импульсы имеют произвольное направление в плоскости движения либо направлены вдоль вектора скорости.

1. *Импульсы направлены произвольно.* В данной задаче зависимость $\mathbf{r}(T)$ от $\mathbf{v}(t)$ нелинейна, поэтому нахождение матрицы влияния сопровождается линеаризацией модели:

$$\mathbf{U}(t) = \frac{\partial \mathbf{r}(T)}{\partial \mathbf{v}(t)} = \begin{bmatrix} \frac{\partial r_x(T)}{\partial v_x(t)} & \frac{\partial r_x(T)}{\partial v_y(t)} \\ \frac{\partial r_y(T)}{\partial v_x(t)} & \frac{\partial r_y(T)}{\partial v_y(t)} \end{bmatrix} = \begin{bmatrix} T-t & 0 \\ 0 & T-t \end{bmatrix} = (T-t) \mathbf{E}_2,$$

где $\mathbf{E}_2$ — единичная матрица размерности 2×2 , $\mathbf{r}(t) = (r_x(t), r_y(t))'$, $\mathbf{v}(t) = (v_x(t), v_y(t))'$.

Тогда в дискретном случае

$$\mathbf{U}_i = \mathbf{U}(t_i) = (T-t_i) \mathbf{E}_2,$$

множества $\mathcal{A}_i$ имеют вид

$$\mathcal{A}_i = \left\{ (T-t_i) \mathbf{v}_i : \mathbf{v}_i \in \mathbb{R}^2, \|\mathbf{v}_i\| = 1 \right\},$$

т.е. при каждом i концы векторов из множества $\mathcal{A}_i$ лежат на окружности радиуса $(T-t_i)$. Если коррекция может проводиться непрерывно в произвольные моменты времени из некоторого

интервала $[t_1, t_n]$, то система множеств $\mathcal{A}_i$ переходит в систему множеств $\mathcal{A}(t)$, $t_1 \leq t \leq t_n$.

На рис. 18 показан геометрический способ нахождения решения. Несложно заметить, что оптимальным будет приложение одного корректирующего импульса в направлении вектора $\mathbf{b}$ в момент времени $t=t_1$.

Найдем величину этого импульса:

$$x_1 \cdot \mathbf{a}_1 = x_1 \cdot (T - t_1) \gamma_1 = \mathbf{b},$$

откуда

$$x_1 = \|\mathbf{u}_1\| = \frac{\|\mathbf{b}\|}{T - t_1}.$$

2. *Импульсы направлены вдоль вектора скорости.* В этом случае, как и в предыдущем, корректирующие импульсы $\mathbf{u}_i$ будут иметь размерность $k_i = 2$. Представим вектор скорости в виде произведения

$$\mathbf{v}(t) = |\mathbf{v}(t)| \cdot \boldsymbol{\tau}(t),$$

где $|\mathbf{v}(t)|$ — длина вектора скорости; $\boldsymbol{\tau}(t) = \frac{\mathbf{v}(t)}{|\mathbf{v}(t)|}$ — единичный вектор, определяющий направление вектора скорости.

Из (9.6) следует, что

$$\mathbf{r}(T) = \mathbf{r}(t) + \mathbf{v}(t)(T - t) - \frac{1}{2}g(T - t)^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

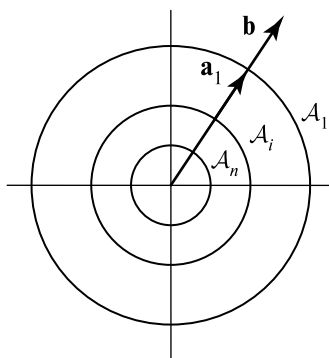


Рис. 18. Геометрический способ решения задачи коррекции

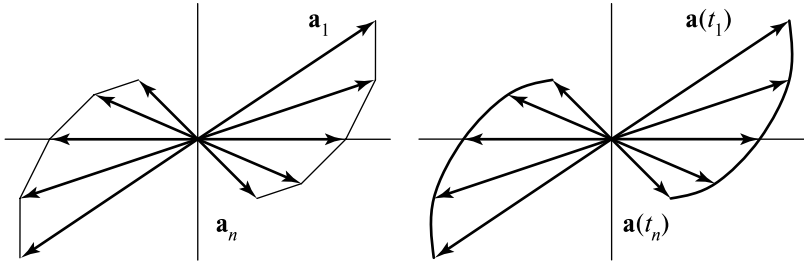


Рис. 19. Годограф множеств $\mathcal{A}_i$ (слева, коррекция проводится в заданные дискретные моменты времени) и множеств $\mathcal{A}(t)$ (справа, коррекция проводится в произвольные моменты времени)

Тогда матрица влияния определяется следующим образом:

$$\begin{aligned} \mathbf{U}(t) &= \frac{\partial \mathbf{r}(T)}{\partial |\mathbf{v}(t)|} = \frac{\partial}{\partial |\mathbf{v}(t)|} \left[\mathbf{r}(t) + \mathbf{v}(t)(T-t) - \frac{1}{2} g(T-t)^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] = \\ &= \frac{\partial}{\partial |\mathbf{v}(t)|} \left[\mathbf{r}(t) + |\mathbf{v}(t)| \boldsymbol{\tau}(t)(T-t) - \frac{1}{2} g(T-t)^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] = (T-t) \boldsymbol{\tau}(t). \end{aligned}$$

В дискретном случае отсюда следует, что

$$\mathbf{U}_i = \mathbf{U}(t_i) = (T-t_i) \boldsymbol{\tau}_i,$$

а множества $\mathcal{A}_i$ имеют вид

$$\mathcal{A}_i = \left\{ (T-t_i) \boldsymbol{\tau}_i \gamma_i : \gamma_i \in \mathbb{R}^1, \gamma_i = \pm 1 \right\},$$

или

$$\mathbf{A}_i = \left\{ (T-t_i) \boldsymbol{\tau}_i, -(T-t_i) \boldsymbol{\tau}_i \right\},$$

т.е. каждое из множеств $\mathcal{A}_i$ состоит из двух векторов. На рис. 19 показаны годографы этих множеств в случае, если коррекция может производиться в заданные дискретные моменты времени, и множеств $\mathcal{A}(t)$ при $t_1 \leq t \leq t_n$, если коррекция может производиться в любой момент времени.

Возможны следующие случаи расположения вектора $\mathbf{b}$.

1. Вектор $\mathbf{b}$ лежит в области годографа множеств $\mathcal{A}_i$ (рис. 20). В этом случае оптимальным решением будет приложение двух корректирующих импульсов (коррекция производится

в заданные дискретные моменты времени (см. рис. 20, слева), оптимальные моменты приложения импульсов — $t=t_1$ и $t=t_2$, импульсы производятся в положительном направлении, т.е. при $\gamma=+1$). Если же коррекция может проводиться в любой момент времени (см. рис. 20, справа), то оптимальным будет приложение одного корректирующего импульса, направление которого определяется вектором $\mathbf{a}(t^*)$.

2. Вектор $\mathbf{b}$ лежит вне области годографа множеств $\mathcal{A}_i$ (рис. 21). В случае, изображенном на левой части рис. 21, для коррекции ЛА необходимо приложить два корректирующих импульса, один — в направлении вектора $\mathbf{a}_1$, т.е. в направлении вектора скорости при $t=t_1$, а второй — в направлении вектора $-\mathbf{a}_4$, т.е. в момент времени $t=t_4$ в направлении, противоположном направлению вектора скорости.

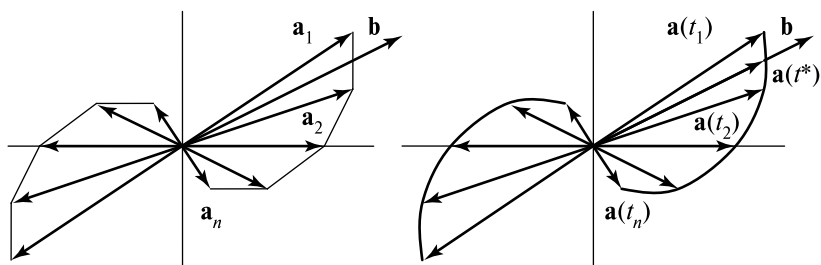


Рис. 20. Построение решения в задаче оптимальной коррекции для дискретного (слева) и непрерывного (справа) случаев: вектор $\mathbf{b}$ лежит в области годографа

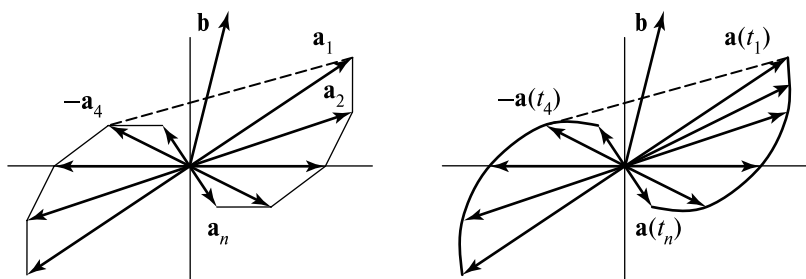


Рис. 21. Построение оптимального решения в задаче коррекции для дискретного (слева) и непрерывного (справа) случаев: вектор $\mathbf{b}$ лежит вне области годографа

В случае, изображенном на правой части рис. 21, для коррекции ЛА также необходимо приложить два корректирующих импульса. Моменты времени и направление импульсов определяются аналогично.

10. ГАРАНТИРОВАННЫЕ ХАРАКТЕРИСТИКИ ТОЧНОСТИ

10.1. КРИТИКА КЛАССИЧЕСКОГО ПОДХОДА К ОЦЕНИВАНИЮ ТОЧНОСТИ

При довольно естественных условиях, налагаемых на матрицу плана измерений и ковариационную матрицу ошибок, дисперсия оценки Гаусса – Маркова стремится к нулю с увеличением числа измерений n [Эльясберг, 1976]. Однако на практике такой картины не наблюдается. Простой пример приведен в книге [Эльясберг, 1983]. Пусть измеряются показания различных часов одной точности и находится среднее арифметическое всех таких показаний. Такая оценка есть оценка МНК, которая совпадает с оценкой Гаусса – Маркова при отсутствии корреляции между наблюдениями и одинаковой дисперсии. При увеличении числа показаний должно было бы получиться сколь угодно точное время, но на практике этого не произойдет, так как существует корреляция между ошибками в точности хода, обусловленная преемственностью способов изготовления часов, и эта корреляция может меняться в зависимости от времени их изготовления. Поэтому классический подход к вычислению точности оценивания может давать слишком оптимистические результаты.

10.2. ГАРАНТИРУЮЩИЙ ПОДХОД К ВЫЧИСЛЕНИЮ ТОЧНОСТИ ОЦЕНИВАНИЯ

Этот подход может дать, наоборот, более пессимистические прогнозы точности, но для практики иногда важно «перестраховаться». Поясним сущность этого подхода на рассмотренном выше примере вычисления дисперсии оценки. Пусть известно

лишь множество $\mathcal{K}$, которому принадлежит ковариационная матрица ошибок измерений. Элементы $\mathbf{K}$ этого множества подчиняются условию неотрицательной определенности, которое будем записывать в виде $\mathbf{K} \geq 0$. Тогда в наихудшем с точки зрения величины дисперсии случае получим *гарантированную дисперсию*

$$D_{\max} = \max_{\mathbf{K} \in \mathcal{K}} D(\delta \mathbf{l}).$$

Отметим, что гарантированная дисперсия вычисляется при заданном *оценителе* $\mathbf{x}$, т.е., например, при заданной весовой матрице, которая уже не может выбираться в соответствии с теоремой Гаусса – Маркова, так как не известна ковариационная матрица ошибок. Вопрос об оптимизации оценителя рассмотрим ниже. Сейчас выпишем $D_{\max}$ для простых случаев задания множества $\mathcal{K}$. Допустим, что множество $\mathcal{K}$ определяется следующим образом:

$$\mathcal{K} = \left\{ \mathbf{K} : \mathbf{K} \geq 0, |k_{ij}| \leq k \leq 1, (i \neq j), k_{ii} = 1 \right\}, \quad (10.1)$$

где k_{ij} — элементы матрицы $\mathbf{K}$, представляющие собой коэффициенты корреляции между ошибками измерений, число k задано, а дисперсии ошибок k_{ij} известны и для простоты записи приняты далее единицами. Тогда

$$D_{\max} \leq D_k,$$

где D_k вычисляется без учета условия $\mathbf{K} \geq 0$. Ниже будет показано, что на самом деле это условие при вычислении гарантированной дисперсии можно не учитывать. Тогда

$$\begin{aligned} D_{\max} \leq D_k &\doteq \max_{k_{ij}} \left\{ \mathbf{x}' \mathbf{K} \mathbf{x} = \sum_{i=1}^n x_i^2 + \sum_{i=1}^n \sum_{j \neq i}^n x_i x_j k_{ij} : |k_{ij}| \leq k \right\} = \\ &= (1-k) \sum_{i=1}^n x_i^2 + k \sum_{i=1}^n x_i^2 + \max_{k_{ij}} \left\{ \sum_{i=1}^n \sum_{j \neq i}^n x_i x_j k_{ij} : |k_{ij}| \leq k \right\} \end{aligned}$$

или, окончательно,

$$D_{\max} \leq D_k = (1-k)D_0 + kD_1, \quad (10.2)$$

где $D_0 = \sum_i x_i^2$ — дисперсия оценки наименьших квадратов при некоррелированных измерениях (это самый оптимистический

случай), $D_1 = \left(\sum_i |x_i| \right)^2$ — гарантированная дисперсия при условии, что корреляция может быть произвольной (самый пессимистический случай), т.е. формально при условии $|k_{ij}| \leq 1$, которое не является ограничением на коэффициенты корреляции, так как выполняется всегда.

Как указывалось в п. 10.1, в неэкзотических случаях $D_0 \rightarrow 0$ при увеличении числа измерений. Поэтому при большом числе измерений в этих случаях $D_{\max} \approx kD_1$. Дисперсии D_0 и D_1 удовлетворяют неравенству Коши* — Шварца** — Буняковского***

* Огюстен Луи Коши (1789–1857) — великий французский математик и механик. Разработал фундамент математического анализа, впервые дав строгое определение таким основным понятиям анализа как *предел*, *непрерывность*, *производная*, *сходимость ряда* и т.д. Внёс огромный вклад в математическую физику, рассмотрев решение краевой задачи с известными начальными условиями (*задача Коши*). Один из основоположников механики сплошных сред. Занимался также алгеброй, теорией чисел, астрономией, оптикой и многими другими разделами физики и математики. Написал свыше 800 работ, полное собрание его сочинений состоит из 27 томов. Член Парижской академии наук, Лондонского королевского общества, Петербургской академии наук и других академий.

** Карл Герман Амандус Шварц (1843–1921) — немецкий математик, член Берлинской академии наук, профессор Галльского, Цюрихского, Гёттингенского и Берлинского университетов. Известен работами по геометрии, теории чисел.

*** Виктор Яковлевич Буняковский (1804–1889) — русский математик, вице-президент Академии наук (1864–1889). Основные труды посвящены теории вероятностей и теории чисел. Активно занимался педагогической деятельностью. Его труды помимо выдающейся научной значимости отличались удивительной ясностью изложения, многие из них были переведены на иностранные языки. Одна из самых известных работ — трактат «Основания математической теории вероятностей», содержащий не только теорию, но и историю возникновения и развития этой науки, обсуждение многих практических приложений.

Пользовался огромным уважением современников как выдающийся ученый и блистательный педагог, лекции которого отличались ясностью, увлекательностью и изяществом изложения. Современники отмечали также его высокие нравственные качества. Любопытно, например, что за все годы преподавания Буняковский не пропустил ни одной своей лекции и ни разу не опоздал.

$$D_1 \leq nD_0.$$

Из изложенного в разделе 8 материала следует, что, если возможно повторение измерений, то при оптимальном плане измерений здесь будет наблюдаться равенство.

Отметим более общий случай [Бахшиян и др., 1980], когда коэффициенты корреляции принадлежат многомерному прямоугольному параллелепипеду, т.е. элементы k_{ij} удовлетворяют неравенствам

$$-w_{ij} \leq k_{ij} - k_{ij}^* \leq w_{ij}, \quad i, j = 1, \dots, n,$$

где k_{ij}^* — элементы некоторой известной номинальной корреляционной матрицы $\mathbf{K}^*$, а неотрицательные элементы w_{ij} задают разброс около $\mathbf{K}^*$ ($k_{ij} = 1$, $w_{ij} = 0$ при всех i). В этом случае

$$D_{\max} \leq D_w = \mathbf{x}' \mathbf{K}^* \mathbf{x} + \mathbf{x}'_+ \mathbf{W} \mathbf{x}_+,$$

где вектор $\mathbf{x}_+$ состоит из компонент $|x_i|$. Критерий достижимости равенства $D_{\max} = D_w$ приведен в книге [Бахшиян и др., 1980], но этот критерий является трудно проверяемым. Однако имеется простое достаточное условие достижимости равенства в последнем соотношении, которое имеет вид

$$\mathbf{K}^* - \mathbf{A} \geq \mathbf{O}, \quad \mathbf{W} + \mathbf{A} \geq \mathbf{O}$$

при некоторой диагональной матрице $\mathbf{A}$. В случае $\mathbf{K}^* = \mathbf{A} = \mathbf{E}$, где $\mathbf{E}$ — единичная матрица, получаем $\mathbf{W} + \mathbf{E} \geq \mathbf{O}$ и $D_{\max} = D_w$. Отсюда следует упомянутое выше равенство $D_{\max} = D_k$.

10.3. СРАВНЕНИЕ РЕШЕНИЙ ЗАДАЧ ОПТИМАЛЬНОГО ОЦЕНИВАНИЯ В ДВУХ ПРОСТЕЙШИХ СЛУЧАЯХ ПРИ ГАРАНТИРУЮЩЕМ И КЛАССИЧЕСКОМ ПОДХОДАХ

Покажем удивительное совпадение состава оптимальных измерений при двух различных подходах к модели ошибок. Рассмотрим сначала классический подход.

10.3.1. ОПТИМИЗАЦИЯ ГАРАНТИРОВАННОЙ ДИСПЕРСИИ D_1

Задача оптимизации алгоритма оценивания с точки зрения минимизации гарантированной дисперсии имеет вид

$$\min_x \{D_{\max} : \mathbf{H}\mathbf{x} = \mathbf{b}\}.$$

Выбирая различные скалярные параметры l_j ($j=1, \dots, k \leq m$), находим для них различные оптимальные векторы-оценители. Составляя из этих векторов матрицу оценителя $\mathbf{X}$ размерности $k \times m$, можно найти, согласно изложенному ранее, множество весовых матриц $\{\mathbf{W}\}$, так что каждой матрице соответствует МНК, дающий ту же оценку, что и оценитель $\mathbf{X}$. Рассмотрим задачу минимизации D_1 (напомним, что D_0 соответствует единичной корреляционной матрице и найдена в п. 10.2). Имеем, согласно (10.2),

$$\sigma_1^* = \min_x \left\{ \sigma_1 \doteq \sqrt{D_1} : \mathbf{H}\mathbf{x} = \mathbf{b} \right\} = \min_{x_i} \left\{ \sum_{i=1}^n |x_i| : \sum_{i=1}^n x_i \mathbf{h}_i = \mathbf{b} \right\},$$

и величина D_1 связана с величиной D_0 соотношением

$$D_0 = \frac{D_1}{n}$$

(см. п. 10.2).

Отсюда следует важный методологический вывод [Бахшиян, 1970]. В случае повторяющихся некоррелированных измерений имеется m оптимальных точек повторения измерения, которые совпадают с оптимальными точками проведения измерений ($x_i \neq 0$) для задачи минимизации гарантированной дисперсии D_1 , полученной при возможности произвольной корреляции между измерениями. При этом $D_1 = nD_0$, а число повторений измерений r_i пропорционально модулю коэффициента x_i , соответствующего i -му измерению в случае произвольной корреляции.

10.3.2. МИНИМАКСНАЯ ЗАДАЧА ОЦЕНИВАНИЯ ПРИ ОГРАНИЧЕННЫХ ПО МОДУЛЮ ОШИБКАХ ИЗМЕРЕНИЙ

Такая задача рассматривалась М.Л. Лидовым [Лидов, 1964] и названа им «схемой бортиков». Предполагается, что ошибки измерений ограничены по модулю:

$$|\varepsilon_i| \leq M_i,$$

где M_i — известные числа. Тогда гарантированная ошибка измерений равна

$$\delta l_{\max} = \max_{\varepsilon_i} \sum_{i=1}^n x_i \varepsilon_i = \sum_{i=1}^n M_i |x_i|,$$

а ее минимизация сводится к задаче линейного программирования того же типа, что и задачи в п. 8.3, 10.3.1:

$$\delta l^* = \min_{x_i} \left\{ \sum_{i=1}^n M_i |x_i| : \mathbf{H}\mathbf{x} = \mathbf{b} \right\}.$$

Рассмотрим частный случай, когда $M_i = M$ при всех i . Этот случай можно трактовать как допущение о пропорциональности максимально возможных ошибок измерений их дисперсиям, которые в п. 10.3.1 приравнены к единицам. Тогда задачи линейного программирования, решаемые при нахождении величин δl^* и σ_1^* (см. п. 10.3.1), дают один и тот же оптимальный оценитель, при этом

$$\delta l^* = M \sigma_1^*.$$

Отметим, что вместо условий на ошибки измерений можно было бы рассматривать аналогичные условия на их математические ожидания:

$$|E(\varepsilon_i)| \leq M_i.$$

При этом гарантированное математическое ожидание равно

$$E_{\max} = \max_i \left\{ |E(\delta l)| : |E(\varepsilon_i)| \leq M_i \right\} = \sum_{i=1}^n M_i |x_i|,$$

а задача минимизации $E_{\max}$ по линейному несмещенному оценителю приводит к той же по виду задаче линейного программирования, что и задача нахождения δl^* (см. выше).

Замечание 10.1. Можно показать, что решение указанной задачи линейного программирования дает оптимальный алгоритм оценивания в классе *всех* (как линейных, так и нелинейных) несмещённых алгоритмов в «схеме бортиков» [Матасов, 1988а, б]. Это позволяет более прагматично отнестись к казалось бы экзотической «схеме бортиков» и философски посмотреть на проблему использования всей информации в том случае, когда информация поставляется «игроком», сопротивляющимся получению хорошей оценки.

10.4. ОПТИМИЗАЦИЯ ГАРАНТИРОВАННЫХ ХАРАКТЕРИСТИК ТОЧНОСТИ МЕТОДОМ ГЕНЕРАЦИИ СТОЛБЦОВ

Рассмотренные выше оптимальные задачи оценивания сводятся либо к задачам линейного программирования, число ограничений которых равно размерности m вектора оцениваемых параметров, либо к задачам обобщенного линейного программирования с числом ограничений, пропорциональным m . Трудоемкость таких задач пропорциональна числу их ограничений. Более сложные практические задачи не могут быть сведены к таким задачам, эффективно решаемым при небольших m .

Рассмотрим общий подход, позволяющий решать достаточно широкий класс задач методом генерации столбцов, упомянутым в разделе 9. Недостатком этого подхода является то, что число ограничений решаемой обобщенной задачи линейного программирования равно $n + 1$, где n — число измерений. Это приводит к большой трудоемкости метода генерации столбцов при больших n .

Рассмотрим сначала задачу минимизации гарантированной ошибки оценивания в более общей постановке, чем ранее. Пусть задано симметричное относительно нуля выпуклое замкнутое ограниченное множество M возможных значений вектора ошибок измерений ϵ . Рассмотрим минимаксную задачу оценивания

$$\delta l^* = \min_x \max_{\epsilon} \{ \mathbf{x}'\epsilon : \mathbf{H}\mathbf{x} = \mathbf{b}, \epsilon \in M \}, \quad (10.3)$$

которая дает минимальное значение ошибки в наименее благоприятном случае. Задачу (10.3) можно переписать также в виде

$$\delta l^* = \min_x \{ f(\mathbf{x}) : \mathbf{H}\mathbf{x} = \mathbf{b} \}, \quad (10.4)$$

где

$$f(\mathbf{x}) = \delta l_{\max} = \max_{\epsilon} \{ \mathbf{x}'\epsilon : \epsilon \in M \}.$$

Будем при этом предполагать, что функция $f(\mathbf{x})$, которая равна гарантированной ошибке оценивания при заданном оценивателе $\mathbf{x}$, может быть вычислена либо аналитически, либо достаточно эффективно численными методами.

Переставим операции минимума и максимума (это возможно согласно теореме фон Неймана) и учтем, что, согласно теории двойственности,

$$\min_x \{ \mathbf{x}'\epsilon : \mathbf{H}\mathbf{x} = \mathbf{b} \} = \max_{\lambda} \{ \mathbf{b}'\lambda : \mathbf{H}'\lambda = \epsilon, \lambda \in \mathbb{R}^m \}.$$

Тогда получим

$$\delta l^* = \max_{\lambda} \{ \mathbf{b}'\lambda : \mathbf{H}'\lambda \in M \}. \quad (10.5)$$

Покажем, что задачу нахождения δl^* удастся решить эффективно методом генерации столбцов. Для этого в соответствии с теоремой Каратеодори представим каждый элемент выпуклого множества M как выпуклую комбинацию неизвестных $n+1$ точек $\boldsymbol{\varepsilon}^{(j)} \in M$, $j = 1, \dots, n+1$. Тогда выписанная выше задача максимизации по λ запишется в виде

$$\delta l^* = \max_{\lambda_i, \boldsymbol{\varepsilon}^{(j)}, \mu_j} \left\{ \mathbf{b}'\lambda : \lambda_i \geq 0, \boldsymbol{\varepsilon}^{(j)} \in M, \mu_j \geq 0 \right\} \quad (10.6)$$

при условиях

$$\sum_{i=1}^n \lambda_i \begin{pmatrix} \mathbf{h}_i \\ 0 \end{pmatrix} + \sum_{j=1}^{n+1} \mu_j \begin{pmatrix} \boldsymbol{\varepsilon}^{(j)} \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix},$$

где $\mathbf{h}_i$ — n -мерные столбцы матрицы $\mathbf{H}'$.

Это есть обобщенная задача линейного программирования, размерность которой (т.е. число уравнений в ограничениях) равна $n+1$. Неограниченные по знаку переменные λ_i можно всегда принимать базисными при решении обобщенной задачи линейного программирования методом генерации столбцов. Это следует из того, что, если соответствующие столбцы попадают в базис, то они оттуда не убираются при использовании метода генерации столбцов.

Пусть $\tilde{\boldsymbol{\pi}}' = (\boldsymbol{\pi}', \rho) \in \mathbb{R}^{n+1}$ — вектор множителей Лагранжа (ρ — скаляр), определяемый из условия (см. п. 1.3):

$$\tilde{\boldsymbol{\pi}}'\mathbf{B} = (\mathbf{b}', 0),$$

где $\mathbf{B}$ — базисная матрица размерности $(n+1) \times (n+1)$, соответствующая текущему базису. Тогда достаточным условием оптимальности текущего базиса является выполнение равенства вида (9.2):

$$-\min_{\boldsymbol{\varepsilon} \in M} (-\boldsymbol{\varepsilon}', 1) \tilde{\boldsymbol{\pi}} = 0 \Rightarrow \max_{\boldsymbol{\varepsilon} \in M} \boldsymbol{\pi}'\boldsymbol{\varepsilon} = \rho, \quad (10.7)$$

и в базис вводится вектор условий $(-\boldsymbol{\varepsilon}', 1)$, для которого $\boldsymbol{\pi}'\boldsymbol{\varepsilon} > \rho$.

Таким образом, для проверки условия оптимальности и введения в базис вектора условий нужно уметь решать подзадачу

$$f(\lambda) = \max_{\varepsilon} \{ \lambda' \varepsilon : \varepsilon \in M \},$$

которая по сделанному выше предположению о функции $f(x)$ решается эффективно (например, аналитически). Поэтому вычисление величины δl^* методом генерации столбцов для задачи (10.6) эффективно. К тому же, согласно (1.23), имеется оценка близости текущего значения $b' \lambda$ целевой функции к оптимуму:

$$b' \lambda^* \leq b' \lambda + \max_{\varepsilon \in M} \pi' \varepsilon - \rho.$$

Можно показать, что скаляр ρ равен оптимальному значению δl^* , а в качестве оптимального оценителя x можно взять вектор π , удовлетворяющий достаточным условиям оптимальности. Это следует из того, что двойственный вектор к двойственной задаче (10.6) является вектором оптимизации в прямой задаче (10.3).

ЛИТЕРАТУРА

- [Авдеев и др., 1965] *Авдеев Ю. Ф., Беляков А. И.* и др. Полет космических аппаратов. Примеры и задачи / Под ред. Г. С. Титова. М.: Машиностроение, 1980.
- [Балк, 1965] *Балк М. Б.* Элементы динамики космического полета. М.: Наука, 1965.
- [Бахшиян, 1970] *Бахшиян Б. Ц.* Выбор оптимальных моментов независимых траекторных измерений // Космич. исследования. 1970. Т. 8. № 1. С. 3–7.
- [Бахшиян, 1983] *Бахшиян Б. Ц.* Представление весовых матриц, определяющих заданную оценку наименьших квадратов // Навигационная привязка и статистическая обработка космической информации М.: Наука. 1983. С. 81–90.
- [Бахшиян, 1989] *Бахшиян Б. Ц.* Критерии оптимальности и алгоритмы решения вырожденной и обобщенной задач линейного программирования // Экономика и математические методы. 1989. Т. 28. № 2. С. 314–324.
- [Бахшиян, 2012] *Бахшиян Б. Ц.* Оценивание и коррекция параметров движущихся систем: Курс лекций. М.: ИКИ РАН, 2012.
- [Бахшиян, Соловьев, 1998] *Бахшиян Б. Ц., Соловьев В. Н.* Теория и алгоритмы решения задач L- и MV-оптимального планирования эксперимента // Автоматика и телемеханика. 1998. № 8. С. 80–96.
- [Бахшиян и др., 1980] *Бахшиян Б. Ц., Назиров Р. Р., Эльясберг П. Е.* Определение и коррекция движения. М.: Наука, 1980.
- [Бахшиян и др., 2000] *Бахшиян Б. Ц., Матасов А. И., Федяев К. С.* О решении вырожденных задач линейного программирования // Автоматика и телемеханика. 2000. № 1. С. 105–117.
- [Белецкий, 2009] Белецкий В. В. Очерки о движении космических тел. 3-е изд. М.: Издво ЛКИ, 2009.
- [Богачев и др., 2001] *Богачев С. А., Мамон П. А., Миронов Ю. В., Ничипорович О. П.* Баллистика и теория полета в примерах и задачах. СПб.: ВИКУ, 2001.
- [Бордовицына, Авдюшев, 2007] *Бордовицына Т. В., Авдюшев В. А.* Теория движения искусственных спутников Земли. Аналитические и численные методы: Учебное пособие. Томск: Изд-во Том. ун-та, 2007.
- [БСЭ, 1969–1978] Большая советская энциклопедия. 3-е изд.: В 30 т. М.: Советская энциклопедия, 1969–1978.

- [Википедия] Википедия. Свободная энциклопедия. [Электрон. ресурс]. Режим доступа: <https://ru.wikipedia.org/wiki>.
- [Данциг, 1966] *Данциг Дж.* Линейное программирование, его применения и обобщения М.: Прогресс, 1966.
- [Иванов, Лысенко, 2004] *Иванов Н. М., Лысенко Л. Н.* Баллистика и навигация космических аппаратов. 2-е изд. М.: Дрофа, 2004.
- [Ильин, Позняк, 1968] *Ильин В. А., Позняк Э. Г.* Аналитическая геометрия. М.: Наука, 1968.
- [Кононович, Мороз, 2009] *Кононович Э. В., Мороз В. И.* Общий курс астрономии: учебное пособие. 3-е изд. М.: Едиториал УРСС, 2009.
- [Лидов, 1964] *Лидов М. Л.* К априорным оценкам точности определения параметров по методу наименьших квадратов // Космич. исслед. 1964. Т. 2. № 5. С. 713–718.
- [Лидов, 1971] *Лидов М. Л.* Математическая аналогия между некоторыми оптимальными задачами коррекции траекторий и выбора состава измерений и алгоритмы их решения // Космич. исслед. 1971. Т. 9. № 5. С. 687–706.
- [Маркеев, 1999] *Маркеев А. П.* Задача трех тел и ее точные решения // Соросовский образовательный журнал+ 1999. Т. 9. С. 112–117.
- [Матасов, 1988а] Матасов А. И. Об оптимальности линейных алгоритмов гарантирующего оценивания. Часть I // Космич. исслед. 1988. Т. 26. № 5. С. 643–653.
- [Матасов, 1988б] Матасов А. И. Об оптимальности линейных алгоритмов гарантирующего оценивания. Часть II // Космич. исслед. 1988. Т. 26. № 6. С. 807–812.
- [Охоцимский, Сихарулидзе, 1990] *Охоцимский Д. Е., Сихарулидзе Ю. Г.* Основы механики космического полета: Учебное пособие. М.: Наука, 1990.
- [Савченко, 2010] *Савченко А. А.* Персоналии в науке. Математики, механики, физики, астрономы, химики. М.: Р. Валент, 2010.
- [Сихарулидзе, 2011] *Сихарулидзе Ю. Г.* Баллистика и наведение летательных аппаратов. М.: БИНОМ. Лаборатория знаний, 2011.
- [Суханов, 2010] *Суханов А. А.* Астродинамика. М.: ИКИ РАН, 2010.
- [Эльясберг, 1965] *Эльясберг П. Е.* Введение в теорию полета искусственных спутников Земли. М.: Наука, 1965.
- [Эльясберг, 1976] *Эльясберг П. Е.* Определение движения. М.: Наука, 1976.
- [Эльясберг, 1983] *Эльясберг П. Е.* Измерительная информация: сколько её нужно? как её обрабатывать? М.: Наука, 1983.
- [Elfving, 1952] *Elfving G.* Optimum allocation in linear programming // Annals of Mathematical Statistics. 1952. V. 23. P. 255.
- [Rao, 1975] *Rao C. R.* On a unified theory of estimation in linear models — a review of recent results // Perspectives in probability and statistics. L. etc.: Acad. Press, 1975. P. 89–104.

055(02)2

Ротапринт ИКИ РАН
117997, Москва, Профсоюзная, 84/32

Подписано к печати 28.02.2013 г.

Заказ 3209

Формат 70×108/32

Тираж 100

5,0 уч.-изд. л.

