Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»

А. Е. Умнов, Е. А. Умнов

АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ И ЛИНЕЙНАЯ АЛГЕБРА

Издание четвертое, исправленное и дополненное

Учебное пособие

МОСКВА
МФТИ
25июн2025г

УДК 517.9(075.8)
ББК 22.161.6я73
У54

Рецензенты:

Доктор физико-математических наук, профессор
межфакультетской кафедры проблем передачи информации
и анализа данных МФТИ *А. В. Назин*

Зав. кафедрой вычислительной техники РТУ МИРЭА
кандидат технических наук, доцент *О. В. Платонова*

Умнов, Александр Евгеньевич,
Умнов, Егор Александрович

У54 Аналитическая геометрия и линейная алгебра. Изд. 4-е, испр. и дополн. : учебное пособие / А. Е. Умнов, Е. А. Умнов ; М-во науки и высшего образования Рос. Федерации, Моск. физ.-техн. ин-т (нац. исслед. ун-т). – Москва : МФТИ, 2025. – 460 с.
ISBN 978-5-7417-0837-8

Представляет собой введение в теорию линейных пространств, предваряемое изложением в терминах этой теории основ евклидовой геометрии.

Рассматриваемые в учебном пособии вопросы соответствуют стандартной университетской программе по предмету «Аналитическая геометрия и линейная алгебра» и могут являться основой для последующего, более глубокого, ознакомления как с теорией, так и с приложениями данного предмета. Изложение материала достаточно подробное и включает описание методов решения некоторых, принципиально важных для успешного освоения курса, типов задач.

Предназначено для студентов высших учебных заведений физико-математического, технического, естественнонаучного и экономического направлений подготовки, программа обучения которых предусматривает изучение базовых тем данного учебного курса.

УДК 517.9(075.8)
ББК 22.161.6я73

Печатается по решению Редакционно-издательского совета Московского физико-технического института (национального исследовательского университета)

ISBN 978-5-7417-0837-8

Оглавление

Предисловие

Отличительной чертой подготовки специалистов в Московском физико-техническом институте (национальном исследовательском университете) — *системы Физтеха* — является сочетание интенсивности обучения с высоким уровнем детализации и глубины изучаемых предметов, в первую очередь естественных наук.

Кафедра высшей математики МФТИ как важный элемент этой системы с момента образования института продолжает вносить существенный вклад в ее формирование и совершенствование. В активе кафедры колоссальный опыт в виде учебных курсов, оригинальных лекций по многим разделам современной математики, систем заданий, методических разработок, приемов, внутрикафедральных материалов, наконец, педагогического фольклора. На кафедре сформировался коллектив преподавателей, педагогически одаренных и обладающих педагогическим мастерством.

Поэтому вполне естественно стремление сделать этот опыт всеобщим достоянием. Многое уже отражено в известных учебниках, задачниках, созданных выдающимися математиками и педагогами, среди которых В. С. Владимиров, С. М. Никольский, Л. Д. Кудрявцев, М. В. Федорюк, О. В. Бесов и многие другие. Без сомнения, эти, ставшие уже классическими, учебные пособия оказали и оказывают существенное влияние на математическое образование как в России, так и за ее пределами.

Вместе с тем есть еще немало того, что, несомненно, будет существенно полезным для улучшения подготовки специалистов. Естественным путем для выявления этого опыта, как нам представляется, могла бы быть серия «Лекции кафедры высшей математики МФТИ», и мы будем благодарны всем, кто окажет поддержку и посильную помощь в осуществлении данного проекта.

В настоящем издании читателю предлагается одна из книг задуманной серии — расширенный курс лекций, подготовленный профессором А. Е. Умновым и доцентом Е. А. Умновым, который ряд лет читался студентам первого курса МФТИ.

По содержанию и стилю изложения материала данная книга рассчитана на студентов физико-математических и технических специальностей высших учебных заведений с углубленной подготовкой по математике. В ней представлены как традиционные разделы аналитической геометрии, теории матриц, теории линейных систем и конечномерных векторных пространств, так и некоторые дополнительные разделы линейной алгебры, важные для студентов физических специальностей.

На кафедре высшей математики МФТИ лекции по аналитической геометрии и линейной алгебре в разное время читали многие выдающиеся ученые и педагоги, такие, как Ф. Р. Гантмахер, В. Б. Лидский, А. А. Абрамов, Д. В. Беклемишев, В. А. Треногин, А. А. Болибрух и другие. Сами авторы, будучи последовательно студентами, аспирантами и преподавателями этой кафедры, не могли не испытать влияния своих учителей. Структура и дух лекций вполне традиционны для кафедры высшей математики МФТИ. В изложении материала авторы успешно сочетают, не злоупотребляя абстракциями, достаточно высокий уровень строгости с простотой и ясностью.

Предлагаемый читателям курс «Аналитическая геометрия и линейная алгебра» рекомендован кафедрой высшей математики МФТИ в качестве учебного пособия для студентов этого института. Книга также может быть использована в качестве учебного пособия и в других учебных заведениях с расширенной подготовкой по высшей математике.

Г. Н. Яковлев, В. Б. Трушин

От авторов

Данное пособие предназначено для студентов физических и технических специальностей высших учебных заведений с углубленной подготовкой по высшей математике. Его основной целью является введение в теорию линейных пространств — математический аппарат, используемый в разнообразных прикладных дисциплинах: от квантовой механики до методов оптимального управления.

Имея в виду особую терминологическую специфику этой теории, ее описание предваряется изложением основ евклидовой геометрии, выполненным при помощи понятий, характерных для теории линейных пространств.

Включенный в пособие материал в основном соответствует программе курсов «Аналитическая геометрия» и «Линейная алгебра», читаемых для студентов первого курса МФТИ. Также рассматриваются некоторые дополнительные вопросы, облегчающие студентам изучение математического аппарата теоретической физики и в первую очередь квантовой механики. Задачи, небольшое число которых включено в текст пособия, по мнению авторов, существенны для понимания курса в целом.

Предполагается, что читатель владеет основными понятиями элементарной геометрии, а также знаком в минимальном объеме с дифференциальным и интегральным исчислением. Используемая система обозначений единообразна для всех разделов пособия, что привело к небольшим отличиям от традиционной системы обозначений, в частности:

- действительные числа, как правило, обозначаются строчными греческими буквами (исключение сделано лишь для декартовых координат x, y и z, целочисленных индексов и некоторых других стандартных обозначений);
- строчные латинские буквы в основном использованы для обозначения более сложных, чем действительные числа, объектов: векторов, комплексных чисел, элементов линейных пространств, функций, функционалов, операторов, а также различных геометрических объектов;
- матрицы обозначаются буквами с двойными вертикальными ограничителями: например, $\|A\|$;
- во избежание конфликтов, для обозначений длин, абсолютных величин, модулей и норм используются одинарные вертикальные ограничители: например, $|a|$, в то время как для обозначения определителей матриц они не применяются, а используются обозначения функционального вида: например, $\det\|A\|$.

Авторы выражают глубокую признательность преподавателям и сотрудникам кафедры высшей математики МФТИ, советы и замечания которых в большой степени способствовали улучшению пособия, и в первую очередь И. А. Чубарову, С. В. Ивановой и В. Б. Трушину.

Предисловие к четвертому изданию

Со времени выхода в свет в 1997 году первого издания были учтены рекомендации, позволившие улучшить структуризацию материала, включенного в пособие, исправлены замеченные опечатки и неточности. Авторы особо благодарны посетителям интернет-сайта `www.umnov.ru` за доброжелательную критику и конструктивные замечания по версии текста, доступной на этом сайте.

1. Векторы и линейные операции с ними

1.1. Матричные объекты

Аналитическое описание геометрических линий, фигур и тел, равно как и операций с ними, может быть в большом числе случаев упрощено при использовании специального математического объекта, называемого *матрицей.*

`Определение 1.1.1`	Матрицей размера $m \times n$ называется упорядоченная прямоугольная таблица (или массив) чисел, содержащая m строк и n столбцов.

Числа, образующие матрицу, называемые ее элементами (или компонентами), характеризуются как своим значением, так и номерами строк и столбцов, в которых они расположены.

Условимся обозначать значение элемента матрицы, расположенного в i-й строке и j-м столбце, как α_{ij}[1] .

`Определение 1.1.2`	Числа m, n и $m \times n$ называются размерами матрицы.

Матрицы обозначаются и записываются перечислением их элементов. Например, матрица с элементами α_{ij} $\forall i = [1, m]$ $\forall j = [1, n]$ может быть записана в *развернутой* форме:

[1]Читается как ’*альфа* $i - j$’ .

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} \qquad \text{или} \qquad \begin{Vmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{Vmatrix},$$

из которых мы будем использовать последнюю. Если же оказывается достаточным *неразвернутое* представление матрицы, то мы записываем ее в виде $\|\alpha_{ij}\|$ или просто $\|A\|$.

Матрицы принято классифицировать по количеству их строк и столбцов.

`Определение 1.1.3`	Если $n = m$, то матрица называется *квадратной, порядка* n . Матрица размера $m \times 1$ называется m-мерным (или m-компонентным) *столбцом.* Матрица размера $1 \times n$ называется n-мерной (или n-компонентной) *строкой.*

Отметим, что, хотя формально для обозначения строк или столбцов следует использовать двухиндексные записи $\|\alpha_{1j}\|$ или $\|\beta_{i1}\|$, неменяющиеся индексы принято опускать, в результате чего обозначения строк или столбцов имеют вид $\|\alpha_j\|$ или $\|\beta_i\|$ соответственно . В этих случаях необходимо явно указывать, о чем идет речь: о строке или о столбце.

Некоторые часто используемые матрицы с особыми значениями элементов имеют специальные названия и обозначения.

`Определение 1.1.4`	Квадратная матрица, для которой $\alpha_{ij} = \alpha_{ji}$ $\forall\, i, j = [1, n]$, называется *симметрической.* Матрица, все элементы которой равны нулю, называется *нулевой*. Нулевую матрицу обозначают как $\|O\|$.

Квадратная матрица порядка n вида

$$\left\| \begin{matrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{matrix} \right\|$$

называется *единичной*. Единичную матрицу принято обозначать $\|E\|$.

Операции с матрицами

Определение 1.1.5 Две матрицы $\|A\|$ и $\|B\|$ считаются *равными* (обозначается: $\|A\| = \|B\|$), если они одинаковых размеров и если их соответствующие компоненты равны, то есть $\alpha_{ij} = \beta_{ij} \;\; \forall i = [1, m]$ и $\forall j = [1, n]$.

Определение 1.1.6 Матрица $\|C\|$ называется *суммой* матриц $\|A\|$ и $\|B\|$ (обозначается: $\|C\| = \|A\| + \|B\|$), если матрицы $\|A\|$, $\|B\|$, $\|C\|$ одинаковых размеров и $\gamma_{ij} = \alpha_{ij} + \beta_{ij} \;\; \forall i = [1, m]$ и $\forall j = [1, n]$, где числа γ_{ij} являются соответствующими компонентами матрицы $\|C\|$.

Определение 1.1.7 Матрица $\|C\|$ называется *произведением* числа λ на матрицу $\|A\|$ (обозначается: $\|C\| = \lambda\|A\|$), если матрицы $\|A\|$, $\|C\|$ одинаковых размеров и $\gamma_{ij} = \lambda\alpha_{ij} \;\; \forall i = [1, m]$ и $\forall j = [1, n]$, где числа γ_{ij} являются соответствующими компонентами матрицы $\|C\|$.

Отметим, что умножать число можно на матрицу любого размера.

Замечание 1.1.1. В качестве всех (или некоторых) элементов матрицы допускается использование не только чисел, но и других математических объектов, для которых подходящим образом определены понятие равенства, операции сложения и умножения на число, например, векторов, функций или тех же матриц.

Определение 1.1.8 *Транспонированием* матрицы называется операция, в результате которой образуется новая матрица, в которой строками служат столбцы исходной, записанные с сохранением порядка их следования (рис. 1.1).

Матрица, которая получается в результате транспонирования матрицы $\|A\|$, обозначается $\|A\|^{\mathrm{T}}$.

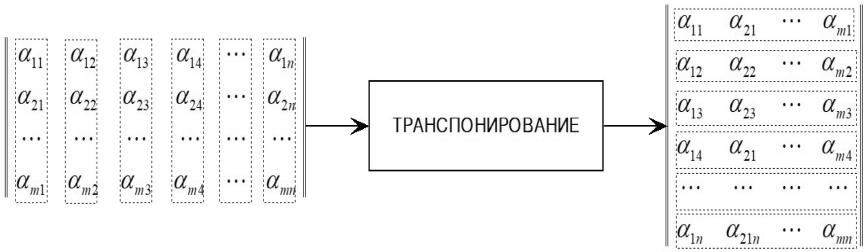


Рис. 1.1. Операция транспонирования

При транспонировании (то есть для элементов транспонированной матрицы $\|\alpha_{ij}\|$) верны равенства $\alpha_{ij}^{\mathrm{T}} = \alpha_{ji} \quad \forall i = [1, m]$ и $\forall j = [1, n]$.

Операция транспонирования, например, не изменяет симметрическую матрицу, но переводит строку размера n в столбец размера n и наоборот [2].

Детерминанты (определители) квадратных матриц 2-го и 3-го порядков

Для квадратных матриц $\|A\|$ имеется специальная числовая характеристика, называемая детерминантом (или определителем) и обозначаемая как $\det \|A\|$. Определение и свойства детерминантов квадратных матриц n-го порядка будут приведены в главе 6, здесь же мы пока ограничимся рассмотрением случаев $n = 2$ и $n = 3$.

[2] В дальнейшем, ради компактности записи формул, содержащих в своей записи столбцы, мы будем иногда представлять последние транспонированными строками.

Определение 1.1.9	*Детерминантом* (*определителем*) квадратной матрицы 2-го порядка $\left\Vert \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\Vert$ называется число $$\det \left\Vert \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\Vert = \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}\,.$$
Определение 1.1.10	*Детерминантом* (*определителем*) квадратной матрицы 3-го порядка $\left\Vert \begin{matrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{matrix} \right\Vert$ называется число $$\det \left\Vert \begin{matrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{matrix} \right\Vert =$$ $$= \alpha_{11}\alpha_{22}\alpha_{33} + \alpha_{13}\alpha_{21}\alpha_{32} + \alpha_{12}\alpha_{23}\alpha_{31} -$$ $$- \alpha_{13}\alpha_{22}\alpha_{31} - \alpha_{11}\alpha_{23}\alpha_{32} - \alpha_{12}\alpha_{21}\alpha_{33}\,.$$

Вычисление значений определителей 3-го порядка иногда удобнее выполнять не по определению 1.1.10, а используя их выражения через определители 2-го порядка.

Для описания этих выражений будут полезны:

Соглашение о суммировании

В тех случаях, когда явная запись суммы некоторого числа слагаемых нецелесообразна или невозможна, но известна $f(k)$ – зависимость значения каждого из слагаемых от его порядкового номера k , допускается использование специальной формы записи операции суммирования:

$$f(m) + f(m+1) + f(m+2) + \ldots + f(n-1) + f(n) = \sum_{k=m}^{n} f(k)$$

(читается: «сумма по k от m до n»), где k — *индекс суммирования*, m — *минимальное* значение индекса суммирования, n — *максимальное* значение индекса суммирования и, наконец, $f(k)$ – слагаемое с номером k.

`Пример 1.1.1:` По соглашению о суммировании будут справедливы следующие равенства:

$$
\begin{aligned}
1^2 + 2^2 + 3^2 + \ldots + 998^2 + 999^2 &= \sum_{k=1}^{999} k^2, \\
\sin\alpha + \sin 2\alpha + \ldots + \sin(n+3)\alpha &= \sum_{k=0}^{n+2} \sin(k+1)\alpha, \\
\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \ldots + \frac{1}{(M-1)M} &= \sum_{k=2}^{M} \frac{1}{(k-1)k}.
\end{aligned}
$$

Дополнительный минор

`Определение 1.1.11` *Минором, дополнительным к элементу* α_{ij}. в матрице $\|A\| = \left\|\begin{matrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{matrix}\right\|$ называется определитель квадратной матрицы 2-го порядка, получаемой из матрицы $\|A\|$ при удалении из нее i-й строки и j-го столбца.

Минор, дополнительный к элементу α_{ij}, принято обозначать $\overline{M}_i^j$. Например, $\overline{M}_2^3 = \det\left\|\begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{31} & \alpha_{32} \end{matrix}\right\|$.

Тогда будет справедлива

`Теорема 1.1.1` **Значение определителя матрицы 3-го порядка $\forall i = 1,2,3$ может быть выражено формулой**

$$
\det\left\|\begin{matrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{matrix}\right\| = \sum_{k=1}^{3} (-1)^{i+k}\, \alpha_{ik}\overline{M}_i^k \ ,
$$

а $\forall j = 1,2,3$ – формулой

$$
\det\left\|\begin{matrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{matrix}\right\| = \sum_{k=1}^{3} (-1)^{k+j}\, \alpha_{kj}\overline{M}_k^j \ .
$$

Доказательство.

Данные формулы проверяются непосредственно при помощи определений 1.1.9, 1.1.10 и 1.1.11.

Теорема доказана.

Первая из формул в теореме 1.1.1 называется *разложением определителя по i-й строке*, а вторая – *разложением по j-му столбцу.*

Пример 1.1.2 В силу теоремы 1.1.1, определитель матрицы 3-го порядка может быть разложен по первой строке и выражен через определители 2-го порядка следующим образом:

$$\det \begin{Vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{Vmatrix} = \alpha_{11} \det \begin{Vmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{Vmatrix} -$$

$$-\alpha_{12} \det \begin{Vmatrix} \alpha_{21} & \alpha_{23} \\ \alpha_{31} & \alpha_{33} \end{Vmatrix} + \alpha_{13} \det \begin{Vmatrix} \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{Vmatrix}.$$

Замечание 1.1.2. Соотношения, аналогичные приведенному в примере 1.1.2, могут быть записаны как для других строк, так и для *любого* из ее столбцов. При этом показатель степени у (-1) в каждом слагаемом равен сумме строкового и столбцового индексов элемента, умножаемого на детерминант. Для вычислений целесообразно выбирать то разложение, которое проще. Например, по строке или по столбцу со значительным числом нулей.

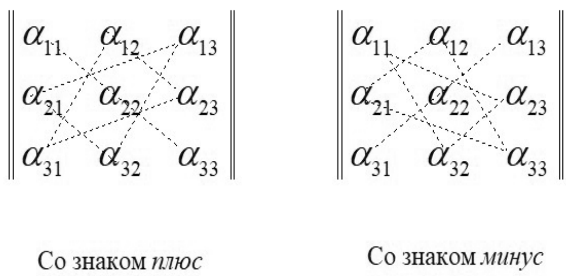


Рис. 1.2. Вычисление определителя 3-го порядка методом треугольников

Замечание 1.1.3. Иногда вычисление определителя матрицы 3-го порядка удобнее выполнить по следующему правилу: каждое слагаемое в определении 1.1.10 есть произведение некоторой тройки элементов матрицы, причем элементы, входящие в произведения, берущиеся со знаком «плюс», соединены в левой части рис. 1.2 штриховыми линиями, а элементы, входящие в произведения, которые берутся со знаком «минус», – в правой.

Непосредственная проверка показывает, что из определений 1.1.9 и 1.1.10 вытекает

Следствие 1.1.1 **При транспонировании квадратных матриц 2-го или 3-го порядков их определители не меняются.**

В терминах определителей матриц второго порядка достаточно удобно формулируется условие однозначной разрешимости системы двух линейных уравнений с двумя неизвестными.

Определение 1.1.12	Квадратные матрицы, имеющие нулевой детерминант, принято называть *вырожденными*, а квадратные матрицы с ненулевым детерминантом – *невырожденными*.

Теорема 1.1.2 (Крамера) **Для того чтобы система двух линейных уравнений имела единственное решение**

$$\begin{cases} \alpha_{11}\xi_1 & + & \alpha_{12}\xi_2 & = & \beta_1 \,, \\ \alpha_{21}\xi_1 & + & \alpha_{22}\xi_2 & = & \beta_2 \,, \end{cases}$$

необходимо и достаточно, чтобы

$$\det \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\| \neq 0 \,,$$

то есть чтобы матрица этого определителя была *невырожденной*.

Доказательство.

Докажем необходимость. Пусть данная система линейных уравнений имеет единственное решение — упорядоченную пару чисел ξ_1 и ξ_2 , тогда должны быть справедливыми вытекающие из ее уравнений соотношения

$$\xi_1 (\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}) = \beta_1\alpha_{22} - \beta_2\alpha_{12} \, ,$$
$$\xi_2 (\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}) = \beta_2\alpha_{11} - \beta_1\alpha_{21} \, ,$$

или $\quad \xi_1\Delta = \Delta_1; \quad \xi_2\Delta = \Delta_2, \quad$ где

$$\Delta = \det \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\| ,$$

$$\Delta_1 = \det \left\| \begin{matrix} \beta_1 & \alpha_{12} \\ \beta_2 & \alpha_{22} \end{matrix} \right\| , \qquad \Delta_2 = \det \left\| \begin{matrix} \alpha_{11} & \beta_1 \\ \alpha_{21} & \beta_2 \end{matrix} \right\| .$$

Равенства $\quad \xi_1\Delta = \Delta_1; \quad \xi_2\Delta = \Delta_2, \quad$ не верны, если

$$\begin{cases} \Delta = 0, \\ \Delta_1 \neq 0 \end{cases} \quad \text{или} \quad \begin{cases} \Delta = 0, \\ \Delta_2 \neq 0. \end{cases}$$

Если же $\Delta = \Delta_1 = \Delta_2 = 0$, то коэффициенты уравнений исходной системы обязаны (проверьте это самостоятельно) быть пропорциональными. Тогда исходная система может иметь решение, но оно не будет единственным.

Поэтому из условий существования и единственности решения следует, что $\Delta \neq 0$.

Докажем достаточность.

Если $\Delta \neq 0$, то числа

$$\xi_1 = \frac{\Delta_1}{\Delta} \qquad \text{и} \qquad \xi_2 = \frac{\Delta_2}{\Delta},$$

являющиеся решениями исходной системы, определяются однозначно значениями коэффициентов

$$\alpha_{11}, \ \alpha_{12}, \ \alpha_{21}, \ \alpha_{22}, \ \beta_1 \ \text{ и } \ \beta_2.$$

Теорема доказана.

1.2. Направленные отрезки

Определение 1.2.1	Отрезок прямой, концами которого служат точки A и B, называется *направленным отрезком*, если указано, какая из этих двух точек является началом и какая – концом отрезка. Направленный отрезок, начало и конец которого совпадают, называется *нулевым* направленным отрезком.

Будем записывать направленный отрезок в виде $\overline{AB}$, полагая, что точка A является началом отрезка, а точка B – его концом. Иногда направленный отрезок представляется просто как $\overline{a}$. Длина отрезка обозначается как $\left|\overline{AB}\right|$ или как $|\overline{a}|$.

Действия с направленными отрезками

Определение 1.2.2	Два ненулевых направленных отрезка называются *равными*, если их начала и их концы могут быть одновременно совмещены параллельным переносом одного из этих отрезков.

Заметим, что в силу данного определения параллельный перенос направленных отрезков не меняет.

Пусть даны два направленных отрезка $\overline{a}$ и $\overline{b}$.

Определение 1.2.3	Совместим начало отрезка $\overline{b}$ с концом $\overline{a}$ параллельным переносом отрезка $\overline{b}$, тогда направленный отрезок $\overline{c}$, начало которого совпадает с началом $\overline{a}$ и конец с концом $\overline{b}$, называется *суммой* направленных отрезков $\overline{a}$ и $\overline{b}$. Эту операцию принято обозначать так: $\overline{c} = \overline{a} + \overline{b}$.

Определение 1.2.3 иногда называют «*правилом треугольника*» (см. рис. 1.3–1)).

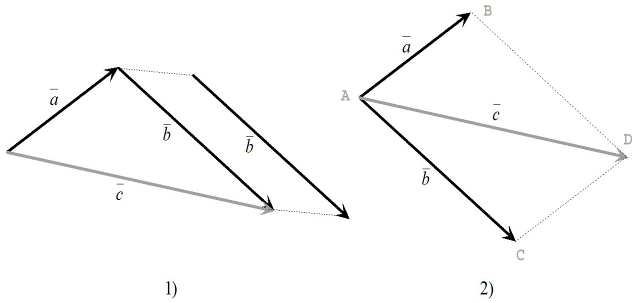


Рис. 1.3. Сложение направленных отрезков

Отметим, что для операции сложения направленных отрезков:

1) операция сложения направленных отрезков может быть выполнена по правилу параллелограмма, равносильному определению 1.2.3 (см. рис. 1.3–2)) ;
2) обобщение правила треугольника на любое число слагаемых носит название правила замыкающей, смысл которого ясен из рис. 1.4;
3) разностью направленных отрезков $\overline{a}$ и $\overline{b}$ называется направленный отрезок $\overline{c}$, удовлетворяющий равенству $\overline{a} = \overline{b} + \overline{c}$;
4) любой направленный отрезок при сложении с нулевым не изменяется.

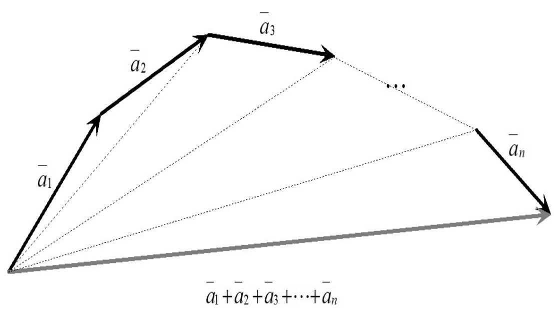


Рис. 1.4. Правило замыкающей

`Определение 1.2.4`	Под произведением числа λ на направленный отрезок $\overline{a}$ понимают: – при $\lambda = 0$ *нулевой* направленный отрезок, – при $\lambda \neq 0$ направленный отрезок $\overline{b}$, для которого *длина* $\left\|\overline{b}\right\| = \|\lambda\| \cdot \|\overline{a}\|$, а *направление* $\overline{b}$ совпадает с направлением $\overline{a}$ для $\lambda > 0$ и противоположно направлению $\overline{a}$ для $\lambda < 0$.

`Пример 1.2.1` Результат применения определения 1.2.4 иллюстрирует рис. 1.5.

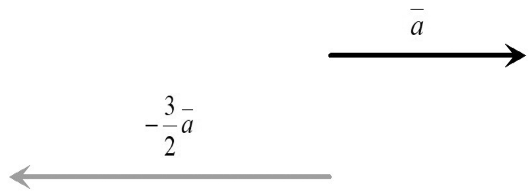


Рис. 1.5. Умножение числа на направленный отрезок

1.3. Определение множества векторов

`Определение 1.3.1`	Совокупность *всех* направленных отрезков, для которой определены описанные в § 1.2 : – понятие равенства (определение 1.2.2); – операция сложения (определение 1.2.3); – операция умножения числа на направленный отрезок (определение 1.2.4), называется *множеством векторов.* Конкретный элемент этого множества будем называть *вектором* и обозначать символом с верхней стрелкой, например, $\vec{a}$.

Нулевой вектор обозначается символом $\vec{o}$.

Имеет место

Теорема 1.3.1 **Операции сложения и умножения на вещественное число на множестве векторов обладают свойствами:**

1. *Коммутативности*

$$\vec{a}+\vec{b}=\vec{b}+\vec{a}.$$

2. *Ассоциативности*

$$\vec{a}+\left(\vec{b}+\vec{c}\right)=\left(\vec{a}+\vec{b}\right)+\vec{c};$$

$$\lambda\left(\mu\vec{a}\right)=\left(\lambda\mu\right)\vec{a}.$$

3. *Дистрибутивности*

$$\lambda\left(\vec{a}+\vec{b}\right)=\lambda\vec{a}+\lambda\vec{b};$$

$$\left(\lambda+\mu\right)\vec{a}=\lambda\vec{a}+\mu\vec{a}$$

для любых векторов $\vec{a}$, $\vec{b}$, $\vec{c}$ и любых чисел λ и μ.

Доказательство.

Данные свойства следуют из определения множества векторов и нуждаются в доказательстве. В качестве примера приведем доказательство свойства *коммутативности.*

Пусть даны два вектора $\vec{a}$ и $\vec{b}$. Совместим начала этих векторов и построим на них параллелограмм (рис. 1.3 – (2)).

Поскольку у параллелограмма противолежащие стороны параллельны и имеют равные длины, то $\vec{CD}=\vec{a}$; $\vec{BD}=\vec{b}$, но тогда, по правилу треугольника, из равных треугольников ΔACD и ΔABD следует, что

$$\vec{AD}=\vec{b}+\vec{CD}: \qquad \vec{AD}=\vec{a}+\vec{BD},$$

то есть

$$\vec{b}+\vec{a}=\vec{a}+\vec{b}.$$

Теорема доказана.

Давая определение понятия *вектора*, следует принять во внимание

`Замечание 1.3.1.` Иногда векторным называют объект, характеризуемый числовой величиной и направлением. Хотя формально такое определение и допустимо, оно неравносильно определению 1.3.1, что иллюстрируется примером, показанном на рис. 1.6.

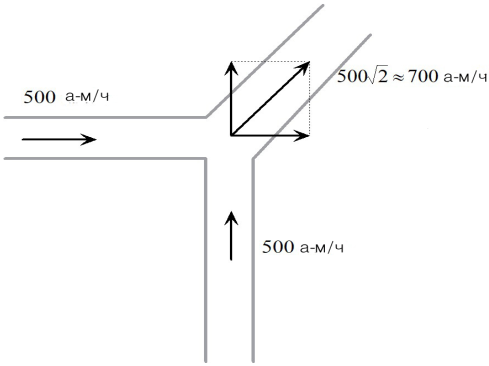


Рис. 1.6. Суммирование потоков

Например, поток автомобилей на некоторой трассе является характеристикой, для описания которой нужно указать как ее величину (число проходящих автомашин мимо наблюдателя за единицу времени), так и ее направление.

Предположим, что этот объект векторный (в смысле определения 1.3.1), и рассмотрим перекресток трех дорог достаточной пропускной способности, показанный на рис. 1.6, на котором сливаются два потока автомобилей по 500 а-м/ч каждый. Если суммировать потоки как векторы, то вместо очевидного результата 1000 а-м/ч мы получим (по правилу параллелограмма) заведомо бессмысленное значение $500\sqrt{2}$ а-м/ч. Откуда следует, что хотя поток автомашин характеризуется числовым значением и направлением, но тем не менее вектором (в смысле определения 1.3.1) он не является.

`Замечание 1.3.2.` С другой стороны, необходимо иметь в виду, что определение множества векторов 1.3.1 допускает их дальнейшую, более тонкую дифференциацию. Например, в некоторых физических и технических приложениях различают векторы полярные и аксиальные. К первым относятся, например, векторы скорости, силы, напряженности электрического поля; ко вторым – векторы момента силы, напряженности магнитного поля. Кроме того, в механике векторы подразделяются на свободные, скользящие и закрепленные, в зависимости от той роли, которую играет точка их приложения.

Замечание 1.3.3. К заключению о векторной природе тех или иных физических характеристик можно прийти путем рассуждений, основанных на определении 1.3.1 и экспериментальных данных.

Например, пусть некоторая материальная точка A , имеющая электрический заряд, перемещается в пространстве под действием электрического поля.

Положение этой точки в пространстве в момент времени t_0 можно задать исходящим из точки наблюдения и направленным в точку A вектором $\vec{r}(t_0)$, а в момент времени t — вектором $\vec{r}(t)$.

Поскольку перемещение $\vec{r}(t) - \vec{r}(t_0)$ (как разность двух векторов) является вектором, то и скорость движения материальной точки будет вектором в силу определения 1.3.1.

Рассуждая аналогично, можно прийти к заключению, что вектором является также и ускорение.

С другой стороны, согласно второму закону Ньютона, ускорение материальной точки пропорционально действующей на нее силе, и, следовательно, сила тоже есть вектор.

Наконец, принимая во внимание пропорциональность силы, действующей на заряженное тело, и напряженности электрического поля, заключаем, что последняя характеристика также векторная.

1.4. Линейная зависимость векторов

Из определения 1.3.1 следует, что все векторы (как элементы множества векторов) обладают одинаковыми свойствами.

Однако если рассмотреть во множестве векторов непустые подмножества, состоящие из конечного числа элементов, то окажется, что каждое такое подмножество (или набор) векторов *обязательно* будет обладать *одним из двух* взаимно исключающих друг друга свойств.

Данные свойства, играющие очень важную роль как в теории, так и в прикладных задачах, имеют названия *линейной зависимости* и *линейной независимости.*

Рассмотрим эти свойства подробнее, введя вначале определения вспомогательных понятий *коллинеарности* и *компланарности* векторов.

Определение 1.4.1

Два ненулевых вектора, параллельные одной и той же прямой, называются *коллинеарными*.
Три ненулевых вектора, параллельные одной и той же плоскости, называются *компланарными*.

Нулевой вектор считается коллинеарным любому вектору. Нулевой вектор считается компланарным любой паре векторов.

Дадим также

Определение 1.4.2

При любом натуральном k выражение вида

$$\lambda_1\vec{a}_1 + \lambda_2\vec{a}_2 + \ldots + \lambda_k\vec{a}_k \quad = \quad \sum_{i=1}^{k} \lambda_i\vec{a}_i\,,$$

где λ_i – некоторые числа, называется *линейной комбинацией* набора векторов $\{\vec{a}_1,\ \vec{a}_2,\ \ldots\ \vec{a}_k\}$.

Если все числа λ_i равны нулю одновременно, что равносильно условию

$$|\lambda_1| + |\lambda_2| + \ldots + |\lambda_k| = 0\,,$$

то в этом случае линейная комбинация называется *тривиальной*.
Если хотя бы одно из чисел λ_i отлично от нуля (то есть $|\lambda_1|+|\lambda_2|+\ldots+|\lambda_k| > 0$), то данная линейная комбинация называется *нетривиальной*.

Теперь можно дать определения линейной зависимости и линейной независимости системы векторов.

Определение 1.4.3

Набор векторов $\{\vec{a}_1,\ \vec{a}_2,\ \ldots\ \vec{a}_k\}$ $\forall k \in \mathbb{N}$ называется *линейно зависимым*, если существует их нетривиальная линейная комбинация, равная нулевому вектору, то есть такая, что

$$\sum_{i=1}^{k} \lambda_i\vec{a}_i = \vec{o}, \quad \text{где} \quad |\lambda_1| + |\lambda_2| + \ldots + |\lambda_k| > 0\,.$$

Определение 1.4.4	Набор векторов $\{\vec{a}_1,\ \vec{a}_2,\ \ldots\ \vec{a}_k\}$ называется *линейно независимым*, если из условия $$\sum_{i=1}^{k}\lambda_i\vec{a}_i=\vec{o}$$ следует тривиальность линейной комбинации $\sum\limits_{i=1}^{k}\lambda_i\vec{a}_i$, то есть равенство $\lambda_1=\lambda_2=\ldots=\lambda_k=0$.

Иначе говоря, если векторы $\vec{a}_1,\ \vec{a}_2,\ \ldots\ \vec{a}_k$ линейно независимы, то для любого набора чисел $\lambda_1,\ \lambda_2,\ \ldots\ ,\lambda_k$, не равных нулю одновременно, линейная комбинация $\sum\limits_{i=1}^{k}\lambda_i\vec{a}_i$ есть не нулевой вектор.

Приведем теперь некоторые свойства линейно зависимых и линейно независимых векторов, следующие из определений 1.4.3 и 1.4.4.

Лемма 1.4.1 **Для линейной зависимости набора векторов $\vec{a}_1,\ \vec{a}_2,\ \ldots\ \vec{a}_k,\quad k\geq 2$, необходимо и достаточно, чтобы один из них был линейной комбинацией остальных.**

Доказательство.

Докажем необходимость.

Пусть векторы $\vec{a}_1,\ \vec{a}_2,\ \ldots\ \vec{a}_k$ линейно зависимы, тогда существуют числа $\lambda_1,\ \lambda_2,\ \ldots\ ,\lambda_k$, одновременно не равные нулю, такие, что $\sum\limits_{i=1}^{k}\lambda_i\vec{a}_i=\vec{o}$.
Для определенности можно считать, что $\lambda_1\neq 0$, но тогда

$$\vec{a}_1=\sum_{i=2}^{k}\left(-\frac{\lambda_i}{\lambda_1}\right)\vec{a}_i\,,$$

что и доказывает необходимость.

Докажем достаточность.

Пусть для определенности $\vec{a}_1 = \sum\limits_{i=2}^{k} \lambda_i \vec{a}_i$, тогда

$$(-1)\vec{a}_1 + \sum_{i=2}^{k} \lambda_i \vec{a}_i = \vec{o},$$

причем $\quad |-1| + |\lambda_2| + \ldots + |\lambda_k| > 0$.

То есть линейная комбинация векторов $\vec{a}_1,\ \vec{a}_2,\ \ldots\ \vec{a}_k$, равная нулевому вектору, нетривиальная, и, значит, они линейно зависимые.

Лемма доказана.

Также справедливы следующие утверждения, связанные с понятиями линейная зависимость и линейная независимость.

Теорема 1.4.1 **Один вектор линейно зависим тогда и только тогда, когда он нулевой.**

Теорема 1.4.2 **Два вектора линейно зависимы тогда и только тогда, когда они коллинеарны.**

Теоремы 1.4.1 и 1.4.2 предлагаются читателю для самостоятельного доказательства.

Теорема 1.4.3 **Три вектора линейно зависимы тогда и только тогда, когда они компланарны.**

Доказательство.

Докажем необходимость. Пусть векторы $\vec{a}_1, \vec{a}_2, \vec{a}_3$ линейно зависимы, т.е. существуют три, не равных нулю одновременно, числа $\lambda_1,\ \lambda_2, \lambda_3$, такие, что

$$\lambda_1\vec{a}_1 + \lambda_2\vec{a}_2 + \lambda_3\vec{a}_3 = \vec{o}.$$

Тогда по лемме 1.4.1 один из векторов есть линейная комбинация двух остальных, и, значит, данные три вектора компланарны. Это и доказывает необходимость.

Докажем достаточность в предположении, что векторы $\vec{a}_1$ и $\vec{a}_2$ неколлинеарны.

Пусть даны три компланарных вектора $\vec{a}_1, \vec{a}_2$ и $\vec{a}_3$. Параллельным переносом совместим начала этих векторов в одной точке.

Через конец вектора $\vec{a}_3$ проведем прямые, параллельные векторам $\vec{a}_1$ и $\vec{a}_2$. При этом получим пару векторов $\vec{b}_1$ и $\vec{b}_2$, таких, что $\vec{a}_3 = \vec{b}_1 + \vec{b}_2$ (рис. 1.7).

Поскольку вектор $\vec{b}_1$ коллинеарен вектору $\vec{a}_1$, а вектор $\vec{b}_2$ коллинеарен вектору $\vec{a}_2$, то по лемме 1.4.1 и теореме 1.4.2 получаем, что

$$\vec{b}_1 = \lambda_1 \vec{a}_1 \qquad \text{и} \qquad \vec{b}_2 = \lambda_2 \vec{a}_2 \, ,$$

но тогда

$$\vec{a}_3 = \lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2$$

и векторы по лемме 1.4.1 линейно зависимы.

Случай коллинеарных $\vec{a}_1$ и $\vec{a}_2$ рассмотрите самостоятельно.

`Теорема доказана.`

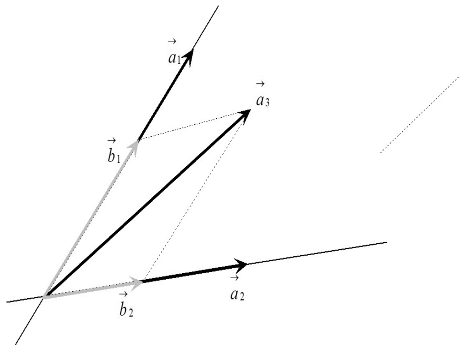


Рис. 1.7. К доказательству теоремы 1.4.3

`Теорема 1.4.4` **Если среди векторов $\{\vec{a}_1,\ \vec{a}_2,\ \dots\ \vec{a}_n\}$ имеется подмножество векторов, которые линейно зависимы, то и все эти векторы $\{\vec{a}_1,\ \vec{a}_2,\ \dots\ \vec{a}_n\}$ линейно зависимы.**

Доказательство.

Без ограничения общности можно считать, что линейно зависимы первые $k < n$ векторов (иначе просто перенумеруем эти векторы), то есть существуют числа $\lambda_1, \ \lambda_2, \ \ldots \ , \lambda_k$, одновременно не равные нулю, такие, что $\sum\limits_{i=1}^{k} \lambda_i \vec{a}_i = \vec{o}$.

Построим нетривиальную линейную комбинацию из векторов $\{\vec{a}_1, \ \vec{a}_2, \ \ldots \ \vec{a}_n\}$, взяв в качестве первых k коэффициентов числа $\lambda_1, \ \lambda_2, \ \ldots \ , \lambda_k$ и нули в качестве остальных. Тогда получим, что

$$\sum_{i=1}^{k} \lambda_i \vec{a}_i + \sum_{i=k+1}^{n} 0 \cdot \vec{a}_i = \vec{o}.$$

То есть исходный набор векторов линейно зависим.

Теорема доказана.

Следствие 1.4.1 **Если среди векторов $\{\vec{a}_1, \ \vec{a}_2, \ \ldots \ \vec{a}_n\}$ имеется хотя бы один нулевой, то векторы $\{\vec{a}_1, \ \vec{a}_2, \ \ldots \ \vec{a}_n\}$ линейно зависимы.**

1.5. Базис. Координаты вектора в базисе

Определение 1.5.1	*Базисом на прямой* называется любой ненулевой вектор, принадлежащий этой прямой. *Базисом на плоскости* называется любая упорядоченная пара линейно независимых векторов, принадлежащих этой плоскости. *Базисом в пространстве* называется любая упорядоченная тройка линейно независимых векторов.

Определение 1.5.2	Базис называется *ортогональным*, если его векторы попарно ортогональны (взаимно перпендикулярны).

Определение 1.5.3	Ортогональный базис называется *ортонормированным*, если образующие его векторы имеют единичную длину.

Пространственный базис, составленный из линейно независимых векторов $\vec{g}_1$, $\vec{g}_2$, $\vec{g}_3$, будем обозначать $\{\vec{g}_1, \vec{g}_2, \vec{g}_3\}$. Ортогональный или ортонормированный базис условимся обозначать как $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$.

`Теорема 1.5.1` **Пусть дан базис $\{\vec{g}_1, \vec{g}_2, \vec{g}_3\}$. Тогда любой вектор $\vec{x}$ в пространстве может быть представлен единственным образом в виде**

$$\vec{x} = \xi_1\vec{g}_1 + \xi_2\vec{g}_2 + \xi_3\vec{g}_3\,, \tag{1.5.1}$$

где ξ_1, ξ_2, ξ_3 — некоторые числа.

`Доказательство.`

1. Докажем вначале существование таких чисел. Совместим начала всех векторов $\vec{g}_1$, $\vec{g}_2$, $\vec{g}_3$ и $\vec{x}$ в точке O и проведем через конец вектора $\vec{x}$ плоскость, параллельную плоскости O, $\vec{g}_1$, $\vec{g}_2$ (рис. 1.8).

   Построим новые векторы $\vec{y}$ и $\vec{z}$ так, чтобы $\vec{x} = \vec{y} + \vec{z}$, а $\vec{z}$ и $\vec{g}_3$ были коллинеарны, тогда в силу коллинеарности векторов $\vec{z}$ и $\vec{g}_3$ имеем $\vec{z} = \xi_3\vec{g}_3$.

   Перенеся затем начало вектора $\vec{y}$ в точку O и рассуждая, как при доказательстве теоремы 1.4.3, получим $\vec{y} = \xi_1\vec{g}_1 + \xi_2\vec{g}_2$ и, следовательно,

$$\vec{x} = \xi_1\vec{g}_1 + \xi_2\vec{g}_2 + \xi_3\vec{g}_3,$$

   что доказывает существование разложения.

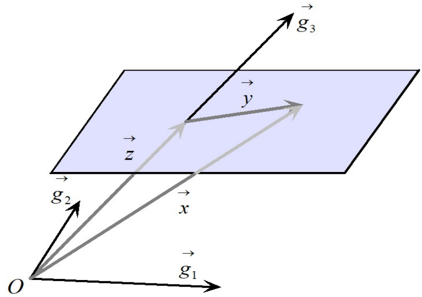


Рис. 1.8. К доказательству теоремы 1.5.1

2. Докажем единственность разложения по базису. Пусть мы имеем $\vec{x} = \xi_1\vec{g}_1 + \xi_2\vec{g}_2 + \xi_3\vec{g}_3$ и допустим, что существует другая тройка чисел ξ_1', ξ_2', ξ_3', таких, что

$$\vec{x} = \xi_1'\vec{g}_1 + \xi_2'\vec{g}_2 + \xi_3'\vec{g}_3.$$

Вычитая почленно эти равенства, получаем

$$\vec{o} = (\xi_1 - \xi_1')\vec{g}_1 + (\xi_2 - \xi_2')\vec{g}_2 + (\xi_3 - \xi_3')\vec{g}_3,$$

где в силу сделанного предположения о неединственности разложения $|\xi_1 - \xi_1'| + |\xi_2 - \xi_2'| + |\xi_3 - \xi_3'| > 0$. Полученное неравенство означает, что линейная комбинация

$$(\xi_1 - \xi_1')\vec{g}_1 + (\xi_2 - \xi_2')\vec{g}_2 + (\xi_3 - \xi_3')\vec{g}_3$$

нетривиальна, векторы $\{\vec{g}_1, \vec{g}_2, \vec{g}_3\}$ линейно зависимы и, следовательно, не могут быть базисом в силу определения 1.5.1.

Полученное противоречие доказывает единственность разложения.

Теорема доказана.

Определение 1.5.4	Числа ξ_1, ξ_2, ξ_3 – коэффициенты в разложении (1.5.1) – называются *координатами* вектора $\vec{x}$ в базисе $\{\vec{g}_1, \vec{g}_2, \vec{g}_3\}$.

Для сокращенной записи *координатного разложения* вектора

$$\vec{x} = \xi_1\vec{g}_1 + \xi_2\vec{g}_2 + \xi_3\vec{g}_3$$

используются формы: $\quad \vec{x}\,(\xi_1, \xi_2, \xi_3)\,, \quad \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}, \quad \left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{matrix} \right\|,$

из которых в дальнейшем мы будем использовать последнюю.

В общем случае утверждение, что *вектор $\vec{x}$ в базисе $\{\vec{g}_1, \vec{g}_2, \vec{g}_3\}$ имеет координатное представление* (или, что то же самое, *имеет координатный столбец*) $\left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{matrix} \right\|$, записывается как $\|\vec{x}\|_g = \left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{matrix} \right\|$.

Иногда, если это не приводит к неоднозначности толкования, будем использовать также нестрогую[3], упрощенную запись $\vec{x} = \left\| \begin{array}{c} \xi_1 \\ \xi_2 \\ \xi_3 \end{array} \right\|$.

Наконец, если вектор $\vec{x}$ в базисе $\{\vec{g}_1, \vec{g}_2\}$ на плоскости может быть представлен как $\vec{x} = \xi_1\vec{g}_1 + \xi_2\vec{g}_2$, то его координатная запись имеет вид $\|\vec{x}\|_g = \left\| \begin{array}{c} \xi_1 \\ \xi_2 \end{array} \right\|$.

Завершим обсуждение значения теоремы 1.5.1 следующим замечанием.

Согласно теоремам 1.4.1, 1.4.2 и 1.4.3 наборы векторов, состоящие из одного, двух или трех векторов, могут быть как линейно зависимыми, так и линейно независимыми.

Из теоремы 1.5.1 следует, что *любой набор векторов в количестве четыре или более может быть только линейно зависимым.*

Действительно, согласно теореме 1.5.1 каждый вектор представи́м в виде линейной комбинации базисных векторов. Значит, любой набор, состоящий из *произвольного* вектора и трех базисных, по лемме 1.4.1 будет линейно зависимым.

Нетрудно заметить, что (опять-таки при помощи теоремы 1.5.1) любые четыре вектора можно привести к формату «произвольный вектор и три базисных». То есть любые четыре вектора в пространстве линейно зависимы.

В случае, когда число векторов рассматриваемого набора больше четырех, каждый такой набор имеет подмножества, состоящие из четырех векторов, которые линейно зависимые. Тогда по теореме 1.4.4 приходим к заключению о линейной зависимости и рассматриваемого набора.

1.6. Действия с векторами в координатном представлении

Поскольку в конкретном базисе каждый вектор полностью и однозначно описывается упорядоченной тройкой чисел – своим координатным представлением, то естественно возникает вопрос о том, как выполняются операции с векторами в координатном представлении.

[3]Понятно, что вектор не может быть равным матричному объекту.

Оказывается, что возможно не только записывать векторы при помощи матриц (столбцов), но и оперировать с ними в матричной форме, поскольку правила действий с векторами в координатной форме совпадают с правилами соответствующих операций с матрицами.

Ответ на вопрос о том, как выполняются в координатном представлении операции с векторами, введенные в определениях § 1.2, дает

Теорема 1.6.1 **В координатном представлении операции с векторами выполняются по следующим правилам:**

1. Два вектора

$$\vec{x} = \xi_1\vec{g}_1 + \xi_2\vec{g}_2 + \xi_3\vec{g}_3 \quad \textbf{и} \quad \vec{y} = \eta_1\vec{g}_1 + \eta_2\vec{g}_2 + \eta_3\vec{g}_3$$

равны тогда и только тогда, когда равны их координатные представления

$$\|\vec{x}\|_g = \|\vec{y}\|_g \quad \textbf{или} \quad \begin{cases} \xi_1 = \eta_1, \\ \xi_2 = \eta_2, \\ \xi_3 = \eta_3. \end{cases}$$

2. Координатное представление суммы векторов равно сумме координатных представлений слагаемых

$$\|\vec{x} + \vec{y}\|_g = \|\vec{x}\|_g + \|\vec{y}\|_g\,.$$

3. Координатное представление произведения числа на вектор равно произведению этого числа на координатное представление вектора

$$\|\lambda\vec{x}\|_g = \lambda\|\vec{x}\|_g\,.$$

Доказательство.

Поскольку рассуждения для всех трех пунктов аналогичны, ибо основаны на определениях координатного представления вектора и действий с матрицами, а также свойствах операций с векторами, то в качестве примера рассмотрим лишь доказательство пункта 2.

Использование свойств операций сложения векторов и умножения числа на вектор, указанных в формулировке теоремы 1.3.1, дает следующую цепочку равенств:

$$
\begin{aligned}
&\|\vec{x}+\vec{y}\|_g = \\
&= \|\ (\xi_1\vec{g}_1+\xi_2\vec{g}_2+\xi_3\vec{g}_3)+(\eta_1\vec{g}_1+\eta_2\vec{g}_2+\eta_3\vec{g}_3)\ \|_g = \\
&= \|\ (\xi_1+\eta_1)\,\vec{g}_1+(\xi_2+\eta_2)\,\vec{g}_2+(\xi_3+\eta_3)\,\vec{g}_3\ \|_g = \\
&= \left\|\begin{array}{c}\xi_1+\eta_1\\ \xi_2+\eta_2\\ \xi_3+\eta_3\end{array}\right\| = \left\|\begin{array}{c}\xi_1\\ \xi_2\\ \xi_3\end{array}\right\| + \left\|\begin{array}{c}\eta_1\\ \eta_2\\ \eta_3\end{array}\right\| = \|\vec{x}\|_g+\|\vec{y}\|_g\,.
\end{aligned}
$$

Теорема доказана.

Следствие 1.6.1 **Координатное представление линейной комбинации $\lambda\vec{x}+\mu\vec{y}$ является той же самой линейной комбинацией координатных представлений векторов $\vec{x}$ и $\vec{y}$:**

$$\|\ \lambda\vec{x}+\mu\vec{y}\ \|_g = \lambda\,\|\vec{x}\|_g+\mu\,\|\vec{y}\|_g\ .$$

Рассмотрим теперь вопрос о том, как в координатном представлении записываются условия линейной зависимости и независимости векторов.

Теорема 1.6.2 **Для того чтобы два вектора $\vec{x}$ и $\vec{y}$ на плоскости были *линейно зависимы*, необходимо и достаточно, чтобы их координатные представления $\|\vec{x}\|_g = \left\|\begin{array}{c}\xi_1\\ \xi_2\end{array}\right\|$ и $\|\vec{y}\|_g = \left\|\begin{array}{c}\eta_1\\ \eta_2\end{array}\right\|$ удовлетворяли условию** $\det\left\|\begin{array}{cc}\xi_1 & \eta_1\\ \xi_2 & \eta_2\end{array}\right\| = 0$**.**

Доказательство.

1. Докажем необходимость. Пусть векторы $\vec{x}$ и $\vec{y}$ линейно зависимы, тогда по лемме 1.4.1 верно соотношение $\vec{x}=\lambda\vec{y}$ или в координатной форме $\left\{\begin{array}{l}\xi_1=\lambda\eta_1,\\ \xi_2=\lambda\eta_2.\end{array}\right.$

   Исключив λ из этих двух скалярных равенств при $\eta_1\eta_2\neq 0$, получим $\xi_1\eta_2-\xi_2\eta_1=0$, которое и означает, что

   $$\det\left\|\begin{array}{cc}\xi_1 & \eta_1\\ \xi_2 & \eta_2\end{array}\right\| = 0\,.$$

2. Докажем достаточность. Пусть $\det\begin{Vmatrix} \xi_1 & \eta_1 \\ \xi_2 & \eta_2 \end{Vmatrix} = 0$, тогда

имеем, что $\dfrac{\xi_1}{\eta_1} = \dfrac{\xi_2}{\eta_2}$ (при $\eta_1\eta_2 \neq 0$), то есть соответствующие координаты векторов $\vec{x}$ и $\vec{y}$ пропорциональны, что доказывает линейную зависимость этих векторов.

Случай $\eta_1\eta_2 = 0$ рассмотрите самостоятельно.

`Теорема доказана.`

`Теорема 1.6.3` **Для того чтобы три вектора в пространстве $\vec{x}$, $\vec{y}$, $\vec{z}$ с координатными представлениями**

$$\|\vec{x}\|_g = \begin{Vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{Vmatrix} \qquad \|\vec{y}\|_g = \begin{Vmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{Vmatrix} \qquad \|\vec{z}\|_g = \begin{Vmatrix} \kappa_1 \\ \kappa_2 \\ \kappa_3 \end{Vmatrix}$$

были *линейно независимы*, необходимо и достаточно, чтобы эти координатные представления удовлетворяли условию

$$\det\begin{Vmatrix} \xi_1 & \eta_1 & \kappa_1 \\ \xi_2 & \eta_2 & \kappa_2 \\ \xi_3 & \eta_3 & \kappa_3 \end{Vmatrix} \neq 0\,.$$

`Доказательство.`

Пусть некоторая линейная комбинация векторов $\vec{x}$, $\vec{y}$, $\vec{z}$ равна нулевому вектору. То есть выполнено равенство

$$\lambda_1\vec{x} + \lambda_2\vec{y} + \lambda_3\vec{z} = \vec{o}\,. \tag{1.6.1}$$

Согласно определению 1.4.4 в случае линейной независимости векторов $\vec{x}$, $\vec{y}$, $\vec{z}$ из равенства (1.6.1) должна следовать тривиальность линейной комбинации, стоящей в левой части этого равенства, то есть

$$\lambda_1 = \lambda_2 = \lambda_3 = 0\,.$$

Воспользуемся этим свойством, записав равенство (1.6.1) в координатной форме.

По следствию 1.6.1 векторное равенство (1.6.1) в координатах имеет вид матричного уравнения

$$\lambda_1 \left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{matrix} \right\| + \lambda_2 \left\| \begin{matrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{matrix} \right\| + \lambda_3 \left\| \begin{matrix} \kappa_1 \\ \kappa_2 \\ \kappa_3 \end{matrix} \right\| = \left\| \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \right\| ,$$

которое, в свою очередь, равносильно однородной системе трех линейных уравнений с тремя неизвестными λ_1, λ_2 и λ_3:

$$\left\{ \begin{array}{rcrcrcl} \xi_1\lambda_1 & + & \eta_1\lambda_2 & + & \kappa_1\lambda_3 & = & 0, \\ \xi_2\lambda_1 & + & \eta_2\lambda_2 & + & \kappa_2\lambda_3 & = & 0, \\ \xi_3\lambda_1 & + & \eta_3\lambda_2 & + & \kappa_3\lambda_3 & = & 0. \end{array} \right. \qquad (1.6.2)$$

Согласно обобщению теоремы 1.1.2 (Крамера), которое будет рассмотрено в § 6.4, система (1.6.2) имеет единственное решение тогда и только тогда, когда определитель ее основной матрицы отличен от нуля, то есть

$$\det \left\| \begin{matrix} \xi_1 & \eta_1 & \kappa_1 \\ \xi_2 & \eta_2 & \kappa_2 \\ \xi_3 & \eta_3 & \kappa_3 \end{matrix} \right\| \neq 0\,. \qquad (1.6.3)$$

Но, с другой стороны, система (1.6.2) очевидно имеет нулевое решение $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Поэтому неравенство (1.6.3) является необходимым и достаточным условием тривиальности линейной комбинации векторов в левой части равенства (1.6.1), а значит, и линейной независимости векторов $\vec{x}$, $\vec{y}$, $\vec{z}$. Заметим, что альтернативная версия доказательства данной теоремы будет рассмотрена в § 2.6.

`Теорема доказана.`

1.7. Декартова система координат

Теперь приступим к решению нашей основной задачи — к созданию инструмента, позволяющего описывать геометрические объекты и исследовать их свойства, не прибегая к использованию каких-либо интуитивно понятных или наглядных представлений.

Зафиксируем в пространстве некоторую точку O и рассмотрим множество всех векторов в пространстве, началами которых является точка O.

Очевидно, что для *любой* точки M в пространстве существует *единственный* вектор, началом которого служит точка O, а концом является точка M. При этом также очевидно, что для *каждого* вектора, исходящего из точки O, существует *единственная* точка M, задаваемая его концом.

В этом случае принято говорить, что *между множеством точек в пространстве и множеством векторов с началом в точке O имеется взаимно однозначное соответствие.*

Определение 1.7.1	Совокупность базиса $\left\{\vec{g}_1, \vec{g}_2, \vec{g}_3\right\}$ и некоторой заранее выбранной точки O («*начало координат*»), в которую помещены начала базисных векторов, называется *общей декартовой системой координат*, которую принято обозначать как $\left\{ O, \vec{g}_1, \vec{g}_2, \vec{g}_3\right\}$.

Определение 1.7.2	Систему координат $\{ O, \vec{e}_1, \vec{e}_2, \vec{e}_3\}$, образованную при помощи ортонормированного базиса $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, называют *нормальной прямоугольной* (или *ортонормированной*) системой координат.

Определение 1.7.3	Вектор $\vec{r}_M = \xi_1\vec{g}_1 + \xi_2\vec{g}_2 + \xi_3\vec{g}_3$ называется *радиусом-вектором точки* M в системе координат $\left\{ O, \vec{g}_1, \vec{g}_2, \vec{g}_3\right\}$.

Согласно теореме 1.5.1 радиус-вектор (как частный вид векторов) будет иметь координатное представление $\|\vec{r}_M\|_g = \|\ \xi_1\ \xi_2\ \xi_3\ \|^{\mathrm{T}}$ и притом единственное для конкретной системы координат.

Это позволяет дать

Определение 1.7.4	Координаты $\vec{r}_M$ – радиуса-вектора точки M – называются *координатами точки* M в системе координат $\left\{ O, \vec{g}_1, \vec{g}_2, \vec{g}_3\right\}$.

Исходя из определений 1.7.3 и 1.7.4, приходим к заключению, что для выбранной конкретной системы координат $\{ O, \vec{e}_1, \vec{e}_2, \vec{e}_3\}$

между множеством точек пространства и множеством их координатных столбцов $\left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{matrix} \right\|$ существует взаимно однозначное соответствие.

А поскольку все геометрические объекты (линии, фигуры, поверхности и тела) являются совокупностями точек, то естественной представляется замена наглядного описания этих объектов формализованным описанием множества координатных столбцов и исследование свойств геометрических объектов в координатном представлении.

В последующих разделах нашего курса мы убедимся, что этот подход может быть успешно использован на практике.

Проиллюстрируем особенности использования векторно-координатного описания геометрических объектов на примере решения следующей задачи.

Задача 1.7.1 *Пусть в некоторой общей декартовой системе координат $\{ O, \vec{g}_1, \vec{g}_2, \vec{g}_3 \}$ две несовпадающие точки M_1 и M_2 имеют координатные представления вида*

$$\left\| \vec{OM_1} \right\|_g = \left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{matrix} \right\| \qquad \text{и} \qquad \left\| \vec{OM_2} \right\|_g = \left\| \begin{matrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{matrix} \right\| .$$

Требуется найти точку N, такую, чтобы для заданного числа λ выполнялось равенство $\vec{M_1N} = \lambda \vec{M_1M_2}$.

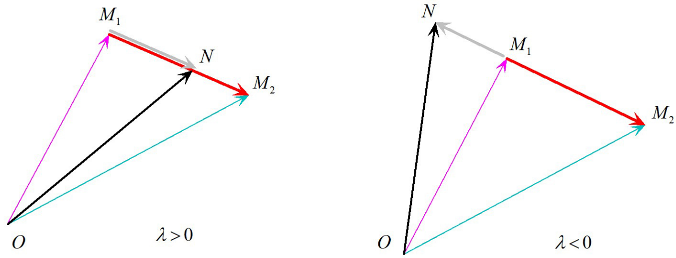


Рис. 1.9. К решению задачи 1.7.1

`Решение.` Найти точку – означает получить ее координатное представление или, проще говоря, найти ее координаты. По правилу треугольника (§ 1.2) получаем равенства

$$\vec{ON} = \vec{OM_1} + \vec{M_1N} \qquad \text{и} \qquad \vec{OM_2} = \vec{M_1M_2} + \vec{OM_1}\,.$$

Поскольку

$$\vec{M_1N} = \lambda\,\vec{M_1M_2}, \quad \text{а} \quad \vec{M_1M_2} = \vec{OM_2} - \vec{OM_1}\,,$$

то

$$\vec{ON} = \vec{OM_1} + \lambda\,\vec{OM_2} - \lambda\,\vec{OM_1} = (1-\lambda)\,\vec{OM_1} + \lambda\,\vec{OM_2}\,.$$

Откуда, по правилам действий с векторами в координатном представлении, окончательно получаем

`Решение получено.`

$$\left\|\vec{ON}\right\|_g = (1-\lambda)\,\left\|\vec{OM_1}\right\|_g + \lambda\,\left\|\vec{OM_2}\right\|_g = \left\|\begin{array}{c} (1-\lambda)\xi_1 + \lambda\,\eta_1 \\ (1-\lambda)\xi_2 + \lambda\,\eta_2 \\ (1-\lambda)\xi_3 + \lambda\,\eta_3 \end{array}\right\|.$$

`Замечание:` в механике к задаче 1.7.1 сводится задача отыскания центра масс системы материальных точек.

1.8. Изменение координат при замене базиса и начала координат

Поскольку выбор системы координат может быть сделан различными способами, вопрос об изменении координат при переходе от одного базиса к другому и при замене начала координат представляет значительный практический интерес.

Действительно, имея возможность решать одну и ту же задачу в разных системах координат, мы должны также иметь возможность сравнивать результаты этих решений. Будучи однозначным по сути, решение задачи может иметь разные координатные представления. Это — очевидный недостаток координатного метода.

Но, с другой стороны, можно ожидать, что в разных системах координат сложность постановки задачи (равно как и сложность методов ее решения) различная. Тогда можно попытаться выбрать такую систему координат, в которой эта сложность окажется минимальной. В последующих разделах нашего курсы мы неоднократно воспользуемся неоднозначностью координатного представления тех или иных математических объектов, для упрощения их описания и/или процедуры исследования их свойств.

Для того чтобы иметь практическую возможность использовать эту идею, прежде всего найдем формулы, выражающие зависимость координат произвольной точки пространства в *одной* системе координат, от координат этой же точки в *другой* декартовой системе координат.

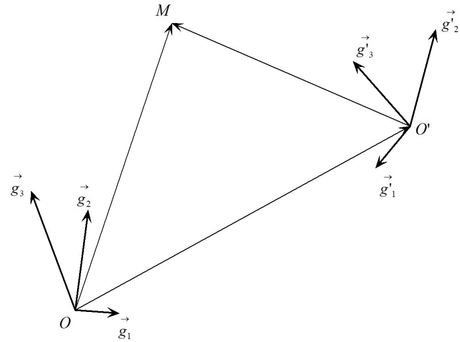


Рис. 1.10. Радиусы-векторы точки M в разных системах координат

Пусть выбраны две декартовы системы координат: *старая* $\{\,O,\,\vec g_1,\,\vec g_2,\,\vec g_3\}$ и *новая*[4] $\{\,O',\,\vec g\,'_1,\,\vec g\,'_2,\,\vec g\,'_3\}$ (рис. 1.10). Для лучшей читаемости все, что относится к *новой* системе, будем помечать штрихами, а для *старой* системы – оставим не помеченным.

Выразим векторы *нового* базиса, а также вектор $\vec{OO'}$ через векторы *старого* базиса. В силу теоремы 1.5.1 это можно сделать всегда и притом единственным образом:

$$\begin{aligned} \vec g\,'_1 &= \sigma_{11}\vec g_1+\sigma_{21}\vec g_2+\sigma_{31}\vec g_3\,,\\ \vec g\,'_2 &= \sigma_{12}\vec g_1+\sigma_{22}\vec g_2+\sigma_{32}\vec g_3\,,\\ \vec g\,'_3 &= \sigma_{13}\vec g_1+\sigma_{23}\vec g_2+\sigma_{33}\vec g_3\,,\\ \vec{OO'} &= \beta_1\vec g_1+\beta_2\vec g_2+\beta_3\vec g_3\,. \end{aligned} \tag{1.8.1}$$

Обратим внимание на два обстоятельства:

[4]С тем же успехом мы могли бы назвать их *первая* и *вторая*, или же – *красная* и *зеленая*.

1. Для полного описания связи между двумя декартовыми системами координат достаточно знать только двенадцать чисел, являющихся параметрами в формулах (1.8.1).

2. Метод индексации (нумерации) этих параметров может быть любым. Однако мы выбираем его так, что итоговые результаты будут иметь наиболее простой и легко запоминающийся вид.

Если использовать обозначения формул (1.8.1), то оказывается справедливой

Теорема 1.8.1 **Координаты произвольной точки пространства в *старой* системе координат связаны с координатами этой точки в *новой* системе координат системой уравнений**

$$\left\{\begin{array}{lcl} \xi_1 & = & \sigma_{11}\xi_1' + \sigma_{12}\xi_2' + \sigma_{13}\xi_3' + \beta_1\,, \\ \xi_2 & = & \sigma_{21}\xi_1' + \sigma_{22}\xi_2' + \sigma_{23}\xi_3' + \beta_2\,, \\ \xi_3 & = & \sigma_{31}\xi_1' + \sigma_{32}\xi_2' + \sigma_{33}\xi_3' + \beta_3\,. \end{array}\right. \qquad (1.8.2)$$

Доказательство.

Пусть радиус-вектор некоторой точки M (см. рис. 1.10) в *старой* системе координат имеет координатное представление и координатное разложение вида

$$\left\|\vec{OM}\right\|_g = \left\|\begin{array}{c} \xi_1 \\ \xi_2 \\ \xi_3 \end{array}\right\| \qquad \text{и} \qquad \vec{OM} = \xi_1\vec{g}_1 + \xi_2\vec{g}_2 + \xi_3\vec{g}_3\,.$$

Соответственно в *новой* системе координат будем иметь

$$\left\|\vec{O'M}\right\|_{g'} = \left\|\begin{array}{c} \xi_1' \\ \xi_2' \\ \xi_3' \end{array}\right\| \qquad \text{и} \qquad \vec{O'M} = \xi_1'\vec{g}_1' + \xi_2'\vec{g}_2' + \xi_3'\vec{g}_3'\,.$$

По правилу треугольника из $\Delta OMO'$ имеем

$$\vec{OM} = \vec{OO'} + \vec{O'M}\,.$$

Использовав соотношения (1.8.1), каждый из трех векторов в последнем равенстве можно представить как линейную комбинацию векторов *старого* базиса. Действительно, для векторов $\vec{OM}$ и $\vec{OO'}$ эти линейные комбинации у нас уже есть. А для $\vec{O'M}$ будет

$$\begin{aligned} \vec{O'M} &= \xi'_1\vec{g}'_1 + \xi'_2\vec{g}'_2 + \xi'_3\vec{g}'_3 = \\ &= \xi'_1\left(\sigma_{11}\vec{g}_1 + \sigma_{21}\vec{g}_2 + \sigma_{31}\vec{g}_3\right) + \\ &+\xi'_2\left(\sigma_{12}\vec{g}_1 + \sigma_{22}\vec{g}_2 + \sigma_{32}\vec{g}_3\right) + \\ &+\xi'_3\left(\sigma_{13}\vec{g}_1 + \sigma_{23}\vec{g}_2 + \sigma_{33}\vec{g}_3\right) . \end{aligned}$$

Подставив все три линейные комбинации в $\vec{OM} = \vec{OO'} + \vec{O'M}$ и сгруппировав в нем слагаемые, имеющие множителями одинаковые векторы *старого* базиса, получим

$$\lambda_1\vec{g}_1 + \lambda_2\vec{g}_2 + \lambda_3\vec{g}_3 = \vec{o}, \quad \text{где}$$

$$\begin{aligned} \lambda_1 &= -\xi_1 + \sigma_{11}\xi'_1 + \sigma_{12}\xi'_2 + \sigma_{13}\xi'_3 + \beta_1 , \\ \lambda_2 &= -\xi_2 + \sigma_{21}\xi'_1 + \sigma_{22}\xi'_2 + \sigma_{23}\xi'_3 + \beta_2 , \\ \lambda_3 &= -\xi_3 + \sigma_{31}\xi'_1 + \sigma_{32}\xi'_2 + \sigma_{33}\xi'_3 + \beta_3 . \end{aligned}$$

Поскольку векторы $\{\vec{g}_1, \vec{g}_2, \vec{g}_3\}$ линейно независимые, то их линейная комбинация, равная $\vec{o}$, обязана быть тривиальной, и потому $\lambda_1 = \lambda_2 = \lambda_3 = 0$ или окончательно

$$\left\{ \begin{aligned} \xi_1 &= \sigma_{11}\xi'_1 + \sigma_{12}\xi'_2 + \sigma_{13}\xi'_3 + \beta_1 , \\ \xi_2 &= \sigma_{21}\xi'_1 + \sigma_{22}\xi'_2 + \sigma_{23}\xi'_3 + \beta_2 , \\ \xi_3 &= \sigma_{31}\xi'_1 + \sigma_{32}\xi'_2 + \sigma_{33}\xi'_3 + \beta_3 . \end{aligned} \right.$$

`Теорема доказана.`

`Определение 1.8.1`	Формулы (1.8.2) называются *формулами перехода* от системы координат $\{O, \vec{g}_1, \vec{g}_2, \vec{g}_3\}$ к системе координат $\{O', \vec{g}'_1, \vec{g}'_2, \vec{g}'_3\}$.

При использовании формул перехода следует обратить внимание на то, что в равенствах (1.8.1) *штрихованные* переменные находятся в *левых* частях, а в равенствах (1.8.2) — в *правых*.

Заметим также, что коэффициенты уравнений в формулах (1.8.2), выражающих *старые* координаты через *новые*, образуют матрицу $\|S\|$, столбцы которой есть координатные представления (столбцы) *новых* базисных векторов в *старом* базисе, а столбец $\left\| \begin{matrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{matrix} \right\|$ образован координатами *нового* начала координат в *старом* базисе.

Определение 1.8.2 Матрица $\|S\| = \left\| \begin{matrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{matrix} \right\|$ называется *матрицей перехода* от *старого* базиса $\{\vec{g}_1, \vec{g}_2, \vec{g}_3\}$ к *новому* базису $\{\vec{g}\,'_1, \vec{g}\,'_2, \vec{g}\,'_3\}$.

Хотя *старая* и *новая* системы координат могут выбираться произвольно, оказывается, что матрицей перехода может служить *не любая* квадратная матрица третьего порядка. Имеет место следующая простая, но очень важная

Теорема 1.8.2 **Для матрицы перехода**

$$\det \|S\| = \det \left\| \begin{matrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{matrix} \right\| \neq 0.$$

Доказательство.

Согласно определению 1.8.2 столбцами матрицы перехода $\|S\| = \left\| \begin{matrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{matrix} \right\|$ служат координатные представления тройки линейно независимых векторов, образующих базис $\{\vec{g}\,'_1, \vec{g}\,'_2, \vec{g}\,'_3\}$.

Тогда из теоремы 1.6.3 следует доказываемое утверждение.

Теорема доказана.

Другими словами: невырожденные матрицы (и только они!) могут являться матрицами перехода между декартовыми системами координат.

Использовать в нашем курсе матрицы перехода мы будем еще не раз. Здесь же проиллюстрируем их важность следующим примером.

Мы установили, что в любой декартовой системе координат каждая точка имеет координатный столбец и притом единственный. Из формул перехода (1.8.2) очевидно следует, что для заданного столбца $\left\|\begin{array}{c}\xi_1'\\ \xi_2'\\ \xi_3'\end{array}\right\|$ столбец $\left\|\begin{array}{c}\xi_1\\ \xi_2\\ \xi_3\end{array}\right\|$ также определяется однозначно.

Но *старая* и *новая* системы координат равноправны. То есть и столбец $\left\|\begin{array}{c}\xi_1\\ \xi_2\\ \xi_3\end{array}\right\|$ должен однозначно определять столбец $\left\|\begin{array}{c}\xi_1'\\ \xi_2'\\ \xi_3'\end{array}\right\|$ через соотношения (1.8.2) .

Проверим это. Предположим, что координатное представление $\left\|\begin{array}{c}\xi_1\\ \xi_2\\ \xi_3\end{array}\right\|$ нам известно, а координатное представление $\left\|\begin{array}{c}\xi_1'\\ \xi_2'\\ \xi_3'\end{array}\right\|$ – нет. И рассмотрим формулы перехода (1.8.2) как систему трех линейных уравнений с тремя неизвестными ξ_1', ξ_2', ξ_3' . Коэффициентами при неизвестных в этой системе служат элементы матрицы перехода $\|S\|$, которая согласно теореме 1.8.2 не вырождена. Тогда, по теореме Крамера, рассматриваемая система линейных уравнений имеет решение и притом единственное.

Чтобы найти формулы, выражающие *новые* координаты через *старые*, решать эту линейную систему не обязательно. Можно снова воспользоваться равноправием *новой* и *старой* декартовых систем координат и вместо соотношений (1.8.1) использовать равенства

$$\begin{array}{rcl}\vec{g}_1 & = & \tau_{11}\vec{g}_1' + \tau_{21}\vec{g}_2' + \tau_{31}\vec{g}_3'\,,\\ \vec{g}_2 & = & \tau_{12}\vec{g}_1' + \tau_{22}\vec{g}_2' + \tau_{32}\vec{g}_3'\,,\\ \vec{g}_3 & = & \tau_{13}\vec{g}_1' + \tau_{23}\vec{g}_2' + \tau_{33}\vec{g}_3'\,,\\ \vec{O'O} & = & \gamma_1\vec{g}_1' + \gamma_2\vec{g}_2' + \gamma_3\vec{g}_3'\,,\end{array}$$

которые следуют из существования и единственности координатных разложений, но уже для *нового* базиса.

Проведя рассуждения, аналогичные использованным при доказательстве теоремы 1.8.1, мы получим искомые соотношения в виде

$$\left\{\begin{array}{rcl}\xi_1' & = & \tau_{11}\xi_1 + \tau_{12}\xi_2 + \tau_{13}\xi_3 + \gamma_1\,,\\ \xi_2' & = & \tau_{21}\xi_1 + \tau_{22}\xi_2 + \tau_{23}\xi_3 + \gamma_2\,,\\ \xi_3' & = & \tau_{31}\xi_1 + \tau_{32}\xi_2 + \tau_{33}\xi_3 + \gamma_3\,.\end{array}\right. \tag{1.8.3}$$

Формулы (1.8.3) естественно назвать формулами перехода от *новой* системы координат к *старой*. Иногда используется термин *формулы обратного перехода*. Наконец, матрицу

$$\|T\| = \left\| \begin{array}{ccc} \tau_{11} & \tau_{12} & \tau_{13} \\ \tau_{21} & \tau_{22} & \tau_{23} \\ \tau_{31} & \tau_{32} & \tau_{33} \end{array} \right\|$$

можно назвать *матрицей обратного перехода* или матрицей перехода от *новой* системы координат к *старой*. Разумеется, что $\det\|T\| \neq 0$, а столбцами матрицы $\|T\|$ являются координатные представления *старых* базисных векторов в *новом* базисе.

Проиллюстрируем построение формул перехода на примере следующей задачи.

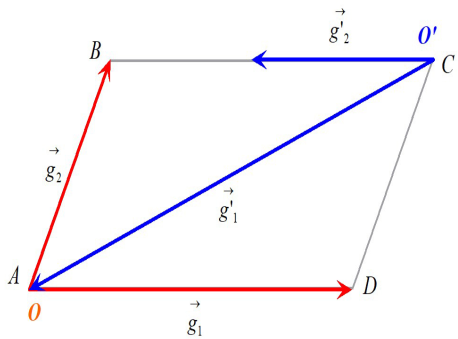


Рис. 1.11. К решению задачи 1.8.1

Задача 1.8.1 *При помощи параллелограмма $ABCD$ на плоскости построены две системы координат (см. рис. 1.11):*

старая $\{\, O,\, \vec{g}_1,\, \vec{g}_2 \}$, *такая что за O принята точка A, $\vec{g}_1 = \vec{AD}$, $\vec{g}_2 = \vec{AB}$*

и **новая** $\{\, O',\, \vec{g}\,'_1,\, \vec{g}\,'_2 \}$, *где за O' принята точка C, $\vec{g}\,'_1 = \vec{CA}$, $\vec{g}\,'_2 = -\frac{1}{2}\vec{BC}$.*

Требуется составить матрицы и формулы перехода, как прямого так и обратного, между старой и новой системами координат.

`Решение.` 1. Разложение векторов *нового* базиса и вектора $\vec{OO'}$ по векторам *старого* имеет вид

$$\left\{\begin{array}{rcrcr} \vec{g}'_1 & = & -\vec{g}_1 & - & \vec{g}_2, \\ \vec{g}'_2 & = & -\frac{1}{2}\vec{g}_1, & & \\ \vec{OO'} & = & \vec{g}_1 & + & \vec{g}_2. \end{array}\right.$$

Использовав коэффициенты этих разложений в качестве столбцов, составляем матрицу *прямого* перехода $\|S\| = \left\|\begin{array}{rr} -1 & -\frac{1}{2} \\ -1 & 0 \end{array}\right\|$ и столбец $\left\|\begin{array}{c} \beta_1 \\ \beta_2 \end{array}\right\| = \left\|\begin{array}{c} 1 \\ 1 \end{array}\right\|$.

Откуда получаем формулы перехода от *старой* системы координат к *новой*:

$$\left\{\begin{array}{rcrcrcr} \xi_1 & = & -\xi'_1 & - & \frac{1}{2}\xi'_2 & + & 1, \\ \xi_2 & = & -\xi'_1 & & & + & 1. \end{array}\right.$$

2. Запишем теперь разложение векторов *старого* базиса и вектора $\vec{O'O}$ по векторам *нового*:

$$\left\{\begin{array}{rcrcr} \vec{g}_1 & = & & - & 2\vec{g}'_2, \\ \vec{g}_2 & = & -\vec{g}'_1 & + & 2\vec{g}'_2, \\ \vec{O'O} & = & \vec{g}'_1. & & \end{array}\right.$$

Из коэффициентов этих разложений составляем матрицу *обратного* перехода $\|T\| = \left\|\begin{array}{rr} 0 & -1 \\ -2 & 2 \end{array}\right\|$ и столбец $\left\|\begin{array}{c} \gamma_1 \\ \gamma_2 \end{array}\right\| = \left\|\begin{array}{c} 1 \\ 0 \end{array}\right\|$.

Откуда получаем формулы перехода от *новой* системы координат к *старой*:

`Решение получено.`

$$\left\{\begin{array}{rcrcrcr} \xi'_1 & = & & - & \xi_2 & + & 1, \\ \xi'_2 & = & -2\xi_1 & + & 2\xi_2. & & \end{array}\right.$$

В заключение обратим внимание на следующий любопытный факт.

В задаче 1.8.1 мы получили $\det\|S\| = -\frac{1}{2}$ и $\det\|T\| = -2$. Откуда следует, что $\det\|S\| \cdot \det\|T\| = 1$.

Это соотношение не случайно. Произведение детерминантов матриц прямого и обратного переходов *всегда* равно 1. Правда, объяснение данного факта мы получим в процессе изучения материала главы 5.

Формулы перехода между ортонормированными системами координат на плоскости

Рассмотрим две *ортонормированные* системы координат $\{O, \vec{e}_1, \vec{e}_2\}$ и $\{O', \vec{e}_1', \vec{e}_2'\}$. Получим формулы перехода для случая, показанного на рис. 1.12А).

Из геометрически очевидных соотношений

$$\vec{e}_1^{\,\prime} = \vec{e}_1 \cos\varphi + \vec{e}_2 \sin\varphi \qquad \text{и} \qquad \vec{e}_2^{\,\prime} = -\vec{e}_1 \sin\varphi + \vec{e}_2 \cos\varphi$$

получаем матрицу перехода $\|S\| = \begin{Vmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{Vmatrix}$.

Тогда, если $\left\|\vec{OO'}\right\|_e = \begin{Vmatrix} \beta_1 \\ \beta_2 \end{Vmatrix}$, то *старые* координаты будут связаны с *новыми* формулами перехода:

$$\begin{cases} \xi_1 = \xi_1' \cos\varphi - \xi_2' \sin\varphi + \beta_1, \\ \xi_2 = \xi_1' \sin\varphi + \xi_2' \cos\varphi + \beta_2. \end{cases}$$

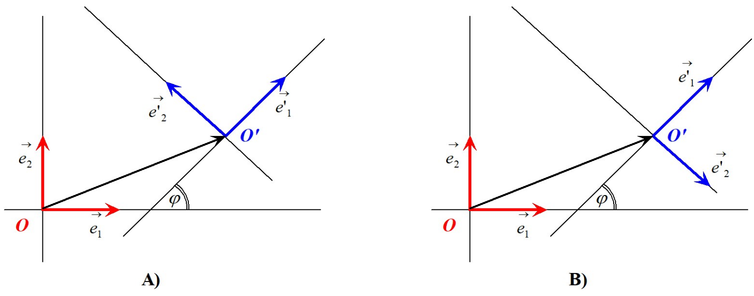


Рис. 1.12. К выводу формул перехода между ортонормированными ДСК

В рассмотренном случае обе системы координат удается совместить последовательным выполнением параллельного переноса *старой* системы на вектор $\vec{OO'}$ и последующего поворота на угол φ вокруг нового начала координат — точки O'.

Однако добиться такого совмещения, используя только параллельный перенос и поворот, вообще говоря, нельзя. Соответствующий случай показан на рис. 1.12B).

Здесь, после совмещения векторов $\vec{e}_1$ и $\vec{e}'_1$, еще потребуется симметричное отражение всей координатной плоскости относительно прямой, проходящей через совмещенные векторы. Формулы перехода будут в этом случае иметь вид

$$\begin{cases} \xi_1 &= \xi'_1\cos\varphi + \xi'_2\sin\varphi + \beta_1, \\ \xi_2 &= \xi'_1\sin\varphi - \xi'_2\cos\varphi + \beta_2. \end{cases}$$

Формально случаи, показанные на рис. 1.12A) и рис. 1.12B), можно различать, используя

Определение 1.8.3	Упорядоченная пара неколлинеарных векторов $\vec{a}$ и $\vec{b}$ на плоскости называется *право-ориентированной*, если кратчайший поворот от вектора $\vec{a}$ к вектору $\vec{b}$, при совмещении их начал, виден выполняющимся *против часовой стрелки*. В противном случае эта пара векторов называется *лево-ориентированной*.

Заметим, что для матрицы перехода $\|S\|$, связывающей два *ортонормированных* базиса, всегда будет $\det\|S\| = \pm 1$.

При этом $\det\|S\| = 1$, если ориентация обеих пар базисных векторов *одинаковая*, то есть если отражения при совмещении не требуется, и $\det\|S\| = -1$ для случая базисных пар *разной* ориентации.

2. Произведения векторов

2.1. Ортогональное проектирование

Эффективность использования векторно-координатного описания геометрических объектов существенно повышается при использовании *метрических характеристик*, таких как: *длина, расстояние, величина угла,* а также связанных с ними [5] специальных операций.

Одной из таких операций является *ортогональное проектирование.*

`Определение 2.1.1`	Прямую L с расположенным на ней ненулевым вектором $\vec{g}$ будем называть *осью.* Вектор $\vec{g}$ называется *направляющим вектором* оси L.

`Определение 2.1.2`	Пусть дана точка M, не лежащая на оси L, тогда основание перпендикуляра, опущенного из M на ось L – точку M^*, будем называть *ортогональной проекцией* точки M на ось L.

Примером оси может служить *ось координат* – прямая, проходящая через начало координат, направляющим вектором которой служит один из базисных векторов.

`Определение 2.1.3`	*Ортогональной проекцией* вектора $\vec{a}$ на ось L называется вектор $\hat{\mathrm{Pr}}_L\vec{a}$, лежащий на оси L, начало которого есть ортогональная проекция начала вектора на ось L, а конец — ортогональная проекция конца вектора $\vec{a}$ (см. рис. 2.1).

[5] Верхний символ $\widehat{\ }$ мы будем использовать в записи идентификаторов операций (или операторов): проектирования, поворота, умножения числа на объект, дифференцирования и т.п.

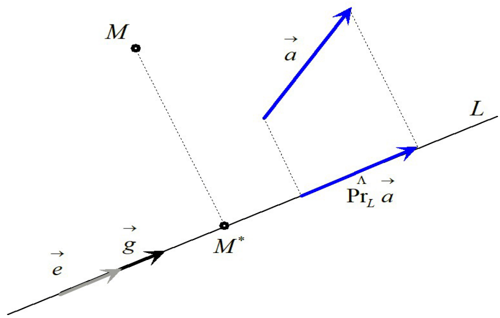


Рис. 2.1. Ортогональное проектирование на прямую

Выполним нормировку направляющего для L вектора $\vec{g}$, то есть заменим его на вектор $\vec{e} = \dfrac{\vec{g}}{|\vec{g}|}$ и рассмотрим нормированный базис $\{\,\vec{e}\}$ на оси L.

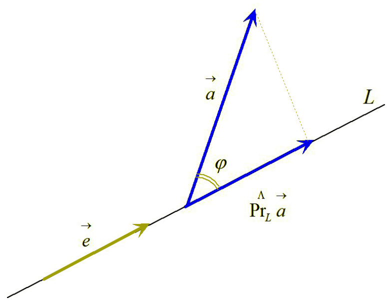


Рис. 2.2. Численное значение ортогональной проекции вектора

Определение 2.1.4	*Численным значением ортогональной проекции* вектора $\vec{a}$ на ось L называется координата вектора $\hat{\Pr}_L\vec{a}$ в базисе $\{\,\vec{e}\}$.

Определение 2.1.5	*Величиной угла* между ненулевыми векторами $\vec{a}$ и $\vec{b}$ называется величина наименьшего из двух углов, образуемых этими векторами при совмещении их начал.

Численное значение ортогональной проекции вектора $\vec{a}$ на ось L обозначим как $\text{Пр}_L\vec{a}$. Из рис. 2.2 очевидно, что $\text{Пр}_L\vec{a} = |\vec{a}|\cos\varphi$, где φ есть угол между $\vec{a}$ и $\vec{e}$.

Свойства ортогональных проекций

2.1. *Проекция суммы двух векторов равна сумме проекций этих векторов*:

$$\hat{\text{Pr}}_L\left(\vec{a}+\vec{b}\right) = \hat{\text{Pr}}_L\vec{a} + \hat{\text{Pr}}_L\vec{b}\,.$$

Данное свойство иллюстрирует рис. 2.3.

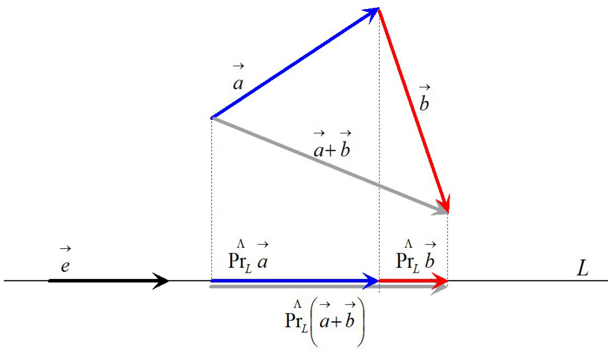


Рис. 2.3. Свойство ортогональной проекции суммы векторов

2.2. *Проекция произведения числа на вектор есть произведение этого числа на проекцию данного вектора*:

$$\hat{\text{Pr}}_L\left(\lambda\vec{a}\right) = \lambda\,\hat{\text{Pr}}_L\vec{a}\,.$$

Заметим, что свойства 2.1 и 2.2 можно объединить в следующее утверждение:

Проекция линейной комбинации векторов равна той же линейной комбинации проекций, образующих исходную линейную комбинацию:

$$\hat{\Pr}_L\left(\lambda\vec{a}+\mu\vec{b}\right)=\lambda\,\hat{\Pr}_L\vec{a}+\mu\,\hat{\Pr}_L\vec{b}\,.$$

Справедливость свойств 2.1 и 2.2 вытекает из определения операции ортогонального проектирования и правил действия с векторами.

Последнее равенство выражает *линейность* операции ортогонального проектирования на множестве векторов.

Свойства численных значений ортогональных проекций

2.3. *Численное значение проекции суммы двух векторов равно сумме численных значений проекций этих векторов*:

$$\text{Пр}_L\left(\vec{a}+\vec{b}\right)=\text{Пр}_L\vec{a}+\text{Пр}_L\vec{b}\,.$$

2.4. *Проекция произведения числа на вектор есть произведение этого числа на проекцию данного вектора*:

$$\text{Пр}_L\left(\lambda\vec{a}\right)=\lambda\,\text{Пр}_L\vec{a}\,.$$

2.5. Объединив 2.3 и 2.4, получим

$$\text{Пр}_L\left(\lambda\vec{a}+\mu\vec{b}\right)=\lambda\,\text{Пр}_L\vec{a}+\mu\,\text{Пр}_L\vec{b}\,.$$

Эти равенства, в свою очередь, следуют из определения численного значения ортогональных проекций и свойств операций с векторами.

Последнее из них выражает *линейность* численного значения ортогональной проекции на множестве векторов.

2.2. Скалярное произведение векторов и его свойства

Определение 2.2.1	*Скалярным произведением* ненулевых векторов $\vec{a}$ и $\vec{b}$ называется число, равное произведению их длин, умноженному на косинус угла между $\vec{a}$ и $\vec{b}$. В случае, когда хотя бы один из сомножителей есть нулевой вектор, скалярное произведение считается равным нулю.

Скалярное произведение векторов $\vec{a}$ и $\vec{b}$ обозначается как $\mathbf{a}\cdot\mathbf{b}$ или $(\vec{a},\vec{b}\,)$. Мы будем использовать последнее из этих обозначений.

Таким образом, для ненулевых векторов $\vec{a}$ и $\vec{b}$:

$$(\vec{a},\vec{b}\,) = \left|\vec{a}\right| \cdot \left|\vec{b}\right| \cdot \cos\varphi\,,$$

где φ — угол между векторами-сомножителями. При этом, согласно определению 2.1.5, $0 \le \varphi \le \pi$. Заметим также, что если $\vec{b} \neq \vec{o}$, то справедливо равенство $(\vec{a},\vec{b}) = \left|\vec{b}\right| \cdot \text{Пр}_{\vec{b}}\,\vec{a}\,.$

Свойства скалярного произведения

1) $(\vec{a},\vec{b}) = 0$ при $\vec{a} \neq \vec{o}$ и $\vec{b} \neq \vec{o}$ тогда и только тогда, когда $\vec{a}$ и $\vec{b}$ взаимно ортогональны.
2) $(\vec{a},\vec{b}) = (\vec{b},\vec{a})$ (*коммутативность*) следует из определений скалярного произведения и угла между векторами.
3) $(\vec{a}_1 + \vec{a}_2,\vec{b}) = (\vec{a}_1,\vec{b}) + (\vec{a}_2,\vec{b})$ (*дистрибутивность*).

> Доказательство.
>
> Если $\vec{b} = \vec{o}$, то 3) очевидно. Пусть $\vec{b} \neq \vec{o}$, тогда
>
> $$(\vec{a}_1 + \vec{a}_2,\vec{b}\,) = \left|\vec{b}\right| \cdot \text{Пр}_{\vec{b}}(\vec{a}_1 + \vec{a}_2) =$$
>
> $$= \left|\vec{b}\right| \text{Пр}_{\vec{b}}\vec{a}_1 + \left|\vec{b}\right| \text{Пр}_{\vec{b}}\vec{a}_2 = (\vec{a}_1,\vec{b}) + (\vec{a}_2,\vec{b})\,.$$
>
> Свойство доказано.

4) $(\lambda\vec{a},\vec{b}) = \lambda(\vec{a},\vec{b})\,.$

5) $(\vec{a},\vec{a}) = \left|\vec{a}\right|^2 \geq 0 \quad \forall\vec{a}$; $\quad \left|\vec{a}\right| = \sqrt{(\vec{a},\vec{a})}$. Отметим, что условия $(\vec{a},\vec{a}) = 0$ и $\vec{a} = \vec{o}$ равносильны.

6) При $\vec{a} \neq \vec{o}$ и $\vec{b} \neq \vec{o}$ $\quad \cos\varphi = \dfrac{(\vec{a},\vec{b})}{\left|\vec{a}\right| \cdot \left|\vec{b}\right|}$, где φ — угол между векторами $\vec{a}$ и $\vec{b}$.

2.3. Выражение скалярного произведения в координатах

Пусть дан базис $\{\vec{g}_1,\, \vec{g}_2,\, \vec{g}_3\}$ и два вектора $\vec{a}$ и $\vec{b}$ с координатными разложениями в этом базисе:

$$\vec{a} = \xi_1\vec{g}_1 + \xi_2\vec{g}_2 + \xi_3\vec{g}_3\,, \qquad \vec{b} = \eta_1\vec{g}_1 + \eta_2\vec{g}_2 + \eta_3\vec{g}_3\,.$$

Тогда по свойствам 3) и 4) скалярного произведения

$$\begin{aligned}(\vec{a},\vec{b}\,) = (\,\xi_1\vec{g}_1 + \xi_2\vec{g}_2 + \xi_3\vec{g}_3,\ \eta_1\vec{g}_1 + \eta_2\vec{g}_2 + \eta_3\vec{g}_3\,) = \\ = \xi_1\eta_1(\vec{g}_1,\vec{g}_1) + \xi_1\eta_2(\vec{g}_1,\vec{g}_2) + \xi_1\eta_3(\vec{g}_1,\vec{g}_3) + \\ + \xi_2\eta_1(\vec{g}_2,\vec{g}_1) + \xi_2\eta_2(\vec{g}_2,\vec{g}_2) + \xi_2\eta_3(\vec{g}_2,\vec{g}_3) + \\ + \xi_3\eta_1(\vec{g}_3,\vec{g}_1) + \xi_3\eta_2(\vec{g}_3,\vec{g}_2) + \xi_3\eta_3(\vec{g}_3,\vec{g}_3) = \\ = \sum_{j=1}^{3}\Big(\xi_j\eta_1(\vec{g}_j,\vec{g}_1) + \xi_j\eta_2(\vec{g}_j,\vec{g}_2) + \xi_j\eta_3(\vec{g}_j,\vec{g}_3)\Big) = \\ = \sum_{j=1}^{3}\sum_{i=1}^{3}\xi_j\eta_i(\vec{g}_j,\vec{g}_i)\,.\end{aligned}$$

В случае *ортонормированного* базиса эта формула упрощается, поскольку для попарных скалярных произведений базисных векторов справедливо равенство $\quad (\vec{e}_j,\vec{e}_i) = \delta_{ji} = \begin{cases} 1, & j = i, \\ 0, & j \neq i, \end{cases}$ где число δ_{ji} называется *символом Кронекера.* Откуда для скалярного произведения векторов в ортонормированном базисе получаем формулу

$$(\vec{a},\vec{b}\,) = \xi_1\eta_1 + \xi_2\eta_2 + \xi_3\eta_3\,,$$

из которой следуют соотношения: $\left|\vec{a}\right| = \sqrt{(\vec{a},\vec{a})} = \sqrt{\xi_1^2+\xi_2^2+\xi_3^2}$, а для случая $\vec{a} \neq \vec{o}$ и $\vec{b} \neq \vec{o}$,

$$\cos\varphi = \frac{(\vec{a},\vec{b})}{\left|\vec{a}\right|\left|\vec{b}\right|} = \frac{\xi_1\eta_1+\xi_2\eta_2+\xi_3\eta_3}{\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}\cdot\sqrt{\eta_1^2+\eta_2^2+\eta_3^2}}.$$

Отметим, что это равенство в сочетании с условием $|\cos\varphi| \leq 1$ приводит к верному $\forall \xi_i,\, \eta_j \quad i,j = 1\ldots 3$ *неравенству Коши – Буняковского*:

$$|\,\xi_1\eta_1+\xi_2\eta_2+\xi_3\eta_3\,| \leq \sqrt{\xi_1^2+\xi_2^2+\xi_3^2}\cdot\sqrt{\eta_1^2+\eta_2^2+\eta_3^2}\,.$$

Задача 2.3.1 *Найти расстояние между двумя точками в ортонормированной системе координат, если известны радиусы-векторы этих точек.*

Решение. Пусть даны: ортонормированная система координат $\{\, O, \vec{e}_1, \vec{e}_2, \vec{e}_3 \}$ и две точки M_1 и M_2, имеющие координатные представления вида

$$\left\|\vec{OM_1}\right\|_e = \begin{Vmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{Vmatrix} \qquad \text{и} \qquad \left\|\vec{OM_2}\right\|_e = \begin{Vmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{Vmatrix}.$$

Тогда, используя решение задачи 1.7.1 и формулу для длины вектора

$$\vec{M_1M_2} = (\xi_1-\eta_1)\vec{e}_1 + (\xi_2-\eta_2)\vec{e}_2 + (\xi_3-\eta_3)\vec{e}_3\,,$$

получим в ортонормированной системе координат

$$\left|\vec{M_1M_2}\right| = \sqrt{(\xi_1-\eta_1)^2+(\xi_2-\eta_2)^2+(\xi_3-\eta_3)^2}\,.$$

Решение получено.

2.4. Векторное произведение векторов и его свойства

Введем предварительно понятие *ориентации в пространстве* для произвольной тройки некомпланарных векторов.

Определение 2.4.1	Упорядоченная тройка некомпланарных векторов $\left\{\vec{a},\,\vec{b},\,\vec{c}\right\}$ называется *правой*, если (после совмещения начал этих векторов) кратчайший поворот от вектора $\vec{a}$ к вектору $\vec{b}$ виден из конца вектора $\vec{c}$, совершающимся против часовой стрелки. В противном случае данная упорядоченная тройка некомпланарных векторов называется *левой*.

Определение 2.4.2	*Векторным произведением* неколлинеарных векторов $\vec{a}$ и $\vec{b}$ называется вектор $\vec{c}$, такой, что 1) $\left\|\vec{c}\right\| = \left\|\vec{a}\right\| \cdot \left\|\vec{b}\right\| \cdot \sin\varphi$, где φ – угол между векторами $\vec{a}$ и $\vec{b}$; 2) вектор $\vec{c}$ ортогонален как вектору $\vec{a}$, так и вектору $\vec{b}$; 3) тройка векторов $\left\{\vec{a},\,\vec{b},\,\vec{c}\right\}$ правая. В случае, когда векторы $\vec{a}$ и $\vec{b}$ коллинеарны (в том числе, когда хотя бы один из сомножителей есть нулевой вектор), векторное произведение считается равным *нулевому вектору*.

Векторное произведение векторов $\vec{a}$ и $\vec{b}$ принято обозначать как $\mathbf{a}\times\mathbf{b}$ или $[\vec{a},\vec{b}]$. Мы будем использовать последнее из этих обозначений. Определение 2.4.2 иллюстрирует рис. 2.4A.

Из определения 2.4.2 следует, что

1) $\left|[\vec{a},\vec{b}]\right|$ равняется площади параллелограмма, построенного на векторах $\vec{a}$ и $\vec{b}$ при совмещении их начал;

2) для коллинеарности ненулевых векторов $\vec{a}$ и $\vec{b}$ необходимо и достаточно, чтобы их векторное произведение было равно нулевому вектору.

Свойства векторного произведения

1) $[\vec{a},\vec{b}] = -[\vec{b},\vec{a}]$ – *антикоммутативность* следует из определения 2.4.2 и нечетности функции $\sin\varphi$.
2) $[\lambda\vec{a},\vec{b}] = \lambda[\vec{a},\vec{b}]$ следует из определения векторного произведения и того факта, что векторы $[\lambda\vec{a},\vec{b}]$ и $[\vec{a},\vec{b}]$ ортогональны одной и той же плоскости при неколлинеарных $\vec{a}$ и $\vec{b}$ и $\lambda \neq 0$.
3) $[\vec{a}+\vec{b},\vec{c}] = [\vec{a},\vec{c}] + [\vec{b},\vec{c}]$ – *дистрибутивность* векторного произведения.

Для доказательства дистрибутивности векторного произведения воспользуемся следующими вспомогательными утверждениями.

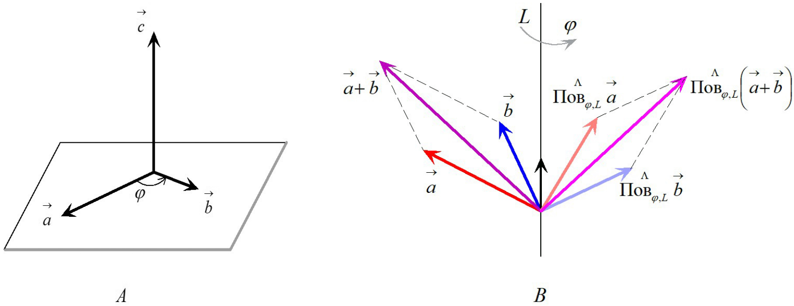


Рис. 2.4. К определению 2.4.2 и доказательству леммы 2.4.1

Лемма 2.4.1 **Пусть даны два вектора $\vec{a}$ и $\vec{b}$, начала которых находятся в общей точке на оси L. Тогда** *результат поворота суммы* **этих векторов на некоторый угол φ вокруг оси L равен** *сумме результатов поворота* **каждого из векторов $\vec{a}$ и $\vec{b}$ вокруг оси L на угол φ.**

Доказательство.

Символически утверждение леммы можно записать так:

$$\hat{\text{Пов}}_{\varphi,L}\left(\vec{a}+\vec{b}\right) = \hat{\text{Пов}}_{\varphi,L}\vec{a} + \hat{\text{Пов}}_{\varphi,L}\vec{b}\,.$$

Его справедливость следует из правила параллелограмма и того факта, что при пространственном повороте вокруг некоторой оси параллелограмм не деформируется (см. рис. 2.4B).

Лемма доказана.

Лемма 2.4.2 **Если** $\left|\vec{e}\right| = 1$ **и** $\vec{p} \neq \vec{o}$, **то вектор** $[\vec{p}, \vec{e}]$ **равен результату** *поворота проекции* **вектора** $\vec{p}$ **на плоскость, перпендикулярную вектору** $\vec{e}$, **вокруг вектора** $\vec{e}$, **на угол** $\frac{\pi}{2}$, **который виден из конца вектора** $\vec{e}$, **выполняющимся по часовой стрелке.**

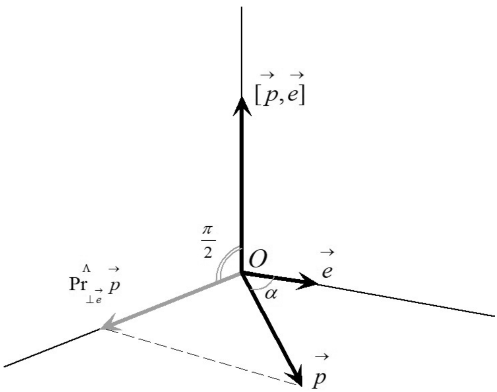


Рис. 2.5. К доказательству леммы 2.4.2

Доказательство.

Проведем две плоскости, одна из которых проходит через точку O — общее начало векторов $\vec{p}$ и $\vec{e}$, перпендикулярно $\vec{e}$, а вторая проходит через векторы $\vec{p}$ и $\vec{e}$.

Ортогональная проекция вектора $\vec{p}$ на плоскость, перпендикулярную $\vec{e}$, будет лежать на линии пересечения построенных плоскостей (см. рис. 2.5), и тогда из определения векторного произведения имеем

$$|\,[\vec{p}, \vec{e}]\,| = |\,\vec{p}\,| \cdot |\,\vec{e}\,| \cdot |\sin\alpha| = |\,\vec{p}\,| \cdot \left|\cos\left(\frac{\pi}{2} - \alpha\right)\right| ,$$

поскольку $|\vec{e}\,| = 1$. Следовательно, в рассматриваемом случае $[\vec{p}, \vec{e}\,] = \hat{\text{Пов}}_{\frac{\pi}{2}, \vec{e}}\left(\hat{\text{Pr}}_{\perp \vec{e}}\, \vec{p}\right)$, где $\hat{\text{Pr}}_{\perp \vec{e}}\, \vec{p}$ обозначает вектор, являющийся результатом ортогонального проектирования вектора $\vec{p}$ на плоскость, перпендикулярную $\vec{e}$.

`Лемма доказана.`

Докажем теперь дистрибутивность векторного произведения.

`Доказательство свойства 3).`

Если $\vec{c} = \vec{o}$, то свойство 3) очевидно. Пусть $\vec{c} \neq \vec{o}$, тогда если обозначить $\vec{e} = \dfrac{\vec{c}}{\left|\vec{c}\right|}$, то в силу утверждений лемм 2.4.1, 2.4.2 и свойства 2.1 из § 2.1 получаем

$$[\vec{a} + \vec{b}, \vec{c}\,] = \left|\vec{c}\right| [\vec{a} + \vec{b}, \frac{\vec{c}}{\left|\vec{c}\right|}] = \hat{\text{Пов}}_{\frac{\pi}{2}, \vec{e}}\left(\hat{\text{Pr}}_{\perp \vec{e}}\left(\vec{a} + \vec{b}\right)\right) =$$

$$= \hat{\text{Пов}}_{\frac{\pi}{2}, \vec{e}}\left(\hat{\text{Pr}}_{\perp \vec{e}}\, \vec{a}\right) + \hat{\text{Пов}}_{\frac{\pi}{2}, \vec{e}}\left(\hat{\text{Pr}}_{\perp \vec{e}}\, \vec{b}\right) =$$

$$= \left|\vec{c}\right| [\,\vec{a}, \frac{\vec{c}}{\left|\vec{c}\right|}] + \left|\vec{c}\right| [\,\vec{b}, \frac{\vec{c}}{\left|\vec{c}\right|}] =$$

$$= [\vec{a}, \vec{c}\,] + [\vec{b}, \vec{c}\,]\,.$$

`Свойство доказано.`

2.5. Выражение векторного произведения в координатах

Пусть задан *правый* базис $\{\vec{g}_1,\, \vec{g}_2,\, \vec{g}_3\}$ (то есть такой, что векторы $\{\vec{g}_1,\, \vec{g}_2,\, \vec{g}_3\}$ образуют *правую* тройку) и пусть в этом базисе векторы $\vec{a}$ и $\vec{b}$ имеют координатные разложения

$$\vec{a} = \xi_1 \vec{g}_1 + \xi_2 \vec{g}_2 + \xi_3 \vec{g}_3 \qquad \text{и} \qquad \vec{b} = \eta_1 \vec{g}_1 + \eta_2 \vec{g}_2 + \eta_3 \vec{g}_3\,.$$

По свойствам 2) и 3) векторного произведения

$$[\vec{a}, \vec{b}\,] = [\,\xi_1 \vec{g}_1 + \xi_2 \vec{g}_2 + \xi_3 \vec{g}_3,\ \eta_1 \vec{g}_1 + \eta_2 \vec{g}_2 + \eta_3 \vec{g}_3\,] =$$

$$
\begin{aligned}
&= \xi_1\eta_1[\vec{g}_1,\vec{g}_1] + \xi_1\eta_2[\vec{g}_1,\vec{g}_2] + \xi_1\eta_3[\vec{g}_1,\vec{g}_3]+\\
&+ \xi_2\eta_1[\vec{g}_2,\vec{g}_1] + \xi_2\eta_2[\vec{g}_2,\vec{g}_2] + \xi_2\eta_3[\vec{g}_2,\vec{g}_3]+\\
&+ \xi_3\eta_1[\vec{g}_3,\vec{g}_1] + \xi_3\eta_2[\vec{g}_3,\vec{g}_2] + \xi_3\eta_3[\vec{g}_3,\vec{g}_3] =\\
&= \sum_{j=1}^{3}\sum_{i=1}^{3}\xi_j\eta_i[\vec{g}_j,\vec{g}_i]\,.
\end{aligned}
$$

Полученную формулу можно упростить до выражения с тремя слагаемыми, если учесть, что

$$
[\vec{g}_j,\vec{g}_j] = \vec{o} \quad \forall j = 1,2,3 \qquad \text{и} \qquad [\vec{g}_j,\vec{g}_i] = -[\vec{g}_i,\vec{g}_j] \quad \forall j,i = 1,2,3\,.
$$

Другими словами, если ввести обозначения

$$
\vec{f}_1 = [\vec{g}_2,\vec{g}_3]\,, \quad \vec{f}_2 = [\vec{g}_3,\vec{g}_1]\,, \quad \vec{f}_3 = [\vec{g}_1,\vec{g}_2]\,,
$$

то мы получим, использовав (после приведения подобных членов) связь определителей квадратных матриц 2-го и 3-го порядков (теорема 1.1.1), легко запоминаемую формулу

$$
\begin{aligned}
[\vec{a},\vec{b}] &= (\xi_2\eta_3 - \xi_3\eta_2)\,\vec{f}_1 - (\xi_1\eta_3 - \xi_3\eta_1)\,\vec{f}_2 + (\xi_1\eta_2 - \xi_2\eta_1)\,\vec{f}_3 =\\
&= \vec{f}_1 \det\begin{Vmatrix} \xi_2 & \xi_3 \\ \eta_2 & \eta_3 \end{Vmatrix} - \vec{f}_2 \det\begin{Vmatrix} \xi_1 & \xi_3 \\ \eta_1 & \eta_3 \end{Vmatrix} + \vec{f}_3 \det\begin{Vmatrix} \xi_1 & \xi_2 \\ \eta_1 & \eta_2 \end{Vmatrix} =\\
&= \det\begin{Vmatrix} \vec{f}_1 & \vec{f}_2 & \vec{f}_3 \\ \xi_1 & \xi_2 & \xi_3 \\ \eta_1 & \eta_2 & \eta_3 \end{Vmatrix}\,.
\end{aligned}
$$

Случай ортонормированного базиса

Пусть исходный базис $\{\vec{e}_1,\, \vec{e}_2,\, \vec{e}_3\}$ *ортонормированный*, образующий *правую* тройку векторов, тогда по определению 2.4.2

$$
\vec{f}_1 = \vec{e}_1\,, \quad \vec{f}_2 = \vec{e}_2\,, \quad \vec{f}_3 = \vec{e}_3\,.
$$

Тогда формула для векторного произведения векторов в *правом ортонормированном* базисе заметно упростится:

$$
[\vec{a},\vec{b}] = \det\begin{Vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ \xi_1 & \xi_2 & \xi_3 \\ \eta_1 & \eta_2 & \eta_3 \end{Vmatrix}\,. \tag{2.5.1}
$$

Из вышеприведенных формул вытекают полезные следствия.

Следствие 2.5.1 **Для того чтобы векторы $\vec{a}$ и $\vec{b}$ были коллинеарны, необходимо и достаточно, чтобы в любом базисе**

$$\det\left\|\begin{matrix} \xi_2 & \xi_3 \\ \eta_2 & \eta_3 \end{matrix}\right\| = \det\left\|\begin{matrix} \xi_1 & \xi_3 \\ \eta_1 & \eta_3 \end{matrix}\right\| = \det\left\|\begin{matrix} \xi_1 & \xi_2 \\ \eta_1 & \eta_2 \end{matrix}\right\| = 0\,,$$

или же (в случае $\vec{b} \neq \vec{o}$) $\quad \dfrac{\xi_1}{\eta_1} = \dfrac{\xi_2}{\eta_2} = \dfrac{\xi_3}{\eta_3}\,.$

Следствие 2.5.2 **В ортонормированном базисе площадь параллелограмма, построенного на векторах $\vec{a}$ и $\vec{b}$, равна**

$$S = \sqrt{\det{}^2\left\|\begin{matrix} \xi_2 & \xi_3 \\ \eta_2 & \eta_3 \end{matrix}\right\| + \det{}^2\left\|\begin{matrix} \xi_1 & \xi_3 \\ \eta_1 & \eta_3 \end{matrix}\right\| + \det{}^2\left\|\begin{matrix} \xi_1 & \xi_2 \\ \eta_1 & \eta_2 \end{matrix}\right\|}\,,$$

а для случая на плоскости $S = \left|\det\left\|\begin{matrix} \xi_2 & \xi_3 \\ \eta_2 & \eta_3 \end{matrix}\right\|\right|\,.$

2.6. Смешанное произведение

Определение 2.6.1	*Смешанным* (или *векторно-скалярным*) произведением векторов $\vec{a}$, $\vec{b}$ и $\vec{c}$ называется число $([\vec{a},\vec{b}\,],\vec{c})\,.$ Смешанное произведение принято обозначать как $(\vec{a},\vec{b},\vec{c})\,.$

Теорема 2.6.1 **Абсолютная величина смешанного произведения $(\vec{a},\vec{b},\vec{c})$ векторов $\vec{a}$, $\vec{b}$ и $\vec{c}$ равна** *объему параллелепипеда*, **построенного на векторах $\vec{a}$, $\vec{b}$ и $\vec{c}$ при совмещении их начал.**

При этом если тройка векторов $\vec{a}$, $\vec{b}$ и $\vec{c}$ некомпланарная и правая, то их смешанное произведение *положительно*, **а если эта тройка левая, то –** *отрицательно*.

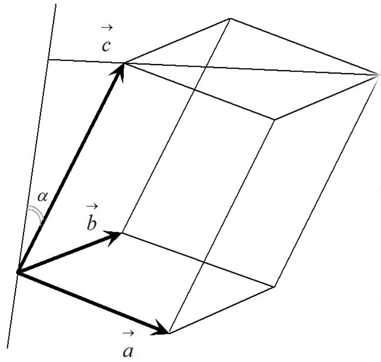


Рис. 2.6. К доказательству теоремы 2.6.1

Доказательство.

Если $\vec{a}$ коллинеарен $\vec{b}$, то утверждение теоремы очевидно.

Пусть $\vec{a}$ неколлинеарен $\vec{b}$, тогда по определению скалярного произведения $(\vec{a},\vec{b},\vec{c}) = \left|[\vec{a},\vec{b}]\right| \cdot \text{Пр}_{[\vec{a},\vec{b}]}\vec{c}$, где $S = \left|[\vec{a},\vec{b}]\right|$ есть площадь параллелограмма, построенного на векторах $\vec{a}$ и $\vec{b}$, а $\left|\text{Пр}_{[\vec{a},\vec{b}]}\vec{c}\right| = \left|\vec{c}\right| |\cos\alpha|$ – высота параллелепипеда с основанием S, откуда $V = \left|(\vec{a},\vec{b},\vec{c})\right|$.

Наконец, по определению 2.6.1

$$(\vec{a},\vec{b},\vec{c}) = ([\vec{a},\vec{b}],\vec{c}) = \left|[\vec{a},\vec{b}]\right| \left|\vec{c}\right| \cos\alpha\,,$$

что и позволяет сделать заключение о знаке смешанного произведения (см. рис. 2.6).

Теорема доказана.

Свойства смешанного произведения

Смешанное произведение:

1. *Коммутативно* при любой циклической перестановке сомножителей и *антикоммутативно* — при не циклической:

$$(\vec{a},\vec{b},\vec{c}) = (\vec{c},\vec{a},\vec{b}) = (\vec{b},\vec{c},\vec{a}) =$$

$$= -(\vec{b}, \vec{a}, \vec{c}) = -(\vec{c}, \vec{b}, \vec{a}) = -(\vec{a}, \vec{c}, \vec{b})\,.$$

2. Допускает вынос скалярного множителя

$$(\lambda\vec{a}, \vec{b}, \vec{c}) = \lambda(\vec{a}, \vec{b}, \vec{c})\,.$$

3. Обладает свойством *дистрибутивности*:

$$(\vec{a_1} + \vec{a_2}, \vec{b}, \vec{c}) = (\vec{a_1}, \vec{b}, \vec{c}) + (\vec{a_2}, \vec{b}, \vec{c})\,.$$

Эти свойства следуют из определения 2.6.1 и теоремы 2.6.1.

Отметим также, что смешанное произведение равно нулю, если среди сомножителей имеется хотя бы одна пара коллинеарных.

2.7. Выражение смешанного произведения в координатах

Пусть задан *правый* базис $\{\vec{g}_1, \vec{g}_2, \vec{g}_3\}$ и пусть в этом базисе векторы $\vec{a}, \vec{b}$ и $\vec{c}$ имеют координатные разложения

$$\vec{a} = \xi_1\vec{g}_1 + \xi_2\vec{g}_2 + \xi_3\vec{g}_3\,, \qquad \vec{b} = \eta_1\vec{g}_1 + \eta_2\vec{g}_2 + \eta_3\vec{g}_3$$

и соответственно $\quad \vec{c} = \kappa_1\vec{g}_1 + \kappa_2\vec{g}_2 + \kappa_3\vec{g}_3\,.$

В §2.5 было показано, что векторное произведение в координатах представимо как

$$[\vec{a}, \vec{b}] = \vec{f}_1 \det\begin{Vmatrix} \xi_2 & \xi_3 \\ \eta_2 & \eta_3 \end{Vmatrix} - \vec{f}_2 \det\begin{Vmatrix} \xi_1 & \xi_3 \\ \eta_1 & \eta_3 \end{Vmatrix} + \vec{f}_3 \det\begin{Vmatrix} \xi_1 & \xi_2 \\ \eta_1 & \eta_2 \end{Vmatrix}\,,$$

где

$$\vec{f}_1 = [\vec{g}_2, \vec{g}_3]\,, \quad \vec{f}_2 = [\vec{g}_3, \vec{g}_1]\,, \quad \vec{f}_3 = [\vec{g}_1, \vec{g}_2]\,.$$

Использовав определение смешанного произведения 2.6.1, нетрудно убедиться, что из последних равенств следуют соотношения

$$(\vec{g}_k, \vec{f}_j) = \begin{cases} (\vec{g}_1, \vec{g}_2, \vec{g}_3) & \text{при} \quad k = j, \\ 0 & \text{при} \quad k \neq j \end{cases} \qquad \forall k, j = 1, 2, 3\,. \tag{2.7.1}$$

Тогда мы получим для смешанного произведения

$$(\vec{a}, \vec{b}, \vec{c}) = ([\vec{a}, \vec{b}], \vec{c}) = \left(\kappa_1 \det\begin{Vmatrix} \xi_2 & \xi_3 \\ \eta_2 & \eta_3 \end{Vmatrix} - \right.$$

$$\left. -\kappa_2 \det\begin{Vmatrix} \xi_1 & \xi_3 \\ \eta_1 & \eta_3 \end{Vmatrix} + \kappa_3 \det\begin{Vmatrix} \xi_1 & \xi_2 \\ \eta_1 & \eta_2 \end{Vmatrix}\right)(\vec{g}_1, \vec{g}_2, \vec{g}_3) =$$

$$= \det \begin{Vmatrix} \xi_1 & \xi_2 & \xi_3 \\ \eta_1 & \eta_2 & \eta_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{Vmatrix} (\vec{g}_1,\, \vec{g}_2,\, \vec{g}_3)\,,$$

поскольку выражение, стоящее в больших круглых скобках, в силу теоремы 1.1.1 есть разложение итогового детерминанта 3-го порядка по его последней строке.

Замечания:

1) Из последней формулы и теоремы 2.6.1 следует (как альтернатива приведенному ранее доказательству) справедливость теоремы 1.6.3.

2) В случае правого ортонормированного базиса мы имеем $(\vec{e}_1,\, \vec{e}_2,\, \vec{e}_3) = 1$. Поэтому в таком базисе

$$(\vec{a}, \vec{b}, \vec{c}) = \det \begin{Vmatrix} \xi_1 & \xi_2 & \xi_3 \\ \eta_1 & \eta_2 & \eta_3 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{Vmatrix} .$$

3) Для тройки векторов $\left\{\vec{f}_1,\, \vec{f}_2,\, \vec{f}_3\right\}$ справедлива

Теорема 2.7.1 **Тройка векторов $\left\{\vec{f}_1,\, \vec{f}_2,\, \vec{f}_3\right\}$ линейно независимая.**

Доказательство.

Для доказательства линейной независимости векторов $\left\{\vec{f}_1,\, \vec{f}_2,\, \vec{f}_3\right\}$ достаточно показать, что из условия

$$\lambda_1 \vec{f}_1 + \lambda_2 \vec{f}_2 + \lambda_3 \vec{f}_3 = \vec{o} \qquad (2.7.2)$$

следует $\lambda_1 = \lambda_2 = \lambda_3 = 0\,.$

Умножив последовательно обе части равенства (2.7.2) скалярно на векторы $\vec{g}_k \quad \forall k = 1, 2, 3\,,$ получим систему трех равенств

$$(\vec{f}_1, \vec{g}_k)\lambda_1 + (\vec{f}_2, \vec{g}_k)\lambda_2 + (\vec{f}_3, \vec{g}_k)\lambda_3 = 0 \quad \forall k = 1, 2, 3\,, \quad (2.7.3)$$

которая при помощи соотношений (2.7.1) упрощается до $(\vec{g}_1, \vec{g}_2, \vec{g}_3)\lambda_k = 0 \ \ \forall k = 1,2,3$.

Векторы $\vec{g}_k \ \forall k = 1,2,3$ линейно независимые как базисные и $(\vec{g}_1, \vec{g}_2, \vec{g}_3) \neq 0 \quad \forall k = 1,2,3$ по теореме 2.6.1. Но тогда $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

Теорема доказана.

Следствие 2.7.1 **Тройка векторов $\left\{\vec{f}_1,\ \vec{f}_2,\ \vec{f}_3\right\}$ образует базис (называемый** *взаимным* **к базису $\{\vec{g}_1,\ \vec{g}_2,\ \vec{g}_3\}$) .**

2.8. Двойное векторное произведение

Определение 2.6.1 *Двойным векторным произведением* векторов $\vec{a},\ \vec{b}$ и $\vec{c}$ называется вектор $[\vec{a},\ [\vec{b},\ \vec{c}]]$.

Для решения ряда задач оказывается полезной

Теорема 2.8.1 **Имеет место равенство**

$$[\vec{a},\ [\vec{b},\ \vec{c}]] = \vec{b}\,(\vec{a},\vec{c}) - \vec{c}\,(\vec{a},\vec{b})\ .$$

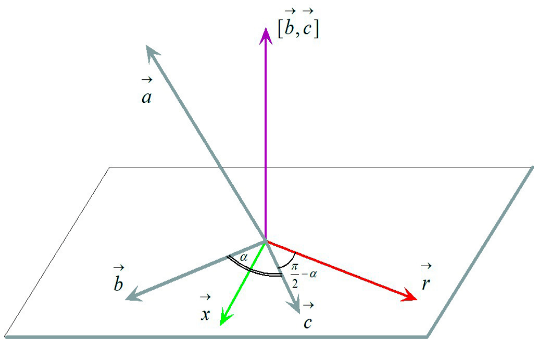


Рис. 2.7. К доказательству теоремы 2.8.1

Доказательство.

Заметим, что если векторы $\vec{a}$, $\vec{b}$ и $\vec{c}$ попарно ортогональны, то доказываемое равенство очевидно, поэтому далее будем предполагать, что числа $(\vec{a},\vec{b})$ и $(\vec{a},\vec{c})$ не равны нулю одновременно.

Обозначим $\vec{x} = [\vec{a},[\vec{b},\vec{c}]]$. По определению векторного произведения вектор $\vec{x}$ ортогонален как вектору $[\vec{b},\vec{c}]$, так и вектору $\vec{a}$.

1. По свойствам смешанного произведения условие ортогональности $\vec{x}$ и $[\vec{b},\vec{c}]$ дает

$$(\vec{x},[\vec{b},\vec{c}]) = (\vec{x},\vec{b},\vec{c}) = 0\,.$$

Это, в свою очередь, означает, что тройка векторов $\vec{x}$, $\vec{b}$, $\vec{c}$ компланарная и в силу леммы 1.4.1 мы имеем равенство $\vec{x} = \lambda\vec{b} + \mu\vec{c}$, где λ и μ – некоторые числа.

2. Из условия $(\vec{x},\vec{a}) = 0$ следует, что $(\lambda\vec{b} + \mu\vec{c},\vec{a}) = 0$ и

$$\lambda(\vec{b},\vec{a}) + \mu(\vec{c},\vec{a}) = 0\,. \tag{2.8.1}$$

Теперь для вычисления вектора $\vec{x}$ достаточно найти одно из чисел λ или μ.

3. Для вычисления значения μ используем вспомогательный вектор $\vec{r}$, удовлетворяющий следующим условиям:

   а) $\vec{r}$, так же как и вектор $\vec{x}$, принадлежит плоскости, проходящей через векторы $\vec{b}$ и $\vec{c}$.

   б) $(\vec{r},\vec{b}) = 0$ и $(\vec{r},\vec{c}) > 0$ (см. рис. 2.7),

   и подсчитаем двумя способами смешанное произведения вида $(\vec{a},\vec{r},[\vec{b},\vec{c}]) = (\vec{a},[\vec{r},[\vec{b},\vec{c}]])$.

4. С одной стороны, по свойствам смешанного произведения и в силу $(\vec{r},\vec{b}) = 0$ имеем

$$(\vec{a},\vec{r},[\vec{b},\vec{c}]) = -(\vec{r},\vec{a},[\vec{b},\vec{c}]) = (\vec{r},[\vec{a},[\vec{b},\vec{c}]]) =$$

$$= -(\vec{r},\vec{x}) = -(\vec{r},\lambda\vec{b} + \mu\vec{c}) =$$

$$= -\lambda(\vec{r},\vec{b}) - \mu(\vec{r},\vec{c}) = -\mu(\vec{r},\vec{c}).$$

5. С другой стороны, вектор $[\vec{r}, [\vec{b}, \vec{c}]]$ сонаправлен с $\vec{b}$, то есть $\exists \kappa > 0$ такое, что $[\vec{r}, [\vec{b}, \vec{c}]] = \kappa \vec{b}$. Поэтому

$$(\vec{a}, \vec{r}, [\vec{b}, \vec{c}]) = \kappa(\vec{a}, \vec{b}).$$

Пусть α угол между векторами $\vec{b}$ и $\vec{c}$, а $\varphi = \frac{\pi}{2}$ – угол между векторами $\vec{r}$ и $[\vec{b}, \vec{c}]$. Значение κ находим при помощи определений векторного и скалярного произведений:

$$\kappa \left|\vec{b}\right| = \left|[\vec{r}, [\vec{b}, \vec{c}]]\right| = \left|\vec{r}\right| \left|\vec{b}\right| \left|\vec{c}\right| \sin\alpha \sin\varphi =$$

$$= \left|\vec{r}\right| \left|\vec{b}\right| \left|\vec{c}\right| \cos\left(\frac{\pi}{2} - \alpha\right) = (\vec{r}, \vec{c}) \left|\vec{b}\right|,$$

поскольку угол между $\vec{r}$ и $[\vec{b}, \vec{c}]$ прямой.

Откуда $\kappa = (\vec{r}, \vec{c})$, что дает

$$(\vec{a}, \vec{r}, [\vec{b}, \vec{c}]) = (\vec{r}, \vec{c})(\vec{a}, \vec{b}).$$

6. Приравняв выражения для $(\vec{a}, \vec{r}, [\vec{b}, \vec{c}])$, полученные в 4 и 5, находим, что $\mu = -(\vec{a}, \vec{b})$ и что, в силу (2.8.1), $\lambda = (\vec{a}, \vec{c})$.

`Теорема доказана.`

Другой способ доказательства этой теоремы приводится в Приложении 4 (см. Прил. 4.5).

2.9. Замечания об инвариантности координатного представления произведений векторов

Операции векторных произведений были введены независимо от координатного представления сомножителей и, значит, независимо от используемого базиса. При этом естественно возникает вопрос о возможности (и соответственно целесообразности) определения операций произведения векторов непосредственно в координатной форме.

В общем случае каждой упорядоченной паре векторов $\vec{a}$ и $\vec{b}$, имеющих в базисе $\left\{\vec{g}_1, \vec{g}_2, \vec{g}_3\right\}$ координатные представления $\left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{matrix} \right\|$ и $\left\| \begin{matrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{matrix} \right\|$, можно поставить в соответствие набор из девяти чисел – попарных произведений координат $\xi_k \eta_j \quad k, j = 1, 2, 3,$ который удобно представляется в виде матрицы

$$\left\| \begin{matrix} \xi_1\eta_1 & \xi_1\eta_2 & \xi_1\eta_3 \\ \xi_2\eta_1 & \xi_2\eta_2 & \xi_2\eta_3 \\ \xi_3\eta_1 & \xi_3\eta_2 & \xi_3\eta_3 \end{matrix} \right\| . \tag{2.9.1}$$

На первый взгляд, зависимость компонент этой матрицы от выбора базиса делает координатный способ введения произведений векторов нецелесообразным, поскольку придется давать их определение для каждого из возможных базисов по отдельности.

Однако было замечено, что существуют некоторые линейные комбинации чисел $\xi_k \eta_j \quad k, j = 1, 2, 3,$ *инвариантные* (то есть не изменяющиеся) при замене базиса, которые, в силу этого свойства, можно было бы принять за определение произведений векторов в координатном представлении.

Покажем в качестве примера, что сумма элементов матрицы 2.9.1, стоящих на ее главной диагонали, не меняется при переходе от одного ортонормированного базиса к другому.

Рассмотрим два *ортонормированных* базиса $\left\{\vec{e}_1, \vec{e}_2, \vec{e}_3\right\}$ и $\left\{\vec{e'}_1, \vec{e'}_2, \vec{e'}_3\right\}$ с матрицей перехода

$$\|S\| = \left\| \begin{matrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{matrix} \right\| .$$

Согласно формулам (1.8.1), в этом случае для базисных векторов имеют место соотношения $\vec{e'}_i = \sum\limits_{p=1}^{3} \sigma_{pi} \vec{e}_p \quad \forall i = 1, 2, 3,$ а для координат соответственно

$$\xi_s = \sum_{i=1}^{3} \sigma_{si} \xi'_i \;\; \forall s = 1, 2, 3 \,, \qquad \eta_s = \sum_{i=1}^{3} \sigma_{si} \eta'_i \;\; \forall s = 1, 2, 3.$$

Пусть $\quad \delta_{ji} = \begin{cases} 1, & j = i, \\ 0, & j \neq i \end{cases}$ — символ Кронекера, тогда из условия ортонормированности базисов $\left\{ \vec{e'}_1, \vec{e'}_2, \vec{e'}_3 \right\}$ и $\left\{ \vec{e}_1, \vec{e}_2, \vec{e}_3 \right\}$ в силу соотношений

$$(\vec{e'}_i, \vec{e'}_t) = \delta_{it} \ \ \forall i, t = 1, 2, 3 \qquad \text{и} \qquad (\vec{e}_s, \vec{e}_p) = \delta_{sp} \ \ \forall s, p = 1, 2, 3$$

имеем

$$(\vec{e'}_i, \vec{e'}_t) = \delta_{it} = (\sum_{s=1}^{3} \sigma_{si} \vec{e}_s, \sum_{p=1}^{3} \sigma_{pt} \vec{e}_p) = \sum_{s=1}^{3} \sum_{p=1}^{3} \sigma_{si} \sigma_{pt} (\vec{e}_s, \vec{e}_p) =$$

$$= \sum_{s=1}^{3} \sum_{p=1}^{3} \sigma_{si} \sigma_{pt} \delta_{sp} = \sum_{s=1}^{3} \sigma_{si} \sigma_{st} \qquad \forall i, t = 1, 2, 3\,.$$

Отметим, что вытекающие из последних соотношений равенства

$$\sum_{s=1}^{3} \sigma_{si} \sigma_{st} = \delta_{it} \qquad \forall i, t = 1, 2, 3$$

являются *свойством* матрицы перехода $\|S\|$ от одного *ортонормированного* базиса к другому.

Теперь, используя это свойство и определение символа Кронекера, найдем выражение для линейной комбинации $\xi_1 \eta_1 + \xi_2 \eta_2 + \xi_3 \eta_3$ в базисе $\left\{ \vec{e'}_1, \vec{e'}_2, \vec{e'}_3 \right\}$ По формулам перехода имеем

$$\sum_{i=1}^{3} \xi_i \eta_i = \sum_{i=1}^{3} \left(\sum_{s=1}^{3} \sigma_{is} \xi'_s \right) \left(\sum_{t=1}^{3} \sigma_{it} \eta'_t \right) = \sum_{s=1}^{3} \sum_{t=1}^{3} \xi'_s \eta'_t \sum_{i=1}^{3} \sigma_{is} \sigma_{it} =$$

$$= \sum_{s=1}^{3} \sum_{t=1}^{3} \xi'_s \eta'_t \delta_{st} = \sum_{t=1}^{3} \xi'_s \eta'_t\,.$$

Полученное равенство доказывает, что числа $\xi_1 \eta_1 + \xi_2 \eta_2 + \xi_3 \eta_3$ не меняются при замене одного ортонормированного базиса другим. Поэтому такая сумма может быть принята за определение *скалярного произведения* векторов $\vec{a}$ и $\vec{b}$, имеющих в любом ортонормированном базисе координатные представления вида

$$\left\| \vec{a} \right\|_e = \left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{matrix} \right\| \qquad \text{и} \qquad \left\| \vec{b} \right\|_e = \left\| \begin{matrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{matrix} \right\|.$$

Используя аналогичные рассуждения, можно показать (проверьте это самостоятельно), что при переходе от одного ортонормированного базиса к другому ортонормированному базису инвариантными также оказываются и линейные комбинации вида $\left\{ \begin{array}{c} \xi_2\eta_3 - \xi_3\eta_2 \\ \xi_3\eta_1 - \xi_1\eta_3 \\ \xi_1\eta_2 - \xi_2\eta_1 \end{array} \right\}$.

Это означает, что инвариантом будет и вектор $\vec{c}$, имеющий координатное представление

$$\left\|\vec{c}\right\|_e = \left\| \begin{array}{c} \xi_2\eta_3 - \xi_3\eta_2 \\ \xi_3\eta_1 - \xi_1\eta_3 \\ \xi_1\eta_2 - \xi_2\eta_1 \end{array} \right\| ,$$

которое можно принять за координатную форму определения *векторного произведения* векторов $\vec{a}$ и $\vec{b}$ в ортонормированном базисе. Сравните это определение с формулой (2.5.1).

3. Прямая и плоскость

Как было показано, с помощью системы координат можно установить взаимно однозначное соответствие между множеством *точек пространства* и множеством их *радиусов-векторов*. Это в свою очередь позволяет свести исследование свойств линий, фигур, поверхностей или тел к изучению множеств радиусов-векторов точек, образующих рассматриваемые геометрические объекты.

Глава 3 посвящена методам описания и исследования свойств простейших геометрических объектов — *прямой* и *плоскости* – средствами векторной алгебры.

В главах 3, 4 и 5 настоящего пособия координата точки по оси $\left\{ O, \vec{g}_1 \right\}$ будет обозначаться через x, координата по оси $\left\{ O, \vec{g}_2 \right\}$ — через y и координата по оси $\left\{ O, \vec{g}_3 \right\}$ — через z. Кроме того, будут использоваться общепринятые форматы записи уравнений.

3.1. Прямая на плоскости

Пусть на плоскости дана система координат $\left\{ O, \vec{g}_1, \vec{g}_2 \right\}$ и задана прямая L, проходящая через точку, имеющую радиус-вектор $\vec{r}_0$, с лежащим на ней *ненулевым* вектором $\vec{a}$.

Определение 3.1.1	Ненулевой вектор $\vec{a}$ называется *направляющим вектором* прямой L.

Теорема 3.1.1 **Множество радиусов-векторов точек прямой L представимо в виде $\vec{r} = \vec{r}_0 + \tau\vec{a}$, где τ – произвольный вещественный параметр.**

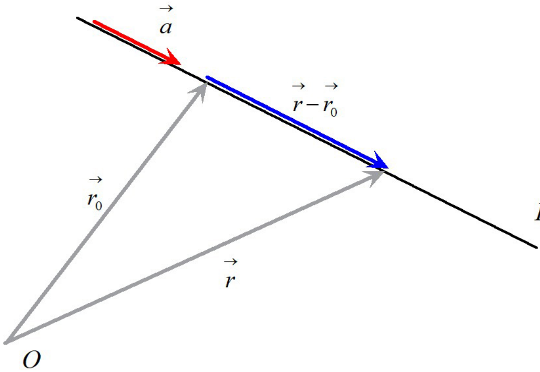


Рис. 3.1. К доказательству теоремы 3.1.1

Доказательство.

Пусть $\vec{r}$ — некоторая точка на прямой L . Ненулевой вектор $\vec{a}$ коллинеарен вектору $\vec{r}-\vec{r}_0$. Поэтому справедливо равенство $\vec{r}-\vec{r}_0=\tau\vec{a}$ (рис. 3.1).
Тогда
$$\vec{r}=\vec{r}_0+\tau\vec{a} \qquad \forall\tau\in(-\infty,+\infty)\,.$$

Теорема доказана.

Найдем теперь координатное представление множества радиусов-векторов точек прямой L. Пусть $\left\|\vec{r}\right\|_g=\left\|\begin{matrix} x \\ y \end{matrix}\right\|$, $\left\|\vec{r}_0\right\|_g=\left\|\begin{matrix} x_0 \\ y_0 \end{matrix}\right\|$ и $\left\|\vec{a}\right\|_g=\left\|\begin{matrix} a_x \\ a_y \end{matrix}\right\|$, тогда будут справедливы следующие теоремы.

Теорема 3.1.2 **Всякая прямая в любой декартовой системе координат может быть задана уравнением вида**
$$Ax+By+C=0\,,\quad |A|+|B|>0\,.$$

Доказательство.

Уравнение $\vec{r} = \vec{r}_0 + \tau\vec{a}$ в силу леммы 1.4.1 означает, что векторы $\vec{r} - \vec{r}_0$ и $\vec{a}$ линейно зависимые.

Условие же линейной зависимости этих векторов в координатной форме имеет вид (теорема 1.4.2):

$$\det \left\| \begin{matrix} x - x_0 & y - y_0 \\ a_x & a_y \end{matrix} \right\| = 0$$

или $a_y(x - x_0) - a_x(y - y_0) = 0$.

Если ввести обозначения

$$A = a_y, \quad B = -a_x, \quad C = -a_y x_0 + a_x y_0 ,$$

то мы получим $Ax + By + C = 0$. При этом неравенство $|A| + |B| > 0$ будет следовать из $\vec{a} \neq \vec{o} \Longleftrightarrow |a_x| + |a_y| > 0$.

Теорема доказана.

Теорема 3.1.3 **Всякое уравнение вида**

$$Ax + By + C = 0 , \quad |A| + |B| > 0$$

в любой декартовой системе координат задает некоторую прямую.

Доказательство.

Пусть дано уравнение вида

$$Ax + By + C = 0 , \quad |A| + |B| > 0 .$$

В этом случае всегда возможно подобрать пару чисел x_0 и y_0 так, чтобы $Ax_0 + By_0 + C = 0$. Почленное вычитание этого равенства из исходного дает при $a_x = -B, \quad a_y = A$

$$a_y(x - x_0) - a_x(y - y_0) = 0 .$$

Возьмем точку

$$\left\| \vec{r}_0 \right\|_g = \left\| \begin{matrix} x_0 \\ y_0 \end{matrix} \right\| \quad \text{и вектор} \quad \left\| \vec{a} \right\|_g = \left\| \begin{matrix} a_x \\ a_y \end{matrix} \right\| .$$

Из теоремы 3.1.2 следует, что прямая, проходящая через точку $\vec{r}_0$ в направлении вектора $\vec{a} \neq \vec{o}$, имеет уравнение вида

$$a_y(x - x_0) - a_x(y - y_0) = 0\,.$$

Следовательно, уравнение, указанное в формулировке теоремы, есть уравнение прямой.

Теорема доказана.

Замечание: из теорем 3.1.2–3.1.3 следует, что каждое линейное уравнение в декартовой системе координат на плоскости задает некоторую конкретную прямую, но, с другой стороны, конкретная прямая на плоскости может быть задана *бесчисленным множеством* линейных уравнений, и естественно возникает вопрос: при каких условиях два разных линейных уравнения задают одну и ту же прямую? Ответ на этот вопрос дает

Теорема 3.1.4 **Для того чтобы уравнения**

$$\begin{array}{ll} A_1x + B_1y + C_1 = 0\,, & |A_1| + |B_1| > 0 \quad \textbf{и} \\ A_2x + B_2y + C_2 = 0\,, & |A_2| + |B_2| > 0 \end{array}$$

являлись уравнениями одной и той же прямой, необходимо и достаточно, чтобы $\exists\lambda \neq 0$ такое, что

$$\left[\begin{array}{lll} A_1 = \lambda A_2 & B_1 = \lambda B_2 & C_1 = \lambda C_2, \\ A_2 = \lambda A_1 & B_2 = \lambda B_1 & C_2 = \lambda C_1. \end{array}\right.$$

Доказательство достаточности.

Пусть коэффициенты уравнений пропорциональны, и имеет место равенство $A_2x + B_2y + C_2 = 0$.
Тогда

$$0 = A_2x + B_2y + C_2 = \frac{1}{\lambda}(A_1x + B_1y + C_1)\ ,$$

но поскольку $\lambda \neq 0$, то $A_1x + B_1y + C_1 = 0$.

Аналогично из равенства $A_1x + B_1y + C_1 = 0$ следует, что и $A_2x + B_2y + C_2 = 0$.

Доказательство необходимости.

Пусть уравнения

$$A_1x + B_1y + C_1 = 0\,, \quad |A_1| + |B_1| > 0 \quad \text{и}$$
$$A_2x + B_2y + C_2 = 0\,, \quad |A_2| + |B_2| > 0$$

задают в некоторой декартовой системе координат одну и ту же прямую. Тогда их направляющие векторы коллинеарны, т.е. $\exists\lambda \neq 0$ такое, что $A_1 = \lambda A_2$, $B_1 = \lambda B_2$.

С другой стороны, из равносильности уравнений

$$\lambda A_2x + \lambda B_2y + C_1 = 0\,, \quad \text{и}$$
$$A_2x + B_2y + C_2 = 0$$

следует также, что и $C_1 = \lambda C_2$.

Теорема доказана.

Замечание: уравнение прямой не в любой системе координат является алгебраическим уравнением первой степени. Например, в *полярной* системе координат (как будет показано в § 4.6) оно может иметь вид $\rho = P\sec(\varphi + \varphi_0)$.

3.2. Способы задания прямой на плоскости

В общей декартовой системе координат $\left\{O, \vec{g}_1, \vec{g}_2\right\}$ существуют различные формы задания прямой на плоскости. Рассмотрим основные из них.

1°. Уравнение прямой, проходящей через две несовпадающие точки $\vec{r}_1$ и $\vec{r}_2$: с $\left\|\vec{r}_1\right\|_g = \left\|\begin{matrix} x_1 \\ y_1 \end{matrix}\right\|$ и $\left\|\vec{r}_2\right\|_g = \left\|\begin{matrix} x_2 \\ y_2 \end{matrix}\right\|$

Поскольку в данном случае в качестве направляющего вектора $\vec{a}$ можно взять $\vec{r}_2 - \vec{r}_1$, а в качестве $\vec{r}_0$ — вектор $\vec{r}_1$, то указанная в теореме 3.1.1 параметрическая форма уравнения прямой примет вид $\vec{r} = \vec{r}_1 + \tau\left(\vec{r}_2 - \vec{r}_1\right)$ или

$$\vec{r} = (1-\tau)\vec{r}_1 + \tau\vec{r}_2 \qquad \forall\tau \in (-\infty, +\infty)\,.$$

Координатный вид полученного векторного уравнения прямой на плоскости можно найти, исключив параметр τ.

В результате получится один из трех следующих случаев:

$$\left[\begin{array}{rcll} \dfrac{x-x_1}{x_2-x_1} & = & \dfrac{y-y_1}{y_2-y_1}, & \text{если} \begin{cases} x_1 \neq x_2\,, \\ y_1 \neq y_2\,, \end{cases} \\ y & = & y_1 \quad \forall x, & \text{если } y_1 = y_2, \\ x & = & x_1 \quad \forall y, & \text{если } x_1 = x_2. \end{array}\right.$$

Проверьте самостоятельно, что эти три случая могут быть описаны одним условием:

$$\det \begin{Vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{Vmatrix} = 0$$

и что справедливо

Следствие 3.2.1 **Три точки $\vec{r}_1 = \begin{Vmatrix} x_1 \\ y_1 \end{Vmatrix}$, $\vec{r}_2 = \begin{Vmatrix} x_2 \\ y_2 \end{Vmatrix}$ и $\vec{r}_3 = \begin{Vmatrix} x_3 \\ y_3 \end{Vmatrix}$ лежат на одной прямой тогда и только тогда, когда их координаты связаны соотношением**

$$\det \begin{Vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{Vmatrix} = 0\,.$$

2°. Векторное уравнение прямой, проходящей через точку $\vec{r}_0 = \begin{Vmatrix} x_0 \\ y_0 \end{Vmatrix}$, перпендикулярно заданному ненулевому вектору $\vec{n} = \begin{Vmatrix} n_x \\ n_y \end{Vmatrix}$.

Здесь в качестве направляющего вектора $\vec{a}$ также можно взять $\vec{r} - \vec{r}_0 = \begin{Vmatrix} x - x_0 \\ y - y_0 \end{Vmatrix}$, где радиус-вектор произвольной точки на прямой L (см. рис. 3.2) $\vec{r} = \begin{Vmatrix} x \\ y \end{Vmatrix}$.

Этот направляющий вектор ортогонален любому ненулевому вектору $\vec{n}$, который перпендикулярен L, и потому уравнение прямой может быть записано в виде

$$(\vec{n}, \vec{r} - \vec{r}_0) = 0 \qquad \text{или же} \qquad (\vec{n}, \vec{r}) = d\,.$$

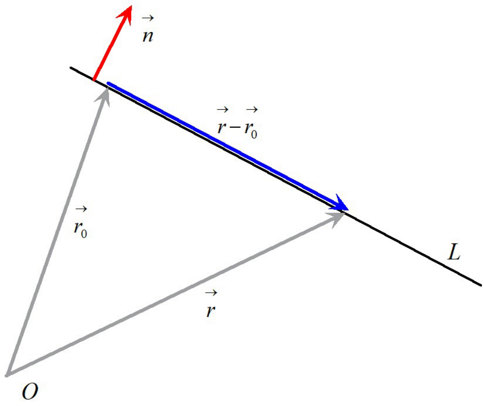


Рис. 3.2. К выводу уравнения $(\vec{n},\vec{r}) = d$

В последней формуле значение числового параметра d определяется по $\vec{r}_0$ и $\vec{n}$ из равенства $d = (\vec{n},\vec{r}_0)$.

При обратном переходе в качестве вектора $\vec{r}_0$ можно использовать вектор $\vec{r}_0 = \dfrac{d}{(\vec{n},\ \vec{n})}\vec{n}$.

Справедливость последней формулы проверьте самостоятельно.

Координатный вид данной формы уравнения прямой на плоскости будет зависеть от выбранной декартовой системы координат. Например, в *ортонормированной* системе координат $\left\{ O,\ \vec{e}_1,\ \vec{e}_2 \right\}$ уравнения

$$(\vec{n},\vec{r}-\vec{r}_0) = 0 \qquad \text{или} \qquad (\vec{n},\vec{r}) = d$$

будут иметь соответственно вид

$$n_x(x - x_0) + n_y(y - y_0) = 0$$

или $\quad n_x x + n_y y = d, \quad$ где $\quad d = n_x x_0 + n_y y_0$.

Сравнивая последнюю запись с общим видом уравнения прямой $Ax + By + C = 0$, приходим к заключению, что в *ортонормированной* системе координат вектор $\vec{n}$, для которого $\left\|\vec{n}\right\|_e = \left\| \begin{matrix} A \\ B \end{matrix} \right\|$, будет ортогонален этой прямой.

Определение 3.2.1 Ненулевой вектор $\vec{n}$ называется *нормальным* вектором прямой L.

Замечание о линейных неравенствах

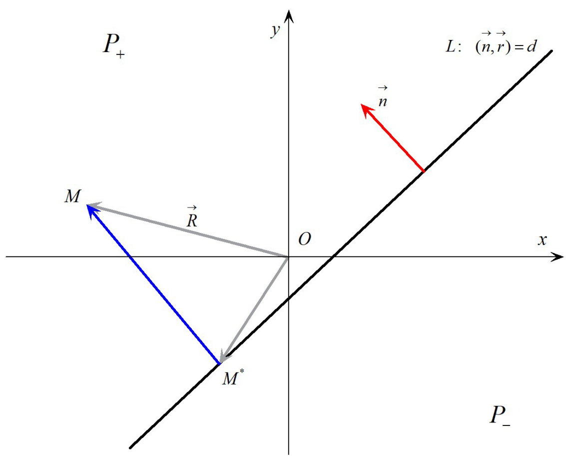


Рис. 3.3. Геометрическая интерпретация линейного неравенства

Аналогично тому, как линейное уравнение задает на плоскости прямую, линейное неравенство

$$Ax + By + C > 0, \quad |A| + |B| > 0$$

определяет часть плоскости (множество точек, координаты которых x и y удовлетворяют данному неравенству), ограниченную прямой

$$Ax + By + C = 0, \quad |A| + |B| > 0\,.$$

Покажем справедливость данного утверждения для случая, когда прямая $L: \quad (\vec{n}, \vec{r}) = d$ выделяет на координатной плоскости P два множества точек, обозначаемых как P_+ и P_- (см. рис. 3.3), согласно следующему определению.

Определение 3.2.2 Будем говорить, что точка M с радиусом-вектором $\vec{R}$ принадлежит полуплоскости P_+ (соответственно P_-) , если существует $\lambda > 0$ (соответственно $\lambda < 0$) такое, что $\overrightarrow{M^*M} = \lambda\vec{n}$, где точка M^* есть ортогональная проекция точки M на прямую L .

Имеет место

Теорема 3.2.1 **Для того чтобы $M \in P_+$, необходимо и достаточно выполнения неравенства** $(\vec{n}, \vec{R}) > d.$

Доказательство необходимости.

Пусть $M \in P_+$, то есть $\exists\lambda > 0: \overrightarrow{M^*M} = \lambda\vec{n}$. Оценим величину $(\vec{n}, \vec{R})$.
Поскольку $M^* \in L$, то $(\vec{n}, \overrightarrow{OM^*}) = d$. Тогда

$$(\vec{n}, \vec{R}) = (\vec{n}, \overrightarrow{OM^*} + \overrightarrow{M^*M}) = d + \lambda(\vec{n}, \vec{n}) > d$$

в силу положительности λ.

Доказательство достаточности.

Пусть $(\vec{n}, \vec{R}) > d$ и $\overrightarrow{M^*M} = \lambda\vec{n}$. Тогда из $(\vec{n}, \overrightarrow{OM^*}) = d$ получаем

$$(\vec{n}, \vec{R}) = (\vec{n}, \overrightarrow{OM^*} + \overrightarrow{M^*M}) = (\vec{n}, \overrightarrow{OM^*}) + (\vec{n}, \overrightarrow{M^*M}) =$$

$$= d + \lambda(\vec{n}, \vec{n}) > d \qquad \Longrightarrow \qquad \lambda(\vec{n}, \vec{n}) > 0\,.$$

Поскольку $\vec{n} \neq \vec{o}$, то $\lambda > 0$ и, значит, $M \in P_+$

Теорема доказана.

Рассмотренные способы координатно-векторного описания точек и прямых на плоскости позволяют решать достаточно сложные геометрические задачи, что иллюстрирует

`Задача 3.2.1` *Дана декартова система координат на плоскости* $\{O, \vec{g}_1, \vec{g}_2\}$ *и прямая* L *с уравнением* $(\vec{n}, \vec{r} - \vec{r}_0) = 0$. *Найти расстояние до этой прямой от точки, радиус-вектор которой* $\vec{R}$ *имеет координатное представление* $\left\|\vec{R}\right\|_g = \left\| \begin{matrix} X \\ Y \end{matrix} \right\|$.

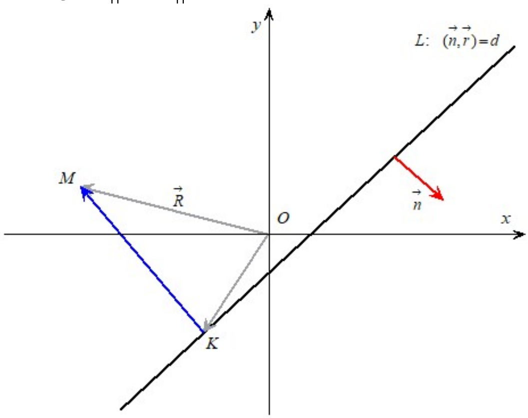


Рис. 3.4. К решению задачи 3.2.1

`Решение.` Пусть $\vec{KM} = \lambda\vec{n}$. Тогда $\vec{R} = \vec{OK} + \lambda\vec{n}$ (см. рис. 3.4). Искомое расстояние в этом случае есть длина вектора $\vec{KM}$.

Точка K принадлежит L, поэтому $(\vec{n}, \vec{OK} - \vec{r}_0) = 0$ или $(\vec{n}, \vec{R} - \lambda\vec{n} - \vec{r}_0) = 0$. Откуда $\lambda = \dfrac{(\vec{n}, \vec{R} - \vec{r}_0)}{\left|\vec{n}\right|^2}$.

Для искомого расстояния имеем

$$\left|\vec{KM}\right| = |\lambda| \left|\vec{n}\right| = \frac{\left|(\vec{n}, \vec{R} - \vec{r}_0)\right|}{\left|\vec{n}\right|}.$$

В заключение найдем ответ в координатной форме. Пусть система координат ортонормированная. Тогда для уравнения $Ax+By+C=0$, $|A|+|B|>0$, как было показано, вектор $\vec{n}=\left\|\begin{array}{c}A\\B\end{array}\right\|$ перпендикулярен прямой L . Поэтому

$$\left|\vec{KM}\right|=\frac{|A(X-x_0)+B(Y-y_0)|}{\sqrt{A^2+B^2}}.$$

Принимая во внимание, что точка $\vec{r}_0$ лежит на прямой L и, следовательно, $Ax_0+By_0+C=0$, окончательный ответ можно записать в виде

$$\left|\vec{KM}\right|=\frac{|AX+BY+C|}{\sqrt{A^2+B^2}}.$$

`Решение получено.`

Координатный метод позволяет строить аналитические описания не только отдельных точек и прямых, но и более сложных объектов.

Примером может служить описание треугольника с вершинами в точках $A(6;0)$, $B(3;3)$ и $C(3;-3)$, являющееся системой трех линейных неравенств

$$\left\{\begin{array}{rcrcr}-x&+&3y&\leq&6\,,\\6x&-&y&\leq&15\,,\\3x&+&8y&\geq&-18\,.\end{array}\right.$$

Чтобы построить эту систему, вначале необходимо найти уравнения сторон треугольника ABC, воспользовавшись, например, результатом пункта 1° этого параграфа.

Затем, применив определение 3.2.2, подберем подходящие знаки неравенств. Для этого воспользуемся тем, что при циклическом обходе по вершинам треугольника против часовой стрелки его внутренние точки должны принадлежать множествам P_- для всех трех границ одновременно.

Наконец, учтем, что неравенства системы нестрогие, поскольку граничные точки должны принадлежать треугольнику.

Другим, полезным для практики, примером такого множества является пучок прямых.

Определение 3.2.3	*Пучком прямых* на плоскости называется совокупность всех прямых, проходящих через некоторую заданную точку, именуемую *вершиной пучка.*

Теорема 3.2.2

Пусть точка, общая для всех прямых пучка, является точкой пересечения непараллельных прямых $A_1x + B_1y + C_1 = 0$ и $A_2x + B_2y + C_2 = 0$. Тогда

1) для любой прямой пучка найдется пара не равных нулю одновременно чисел α и β, таких, что

$$\alpha(A_1x + B_1y + C_1) + \beta(A_2x + B_2y + C_2) = 0$$

есть уравнение этой прямой;

2) при любых, не равных нулю одновременно α и β, уравнение

$$\alpha(A_1x + B_1y + C_1) + \beta(A_2x + B_2y + C_2) = 0$$

есть уравнение некоторой прямой данного пучка.

Доказательство.

1. Возьмем некоторую точку $\vec{r}^{\,*} = \left\| \begin{matrix} x^* \\ y^* \end{matrix} \right\|$, не совпадающую с вершиной пучка, и примем в качестве параметров числа $\alpha = A_1x^* + B_1y^* + C_1$ и $\beta = -(A_2x^* + B_2y^* + C_2)$. Заметим, что при этом $|\alpha| + |\beta| > 0$, поскольку точка $\vec{r}^{\,*}$ не принадлежит данным прямым одновременно.

   Кроме того, прямая

$$(A_2x^* + B_2y^* + C_2)(A_1x + B_1y + C_1) +$$

$$+(A_1x^* + B_1y^* + C_1)(A_2x + B_2y + C_2) = 0$$

   проходит как через точку $\vec{r}^{\,*}$, так и через вершину пучка и, следовательно, принадлежит пучку.

2. Пусть $A_1x+B_1y+C_1=0$ и $A_2x+B_2y+C_2=0$ — пара пересекающихся прямых из рассматриваемого пучка, тогда очевидно, что

$$\alpha(A_1x+B_1y+C_1)+\beta(A_2x+B_2y+C_2)=0\,.$$

При этом уравнение

$$(\alpha A_1+\beta A_2)x+(\alpha B_1+\beta B_2)y+(\alpha C_1+\beta C_2)=0$$

является уравнением прямой, поскольку из $|A_1|+|B_1|>0$, $|A_2|+|B_2|>0$ и $|\alpha|+|\beta|>0$ следует, что

$$|\alpha A_1+\beta A_2|+|\alpha B_1+\beta B_2|>0\,.$$

Действительно, допустим противное:

$$\left\{\begin{array}{l}\alpha A_1+\beta A_2=0,\\ \alpha B_1+\beta B_2=0.\end{array}\right. \tag{3.2.1}$$

Прямые $A_1x+B_1y+C_1=0$ и $A_2x+B_2y+C_2=0$ по построению имеют, по крайней мере, одну общую точку. Поэтому они либо совпадают, либо пересекаются. По теореме 3.1.4 они совпадают тогда и только тогда, когда существует $\lambda\neq 0$, для которого $A_1=\lambda A_2$ и $B_1=\lambda B_2$.

А последние два равенства по теореме 1.6.2 равносильны условию

$$\det\left\|\begin{array}{cc}A_1 & A_1\\ B_2 & B_2\end{array}\right\|=0\,.$$

Но в рассматриваемом случае прямые пересекаются, поэтому $\det\left\|\begin{array}{cc}A_1 & A_1\\ B_2 & B_2\end{array}\right\|\neq 0$, и в силу теоремы 1.1.2 система линейных уравнений 3.2.1 может иметь лишь единственное решение.

С другой стороны, очевидно, что эта система имеет тривиальное решение $\alpha=\beta=0$, что противоречит неравенству $|\alpha|+|\beta|>0$. Следовательно,

$$|\alpha A_1+\beta A_2|+|\alpha B_1+\beta B_2|>0\,.$$

`Теорема доказана.`

Определение 3.2.4	Уравнение $$\alpha(A_1x + B_1y + C_1) + \beta(A_2x + B_2y + C_2) = 0\,,$$ где $\|\alpha\| + \|\beta\| > 0$, называется *уравнением пучка прямых на плоскости.*

3.3. Плоскость в пространстве

Пусть дана система координат $\left\{ O, \vec{g}_1, \vec{g}_2, \vec{g}_3 \right\}$ и выбрана плоскость S, проходящая через точку, имеющую радиус-вектор $\vec{r}_0$, и лежащими на S *неколлинеарными* векторами $\vec{p}$ и $\vec{q}$.

Определение 3.3.1	Неколлинеарные векторы $\vec{p}$ и $\vec{q}$ называются *направляющими векторами* плоскости S.

Теорема 3.3.1 **Множество радиусов-векторов точек плоскости S представимо в виде $\vec{r} = \vec{r}_0 + \varphi\vec{p} + \theta\vec{q}$, где φ и θ — произвольные вещественные параметры.**

Доказательство.

Пусть $\vec{r}$ — некоторая точка на плоскости S .
Тройка векторов $\vec{p}$, $\vec{q}$ и $\vec{r} - \vec{r}_0$ очевидно компланарная. Поэтому вектор $\vec{r}-\vec{r}_0$ может быть представлен в виде линейной комбинации векторов $\vec{p}$ и $\vec{q}$:

$$\vec{r} - \vec{r}_0 = \varphi\vec{p} + \theta\vec{q} \qquad \forall\varphi, \theta \in (-\infty, +\infty)\,.$$

И, следовательно, векторно-параметрическое уравнение плоскости можно записать в виде

$$\vec{r} = \vec{r}_0 + \varphi\vec{p} + \theta\vec{q} \qquad \forall\varphi, \theta \in (-\infty, +\infty)\,.$$

Теорема доказана.

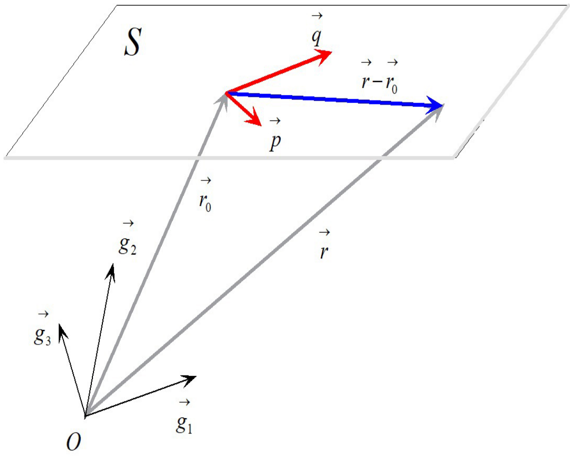


Рис. 3.5. К доказательству теоремы 3.3.1

Иными словами, каждая *упорядоченная* пара чисел φ и θ определяет некоторую точку плоскости S , а для каждой точки этой плоскости существует единственная упорядоченная пара чисел φ и θ, определяющая ее радиус-вектор.

Найдем теперь координатное представление множества радиусов-векторов точек плоскости S . Введем обозначения $\left\|\vec{r}\right\|_g = \left\|\begin{matrix} x \\ y \\ z \end{matrix}\right\|$, $\left\|\vec{r}_0\right\|_g = \left\|\begin{matrix} x_0 \\ y_0 \\ z_0 \end{matrix}\right\|$, $\left\|\vec{p}\right\|_g = \left\|\begin{matrix} p_x \\ p_y \\ p_z \end{matrix}\right\|$ и $\left\|\vec{q}\right\|_g = \left\|\begin{matrix} q_x \\ q_y \\ q_z \end{matrix}\right\|$, тогда будут справедливы следующие теоремы.

Теорема 3.3.2 **Всякая плоскость в любой декартовой системе координат может быть задана уравнением вида**

$$Ax + By + Cz + D = 0\,, \quad |A| + |B| + |C| > 0\,.$$

Доказательство.

Уравнение $\vec{r} = \vec{r}_0 + \varphi\vec{p} + \theta\vec{q}$, в силу леммы 1.4.1, означает, что три вектора $\vec{r} - \vec{r}_0$, $\vec{p}$ и $\vec{q}$ линейно зависимые.

Условие же линейной зависимости этих векторов в координатной форме имеет вид (теорема 1.4.3):

$$\det\left\|\begin{array}{ccc} x-x_0 & y-y_0 & z-z_0 \\ p_x & p_y & p_z \\ q_x & q_y & q_z \end{array}\right\| = 0\,. \qquad (3.3.1)$$

Тогда, разложив детерминант по первой строке, получим

$$A(x-x_0)+B(y-y_0)+C(z-z_0)=0\,,$$

что в итоге дает $\quad Ax+By+Cz+D=0\,.$

Действительно, если ввести обозначения

$$A=\det\left\|\begin{array}{cc} p_y & p_z \\ q_y & q_z \end{array}\right\|, \qquad B=-\det\left\|\begin{array}{cc} p_x & p_z \\ q_x & q_z \end{array}\right\|,$$

$$C=\det\left\|\begin{array}{cc} p_x & p_y \\ q_x & q_y \end{array}\right\|$$

и $D=-Ax_0-By_0-Cz_0$, то получается $Ax+By+Cz+D=0.$

При этом неравенство $|A|+|B|+|C|>0$ будет следовать из условия неколлинеарности векторов $\vec{p}$ и $\vec{q}$, записаного в форме, указанной в следствии 2.5.1.

Теорема доказана.

Теорема 3.3.3 **Всякое уравнение вида**

$$Ax+By+Cz+D=0\,, \qquad |A|+|B|+|C|>0\,,$$

в любой декартовой системе координат есть уравнение некоторой плоскости.

Доказательство.

Непосредственной проверкой убеждаемся, что уравнение

$$Ax+By+Cz+D=0\,, \quad |A|+|B|+|C|>0\,,$$

при $C\neq 0$ может быть записано в виде

$$\det \left\| \begin{array}{ccc} x+\dfrac{DA}{A^2+B^2+C^2} & y+\dfrac{DB}{A^2+B^2+C^2} & z+\dfrac{DC}{A^2+B^2+C^2} \\ 0 & -C & B \\ C & 0 & -A \end{array} \right\| = 0, \tag{3.3.2}$$

а при $C=0$ в виде

$$\det \left\| \begin{array}{ccc} x+\dfrac{DA}{A^2+B^2} & y+\dfrac{DB}{A^2+B^2} & z+0 \\ -B & A & 0 \\ 0 & 0 & 1 \end{array} \right\| = 0. \tag{3.3.3}$$

Тогда в любой декартовой системе координат в качестве векторов $\vec{p}$ и $\vec{q}$ при $C \neq 0$ можно взять

$$\left\|\vec{p}\right\|_g = \left\| \begin{array}{r} 0 \\ -C \\ B \end{array} \right\| \qquad \text{и} \qquad \left\|\vec{q}\right\|_g = \left\| \begin{array}{r} C \\ 0 \\ -A \end{array} \right\| ,$$

а в случае $C=0$

$$\left\|\vec{p}\right\|_g = \left\| \begin{array}{r} -B \\ A \\ 0 \end{array} \right\| \qquad \text{и} \qquad \left\|\vec{q}\right\|_g = \left\| \begin{array}{c} 0 \\ 0 \\ 1 \end{array} \right\| ,$$

поскольку из сопоставления уравнений (3.3.2) и (3.3.3) с формулой (3.3.1) следует, что как (3.3.2), так и (3.3.3) суть уравнения плоскости, проходящей через заданную точку параллельно двум неколлинеарным векторам.

Теорема доказана.

Замечание: уравнение (3.3.1) может быть также получено, если условие компланарности векторов $\vec{r}-\vec{r}_0$, $\vec{p}$ и $\vec{q}$ записать в виде равенства нулю их смешанного произведения:

$$(\vec{r}-\vec{r}_0, \vec{p}, \vec{q}) = 0\,.$$

Задача 3.3.1 *В системе координат* $\{O, \vec{g}_1, \vec{g}_2, \vec{g}_3\}$ *составить уравнение плоскости, проходящей через три заданные, не лежащие на одной прямой точки:*

$$\|\vec{r_1}\|_g = \left\| \begin{matrix} x_1 \\ y_1 \\ z_1 \end{matrix} \right\|, \quad \|\vec{r_2}\|_g = \left\| \begin{matrix} x_2 \\ y_2 \\ z_2 \end{matrix} \right\|, \quad \|\vec{r_3}\|_g = \left\| \begin{matrix} x_3 \\ y_3 \\ z_3 \end{matrix} \right\|,$$

Решение. Из условия задачи следует, что неколлинеарные векторы $\vec{r}_2 - \vec{r}_1$ и $\vec{r}_3 - \vec{r}_1$ параллельны искомой плоскости. Кроме того, для радиуса-вектора любой принадлежащей этой плоскости точки $\vec{r}$ вектор $\vec{r} - \vec{r}_1$ также будет ей параллелен.

Из условия компланарности тройки векторов

$$\vec{r} - \vec{r}_1, \quad \vec{r}_2 - \vec{r}_1 \quad \text{и} \quad \vec{r}_3 - \vec{r}_1$$

получаем уравнение искомой плоскости, которое будет иметь вид $(\vec{r} - \vec{r}_1, \vec{r}_2 - \vec{r}_1, \vec{r}_3 - \vec{r}_1) = 0$ или же в координатной форме (согласно § 2.7):

$$\det \left\| \begin{matrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{matrix} \right\| (\vec{g}_1, \vec{g}_2, \vec{g}_3) = 0\,.$$

Наконец, учитывая, что смешанное произведение тройки линейно независимых (поскольку они суть базисные) векторов $\vec{g}_1, \vec{g}_2, \vec{g}_3$ не равно нулю, то окончательно получаем искомое уравнение (в *любой* декартовой системе координат):

$$\det \left\| \begin{matrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{matrix} \right\| = 0\,.$$

Решение получено.

Задача 3.3.2 *В системе координат* $\{O, \vec{g}_1, \vec{g}_2, \vec{g}_3\}$ *составить уравнение плоскости, проходящей через заданную точку* $\left\|\vec{r}_0\right\|_g = \begin{Vmatrix} x_0 \\ y_0 \\ z_0 \end{Vmatrix}$ *перпендикулярно ненулевому вектору* $\left\|\vec{n}\right\|_g = \begin{Vmatrix} n_x \\ n_y \\ n_z \end{Vmatrix}$.

Решение. По условию задачи для любой точки, принадлежащей этой плоскости, с радиусом-вектором $\vec{r}$ векторы $\vec{n}$ и $\vec{r} - \vec{r}_0$ будут ортогональны, т.е. $(\vec{n}, \vec{r} - \vec{r}_0) = 0$.

В *ортонормированной* системе $\{O, \vec{e}_1, \vec{e}_2, \vec{e}_3\}$ это условие имеет координатный вид

$$n_x(x - x_0) + n_y(y - y_0) + n_z(z - z_0) = 0$$

или если обозначить $A = n_x$, $B = n_y$, $C = n_z$ и соответственно $D = -n_x x_0 - n_y y_0 - n_z z_0$, то получаем знакомое равенство $Ax + By + Cz + D = 0$.

Решение получено.

Определение 3.3.2	Ненулевой вектор $\vec{n}$ называется *нормальным* вектором плоскости $(\vec{n}, \vec{r} - \vec{r}_0) = 0$.

Определение 3.3.3	Вектор $\begin{Vmatrix} A \\ B \\ C \end{Vmatrix}$ называется *главным* вектором плоскости $Ax+By+Cz+D = 0\,, \quad \lvert A\rvert+\lvert B\rvert+\lvert C\rvert > 0\,.$

В *ортонормированной* системе координат главный вектор плоскости является также и ее нормальным вектором.

Задача 3.3.3 *В ортонормированной системе координат* $\{O, \vec{e}_1, \vec{e}_2, \vec{e}_3\}$ *найти расстояние от точки* M *с радиусом-вектором* $\left\|\vec{R}\right\|_e = \begin{Vmatrix} X \\ Y \\ Z \end{Vmatrix}$ *до плоскости* $(\vec{n}, \vec{r} - \vec{r}_0) = 0$.

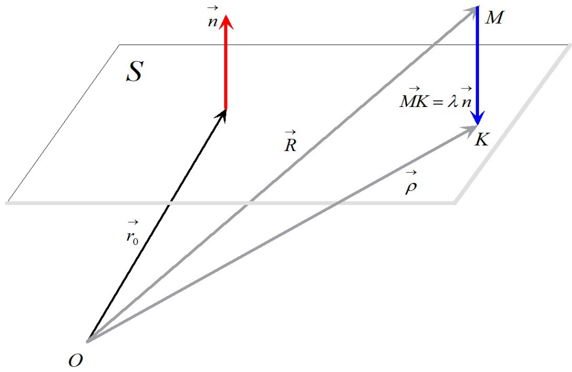


Рис. 3.6. К решению задачи 3.3.3

`Решение.` Пусть K есть ортогональная проекция точки M на данную плоскость, тогда $\vec{MK} = \lambda\vec{n}$ и $\vec{\rho} = \vec{R} + \lambda\vec{n}_0$. Точка K принадлежит данной плоскости, поэтому имеет место соотношение $(\vec{n}, \vec{R} + \lambda\vec{n} - \vec{r}_0) = 0$, из которого получаем $\lambda = -\dfrac{(\vec{n}, \vec{R} - \vec{r}_0)}{|\vec{n}|^2}$. Тогда искомое расстояние будет равно

$$|\vec{MK}| = \left|\frac{(\vec{n}, \vec{R} - \vec{r}_0)}{|\vec{n}|}\right| = \frac{\left|(\vec{n}, \vec{R} - \vec{r}_0)\right|}{|\vec{n}|}.$$

В *ортонормированной* системе вектор $\vec{n} = \left\| \begin{matrix} A \\ B \\ C \end{matrix} \right\|^{\mathrm{T}}$ — нормальный вектор плоскости $Ax + By + Cz + D = 0$.

Откуда $\quad |\vec{MK}| = \dfrac{|A(X - x_0) + B(Y - y_0) + C(Z - z_0)|}{\sqrt{A^2 + B^2 + C^2}}$.

Точка $\vec{r}_0 \in S$, следовательно, $Ax_0 + By_0 + Cz_0 + D = 0$. Поэтому ответ задачи можно записать так:

$$|\vec{MK}| = \frac{|\ AX + BY + CZ + D\ |}{\sqrt{A^2 + B^2 + C^2}}.$$

`Решение получено.`

Теорема 3.3.4 **Для того чтобы плоскости**

$$A_1x + B_1y + C_1z + D_1 = 0\,, \qquad |A_1| + |B_1| + |C_1| > 0\,,$$

и

$$A_2x + B_2y + C_2z + D_2 = 0\,, \qquad |A_2| + |B_2| + |C_2| > 0\,,$$

были параллельны или совпадали, необходимо и достаточно, чтобы их главные векторы были коллинеарны.

Доказательство.

Докажем достаточность.

Если главные векторы плоскостей коллинеарны, то $\exists\lambda \neq 0$ такое, что $A_1 = \lambda A_2,\ \ B_1 = \lambda B_2,\ \ C_1 = \lambda C_2$.

Тогда система уравнений

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0 \end{cases}$$

может быть записана в виде

$$\begin{cases} A_1x + B_1y + C_1z + \quad D_1 = 0, \\ A_1x + B_1y + C_1z + \lambda D_2 = 0. \end{cases}$$

Из этой системы следует, что при $D_1 \neq \lambda D_2$ общих точек у рассматриваемых плоскостей нет, а при $D_1 = \lambda D_2$ все точки общие. Таким образом, достаточность доказана.

Теперь докажем необходимость.

Пусть плоскости

$$A_1x + B_1y + C_1z + D_1 = 0\,, \quad |A_1| + |B_1| + |C_1| > 0\,. \quad \text{и}$$

$$A_2x + B_2y + C_2z + D_2 = 0\,, \quad |A_2| + |B_2| + |C_2| > 0$$

параллельны. Тогда они должны пересекать каждую из координатных плоскостей по параллельным прямым.

Для определенности будем считать, что это плоскости, для которых $x = 0$ и $z = 0$. Линии пересечения с первой из этих координатных плоскостей будут задаваться системами уравнений

$$\begin{cases} x = 0, \\ B_1 y + C_1 z + D_1 = 0 \end{cases} \quad \text{и} \quad \begin{cases} x = 0, \\ B_2 y + C_2 z + D_2 = 0. \end{cases}$$

Параллельность этих прямых означает, что $\exists \lambda \neq 0$ такое, что

$$B_1 = \lambda B_2, \;\; C_1 = \lambda C_2.$$

Для случая $z = 0$ системы линейных уравнений, определяющие линии пересечения, имеют аналогичный вид:

$$\begin{cases} z = 0, \\ A_1 y + B_1 y + D_1 = 0 \end{cases} \quad \text{и} \quad \begin{cases} z = 0, \\ A_2 x + B_2 y + D_2 = 0. \end{cases}$$

Теперь из условия параллельности секущих плоскостей и полученного ранее равенства $B_1 = \lambda B_2$ находим, что также и $A_1 = \lambda A_2$.

Теорема доказана.

Следствие 3.3.2 **Для того чтобы уравнения**

$$A_1 x + B_1 y + C_1 z + D_1 = 0 \,, \qquad |A_1| + |B_1| + |C_1| > 0 \,,$$

и

$$A_2 x + B_2 y + C_2 z + D_2 = 0 \,, \qquad |A_2| + |B_2| + |C_2| > 0 \,,$$

были уравнениями одной и той же плоскости, необходимо и достаточно, чтобы $\exists \lambda \neq 0$ такое, что

$$A_1 = \lambda A_2, \;\; B_1 = \lambda B_2, \;\; C_1 = \lambda C_2, \;\; D_1 = \lambda D_2.$$

Определение 3.3.4 *Пучком плоскостей* в пространстве называется совокупность всех плоскостей, проходящих через данную прямую.

Определение 3.3.5	*Уравнением пучка плоскостей*, проходящих через прямую, определяемую пересечением пары непараллельных плоскостей $A_1x + B_1y + C_1z + D_1 = 0\,, \;\; \lvert A_1\rvert + \lvert B_1\rvert + \lvert C_1\rvert > 0$ и $A_2x + B_2y + C_2z + D_2 = 0\,, \;\; \lvert A_2\rvert + \lvert B_2\rvert + \lvert C_2\rvert > 0\,,$ называется уравнение вида $\alpha(A_1x+B_1y+C_1z+D_1)+\beta(A_2x+B_2y+C_2z+D_2) = 0\,,$ где $\lvert\alpha\rvert + \lvert\beta\rvert > 0\,.$

Определение 3.3.6	*Связкой плоскостей* в пространстве называется совокупность всех плоскостей, проходящих через данную точку.

Определение 3.3.7	Если точка P, принадлежащая одновременно трем плоскостям: $A_1x+B_1y+C_1z+D_1 = 0\,, \;\; \lvert A_1\rvert+\lvert B_1\rvert+\lvert C_1\rvert > 0\,,$ $A_2x+B_2y+C_2z+D_2 = 0\,, \;\; \lvert A_2\rvert+\lvert B_2\rvert+\lvert C_2\rvert > 0$ и $A_3x+B_3y+C_3z+D_3 = 0\,, \;\; \lvert A_3\rvert+\lvert B_3\rvert+\lvert C_3\rvert > 0$ единственная, то уравнение вида $\alpha(A_1x + B_1y + C_1z + D_1)+$ $+\beta(A_2x + B_2y + C_2z + D_2)+$ $+\gamma(A_3x + B_3y + C_3z + D_3) = 0$ называется *уравнением связки плоскостей*, проходящих через точку P.

Для пучка и связки плоскостей в пространстве справедливы теоремы, аналогичные теореме 3.2.1 для пучка прямых на плоскости.

3.4. Способы задания прямой в пространстве

Существуют различные способы задания прямой в пространстве в декартовой системе координат $\left\{ O, \vec{g}_1, \vec{g}_2, \vec{g}_3 \right\}$.

`1°. Уравнение прямой в пространстве в векторной параметрической форме`

Возьмем произвольную точку с радиусом-вектором $\vec{r} = \left\| \begin{matrix} x \\ y \\ z \end{matrix} \right\|$, принадлежащую прямой L, имеющей ненулевой направляющий вектор $\vec{a} = \left\| \begin{matrix} a_x \\ a_y \\ a_z \end{matrix} \right\|$ и проходящей через заданную точку с радиусом-вектором $\vec{r}_0 = \left\| \begin{matrix} x_0 \\ y_0 \\ z_0 \end{matrix} \right\|$.

Тогда из коллинеарности векторов $\vec{a}$ и $\vec{r} - \vec{r}_0$ следует, что уравнение прямой имеет вид

$$\vec{r} = \vec{r}_0 + \tau\vec{a} \qquad \forall\tau \in (-\infty, +\infty)\,.$$

См. рис. 3.1.

`2°. Уравнение прямой в канонической форме`

Рассмотрим координатную запись уравнения $\vec{r} = \vec{r}_0 + \tau\vec{a}$:

$$\begin{cases} x = x_0 + \tau a_x, \\ y = y_0 + \tau a_y, \\ z = z_0 + \tau a_z. \end{cases}$$

Если в этой системе уравнений исключить параметр τ, то при $a_x a_y a_z \neq 0$ получается так называемое *каноническое уравнение прямой*:

$$\frac{x - x_0}{a_x} = \frac{y - y_0}{a_y} = \frac{z - z_0}{a_z},$$

хотя здесь правильнее говорить о *системе уравнений*.

Случай $a_x a_y a_z = 0$ рассматривается так же, как в § 3.2.

3°. Уравнение прямой, проходящей через две несовпадающие точки $\vec{r}_1$, и $\vec{r}_2$ с $\left\|\vec{r}_1\right\|_g = \left\| \begin{matrix} x_1 \\ y_1 \\ z_1 \end{matrix} \right\|$ **и** $\left\|\vec{r}_2\right\|_g = \left\| \begin{matrix} x_2 \\ y_2 \\ z_2 \end{matrix} \right\|$

Поскольку в данном случае в качестве направляющего вектора $\vec{a}$ можно взять $\vec{r}_2 - \vec{r}_1$, а в качестве $\vec{r}_0$ — вектор $\vec{r}_1$, то векторная параметрическая форма уравнения прямой L примет вид $\vec{r} = \vec{r}_1 + \tau\left(\vec{r}_2 - \vec{r}_1\right)$ или

$$\vec{r} = (1-\tau)\vec{r}_1 + \tau\vec{r}_2 \quad \forall\tau \in (-\infty, +\infty)\,.$$

Координатный вид полученного векторного уравнения прямой в пространстве можно найти, исключив параметр τ.

В результате получится система вида

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}.$$

4°. Уравнение прямой в пространстве в 1-й векторной форме

Прямая в пространстве может быть задана как линия пересечения двух плоскостей с неколлинеарными нормальными векторами, т.е. плоскостей заданных уравнениями $(\vec{n}_1, \vec{r}) = d_1$ и $(\vec{n}_2, \vec{r}) = d_2$, т.е. системой

$$\begin{cases} (\vec{n}_1, \vec{r}) = d_1, \\ (\vec{n}_2, \vec{r}) = d_2, \end{cases}$$

где векторы $\vec{n}_1$ и $\vec{n}_2$ линейно независимые, $\vec{r}$ — радиус-вектор произвольной точки на прямой L , а d_1 и d_2 — некоторые известные числа. Или же, если дан $\vec{r}_0$ — радиус-вектор точки, через которую проходит прямая, то радиус-вектор любой точки этой прямой удовлетворяет следующей системе уравнений:

$$\begin{cases} (\vec{n}_1, \vec{r} - \vec{r}_0) = 0, \\ (\vec{n}_2, \vec{r} - \vec{r}_0) = 0. \end{cases}$$

Наконец, в координатной форме это будет система линейных уравнений

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases}$$

5°. Уравнение прямой в пространстве во 2-й векторной форме

Прямая в пространстве может быть задана при помощи условия *коллинеарности* векторов $\vec{a}$ и $\vec{r} - \vec{r}_0$ в виде уравнения $[\vec{a}, \vec{r} - \vec{r}_0] = \vec{o}$. Или же $[\vec{a}, \vec{r}] = \vec{b}$, где $\vec{b} = [\vec{a}, \vec{r}_0]$.

В ортонормированной системе координат уравнение этой прямой принимает вид векторного равенства

$$\det \begin{Vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ a_x & a_y & a_z \\ x & y & z \end{Vmatrix} = \vec{b},$$

которое в координатной форме будет

$$\begin{cases} a_y z - a_z y = b_x, \\ a_z x - a_x z = b_y, \\ a_x y - a_y x = b_z. \end{cases}$$

Отметим, что в последней системе скалярных условий только два уравнения из трех независимые, то есть любое из этих уравнений является следствием двух других.

Действительно, умножив первое уравнение на a_x , второе на a_y и третье на a_z и сложив затем полученные равенства почленно, приходим к тождеству вида $0 = 0$, поскольку числа a_x , a_y и a_z не равны нулю одновременно, а $\begin{cases} b_x = a_y z_0 - a_z y_0, \\ b_y = a_z x_0 - a_x z_0, \\ b_z = a_x y_0 - a_y x_0. \end{cases}$

Наконец, покажем, что расстояние h в пространстве от некоторой точки с радиусом-вектором $\vec{R}$ до прямой $\vec{r} = \vec{r}_0 + \tau\vec{a}$ можно найти, воспользовавшись свойством, что S – площадь параллелограмма, построенного на паре векторов, равна длине векторного произведения этих векторов. Из рис. 3.7 получаем формулу $h = \dfrac{S}{|\vec{a}|} = \dfrac{\left|\,[\vec{R} - \vec{r}_0, \vec{a}]\,\right|}{|\vec{a}|}$.

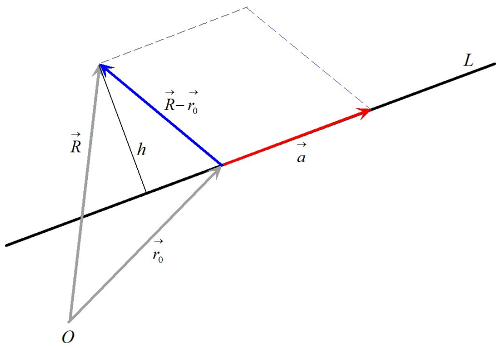


Рис. 3.7. К вычислению расстояния от точки до прямой в пространстве

3.5. Решение геометрических задач методами векторной алгебры

Эффективность использования методов векторной алгебры при решении геометрических задач во многом зависит от выбора формы представления геометрических условий в векторной форме, который нередко оказывается неоднозначным.

Например, вычисление углов, определяющих взаимное расположение прямых и плоскостей в пространстве, может быть сведено к нахождению скалярных и/или векторных произведений соответствующих нормальных и направляющих векторов.

Определение 3.5.1	*Углом между плоскостями* $(\vec{n}_1, \vec{r} - \vec{r}_{01}) = 0$ и $(\vec{n}_2, \vec{r} - \vec{r}_{02}) = 0$ называется угол между их нормальными векторами $\vec{n}_1$ и $\vec{n}_2$.
Определение 3.5.2	*Углом между плоскостью* $(\vec{n}, \vec{r} - \vec{r}_0) = 0$ *и прямой* $\vec{r} = \vec{r}_0 + \tau\vec{a}$ называется угол $\pi/2 - \alpha$, где α – угол между векторами $\vec{n}$ и $\vec{a}$.

В табл. 3.5.1—3.5.4 приведены некоторые из возможных форм записи геометрических условий при помощи векторных операций.

Таблица 3.5.1

Геометрическое условие	Возможное представление
Коллинеарность прямых $\vec{r}=\vec{r}_{01}+\tau\vec{a}_1$ и $\vec{r}=\vec{r}_{02}+\tau\vec{a}_2$	$1^\circ.\quad \exists\lambda\neq 0:\quad \vec{a}_1=\lambda\vec{a}_2$ $2^\circ.\quad [\vec{a}_1,\vec{a}_2]=\vec{o}$
Ортогональность прямых $\vec{r}=\vec{r}_{01}+\tau\vec{a}_1$ и $\vec{r}=\vec{r}_{02}+\tau\vec{a}_2$	$(\vec{a}_1,\vec{a}_2)=0$
Коллинеарность прямых $\vec{r}=\vec{r}_0+\tau\vec{a}$ и $\begin{cases}(\vec{n}_1,\vec{r})=d_1,\\(\vec{n}_2,\vec{r})=d_2,\end{cases}$	$1^\circ.\quad \exists\lambda\neq 0:\quad \vec{a}=\lambda[\vec{n}_1,\vec{n}_2]$ $2^\circ.\quad [\vec{a},[\vec{n}_1,\vec{n}_2]]=\vec{o}$
Ортогональность прямых $\vec{r}=\vec{r}_0+\tau\vec{a}$ и $\begin{cases}(\vec{n}_1,\vec{r})=d_1,\\(\vec{n}_2,\vec{r})=d_2,\end{cases}$	$(\vec{a},\vec{n}_1,\vec{n}_2)=0$
Совпадение прямых $\vec{r}=\vec{r}_{01}+\tau\vec{a}_1$ и $\vec{r}=\vec{r}_{02}+\tau\vec{a}_2$	$1^\circ.\ \exists\lambda\neq 0:\vec{a}_1=\lambda\vec{a}_2,$ $\exists\mu\neq 0:\vec{r}_{01}-\vec{r}_{02}=\mu\vec{a}_1$ $2^\circ.\quad [\vec{a}_1,\vec{a}_2]=\vec{o}$ и $[\vec{r}_{01}-\vec{r}_{02},\vec{a}_1]=\vec{o}$

Таблица 3.5.2

Геометрическое условие	Возможное представление
Пересечение прямых $\vec{r} = \vec{r}_{01} + \tau\vec{a}_1$ и $\vec{r} = \vec{r}_{02} + \tau\vec{a}_2$	$[\vec{a}_1, \vec{a}_2] \neq \vec{o}$ и $(\vec{r}_{01} - \vec{r}_{02}, \vec{a}_1, \vec{a}_2) = 0$
Скрещивание прямых $\vec{r} = \vec{r}_{01} + \tau\vec{a}_1$ и $\vec{r} = \vec{r}_{02} + \tau\vec{a}_2$	$[\vec{a}_1, \vec{a}_2] \neq \vec{o}$ и $(\vec{r}_{01} - \vec{r}_{02}, \vec{a}_1, \vec{a}_2) \neq 0$
Параллельность плоскостей $\vec{r} = \vec{r}_{01} + \varphi\vec{p}_1 + \theta\vec{q}_1$ и $\vec{r} = \vec{r}_{02} + \varphi\vec{p}_2 + \theta\vec{q}_2$	1°. $\exists\lambda \neq 0 : [\vec{p}_1, \vec{q}_1] = \lambda[\vec{p}_2, \vec{q}_2]$ и $(\vec{r}_{01} - \vec{r}_{02}, \vec{p}_1, \vec{q}_2) \neq 0$ 2°. $[[\vec{p}_1, \vec{q}_1], [\vec{p}_2, \vec{q}_2]] = \vec{o}$ и $(\vec{r}_{01} - \vec{r}_{02}, \vec{p}_1, \vec{q}_2) \neq 0$
Совпадение плоскостей $\vec{r} = \vec{r}_{01} + \varphi\vec{p}_1 + \theta\vec{q}_1$ и $\vec{r} = \vec{r}_{02} + \varphi\vec{p}_2 + \theta\vec{q}_2$	1°. $\exists\lambda \neq 0 : [\vec{p}_1, \vec{q}_1] = \lambda[\vec{p}_2, \vec{q}_2]$ и $(\vec{r}_{01} - \vec{r}_{02}, \vec{p}_1, \vec{q}_2) = 0$ 2°. $[[\vec{p}_1, \vec{q}_1], [\vec{p}_2, \vec{q}_2]] = \vec{o}$ и $(\vec{r}_{01} - \vec{r}_{02}, \vec{p}_1, \vec{q}_2) = 0$
Ортогональность плоскостей $\vec{r} = \vec{r}_{01} + \varphi\vec{p}_1 + \theta\vec{q}_1$ и $\vec{r} = \vec{r}_{02} + \varphi\vec{p}_2 + \theta\vec{q}_2$	$([\vec{p}_1, \vec{q}_1], [\vec{p}_2, \vec{q}_2]) = 0$

Таблица 3.5.3

Геометрическое условие	Возможное представление
Параллельность плоскостей $\vec{r} = \vec{r}_0 + \varphi\vec{p} + \theta\vec{q}$ и $(\vec{n}, \vec{r}) = d$	$\begin{cases} (\vec{p}, \vec{n}) = 0, \\ (\vec{q}, \vec{n}) = 0 \end{cases}$ при $(\vec{n}, \vec{r}_0) \neq d$
Совпадение плоскостей $\vec{r} = \vec{r}_{01} + \varphi\vec{p}_1 + \theta\vec{q}_1$ и $(\vec{n}, \vec{r}) = d$	$\begin{cases} (\vec{p}, \vec{n}) = 0, \\ (\vec{q}, \vec{n}) = 0 \end{cases}$ при $(\vec{n}, \vec{r}_0) = d$
Ортогональность плоскостей $\vec{r} = \vec{r}_0 + \varphi\vec{p} + \theta\vec{q}$ и $(\vec{n}, \vec{r}) = d$	$1^\circ.\quad \exists\lambda \neq 0 : \quad [\vec{p}, \vec{q}] = \lambda\vec{n}$ $2^\circ.\quad [[\vec{p}, \vec{q}], \vec{n}] = \vec{o}$
Ортогональность плоскостей $\vec{r} = \vec{r}_{01} + \varphi\vec{p}_1 + \theta\vec{q}_1$ и $(\vec{n}, \vec{r}) = d$	$(\vec{a}, \vec{n}_1, \vec{n}_2) = 0$
Параллельность прямой $\vec{r} = \vec{r}_{01} + \tau\vec{a}$ и плоскости $\vec{r} = \vec{r}_{02} + \varphi\vec{p} + \theta\vec{q}$	$\begin{cases} (\vec{a}, \vec{p}, \vec{q}) = 0, \\ (\vec{r}_{01} - \vec{r}_{02}, \vec{p}, \vec{q}) \neq 0. \end{cases}$

Отметим, что в табл. 3.5.1—3.5.4 сохранены введенные ранее обозначения и ограничения на значения параметров.

Таблица 3.5.4

Геометрическое условие	**Возможное представление**
Ортогональность прямой $\vec{r}=\vec{r}_{01}+\tau\vec{a}$ и плоскости $\vec{r}=\vec{r}_{02}+\varphi\vec{p}+\theta\vec{q}$	1°. $\exists\lambda\neq 0: \quad \vec{a}=\lambda[\vec{p},\vec{q}]$ 2°. $[[\vec{p},\,\vec{q}],\vec{a}]=\vec{o}$
Параллельность прямой $\vec{r}=\vec{r}_0+\tau\vec{a}$ и плоскости $(\vec{n},\vec{r})=d$	$(\vec{a},\vec{n})=0$ при условии $(\vec{n},\vec{r}_0)\neq d$
Принадлежность прямой $\vec{r}=\vec{r}_0+\tau\vec{a}$ плоскости $(\vec{n},\vec{r})=d$	$\begin{cases}(\vec{a},\vec{n})=0,\\ (\vec{n},\vec{r}_0)=d\end{cases}$
Ортогональность прямой $\vec{r}=\vec{r}_0+\tau\vec{a}$ и плоскости $(\vec{n},\vec{r})=d$	$[\vec{a},\,\vec{n}]=\vec{o}$
Ортогональность прямой $\begin{cases}(\vec{n}_1,\vec{r})=d_1,\\ (\vec{n}_2,\vec{r})=d_2\end{cases}$ и плоскости $(\vec{n},\vec{r})=d$	$[[\vec{n}_1,\,\vec{n}_2],\vec{n}]=\vec{o}$

При решении геометрических задач методами векторной алгебры также важно уметь переводить эти представления из одной формы в другую, эквивалентную первой [6].

[6]Следует иметь в виду, что использование различных векторных представле-

Найдем, например, для прямой, заданной в пространстве пересечением двух непараллельных плоскостей $\begin{cases} (\vec{n}_1, \vec{r}) = d_1, \\ (\vec{n}_2, \vec{r}) = d_2, \end{cases}$ ее уравнение в параметрическом виде $\vec{r} = \vec{r}_0 + \tau\vec{a}$.

Нетрудно убедиться, что в качестве направляющего вектора данной прямой можно взять $\vec{a} = [\vec{n}_1, \vec{n}_2]$, а радиус-вектор точки $\vec{r}_0$ выразить как линейную комбинацию векторов $\vec{n}_1$ и $\vec{n}_2$. Действительно, пусть $\vec{r}_0 = \xi\vec{n}_1 + \eta\vec{n}_2$, тогда из системы линейных уравнений $\begin{cases} (\vec{n}_1, \vec{r}_0) = d_1, \\ (\vec{n}_2, \vec{r}_0) = d_2 \end{cases}$ находим $\xi = \dfrac{\Delta_\xi}{\Delta}$ и $\eta = \dfrac{\Delta_\eta}{\Delta}$, где

$$\Delta = \det\left\| \begin{matrix} (\vec{n}_1, \vec{n}_1) & (\vec{n}_1, \vec{n}_2) \\ (\vec{n}_2, \vec{n}_1) & (\vec{n}_2, \vec{n}_2) \end{matrix} \right\|, \quad \Delta_\xi = \det\left\| \begin{matrix} d_1 & (\vec{n}_1, \vec{n}_2) \\ d_2 & (\vec{n}_2, \vec{n}_2) \end{matrix} \right\|$$

и

$$\Delta_\eta = \det\left\| \begin{matrix} (\vec{n}_1, \vec{n}_1) & d_1 \\ (\vec{n}_2, \vec{n}_1) & d_2 \end{matrix} \right\|$$

(см. теорему 1.1.2).

Покажите самостоятельно, что условие неколлинеарности нормальных векторов $\vec{n}_1$ и $\vec{n}_2$ равносильно условию $\Delta \neq 0$.

Аналогично может быть выполнен и обратный переход. Пусть уравнение прямой в пространстве имеет вид $\vec{r} = \vec{r}_0 + \tau\vec{a}$, причем предположим, что $\vec{r}_0$ и $\vec{a}$ неколлинеарны.

Тогда в качестве нормальных векторов плоскостей, которые пересекаются по данной прямой, можно взять $\vec{n}_1 = [\vec{a}, \vec{r}_0]$ и $\vec{n}_2 = [\vec{a}, \vec{n}_1]$.

Из второго равенства, используя формулу для двойного векторного произведения (см. § 2.8), получаем

$$\vec{n}_2 = [\vec{a}, \vec{n}_1] = [\vec{a}, [\vec{a}, \vec{r}_0]] = (\vec{a}, \vec{r}_0)\vec{a} - |\vec{a}|^2\vec{r}_0 \,.$$

В качестве d_1 и d_2, очевидно, можно взять числа $d_1 = (\vec{n}_1, \vec{r}_0)$ и $d_2 = (\vec{n}_2, \vec{r}_0)$. Случай коллинеарных векторов $\vec{r}_0$ и $\vec{a}$ рассмотрите самостоятельно.

В заключение приведем в качестве примеров решения некоторых стереометрических задач методами векторной алгебры.

ний одного и то же геометрического условия может приводить к различным, но, естественно, эквивалентным формам записи решения (см., например, задачу 3.5.2).

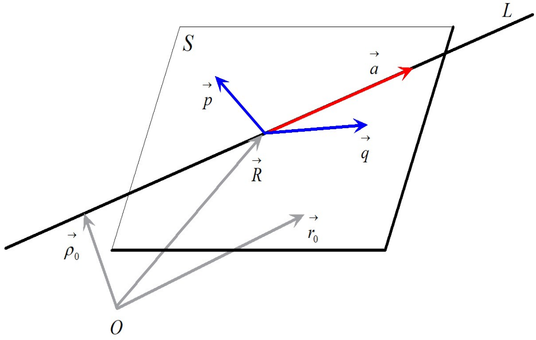


Рис. 3.8. К решению задачи 3.5.1

Задача 3.5.1 *Даны плоскость* $(\vec{r}-\vec{r}_0,\vec{p},\vec{q}) = 0$ *и пересекающая ее прямая* $\vec{r} = \vec{\rho}_0 + \tau\vec{a}$. *Найти в векторной форме радиус-вектор точки пересечения этих прямой и плоскости. Решить задачу в общей декартовой системе координат, уравнения плоскости и прямой в которой имеют вид соответственно* $x + 4z + 7 = 0$ *и*

$$\begin{cases} x = -3 + 4\tau, \\ y = 1 - 4\tau, \\ z = -5 + \tau. \end{cases}$$

Решение 1) Заметим, что если $(\vec{a},\vec{p},\vec{q}) = 0$, то либо решений нет, либо вся прямая целиком принадлежит данной плоскости. Поэтому можно считать, что $(\vec{a},\vec{p},\vec{q}) \neq 0$,

2) Имеем $\vec{R} = \vec{\rho}_0 + \tau_0\vec{a}$, где $\vec{R}$ — радиус-вектор искомой точки пересечения прямой и плоскости, а τ_0 — значение параметра τ, соответствующее этой точке (см. рис. 3.8). Но поскольку точка пересечения принадлежит также и плоскости, то имеем равенство $(\vec{R}-\vec{r}_0,\vec{p},\vec{a}) = 0$ или $(\vec{\rho}_0 + \tau_0\vec{a} - \vec{r}_0,\vec{p},\vec{a}) = 0$. Откуда $\tau_0 = -\dfrac{(\vec{\rho}_0-\vec{r}_0,\vec{p},\vec{q})}{(\vec{a},\vec{p},\vec{q})}$ и, наконец, $\vec{R} = \vec{\rho}_0 - \dfrac{(\vec{\rho}_0-\vec{r}_0,\vec{p},\vec{q})}{(\vec{a},\vec{p},\vec{q})}\,\vec{a}\,.$

3). Найдем теперь решение в координатной форме.
Поскольку координатное представление каждого смешанного произведения в базисе $\{\vec{g}_1, \vec{g}_2, \vec{g}_3\}$ (см. § 2.7) имеет ненулевой сомножитель $(\vec{g}_1, \vec{g}_2, \vec{g}_3)$, то координатное представление радиуса-вектора $\vec{R}$ будет одинаковым в любой декартовой системе координат.
Из условия задачи и теоремы 3.3.3 имеем

$$\left\|\vec{p}\right\| = \left\| \begin{array}{r} 0 \\ -4 \\ 0 \end{array} \right\|, \quad \left\|\vec{q}\right\| = \left\| \begin{array}{r} 4 \\ 0 \\ -1 \end{array} \right\|, \quad \left\|\vec{\rho}_0\right\| = \left\| \begin{array}{r} -3 \\ 1 \\ -5 \end{array} \right\| \quad \text{и}$$

$$\left\|\vec{a}\right\| = \left\| \begin{array}{r} 4 \\ -4 \\ 1 \end{array} \right\| .$$

Координаты точки $\vec{r}_0$, через которую проходит плоскость $x+4z+7=0$, подберем произвольно: например, возьмем $\left\|\vec{r}_0\right\| = \left\| \begin{array}{r} -3 \\ 0 \\ -1 \end{array} \right\|$.

Таким образом, имеем

$$(\vec{\rho}_0 - \vec{r}_0, \vec{p}, \vec{q}) = \det \left\| \begin{array}{rrr} 0 & 1 & -4 \\ 0 & -4 & 0 \\ 4 & 0 & -1 \end{array} \right\| (\vec{g}_1, \vec{g}_2, \vec{g}_3) .$$

Аналогично $\quad (\vec{a}, \vec{p}, \vec{q}) = \det \left\| \begin{array}{rrr} 4 & -4 & 1 \\ 0 & -4 & 0 \\ 4 & 0 & -1 \end{array} \right\| (\vec{g}_1, \vec{g}_2, \vec{g}_3) .$

Откуда $\quad \tau_0 = -\dfrac{(\vec{\rho}_0 - \vec{r}_0, \vec{p}, \vec{q})}{(\vec{a}, \vec{p}, \vec{q})} = -\dfrac{64 \cdot (\vec{g}_1, \vec{g}_2, \vec{g}_3)}{(-32)(\vec{g}_1, \vec{g}_2, \vec{g}_3)} = 2,$
и, в силу $\vec{R} = \vec{\rho}_0 + \tau_0 \, \vec{a}$,

$$\left\|\vec{R}\right\| = \left\| \begin{array}{c} X \\ Y \\ Z \end{array} \right\| = \left\| \begin{array}{r} -3 \\ 1 \\ -5 \end{array} \right\| + 2 \left\| \begin{array}{r} 4 \\ -4 \\ 1 \end{array} \right\| = \left\| \begin{array}{r} 5 \\ -7 \\ -3 \end{array} \right\| .$$

Решение получено.

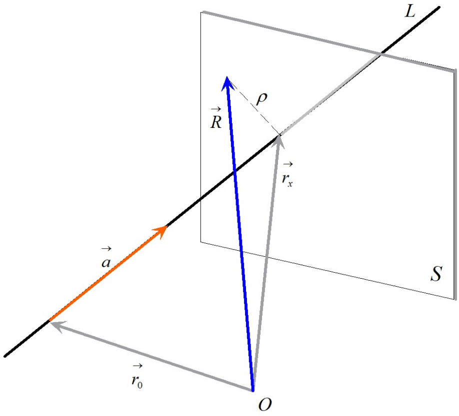


Рис. 3.9. К решению задачи 3.5.2

Задача 3.5.2 *Пусть заданы точка с радиусом-вектором $\vec{R}$ и прямая $\vec{r} = \vec{r}_0 + \tau\vec{a}$. Найти расстояние от этой точки до данной прямой, не используя операцию векторного произведения.*

Решение Проведем через данную точку плоскость, перпендикулярную прямой (рис. 3.9). Обозначим через $\vec{r}_x$ радиус-вектор точки пересечения прямой и плоскости. Тогда искомое расстояние будет равно $\rho = |\vec{R} - \vec{r}_x|$.

Точка $\vec{r}_x$ принадлежит одновременно как прямой, так и плоскости. Поэтому $\vec{r}_x$ будет удовлетворять системе, состоящей из их уравнений:

$$\begin{cases} \vec{r} = \vec{r}_0 + \tau\vec{a}, \\ (\vec{a}, \vec{R} - \vec{r}_x) = 0. \end{cases}$$

но тогда, исключая параметр τ, находим, что

$$\vec{r}_x = \vec{r}_0 + \frac{(\vec{\rho}_0 - \vec{r}_0, \vec{a})}{|\vec{a}|^2}\,\vec{a}\,.$$

Тогда получаем

$$\rho = |\vec{R} - \vec{r}_x| = \sqrt{(\vec{R} - \vec{r}_x, \vec{R} - \vec{r}_x)} =$$

$$= \sqrt{\left(\vec{R} - \vec{r}_0 - \frac{(\vec{\rho}_0 - \vec{r}_0, \vec{a})}{|\vec{a}|^2}\,\vec{a}, \vec{R} - \vec{r}_0 - \frac{(\vec{\rho}_0 - \vec{r}_0, \vec{a})}{|\vec{a}|^2}\,\vec{a}\right)}\,.$$

А после упрощений находим, что

$$\rho = \sqrt{|\vec{R} - \vec{r}_0|^2 - \frac{(\vec{\rho}_0 - \vec{r}_0, \vec{a})^2}{|\vec{a}|^2}}\,.$$

Наконец, заметим, что, в силу легко проверяемого равенства $|\vec{p}|^2|\vec{q}|^2 = (\vec{p}, \vec{q})^2 + |[\vec{p}, \vec{q}]|^2$, мы получим формулу для ρ, совпадающую с приведенной в конце § 3.4.

`Решение получено.`

`Задача 3.5.3` *Найти расстояние между прямыми* $\vec{r} = \vec{r}_{01} + \tau\vec{a}_1$ *и* $\vec{r} = \vec{r}_{02} + \tau\vec{a}_2$.

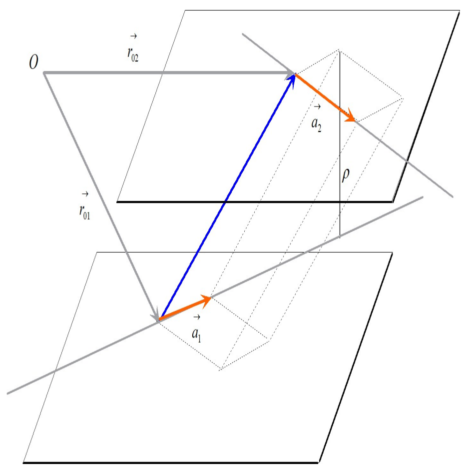


Рис. 3.10. К решению задачи 3.5.3

Решение Будем предполагать, что $\vec{r}_{01} \neq \vec{r}_{02}$. Если векторы $\vec{a}_1$ и $\vec{a}_2$ коллинеарны, то решение задачи сводится, например, к нахождению расстояния от точки $\vec{r}_{01}$ до второй прямой $\vec{r} = \vec{r}_{02} + \tau\vec{a}_2$.

Пусть векторы $\vec{a}_1$ и $\vec{a}_2$ не коллинеарны, тогда построим пару плоскостей, параллельных этим векторам, одна из которых проходит через точку $\vec{r}_{01}$, а другая — через $\vec{r}_{02}$. Объем параллелепипеда, построенного на некомпланарной тройке векторов $\vec{a}_1, \vec{a}_2$ и $\vec{r}_{01} - \vec{r}_{02}$, равен, с одной стороны, произведению площади параллелограмма, находящегося в основании, на искомую величину ρ и числу $\left|(\vec{r}_{01} - \vec{r}_{02}, \vec{a}_1, \vec{a}_2)\right|$, с другой стороны. Откуда находим, что $\rho = \dfrac{\left|(\vec{r}_{02} - \vec{r}_{01}, \vec{a}_1, \vec{a}_2)\right|}{\left|[\vec{a}_1, \vec{a}_2]\right|}$.

Решение получено.

Задача 3.5.4 *Даны плоскость* $(\vec{n}, \vec{r}) = d$ *и прямая* $[\vec{a}, \vec{r}] = \vec{b}$. *Найти* $\vec{R}$ *— радиус-вектор точки их пересечения.*

Решение Умножив обе части уравнения прямой векторно слева на $\vec{n}$, получим $[\vec{n}, [\vec{a}, \vec{r}]] = [\vec{n}, \vec{b}]$.

Подставляя в это соотношение искомый вектор $\vec{R}$ и применяя формулу для двойного векторного произведения (см. теорему 2.8.1), приходим к равенству

$$(\vec{n}, \vec{R})\vec{a} - (\vec{n}, \vec{a})\vec{R} = [\vec{n}, \vec{b}].$$

Поскольку точка $\vec{R}$ принадлежит данной плоскости, то $(\vec{n}, \vec{R}) = d$ и тогда при ограничении $(\vec{n}, \vec{a}) \neq 0$ получаем

$$\vec{R} = \frac{d\vec{a} - [\vec{n}, \vec{b}]}{(\vec{n}, \vec{a})}.$$

Решение получено.

4. Нелинейные объекты на плоскости и в пространстве

4.1. Линии на плоскости и в пространстве

Пусть дана декартова система координат $\{ O, \vec{g}_1, \vec{g}_2\}$ на плоскости и числовое множество Ω, являющееся промежутком (возможно, бесконечным) вещественной оси.

Определение 4.1.1	Будем говорить, что на плоскости задана *параметрически линия* L вектор-функцией $\vec{r} = \vec{F}(\tau)$ (или в координатной форме $$\| \vec{r} \|_g = \left\| \begin{array}{c} x \\ y \end{array} \right\| = \left\| \begin{array}{c} F_x(\tau) \\ F_y(\tau) \end{array} \right\| ,$$ где $F_x(\tau)$, $F_y(\tau)$ — непрерывные, скалярные функции аргумента $\tau \in \Omega$), если 1) $\forall \tau_0 \in \Omega$ точка $\vec{r}_0 = \vec{F}(\tau_0)$ лежит на L; 2) для любой точки $\vec{r}_0$, лежащей на L, $\exists \tau_0 \in \Omega$ такое, что выполнено равенство $\vec{r}_0 = \vec{F}(\tau_0)$.

Иногда линия на плоскости задается в виде уравнения $G(x, y) = 0$, которое получается при исключении параметра τ из системы уравнений

$$\begin{cases} x = F_x(\tau), \\ y = F_y(\tau) \end{cases} \qquad \tau \in \Omega\,.$$

Пример 4.1.1

1) Прямая линия, например, задается вектор-функцией $\vec{r} = \vec{r}_0 + \tau\vec{a}$, где $\vec{a}$ — направляющий вектор, а $\vec{r}_0$ — радиус-вектор одной из точек этой прямой.

Скалярная форма задания прямой в этом случае имеет вид

$$\begin{cases} x = x_0 + \tau a_x, \\ y = y_0 + \tau a_y \end{cases} \qquad \tau \in (-\infty, +\infty)\,,$$

то есть

$$\begin{cases} F_x(\tau) = x_0 + \tau a_x, \\ F_y(\tau) = y_0 + \tau a_y \end{cases} \qquad \tau \in (-\infty, +\infty)\,,$$

или же $\qquad Ax + By + C = 0\,, \quad |A| + |B| > 0\,,$

с $\;G(x,y) = Ax + By + C\,.$

2) В прямоугольной система координат окружность радиусом R с центром в точке $\left\|\begin{matrix} x_0 \\ y_0 \end{matrix}\right\|$ параметрически может быть задана так:

$$\begin{cases} x = x_0 + R\cos\tau, \\ y = y_0 + R\sin\tau \end{cases} \qquad \tau \in [0, 2\pi)\,,$$

то есть

$$\begin{cases} F_x(\tau) = x_0 + R\cos\tau, \\ F_y(\tau) = y_0 + R\sin\tau \end{cases} \qquad ,$$

или же

$$(x - x_0)^2 + (y - y_0)^2 = R^2 \qquad ,$$

с $\;G(x,y) = (x - x_0)^2 + (y - y_0)^2 - R^2\,.$

Определение 4.1.2

Линия называется *алгебраической*, если ее уравнение в декартовой системе координат имеет вид

$$\sum_{k=0}^{m} \alpha_k x^{p_k} y^{q_k} = 0\,,$$

где p_k и q_k — целые неотрицательные числа, а числа α_k не равны нулю одновременно.

Определение 4.1.3	Число $N = \max\limits_{k=[0,m]} \{p_k + q_k\}$ называется *порядком алгебраического уравнения*, указанного в определении 4.1.2, где максимум находится по всем k, для которых $\alpha_k \neq 0$. *Наименьший* из порядков алгебраических уравнений, задающих данную алгебраическую линию, называется *порядком алгебраической линии.*

Примеры алгебраических линий на плоскости различных порядков приведены в табл. 4.1.1.

Т а б л и ц а 4.1.1

Название	**Уравнение**	**Порядок**
Прямая	$Ax + By + C = 0$, $\vert A\vert + \vert B\vert > 0$	1
Квадратная парабола	$x - y^2 = 0$	2
Гипербола	$xy - 1 = 0$	2
«Декартов лист»	$x^3 + y^3 - 3xy = 0$	3

Теорема 4.1.1 **Порядок алгебраической линии не зависит от выбора декартовой системы координат.**

Доказательство.

Пусть алгебраическая линия L имеет в исходной («старой») системе координат $\{ O, \vec{g}_1, \vec{g}_2\}$ порядок N и уравнение $G(x,y)=0$.

Перейдем к «новой» системе координат $\left\{ O', \vec{g'}_1, \vec{g'}_2\right\}$. Формулы перехода, согласно соотношениям (1.8.2), имеют вид

$$\left\{ \begin{array}{rcl} x_1 & = & \sigma_{11}x'_1 + \sigma_{12}y'_2 + \beta_1\,, \\ x_2 & = & \sigma_{21}x'_1 + \sigma_{22}y'_2 + \beta_2\,, \end{array} \right.$$

поэтому уравнение линии L в «новой» системе координат будет

$$G(\sigma_{11}x'_1 + \sigma_{12}y'_2 + \beta_1,\ \sigma_{21}x'_1 + \sigma_{22}y'_2 + \beta_2) = 0\,,$$

порядок которого N'. Откуда в силу определения 4.1.2 следует, что $N \geq N'$, то есть при переходе к «новой» системе координат порядок алгебраической кривой повыситься не может.

Применяя аналогичные рассуждения для обратного перехода от системы координат $\left\{ O', \vec{g'}_1, \vec{g'}_2\right\}$ к системе $\{ O, \vec{g}_1, \vec{g}_2\}$, получим $N' \geq N$ и окончательно $N' = N$.

Теорема доказана.

Замечание 4.1.1. Фигуры на плоскости можно задавать, используя ограничения типа *неравенств*.

Пример 4.1.2

1) В прямоугольной системе координат набор условий

$$\left\{ \begin{array}{l} x \geq 0, \\ y \geq 0, \\ x + y - 3 \leq 0 \end{array} \right.$$

задает прямоугольный треугольник, катеты которого лежат на осях координат и имеют длины, равные 3.

2) В *ортонормированной* системе координат неравенство вида $x^2 + y^2 - 4 \leq 0$ определяет круг радиусом 2 с центром в начале координат.

Рассмотрим теперь случай линии в пространстве. Пусть дана пространственная система координат $\{ O, \vec{g}_1, \vec{g}_2, \vec{g}_3 \}$.

Определение 4.1.4

Будем говорить, что в пространстве задана *параметрически линия* L вектор-функцией $\vec{r} = \vec{F}(\tau)$ (или в координатной форме

$$\| \vec{r} \|_g = \left\| \begin{matrix} x \\ y \\ z \end{matrix} \right\| = \left\| \begin{matrix} F_x(\tau) \\ F_y(\tau) \\ F_y(\tau) \end{matrix} \right\| ,$$

где $F_x(\tau)$, $F_y(\tau)$, $F_z(\tau)$ — непрерывные, скалярные функции аргумента $\tau \in \Omega$), если

1) $\forall \tau_0 \in \Omega$ точка $\vec{r}_0 = \vec{F}(\tau_0)$ лежит на L;

2) для любой точки $\vec{r}_0$, лежащей на L, $\exists \tau_0 \in \Omega$ такое, что выполнено равенство $\vec{r}_0 = \vec{F}(\tau_0)$.

Иногда линия в пространстве задается системой уравнений

$$\begin{cases} G(x,y,z) = 0\,, \\ H(x,y,z) = 0\,, \end{cases}$$

которая получается исключением параметра τ из соотношений

$$\begin{cases} x = F_x(\tau), \\ y = F_y(\tau), \qquad \tau \in \Omega, \\ z = F_z(\tau), \end{cases}$$

или же равносильным уравнением, например, вида

$$G^2(x,y,z) + H^2(x,y,z) = 0\,.$$

Пример 4.1.3

1) В декартовой системе координат алгебраическая линия второго порядка $x^2 + y^2 = 0 \quad \forall z$ является прямой в пространстве.

2) В ортонормированной системе координат *винтовая линия* (пространственная спираль) радиусом R и с шагом $2\pi a$ может быть задана в параметрическом виде

$$\begin{cases} x = R\cos\tau, \\ y = R\sin\tau, \qquad \tau \in [0, 2\pi)\,, \\ z = a\tau, \end{cases}$$

$$\text{или же системой} \quad \begin{cases} x = R\cos\dfrac{z}{a}, \\ y = R\sin\dfrac{z}{a}. \end{cases}$$

4.2. Поверхности в пространстве

Пусть имеется пространственная система координат $\{\, O,\, \vec{g}_1,\, \vec{g}_2,\, \vec{g}_3 \}$ и Ω — множество упорядоченных пар чисел $\{\, \varphi;\, \theta \,\}$, заданное условиями: $\alpha \le \varphi \le \beta,\ \ \gamma \le \theta \le \delta$.

Определение 4.2.1	Будем говорить, что в пространстве задана *параметрически поверхность* S вектор-функцией $\vec{r} = \vec{F}(\varphi, \theta)$ (или в координатной форме $$\|\,\vec{r}\,\|_g = \begin{Vmatrix} x \\ y \\ z \end{Vmatrix} = \begin{Vmatrix} F_x(\varphi,\theta) \\ F_y(\varphi,\theta) \\ F_z(\varphi,\theta) \end{Vmatrix},$$ где $F_x(\varphi,\theta)$, $F_y(\varphi,\theta)$, $F_z(\varphi,\theta)$ — непрерывные, скалярные функции, зависящие от двух переменных $\{\varphi,\theta\} \in \Omega$), если 1) для любой упорядоченной пары чисел $\{\varphi,\theta\} \in \Omega$ точка $\vec{r}_0 = \vec{F}(\varphi_0,\theta_0)$ лежит на поверхности S; 2) для любой точки $\vec{r}_0$, лежащей на S, существует пара чисел $\{\varphi_0, \theta_0 \in \Omega\}$ таких, что $\vec{r}_0 = \vec{F}(\varphi_0,\theta_0)$.

Иногда поверхность в пространстве задается в виде уравнения $G(x,y,z) = 0$, которое получается при исключении φ_0, θ_0 из системы уравнений

$$\begin{cases} x = F_x(\varphi_0,\theta_0), \\ y = F_y(\varphi_0,\theta_0), \\ z = F_z(\varphi_0,\theta_0), \end{cases} \qquad \{\varphi_0,\theta_0\} \in \Omega\,.$$

Пример 4.2.1

1) Плоскость S в пространстве задается, например, вектор-функцией $\vec{r} = \vec{r}_0 + \varphi\vec{p} + \theta\vec{q}$, где $\vec{p}$ и $\vec{q}$ — *направляющие* (т.е. неколлинеарные и параллельные S) векторы, а $\vec{r}_0$ — радиус-вектор одной из точек этой плоскости. Скалярная форма задания плоскости в этом случае имеет вид

$$\begin{cases} x = x_0 + \varphi p_x + \theta q_x, \\ y = y_0 + \varphi p_y + \theta q_y, \\ z = z_0 + \varphi p_z + \theta q_z, \end{cases} \qquad \varphi, \theta \in (-\infty, +\infty) .$$

2) В *прямоугольной* системе координат сфера радиусом R с центром в точке $\left\| \begin{matrix} x_0 \\ y_0 \\ z_0 \end{matrix} \right\|$ параметрически может быть задана так:

$$\begin{cases} x = x_0 + R\cos\varphi\sin\theta, \\ y = y_0 + R\,\sin\varphi\sin\theta, \\ z = z_0 + R\cos\theta, \end{cases} \qquad \varphi \in [0, 2\pi), \quad \theta \in [0, \pi] ,$$

или же уравнением $\quad (x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2 = R^2 .$

Определение 4.2.2

Поверхность называется *алгебраической*, если ее уравнение в декартовой системе координат имеет вид

$$\sum_{k=0}^{m} \alpha_k x^{p_k} y^{q_k} z^{r_k} = 0 ,$$

где p_k, q_k и r_k — целые неотрицательные числа, а числа α_k не равны нулю одновременно.

Определение 4.2.3

Число $N = \max\limits_{k=[0,m]} \{p_k + q_k + r_k\}$ называется *порядком алгебраического уравнения*, указанного в определении 4.2.2, где максимум находится по всем k, для которых $\alpha_k \neq 0$.

Наименьший из порядков алгебраических уравнений, задающих данную алгебраическую поверхность, называется *порядком алгебраической поверхности*.

Примеры алгебраических поверхностей различных порядков приведены в табл. 4.2.1.

Т а б л и ц а 4.2.1

Название	**Уравнение**	**Порядок**
Плоскость	$Ax + By + Cz + D = 0,$ $\lvert A\rvert + \lvert B\rvert + \lvert C\rvert > 0$	1
Прямой круговой цилиндр	$x^2 + y^2 - 1 = 0 \ \ \forall z$	2
Сфера	$x^2 + y^2 + z^2 - 9 = 0$	2
Гиперболический параболоид	$xy + z = 0$	2

Теорема 4.2.1 **Порядок алгебраической поверхности не зависит от выбора декартовой системы координат.**

Доказательство.

Аналогично доказательству теоремы 4.1.1.

Теорема доказана.

Замечание 4.2.1. Тела в пространстве (равно как и фигуры на плоскости) можно задавать, используя условия типа *неравенств*.

Например, шар радиусом R с центром в начале координат может быть задан неравенством

$$x^2 + y^2 + z^2 \le R^2 .$$

4.3. Цилиндрические и конические поверхности

Пусть в пространстве заданы система координат $\{\ O,\ \vec{g}_1,\ \vec{g}_2,\ \vec{g}_3\}$ и некоторая линия $\vec{r} = \vec{F}(\varphi) \quad \varphi \in \Omega$, которую будем называть *направляющей*. Проведем через каждую точку направляющей линии прямую, называемую *образующей*, параллельную некоторому фиксированному ненулевому вектору $\vec{a}$.

Определение 4.3.1	Совокупность всех точек пространства, лежащих на образующих данного вида, называется *цилиндрической поверхностью*.

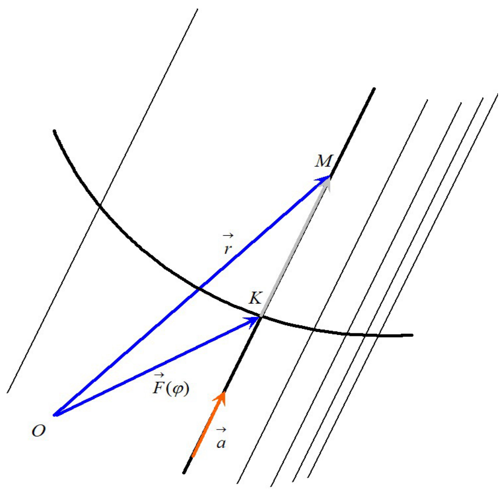


Рис. 4.1. Цилиндрическая поверхность

Составим уравнение цилиндрической поверхности в общем виде. Во введенных обозначениях (см. рис. 4.1):

$$\vec{r} = \vec{F}(\varphi) + \vec{KM}\ ,$$

а по определению цилиндрической поверхности $\vec{KM} = \theta\,\vec{a}$ и, следовательно, уравнение этой поверхности в векторной форме имеет вид

$$\vec{r}(\varphi,\theta) = \vec{F}(\varphi) + \theta\,\vec{a} \qquad \varphi \in \Omega,\ \theta \in (-\infty, +\infty)\,.$$

Пусть в координатах $\left\|\vec{F}(\varphi)\right\|_g = \left\|\begin{matrix} F_x(\varphi) \\ F_y(\varphi) \\ F_y(\varphi) \end{matrix}\right\|$ и $\left\|\vec{a}\right\|_g = \left\|\begin{matrix} a_x \\ a_y \\ a_z \end{matrix}\right\|$.

Тогда после исключения θ получаем

$$\frac{x - F_x(\varphi)}{a_x} = \frac{y - F_y(\varphi)}{a_y} = \frac{z - F_z(\varphi)}{a_z}\,.$$

Пример 4.3.1 *Прямая круговая цилиндрическая поверхность*, для которой в ортонормированной системе координат

— направляющей служит окружность радиуса 3, лежащая в плоскости, перпендикулярной оси Oz, с центром в начале координат,

— а образующими являются прямые, перпендикулярные этой плоскости,

задается системой условий $\begin{cases} x = 3\cos\varphi, \\ y = 3\sin\varphi, \\ z = \theta, \end{cases}$ поскольку в рассматриваемом случае

$$\left\|\vec{F}(\varphi)\right\|_e = \left\|\begin{matrix} 3\cos\varphi \\ 3\sin\varphi \\ 0 \end{matrix}\right\| \text{ и } \left\|\vec{a}\right\|_e = \left\|\begin{matrix} 0 \\ 0 \\ 1 \end{matrix}\right\|.$$

Заметим, что если из данной системы исключить и параметр φ, то получится уравнение вида $x^2 + y^2 = 9\ \ \forall z$. Откуда следует, что данная поверхность алгебраическая, порядка $N = 2$.

Проведем через каждую точку *направляющей* линии *образующую* прямую иначе, чем в цилиндрическом случае: пусть образующая прямая проходит через некоторую, не принадлежащую направляющей линии, фиксированную точку A (называемую *вершиной*), радиус-вектор которой $\vec{r}_0$.

Определение 4.3.2	Совокупность всех точек пространства, лежащих на образующих данного вида, называется *конической поверхностью*.

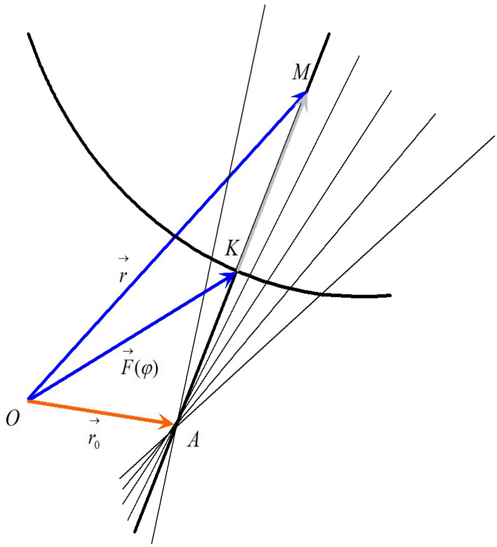


Рис. 4.2. Коническая поверхность

Составим уравнение конической поверхности в общем виде. Аналогично цилиндрическому случаю, во введенных обозначениях

$$\vec{r} = \vec{F}(\varphi) + \vec{KM}$$

(см. рис. 4.2), а по определению конической поверхности

$$\vec{KM} = \theta \cdot \left(\vec{r}_0 - \vec{F}(\varphi)\right)$$

и, следовательно, уравнение этой поверхности в векторной форме имеет вид

$$\vec{r}(\varphi, \theta) = \vec{F}(\varphi) + \theta \left(\vec{r}_0 - \vec{F}(\varphi)\right) \quad \varphi \in \Omega, \ \theta \in (-\infty, +\infty)\,.$$

Пусть в координатах $\left\|\vec{F}(\varphi)\right\|_g = \left\| \begin{matrix} F_x(\varphi) \\ F_y(\varphi) \\ F_y(\varphi) \end{matrix} \right\|$ и $\left\|\vec{r}_0\right\|_g = \left\| \begin{matrix} x_0 \\ y_0 \\ z_0 \end{matrix} \right\|$, тогда после исключения θ получаем

$$\frac{x - F_x(\varphi)}{x_0 - F_x(\varphi)} = \frac{y - F_y(\varphi)}{y_0 - F_y(\varphi)} = \frac{z - F_z(\varphi)}{z_0 - F_z(\varphi)}.$$

Пример 4.3.2

Прямая круговая коническая поверхность, для которой в ортонормированной системе координат

— направляющей служит окружность радиусом 3, лежащая в плоскости, перпендикулярной оси аппликат, с центром в начале координат,

— а образующими являются прямые, проходящие через вершину $\left\|\vec{r}_0\right\|_e = \left\| \begin{matrix} 0 \\ 0 \\ -1 \end{matrix} \right\|$,

задается системой условий (см. пример 4.3.1):

$$\frac{x - 3\cos\varphi}{-3\cos\varphi} = \frac{y - 3\sin\varphi}{-3\sin\varphi} = \frac{z}{-1}\,, \qquad \varphi \in [0, 2\pi)\,.$$

Заметим, что если из полученных соотношений также исключить и параметр φ, то получится уравнение вида

$$\frac{x^2}{9} + \frac{y^2}{9} - (z+1)^2 = 0\,.$$

Откуда следует, что данная поверхность алгебраическая, порядка $N = 2$.

4.4. Линии второго порядка на плоскости

Выбор базиса или системы координат является субъективным фактором, то есть он зависит от преференций пользователя координатного метода.

Поэтому в случае неоднозначности выбора системы координат возникает естественная потребность в методах, позволяющих сравнивать описания геометрических объектов, выполненные в разных системах. Примером подобного инструментария являются *формулы перехода*, рассмотренные в § 1.8.

Понятно, что использование формул перехода приводит к дополнительным затратам вычислительных ресурсов, и потому данное обстоятельство можно считать недостатком координатного метода.

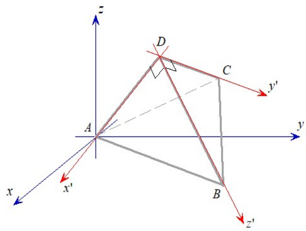


Рис. 4.3

Однако, с другой стороны, свободу выбора системы координат можно использовать для упрощения постановки задачи и/или метода решения.

Например, объем пирамиды $ABCD$, у которой известны координаты вершин и плоские углы при вершине D прямые, проще находится в системе координат $\{Dx'y'z'\}$, чем в системе $\{Axyz\}$ (см. рис. 4.3).

В этом параграфе мы рассмотрим метод, позволяющий находить ортонормированную систему координат, в которой аналитическое описание геометрического объекта – плоской линии второго порядка — существенно упрощено по сравнению с общим случаем. Что, в свою очередь, позволяет облегчить исследование свойств такого объекта.

Пусть на плоскости дана ортонормированная система координат $\{ O, \vec{e}_1, \vec{e}_2 \}$ и некоторая линия L.

Определение 4.4.1	Если линия L является алгебраической линией второго порядка, то (в соответствии с определениями 4.1.2 и 4.1.3) ее уравнение в данной системе координат может иметь вид $$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0\,, \quad (4.4.1)$$ где числа A, B и C не равны нулю одновременно, а x и y суть координаты радиуса-вектора точки, принадлежащей L.

Поскольку коэффициенты уравнения 4.4.1 зависят от выбора системы координат, при исследовании свойств линий второго порядка целесообразно предварительно перейти к другой системе координат $\left\{ O', \vec{e}'_1, \vec{e}'_2 \right\}$, в которой вид уравнения линии оказывается *наиболее простым.*

`Теорема 4.4.1` **Для любой линии второго порядка существует ортонормированная система координат, в которой уравнение этой линии (при $a > 0$, $b > 0$, $p > 0$) имеет один из следующих девяти (называемых *каноническими*) видов:**

Таблица 4.4.1

Тип линии **Вид линии**	Эллиптический $\Delta > 0$	Гиперболический $\Delta < 0$	Параболический $\Delta = 0$
Пустые множества	$\dfrac{x'^2}{a^2} + \dfrac{y'^2}{b^2} = -1$		$y'^2 = -a^2 \quad \forall x'$
Изолированные точки	$\dfrac{x'^2}{a^2} + \dfrac{y'^2}{b^2} = 0$		
Совпадающие прямые			$y'^2 = 0 \quad \forall x'$
Несовпадающие прямые		$\dfrac{x'^2}{a^2} - \dfrac{y'^2}{b^2} = 0$	$y'^2 = a^2 \quad \forall x'$
Кривые	*Эллипс* $\dfrac{x'^2}{a^2} + \dfrac{y'^2}{b^2} = 1$	*Гипербола* $\dfrac{x'^2}{a^2} - \dfrac{y'^2}{b^2} = 1$	*Парабола* $y'^2 = 2px'$

где $\Delta = \det\begin{Vmatrix} A & B \\ B & C \end{Vmatrix} = AC - B^2$. **Мы также потребуем, чтобы для канонического уравнения** *эллипса* **выполнялось** $a \geq b$**.**

`Доказательство.`

1. Предварительно заметим, что без потери общности можно считать выполненными условия: $B \geq 0$ и $A \geq C$.

   Действительно, если $B < 0$, то можно изменить знаки всех коэффициентов в уравнении 4.4.1.

   Если же $A < C$, то, перейдя к новой ортонормированной системе координат, для которой $\vec{e'}_1 = \vec{e}_2,\ \vec{e'}_2 = \vec{e}_1,\ \vec{OO'} = \vec{o}$, мы получим желаемое соотношение, поскольку при таком переходе справедливы равенства $x = y',\ y = x'$ в силу утверждений § 1.8. Заметим также, что при этой замене Δ не меняется, поскольку

$$\det\left\|\begin{array}{cc} C & B \\ B & A \end{array}\right\| = \det\left\|\begin{array}{cc} A & B \\ B & C \end{array}\right\| = \Delta\,.$$

2. Если $B = 0$, то переходим к пункту 4 на с. 125.

   Если же $B > 0$, то выбираем новую ортонормированную систему координат $\left\{ O',\, \vec{e'}_1,\, \vec{e'}_2 \right\}$, получаемую из исходной поворотом против часовой стрелки вокруг точки O на угол $0 < \alpha \leq \frac{\pi}{4}$, такой, чтобы коэффициент при произведении $x'y'$ оказался равным нулю.

   Выведем правило выбора этого угла. Для такого поворота (см. § 1.8):

$$\left\{\begin{array}{rcl} \vec{e'}_1 & = & \vec{e}_1\cos\alpha + \vec{e}_2\sin\alpha, \\ \vec{e'}_2 & = & -\vec{e}_1\sin\alpha + \vec{e}_2\cos\alpha, \\ \vec{OO'} & = & \vec{o}\,. \end{array}\right.$$

   Тогда формулы перехода от $\left\{ O,\, \vec{e}_1,\, \vec{e}_2 \right\}$ к $\left\{ O',\, \vec{e'}_1,\, \vec{e'}_2 \right\}$ будут иметь вид

$$\left\{\begin{array}{rcl} x & = & x'\cos\alpha - y'\sin\alpha, \\ y & = & x'\sin\alpha + y'\cos\alpha. \end{array}\right.$$

   Подставляя выражения «старых» координат через «новые», получаем уравнение 4.4.1 в виде

$$A\Big(x'\cos\alpha - y'\sin\alpha\Big)^2 + 2B\Big(x'\cos\alpha - y'\sin\alpha\Big)\Big(x'\sin\alpha + y'\cos\alpha\Big) +$$

$$+C\Big(x'\sin\alpha + y'\cos\alpha\Big)^2 + 2D\Big(x'\cos\alpha - y'\sin\alpha\Big) +$$

$$+2E\Big(x'\sin\alpha + y'\cos\alpha\Big) + F = 0$$

или же

$$A'x'^2 + 2B'x'y' + C'y'^2 + 2D'x' + 2E'y' + F' = 0\,,$$

где

$$\left\{\begin{array}{rcl} A' & = & A\cos^2\alpha + 2B\cos\alpha\sin\alpha + C\sin^2\alpha, \\ 2B' & = & -2A\sin\alpha\cos\alpha + 2B\cos^2\alpha - 2B\sin^2\alpha + 2C\sin\alpha\cos\alpha, \\ C' & = & A\sin^2\alpha - 2B\sin\alpha\cos\alpha + C\cos^2\alpha\,. \end{array}\right.$$

Из условия $B' = 0$ следует: $\quad 2B\cos 2\alpha - (A-C)\sin 2\alpha = 0\,,$

т.е. окончательно: $\alpha = \dfrac{\pi}{4}$ при $A = C$ или же при $A > C$

$$\operatorname{tg} 2\alpha = \frac{2B}{A-C} \quad\Longrightarrow\quad \alpha = \frac{1}{2}\operatorname{acrtg}\frac{2B}{A-C}\,,$$

то есть искомый угол найден.

Заметим, что угол также может быть найден из равносильного уравнения $\quad \operatorname{tg}^2\alpha + \dfrac{A-C}{B}\operatorname{tg}\alpha - 1 = 0\,.$

3. Убедимся, что при такой замене координат величины Δ и $A + C$ сохраняют свои значения. Действительно, из соотношений

$$1 + \operatorname{ctg}^2 2\alpha = 1 + \left(\frac{A-C}{2B}\right)^2 = \frac{1}{\sin^2 2\alpha}$$

и неравенства $0 < \alpha \le \dfrac{\pi}{4}$ получаем

$$\sin 2\alpha = \frac{2B}{\sqrt{4B^2 + (A-C)^2}} \quad \text{и} \quad \cos 2\alpha = \frac{A-C}{\sqrt{4B^2 + (A-C)^2}}\,.$$

Учитывая полученные равенства, из выражения для значения A' имеем

$$A' = A\frac{1+\cos 2\alpha}{2} + B\sin 2\alpha + C\frac{1-\cos 2\alpha}{2} =$$

$$= \frac{A+C}{2} + \frac{1}{2}\sqrt{4B^2 + (A-C)^2}\,.$$

Аналогичным образом получаем, что

$$C' = \frac{A+C}{2} - \frac{1}{2}\sqrt{4B^2 + (A-C)^2}\,.$$

Теперь находим

$$\Delta' = \det\left\|\begin{array}{cc} A' & 0 \\ 0 & C' \end{array}\right\| = A'C' =$$

$$= \left(\frac{A+C}{2}\right)^2 - \frac{1}{4}\Big(4B^2 + (A-C)^2\Big) =$$

$$= AC - B^2 = \det\left\|\begin{array}{cc} A & B \\ B & C \end{array}\right\| = \Delta\,,$$

то есть величина Δ не меняется при выполненной замене системы координат. Также очевидно, что при этом выполняется равенство $A' + C' = A + C\,.$

4. В дальнейших рассуждениях будем полагать, что $B = 0$, и рассмотрим отдельно случаи $\Delta \neq 0$ и $\Delta = 0$ для уравнения вида $Ax^2 + Cy^2 + 2Dx + 2Ey + F = 0\,.$

   Пусть $\Delta \neq 0$. Это означает, что $A \neq 0$ и $C \neq 0$ и уравнение линии может быть переписано в виде

$$A\left(x + \frac{D}{A}\right)^2 + C\left(y + \frac{E}{C}\right)^2 = \frac{D^2}{A} + \frac{E^2}{C} - F\,.$$

   Обозначим $P = \dfrac{D^2}{A} + \dfrac{E^2}{C} - F\,,$ тогда, перейдя к новой ортонормированной системе координат

$$
\left\{\begin{array}{lcl} \vec{e'}_1 & = & \vec{e}_1\,, \\ \vec{e'}_2 & = & \vec{e}_2\,, \\ \vec{OO'} & = & -\dfrac{D}{A}\vec{e}_1 - \dfrac{E}{C}\vec{e}_2 \end{array}\right. \qquad \left\{\begin{array}{lcl} x & = & x' - \dfrac{D}{A}\,, \\ \\ y & = & y' - \dfrac{E}{C}\,, \end{array}\right.
$$

получим $\quad Ax'^2 + Cy'^2 = P$. Откуда следует, что

$$
\pm\frac{x'^2}{\left|\dfrac{P}{A}\right|} \pm \frac{y'^2}{\left|\dfrac{P}{C}\right|} = \pm 1 \qquad \text{при} \quad P \neq 0\,,
$$

$$
\pm\frac{x'^2}{|\,C\,|} \pm \frac{y'^2}{|\,A\,|} = 0 \qquad \text{при} \quad P = 0\,,
$$

и мы приходим, таким образом, к одному из шести следующих уравнений:

$$
\frac{x'^2}{a^2} + \frac{y'^2}{b^2} = 0\,, \quad \frac{x'^2}{a^2} + \frac{y'^2}{b^2} = 1\,, \quad \frac{x'^2}{a^2} + \frac{y'^2}{b^2} = -1 \quad \text{для } \Delta > 0\,,
$$

$$
\frac{x'^2}{a^2} - \frac{y'^2}{b^2} = 0\,, \quad \frac{x'^2}{a^2} - \frac{y'^2}{b^2} = 1\,, \quad \frac{x'^2}{a^2} - \frac{y'^2}{b^2} = -1 \quad \text{для } \Delta < 0\,.
$$

Первые пять из этих случаев содержатся в формулировке теоремы, а шестой сводится к пятому умножением обеих частей уравнения на -1 с последующим переобозначением: $x = y'$ и $y = x'$.

5. Пусть $\Delta = 0$. Это означает, что $AC = 0$, то есть либо $A = 0$, либо $C = 0$ (но не вместе!). Пусть $A = 0$ (если это не так, то взаимно переобозначим переменные x' и y'), тогда уравнение линии $Cy^2 + 2Dx + 2Ey + F = 0$ может быть в силу $C \neq 0$ записано в виде

$$
C\left(y + \frac{E}{C}\right)^2 = \frac{E^2}{C} - F - 2Dx\,.
$$

При $D = 0$ получаем $\quad C\left(y + \dfrac{E}{C}\right)^2 = \dfrac{E^2}{C} - F$, то есть одно из трех уравнений $\quad y'^2 = a^2\,,\;\; y'^2 = 0\,,\;\; y'^2 = -a^2\,.$

Если же $D \neq 0$, то уравнение можно привести к виду

$$\left(y + \frac{E}{C}\right)^2 = -\frac{2D}{C}\left(x - \frac{1}{2D}\left(\frac{E^2}{C} - F\right)\right)$$

и, таким образом, либо $\; y'^2 = 2px'$, либо $\; y'^2 = -2px'$, где $p > 0$.

Первый из этих случаев указан в формулировке теоремы, а второй сводится к первому заменой координат:

$$\left\{\begin{array}{rcr} \vec{e'}_1 & = & -\vec{e}_1\,, \\ \vec{e'}_2 & = & \vec{e}_2\,, \\ \vec{OO'} & = & \vec{o} \end{array}\right. \qquad \left\{\begin{array}{rcr} x & = & -x'\,, \\ y & = & y'\,, \end{array}\right.$$

и, таким образом, для линии второго порядка на плоскости получены все девять канонических случаев.

Теорема доказана.

Замечания. 1. В теореме 4.1.1 было показано, что порядок алгебраической линии, в том числе и для рассматриваемых в теореме 4.4.1 случаев, не меняется при замене системы координат.

2. Из доказательства теоремы также следует, что поворот и параллельный перенос ортонормированной системы координат не допускают перемещения уравнения линии второго порядка из одной строки таблицы, приведенной в формулировке теоремы 4.4.1, в другую. Более того, в дальнейшем будет показано (см. § 5.4), что никакой заменой общей декартовой системы координат нельзя переместить линию второго порядка, находящуюся в одной из клеток таблицы в условии теоремы 4.4.1, в другую клетку.

3. Пустое множество эллиптического типа иногда называют *мнимым эллипсом*, а пустое множество параболического типа – *парой мнимых параллельных прямых*.

4. Алгоритм доказательства теоремы 4.4.1 можно использовать как для нахождения канонического вида

уравнения линии второго порядка, так и для построения канонической системы координат, то есть системы координат, в которой данная линия второго порядка имеет канонический вид.

Исследование конкретных свойств различных типов линий второго порядка приводится в приложении 1.

4.5. Поверхности второго порядка в пространстве

Пусть в пространстве даны ортонормированная система координат $\{ O, \vec{e}_1, \vec{e}_2, \vec{e}_3 \}$ и некоторая поверхность S.

Определение 4.5.1	Пусть поверхность S является алгебраической поверхностью второго порядка, тогда (в соответствии с определениями 4.2.2 и 4.2.3) ее уравнение в данной системе координат может иметь вид $$\begin{aligned} & A_{11}x^2 + A_{22}y^2 + A_{33}z^2 + \\ & +2A_{12}xy + 2A_{13}xz + 2A_{23}yz + \\ & +2A_{14}x + 2A_{24}y + 2A_{34}z + A_{44} = 0 \,, \end{aligned} \qquad (4.5.1)$$ где числа A_{11}, A_{22}, A_{33}, A_{12}, A_{13}, и A_{23} не равны нулю одновременно, а x, y и z суть координаты радиуса-вектора точки, принадлежащей S.

Как и в плоском случае, коэффициенты уравнения (4.5.1) зависят от выбора системы координат, поэтому при исследовании свойств поверхностей второго порядка целесообразно предварительно перейти в ту систему координат, для которой уравнение поверхности оказывается наиболее простым.

Теорема 4.5.1 **Для любой поверхности второго порядка существует ортонормированная система координат, в которой уравнение этой поверхности имеет один из следующих семнадцати (называемых *каноническими*) видов:**

Таблица 4.5.1

Пустое множество $\dfrac{x'^2}{a^2}+\dfrac{y'^2}{b^2}+\dfrac{z'^2}{c^2}=-1$	Изолированная точка $\dfrac{x'^2}{a^2}+\dfrac{y'^2}{b^2}+\dfrac{z'^2}{c^2}=0$	Эллиптический цилиндр $\dfrac{x'^2}{a^2}+\dfrac{y'^2}{b^2}=1 \quad \forall z'$
Пустое множество $\dfrac{x'^2}{a^2}+\dfrac{y'^2}{b^2}=-1 \quad \forall z'$	Прямая $\dfrac{x'^2}{a^2}+\dfrac{y'^2}{b^2}=0 \quad \forall z'$	Гиперболический цилиндр $\dfrac{x'^2}{a^2}-\dfrac{y'^2}{b^2}=1 \quad \forall z'$
Пустое множество $x'^2=-a^2 \quad \forall y',\, z'$	Пара пересекающихся плоскостей $\dfrac{x'^2}{a^2}-\dfrac{y'^2}{b^2}=0 \quad \forall z'$	Параболический цилиндр $y'^2=2px' \quad \forall z'$
	Пара совпадающих $x'^2=0 \qquad \forall y', z'$ или параллельных плоскостей $x'^2=a^2 \qquad \forall y', z'$	Конус $\dfrac{x'^2}{a^2}+\dfrac{y'^2}{b^2}-\dfrac{z'^2}{c^2}=0$

Таблица 4.5.2

Эллипсоид $\dfrac{x'^2}{a^2}+\dfrac{y'^2}{b^2}+\dfrac{z'^2}{c^2}=1$	Эллиптический параболоид $\dfrac{x'^2}{a^2}+\dfrac{y'^2}{b^2}=2z'$	Однополостный гиперболоид $\dfrac{x'^2}{a^2}+\dfrac{y'^2}{b^2}-\dfrac{z'^2}{c^2}=1$
	Гиперболический параболоид $\dfrac{x'^2}{a^2}-\dfrac{y'^2}{b^2}=2z'$	Двуполостный гиперболоид $\dfrac{x'^2}{a^2}-\dfrac{y'^2}{b^2}-\dfrac{z'^2}{c^2}=1$

где $a > 0,\ b > 0,\ c > 0,\ p > 0$.

Доказательство.

Хотя возможно доказать существование ортонормированной системы координат с требуемыми свойствами, применив подход, аналогичный использованному при доказательстве теоремы 4.4.1, представляется целесообразным рассмотреть этот вопрос в рамках теории евклидовых пространств, где утверждение теоремы 4.5.1 непосредственно вытекает из более общего случая, рассмотренного в § 12.2.

Теорема доказана.

Исследование свойств конкретных типов поверхностей второго порядка приводится в приложении 2.

4.6. Альтернативные системы координат

В ряде практических приложений оказывается целесообразным использование систем координат, отличных от декартовой.

Полярная система координат

Примером альтернативной системы координат на плоскости является *полярная система координат.*

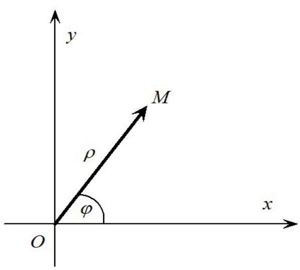


Рис. 4.4

Положение точки на плоскости в этой системе координат задается парой упорядоченных чисел $\{\rho, \varphi\}$, где $\rho = \left|\vec{OM}\right|$, $\varphi = \angle OMx$ удовлетворяют ограничениям $\rho \geq 0$ и $0 \leq \varphi < 2\pi$.

Точка O называется *полюсом*, а луч Ox — *полярной осью.* Угол φ отсчитывается против часовой стрелки (рис. 4.4). Для полюса этот угол не определяется.

Формулы перехода от ортонормированной декартовой системы координат к полярной и обратно имеют следующий вид:

$$\left\{\begin{array}{l} x = \rho\cos\varphi\,, \\ \\ y = \rho\sin\varphi\,, \end{array}\right. \qquad \left\{\begin{array}{c} \rho = \sqrt{x^2+y^2}\,, \\ \cos\varphi = \dfrac{x}{\sqrt{x^2+y^2}}, \quad \sin\varphi = \dfrac{y}{\sqrt{x^2+y^2}}. \end{array}\right.$$

Использование полярной системы координат позволяет упростить описание объектов, обладающих точечной симметрией. Например, окружность единичного радиуса с центром в начале координат, имеющая в ортонормированной декартовой системе координат уравнение $x^2 + y^2 = 1$, в полярной системе координат задается условием $\rho = 1$.

Более того, в приложении 1 показано, что в полярной системе координат три различных типа линий второго порядка — эллипс, гипербола и парабола – задаются одним и тем же уравнением

$$\rho(1 - \varepsilon\cos\varphi) - p = 0\,, \tag{4.6.1}$$

где $p > 0$ и $\varepsilon > 0$ — некоторые константы, называемые соответственно *эксцентриситетом* и *фокальным параметром*, и что для различных значений ε при фиксированном p получаются различные типы кривых: эллипсы при $0 < \varepsilon < 1$, параболы при $\varepsilon = 1$ и гиперболы при $\varepsilon > 1$. Соответствующие случаи показаны на рис. 4.5.

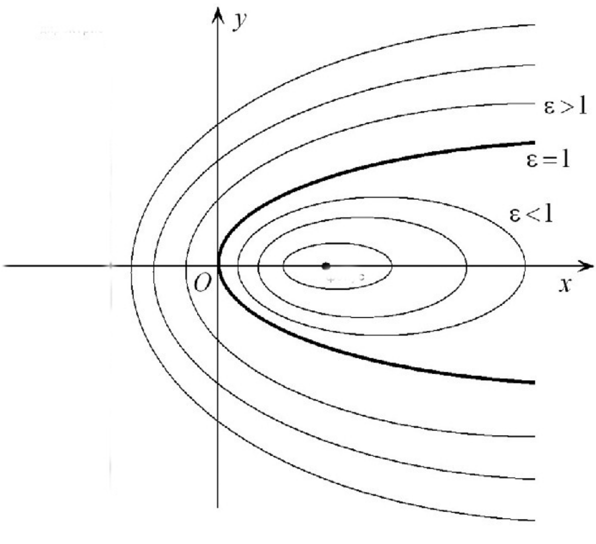


Рис. 4.5. Линии второго порядка с разными значениями ε

Проверим справедливость этого утверждения, выполнив в уравнении (4.6.1) переход от полярной к ортонормированной декартовой системе координат. Действительно, поскольку

$$\rho = \sqrt{x^2 + y^2} \qquad \text{и} \qquad \cos\varphi = \frac{x}{\sqrt{x^2 + y^2}}\,,$$

то уравнение (4.6.1) может быть преобразовано к виду

$$(1-\varepsilon^2)x^2 + y^2 = 2\varepsilon p x + p^2 \,.$$

Если $\varepsilon = 1$, то это уравнение параболы. Если же $\varepsilon \neq 1$, то, выделив полные квадраты, получим

$$\left(x - \frac{\varepsilon p}{1-\varepsilon^2}\right)^2 + \frac{1}{1-\varepsilon^2}y^2 = \frac{p^2}{(1-\varepsilon^2)^2}\,.$$

Теперь нетрудно заметить, что при $0 < \varepsilon < 1$ это эллипс, а при $\varepsilon > 1$ — гипербола.

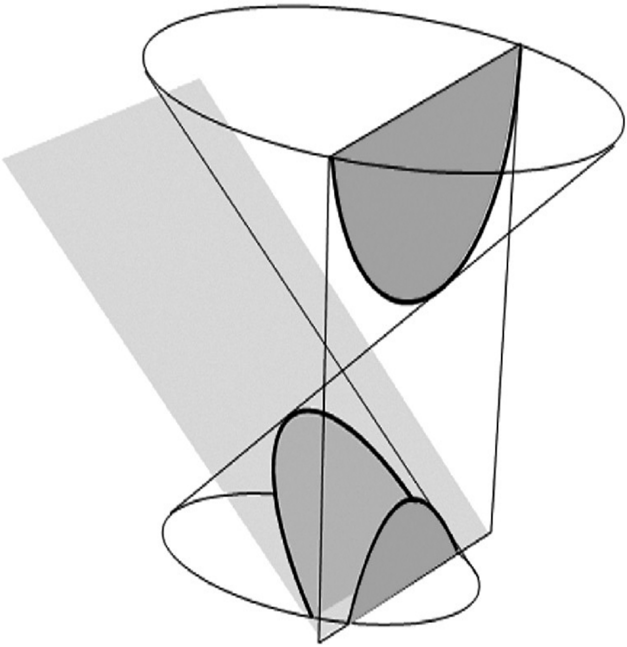


Рис. 4.6. Конические сечения

Далее, если ослабить ограничения на параметры уравнения (4.6.1), разрешив им принимать (в смысле предельного перехода) как нулевые, так и бесконечно большие положительные значения, то можно получить и другие виды линий второго порядка, указанные в формулировке теоремы 4.4.1.

Так, например, при $\varepsilon = 0$ мы имеем для $p > 0$ окружность, а для $p = 0$ — изолированную точку.

Наконец, поскольку линии пересечения конической поверхности некоторой плоскостью являются линиями второго порядка (это иллюстрирует рис. 4.6), можно дать

Определение 4.6.1	Линия, уравнение которой в полярной системе координат имеет вид $$\rho(1-\varepsilon\cos\varphi) - p = 0\,, \quad \text{где } p \geq 0 \text{ и } \varepsilon \geq 0\,,$$ называется *коническим сечением.*

Сферическая система координат

В ряде практических приложений, требующих аналитического исследования пространственных объектов, используется так называемая *сферическая система координат.*

Положение точки M в пространстве в этой системе однозначно задается при помощи тройки чисел $\{\rho,\ \varphi,\ \theta\}$, где (см. рис. 4.7):

$$\rho = \left|\vec{OM}\right|\,, \qquad \varphi = \angle POx\,,$$

$$\theta = \angle MOz\,.$$

Эти числа удовлетворяют ограничениям

$$\rho \geq 0, \qquad 0 \leq \varphi < 2\pi\,, \qquad 0 \leq \theta \leq \pi\,.$$

Использование сферической системы координат иногда позволяет получить более простое аналитическое описание геометрических объектов, обладающих точечной симметрией.

Например, уравнение сферы единичного радиуса с центром в начале координат в сферической системе будет иметь вид $\rho = 1$.

Формулы перехода между ортонормированной декартовой системой координат и сферической имеют следующий вид:

$$\begin{cases} x = \rho\cos\varphi\sin\theta\,, \\ y = \rho\,\sin\varphi\sin\theta\,, \\ z = \rho\,\cos\theta\,, \end{cases}$$

а для обратного перехода соответственно

$$\begin{cases} \rho = \sqrt{x^2+y^2+z^2}\,, \\ \cos\varphi = \dfrac{x}{\sqrt{x^2+y^2}}\,, \quad \sin\varphi = \dfrac{y}{\sqrt{x^2+y^2}}\,, \\ \cos\theta = \dfrac{z}{\sqrt{x^2+y^2+z^2}}\,. \end{cases}$$

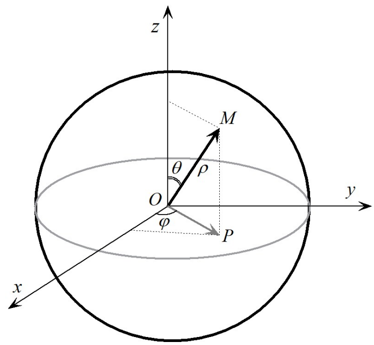


Рис. 4.7. Сферическая система координат

Цилиндрическая система координат

В тех случаях, когда исследуемый пространственный объект обладает осевой симметрией, может оказаться удобным применение *цилиндрической системы координат.*

В тех случаях, когда исследуемый пространственный объект обладает осевой симметрией, может оказаться удобным применение цилиндрической системы координат.

Положение точки M в пространстве в этой системе однозначно задается при помощи упорядоченной тройки чисел $\{\rho,\, \varphi,\, h\}$, где (см. рис. 4.8)

$$\rho = \left|\vec{OP}\right|, \qquad \varphi = \angle POx\,,$$

а h есть z-координата точки M.

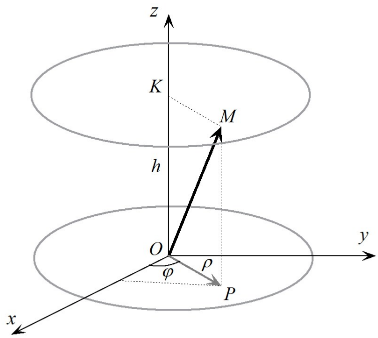


Рис. 4.8. Цилиндрическая система координат

Значения этих чисел удовлетворяют ограничениям

$$\rho \geq 0, \qquad 0 \leq \varphi < 2\pi\,, \qquad h \in (-\infty, +\infty)\,.$$

Формулы перехода от ортонормированной декартовой системы координат к цилиндрической и обратно имеют вид

$$\left\{\begin{array}{l} x = \rho \cos\varphi\,, \\ y = \rho\sin\varphi\,, \\ z = h\,, \end{array}\right. \qquad \left\{\begin{array}{rcl} \rho & = & \sqrt{x^2+y^2}\,, \\ \cos\varphi & = & \dfrac{x}{\sqrt{x^2+y^2}}, \quad \sin\varphi = \dfrac{y}{\sqrt{x^2+y^2}}\,, \\ h & = & z\,. \end{array}\right.$$

5. Преобразования плоскости

5.1. Умножение матриц

Определение 5.1.1	Матрица $\|C\|$ размера $m \times n$ с элементами $$\gamma_{ij} \quad \forall i = \overline{1,m}\,,\ \forall j = \overline{1,n}$$ называется *произведением* матрицы $\|A\|$ размера $m \times l$ с элементами $$\alpha_{ik} \quad \forall i = \overline{1,m}\,,\ \forall k = \overline{1,l}$$ на матрицу $\|B\|$ размера $l \times n$ с элементами $$\beta_{kj} \quad \forall k = \overline{1,l}\,,\ \forall j = \overline{1,n}\,,$$ где $$\gamma_{ij} = \sum_{k=1}^{l} \alpha_{ik}\beta_{kj} \quad \forall i = \overline{1,m}\,,\ \forall j = \overline{1,n}\,.$$

Результат *умножения матриц* — матрица $\|C\|$ — есть матрица размера $m \times n$ при любом натуральном l. Произведение матриц обозначается как $\|C\| = \|A\|\,\|B\|$. Правило вычисления компонентов произведения по компонентам сомножителей иллюстрирует рис. 5.1.

Пример 5.1.1

Приведем результаты умножения матриц, имеющих не более чем пару строк или столбцов.

1. Пусть размер $\|A\|$ есть 2×2, а размер $\|B\|$ — 2×1, тогда размер $\|C\|$ будет 2×1.

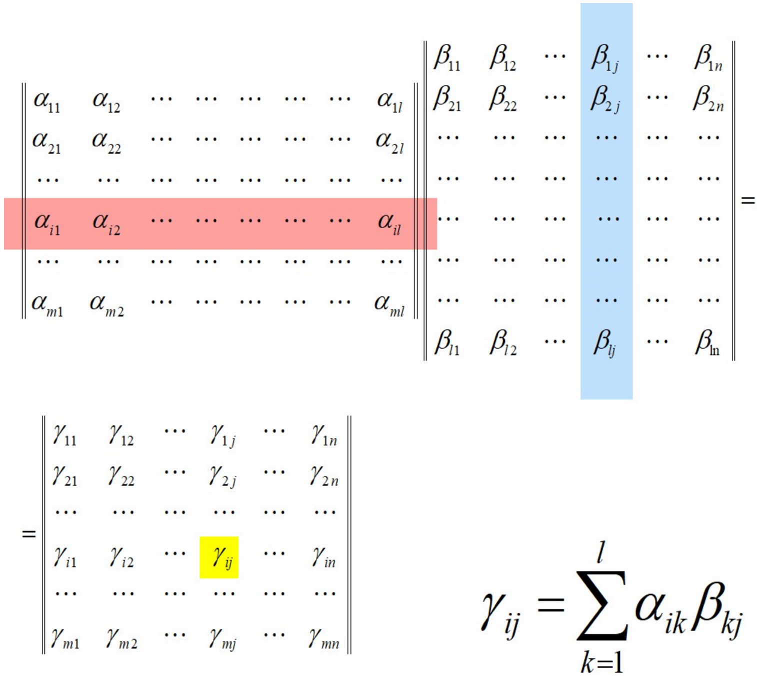


Рис. 5.1. Умножение матриц

Конкретно в этом случае

$$\|C\| = \|A\|\,\|B\| =$$

$$= \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\| \left\| \begin{matrix} \beta_{11} \\ \beta_{21} \end{matrix} \right\| = \left\| \begin{matrix} \alpha_{11}\beta_{11} + \alpha_{12}\beta_{21} \\ \alpha_{21}\beta_{11} + \alpha_{22}\beta_{21} \end{matrix} \right\| .$$

2. Если размер $\|A\|$ есть 1×2, а размер $\|B\|$ — 2×2, тогда размер $\|C\|$ будет 1×2. Тогда

$$\|C\| = \|A\|\,\|B\| = \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \end{matrix} \right\| \left\| \begin{matrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{matrix} \right\| =$$

$$= \left\| \begin{matrix} \alpha_{11}\beta_{11} + \alpha_{12}\beta_{21} & \alpha_{11}\beta_{12} + \alpha_{12}\beta_{22} \end{matrix} \right\| .$$

3. Наконец, пусть размеры $\|A\|$ и $\|B\|$ одинаковые. Для размера 1×1 очевидно, что

$$\|C\| = \|A\|\,\|B\| = \|\,\alpha_{11}\beta_{11}\,\|\ .$$

Заметим, что в силу определения 5.1.1 матрицы такого размера по алгебраическим свойствам не будут отличаться от вещественных чисел.

Если же размеры $\|A\|$ и $\|B\|$ есть 2×2, то

$$\|C\| = \|A\|\,\|B\| = \left\|\begin{array}{cc} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{array}\right\| \left\|\begin{array}{cc} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{array}\right\| =$$

$$= \left\|\begin{array}{cc} \alpha_{11}\beta_{11} + \alpha_{12}\beta_{21} & \alpha_{11}\beta_{12} + \alpha_{12}\beta_{22} \\ \alpha_{21}\beta_{11} + \alpha_{22}\beta_{21} & \alpha_{21}\beta_{12} + \alpha_{22}\beta_{22} \end{array}\right\| .$$

Замечания об умножении матриц

Из определения произведения матриц непосредственно следует, что для матриц подходящих размеров:

1) умножение матриц *некоммутативно*, то есть в общем случае

$$\|A\|\|B\| \neq \|B\|\|A\|\,,$$

2) умножение матриц *ассоциативно*

$$\|A\|\Big(\|B\|\|C\|\Big) = \Big(\|A\|\|B\|\Big)\|C\|\,,$$

3) умножение матриц обладает свойством *дистрибутивности*.

$$\|A\|\Big(\|B\| + \|C\|\Big) = \|A\|\|B\| + \|A\|\|C\|\,.$$

Легко убедиться, что умножение (как справа, так и слева) любой матрицы на подходящего размера *единичную* матрицу $\|E\|$ дает в результате ту же самую матрицу:

$$\|A\|\|E_1\| = \|A\| \qquad \text{или} \qquad \|E_2\|\|A\| = \|A\|\,.$$

Определение 5.1.2 Матрица $\|A\|^{-1}$ называется *обратной* к квадратной матрице $\|A\|$, если выполнены равенства

$$\|A\|^{-1}\|A\| = \|A\|\|A\|^{-1} = \|E\|\,.$$

Обратная матрица существует *не для произвольной* квадратной матрицы. Для существования матрицы, обратной к $\|A\|$, необходимо и достаточно, чтобы выполнялось условие $\det\|A\| \neq 0$ [7].

В связи с этим дадим определение обобщающее определение 1.1.12.

Определение 5.1.3 Квадратная матрица $\|A\|$, для которой ее определитель (детерминант) равен нулю, называется *вырожденной*, а квадратная матрица, для которой $\det\|A\| \neq 0$, — *невырожденной.*

Лемма 5.1.1 **Если обратная матрица существует, то она единственна.**

Доказательство.

Предположим, что невырожденная матрица $\|A\|$ имеет две обратные: $\|A\|_1^{-1}$ и $\|A\|_2^{-1}$. Тогда из равенств

$$\|A\|\|A\|_1^{-1} = \|E\| \qquad \text{и} \qquad \|A\|\|A\|_2^{-1} = \|E\|$$

следует, что

$$\|A\|\Big(\|A\|_1^{-1} - \|A\|_2^{-1}\Big) = \|A\|\|A\|_1^{-1} - \|A\|\|A\|_2^{-1} = \|O\|\,.$$

Умножив слева обе части данного равенства на $\|A\|_1^{-1}$, получим

$$\|A\|_1^{-1}\|A\|\Big(\|A\|_1^{-1} - \|A\|_2^{-1}\Big) = \|O\|$$

и, учтя, что $\|A\|_1^{-1}\|A\| = \|E\|$, приходим к равенству

$$\|A\|_1^{-1} - \|A\|_2^{-1} = \|O\|\,.$$

Лемма доказана.

[7] Определение и свойства детерминанта квадратной матрицы порядка n рассматриваются в гл. 6 .

В частном случае, когда $\|A\| = \left\| \begin{array}{cc} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{array} \right\|$ и если $\det\|A\| \neq 0$, ее обратная матрица имеет вид

$$\|A\|^{-1} = \frac{1}{\det\|A\|} \left\| \begin{array}{rr} \alpha_{22} & -\alpha_{12} \\ -\alpha_{21} & \alpha_{11} \end{array} \right\| . \tag{5.1.1}$$

Отметим, что для квадратных матриц порядка n справедливы[8] следующие равенства:

$$\det\bigl(\|A\|\,\|B\|\bigr) = \det\|A\| \; \det\|B\| ,$$

$$\det\|A\|^{-1} = \frac{1}{\det\|A\|}, \quad \text{если} \ \ \det\|A\| \neq 0 .$$

Пример 5.1.2 Используя матричные операции, систему линейных уравнений

$$\left\{ \begin{array}{ccccc} \alpha_{11}\xi_1 & + & \alpha_{12}\xi_2 & = & \beta_1 , \\ \alpha_{21}\xi_1 & + & \alpha_{22}\xi_2 & = & \beta_2 \end{array} \right.$$

можно записать в виде $\|A\|\,\|x\| = \|b\|$, где

$$\|x\| = \left\| \begin{array}{c} \xi_1 \\ \xi_2 \end{array} \right\| , \quad \|b\| = \left\| \begin{array}{c} \beta_1 \\ \beta_2 \end{array} \right\| , \quad \|A\| = \left\| \begin{array}{cc} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{array} \right\| ,$$

а ее решение при $\det\|A\| \neq 0$, как $\|x\| = \|A\|^{-1}\,\|b\|$.

Пример 5.1.3 Формулы перехода (1.8.2) от одной декартовой системы координат к другой с помощью матричных операций могут быть записаны в виде

$$\left\| \begin{array}{c} \vec{g}\,'_1 \\ \vec{g}\,'_2 \\ \vec{g}\,'_3 \end{array} \right\| = \|S\|^{\mathrm{T}} \left\| \begin{array}{c} \vec{g}_1 \\ \vec{g}_2 \\ \vec{g}_3 \end{array} \right\| , \quad \left\| \begin{array}{c} \xi_1 \\ \xi_2 \\ \xi_3 \end{array} \right\| = \|S\| \left\| \begin{array}{c} \xi'_1 \\ \xi'_2 \\ \xi'_3 \end{array} \right\| + \left\| \begin{array}{c} \beta_1 \\ \beta_2 \\ \beta_3 \end{array} \right\| ,$$

где $\|S\|$ — матрица перехода.

Теорема 5.1.1 **Справедливо равенство** $\bigl(\|A\|\,\|B\|\bigr)^{\mathrm{T}} = \|B\|^{\mathrm{T}}\,\|A\|^{\mathrm{T}}$.

[8] Для $n = 2$ эти равенства проверяются непосредственно по определению 1.1.9. Для квадратных матриц порядка n их справедливость будет доказана в гл. 6.

Доказательство.

Будем предполагать, что размеры матриц $\|A\|$ и $\|B\|$ таковы, что произведения матриц, указанные в формулировке теоремы, существуют.

Пусть числа α_{ik}, β_{kj}, γ_{ij} суть элементы матриц $\|A\|$, $\|B\|$ и $\|C\| = \|A\|\|B\|$ соответственно. Тогда, согласно определению 5.1.1,

$$\gamma_{ij} = \sum_{k=1}^{l} \alpha_{ik}\beta_{kj}\,.$$

Но, с другой стороны, по определению 1.1.8 (операции транспонирования):

$$\gamma_{ij}^{\mathrm{T}} = \gamma_{ji} = \sum_{k=1}^{l} \alpha_{jk}\beta_{ki} = \sum_{k=1}^{l} \alpha_{kj}^{\mathrm{T}}\beta_{ik}^{\mathrm{T}} = \sum_{k=1}^{l} \beta_{ik}^{\mathrm{T}}\alpha_{kj}^{\mathrm{T}}\,.$$

Откуда, в силу определения 5.1.1, следует заключение о справедливости утверждения теоремы.

Теорема доказана.

Пример 5.1.4 Согласно правилу транспонирования произведения матриц равенство из примера 5.1.3

$$\left\|\begin{array}{c} \vec{g}\,'_1 \\ \vec{g}\,'_2 \\ \vec{g}\,'_3 \end{array}\right\| = \|S\|^{\mathrm{T}} \left\|\begin{array}{c} \vec{g}_1 \\ \vec{g}_2 \\ \vec{g}_3 \end{array}\right\|$$

может быть записано в виде

$$\|\, \vec{g}\,'_1\ \vec{g}\,'_2\ \vec{g}\,'_3 \,\| = \|\, \vec{g}_1\ \vec{g}_2\ \vec{g}_3 \,\|\, \|S\|\,.$$

Для дальнейших рассуждений нам будет полезно следующее вспомогательное утверждение.

Лемма 5.1.2 **Если произведение квадратной порядка n матрицы $\|Q\|$ на любой n-компонентный столбец $\|x\|$ есть нулевой n-компонентный столбец, то матрица $\|Q\|$ нулевая.**

Доказательство.

Пусть $\|Q\| = \left\| \begin{array}{cccc} \omega_{11} & \omega_{12} & \cdots & \omega_{1n} \\ \omega_{21} & \omega_{22} & \cdots & \omega_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \omega_{n1} & \omega_{n2} & \cdots & \omega_{nn} \end{array} \right\|$. Выберем в качестве столбца $\|x\|$ столбец вида $\|x\| = \| \ 0 \ \ldots \ 1 \ \ldots \ 0 \ \|^{\mathrm{T}}$, где единица стоит в строке с номером k.

Тогда получим $\|Q\| \left\| \begin{array}{c} 0 \\ \cdots \\ 1 \\ \cdots \\ 0 \end{array} \right\| = \left\| \begin{array}{c} \omega_{1k} \\ \cdots \\ \omega_{kk} \\ \cdots \\ \omega_{nk} \end{array} \right\| = \left\| \begin{array}{c} 0 \\ \cdots \\ 0 \\ \cdots \\ 0 \end{array} \right\|$, то есть k-й столбец матрицы $\|Q\|$ нулевой. В силу произвольности k приходим к заключению о справедливости утверждения леммы.

Лемма доказана.

Теорема 5.1.2 **Для невырожденных одинакового размера квадратных матриц $\|A\|$ и $\|B\|$ справедливо равенство** $\left(\|A\| \, \|B\|\right)^{-1} = \|B\|^{-1} \, \|A\|^{-1}$.

Доказательство.

1. Пусть произведение матрицы $\left(\|A\| \, \|B\|\right)^{-1}$ на некоторый n-компонентный столбец $\|x\|$ есть столбец $\|c\|$, то есть $\left(\|A\| \, \|B\|\right)^{-1}\|x\| = \|c\|$ или $\|x\| = \|A\| \, \|B\| \, \|c\|$.
2. Из последнего равенства получаем, что
$$\|A\|^{-1}\|x\| = \|B\| \, \|c\| \qquad \Longrightarrow \qquad \|B\|^{-1}\|A\|^{-1}\|x\| = \|c\| .$$
3. Вычитая почленно равенства $\left(\|A\| \, \|B\|\right)^{-1}\|x\| = \|c\|$ и $\|B\|^{-1}\|A\|^{-1}\|x\| = \|c\|$, получаем в силу дистрибутивности матричного произведения
$$\Big(\left(\|A\| \, \|B\|\right)^{-1} - \|B\|^{-1}\|A\|^{-1}\Big)\|x\| = \|o\| .$$

Откуда на основании леммы 5.1.2 ввиду произвольности столбца $\|x\|$ заключаем, что матрица

$$\bigl(\|A\|\,\|B\|\bigr)^{-1} - \|B\|^{-1}\|A\|^{-1}$$

нулевая.

Теорема доказана.

Задача 5.1.1 *Проверить тождество* $\bigl(\,\|A\|^{-1}\bigr)^{\mathrm{T}} = \bigl(\,\|A\|^{\mathrm{T}}\bigr)^{-1}$.

Определение 5.1.4 Невырожденная квадратная матрица $\|Q\|$, для которой $\|Q\|^{-1} = \|Q\|^{\mathrm{T}}$, называется *ортогональной.*

Свойства ортогональных матриц, играющих важную роль во многих приложениях, можно сформулировать в виде следующих теорем.

Теорема 5.1.3 **Для ортогональной матрицы** $\|Q\|$ **справедливо равенство** $\det\|Q\| = \pm 1$.

Доказательство.

Умножая равенство $\|Q\|^{-1} = \|Q\|^{\mathrm{T}}$ последовательно слева и справа на $\|Q\|$, в силу определения 5.1.2, приходим к соотношению

$$\|Q\|\,\|Q\|)^{\mathrm{T}} = \|Q\|^{\mathrm{T}}\|Q\| = \|E\|\,.$$

Откуда находим, что $\det^2\|Q\| = 1$, поскольку

— определитель произведения квадратных матриц одинакового размера равен произведению определителей сомножителей;

— определитель матрицы не меняется при ее транспонировании;

— $\det\|E\| = 1$.

Теорема доказана.

Теорема 5.1.4 **Каждая ортогональная матрица второго порядка $\|Q\|$, для которой $\det\|Q\| = 1$, может быть представлена в виде $\begin{Vmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{Vmatrix}$, где φ — некоторое число, а каждая ортогональная матрица с $\det\|Q\| = -1$ — в виде $\begin{Vmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{Vmatrix}$.**

Доказательство.

Пусть матрица $\|Q\| = \begin{Vmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{Vmatrix}$ ортогональная, тогда должны быть справедливы равенства

$$\|Q\|\|Q\|^{\mathrm{T}} = \begin{Vmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{Vmatrix} \begin{Vmatrix} \omega_{11} & \omega_{21} \\ \omega_{12} & \omega_{22} \end{Vmatrix} = \|E\|$$

и, следовательно,

$$\begin{Vmatrix} \omega_{11}^2 + \omega_{12}^2 & \omega_{11}\omega_{21} + \omega_{12}\omega_{22} \\ \omega_{11}\omega_{21} + \omega_{12}\omega_{22} & \omega_{21}^2 + \omega_{22}^2 \end{Vmatrix} = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}.$$

Последнее матричное равенство может быть записано в виде системы скалярных условий

$$\begin{cases} \omega_{11}^2 + \omega_{12}^2 & = 1\,, \\ \omega_{11}\omega_{21} + \omega_{12}\omega_{22} & = 0\,, \\ \omega_{21}^2 + \omega_{22}^2 & = 1\,, \end{cases}$$

причем из этих равенств, как было показано при доказательстве теоремы 5.1.3, следует, что $\det\|Q\| = \pm 1$. Рассмотрим вначале случай $\det\|Q\| = 1$.

Если из суммы первого и третьего уравнений системы вычесть удвоенное равенство $\det\|Q\| = 1$, то есть равенство $\omega_{11}\omega_{22} - \omega_{12}\omega_{21} = 1$, то мы получим

$$\left(\omega_{11}^2 + \omega_{12}^2\right) + \left(\omega_{21}^2 + \omega_{22}^2\right) - 2\left(\omega_{11}\omega_{22} - \omega_{12}\omega_{21}\right) = 0$$

или $\quad \left(\omega_{11}-\omega_{22}\right)^2+\left(\omega_{12}+\omega_{21}\right)^2=0, \quad$ откуда следует, что

$$\left\{\begin{array}{rcr} \omega_{11} & = & \omega_{22}\,, \\ \omega_{12} & = & -\omega_{21}\,. \end{array}\right.$$

Наконец, из условий

$$\omega_{11}^2+\omega_{12}^2=1 \qquad \text{и} \qquad \omega_{21}^2+\omega_{22}^2=1$$

получаем оценки:

$$0 \le \omega_{11}^2 \le 1 \qquad \text{и} \qquad 0 \le \omega_{21}^2 \le 1\,,$$

из которых следует существование числа φ, такого, что

$$\left\{\begin{array}{rcr} \omega_{11} & = & \cos\varphi\,, \\ \omega_{21} & = & -\sin\varphi\,, \end{array}\right.$$

приводящие к требуемому виду матрицы $\|Q\|$, поскольку из полученных соотношений также, очевидно, следует, что $\omega_{11}^2+\omega_{21}^2=1$.

Случай $\det\|Q\|=-1$ рассматривается аналогично.

Теорема доказана.

Следствие 5.1.1 **Матрица перехода от одного ортонормированного базиса на плоскости к другому ортогональная.**

Доказательство.

В § 1.8 было показано, что $\|S\|$ — матрица перехода от одной ортонормированной системы координат на плоскости к другой может иметь один из двух следующих видов:

$$\left\|\begin{array}{rr} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{array}\right\| \qquad \text{или} \qquad \left\|\begin{array}{rr} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{array}\right\|,$$

где φ — угол между первыми базисными векторами. Но тогда матрица перехода ортогональная в силу теоремы 5.1.4.

Следствие доказано.

5.2. Операторы и функционалы. Отображения и преобразования плоскости

Вводимое в курсе математического анализа понятие *функции* (как правила, устанавливающего однозначное соответствие между числом, принадлежащим области определения, и числом, принадлежащим множеству значений) может быть естественным образом обобщено на случай, когда область определения и область значений не являются числовыми множествами.

Определение 5.2.1	Будем говорить, что задан *оператор* $\hat{A}$, действующий на множестве Ω со значениями в множестве Θ, если указано правило, по которому каждому элементу множества Ω поставлен в соответствие единственный элемент из множества Ω.

Символически результат действия оператора $\hat{A}$ обозначается так: $y = \hat{A}x$, $x \in \Omega$, $y \in \Theta$. Элемент y в этом случае называется *образом* элемента x, а элемент x — *прообразом* элемента y.

Определение 5.2.2	Если Θ — область значений некоторого оператора — является числовым множеством, то говорят, что на множестве Ω задан *функционал*.

Функционалы обычно обозначаются так же, как и функции: например, $y = \Phi(x) \quad x \in \Omega$.

Пример 5.2.1

1. Если каждому вектору $\vec{x}$ в пространстве поставлен в соответствие вектор $\vec{y}$, являющийся ортогональной проекцией вектора $\vec{x}$ на некоторую ось l, то говорят, что в пространстве задан оператор $\vec{y} = \hat{\mathrm{Pr}}_l\,\vec{x}$ ортогонального проектирования векторов на ось l. В этом случае можно записать, что $\vec{y} = \hat{A}\vec{x}$, где символически $\hat{A} = \hat{\mathrm{Pr}}_l$.

2. Каждой дифференцируемой на $[\alpha, \beta]$ функции $f(\tau)$ можно поставить в однозначное соответствие $f'(\tau)$ — ее производную функцию, поэтому можно говорить об *операторе дифференцирования*, действующего по формуле $f'(x) = \hat{A}\, f(x)$ и обозначаемом символически как $\hat{A} = \dfrac{d}{d\tau}$.
3. Каждому вектору $\vec{x}$ в пространстве можно поставить в однозначное соответствие его длину — число $\left|\vec{x}\right|$. Согласно определению 5.2.2 данная зависимость является *функционалом*, определенным на множестве всех векторов.
4. На множестве квадратных матриц третьего порядка сопоставление матрицы и ее детерминанта является функционалом.
5. Если каждой непрерывной на $[\alpha, \beta]$ функции $f(x)$ поставить в соответствие ее определенный интеграл $\Phi = \int\limits_{\alpha}^{\beta} f(\tau)\, d\tau$, то можно говорить о функционале вида $\Phi(f)$, заданном на множестве функций, непрерывных на $[\alpha, \beta]$.

Определение 5.2.3	Оператором $\hat{A}$, отображающим плоскость (или, просто, *отображением* плоскости) P на плоскость Q, называется правило, по которому каждой точке плоскости P поставлена в соответствие единственная точка плоскости Q.

Отображение плоскости принято обозначать следующим образом: $\hat{A}:\ P \longrightarrow Q$.

Если точка M плоскости P отображается в точку M^* плоскости Q, то это представляется как $M^* = \hat{A}\, M$ (что также иногда записывают в виде $M^* = \hat{A}(M)$), при этом точка M^* является образом точки M, а точка M — прообразом точки M^*.

Определение 5.2.4	Отображение $\hat{A}:\ P \longrightarrow Q$ называется *взаимно однозначным*, если каждая точка плоскости Q имеет прообраз и притом единственный.

Определение 5.2.5 Отображение $\hat{A}$ плоскости P в саму себя называется *преобразованием* плоскости P.

Определение 5.2.6 Последовательное выполнение преобразований $M^* = \hat{A}\,M$ и $M^{**} = \hat{B}\,M^*$ называется *произведением* (или *композицией*) этих преобразований.

Произведение операторов записывается в виде $M^{**} = \hat{B}\,\hat{A}\,M$. Заметим, что в общем случае эта операция *не коммутативна*, но *ассоциативна*.

Определение 5.2.7 Преобразованием, *обратным* взаимно однозначному преобразованию $\hat{A}: P \longrightarrow P$, называется оператор $\hat{A}^{-1}: P \longrightarrow P$, такой, что для каждой точки M плоскости P имеет место

$$\hat{A}^{-1}\hat{A}\,M = \hat{A}\,\hat{A}^{-1}\,M = M\,.$$

Определение 5.2.8 Точка плоскости P, переводимая преобразованием $\hat{A}$ сама в себя, называется *неподвижной* точкой для $\hat{A}$.

Множество на P, состоящее из неподвижных точек для $\hat{A}$, называется *неподвижным* для $\hat{A}$.

Множество точек на P, переходящее при действии $\hat{A}$ само в себя, называется *инвариантным* множеством этого преобразования.

5.3. Линейные преобразования плоскости

Векторно-координатный метод представления геометрических объектов оказывается удобным как для описания, так и для исследования преобразований плоскости.

Пусть на плоскости с декартовой системой координат $\{O\,,\ \vec{g}_1\,,\ \vec{g}_2\}$ каждой ее точке M поставлена в однозначное соответствие точка M^*, то есть согласно определению 5.2.6 задано преобразование этой плоскости $M^* = \hat{A}\,M$. И пусть координатные представления (координатные столбцы) радиусов-векторов этих точек суть соответственно

$$\left\|\vec{r}_M\right\|_g = \left\| \begin{matrix} x \\ y \end{matrix} \right\| \qquad \text{и} \qquad \left\|\vec{r}_{M^*}\right\|_g = \left\| \begin{matrix} x^* \\ y^* \end{matrix} \right\| ,$$

тогда координаты x^* и будут x^* некоторыми *функциями* от x и y : $x^* = F_x(x,y)$ и $y^* = F_y(x,y)$.

Поэтому систему уравнений или равносильное ей матричное равенство

$$\left\{ \begin{array}{l} x^* = F_x(x,y)\,, \\ y^* = F_y(x,y) \end{array} \right. \qquad \Longleftrightarrow \qquad \left\| \begin{matrix} x^* \\ y^* \end{matrix} \right\| = \left\| \begin{matrix} F_x(x,y) \\ F_y(x,y) \end{matrix} \right\|$$

можно рассматривать как описание (или способ задания) оператора $\vec{r}_{M^*} = \hat{A}\vec{r}_M$ в системе координат $\left\{O\,,\ \vec{g}_1\,,\ \vec{g}_2\right\}$.

Далее мы будем рассматривать частные, но важные для приложений, виды функций $F_x(x,y)$ и $F_y(x,y)$.

Определение 5.3.1	Оператор $\vec{r}_{M^*} = \hat{A}\vec{r}_M$ называется *линейным оператором*, если в каждой декартовой системе координат $\left\{O\,,\ \vec{g}_1\,,\ \vec{g}_2\right\}$ он задается формулами $$\left\{ \begin{array}{l} x^* = \alpha_{11}x + \alpha_{12}y + \beta_1\,, \\ y^* = \alpha_{21}x + \alpha_{22}y + \beta_2\,. \end{array} \right.$$

При помощи матричных операций линейный оператор может быть записан в виде

$$\left\| \begin{matrix} x^* \\ y^* \end{matrix} \right\| = \left\|\hat{A}\right\|_g \left\| \begin{matrix} x \\ y \end{matrix} \right\| + \left\| \begin{matrix} \beta_1 \\ \beta_2 \end{matrix} \right\| ,$$

где матрица $\left\|\hat{A}\right\|_g = \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\|$ называется *матрицей линейного оператора* (*координатным представлением* $\hat{A}$) в системе координат $\left\{O\,,\ \vec{g}_1\,,\ \vec{g}_2\right\}$.

Определение 5.3.2

Оператор $\vec{r}_{M^*} = \hat{A}\vec{r}_M$ называется *линейным однородным оператором*, если он удовлетворяет определению 5.3.1 и, кроме того, $\beta_1 = \beta_2 = 0$.

Если же $|\beta_1|+|\beta_2| > 0$, то оператор $\hat{A}$ называется *неоднородным*.

Пример 5.3.1

К линейным однородным операторам относятся:

1. Оператор $\hat{A}$, действие которого сводится к умножению координат радиуса-вектора прообраза на фиксированные положительные числа, называемый «оператором сжатия к осям», или просто «сжатием к осям», имеющий матрицу

$$\left\|\hat{A}\right\|_g = \left\| \begin{matrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{matrix} \right\|,$$

где числа κ_1 и κ_2 — коэффициенты сжатия.

2. Оператор ортогонального проектирования радиусов-векторов точек плоскости на некоторую заданную ось, проходящую через начало координат.

3. Гомотетия с коэффициентом κ и с центром в начале координат.

Теорема 5.3.1

Для линейного однородного оператора $\hat{A}$ справедливы соотношения:

$$\textbf{1)}\quad \hat{A}(\vec{r}_1 + \vec{r}_2) = \hat{A}\vec{r}_1 + \hat{A}\vec{r}_2 \quad \forall \vec{r}_1, \vec{r}_2 \ ,$$

$$\textbf{2)}\quad \hat{A}(\lambda\vec{r}) = \lambda\hat{A}\vec{r} \quad \forall \lambda, \vec{r} \ .$$

Доказательство.

В справедливости утверждения теоремы убедимся непосредственной проверкой, используя правила действия с матрицами. Например, для 1) имеем

$$\hat{A}(\vec{r}_1+\vec{r}_2) = \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\| \left(\left\| \begin{matrix} x_1 \\ y_1 \end{matrix} \right\| + \left\| \begin{matrix} x_2 \\ y_2 \end{matrix} \right\| \right) =$$

$$= \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\| \left\| \begin{matrix} x_1 \\ y_1 \end{matrix} \right\| + \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\| \left\| \begin{matrix} x_2 \\ y_2 \end{matrix} \right\| = \hat{A}\vec{r}_1 + \hat{A}\vec{r}_2 \,.$$

Теорема доказана.

Теорема 5.3.2 **Если для некоторого оператора $\hat{A}$ справедливы соотношения**

$$\textbf{1)}\quad \hat{A}(\vec{r}_1+\vec{r}_2) = \hat{A}\vec{r}_1 + \hat{A}\vec{r}_2 \quad \forall \vec{r}_1, \vec{r}_2 \ ,$$

$$\textbf{2)}\quad \hat{A}(\lambda\vec{r}) = \lambda\hat{A}\vec{r} \quad \forall \lambda, \vec{r} \ ,$$

то этот оператор линейный и однородный.

Доказательство.

Пусть $\vec{r} = x\,\vec{g}_1 + y\,\vec{g}_2$ и $\hat{A}\vec{r} = x^*\,\vec{g}_1 + y^*\,\vec{g}_2$ суть соответственно координатные разложения для прообраза и образа, тогда

$$x^*\,\vec{g}_1 + y^*\,\vec{g}_2 = \hat{A}\big(x\,\vec{g}_1 + y\,\vec{g}_2\big) = x\,\hat{A}\vec{g}_1 + y\,\hat{A}\vec{g}_2 \,.$$

По теореме 1.5.1 существуют числа $\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}$, такие, что

$$\hat{A}\vec{g}_1 = \alpha_{11}\,\vec{g}_1 + \alpha_{21}\,\vec{g}_2 \qquad \text{и} \qquad \hat{A}\vec{g}_2 = \alpha_{12}\,\vec{g}_1 + \alpha_{22}\,\vec{g}_2 \,.$$

Тогда получаем

$$x^*\,\vec{g}_1 + y^*\,\vec{g}_2 = x\,\hat{A}\vec{g}_1 + y\,\hat{A}\vec{g}_2 =$$

$$= \big(\alpha_{11}x + \alpha_{12}y\big)\,\vec{g}_1 + \big(\alpha_{21}x + \alpha_{22}y\big)\,\vec{g}_2 \,.$$

И в силу линейной независимости векторов $\vec{g}_1$ и $\vec{g}_2$ имеем

$$\left\{ \begin{array}{l} x^* = \alpha_{11}x + \alpha_{12}y \,, \\ y^* = \alpha_{21}x + \alpha_{22}y \end{array} \right. \qquad \text{или} \qquad \left\| \begin{matrix} x^* \\ y^* \end{matrix} \right\| = \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\| \left\| \begin{matrix} x \\ y \end{matrix} \right\| .$$

Теорема доказана.

Для вектора $\vec{a}$, имеющего в базисе $\{\vec{g}_1\,,\ \vec{g}_2\}$ координатное представление $\|\,\vec{a}\,\|_g = \left\| \begin{matrix} a_x \\ a_y \end{matrix} \right\|$, при любом линейном преобразовании образ $\hat{A}\,\vec{a}$ есть вектор с координатным представлением

$$\left\|\hat{A}\,\vec{a}\right\|_g = \left\| \begin{matrix} a_x^* \\ a_y^* \end{matrix} \right\| = \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\| \left\| \begin{matrix} a_x \\ a_y \end{matrix} \right\| .$$

Отметим, что в эти формулы не входят β_1 и β_2, поскольку вектор $\vec{a}$ равен разности радиусов-векторов его конца и начала.

Из теорем 5.3.1 и 5.3.2 вытекают важные следствия.

`Следствие 5.3.1` **Столбцами матрицы линейного оператора $\hat{A}$ в базисе $\{\vec{g}_1\,,\ \vec{g}_2\}$ являются координатные представления векторов $\hat{A}\,\vec{g}_1$ и $\hat{A}\,\vec{g}_2$.**

`Следствие 5.3.2` **Каждому линейному однородному оператору преобразования плоскости в конкретном базисе соответствует однозначно определяемая квадратная матрица второго порядка, а каждая квадратная матрица второго порядка задает в этом базисе некоторый линейный однородный оператор.**

`Задача 5.3.1` *Показать, что для линейных однородных операторов на плоскости справедливы утверждения:*

1. *Матрица произведения линейных однородных операторов равна произведению матриц операторов-сомножителей:* $\left\|\hat{A}\,\hat{B}\right\|_g = \left\|\hat{A}\right\|_g \left\|\hat{B}\right\|_g$.
2. *Если оператор, обратный линейному однородному оператору $\hat{A}$, существует, то* $\left\|\hat{A}^{-1}\right\|_g = \left\|\hat{A}\right\|_g^{-1}$.

Выясним теперь, как изменится матрица линейного однородного оператора при замене базиса. Имеет место

Теорема 5.3.3 **Пусть в исходной системе координат $\{O,\, \vec{g}_1,\, \vec{g}_2\}$ однородный линейный оператор имеет матрицу $\left\| \hat{A} \right\|_g$. Тогда в системе координат $\{O',\, \vec{g}\,'_1,\, \vec{g}\,'_2\}$ этот оператор будет иметь матрицу**

$$\left\| \hat{A} \right\|_{g'} = \| S \|^{-1} \left\| \hat{A} \right\|_g \| S \| ,$$

где $\| S \|$ – матрица перехода между системами координат $\{O,\, \vec{g}_1,\, \vec{g}_2\}$ и $\{O',\, \vec{g}\,'_1,\, \vec{g}\,'_2\}$.

Доказательство.

Пусть в исходной системе координат действие линейного оператора описывается формулой

$$\| \vec{r}^{\,*} \|_g = \left\| \hat{A} \right\|_g \| \vec{r} \|_g , \tag{5.3.1}$$

а в новой системе координат формулой

$$\| \vec{r}^{\,*} \|_{g'} = \left\| \hat{A} \right\|_{g'} \| \vec{r} \|_{g'} . \tag{5.3.2}$$

Согласно теореме 1.8.1 при переходе от системы $\{O,\, \vec{g}_1,\, \vec{g}_2\}$ к системе $\{O',\, \vec{g}\,'_1,\, \vec{g}\,'_2\}$ справедливы равенства

$$\| \vec{r} \|_g = \|S\| \| \vec{r} \|_{g'} \quad \text{и} \quad \| \vec{r}^{\,*} \|_g = \|S\| \| \vec{r}^{\,*} \|_{g'} .$$

Подставив эти соотношения в равенство (5.3.1), получим

$$\|S\| \| \vec{r}^{\,*} \|_{g'} = \left\| \hat{A} \right\|_g \|S\| \| \vec{r} \|_{g'} \qquad \Longrightarrow$$

$$\Longrightarrow \quad \| \vec{r}^{\,*} \|_{g'} = \|S\|^{-1} \left\| \hat{A} \right\|_g \|S\| \| \vec{r} \|_{g'} .$$

Почленное вычитание этого равенства из (5.3.2) дает

$$\left(\left\| \hat{A} \right\|_{g'} - \|S\|^{-1} \left\| \hat{A} \right\|_g \|S\| \right) \| \vec{r} \|_{g'} = \| \vec{o} \| .$$

Последнее равенство верно для любого вектора $\| \vec{r} \|_{g'}$. Поэтому, в силу леммы 5.1.2, выражение, стоящее в круглых скобках, есть нулевая матрица. Откуда следует

$$\left\| \hat{A} \right\|_{g'} = \| S \|^{-1} \left\| \hat{A} \right\|_{g} \| S \| .$$

Теорема доказана.

Следствие 5.3.3 **Величина** $\det \left\| \hat{A} \right\|_{g}$ **не зависит от выбора базиса.**

Доказательство.

Поскольку определитель произведения матриц равен произведению определителей сомножителей, то в силу теоремы 5.3.3 и невырожденности матрицы перехода $\| S \|$ имеем

$$\det \left\| \hat{A} \right\|_{g'} = \det \left(\| S \|^{-1} \left\| \hat{A} \right\|_{g} \| S \| \right) =$$

$$= \det \| S \|^{-1} \det \left\| \hat{A} \right\|_{g} \det \| S \| =$$

$$= \frac{1}{\det \| S \|} \det \left\| \hat{A} \right\|_{g} \det \| S \| = \det \left\| \hat{A} \right\|_{g} .$$

Следствие доказано.

Задача 5.3.2 *В ортонормированной системе координат найти формулы, задающие оператор ортогонального проектирования точек координатной плоскости на прямую*

$$x + 3y - 2 = 0 .$$

Решение. Пусть точка-прообраз M имеет радиус-вектор с координатным представлением $\| \vec{r}_0 \| = \left\| \begin{matrix} x_0 \\ y_0 \end{matrix} \right\|$, а точка M^* — образ точки M — радиус-вектор $\| \vec{r}_0^{\,*} \| = \left\| \begin{matrix} x_0^* \\ y_0^* \end{matrix} \right\|$.

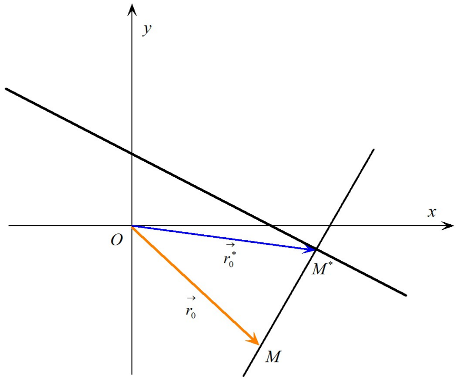


Рис. 5.2. К решению задачи 5.3.2

Из определения ортогональной проекции следует, что M^* есть точка пересечения прямой $x + 3y - 2 = 0$ и перпендикуляра к ней, проходящего через M. См. рис. 5.2.

Поскольку нормальный вектор прямой $x + 3y - 2 = 0$ является направляющим вектором этого перпендикуляра, то уравнение последнего в параметрической форме будет иметь вид

$$\left\| \begin{matrix} x \\ y \end{matrix} \right\| = \left\| \begin{matrix} x_0 \\ y_0 \end{matrix} \right\| + \tau \left\| \begin{matrix} 1 \\ 3 \end{matrix} \right\| \qquad \forall \tau \in \mathbb{R}.$$

Откуда следует, что координаты радиуса-вектора точки M^* будут удовлетворять системе уравнений

$$\begin{cases} x_0^* = x_0 + \ \tau\,, \\ y_0^* = \ y_0 + 3\tau\,, \\ x_0^* + 3y_0^* - 2 = 0\,. \end{cases}$$

Исключив из уравнений этой системы параметр τ, получим

$$\begin{cases} x_0^* \ = \ \frac{9}{10}\,x_0 - \frac{3}{10}\,y_0 + \frac{1}{5}\,, \\ y_0^* \ = \ -\frac{3}{10}\,x_0 + \frac{1}{10}\,y_0 + \frac{3}{5}\,. \end{cases} \tag{5.3.3}$$

Формулы (5.3.3) выражают координаты точки $M*$ — ортогональной проекции произвольной точки M на заданную прямую — через координаты точки M. Следовательно, эти формулы являются искомыми.

В матричном виде ответ можно представить так:

$$\left\| \begin{matrix} x \\ y \end{matrix} \right\| = \left\| \hat{A} \right\|_e \left\| \begin{matrix} x_0 \\ y_0 \end{matrix} \right\| + \left\| \begin{matrix} \beta_1 \\ \beta_2 \end{matrix} \right\|,$$

где

$$\left\| \hat{A} \right\|_e = \left\| \begin{matrix} \frac{9}{10} & -\frac{3}{10} \\ -\frac{3}{10} & \frac{1}{10} \end{matrix} \right\|, \qquad \left\| \begin{matrix} \beta_1 \\ \beta_2 \end{matrix} \right\| = \left\| \begin{matrix} \frac{1}{5} \\ \frac{3}{5} \end{matrix} \right\|.$$

В заключение отметим, что из формул (5.3.3) следует, что оператор $\hat{A}$ ортогонального проектирования точки на прямую является линейным и что он не имеет обратного (проверьте это самостоятельно).

`Решение получено.`

5.4. Аффинные преобразования и их свойства

Линейные операторы, преобразующие плоскость саму в себя (то есть линейные операторы вида $\hat{A} : P \to P$) и *имеющие обратный оператор*, играют важную с практической точки зрения роль и потому выделяются в специальный класс.

`Определение 5.4.1`

Линейный оператор $\begin{cases} x^* = \alpha_{11}x + \alpha_{12}y + \beta_1 , \\ y^* = \alpha_{21}x + \alpha_{22}y + \beta_2 , \end{cases}$ отображающий плоскость P саму на себя, с матрицей $\left\| \hat{A} \right\|_g = \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\|$, для которой в любом базисе $\det \left\| \hat{A} \right\|_g \neq 0$, называется *аффинным преобразованием плоскости.*

Теорема 5.4.1 (признак аффинности) **Если для линейного преобразования плоскости $\det\left\|\hat{A}\right\|_g \neq 0$ в некоторой декартовой системе координат, то это преобразование аффинное.**

Доказательство.

По следствию 5.3.3 определитель матрицы линейного оператора не зависит от выбора базиса, поэтому для аффинности линейного преобразования достаточно, чтобы хотя бы в одном базисе $\det\left\|\hat{A}\right\|_g \neq 0$.

Теорема доказана.

Теорема 5.4.2 **Каждое аффинное преобразование имеет единственное обратное, которое также является аффинным.**

Доказательство.

Поскольку $\det\left\|\hat{A}\right\|_g \neq 0$, то матрица $\left\|\hat{A}\right\|_g^{-1}$ существует, единственна и невырожденная (см. § 5.1), а в силу теоремы 1.1.2 (Крамера) система линейных уравнений

$$\left\|\hat{A}\right\|_g \left\| \begin{matrix} x \\ y \end{matrix} \right\| = \left\| \begin{matrix} x^* \\ y^* \end{matrix} \right\| - \left\| \begin{matrix} \beta_1 \\ \beta_2 \end{matrix} \right\|$$

всегда имеет единственное решение $\left\| \begin{matrix} x \\ y \end{matrix} \right\|$ для любого столбца $\left\| \begin{matrix} x^* \\ y^* \end{matrix} \right\|$. Но это означает, что между образами и прообразами аффинного преобразования существует взаимно однозначное соответствие, то есть для $\hat{A}$ существует единственное обратное аффинное преобразование, задаваемое формулой

$$\left\| \begin{matrix} x \\ y \end{matrix} \right\| = \left\|\hat{A}\right\|_g^{-1} \left\| \begin{matrix} x^* \\ y^* \end{matrix} \right\| - \left\|\hat{A}\right\|_g^{-1} \left\| \begin{matrix} \beta_1 \\ \beta_2 \end{matrix} \right\| .$$

Теорема доказана.

Задача 5.4.1 *Определитель матрицы, полученной при решении задачи 5.3.2, оказался равным нулю. Останется ли это равенство верным, если заменить прямую, на которую выполняется ортогональное проектирование, на некоторую другую?*

Теорема 5.4.3 **При аффинном преобразовании всякий базис переходит в базис, а для любых двух базисов существует единственное аффинное преобразование, переводящее первый базис во второй.**

Доказательство.

Пусть аффинное преобразование $\hat{A}$ задано формулами

$$\begin{cases} x^* = \alpha_{11}x + \alpha_{12}y + \beta_1\,, \\ y^* = \alpha_{21}x + \alpha_{22}y + \beta_2\,, \end{cases}$$

тогда образами первой пары базисных векторов будут векторы

$$\vec{g}_1^{\,*} = \alpha_{11}\,\vec{g}_1 + \alpha_{21}\,\vec{g}_2 \qquad \text{и} \qquad \vec{g}_2^{\,*} = \alpha_{12}\,\vec{g}_1 + \alpha_{22}\,\vec{g}_2\,.$$

А поскольку $\det\left\|\hat{A}\right\|_g = \det\left\|\begin{matrix}\alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22}\end{matrix}\right\| \neq 0$, то векторы $\vec{g}_1^{\,*}$ и g_2^* линейно независимы (теорема 1.6.2) и из них можно образовать базис.

Сопоставляя определение 1.8.2 (определение матрицы перехода) и следствие 5.3.1 (о строении матрицы линейного преобразования), замечаем, что в том случае, когда базис $\left\{\vec{g}_1^{\,*}, \vec{g}_2^{\,*}\right\}$ является образом базиса $\left\{\vec{g}_1, \vec{g}_2\right\}$ при аффинном преобразовании $\hat{A}$, матрица преобразования и матрица перехода от базиса $\left\{\vec{g}_1, \vec{g}_2\right\}$ к базису $\left\{\vec{g}_1^{\,*}, \vec{g}_2^{\,*}\right\}$ совпадают, то есть $\|S\| = \left\|\hat{A}\right\|_g$.

Но поскольку для любой пары базисов матрица перехода существует, единственна и невырождена, то и преобразование, переводящее первый базис во второй, существует, аффинное и единственное.

Теорема доказана.

Рассмотрим теперь вопрос о том, что происходит с различными геометрическими объектами на плоскости при ее аффинном преобразовании.

Теорема 5.4.4 **При аффинном преобразовании образом прямой линии является прямая.**

Доказательство.

Пусть даны прямая $\begin{cases} x = x_0 + \tau p\,, \\ y = y_0 + \tau q\,, \end{cases}$ где p и q — не равные нулю одновременно, координаты направляющего вектора прямой, и аффинное преобразование

$$\begin{cases} x^* = \alpha_{11}x + \alpha_{12}y + \beta_1\,, \\ y^* = \alpha_{21}x + \alpha_{22}y + \beta_2\,. \end{cases}$$

Тогда образом прямой будет множество точек плоскости с координатами

$$\begin{cases} x^* = \bigl(\alpha_{11}x_0 + \alpha_{12}y_0 + \beta_1\bigr) + \tau\bigl(\alpha_{11}p + \alpha_{12}q\bigr)\,, \\ y^* = \bigl(\alpha_{21}x_0 + \alpha_{22}y_0 + \beta_2\bigr) + \tau\bigl(\alpha_{21}p + \alpha_{22}q\bigr)\,. \end{cases}$$

Заметим, что если $\left|\,\alpha_{11}p + \alpha_{12}q\,\right| + \left|\,\alpha_{21}p + \alpha_{22}q\,\right| > 0$, то это множество — прямая.

Предположим противное, пусть

$$\begin{cases} \alpha_{11}p + \alpha_{12}q = 0\,, \\ \alpha_{21}p + \alpha_{22}q = 0\,, \end{cases}$$

тогда в силу аффинности преобразования, то есть условия $\det\begin{Vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{Vmatrix} \neq 0$, и по теореме 1.1.2 (Крамера) решение $p = q = 0$ будет единственным для этой системы уравнений, что противоречит условию.

Теорема доказана.

Теорема 5.4.5 **При аффинном преобразовании образом параллельных прямых являются параллельные прямые, общая точка пересекающихся прямых-прообразов переходит в точку пересечения их образов.**

Доказательство.

Предположим, что пара параллельных прямых переведена аффинным преобразованием в пересекающиеся или совпадающие прямые. Рассмотрим одну из точек, общих для образов прямых. Поскольку аффинное преобразование взаимно однозначно, то прообраз общей точки единственный и должен принадлежать одновременно каждой из прямых-прообразов. Однако таких точек нет, ибо прямые-прообразы параллельны. Следовательно, образы параллельных прямых также параллельны.

Если же прямые-прообразы пересекаются, то в силу однозначности аффинного преобразования образом их точки пересечения может быть только точка пересечения образов этих прямых.

Теорема доказана.

Проверьте самостоятельно, что при аффинном преобразовании трапеция переходит в трапецию.

Теорема 5.4.6 **При аффинном преобразовании сохраняется деление отрезка в данном отношении.**

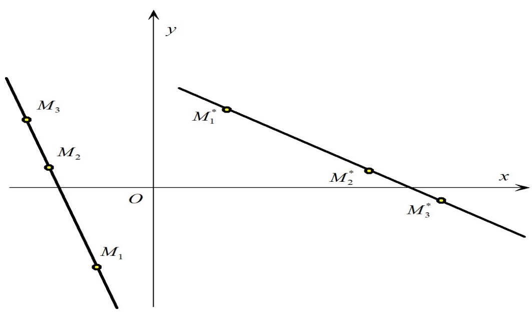


Рис. 5.3. К теореме 5.4.6

`Доказательство.`

Пусть точки M_k^*, $k = 1,2,3$ с координатами $\left\| \begin{matrix} x_k^* \\ y_k^* \end{matrix} \right\|$ являются образами попарно различных точек M_k, $k = 1,2,3$ (см. рис. 5.3) с координатами $\left\| \begin{matrix} x_k \\ y_k \end{matrix} \right\|$ соответственно. И пусть дано, что

$$\frac{x_2 - x_1}{x_3 - x_2} = \frac{y_2 - y_1}{y_3 - y_2} = \lambda, \tag{5.4.1}$$

где $\lambda \neq -1$. Если преобразование аффинное, точки-образы также попарно различны. Кроме того, мы предположили, что одноименные координаты точек попарно различны. Случай, когда это не так, рассмотрите самостоятельно.

Нам нужно показать, что

$$\frac{x_2^* - x_1^*}{x_3^* - x_2^*} = \frac{y_2^* - y_1^*}{y_3^* - y_2^*} = \lambda.$$

Если аффинное преобразование задано в виде

$$\begin{cases} x^* = \alpha_{11}x + \alpha_{12}y + \beta_1\,, \\ y^* = \alpha_{21}x + \alpha_{22}y + \beta_2\,, \end{cases}$$

то, учитывая (5.4.1), имеем

$$\frac{x_2^* - x_1^*}{x_3^* - x_2^*} = \frac{\alpha_{11}(x_2 - x_1) + \alpha_{12}(y_2 - y_1)}{\alpha_{11}(x_3 - x_2) + \alpha_{12}(y_3 - y_2)} =$$

$$= \frac{\alpha_{11}\lambda(x_3 - x_2) + \alpha_{12}\lambda(y_3 - y_2)}{\alpha_{11}(x_3 - x_2) + \alpha_{12}(y_3 - y_2)} = \lambda\,.$$

Аналогично показывается, что $\dfrac{y_2^* - y_1^*}{y_3^* - y_2^*} = \lambda\,.$

Из полученных соотношений также следует равенство отношения длин образов и отношения длин прообразов отрезков, лежащих на одной прямой.

Проверим справедливость этих утверждений для случая ортонормированной системы координат.

В силу равенств (5.4.1) имеем

$$\frac{\left|\overrightarrow{M_1^* M_2^*}\right|}{\left|\overrightarrow{M_2^* M_3^*}\right|} = \frac{\sqrt{(x_2^* - x_1^*)^2 + (y_2^* - y_1^*)^2}}{\sqrt{(x_3^* - x_2^*)^2 + (y_3^* - y_2^*)^2}} =$$

$$= \frac{|\lambda|\sqrt{(x_3^* - x_2^*)^2 + (y_3^* - y_2^*)^2}}{\sqrt{(x_3^* - x_2^*)^2 + (y_3^* - y_2^*)^2}} = |\lambda| = \frac{|\lambda|\sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2}}{\sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2}} =$$

$$= \frac{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}}{\sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2}} = \frac{\left|\overrightarrow{M_1 M_2}\right|}{\left|\overrightarrow{M_2 M_3}\right|}.$$

Теорема доказана.

Отметим также, что из теоремы 5.4.6 непосредственно вытекает, что при аффинном преобразовании отрезок прямой переходит в отрезок.

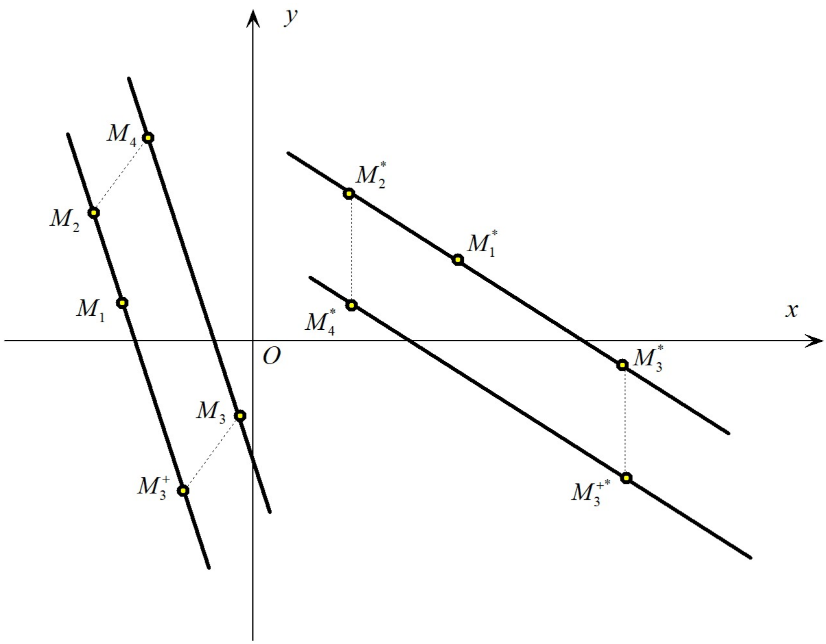


Рис. 5.4. К теореме 5.4.7

Теорема 5.4.7 **При аффинном преобразовании отношение длин образов двух отрезков, лежащих на параллельных прямых, равно отношению длин прообразов этих отрезков.**

Доказательство.

Пусть попарно несовпадающие точки M_1, M_2 и M_3, M_4 лежат на параллельных прямых, а точки M_1^*, M_2^* и M_3^*, M_4^* соответственно их образы при некотором аффинном преобразовании (см. рис. 5.4) и пусть известно, что $\dfrac{\left|\overrightarrow{M_1M_2}\right|}{\left|\overrightarrow{M_3M_4}\right|} = |\lambda|$.

Проведем прямую $M_3M_3^+$, параллельную M_2M_4. Поскольку при аффинном преобразовании образы параллельных прямых параллельны, то согласно теореме 5.4.6 $M_4M_2M_3^+M_3$ и $M_4^*M_2^*M_3^{+*}M_3^*$ — параллелограммы. Следовательно, $\left|\overrightarrow{M_2^*M_4^*}\right| = \left|\overrightarrow{M_3^*M_3^{+*}}\right|$.

Наконец, еще раз используя теорему 5.4.6, получаем

$$\frac{\left|\overrightarrow{M_1^*M_2^*}\right|}{\left|\overrightarrow{M_3^*M_4^*}\right|} = \frac{\left|\overrightarrow{M_1^*M_2^*}\right|}{\left|\overrightarrow{M_3^{+*}M_2^*}\right|} = \frac{\left|\overrightarrow{M_1M_2}\right|}{\left|\overrightarrow{M_3^+M_2}\right|} = \frac{\left|\overrightarrow{M_1M_2}\right|}{\left|\overrightarrow{M_3M_4}\right|} = |\lambda|.$$

Теорема доказана.

Теорема 5.4.8 **При аффинном преобразовании всякая декартова система координат переходит в декартову систему координат, причем координаты образа каждой точки плоскости в новой системе координат будут совпадать с координатами прообраза в исходной.**

Доказательство.

Пусть исходная система координат образована базисом $\{\vec{g}_1, \vec{g}_2\}$ и началом координат O. Согласно теореме 5.4.3 при аффинном преобразовании базис переходит в базис. Дополняя преобразованный базис образом начала координат O^*, мы получаем преобразованную систему координат $\{O^*, \vec{g}_1^*, \vec{g}_2^*\}$.

Пусть в исходной системе координаты точки-прообраза M суть x и y, а в преобразованной системе координаты точки-образа M^* суть x^* и y^*, (рис. 5.5), тогда в силу теоремы 5.4.6 будут справедливы соотношения

$$|\,x\,| = \frac{\left|O\vec{M}_1\right|}{|\,\vec{g}_1|} = \frac{\left|O^*\vec{M}_1^*\right|}{|\,\vec{g}_1^*|} = |\,x^*|\,,$$

$$|\,y\,| = \frac{\left|O\vec{M}_2\right|}{|\,\vec{g}_2|} = \frac{\left|O^*\vec{M}_2^*\right|}{|\,\vec{g}_2^*|} = |\,y^*|\,.$$

Наконец, интерпретируя с помощью определения 1.2.4 знак координаты как сонаправленость (или противонаправленность) для каждой из пар векторов

$$\left\{\vec{g}_1,\, O\vec{M}_1\right\},\ \left\{\vec{g}_2,\, O\vec{M}_2\right\},\ \left\{\vec{g}_1^*,\, O^*\vec{M}_1^*\right\} \text{ и } \left\{\vec{g}_2^*,\, O^*\vec{M}_2^*\right\},$$

приходим к заключению о справедливости утверждения теоремы.

Теорема доказана.

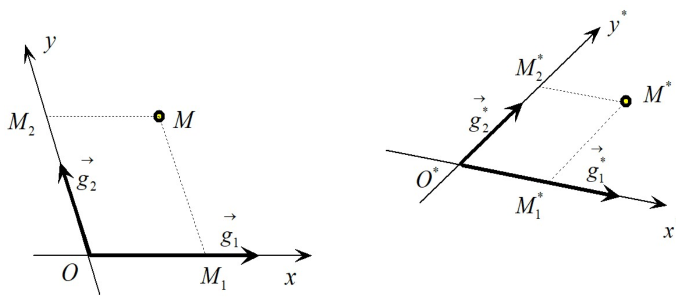


Рис. 5.5. К теореме 5.4.8

При выяснении геометрического смысла модуля и знака детерминанта матрицы аффинного преобразования оказывается удобным использование (альтернативного определению 1.8.3) определения ориентации пары неколлинеарных векторов на плоскости.

Определение 5.4.2

Пусть $\vec{n}$ есть нормальный вектор плоскости P, направленный в сторону наблюдателя. Тогда пару неколлинеарных векторов $\vec{a}$ и $\vec{b}$, лежащих в этой плоскости, назовем *правоориентированной*, если существует $\lambda > 0$ (и соответственно *левоориентированной*, если существует $\lambda < 0$) такое, что $[\vec{a}, \vec{b}] = \lambda\,\vec{n}\,.$

Тогда будет справедлива

Теорема 5.4.9

1. При аффинном преобразовании с матрицей

$$\|\hat{A}\,\|_g = \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\|$$

величины S^* — площади образа параллелограмма и S – площади прообраза параллелограмма связаны соотношением

$$S^* = \left|\,\det\|\hat{A}\,\|_g\right| \cdot S\,.$$

2. При аффинном преобразовании ориентация образов пары неколлинеарных векторов совпадает с ориентацией прообразов, если $\det\|\hat{A}\,\|_g > 0\,,$ и меняется на противоположную, если $\det\|\hat{A}\,\|_g < 0\,.$

Доказательство.

При аффинном преобразовании параллелограмм переходит в параллелограмм.

Рассмотрим некоторый базис, образованный векторами $\vec{g}_1$ и $\vec{g}_2$, образы которых при аффинном преобразовании $\hat{A}$, согласно следствию 5.3.2, имеют вид

$$\vec{g}_1^{\,*} = \hat{A}\vec{g}_1 = \alpha_{11}\vec{g}_1 + \alpha_{21}\vec{g}_2 \qquad \text{и} \qquad \vec{g}_2^{\,*} = \hat{A}\vec{g}_2 = \alpha_{12}\vec{g}_1 + \alpha_{22}\vec{g}_2\,.$$

По свойству векторного произведения (см. § 2.4) площадь параллелограмма, построенного на базисных векторах $\vec{g}_1$ и $\vec{g}_2$, равна $S = \left| [\,\vec{g}_1,\, \vec{g}_2\,] \right|$, а площадь параллелограмма, построенного на их образах, будет $S^* = \left| [\,\vec{g}_1^*,\, \vec{g}_2^*\,] \right|$.

Тогда $[\,\vec{g}_1^*,\, \vec{g}_2^*\,] = [\,\alpha_{11}\vec{g}_1 + \alpha_{21}\vec{g}_2,\ \alpha_{12}\vec{g}_1 + \alpha_{22}\vec{g}_2\,] =$

$$= (\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21})[\,\vec{g}_1,\, \vec{g}_2\,] \qquad \text{и} \qquad (5.4.2)$$

$$\left|[\,\vec{g}_1^*,\, \vec{g}_2^*\,]\right| = \left|\det\|\hat{A}\,\|_g\right| \cdot \left|[\,\vec{g}_1,\, \vec{g}_2\,]\right| \quad \Rightarrow \quad S^* = \left|\det\|\hat{A}\,\|_g\right| \cdot S.$$

Из соотношения (5.4.2) согласно определению 5.4.2 также следует, что ориентация пары векторов $\vec{g}_1^*$ и $\vec{g}_2^*$ совпадает с ориентацией пары $\vec{g}_1$ и $\vec{g}_2$ при $\det\|\hat{A}\,\|_g > 0$, и меняется на противоположную при $\det\|\hat{A}\,\|_g < 0$.

Наконец, отметим, что полученные соотношения будут выполнены для любого базиса, а значит, и для любого параллелограмма.

Теорема доказана.

Теорема 5.4.10 **Для любой линии второго порядка, указанной в формулировке теоремы 4.4.1:**

1. При аффинном преобразовании ее тип и вид не могут измениться.

2. Найдется аффинное преобразование, переводящее ее в любую другую линию второго порядка этого же типа и вида.

Доказательство.

Рассмотрим первое утверждение теоремы.

1. В силу теорем 5.4.6 и 5.4.8 параллелограмм вместе со своей внутренней частью переходит в параллелограмм, и, значит, ограниченная линия перейдет в ограниченную. Отсюда следует, что эллипсы и точки могут переходить только в эллипсы и точки. С другой стороны, точка не может переходить в эллипс и наоборот, поскольку это противоречит взаимной однозначности аффинного преобразования.

2. Среди линий второго порядка только гиперболы и параллельные прямые имеют несвязанные ветви, то есть существует прямая, не пересекающая линию второго порядка, такая, что ветви этой линии расположены по разные стороны от прямой. Сохранение данного свойства при аффинном преобразовании очевидно. Параллельные же прямые не могут перейти в ветви гиперболы в силу теоремы 5.4.5.
3. Среди непрямых линий второго порядка только парабола является неограниченной связной кривой. Следовательно, при аффинном преобразовании парабола может перейти только в параболу.
4. Если линия второго порядка есть точка, прямая или же пара параллельных или пересекающихся прямых, то из утверждений теорем 5.4.4 и 5.4.5 вытекает, что их тип не может измениться.

Рассмотрим второе утверждение теоремы.

Из теорем 4.4.1 и 5.4.3 следует, что для каждой линии второго порядка может быть построено аффинное преобразование, приводящее уравнение линии к одной из следующих девяти форм:

$$\begin{array}{llllll} x'^2+y'^2 & =\pm 1\,, & x'^2-\ y'^2 & =1\,, & x'^2\pm y'^2 & =0\,, \\ y'^2\pm 1 & =0\,, & y'^2-2x' & =0\,, & y'^2 & =0\,. \end{array} \qquad (5.4.3)$$

Пусть линии второго порядка L_1 и L_2 приводятся к одной и той же форме в списке (5.4.3) соответственно аффинными преобразованиями $\hat{A}_1$ и $\hat{A}_2$. Тогда произведение преобразований $\hat{A}_2^{-1}\hat{A}_1$ очевидно переведет линию L_1 и L_2.

Произведение аффинных преобразований также аффинное. Поэтому приходим к заключению о справедливости второго утверждения теоремы.

`Теорема доказана.`

`Замечание 5.4.1.` Изменение при аффинном преобразовании типа линии второго порядка оказывается также невозможным и для случая «пустых множеств». Справедливость этого утверждения будет показана в § 9.4 (теорема 9.4.1). Таким образом, аффинным преобразованием нельзя переместить линию второго порядка из одной клетки таблицы в формулировке теоремы 4.4.1 в другую.

Теорема 5.4.11 **Для всякого аффинного преобразования существует пара взаимно ортогональных направлений, которые переводятся данным аффинным преобразованием во взаимно ортогональные.**

Доказательство.

Пусть в ортонормированной системе координат задана пара взаимно ортогональных направлений ненулевыми векторами $\vec{p}$ и $\vec{q}$ с координатными представлениями $\left\| \vec{p} \right\|_e = \left\| \begin{matrix} \xi \\ \eta \end{matrix} \right\|$ и $\left\| \vec{q} \right\|_e = \left\| \begin{matrix} \eta \\ -\xi \end{matrix} \right\|$.

Потребуем, чтобы их образы (*неколлинеарные* в силу аффинности $\widehat{A}$)

$$\left\| \vec{p}^{\,*} \right\|_e = \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\| \left\| \begin{matrix} \xi \\ \eta \end{matrix} \right\| = \left\| \begin{matrix} \alpha_{11}\xi + \alpha_{12}\eta \\ \alpha_{21}\xi + \alpha_{22}\eta \end{matrix} \right\|$$

и

$$\left\| \vec{q}^{\,*} \right\|_e = \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\| \left\| \begin{matrix} \eta \\ -\xi \end{matrix} \right\| = \left\| \begin{matrix} \alpha_{11}\eta - \alpha_{12}\xi \\ \alpha_{21}\eta - \alpha_{22}\xi \end{matrix} \right\|$$

были также взаимно ортогональны.

Условие ортогональности векторов $\vec{p}^{\,*}$ и $\vec{q}^{\,*}$ в ортонормированном базисе $\left\{ \vec{e}_1, \vec{e}_2 \right\}$ имеет вид

$$(\alpha_{11}\xi + \alpha_{12}\eta)(\alpha_{11}\eta - \alpha_{12}\xi) + (\alpha_{21}\xi + \alpha_{22}\eta)(\alpha_{21}\eta - \alpha_{22}\xi) = 0$$

или

$$-\left(\alpha_{11}\alpha_{12} + \alpha_{21}\alpha_{22}\right)\xi^2 + \left(\alpha_{11}^2 - \alpha_{12}^2 + \alpha_{21}^2 - \alpha_{22}^2\right)\xi\eta +$$
$$+\left(\alpha_{11}\alpha_{12} + \alpha_{21}\alpha_{22}\right)\eta^2 = 0\,.$$

После очевидного переобозначения коэффициентов приходим к уравнению

$$U\xi^2 - 2V\xi\eta - U\eta^2 = 0\,.$$

Рассмотрим следующие случаи:

1. Если $U = V = 0$, то любая пара взаимно ортогональных векторов данным аффинным преобразованием переводится во взаимно ортогональную пару векторов.
2. Если $U = 0$ и $V \neq 0$, то $\xi\eta = 0$, значит, искомая пара векторов — базисная.
3. Наконец, если $U \neq 0$, то отношение координат векторов $\vec{p}$ и $\vec{q}$ находится из квадратного уравнения

$$\left(\frac{\xi}{\eta}\right)^2 - \frac{2V}{U}\left(\frac{\xi}{\eta}\right) - 1 = 0\,,$$

имеющего действительные решения

$$\left(\frac{\xi}{\eta}\right)_{1,2} = \frac{V}{U} \pm \sqrt{\frac{V^2}{U^2} + 1}$$

при любом ненулевом U.

Теорема доказана.

5.5. Ортогональные преобразования плоскости

Определение 5.5.1

Ортогональным преобразованием плоскости P называется линейный оператор вида

$$\left\| \begin{matrix} x^* \\ y^* \end{matrix} \right\| = \left\| \hat{Q} \right\|_e \left\| \begin{matrix} x \\ y \end{matrix} \right\| + \left\| \begin{matrix} \beta_1 \\ \beta_2 \end{matrix} \right\|,$$

матрица которого $\left\| \hat{Q} \right\|_e = \left\| \begin{matrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{matrix} \right\|$, *ортогональная* [9] в *любой ортонормированной* системе координат.

[9] См. определение 5.1.4.

Заметим, что ортогональное преобразование является частным случаем аффинного преобразования, поскольку в силу теоремы 5.1.3 имеет место либо $\det \left\| \hat{Q} \right\|_e = 1$, либо $\det \left\| \hat{Q} \right\|_e = -1$.

Помимо приведенных в § 5.4 аффинных свойств, ортогональные преобразования обладают своими специфическими особенностями. Рассмотрим основные из них.

Признак того, что некоторый линейный оператор является ортогональным, может быть сформулирован как

Теорема 5.5.1 **Линейный оператор на плоскости является ортогональным, если его матрица ортогональная хотя бы в одной ортонормированной системе координат.**

Доказательство.

Пусть на плоскости имеются два ортонормированных базиса $\left\{\vec{e}_1, \vec{e}_2\right\}$ и $\left\{\vec{e}\,'_1, \vec{e}\,'_2\right\}$ с матрицей перехода $\| S \|$. Согласно следствию 5.1.1, эта матрица ортогональная и для нее справедливо равенство $\| S \|^{-1} = \| S \|^{\mathrm{T}}$.

Пусть матрица оператора $\hat{Q}$ ортогональна в $\left\{\vec{e}_1, \vec{e}_2\right\}$, то есть для нее $\left\| \hat{Q} \right\|_e^{-1} = \left\| \hat{Q} \right\|_e^{\mathrm{T}}$.

Перейдем к базису $\left\{\vec{e}\,'_1, \vec{e}\,'_2\right\}$, в котором матрица линейного оператора $\hat{Q}$, согласно теореме 5.3.3 будет иметь вид

$$\left\| \hat{Q} \right\|_{e'} = \| S \|^{-1} \left\| \hat{Q} \right\|_e \| S \| .$$

Найдем в новом базисе матрицу $\left\| \hat{Q} \right\|_{e'}^{-1}$. Использовав теоремы 5.1.1 и 5.1.2, а также ортогональность матриц $\| S \|$ и $\left\| \hat{Q} \right\|_e$, получим

$$\left\| \hat{Q} \right\|_{e'}^{-1} = \left(\| S \|^{-1} \left\| \hat{Q} \right\|_e \| S \| \right)^{-1} = \| S \|^{-1} \left\| \hat{Q} \right\|_e^{-1} \left(\| S \|^{-1} \right)^{-1} =$$

$$= \| S \|^{-1} \left\| \hat{Q} \right\|_e^{-1} \| S \| = \| S \|^{\mathrm{T}} \left\| \hat{Q} \right\|_e^{\mathrm{T}} \left(\| S \|^{\mathrm{T}} \right)^{\mathrm{T}} =$$

$$= \left(\| S \|^{\mathrm{T}} \left\| \hat{Q} \right\|_e \| S \| \right)^{\mathrm{T}} = \left(\| S \|^{-1} \left\| \hat{Q} \right\|_e \| S \| \right)^{\mathrm{T}} = \left\| \hat{Q} \right\|_{e'}^{\mathrm{T}} .$$

Но равенство $\left\| \hat{Q} \right\|_{e'}^{-1} = \left\| \hat{Q} \right\|_{e'}^{\mathrm{T}}$ означает, что матрица линейного оператора $\hat{Q}$ ортогональная и в базисе $\left\{ \vec{e}\,'_1, \vec{e}\,'_2 \right\}$.

Теорема доказана.

Теорема 5.5.2 **В ортонормированной системе координат ортогональное преобразование плоскости сохраняет:**

1. **Скалярное произведение векторов.**
2. **Длины векторов и расстояния между точками плоскости.**
3. **Величины углов между прямыми.**

Доказательство.

1. Пусть дано ортогональное преобразование плоскости $\hat{Q}$ с матрицей $\left\| \hat{Q} \right\|_e$ в ортонормированной системе координат $\left\{ O, \vec{e}_1, \vec{e}_2 \right\}$.

   Тогда, как было показано в § 2.3, скалярное произведение векторов $\vec{a}$ и $\vec{b}$ с координатными представлениями

$$\| \vec{a} \|_e = \left\| \begin{matrix} \xi_1 \\ \xi_2 \end{matrix} \right\| \qquad \text{и} \qquad \left\| \vec{b} \right\|_e = \left\| \begin{matrix} \eta_1 \\ \eta_2 \end{matrix} \right\|$$

   в ОНБ выражается в следующем виде:

$$(\vec{a}, \vec{b}) = \xi_1 \eta_1 + \xi_2 \eta_2 = \| \xi_1 \; \xi_2 \| \left\| \begin{matrix} \eta_1 \\ \eta_2 \end{matrix} \right\| = \| \vec{a} \|_e^{\mathrm{T}} \left\| \vec{b} \right\|_e .$$

   Для скалярного произведения векторов $\vec{a}$ и $\vec{b}$, учитывая ортогональность матрицы $\left\| \hat{Q} \right\|_e$, получаем

$$(\hat{Q}\vec{a}, \hat{Q}\vec{b}) = \left\| \hat{Q}\vec{a} \right\|_e^{\mathrm{T}} \left\| \hat{Q}\vec{b} \right\|_e = \left(\left\| \hat{Q} \right\|_e \| \vec{a} \|_e \right)^{\mathrm{T}} \left\| \hat{Q} \right\|_e \left\| \vec{b} \right\|_e =$$

$$= \|\,\vec{a}\,\|_e^{\mathrm{T}} \left\| \hat{Q} \right\|_e^{\mathrm{T}} \left\| \hat{Q} \right\|_e \left\| \vec{b} \right\|_e = \|\,\vec{a}\,\|_e^{\mathrm{T}} \left\| \hat{Q} \right\|_e^{-1} \left\| \hat{Q} \right\|_e \left\| \vec{b} \right\|_e =$$

$$= \|\,\vec{a}\,\|_e^{\mathrm{T}} \left\| \hat{E} \right\| \left\| \vec{b} \right\|_e = \|\,\vec{a}\,\|_e^{\mathrm{T}} \left\| \vec{b} \right\|_e = (\,\vec{a},\,\vec{b}).$$

Равенство $(\hat{Q}\vec{a},\,\hat{Q}\vec{b}) = (\,\vec{a},\,\vec{b})$ означает, что при ортогональном преобразовании плоскости скалярное преобразование сохраняется в любом ортонормированном базисе.

2. Из сохранения при ортогональном преобразовании скалярного произведения для любой пары векторов следует сохранение длин векторов, поскольку

$$\left|\hat{Q}\vec{a}\right| = \sqrt{(\hat{Q}\vec{a},\hat{Q}\vec{a})} = \sqrt{(\vec{a},\vec{a})} = |\,\vec{a}\,|\,.$$

3. В силу 2 при ортогональном преобразовании равные треугольники переходят в равные (третий признак равенства треугольников), и величины углов между векторами на плоскости будут сохраняться.

Теорема доказана.

Используя свойства ортогональных преобразований, покажем, что для аффинных преобразований справедлива следующая важная теорема.

Теорема 5.5.3 **Каждое аффинное преобразование может быть представлено в виде произведения ортогонального преобразования и двух сжатий по взаимно ортогональным направлениям.**

Доказательство.

1. В силу следствия 5.3.2, а также справедливости утверждений задачи 5.3.1 и примера 5.3.1 нам достаточно убедиться, что матрица каждого аффинного преобразования в любом ортонормированном базисе $\left\{\vec{e}_1,\,\vec{e}_2\right\}$ может быть представлена в виде произведения ортогональной матрицы и диагональной матрицы с положительными значениями диагональных элементов.

2. По теореме 5.4.11 существует ортогональный (но, вообще говоря, не нормированный) базис $\left\{ \vec{\varepsilon}_1, \vec{\varepsilon}_2 \right\}$, в который данное аффинное преобразование $\hat{A}$ переведет исходный ортонормированный базис $\left\{ \vec{e}_1, \vec{e}_2 \right\}$. При этом существуют положительные нормирующие множители κ_1 и κ_2, такие, что

$$\vec{e}\,'_1 = \frac{\vec{\varepsilon}_1}{\kappa_1}\,, \qquad \vec{e}\,'_2 = \frac{\vec{\varepsilon}_2}{\kappa_2}\,, \qquad \kappa_1 = |\,\vec{\varepsilon}_1\,|\,, \qquad \kappa_2 = |\,\vec{\varepsilon}_2\,|\,,$$

то есть $\left\{ \vec{e}\,'_1, \vec{e}\,'_2 \right\}$ — ортонормированный базис.

3. С другой стороны, линейное преобразование $\hat{Q}$, переводящее ортонормированный базис $\left\{ \vec{e}_1, \vec{e}_2 \right\}$ в ортонормированный базис $\left\{ \vec{e}\,'_1, \vec{e}\,'_2 \right\}$, очевидно, ортогональное и имеет в исходном базисе ортогональную матрицу $\|\hat{Q}\|_e$. Тогда будут справедливы соотношения

$$\left\| \begin{array}{c} \vec{\varepsilon}_1 \\ \vec{\varepsilon}_2 \end{array} \right\| = \left\| \begin{array}{cc} \kappa_1 & 0 \\ 0 & \kappa_2 \end{array} \right\| \left\| \begin{array}{c} \vec{e}\,'_1 \\ \vec{e}\,'_2 \end{array} \right\|,$$

$$\left\| \begin{array}{c} \vec{e}\,'_1 \\ \vec{e}\,'_2 \end{array} \right\| = \left\|\hat{Q}\right\|_e^{\mathrm{T}} \left\| \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \end{array} \right\|, \qquad \left\| \begin{array}{c} \vec{\varepsilon}_1 \\ \vec{\varepsilon}_2 \end{array} \right\| = \left\|\hat{A}\right\|_e^{\mathrm{T}} \left\| \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \end{array} \right\|,$$

из которых следует равенство

$$\left(\left\|\hat{A}\right\|_e^{\mathrm{T}} - \left\| \begin{array}{cc} \kappa_1 & 0 \\ 0 & \kappa_2 \end{array} \right\| \left\|\hat{Q}\right\|_e^{\mathrm{T}} \right) \left\| \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \end{array} \right\| = \left\| \begin{array}{c} \vec{o} \\ \vec{o} \end{array} \right\|.$$

Тогда в силу линейной независимости базисных векторов $\left\{ \vec{e}_1, \vec{e}_2 \right\}$ мы имеем $\left\|\hat{A}\right\|_e^{\mathrm{T}} = \left\| \begin{array}{cc} \kappa_1 & 0 \\ 0 & \kappa_2 \end{array} \right\| \left\|\hat{Q}\right\|_e^{\mathrm{T}}$ или после транспонирования обеих частей этого равенства

$$\left\|\hat{A}\right\|_e = \left\|\hat{Q}\right\|_e \left\| \begin{array}{cc} \kappa_1 & 0 \\ 0 & \kappa_2 \end{array} \right\|.$$

Таким образом, аффинное преобразование представимо в виде произведения ортогонального преобразования и оператора «сжатия к осям» (см. пример 5.3.1).

`Теорема доказана.`

5.6. Понятие группы

Определение 5.6.1

Множество G называется *группой по отношению к заданной операции*, если любым двум его элементам x и y этой операцией поставлен в соответствие третий элемент этого же множества, называемый *произведением* и обозначаемый xy , и если выполняются следующие условия:

1) $\forall x, y, z \in G : \quad x(yz) = (xy)z$,
2) существует элемент e, такой, что $\forall x \in G : \quad xe = ex = x$,
3) $\forall x \in G$ существует элемент x^{-1}, такой, что $x^{-1}x = e$.

Если, кроме того, $\forall x, y \in G : \quad xy = yx$, то группа называется *коммутативной*, или *абелевой*.

Пример 5.6.1

Группами являются, например:

1) множество *вещественных чисел относительно операции сложения*, где e – число 0;
2) множество *положительных вещественных чисел относительно операции умножения*, где e – число 1;
3) множество *поворотов плоскости вокруг фиксированной точки относительно операции композиции*;
4) множество *аффинных преобразований плоскости относительно операции композиции.*

6. Системы линейных уравнений

6.1. Определители

Рассмотрим множество, состоящее из n натуральных чисел $1, 2, 3, \ldots, n$. Будем обозначать *перестановки* этих чисел (то есть последовательную их запись в некотором порядке без пропусков и повторений) как $\{k_1, k_2, k_3, \ldots, k_n\}$. Напомним, что полное число таких различных перестановок равно $n!$.

Определение 6.1.1	Будем говорить, что числа k_i и k_j образуют в перестановке *беспорядок* (*нарушение порядка*, или *инверсию*), если при $i > j$ имеет место $k_i < k_j$.

Полное число беспорядков в перестановке $\{k_1, k_2, k_3, \ldots, k_n\}$ будем обозначать $\text{Б}(k_1, k_2, k_3, \ldots, k_n)$. Например, $\text{Б}(3, 1, 4, 2) = 3$.

Пусть дана квадратная матрица

$$\|A\| = \left\| \begin{matrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \cdots & \alpha_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \alpha_{n3} & \cdots & \alpha_{nn} \end{matrix} \right\| = \|\alpha_{ij}\| \qquad i, j = \overline{1, n}\,.$$

Определение 6.1.2

Детерминантом (или *определителем*) квадратной матрицы $\|A\|$ размера $n \times n$ называется число $\det \|A\|$, получаемое по формуле

$$\det \|A\| =$$

$$= \sum_{\{k_1, k_2, \ldots, k_n\}} (-1)^{\text{Б}(k_1, k_2, \ldots, k_n)} \alpha_{1k_1} \alpha_{2k_2} \cdots \alpha_{nk_n},$$

где $\{k_1, k_2, \ldots, k_n\}$ — всевозможные различные перестановки, образованные из натуральных чисел $1, 2, 3, \ldots, n$.

Поскольку в данном определении указано, что сумма берется по всем возможным различным перестановкам, то число слагаемых равно $n!$.

Из определения 6.1.2 также вытекает, что каждое слагаемое содержит в качестве сомножителя по одному элементу матрицы из каждого столбца и каждой строки.

Задача 6.1.1

Проверить совпадение определения 6.1.2 и определения детерминантов матриц второго и третьего порядков 1.1.9 и 1.1.10.

6.2. Свойства определителей

Теорема 6.2.1

При транспонировании квадратной матрицы ее определитель не меняется.

Доказательство.

Общий вид слагаемого в формуле определителя транспонированной матрицы $\|B\| = \|A\|^{\text{T}}$ будет

$$(-1)^{\text{Б}(m_1, m_2, \ldots, m_n)} \beta_{1m_1} \beta_{2m_2} \cdots \beta_{nm_n},$$

а учитывая, что $\beta_{km_k} = \alpha_{m_k k}$, получим

$$\det \|A\|^{\mathrm{T}} =$$

$$= \sum_{\{m_1, m_2, \ldots, m_n\}} (-1)^{\text{Б}(m_1, m_2, \ldots, m_n)} \alpha_{m_1 1} \alpha_{m_2 2} \cdots \alpha_{m_n n} \,.$$

Упорядочим сомножители каждого слагаемого по номерам строк, то есть приведем их к виду

$$(-1)^{\text{Б}(m_1, m_2, \ldots, m_n)} \alpha_{1k_1} \alpha_{2k_2} \cdots \alpha_{nk_n} \,,$$

где $1, 2, \ldots, n$ — номера строк, а $\{k_1, k_2, \ldots, k_n\}$ — номера соответствующих столбцов.

Отметим, что для введенных обозначений имеет место очевидное равенство:

$$k_{m_j} = j \ \ \forall j = \overline{1, n} \,, \tag{6.2.1}$$

поскольку транспонирование транспонированной матрицы дает матрицу исходную.
Покажем, что тогда

$$\text{Б}\,(m_1, m_2, \ldots, m_n) = \text{Б}\,(k_1, k_2, k_3, \ldots, k_n) \,.$$

Действительно, пусть m_i и m_j дают беспорядок, то есть $m_i > m_j$ при $i < j$, тогда дают беспорядок и числа k_{m_i} и k_{m_j}, поскольку в силу (6.2.1) $k_{m_i} = i < j = k_{m_j}$ при $m_i > m_j$. Заметим, что верно и обратное утверждение.

Итак, при перестановке сомножителей абсолютная величина каждого слагаемого в формуле детерминанта не изменится, а знак останется тем же, поскольку не меняется число беспорядков. Окончательно получаем

$$\det \|A\|^{\mathrm{T}} =$$

$$= \sum_{\{k_1, k_2, \ldots, k_n\}} (-1)^{\text{Б}(k_1, k_2, \ldots, k_n)} \alpha_{1k_1} \alpha_{2k_2} \cdots \alpha_{nk_n} = \det \|A\| \,.$$

`Теорема доказана.`

`Замечание 6.2.1.` Утверждение теоремы 6.2.1 допускает следующую наглядную интерпретацию. Выделим в матрице элементы, входящие в некоторое слагаемое опреде-

ления 6.1.2, и соединим их отрезками прямых, как показано на рис. 6.1.
Заметим, что пара элементов α_{ik_i} и α_{jk_j} дает беспорядок, если соединяющий их отрезок имеет «положительный» наклон, то есть правый конец отрезка расположен выше левого.

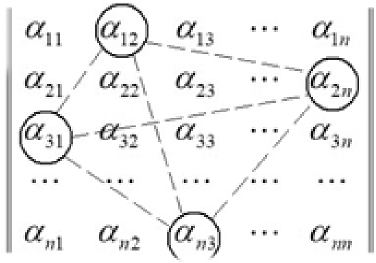


Рис. 6.1. К замечанию 6.2.1

Очевидно, что при транспонировании квадратной матрицы число отрезков с «положительным» наклоном не меняется, поэтому не меняется и знак каждого слагаемого формулы в определении 6.1.2, и, следовательно, значение определителя сохраняется.

Следствие 6.2.1 **Всякое свойство определителя матрицы, сформулированное для ее столбцов, справедливо для ее строк, и наоборот.**

Теорема 6.2.2 **При перестановке двух столбцов матрицы знак ее определителя меняется на противоположный.**

Доказательство.

Рассмотрим вначале случай, когда переставляются соседние столбцы. Поскольку общий вид слагаемых в выражении для определителя дается формулой

$$\sum_{\{k_1,k_2,\ldots,k_n\}} (-1)^{\text{Б}(k_1,k_2,\ldots,k_n)} \alpha_{1k_1}\alpha_{2k_2}\cdots\alpha_{nk_n} \,,$$

то достаточно показать, что число беспорядков изменится при перестановке соседних столбцов на единицу.

Рассмотрим перестановку чисел

$$\left\{k_1, \ldots, k_{i-1}, k_i, k_{i+1}, k_{i+2}, \ldots, k_n\right\}.$$

Если в ней поменять местами числа k_i и k_{i+1}, то число беспорядков, образуемых числами $\left\{k_1, \ldots, k_{i-1}, k_{i+2}, \ldots, k_n\right\}$, останется прежним, а за счет изменения порядка следования чисел k_i и k_{i+1} общее число беспорядков изменится на единицу. Это означает, что знак каждого слагаемого в формуле определителя изменится на противоположный и, следовательно, изменит знак и весь определитель.

Наконец, если требуется поменять местами столбцы, между которыми находится L столбцов, то для этого потребуется $L+L+1 = 2L+1$ перестановка соседних столбцов, и, поскольку $(-1)^{2L+1} = -1$, знак определителя опять-таки изменится на противоположный.

Теорема доказана.

Следствие 6.2.2 **Определитель матрицы, содержащей два одинаковых столбца, равен нулю.**

Доказательство.

При перестановке одинаковых столбцов значение определителя, с одной стороны, не меняется, но, с другой стороны, это значение должно изменить знак. Поэтому данный определитель может равняться только нулю.

Следствие доказано.

Теорема 6.2.3 (линейное свойство определителя) **Если k-й столбец матрицы задан в виде линейной комбинации некоторых «новых» столбцов, то ее определитель представим в виде той же линейной комбинации определителей матриц, k-ми столбцами которых являются соответствующие «новые» столбцы из исходной линейной комбинации.**

Доказательство.

Пусть в матрице $\|A\|_\alpha$ k-й столбец состоит из элементов $\alpha_{ik} = \lambda\beta_{ik} + \mu\gamma_{ik}$, где $i = 1, 2, \ldots, n$. Тогда справедливы равенства

$$(-1)^{\text{Б}(k_1,k_2,\ldots,k_n)}\alpha_{1k_1}\alpha_{2k_2}\ldots\alpha_{ik}\ldots\alpha_{nk_n} =$$

$$= (-1)^{\text{Б}(k_1,k_2,\ldots,k_n)}\alpha_{1k_1}\alpha_{2k_2}\ldots(\lambda\beta_{ik} + \mu\gamma_{ik})\ldots\alpha_{nk_n} =$$

$$= \lambda(-1)^{\text{Б}(k_1,k_2,\ldots,k_n)}\alpha_{1k_1}\alpha_{2k_2}\ldots(\beta_{ik})\ldots\alpha_{nk_n} +$$

$$+\mu(-1)^{\text{Б}(k_1,k_2,\ldots,k_n)}\alpha_{1k_1}\alpha_{2k_2}\cdots(\gamma_{ik})\ldots\alpha_{nk_n}\,.$$

А поскольку каждое из $n!$ слагаемых в формуле для $\det\|A\|_\alpha$ содержит точно по одному элементу из k-го столбца, то $\det\|A\|_\alpha = \lambda\det\|A\|_\beta + \mu\det\|A\|_\gamma$, где k-ые столбцы матриц $\|A\|_\beta$ и $\|A\|_\gamma$ соответственно состоят из элементов β_{ik} и γ_{ik}, $i = 1, 2, \ldots, n$.

Теорема доказана.

Следствие 6.2.3 **При вычислении определителя из столбца его матрицы можно выносить общий множитель.**

Следствие 6.2.4 **Если к некоторому столбцу матрицы прибавить линейную комбинацию остальных ее столбцов, то определитель матрицы не изменится.**

Доказательство.

Действительно, определитель, получившийся в результате данной операции с матрицей, можно (по теореме 6.2.3) представить в виде линейной комбинации исходного определителя и линейной комбинации определителей матриц, имеющих одинаковые столбцы. Последние равны нулю по следствию 6.2.2.

Следствие доказано.

Теорема 6.2.4 **Определитель произведения матриц размера $n\times n$ равен произведению их определителей, то есть**

$$\det\bigl(\|A\|\cdot\|B\|\bigr) = \det\|A\|\cdot\det\|B\|\,.$$

Доказательство.

1°. Обозначим $\|C\| = \|A\| \cdot \|B\|$. Пусть матрицы $\|A\|$, $\|B\|$ и $\|C\|$ имеют соответственно элементы α_{ij}, β_{kl} и γ_{pq}. Тогда по определению 5.1.1 $\quad \gamma_{pq} = \sum\limits_{j=1}^{n} \alpha_{pj}\beta_{jq}$, и потому

$$\det\|C\| =$$

$$= \det \left\| \begin{array}{ccc} \alpha_{11}\beta_{11} + \ldots + \alpha_{1n}\beta_{n1} & \cdots & \alpha_{11}\beta_{1n} + \ldots + \alpha_{1n}\beta_{nn} \\ \alpha_{21}\beta_{11} + \ldots + \alpha_{2n}\beta_{n1} & \cdots & \alpha_{21}\beta_{1n} + \ldots + \alpha_{2n}\beta_{nn} \\ \cdots & \cdots & \cdots \\ \alpha_{n1}\beta_{11} + \ldots + \alpha_{nn}\beta_{n1} & \cdots & \alpha_{n1}\beta_{1n} + \ldots + \alpha_{nn}\beta_{nn} \end{array} \right\| .$$

Введем в рассмотрение обобщенный тип перестановок натуральных чисел 1, 2, ..., n, в которых допускаются повторения одинаковых чисел. Такие перестановки условимся обозначать как $[1,\ 2,\ \ldots,\ n]$. По линейному свойству определителя (теорема 6.2.3):

$$\det\|C\| =$$

$$= \sum_{[l_1,\, l_2,\, \ldots,\, l_n]} \beta_{l_1 1}\beta_{l_2 2}\ldots\beta_{l_n n} \cdot \det \left\| \begin{array}{cccc} \alpha_{1l_1} & \alpha_{1l_2} & \cdots & \alpha_{1l_n} \\ \alpha_{2l_1} & \alpha_{2l_2} & \cdots & \alpha_{2l_n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{nl_1} & \alpha_{nl_2} & \cdots & \alpha_{nl_n} \end{array} \right\| =$$

$$= \sum_{[l_1,\, l_2,\, \ldots,\, l_n]} \beta_{l_1 1}\beta_{l_2 2}\ldots\beta_{l_n n} \cdot \det\|A^*\|_{[l_1,\, l_2,\, \ldots,\, l_n]} .$$

Поскольку перестановки $[\,l_1,\ l_2,\ \ldots,\ l_n]$ (в отличие от $\{\,l_1,\ l_2,\ \ldots,\ l_n\}$) могут содержать одинаковые числа, то общее число слагаемых в полученной сумме равно n^n, но не равных нулю среди них в силу следствия 6.2.2 оказывается только $n!$.

2°. Заметим, что, поскольку матрицы $\|A^*\|_{[l_1,\, l_2,\, \ldots,\, l_n]}$ составлены из тех же столбцов, что и $\|A\|$, но записанных в разном порядке, то их определители могут отличаться друг от друга в силу теоремы 6.2.2 только знаком.

Перестроим каждую из матриц $\|A^*\|_{[l_1,\, l_2,\, \ldots,\, l_n]}$, переставив ее столбцы так, чтобы каждый столбец с индексом $l_k \quad k = [1, n]$ был расположен слева от столбцов с большими индексами.

В итоге этой операции столбцы будут полностью упорядочены, для чего потребуется число перестановок столбцов, равное числу беспорядков в перестановке $\{ l_1, l_2, \ldots, l_n\}$, и, следовательно, для каждой матрицы $\|A^*\|_{[l_1, l_2, \ldots, l_n]}$ будет справедливо соотношение

$$\det \|A^*\|_{[l_1, l_2, \ldots, l_n]} = (-1)^{\text{Б}(l_1, l_2, \ldots, l_n)} \cdot \det \|A\| \ .$$

3°. Подставляя это соотношение в выражение для $\det \|C\|$, получаем

$$\det \|C\| =$$

$$= \det \|A\| \sum_{[l_1, l_2, \ldots, l_n]} (-1)^{\text{Б}(l_1, l_2, \ldots, l_n)} \beta_{l_1 1}\beta_{l_2 2} \ldots \beta_{l_n n} =$$

$$= \det \|A\| \cdot \det \|B\|^{\text{T}} ,$$

что в силу теоремы 6.2.1 означает

$$\det\bigl(\|A\| \cdot \|B\|\bigr) = \det \|A\| \cdot \det \|B\| .$$

Теорема доказана.

6.3. Разложение определителей

Выберем в квадратной матрице n-го порядка $\|A\|$ строки с номерами $i_1, i_2, \ldots, i_k$ и столбцы с номерами $j_1, j_2, \ldots, j_k$, где $1 \le k \le n$. Заметим, что выбор строк и выбор столбцов выполняется *независимо друг от друга*.

Определение 6.3.1	Детерминант квадратной подматрицы порядка k, образованной элементами, стоящими на пересечении строк $i_1, i_2, \ldots, i_k$ и столбцов $j_1, j_2, \ldots, j_k$, называется *минором* k-го порядка матрицы $\|A\|$ и обозначается $M^{j_1, j_2, \ldots, j_k}_{i_1, i_2, \ldots, i_k}$.

Определение 6.3.2	Детерминант квадратной подматрицы порядка $n-k$, образованной элементами, остающимися после удаления строк $i_1, i_2, \ldots, i_k$ и столбцов $j_1, j_2, \ldots, j_k$, называется *дополнительным минором* (к минору $M^{j_1, j_2, \ldots, j_k}_{i_1, i_2, \ldots, i_k}$) и обозначается как $\overline{M}^{j_1, j_2, \ldots, j_k}_{i_1, i_2, \ldots, i_k}$.

Выберем в матрице $\|A\|$ i-ю строку и j-й столбец, на пересечении которых расположен элемент α_{ij}. Удалим из $\|A\|$ выбранные строку и столбец, рассмотрим квадратную подматрицу $\|A^*\|$ порядка $n-1$.

Определение 6.3.3	Детерминант матрицы $\|A^*\|$ называется *дополнительным минором элемента* α_{ij} и обозначается как $\overline{M}^{j}_{i}$.

Сгруппируем в определении 6.1.2 – детерминанта матрицы $\|A\|$ — *все* слагаемые, содержащие элемент α_{ij}, и вынесем его за скобки. Получим равенство вида

$$\det \|A\| = \alpha_{ij} D_{ij} + \ldots$$

Проверьте самостоятельно, что число сгруппированных слагаемых равно $(n-1)!$.

Определение 6.3.4	Число D_{ij} называется *алгебраическим дополнением элемента* α_{ij}.

Из определений 6.1.2 и 6.3.4 очевидны равенства

$$\begin{aligned} \det \|A\| &= \sum_{j=1}^{n} \alpha_{ij} D_{ij} \qquad \forall i = [1, n], \\ \det \|A\| &= \sum_{i=1}^{n} \alpha_{ij} D_{ij} \qquad \forall j = [1, n], \end{aligned} \tag{6.3.1}$$

которые можно использовать для вычисления определителей квадратных матриц, находя значения алгебраических дополнений при помощи соотношений, которые устанавливает

Теорема 6.3.1 **Справедливы равенства**

$$D_{ij} = (-1)^{i+j}\, \overline{M}_i^j \qquad \forall i, j = [1, n]\,.$$

Доказательство.

1. По определению детерминанта 6.1.2

$$\det \|A\| =$$

$$= \alpha_{11} \sum_{\{1,k_2,\ldots,k_n\}} (-1)^{\text{Б}(1,k_2,\ldots,k_n)} \alpha_{2k_2} \cdots \alpha_{nk_n} + \ldots ,$$

то есть

$$D_{11} = \sum_{\{k_2,\ldots,k_n\}} (-1)^{\text{Б}(k_2,\ldots,k_n)} \alpha_{2k_2} \alpha_{3k_3} \cdots \alpha_{nk_n} + \ldots ,$$

поскольку очевидно, что $\text{Б}\,(1, k_2, \ldots, k_n) = \text{Б}\,(k_2, \ldots, k_n)$, но тогда выражение для D_{11} совпадает с формулой определителя матрицы порядка $n - 1$, получаемой из $\|A\|$ удалением первого столбца и первой строки. Следовательно, $D_{11} = \overline{M}_1^1$.

2. Построим новую матрицу $\|A'\|$, переместив элемент α_{ij} матрицы $\|A\|$ в ее левый верхний угол, переставив i-ю строку на первое место, для чего потребуется $i-1$ перестановка строк, и переставим на первое место j-й столбец, что потребует выполнения $j - 1$ перестановок столбцов. Тогда определитель перестроенной матрицы $\|A'\|$ равен

$$\det \|A'\| = (-1)^{i-1+j-1} \det \|A\| = (-1)^{i+j} \det \|A\|.$$

Согласно линейному свойству определителя (теорема 6.2.3) данное соотношение будет также выполняться и для каждого из его слагаемых, а значит, в силу формул (6.3.1) и для каждого алгебраического дополнения. Поэтому справедливо равенство $D_{ij} = (-1)^{i+j}\, D_{11}$.

3. Наконец, очевидно, что значение дополнительного к α_{ij} минора не зависит от его положения в матрице, т.е. $\overline{M}_i^j = \overline{M'}_1^1$. Учитывая соотношения $\overline{M}_i^j = \overline{M'}_1^1 = D'_{11} = (-1)^{i+j} D_{ij}$, приходим к доказываемому равенству $D_{ij} = (-1)^{i+j}\, \overline{M}_i^j$.

Теорема доказана.

Следствие 6.3.1 **Разложение определителя по j-му столбцу имеет вид**

$$\det\|A\| = \sum_{k=1}^{n}(-1)^{k+j}\alpha_{kj}\overline{M}^j_k,$$

или
$$\det\|A\| = \sum_{k=1}^{n}(-1)^{k+j}M^j_k\overline{M}^j_k.$$

Для практических приложений особо полезной является обобщающая теорему 1.1.1

Теорема 6.3.2 **Для любой квадратной матрицы $\|A\|$ имеет место равенство**

$$\sum_{i=1}^{n}\alpha_{ij}D_{ik} = \delta_{jk}\cdot\det\|A\|,$$

где $\delta_{jk} = \begin{cases} 1, & j=k, \\ 0, & j\neq k \end{cases}$ **— символ Кронекера (см. § 2.2).**

Доказательство.

По определению 6.3.4 алгебраического дополнения имеем $\det\|A\| = \alpha_{1j}D_{1j} + \alpha_{2j}D_{2j} + \cdots + \alpha_{nj}D_{nj}$, то есть утверждение теоремы для случая $i=j$ справедливо.

Пусть теперь $i\neq j$. Тогда выражение

$$\alpha_{1j}D_{1k} + \alpha_{2j}D_{2k} + \cdots + \alpha_{nj}D_{nk}$$

можно рассматривать как разложение по k-у столбцу определителя матрицы, у которой k-й столбец совпадает с j-м столбцом. Но такой определитель равен нулю по следствию 6.2.2.

Теорема доказана.

Следствие 6.3.2 **Если квадратная матрица $\|A\|$ невырожденна, то элементами ее обратной матрицы $\|A\|^{-1}$ являются числа $\beta_{ij} = \dfrac{1}{\Delta}(-1)^{i+j}\overline{M}^i_j$, где $\Delta = \det\|A\|$.**

Доказательство.

Найдем произведение матриц $\|A\|$ и $\|B\|$, элементы которых α_{ij} и β_{ij} $\forall i, j = [1, n]$. Пусть γ_{ij} — элемент произведения $\|A\|$ и $\|B\|$, тогда, согласно определению 5.1.1 и теореме 6.3.2, имеем

$$\gamma_{pq} = \sum_{j=1}^{n} \alpha_{pj}\beta_{jq} = \sum_{j=1}^{n} \alpha_{pj} \frac{1}{\Delta}(-1)^{j+q}\overline{M}_q^j = \frac{1}{\Delta}\sum_{j=1}^{n} \alpha_{pj} D_{qj} =$$

$$= \frac{1}{\Delta} \cdot \Delta \cdot \delta_{pq} = \delta_{pq}\,.$$

Аналогичное соотношение получается и для произведения $\|B\|\|A\|$ и по определению 1.1.4:

$$\|A\|\|B\| = \|B\|\|A\| = \|E\|,$$

но тогда, по определению 5.1.2 и лемме 5.1.1, $\|B\| = \|A\|^{-1}$.

Следствие доказано.

Проверьте самостоятельно справедливость формулы (5.1.1).

Обозначим $I = i_1 + i_2 + \ldots + i_k$ и $J = j_1 + j_2 + \ldots + j_k$ $\forall k = [1, n]$, тогда оказывается справедливой обобщающая следствие 6.3.1

Теорема 6.3.3 (Лапласа) **Для фиксированного набора столбцов** $j_1, j_2, \ldots, j_k$ **имеет место равенство**

$$\det\|A\| = \sum_{\{i_1, i_2, \ldots, i_k\}} (-1)^{I+J}\, M_{i_1, i_2, \ldots, i_k}^{j_1, j_2, \ldots, j_k}\, \overline{M}_{i_1, i_2, \ldots, i_k}^{j_1, j_2, \ldots, j_k}\,.$$

Отметим, что в этой формуле суммирование выполняется по всем возможным перестановкам номеров строк $\{i_1, i_2, \ldots, i_k\}$.

Задача 6.3.1 *Найти определитель матрицы n-го порядка*

$$\Delta_n = \det\begin{Vmatrix} x & a & a & \cdots & a \\ a & x & a & \cdots & a \\ a & a & x & \cdots & a \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a & a & a & \cdots & x \end{Vmatrix}\,.$$

`Решение.` 1°. Заметим, что в данной матрице сумма элементов каждого столбца одинакова и равна $x+a(n-1)$. Поэтому, прибавив к первой строке сумму остальных строк и вынося общий множитель из первой строки, мы получим матрицу с тем же определителем (см. следствия 6.2.4 и 6.2.3):

$$\Delta_n = \big(x+a(n-1)\big)\cdot\det\left\|\begin{array}{ccccc} 1 & 1 & 1 & \cdots & 1 \\ a & x & a & \cdots & a \\ a & a & x & \cdots & a \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a & a & a & \cdots & x \end{array}\right\|.$$

2°. Вычитая последовательно из каждой строки, начиная со второй, первую строку, умноженную на a, получим

$$\Delta_n = \big(x+a(n-1)\big)\cdot\det\left\|\begin{array}{ccccc} 1 & 1 & 1 & \cdots & 1 \\ 0 & x-a & 0 & \cdots & 0 \\ 0 & 0 & x-a & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & x-a \end{array}\right\|.$$

3°. Последовательно применив $n-1$ раз следствие 6.3.1 для разложения определителя по первому столбцу, приходим к выражению

`Решение получено.`

$$\Delta_n = \big(x+a(n-1)\big)(x-a)^{n-1}.$$

6.4. Правило Крамера

Рассмотрим неоднородную систему n линейных уравнений с n неизвестными:

$$\left\{\begin{array}{ccccccccc} \alpha_{11}\xi_1 & + & \alpha_{12}\xi_2 & + & \cdots & + & \alpha_{1n}\xi_n & = & \beta_1 \\ \alpha_{21}\xi_1 & + & \alpha_{22}\xi_2 & + & \cdots & + & \alpha_{2n}\xi_n & = & \beta_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1}\xi_1 & + & \alpha_{n2}\xi_2 & + & \cdots & + & \alpha_{nn}\xi_n & = & \beta_n, \end{array}\right. \qquad (6.4.1)$$

записываемую в неразвернутом виде $\sum\limits_{j=1}^{n} \alpha_{ij}\xi_j = \beta_i, \quad i = [1, n],$ или же в матричной форме $\|A\|\|X\| = \|B\|$, где квадратная матрица $\|A\|$ имеет компоненты α_{ij}, а столбцы $\|X\|$ и $\|B\|$ — соответственно ξ_j и β_i.

Определение 6.4.1	Упорядоченный набор чисел $\{\xi_1, \xi_2, \dots, \xi_n\}$ будем называть *частным решением* (или просто *решением*) системы линейных уравнений (6.4.1), если при подстановке этих чисел в каждое из уравнений системы мы получаем тождество.

Имеет место

Теорема 6.4.1 (правило Крамера) **Для того чтобы система линейных уравнений (6.4.1) имела единственное решение, необходимо и достаточно, чтобы $\Delta = \det \|A\| \neq 0$.**
В этом случае решение системы будет иметь вид

$$\xi_j = \frac{\Delta_j}{\Delta} \quad \forall j = [1, n],$$

где Δ_j – определитель матрицы, получаемой из матрицы $\|A\|$ заменой ее j-го столбца на столбец свободных членов:

$$\Delta_j = \det \begin{Vmatrix} \alpha_{11} & \alpha_{12} & \cdots & \beta_1 & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \beta_2 & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \beta_n & \cdots & \alpha_{nn} \end{Vmatrix}$$

$\uparrow$ – j-й столбец.

Доказательство.

1°. Проверим вначале утверждение теоремы в предположении, что система линейных уравнений (6.4.1) имеет единственное решение $\|X\| = \|\,\xi_1, \xi_2, \dots, \xi_n\|^{\mathrm{T}}$, то есть когда выполняются равенства

$$\sum_{j=1}^{n} \alpha_{ij}\xi_j = \beta_i \qquad \forall i = [1, n].$$

Умножив последовательно для всех $i = [1, n]$ обе части этих равенств на алгебраическое дополнение D_{ik} и просуммировав по i результаты умножения, получим

$$\sum_{i=1}^{n} D_{ik} \left(\sum_{j=1}^{n} \alpha_{ij} \xi_j \right) = \sum_{i=1}^{n} \beta_i D_{ik} \qquad \forall k = [1, n].$$

Изменим порядок суммирования (то есть выполним перегруппировку слагаемых) в левой части этого равенства:

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{n} \alpha_{ij} D_{ik} \right) \xi_j = \sum_{i=1}^{n} \beta_i D_{ik} \qquad \forall k = [1, n].$$

Но выражение в круглых скобках равно $\Delta \cdot \delta_{jk}$ (по теореме 6.3.2), поэтому, учитывая, что

$$\Delta \sum_{j=1}^{n} \xi_j \delta_{jk} = \Delta \cdot \xi_k \qquad \text{и} \qquad \sum_{i=1}^{n} \beta_i D_{ik} = \Delta_k \ ,$$

получаем $\quad \Delta \cdot \xi_k = \Delta_k \quad \forall k = [1, n].$

Поскольку уравнения вида $\quad \Delta \cdot \xi_k = \Delta_k \quad \forall k = [1, n]$ имеют единственное решение тогда и только тогда, когда $\Delta \neq 0$, то необходимость доказана. При этом также очевидно, что

$$\xi_k = \frac{\Delta_k}{\Delta} \quad \forall k = [1, n]. \tag{6.4.2}$$

2°. Докажем теперь, что в условиях теоремы однозначно заданный набор чисел

$$\xi_j = \frac{\Delta_j}{\Delta} \quad \forall j = [1, n]$$

есть решение данной системы линейных уравнений. Убедимся в этом, подставив значения ξ_j в левые части исходной системы линейных уравнений (6.4.1):

$$\sum_{j=1}^{n} \alpha_{ij} \frac{\Delta_j}{\Delta} = \frac{1}{\Delta} \sum_{j=1}^{n} \alpha_{ij} \left(\sum_{k=1}^{n} \beta_k D_{kj} \right) = \frac{1}{\Delta} \sum_{k=1}^{n} \beta_k \left(\sum_{j=1}^{n} \alpha_{ij} D_{kj} \right) =$$

$$= \frac{1}{\Delta} \sum_{k=1}^{n} \beta_k \delta_{ik} \Delta = \beta_i \qquad \forall i = [1, n] \,.$$

Для получения последнего равенства мы снова изменили порядок суммирования и воспользовались теоремой 6.3.2.

`Теорема доказана.`

6.5. Ранг матрицы

Согласно определению 6.1.2, детерминант является числовой характеристикой *квадратных* матриц.

Рассмотрим теперь матрицу $\|A\|$ размера $m \times n$. Пусть число k такое, что $1 \leq k \leq \min\{m, n\}$. Выберем некоторым способом в $\|A\|$ k столбцов и k строк, на пересечении которых стоят элементы, образующие квадратную матрицу минора порядка k. Пусть при данном k все миноры k-го порядка равны нулю, тогда будут равны нулю и все миноры порядка выше, чем k, поскольку каждый минор $k+1$-го порядка представим в виде линейной комбинации миноров порядка k (см. следствие 6.3.1).

`Определение 6.5.1`	Наибольший из порядков миноров матрицы $\|A\|$, отличных от нуля, называется *рангом матрицы* и обозначается $\mathrm{rg}\|A\|$.
`Определение 6.5.2`	Любой ненулевой минор матрицы, порядок которого равен ее рангу, называется *базисным минором.*
`Определение 6.5.3`	Столбцы (строки) матрицы, входящие в матрицу базисного минора, называются *базисными.*

Далее рассмотрим n m-компонентных столбцов вида

$$\|a_1\| = \left\| \begin{array}{c} \alpha_{11} \\ \alpha_{21} \\ \cdots \\ \alpha_{m1} \end{array} \right\| , \quad \|a_2\| = \left\| \begin{array}{c} \alpha_{12} \\ \alpha_{22} \\ \cdots \\ \alpha_{m2} \end{array} \right\| , \quad \cdots , \quad \|a_n\| = \left\| \begin{array}{c} \alpha_{1n} \\ \alpha_{2n} \\ \cdots \\ \alpha_{mn} \end{array} \right\|$$

и столбцы $\quad \|b\| = \left\| \begin{matrix} \beta_1 \\ \beta_2 \\ \cdots \\ \beta_m \end{matrix} \right\| , \quad \|o\| = \left\| \begin{matrix} 0 \\ 0 \\ \cdots \\ 0 \end{matrix} \right\| .$

Поскольку для столбцов (как частного случая матриц) определены операции сравнения, сложения и умножения на число, то будем говорить, что столбец $\|b\|$ есть *линейная комбинация* столбцов

$$\|a_1\|\,,\ \|a_2\|\,,\ \ldots\,,\|a_n\|\,,$$

если существуют числа $\lambda_1, \lambda_2, \ldots, \lambda_n,$ такие, что $\quad \|b\| = \sum\limits_{j=1}^{n} \lambda_j \|a_j\|\,.$

Теорема 6.5.1 (о базисном миноре) **Всякий столбец (строка) матрицы есть линейная комбинация базисных столбцов (строк) этой матрицы.**

Доказательство.

1°. Пусть ранг матрицы равен r. Без ограничения общности можно считать, что матрица базисного минора расположена в левом верхнем углу матрицы $\|A\|$.
Окаймим матрицу базисного минора фрагментами i-й строки и j-го столбца и рассмотрим определитель построенной матрицы

$$\Delta = \det \left\| \begin{array}{ccc|c} \alpha_{11} & \cdots & \alpha_{1r} & \alpha_{1j} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{r1} & \cdots & \alpha_{rr} & \alpha_{rj} \\ \hline \alpha_{i1} & \cdots & \alpha_{ir} & \alpha_{ij} \end{array} \right\| ,$$

который равен нулю как минор порядка $r+1$ в матрице, имеющей ранг r.

2°. Разложив определитель Δ по последней строке, получим

$$\alpha_{i1}K_1 + \alpha_{i2}K_2 + \cdots + \alpha_{ir}K_r + \alpha_{ij}M = 0,$$

где $M \neq 0$ — базисный минор, а $K_1, \ldots, K_r$ – некоторые константы, *одинаковые* для любых i, то есть *для всего* j-го столбца. Следовательно,

$$\alpha_{ij} = \lambda_1\alpha_{i1} + \lambda_2\alpha_{i2} + \ldots + \lambda_r\alpha_{ir}\,, \quad \text{где} \quad \forall i: \ \ \lambda_i = -K_i/M\,.$$

Теорема доказана.

Определение 6.5.4

Столбцы $\|a_1\|\,,\ \|a_2\|\,,\ \ldots\,,\|a_k\|$ будем называть *линейно зависимыми*, если существуют не равные нулю одновременно числа $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_k$, такие, что

$$\sum_{j=1}^{k}\lambda_j\|a_j\| = \|o\|\,. \qquad (6.5.1)$$

Столбцы $\|a_1\|\,,\ \|a_2\|\,,\ \ldots\,,\|a_k\|$ будем называть *линейно независимыми*, если из равенства (6.5.1) следует, что $\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0\,.$

Лемма 6.5.1

Для того чтобы столбцы (строки) матрицы были линейно зависимыми, необходимо и достаточно, чтобы один из них был линейной комбинацией остальных.

Доказательство.

Совпадает с доказательством леммы 1.4.1.

Лемма доказана.

Лемма 6.5.2

Если в наборе столбцов есть линейно зависимое подмножество, то множество всех столбцов этого набора также линейно зависимое.

Доказательство.

Допустим, что линейно зависимыми являются первые m столбцов, то есть для них существует нетривиальная линейная комбинация, равная нулевому столбцу:

$$\sum_{j=1}^{m}\lambda_j\|a_j\| = \|o\|\,, \quad \text{где} \quad \sum_{j=1}^{m}|\lambda_j| > 0\,.$$

Тогда очевидно, что будет верным равенство для также нетривиальной линейной комбинации

$$\sum_{j=1}^{m}\lambda_j\|a_j\| + \sum_{j=m+1}^{k}0\cdot\|a_j\| = \|o\|\,.$$

Лемма доказана.

Теорема 6.5.2 **Для того чтобы определитель был равен нулю, необходимо и достаточно, чтобы столбцы (строки) его матрицы были линейно зависимыми.**

Доказательство.

Доказательство необходимости.

Пусть определитель квадратной матрицы размера $n \times n$ равен нулю, то есть равный ему единственный минор порядка n нулевой, тогда ранг матрицы меньше n. По теореме о базисном миноре всякий столбец есть линейная комбинация базисных столбцов и тогда по лемме 6.5.1 столбцы матрицы линейно зависимы.

Доказательство достаточности.

Пусть столбцы матрицы линейно зависимы. По лемме 6.5.1 один из столбцов есть линейная комбинация остальных. Допустим, что этот столбец последний в матрице, то есть $\|a_n\| = \sum\limits_{j=1}^{n-1} \lambda_j \|a_j\|$.

Умножим последовательно $\forall j = [1, n-1]$ j-й столбец на число λ_j и сложим все их. Вычитание полученной суммы из n-го столбца не изменит величины определителя, но поскольку при этом мы получим нулевой столбец, то определитель равен нулю.

Теорема доказана.

Теорема 6.5.3 (о ранге матрицы) **Максимальное число линейно независимых столбцов в матрице равно максимальному числу линейно независимых строк и равно рангу этой матрицы.**

Доказательство.

1°. Если ранг матрицы нулевой, то все ее элементы нулевые и среди них нет линейно независимых.

Пусть ранг матрицы равен $r > 0$. Рассмотрим новую матрицу, составленную из r базисных столбцов исходной матрицы. Она имеет ненулевой минор r-го порядка и, следовательно, ее столбцы линейно независимы.

2°. Выберем $k > r$ столбцов исходной матрицы и покажем, что эти столбцы линейно зависимы. Построим из выбранных столбцов матрицу $\|A^*\|$. Ее ранг $r^* \leq r$, поскольку является частью матрицы $\|A\|$.
Следовательно, $r^* \leq r < k$ и в матрице $\|A^*\|$ есть, по крайней мере, один небазисный столбец, а тогда и столбцы матрицы $\|A\|$ линейно зависимы по лемме 6.5.2.

Теорема доказана.

6.6. Системы m линейных уравнений с n неизвестными

Рассмотрим систему m линейных уравнений с n неизвестными:

$$\left\{ \begin{array}{ccccccccc} \alpha_{11}\xi_1 & + & \alpha_{12}\xi_2 & + & \cdots & + & \alpha_{1n}\xi_n & = & \beta_1 \\ \alpha_{21}\xi_1 & + & \alpha_{22}\xi_2 & + & \cdots & + & \alpha_{2n}\xi_n & = & \beta_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{m1}\xi_1 & + & \alpha_{m2}\xi_2 & + & \cdots & + & \alpha_{mn}\xi_n & = & \beta_m . \end{array} \right. \qquad (6.6.1)$$

Она записывается в неразвернутом виде $\sum\limits_{j=1}^{n} \alpha_{ij}\xi_j = \beta_i \quad i = [1, m]$ или же в матричной форме $\|A\|\|X\| = \|B\|$, где матрица $\|A\|$ размера $m \times n$ имеет компоненты α_{ij}, а столбцы $\|X\|$ и $\|B\|$ — соответственно компоненты $\xi_j \quad \forall j = [1, n]$ и $\beta_i \quad \forall i = [1, m]$.

Определение 6.6.1

Упорядоченный набор чисел $\{\xi_1, \xi_2, \ldots, \xi_n\}$ будем называть *частным решением* (или просто *решением*) системы линейных уравнений (6.6.1), если при подстановке этих чисел в каждое из уравнений системы мы получаем тождество.
Частное решение системы линейных уравнений (6.6.1) также может быть записано в виде столбца

$$\|X^0\| = \|\ \xi_1^0\ \xi_2^0\ \ldots\ \xi_n^0\ \|^{\mathrm{T}}.$$

Совокупность всех частных решений системы линейных уравнений (6.6.1) назовем *общим решением* системы (6.6.1).

Определение 6.6.2	Если система (6.6.1) имеет хотя бы одно частное решение, то она называется *совместной*, в противном случае – *несовместной* системой уравнений.
Определение 6.6.3	Матрица $\|A\| = \begin{Vmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{Vmatrix}$ называется *основной матрицей* системы (6.6.1), а матрица $$\|A\|B\| = \left\Vert \begin{array}{cccc\|c} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} & \beta_1 \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} & \beta_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} & \beta_m \end{array} \right\Vert$$ называется *расширенной матрицей* этой системы.
Определение 6.6.4	Система (6.6.1) называется *однородной*, если $$\beta_i = 0 \qquad \forall i = [1, m],$$ в противном случае — *неоднородной* системой уравнений.

Теорема 6.6.1 (Кронекера - Капелли) **Для того чтобы система (6.6.1) была совместна, необходимо и достаточно, чтобы ранг ее основной матрицы был равен рангу расширенной.**

Доказательство.

Доказательство необходимости.

Пусть существует решение системы (6.6.1) $\{\xi_1, \xi_2, \ldots, \xi_n\}$, тогда эту систему можно представить в виде следующего матричного равенства:

$$\xi_1\|a_1\| + \xi_2\|a_2\| + \ldots + \xi_n\|a_n\| = \|B\|,$$

где $\qquad ||a_j\| = \|\alpha_{1j}\, \alpha_{2j} \ldots \alpha_{mj}\,\|^{\mathrm{T}} \qquad \forall j = [1, n].$

Поскольку в этом случае столбец свободных членов есть линейная комбинация столбцов, образующих основную матрицу, то число линейно независимых столбцов основной и расширенной матриц будет одинаковым. Следовательно, по теореме 6.5.3 (о ранге матрицы) $\mathrm{rg}\|A\| = \mathrm{rg}\|A|B\|$.

Доказательство достаточности.

Пусть ранг основной матрицы равен рангу расширенной матрицы и равен r. Без ограничения общности предположим, что базисный минор расположен в левом верхнем углу расширенной матрицы, но тогда по теореме 6.5.1 (о базисном миноре) имеет место равенство

$$\|B\| = \sum_{j=1}^{r} \lambda_j \|a_j\|\,,$$

которое можно переписать в виде

$$\|B\| = \sum_{j=1}^{r} \lambda_j \|a_j\| + \sum_{j=r+1}^{n} 0 \cdot \|a_j\|\,.$$

Однако последнее означает, что система (6.6.1) имеет решение $\{\lambda_1, \lambda_2, \ldots, \lambda_r, \underbrace{0, 0, \ldots, 0}_{n-r}\}$, то есть она совместна.

Теорема доказана.

Задача 6.6.1 *Справедливо ли следующее утверждение?*

Для того чтобы прямые

$$A_k x + B_k y + C_k = 0, \quad |A_k| + |B_k| > 0 \qquad \forall k = [1, m]$$

пересекались в одной и той же точке плоскости, необходимо и достаточно, чтобы

$$\mathrm{rg}\left\|\begin{array}{cc} A_1 & B_1 \\ A_2 & B_2 \\ \ldots & . \\ A_m & B_m \end{array}\right\| = \mathrm{rg}\left\|\begin{array}{ccc} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ \ldots & \ldots & \ldots \\ A_m & B_m & C_m \end{array}\right\|.$$

Ответ обоснуйте.

6.7. Фундаментальная система решений

В § 6.6 было показано, что факт совместности или несовместности системы (6.6.1) можно установить, сравнив ранги ее основной и расширенной матриц. Рассмотрим теперь случай, когда система (6.6.1) совместна и найдем все ее решения, то есть ее общее решение.

При выводе формулы общего решения системы (6.6.1) окажутся полезными следующие утверждения.

Лемма 6.7.1 **Любая линейная комбинация частных решений** *однородной* **системы (6.6.1) также является ее частным решением.**

Доказательство.

Пусть $\|x^p\| = \|\,\xi_1^p\,\xi_2^p\,\ldots\,\xi_n^p\,\|^{\mathrm{T}}\ \ \forall p = [1,k]$ — частные решения однородной системы, т.е. $\|A\|\|x^p\| = \|o\|\ \ \forall p = [1,k]$.

Рассмотрим столбец $\|y\| = \sum\limits_{p=1}^{k} \lambda_p \|x^p\|$. По правилам действий с матрицами для него справедливы равенства

$$\|A\|\|y\| = \|A\| \left(\sum_{p=1}^{k} \lambda_p \|x^p\| \right) = \sum_{p=1}^{k} \lambda_p \Big(\|A\|\|x^p\| \Big) = \|o\|\,.$$

Лемма доказана.

Лемма 6.7.2 **Сумма некоторого частного решения** *однородной* **системы (6.6.1) и некоторого частного решения** *неоднородной* **системы является частным решением** *неоднородной* **системы (6.6.1).**

Доказательство.

Пусть $\|x\|$ — частное решение однородной системы, а $\|y\|$ — некоторое частное решение неоднородной, то есть

$$\|A\|\|x\| = \|o\|\,, \qquad \|A\|\|y\| = \|b\|\,.$$

Тогда по правилам действий с матрицами справедливы равенства

$$\|A\|\big(\|x\| + \|y\|\big) = \|A\|\|x\| + \|A\|\|y\| = \|o\| + \|b\| = \|b\|\,.$$

Лемма доказана.

Лемма 6.7.3 **Разность двух некоторых частных решений *неоднородной* системы (6.6.1) является частным решением *однородной* системы (6.6.1).**

Доказательство.

Пусть $\|x\|$ и $\|y\|$ — частные решения неоднородной системы, то есть

$$\|A\|\|x\| = \|b\| , \qquad \|A\|\|y\| = \|b\| .$$

Тогда по правилам действий с матрицами справедливы равенства

$$\|A\|\bigl(\|x\| - \|y\|\bigr) = \|A\|\|x\| - \|A\|\|y\| = \|b\| - \|b\| = \|o\| .$$

Лемма доказана.

Замечания. 1°. Из лемм 6.7.1–6.7.3. вытекает (см. следст. 6.7.1), что

общее решение неоднородной системы уравнений есть общее решение однородной плюс некоторое (*любое*!) частное решение неоднородной,

и поэтому целесообразно вначале изучить вопрос о нахождении общего решения однородной системы линейных уравнений.

2°. Однородная система линейных уравнений всегда совместна, поскольку у нее есть, по крайней мере, одно частное, называемое тривиальным, решение, для которого все неизвестные имеют нулевое значение.

3°. Поскольку частные решения системы линейных уравнений представимы в виде столбцов, то, используя понятие равенства, операции сложения и умножения на число для столбцов, а также лемму 6.7.1, можно ввести понятие *линейной зависимости решений* однородной системы линейных уравнений, аналогично определению 6.5.4.

Теорема 6.7.1 **Однородная система линейных уравнений (6.6.1) имеет $n - \mathrm{rg}\|A\|$ линейно независимых частных решений.**

Доказательство.

1°. Рассмотрим вначале совместную *неоднородную* систему (6.6.1):

$$\left\{\begin{array}{ccccccccc} \alpha_{11}\xi_1 & + & \alpha_{12}\xi_2 & + & \cdots & + & \alpha_{1n}\xi_n & = & \beta_1 \\ \alpha_{21}\xi_1 & + & \alpha_{22}\xi_2 & + & \cdots & + & \alpha_{2n}\xi_n & = & \beta_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{m1}\xi_1 & + & \alpha_{m2}\xi_2 & + & \cdots & + & \alpha_{mn}\xi_n & = & \beta_m \end{array}\right.$$

и предположим, что матрица базисного минора расширенной матрицы $\|A|b\|$, ранга r, расположена в левом верхнем углу последней.

По теореме 6.5.1 (о базисном миноре) последние $m-r$ уравнений являются линейными комбинациями первых r уравнений, и, следовательно, их можно отбросить, поскольку они будут тождественно удовлетворяться решениями первых r уравнений.

В оставшихся уравнениях перенесем в правые части слагаемые, содержащие неизвестные $\xi_{r+1}, \xi_{r+2}, \ldots, \xi_n$, и получим т.н. *упрощенную* систему линейных уравнений, равносильную системе (6.6.1):

$$\sum_{j=1}^{r} \alpha_{ij}\xi_j = \beta_i - \sum_{j=r+1}^{n} \alpha_{ij}\xi_j \qquad i = [1, r] \, .$$

Неизвестные $\xi_1, \xi_2, \ldots, \xi_r$ будем называть *основными* (*главными, зависимыми, базисными*), а неизвестные $\xi_{r+1}, \xi_{r+2}, \ldots, \xi_n$ — *свободными* (*параметрическими, независимыми, небазисными*).

Присвоим свободным неизвестным некоторые конкретные значения $\xi_{r+1} = \mu_1, \xi_{r+2} = \mu_2, \ldots, \xi_n = \mu_{n-r}$ и рассчитаем по правилу Крамера (теорема 6.4.1) соответствующие им значения основных неизвестных:

$$\xi_j = \frac{1}{M} \det \left\| \begin{array}{ccccccc} \alpha_{11} & \alpha_{12} & \cdots & \beta_1 - \sum\limits_{k=r+1}^{n} \alpha_{1k}\mu_{k-r} & \cdots & \alpha_{1r} \\ \alpha_{21} & \alpha_{22} & \cdots & \beta_2 - \sum\limits_{k=r+1}^{n} \alpha_{2k}\mu_{k-r} & \cdots & \alpha_{2r} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{r1} & \alpha_{r2} & \cdots & \beta_r - \sum\limits_{k=r+1}^{n} \alpha_{rk}\mu_{k-r} & \cdots & \alpha_{rr} \end{array} \right\| \tag{6.7.1}$$

$\uparrow$ – j-й столбец.

В формуле (6.7.1) $j = [1, r]$, а M — базисный минор. Заметим, что из соотношений (6.7.1), положив, например, $\mu_k = 0 \quad \forall k = [1, n - r]$, можно найти частное решение неоднородной системы (6.6.1).

2°. Теперь рассмотрим *однородную* систему, заменив в (6.6.1) все $\beta_i \quad i = [1, m]$ нулями. По линейному свойству определителей (теорема 6.2.3) получаем выражения для значений основных неизвестных:

$$\begin{aligned} \xi_j &= \sum_{k=1}^{n-r} \kappa_{jk}\mu_k && \forall j = [1, r] \\ \xi_{r+i} &= \mu_i && \forall i = [1, n-r]\,, \end{aligned} \tag{6.7.2}$$

где

$$\kappa_{jk} = \frac{1}{M} \det \left\| \begin{matrix} \alpha_{11} & \alpha_{12} & \cdots & -\alpha_{1,r+k} & \cdots & \alpha_{1r} \\ \alpha_{21} & \alpha_{22} & \cdots & -\alpha_{2,r+k} & \cdots & \alpha_{2r} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{r1} & \alpha_{r2} & \cdots & -\alpha_{r,r+k} & \cdots & \alpha_{rr} \end{matrix} \right\|$$

$\uparrow$ – j-й столбец

$$j = [1, r], \quad k = [1, n - r].$$

Наконец, в матричной форме соотношения (6.7.2) могут быть записаны в виде

$$\left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \cdots \\ \xi_r \end{matrix} \right\| = \left\| \begin{matrix} \kappa_{11} & \kappa_{12} & \cdots & \kappa_{1r} \\ \kappa_{21} & \kappa_{22} & \cdots & \kappa_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ \kappa_{r1} & \kappa_{r2} & \cdots & \kappa_{rr} \end{matrix} \right\| \left\| \begin{matrix} \xi_{r+1} \\ \xi_{r+2} \\ \cdots \\ \xi_n \end{matrix} \right\| \tag{6.7.3}$$

или

$$\left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \cdots \\ \xi_r \\ \xi_{r+1} \\ \xi_{r+2} \\ \cdots \\ \xi_n \end{matrix} \right\| = \left\| \begin{matrix} \kappa_{11} & \kappa_{12} & \cdots & \kappa_{1r} \\ \kappa_{21} & \kappa_{22} & \cdots & \kappa_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ \kappa_{r1} & \kappa_{r2} & \cdots & \kappa_{rr} \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{matrix} \right\| \left\| \begin{matrix} \mu_1 \\ \mu_2 \\ \cdots \\ \mu_{n-r} \end{matrix} \right\| .$$

3°. Полагая $\mu_1 = 1, \mu_2 = \mu_3 = \ldots = \mu_{n-r} = 0$, получим решение $\{\, \xi_{1(1)}, \xi_{2(1)}, \ldots, \xi_{r(1)}, 1, 0, \ldots, 0 \,\}$ [10] .
Аналогично, при $\mu_1 = 0, \mu_2 = 1, \mu_3 = \ldots = \mu_{n-r} = 0$ найдем решение $\{\, \xi_{1(2)}, \xi_{2(2)}, \ldots, \xi_{r(2)}, 0, 1, \ldots, 0 \,\}$.
И, продолжая этот процесс, на последнем шаге при $\mu_1 = 0, \mu_2 = 0, \mu_3 = 0 \ldots, \mu_{n-r} = 1$ найдем решение $\{\, \xi_{1(n-r)}, \xi_{2(n-r)}, \ldots, \xi_{r(n-r)}, 0, 0, \ldots, 1 \,\}$.
Полученные решения будем называть *нормальными фундаментальными решениями.*

4°. Покажем теперь, что $n-r$ построенных частных решений однородной системы уравнений (6.6.1) являются линейно независимыми. Действительно, записав эти решения как строки, получим матрицу вида

$$\left\| \begin{matrix} \xi_{1(1)} & \xi_{2(1)} & \ldots & \xi_{r(1)} & 1 & 0 & \ldots & 0 \\ \xi_{1(2)} & \xi_{2(2)} & \ldots & \xi_{r(2)} & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \xi_{1(n-r)} & \xi_{2(n-r)} & \ldots & \xi_{r(n-r)} & 0 & 0 & \ldots & 1 \end{matrix} \right\| \tag{6.7.4}$$

Заметим, что ее ранг, с одной стороны, не меньше, чем $n-r$, поскольку содержит ненулевой минор этого порядка, но, с другой стороны, не больше, чем число строк в этой матрице, равное $n-r$, и потому ранг в точности равен $n-r$, что доказывает линейную независимость построенных частных решений.

Теорема доказана.

Определение 6.7.1 *Фундаментальной системой решений* для *однородной* системы линейных уравнений (6.6.1) называется упорядоченная совокупность любых ее $n - \mathrm{rg}\|A\|$ частных, линейно независимых решений, где n — число неизвестных системы (6.6.1), а $\|A\|$ — ее основная матрица.

Матрицу (6.7.4), которая определена только при $n > \mathrm{rg}\|A\|$, принято называть *фундаментальной матрицей* однородной системы (6.6.1).

[10] Здесь и далее, нижний индекс в круглых скобках является порядковым номером объекта в некотором множестве (тут это — номер частного решения), нижний индекс без скобок есть номер компоненты вектора или матрицы.

Теорема 6.7.2 **Каждое частное решение однородной системы (6.6.1) может быть представлено в виде линейной комбинации частных решений, образующих нормальную фундаментальную систему решений.**

Доказательство.

Пусть дано $\| \, \xi_1 \, \xi_2 \, \dots \, \xi_n \, \|$ — некоторое решение однородной системы (6.6.1). Используем это решение как первую строку в матрице размера $(n-r+1) \times n$

$$\left\| \begin{array}{cccccccc} \xi_1 & \xi_2 & \dots & \xi_r & \xi_{r+1} & \xi_{r+2} & \dots & \xi_n \\ \xi_{1(1)} & \xi_{2(1)} & \dots & \xi_{r(1)} & 1 & 0 & \dots & 0 \\ \xi_{1(2)} & \xi_{2(2)} & \dots & \xi_{r(2)} & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \xi_{1(n-r)} & \xi_{2(n-r)} & \dots & \xi_{r(n-r)} & 0 & 0 & \dots & 1 \end{array} \right\| , \qquad (6.7.5)$$

остальные строки в которой суть нормальные фундаментальные решения. Ранг матрицы (6.7.5), с одной стороны, очевидно, не меньше, чем $n-r$.

С другой стороны, первые r столбцов этой матрицы являются линейными комбинациями (заданными соотношениями (6.7.3)) последних $n-r$ столбцов.

Действительно, эти соотношения, связывающие значения свободных и основных переменных, *одни и те же для всех строк* матрицы (6.7.5), и потому в этой матрице каждый из первых r столбцов есть линейная комбинация последних $n-r$. Значит, ранг матрицы не превосходит $n-r$ и, следовательно, равен в точности $n-r$.

Значит, по теореме 6.5.1 о базисном миноре, последние $n-r$ строк матрицы (6.7.5) суть базисные. Тогда первая строка матрицы (6.7.5) есть некоторая линейная комбинация остальных, и, следовательно, произвольное частное решение однородной системы (6.6.1) может быть записано в виде

$$\left\| \begin{array}{c} \xi_1 \\ \xi_2 \\ \dots \\ \xi_r \\ \xi_{r+1} \\ \xi_{r+2} \\ \dots \\ \xi_n \end{array} \right\| = \lambda_1 \left\| \begin{array}{c} \xi_{1(1)} \\ \xi_{2(1)} \\ \dots \\ \xi_{r(1)} \\ 1 \\ 0 \\ \dots \\ 0 \end{array} \right\| + \lambda_2 \left\| \begin{array}{c} \xi_{1(2)} \\ \xi_{2(2)} \\ \dots \\ \xi_{r(2)} \\ 0 \\ 1 \\ \dots \\ 0 \end{array} \right\| + \dots + \lambda_{n-r} \left\| \begin{array}{c} \xi_{1(n-r)} \\ \xi_{2(n-r)} \\ \dots \\ \xi_{r(n-r)} \\ 0 \\ 0 \\ \dots \\ 1 \end{array} \right\| ,$$

где $\lambda_k \quad k = [1, n-k]$ — произвольные константы.

Теорема доказана.

Следствие 6.7.1 **Общее решение неоднородной системы (6.6.1) может быть дано формулой**

$$\left\|\begin{array}{c} \xi_1 \\ \xi_2 \\ \cdots \\ \xi_r \\ \xi_{r+1} \\ \xi_{r+2} \\ \cdots \\ \xi_n \end{array}\right\| = \lambda_1 \left\|\begin{array}{c} \xi_{1(1)} \\ \xi_{2(1)} \\ \cdots \\ \xi_{r(1)} \\ 1 \\ 0 \\ \cdots \\ 0 \end{array}\right\| + \cdots + \lambda_{n-r} \left\|\begin{array}{c} \xi_{1(n-r)} \\ \xi_{2(n-r)} \\ \cdots \\ \xi_{r(n-r)} \\ 0 \\ 0 \\ \cdots \\ 1 \end{array}\right\| + \left\|\begin{array}{c} \xi_{1(0)} \\ \xi_{2(0)} \\ \cdots \\ \xi_{r(0)} \\ \xi_{r+1(0)} \\ \xi_{r+2(0)} \\ \cdots \\ \xi_{n(0)} \end{array}\right\|,$$

где $\|\, \xi_{1(0)} \ \xi_{2(0)} \ \ldots \ \xi_{r(0)} \ \xi_{r+1(0)} \ \xi_{r+2(0)} \ \ldots \ \xi_{n(0)} \,\|^{\mathrm{T}}$ **является некоторым частным решением неоднородной системы (6.6.1), а числа** $\lambda_k \in \mathbb{R} \quad k = [1, n-k]$**.**

Доказательство.

Пусть $\|\, x^0 \|$ — некоторое *частное* решение *неоднородной* системы (6.6.1), а $\|x\|$ — ее *произвольное* решение.

Тогда по лемме 6.7.3 разность $\|y\| = \|x\| - \|x^0\|$ будет частным решением *однородной* системы для любого $\|x\|$.

С другой стороны, для любого частного решения *однородной* системы $\|y\|$ по лемме 6.7.2 имеем, что $\|y\|+\|x^0\|$ есть частное решение *неоднородной.*

Откуда заключаем, что утверждение доказываемого следствия справедливо.

Следствие доказано.

Из теорем 6.7.1 и 6.7.2 непосредственно вытекает

Следствие 6.7.2 **Для того чтобы** *однородная* **система (6.6.1) имела** *ненулевое* **решение, необходимо и достаточно, чтобы ранг ее основной матрицы удовлетворял условию** $\mathrm{rg}\|A\| < \min\{n, m\}$**.**

В случае, когда основная матрица однородной системы (6.6.1) квадратная, условие существования нетривиального решения равносильно равенству $\det \|A\| = 0$ **.**

Иное, полезное для приложений условие совместности системы линейных уравнений, дает

`Теорема 6.7.3 (Фредгольма)` **Для того чтобы система (6.6.1) была совместной, необходимо и достаточно, чтобы** *каждое* **решение** $\|y\| = \|\,\eta_1\,\eta_2\,\ldots\eta_m\,\|^{\mathrm{T}}$ **сопряженной системы** $\|A\|^{\mathrm{T}}\|y\| = \|o\|$ **удовлетворяло условию** $\sum\limits_{i=1}^{m}\beta_i\eta_i = 0$ **(или в матричном виде** $\|b\|^{\mathrm{T}}\|y\| = 0$ **).**

`Доказательство.`

Доказательство необходимости.

Пусть система уравнений (6.6.1) совместна, то есть для каждого ее решения $\|x\|$ справедливо равенство $\|b\| = \|A\|\|x\|$. Тогда, вычисляя произведение $\|b\|^{\mathrm{T}}\|y\|$ в предположении, что $\|A\|^{\mathrm{T}}\|y\| = \|o\|$, получаем

$$\|b\|^{\mathrm{T}}\|y\| = \big(\|A\|\|x\|\big)^{\mathrm{T}}\|y\| = \|x\|^{\mathrm{T}}\big(\|A\|^{\mathrm{T}}\|y\|\big) = \|x\|^{\mathrm{T}}\|o\| = 0\,.$$

Доказательство достаточности.

Пусть $\|b\|^{\mathrm{T}}\|y\| = 0$ для любого решения системы линейных уравнений $\|A\|^{\mathrm{T}}\|y\| = \|o\|$. Тогда общие решения систем линейных уравнений

$$\|A\|^{\mathrm{T}}\|y\| = \|o\| \qquad \text{и} \qquad \begin{cases} \|A\|^{\mathrm{T}}\|y\| = \|o\| \\ \|b\|^{\mathrm{T}}\|y\| = 0 \end{cases}$$

совпадают, и для этих систем максимальное число линейно независимых частных решений одинаково. Поэтому, согласно теоремам 6.7.1 и 6.7.2,

$$m - \mathrm{rg}\|A\|^{\mathrm{T}} = m - \mathrm{rg}\left\|\frac{A}{b}\right\|^{\mathrm{T}} \qquad \text{или} \qquad \mathrm{rg}\|A\|^{\mathrm{T}} = \mathrm{rg}\left\|\frac{A}{b}\right\|^{\mathrm{T}},$$

но поскольку ранг матрицы не меняется при ее транспонировании, то имеет место равенство $\mathrm{rg}\|A\| = \mathrm{rg}\|A\,|\,b\,\|$, означающее в силу теоремы 6.6.1 (Кронекера – Капелли) совместность системы линейных уравнений (6.6.1).

`Теорема доказана.`

Альтернативное доказательство теоремы Фредгольма приведено в главе 10 (см. теоремы 10.6.4 и 10.6.5).

6.8. Элементарные преобразования матриц. Метод Гаусса

Практическое применение теорем 6.7.3 и 6.7.4 затрудняется тем, что заранее, как правило, неизвестно, совместна ли решаемая система. Определение же рангов основной и расширенной матриц независимо от поиска решений оказывается весьма нерациональной (с точки зрения расходования вычислительных ресурсов) процедурой. Более эффективным вычислительным алгоритмом, позволяющим либо находить общее решение системы (6.6.1), либо устанавливать факт ее несовместности, является *метод Гаусса.*

Суть этого метода заключается в приведении расширенной матрицы системы линейных уравнений к наиболее простому виду последовательностью так называемых *элементарных преобразований*, каждое из которых не меняет общего решения системы уравнений. Под «наиболее простым» видом расширенной матрицы мы будем понимать *верхнюю треугольную форму* (т.е. случай, когда $\alpha_{ij} = 0$ при $i > j$), для которой возможно рекуррентное нахождение неизвестных путем решения на каждом шаге процедуры лишь линейного уравнения *с одним* неизвестным. Ниже приведен пример матрицы размера $m \times n$, имеющей верхнюю треугольную форму:

$$\left\|\begin{array}{cccccccccc} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1,m-2} & \alpha_{1,m-1} & \alpha_{1,m} & \alpha_{1,m+1} & \cdots & \alpha_{1,n} \\ 0 & \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2,m-2} & \alpha_{2,m-1} & \alpha_{2,m} & \alpha_{2,m+1} & \cdots & \alpha_{2,n} \\ 0 & 0 & \alpha_{33} & \cdots & \alpha_{3,m-2} & \alpha_{3,m-1} & \alpha_{3,m} & \alpha_{3,m+1} & \cdots & \alpha_{3,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & \alpha_{m-1,m-1} & \alpha_{m-1,m} & \alpha_{m-1,m+1} & \cdots & \alpha_{m-1,n} \\ 0 & 0 & 0 & \cdots & 0 & 0 & \alpha_{m,m} & \alpha_{m,m+1} & \cdots & \alpha_{m,n} \end{array}\right\|.$$

К элементарным преобразованиям матрицы относятся:

- перестановка строк (перенумерация уравнений);
- перестановка столбцов основной матрицы (перенумерация неизвестных);
- удаление нулевой строки (исключение уравнений, тождественно удовлетворяющихся любыми значениями неизвестных);

— умножение строки на ненулевое число (нормирование уравнений);

— сложение строки с линейной комбинацией остальных строк, с записью результата на место исходной строки (замена одного из уравнений системы следствием ее уравнений, получаемым при помощи линейных операций).

Решение неоднородной системы уравнений (равно как и ранг ее матрицы) не изменится также и при любой комбинации элементарных операций с расширенной матрицей этой системы..

Непосредственной проверкой можно убедиться, что элементарные преобразования любой матрицы могут быть выполнены при помощи умножения ее на матрицы следующего специального вида. Например:

— перестановка столбцов с номерами i и j матрицы $\|A\|$ размера $m\times n$ осуществляется путем ее умножения справа на матрицу $\|S_1\|$ размера $n \times n$, которая в свою очередь получается из единичной матрицы $\|E\|$ n-го порядка путем перестановки в последней i-го и j-го столбцов;

— умножение i-й строки матрицы $\|A\|$ на число $\lambda \neq 0$ осуществляется путем умножения слева на матрицу $\|S_2\|$, которая получается из единичной размера $m\times m$ матрицы $\|E\|$ путем замены в последней i-го диагонального элемента (равного единице) на λ;

— сложение строк с номерами i и j матрицы $\|A\|$ осуществляется путем ее умножения слева на матрицу $\|S_3\|$ размера $m\times m$, которая получается из единичной матрицы $\|E\|$ порядка m путем замены в последней нулевого элемента, стоящего в i-й строке и j-м столбце, на единицу. При этом результат суммирования строк i и j матрицы $\|A\|$ окажется i-й строкой результирующей матрицы $\|S_3\| \cdot \|A\|$.

В дальнейшем (см. теорему 8.4.3) будет показано, что если матрица $\|S\|$ квадратная и невырожденная, и возможно умножение матрицы $\|S\|$ на матрицу $\|A\|$, то справедливо равенство

$$\operatorname{rg}\big(\|S\|\|A\|\big) = \operatorname{rg}\|A\|\,.$$

В рассмотренных выше примерах $\det\|S_1\| = \det\|S_3\| = -1$ и $\det\|S_2\| = \lambda \neq 0$. Поэтому ранг матрицы $\|A\|$ при таких элементарных преобразованиях не меняется.

Проверьте самостоятельно, что будут также справедливы следующие теоремы.

Теорема 6.8.1 **Последовательное применение нескольких элементарных преобразований также есть** *элементарное преобразование*, **матрица которого равна произведению матриц данных элементарных преобразований.**

Теорема 6.8.2 **Если умножение матрицы $\|A\|$ слева на квадратную матрицу $\|S\|$ реализует некоторое преобразование над строками $\|A\|$, то умножение $\|A\|$ справа на $\|S\|$ реализует то же самое преобразование матрицы $\|A\|$, но выполненное над ее столбцами.**

Отмеченные свойства элементарных преобразований достаточно часто позволяют упрощать вычислительные процедуры с матричными выражениями. Пусть, например, $\|S^*\|$ есть матрица преобразования, переводящего невырожденную квадратную матрицу $\|A\|$ в единичную. Тогда преобразование с матрицей $\|S^*\|$ переведет единичную матрицу$\|E\|$ в матрицу $\|A\|^{-1}$, поскольку в силу $\|E\| = \|S^*\|\|A\|$ и невырожденности $\|A\|$ справедливы равенства

$$\|E\|\|A\|^{-1} = \|S^*\|\|A\|\|A\|^{-1} \quad \text{или} \quad \|A\|^{-1} = \|S^*\|\|E\|\,.$$

Из этих соотношений следует, что вычисление произведения квадратных матриц $\|A\|^{-1}\|B\|$ может быть сведено к последовательности элементарных преобразований матрицы $\|A|B\|$ (то есть матрицы, образованной добавлением столбцов матрицы $\|B\|$ к матрице $\|A\|$), приводящих подматрицу $\|A\|$ к единичной. В результате искомое произведение оказывается на месте подматрицы $\|B\|$.

Проиллюстрируем применение метода Гаусса на примере решения следующей системы линейных уравнений.

Задача 6.8.1 *Решить систему уравнений*

$$\left\{\begin{array}{rcrcrcrcrcr} \xi_1 & + & \xi_2 & + & \xi_3 & + & \xi_4 & + & \xi_5 & = & 7, \\ 3\xi_1 & + & 2\xi_2 & + & \xi_3 & + & \xi_4 & - & 3\xi_5 & = & -2, \\ & & 2\xi_2 & + & 2\xi_3 & + & \xi_4 & + & 6\xi_5 & = & 23, \\ 5\xi_1 & + & 4\xi_2 & + & 3\xi_3 & + & 3\xi_4 & - & \xi_5 & = & 12. \end{array}\right.$$

`Решение.` 1°. Составляем расширенную матрицу решаемой системы

$$\left\|\begin{array}{rrrrr|r} 1 & 1 & 1 & 1 & 1 & 7 \\ 3 & 2 & 1 & 1 & -3 & -2 \\ 0 & 1 & 2 & 2 & 6 & 23 \\ 5 & 4 & 3 & 3 & -1 & 12 \end{array}\right\| .$$

2°. Преобразуем ее к верхнему треугольному виду. Для этого

а) преобразуем в нули все элементы первого столбца, кроме элемента, стоящего в первой строке.

Например, для зануления элемента, стоящего во второй строке первого столбца, заменим вторую строку матрицы строкой, которая является суммой первой строки, умноженной на -3, и второй строки. Символически это выглядит как $(-3)\cdot\boxed{1}+\boxed{2} \longrightarrow \boxed{2}$.

Аналогично поступаем с четвертой строкой: ее заменяем линейной комбинацией первой и четвертой строк с коэффициентами -5 и 1 соответственно. Третью, естественно, не меняем: там уже имеется необходимый для верхнего треугольного вида ноль. В итоге матрица приобретает вид

$$\left\|\begin{array}{rrrrr|r} 1 & 1 & 1 & 1 & 1 & 7 \\ 0 & -1 & -2 & -2 & -6 & -23 \\ 0 & 1 & 2 & 2 & 6 & 23 \\ 0 & -1 & -2 & -2 & -6 & -23 \end{array}\right\| ;$$

б) выполняем теперь операцию зануления элементов второго столбца, стоящих в его третьей и четвертой строках. Для этого третью строку матрицы заменяем суммой второй и третьей, а четвертую – разностью второй и четвертой. Получаем

$$\left\|\begin{array}{rrrrr|r} 1 & 1 & 1 & 1 & 1 & 7 \\ 0 & -1 & -2 & -2 & -6 & -23 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right\| ;$$

в) поскольку в данном конкретном случае элемент, расположенный в четвертой строке третьего столбца, оказался равным нулю, то приведение расширенной матрицы к верхнему треугольному виду завершено.

3°. Полученная матрица является расширенной матрицей упрощенной системы линейных уравнений, равносильной исходной системе. Ранг этой матрицы совпадает с рангом исходной. Потому заключаем, что

а) решаемая система совместна по теореме 6.6.1 (Кронекера – Капелли), поскольку ранг основной матрицы равен рангу расширенной и равен 2;

б) однородная система уравнений будет иметь по теореме 6.7.1 $\quad n - \mathrm{rg}\|A\| = 5 - 2 = 3 \quad$ линейно независимых решения.

4°. Поскольку общее решение неоднородной системы есть общее решение однородной плюс частное решение неоднородной, то нам достаточно найти три любых линейно независимых решения однородной системы и какое-нибудь одно решение неоднородной.

Перепишем исходную систему в преобразованном упрощенном виде (как в доказательстве теоремы 6.7.1), приняв первое и второе неизвестные за основные, а третье, четвертое и пятое – за свободные:

$$\left\{\begin{array}{ccccccccccc} \xi_1 & + & \xi_2 & = & 7 & - & \xi_3 & - & \xi_4 & - & \xi_5, \\ & & \xi_2 & = & 23 & - & 2\xi_3 & - & 2\xi_4 & - & 6\xi_5. \end{array}\right. \tag{6.8.1}$$

Второе уравнение для удобства вычислений мы умножили на -1, а третье и четвертое уравнения отбросили как удовлетворяющиеся тождественно.

Положив теперь в системе (6.8.1) свободные неизвестные равными нулю, находим частное решение неоднородной системы $\quad \| \ -16 \ \ 23 \ \ 0 \ \ 0 \ \ 0 \ \|^{\mathrm{T}} \quad$ из легко решаемой системы линейных уравнений

$$\left\{\begin{array}{ccccc} \xi_1 & + & \xi_2 & = & 7, \\ & & \xi_2 & = & 23. \end{array}\right.$$

Отметим, что основная матрица этой системы (это гарантируется методом Гаусса) невырожденная и верхняя треугольная.

Для однородной системы (6.8.1), т.е. системы (6.8.1), в которой константы 7 и 23 заменены нулями:

$$\left\{\begin{array}{ccccccccccc} \xi_1 & + & \xi_2 & = & 0 & - & \xi_3 & - & \xi_4 & - & \xi_5, \\ & & \xi_2 & = & 0 & - & 2\xi_3 & - & 2\xi_4 & - & 6\xi_5, \end{array}\right.$$

строим нормальную фундаментальную систему решений по схеме, использованной при доказательстве теоремы 6.7.1.

Первое независимое решение $\|\, 1 \quad -2 \quad 1 \quad 0 \quad 0 \,\|^{\mathrm{T}}$ находится из системы

$$\left\{\begin{array}{ccccc} \xi_1 & + & \xi_2 & = & -1, \\ & & \xi_2 & = & -2. \end{array}\right.$$

Аналогичным образом находятся $\|\, 1 \ -2 \ 0 \ 1 \ 0 \,\|^{\mathrm{T}}$ и $\|\, 5 \ -6 \ 01 \ 0 \ 1 \,\|^{\mathrm{T}}$ — второе и третье фундаментальные решения (проверьте это самостоятельно).

Окончательно общее решение исходной неоднородной системы в матричном виде может быть записано как

$$\left\|\begin{array}{c} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \end{array}\right\| = \lambda_1 \left\|\begin{array}{r} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{array}\right\| + \lambda_2 \left\|\begin{array}{r} 1 \\ -2 \\ 0 \\ 1 \\ 0 \end{array}\right\| + \lambda_3 \left\|\begin{array}{r} 5 \\ -6 \\ 0 \\ 0 \\ 1 \end{array}\right\| + \left\|\begin{array}{r} -16 \\ -23 \\ 0 \\ 0 \\ 0 \end{array}\right\|$$

Решение получено.

$\forall \lambda_1,\ \lambda_2,\ \lambda_3 \in \mathbb{R}$.

Замечание 6.8.1. Поскольку существует свобода выбора как частного решения неоднородной системы, так и линейно независимых (фундаментальных) решений однородной, то общее решение системы может быть записано в различных, но, естественно, равносильных формах.

Просто, каждое частное решение будет получаться для различных форм записи общего решения при разных значениях параметров $\lambda_1,\ \lambda_2,\ \lambda_3 \in \mathbb{R}$.

7. Линейное пространство

7.1. Определение линейного пространства

`Определение 7.1.1`	Множество Λ, состоящее из элементов x, y, $z, \ldots$, к которым применимы как понятия *равенства* (вида $x = y$), так и *не равенства* ($x \neq y$), называется *линейным пространством*, если 1°. Каждой паре элементов $x, y \in \Lambda$ поставлен в соответствие третий элемент этого же множества, называемый их *суммой* и обозначаемый $x+y$, таким образом, что выполнены аксиомы а) $x + y = y + x$; б) $x + (y + z) = (x + y) + z$; в) существует *нулевой* элемент o, такой, что $\forall x \in \Lambda$ имеет место $x + o = x$; г) $\forall x \in \Lambda$ имеется *противоположный* элемент $\tilde{x}$, такой, что $x + \tilde{x} = o$. 2°. Для любого числа λ и $\forall x \in \Lambda$ существует такой принадлежащий Λ элемент, обозначаемый λx и называемый *произведением числа на элемент*, что выполнены аксиомы: а) $\forall x \in \Lambda : \qquad 1\,x = x$; б) $\forall x \in \Lambda$ и любых чисел λ и μ : $(\lambda\mu)x = \lambda(\mu x)$.

3°. Для операций сложения элементов и умножения числа на элемент $\forall x, y \in \Lambda$ и для любых чисел λ, μ выполнены аксиомы *дистрибутивности*:

а) $(\lambda + \mu)x = \lambda x + \mu x$;

б) $\lambda(x + y) = \lambda x + \lambda y$.

`Замечание 7.1.1.` 1°. Под «числами» в аксиомах второй и третьей групп определения 7.1.1 подразумеваются действительные или комплексные числа.

2°. Первые четыре аксиомы равносильны требованию, чтобы Λ являлось абелевой группой относительно операции сложения (см. § 5.6).

`Пример 7.1.1` Линейным пространством (в предположении, что операции введены стандартным образом) являются:

1°. Множество всех векторов в пространстве.

2°. Множество всех вещественных чисел $\mathbb{R}$.

3°. Множество всех n-компонентных столбцов.

4°. Множество всех алгебраических многочленов степени не выше, чем n.

5°. Множество всех матриц одного размера $m \times n$.

6°. $C[\alpha, \beta]$ — множество всех функций, непрерывных на $[\alpha, \beta]$.

7°. Общее решение однородной системы m линейных уравнений с n неизвестными.

`Задача 7.1.1` *Показать, что в общем случае множество радиусов-векторов точек в пространстве, принадлежащих плоскости* $(\vec{n}, \vec{r}) = d$, *не является линейным пространством. Выяснить, при каких значениях параметра* d *данное множество будет линейным пространством.*

`Задача 7.1.2` *Показать, что множество, состоящее из одного нулевого элемента, является линейным пространством.*

Задача 7.1.3 *Будет ли линейным пространством множество всех положительных чисел $\mathbb{R}^+$?*

Решение. Ответ зависит от способа введения операций *сложения* и *умножения на число* элементов рассматриваемого множества.

1°. Пусть операции вводятся стандартным образом. В этом случае множество положительных чисел не образует линейного пространства, поскольку в нем, например, отсутствует нулевой элемент.

2°. Если же операцию *сложения* определить как обычное умножение двух чисел, а *умножение числа* λ на элемент x определить как возведение положительного числа x в степень $\lambda \in \mathbb{R}$:

$$\begin{array}{lccl} \textit{сложение } x+y & := x \cdot y & & x>0, y>0, \\ \textit{умножение } \lambda x & := & x^\lambda & x>0, \lambda \in \mathbb{R}, \end{array}$$

то множество положительных чисел будет являться линейным пространством, в котором роль нулевого элемента играет число 1.

Решение получено.

Из аксиоматики линейного пространства вытекают свойства, описываемые следующими теоремами.

Теорема 7.1.1 **В линейном пространстве имеется единственный нулевой элемент.**

Доказательство.

Пусть существуют два различных нулевых элемента o_1 и o_2. Тогда, согласно аксиоме 1°в) из определения 7.1.1 линейного пространства, будут справедливы равенства

$$o_1 + o_2 = o_1 \qquad \text{и} \qquad o_2 + o_1 = o_2 \, .$$

Откуда в силу аксиомы 1°а) — коммутативности операции сложения — получаем $o_1 = o_2$.

Теорема доказана.

Теорема 7.1.2 $\forall x \in \Lambda$ **имеет место равенство** $0x = o$**.**

Доказательство.

Из аксиоматики линейного пространства имеем

$$x = 1x = (1+0)x = 1x + 0x = x + 0x\,.$$

Прибавляя к обеим частям равенства $x = 0x + x$ элемент $\tilde{x}$, противоположный элементу x, в силу $x + \tilde{x} = o$ получаем, что $0x = o$.

Теорема доказана.

Теорема 7.1.3 $\forall x \in \Lambda$ **существует единственный противоположный элемент.**

Доказательство.

Пусть для элемента $x \in \Lambda$ существуют два различных противоположных элемента $\tilde{x}_1$ и $\tilde{x}_2$. Тогда, согласно аксиоме 1°г), будут справедливы равенства $x + \tilde{x}_1 = o$ и $x + \tilde{x}_2 = o$. Прибавим к обеим частям первого равенства элемент $\tilde{x}_2$, в силу ассоциативности операции сложения и первого равенства, получим

$$x + \tilde{x}_1 + \tilde{x}_2 = (x + \tilde{x}_1) + \tilde{x}_2 = \tilde{x}_2\,.$$

Но, с другой стороны,

$$x + \tilde{x}_1 + \tilde{x}_2 = (x + \tilde{x}_2) + \tilde{x}_1 = \tilde{x}_1\,,$$

то есть $\tilde{x}_1 = \tilde{x}_2$.

Теорема доказана.

Теорема 7.1.4 $\forall x \in \Lambda$ **противоположным элементом служит элемент** $\tilde{x} = (-1)x$**.**

Доказательство.

Из аксиоматики линейного пространства и в силу теорем 7.1.2—7.1.3 имеем

$$o = 0x = (1-1)x = 1x + (-1)x = x + (-1)x\,.$$

Данное равенство и означает, что противоположный к x элемент имеет вид $(-1)x$.

Теорема доказана.

7.2. Линейная зависимость, размерность и базис в линейном пространстве

Определение 7.2.1

1°. Выражение $\sum\limits_{i=1}^{k} \lambda_i x_i$ называется *линейной комбинацией* элементов $x_1, x_2, \ldots, x_k$ линейного пространства Λ.

2°. Элементы $x_1, x_2, \ldots, x_k$ линейного пространства Λ называются *линейно зависимыми*, если существуют числа $\lambda_1, \lambda_2, \ldots, \lambda_k$, не равные нулю одновременно, такие, что $\sum\limits_{i=1}^{k} \lambda_i x_i = o$.

3°. Элементы линейного пространства называются *линейно независимыми*, если из равенства $\sum\limits_{i=1}^{k} \lambda_i x_i = o$ следует, что $\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0$.

Лемма 7.2.1

Для того чтобы некоторое множество элементов линейного пространства было линейно зависимым, необходимо и достаточно, чтобы один из этих элементов являлся линейной комбинацией остальных.

Доказательство.

Доказательство совпадает с доказательством леммы 1.4.1, в котором слово *вектор* заменено словом *элемент*.

Лемма доказана.

Лемма 7.2.2 **Если некоторое подмножество множества элементов линейно зависимо, то линейно зависимы и сами элементы.**

Доказательство.

Без ограничения общности можно предположить, что линейно зависимое подмножество состоит их первых $j < k$ элементов множества $x_1, x_2, \ldots, x_k$. Тогда существуют не равные нулю одновременно числа $\lambda_1, \lambda_2, \ldots, \lambda_k$, такие, что $\sum\limits_{i=1}^{j} \lambda_i x_i = o$. Но это равенство можно записать в виде

$$\sum_{i=1}^{j} \lambda_i x_i + \sum_{i=j+1}^{k} 0 \cdot x_i = o\,,$$

тогда из нетривиальности первой линейной комбинации, стоящей в левой части этого равенства, следует линейная зависимость набора элементов $x_1, x_2, \ldots, x_k$.

Лемма доказана.

Определение 7.2.2 *Базисом* в линейном пространстве Λ называется любой упорядоченный набор его n элементов $g_1, g_2, \ldots, g_n$, если

1) этот набор линейно независимый;
2) $\forall x \in \Lambda$ множество $\{g_1, g_2, \ldots, g_n, x\}$ линейно зависимое.

Определение 7.2.3 Линейное пространство Λ называется n-*мерным* и обозначается Λ^n, если в нем существует базис, состоящий из n элементов.
В этом случае число n называется *размерностью* линейного пространства Λ^n и обозначается $\dim \Lambda^n$.

Теорема 7.2.1 **Для каждого элемента линейного пространства Λ^n *существует единственное* представление в виде линейной комбинации базисных элементов.**

Доказательство.

Пусть в конечномерном линейном пространстве Λ^n заданы базис $\{ g_1, g_2, \ldots, g_n\}$ и произвольный элемент x. Тогда, по определению базиса, система элементов $\{ x, g_1, g_2, \ldots, g_n \}$ линейно зависима, то есть существуют не равные нулю одновременно числа $\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_n$, такие, что

$$\lambda_0 x + \sum_{i=1}^{n} \lambda_i g_i = o\,.$$

Покажите самостоятельно, что число $\lambda_0 \neq 0$, поскольку это противоречило бы линейной независимости базисных элементов. Поэтому

$$x = \sum_{i=1}^{n} \left(-\frac{\lambda_i}{\lambda_0} \right) g_i\,,$$

и существование разложения, таким образом, доказано.

Докажем теперь единственность разложения. Допустим, что существуют два различных разложения x по базису

$$x = \sum_{i=1}^{n} \xi_i g_i \qquad \text{и} \qquad x = \sum_{i=1}^{n} \eta_i g_i\,.$$

Тогда, вычитая эти равенства почленно, получаем, что

$$o = \sum_{i=1}^{n} (\xi_i - \eta_i) g_i\,.$$

Поскольку линейная комбинация линейно независимых элементов может равняться нулевому элементу, только если она тривиальная, то $\xi_i = \eta_i \quad \forall i = [1, n]$. Но это и означает, что разложение элемента x по базису единственно.

Теорема доказана.

В общем случае линейное пространство может не иметь базиса. Таким свойством обладает, например, линейное пространство, состоящее из одного нулевого элемента, поскольку в нем нет ни одного линейно независимого элемента.

Однако базиса может не быть и в линейном пространстве, имеющем линейно независимые элементы. Соответствующий пример можно найти в следующей таблице 7.2.1.

Т а б л и ц а 7.2.1

Линейное пространство	**Размерность**	**Пример базиса**
Множество всех векторов в пространстве	3	Любая упорядоченная некомпланарная тройка векторов
Множество всех n-компонентных столбцов	n	n столбцов вида $\left\lVert\begin{array}{c}1\\0\\\dots\\0\end{array}\right\rVert, \left\lVert\begin{array}{c}0\\1\\\dots\\0\end{array}\right\rVert, \dots, \left\lVert\begin{array}{c}0\\0\\\dots\\1\end{array}\right\rVert$
Множество всех алгебраических многочленов степени не выше, чем n	$n+1$	Набор из $n+1$ одночлена вида $P_0(\tau)=1,\ P_1(\tau)=\tau,\ \dots, P_n(\tau)=\tau^n$
Множество всех матриц размера $m\times n$	$m\cdot n$	$m\cdot n$ всевозможных различных матриц размера $m\times n$, все элементы которых равны нулю, кроме одного, равного 1
Множество всех функций $f(\tau)$, непрерывных на $[0,1]$	Базиса нет	Пояснение см. ниже
Множество решений однородной системы m линейных уравнений с n неизвестными и с рангом основной матрицы, равным r	$n-r$	Нормальная фундаментальная система решений

Заметим, что базиса в линейном пространстве функций, непрерывных на отрезке $[0,1]$, нет, поскольку $\forall n\in\mathbb{N}$ для каждого набора линейно независимых непрерывных функций $\{f_k(\tau)=\tau^k\quad k\in[1,n]\ \}$ имеется непрерывная функция $f_{n+1}(\tau)=\tau^{n+1}$, такая, что набор вида $\{\ f_k(\tau)=\tau^k\quad k\in[1,n+1]\ \}$ оказывается линейно независимым. А это противоречит определению 7.2.2.

7.3. Подмножества линейного пространства

Подпространство

Определение 7.3.1	Непустое множество Ω, образованное из элементов линейного пространства Λ, называется *подпространством* этого линейного пространства, если для любых $x, y \in \Omega$ и любого числа λ 1) $x + y \in \Omega$, 2) $\lambda x \in \Omega$.

Замечание 7.3.1: из определения 7.3.1 следует, что множество Ω само является *линейным пространством* [11], поскольку для него, очевидно, выполняются все аксиомы определения линейного пространства.

Пример 7.3.1

Подпространства линейного пространства:

1°. Множество радиусов-векторов всех точек, лежащих на некоторой плоскости, проходящей через начало координат, является подпространством во множестве радиусов-векторов всех точек трехмерного геометрического пространства.

2°. Множество всех многочленов степени не выше, чем n, есть подпространство в линейном пространстве непрерывных на $[\alpha, \beta]$ функций.

3°. В пространстве n-мерных столбцов совокупность частных решений однородной системы m линейных уравнений с n неизвестными и с основной матрицей ранга r образует подпространство размерности $n - r$.

4°. Подпространством любого линейного пространства будет:

а) само линейное пространство;

б) множество, состоящее из одного нулевого элемента.

[11] По умолчанию считается, что набор операций в подпространстве тот же, что и в исходном пространстве. Например, пространство $\mathbb{R}^+$ в задаче 7.1.3 не является попространством пространства $\mathbb{R}$ в 2° примера 7.1.1.

Определение 7.3.2

Пусть даны два подпространства Ω_1 и Ω_2 линейного пространства Λ. Тогда

1°. *Объединением* подпространств Ω_1 и Ω_2 называется множество элементов $x \in \Lambda$, таких, что либо $x \in \Omega_1$, либо $x \in \Omega_2$. Объединение подпространств Ω_1 и Ω_2 обозначается $\Omega_1 \cup \Omega_2$.

2°. *Пересечением* подпространств Ω_1 и Ω_2 называется множество элементов $x \in \Lambda$, таких, что $x \in \Omega_1$ и $x \in \Omega_2$. Пересечение подпространств Ω_1 и Ω_2 обозначается $\Omega_1 \cap \Omega_2$.

3°. *Суммой* подпространств Ω_1 и Ω_2 называется совокупность всех элементов $x = x_1 + x_2 \in \Lambda$ при условии, что $x_1 \in \Omega_1$ и $x_2 \in \Omega_2$. Сумма подпространств Ω_1 и Ω_2 обозначается $\Omega_1 + \Omega_2$.

4°. *Прямой суммой* подпространств Ω_1 и Ω_2 называется совокупность всех элементов x таких, что $x = x_1 + x_2 \in \Lambda$ при условиях: $x_1 \in \Omega_1$, $x_2 \in \Omega_2$ и $\Omega_1 \cap \Omega_2 = o$. Прямая сумма подпространств Ω_1 и Ω_2 обозначается $\Omega_1 \oplus \Omega_2$.

Покажите самостоятельно, что справедлива

Теорема 7.3.1 **Как сумма, так и пересечение подпространств Ω_1 и Ω_2 в Λ суть также подпространства в Λ.**

Теорема 7.3.2 **Размерность суммы подпространств Ω_1 и Ω_2 равна** $\dim(\Omega_1 + \Omega_2) = \dim \Omega_1 + \dim \Omega_2 - \dim(\Omega_1 \cap \Omega_2)$.

Доказательство.

1°. Пусть подпространство $\Omega_1 \cap \Omega_2$ имеет размерность k и базис $\{ g_1, g_2, \ldots, g_k \}$. Дополним этот базис элементами $\{ g'_1, g'_2, \ldots, g'_l \}$ до базиса в Ω_1 и элементами $\{ g''_1, g''_2, \ldots, g''_m \}$ до базиса в Ω_2.

В этом случае каждый элемент может быть разложен по системе элементов:

$$\left\{ g_1, g_2, \ldots, g_k, g'_1, g'_2, \ldots, g'_l, g''_1, g''_2, \ldots, g''_m \right\}.$$

2°. Покажем теперь, что набор элементов

$$\left\{ g_1,\, g_2,\, \ldots,\, g_k,\, g'_1,\, g'_2,\, \ldots,\, g'_l,\, g''_1,\, g''_2,\, \ldots,\, g''_m \right\}$$

линейно независим в Λ. Рассмотрим некоторую, равную нулевому элементу, линейную комбинацию этих элементов:

$$\sum_{i=1}^{l} \lambda'_i g'_i + \sum_{j=1}^{k} \lambda_j g_j + \sum_{p=1}^{m} \lambda''_p g''_p = o. \tag{7.3.1}$$

Заметим, что по построению элемент

$$\tilde{x} = \sum_{p=1}^{m} \lambda''_p g''_p \quad \in \quad \Omega_2,$$

но, с другой стороны, этот же элемент в силу (7.3.1)

$$\tilde{x} = -\left(\sum_{i=1}^{l} \lambda'_i g'_i + \sum_{j=1}^{k} \lambda_j g_j \right) \quad \in \quad \Omega_1.$$

Это означает, что $\tilde{x} \in \Omega_1 \cap \Omega_2$ и, следовательно, в равенстве (7.3.1) все $\lambda'_i = 0 \ \forall i = [1, l] \qquad \lambda''_p = 0 \ \forall p = [1, m]$. А поскольку $\{ g_1,\, g_2,\, \ldots,\, g_k \}$ — базис в $\Omega_1 \cap \Omega_2$, то и все $\lambda_j = 0 \ \forall i = [1, k]$ и линейная комбинация, стоящая в левой части равенства (7.3.1), тривиальная. Следовательно, $\left\{ g_1,\, g_2,\, \ldots,\, g_k,\, g'_1,\, g'_2,\, \ldots,\, g'_l,\, g''_1,\, g''_2,\, \ldots,\, g''_m \right\}$ — линейно независимая система элементов.

3°. Из пункта 2° следует, что набор элементов

$$\left\{ g_1,\, g_2,\, \ldots,\, g_k,\, g'_1,\, g'_2,\, \ldots,\, g'_l,\, g''_1,\, g''_2,\, \ldots,\, g''_m \right\}$$

является базисом в $\Omega_1 + \Omega_2$. Размерность подпространства при этом равна

$$\dim(\Omega_1 + \Omega_2) = k + l + m = (k + l) + (k + m) - k =$$

$$= \dim \Omega_1 + \dim \Omega_2 - \dim(\Omega_1 \cap \Omega_2)\,.$$

Теорема доказана.

Следствие 7.3.1 **В случае прямой суммы подпространств Ω_1 и Ω_2**

$$\dim(\Omega_1 \oplus \Omega_2) = \dim \Omega_1 + \dim \Omega_2$$

и каждый элемент $x \in (\Omega_1 \oplus \Omega_2)$ представим *единственным* **образом в виде $x = x_1 + x_2$ так, что $x_1 \in \Omega_1$ и $x_2 \in \Omega_2$, поскольку набор элементов**

$$\left\{ g'_1,\, g'_2,\, \ldots,\, g'_l,\, g''_1,\, g''_2,\, \ldots,\, g''_m \right\}$$

является базисом в $\Omega_1 \oplus \Omega_2$.

Линейная оболочка набора элементов

Определение 7.3.3 Совокупность всевозможных линейных комбинаций некоторого множества элементов $\left\{ x_1,\, x_2,\, \ldots,\, x_k \right\}$ линейного пространства Λ называется *линейной оболочкой* этого множества и обозначается $L\left\{ x_1,\, x_2,\, \ldots,\, x_k \right\}$.

Пример 7.3.2 Множество многочленов степени не выше, чем n, является линейной оболочкой набора одночленов вида $\left\{ 1,\, \tau,\, \tau^2 \ldots,\, \tau^n \right\}$ в линейном пространстве непрерывных на $[\alpha, \beta]$ функций $f(\tau)$.

Пусть задан набор элементов $\left\{ x_1,\, x_2,\, \ldots,\, x_k \right\} \subseteq \Lambda$, порождающих линейную оболочку $L\left\{ x_1,\, x_2,\, \ldots,\, x_k \right\}$, тогда любой элемент этой линейной оболочки имеет вид $x = \sum\limits_{i=1}^{k} \lambda_i x_i$ и справедлива

Теорема 7.3.3 **Множество всех элементов, принадлежащих линейной оболочке $L\left\{ x_1,\, x_2,\, \ldots,\, x_k \right\}$, является в Λ** *подпространством размерности* m**, где m — максимальное число линейно независимых элементов в наборе $\left\{ x_1,\, x_2,\, \ldots,\, x_k \right\}$.**

Доказательство.

1°. Непосредственной проверкой убеждаемся, что для совокупности элементов вида $x = \sum\limits_{i=1}^{k} \lambda_i x_i$ (в предположении, что λ_i суть произвольные числа) справедливы все аксиомы из определения 7.1.1, то есть рассматриваемая линейная оболочка является линейным пространством.

2°. Пусть максимальное число линейно независимых элементов в наборе $\left\{ x_1, x_2, \ldots, x_k \right\}$ равно $m \le k$. Без ограничения общности можно считать, что этими элементами являются $x_1, x_2, \ldots, x_m$. Тогда

$$x_j = \sum_{i=m+1}^{k} \alpha_{ji} x_i \qquad \forall j = [1, m]\,,$$

и любой элемент линейной оболочки может быть представлен в виде линейной комбинации элементов $x_1, x_2, \ldots, x_m$.

3°. Покажем теперь, что любой набор из l ($l > m$) элементов данной линейной оболочки будет линейно зависимым. Для этого выберем l элементов $y_1, y_2, \ldots, y_l$, принадлежащих линейной оболочке, и выразим их через элементы $x_1, x_2, \ldots, x_m$, получим

$$y_j = \sum_{i=1}^{m} \beta_{ji} x_i \qquad \forall j = [1, l]\,.$$

Приравняем нулевому элементу произвольную линейную комбинацию выбранного набора $y_1, y_2, \ldots, y_l$:

$$\sum_{j=1}^{l} \mu_j y_j = \sum_{j=1}^{l} \mu_j \left(\sum_{i=1}^{m} \beta_{ji} x_i \right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{l} \beta_{ji} y_j \right) x_i = o\,.$$

Поскольку элементы $x_1, x_2, \ldots, x_m$ линейно независимые, то коэффициенты $\mu_1, \mu_2, \ldots, \mu_m$ должны удовлетворять следующей однородной системе линейных уравнений:

$$\sum_{j=1}^{l} \beta_{ji} y_j = 0 \qquad \forall i = [1, m]\,.$$

Пусть ранг основной матрицы этой системы равен r. Поскольку $r \leq m$, то она имеет в силу теоремы 6.7.1 $l-r \geq l-m > 0$ линейно независимых и следовательно, ненулевых решений. Принимая во внимание, что l и m суть не равные друг другу натуральные числа, получаем $l-m \geq 1$, то есть существует нетривиальная линейная комбинация элементов $y_1, y_2, \ldots, y_l$, равная o.

Теорема доказана.

Гиперплоскость

Определение 7.3.4 Множество Γ, образованное из элементов вида $x+x_0$, где x_0 есть произвольный фиксированный элемент линейного пространства Λ, а x — любой элемент некоторого подпространства $\Omega \subseteq \Lambda$, называется *гиперплоскостью* (или *линейным многообразием*) в линейном пространстве Λ.

Замечание 7.3.1. 1°. В общем случае гиперплоскость не является подпространством.

2°. Если $\dim\Omega = k$, то говорят о k-мерной гиперплоскости.

Например, общее решение совместной *неоднородной* системы линейных уравнений с n неизвестными является гиперплоскостью в линейном пространстве n-компонентных столбцов.

Задача 7.3.1 *Показать, что если элементы x и y принадлежат гиперплоскости Γ, то ей будет принадлежать и элемент $z = \alpha x + (1-\alpha)y$, где α –– любое число.*

7.4. Координатное представление элементов линейного пространства

Определение 7.4.1 Пусть $\{ g_1, g_2, \ldots, g_n \}$ — базис в Λ^n. Тогда числа $\xi_1, \xi_2, \ldots, \xi_n$ в формуле $x = \sum\limits_{i=1}^{n} \xi_i g_i$ называются *координатами* элемента x в базисе $\{ g_1, g_2, \ldots, g_n \}$.

Напомним, что в силу теоремы 7.2.1 элемент x линейного пространства Λ^n в базисе $\{\, g_1,\, g_2,\, \ldots,\, g_n \,\}$ *однозначно* представляется n-компонентным столбцом

$$\|x\|_g = \|\, \xi_1,\, \xi_2,\, \ldots,\, \xi_n \,\|^{\mathrm{T}} \,,$$

называемым *координатным представлением* или *координатным столбцом* элемента в этом базисе.

В Λ^n базис может быть выбран *не единственным* способом, и потому прежде всего необходимо установить правило изменения координат элемента линейного пространства при переходе от одного базиса к другому.

Пусть в Λ^n даны два базиса: «старый» $\{\, g_1,\, g_2,\, \ldots,\, g_n \,\}$ и «новый» $\{\, g'_1,\, g'_2,\, \ldots,\, g'_n \,\}$ с соответствующими координатными разложениями

$$x = \sum_{i=1}^{n} \xi_i\, g_i \qquad \text{и} \qquad x = \sum_{i=1}^{n} \xi'_i\, g'_i$$

и координатными представлениями элемента x:

$$\|x\|_g = \|\, \xi_1,\, \xi_2,\, \ldots,\, \xi_n \,\|^{\mathrm{T}} \;\text{ и }\; \|x\|_{g'} = \|\, \xi'_1,\, \xi'_2,\, \ldots,\, \xi'_n \,\|^{\mathrm{T}} \,.$$

Пусть, кроме того, известны разложения элементов «нового» базиса по элементам «старого»:

$$g'_j = \sum_{i=1}^{n} \sigma_{ij} g_i \qquad \forall j = [1, n] \,. \tag{7.4.1}$$

Дадим

Определение 7.4.2

Матрица $\|S\| = \left\| \begin{matrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \ldots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \ldots & \sigma_{2n} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \ldots & \sigma_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \ldots & \sigma_{nn} \end{matrix} \right\|,$

j-й столбец которой состоит из коэффициентов координатного разложения j-го элемента «нового» базиса по элементам «старого», называется матрицей перехода от базиса $\{\, g_1,\, g_2,\, \ldots,\, g_n \,\}$ к базису $\{\, g'_1,\, g'_2,\, \ldots,\, g'_n \,\}$

Это определение является обобщением определения 1.8.2, и справедлива

Теорема 7.4.1 **Координаты $\xi_1, \xi_2, \ldots, \xi_n$ и $\xi'_1, \xi'_2, \ldots, \xi'_n$ связаны соотношениями** $\xi_i = \sum\limits_{j=1}^{n} \sigma_{ij}\xi'_j \quad \forall i = [1,n]$, **называемыми** *формулами перехода*, **где коэффициенты σ_{ij} – элементы матрицы перехода $\|S\|$.**

Доказательство.

В силу соотношений (7.4.1) будут справедливы равенства

$$\sum_{i=1}^{n} \xi_i\, g_i = x = \sum_{i=1}^{n} \xi'_i\, g'_i = \sum_{i=1}^{n} \xi'_i \left(\sum_{j=1}^{n} \sigma_{ji} g_j \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \sigma_{ji} \xi'_i \right) g_j .$$

Значение суммы не зависит от того, каким символом обозначается индекс суммирования. Поэтому если в самой правой части заменить i на j, а j — на i, то мы получим

$$\sum_{i=1}^{n} \xi_i\, g_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \sigma_{ij} \xi'_j \right) g_i \quad \Longrightarrow \quad \sum_{i=1}^{n} \left(-\xi_i + \sum_{j=1}^{n} \sigma_{ij} \xi'_j \right) g_i = o\,.$$

Но если линейная комбинация линейно независимых (в данном случае базисных) элементов равна нулевому элементу, то она тривиальная. Откуда получаем, что

$$\xi_i = \sum_{j=1}^{n} \sigma_{ij} \xi'_j \qquad \forall i = [1,n]\,.$$

Теорема доказана.

Заметим, что если столбец элементов «нового» базиса выражается через столбец элементов «старого» при помощи умножения последнего слева на транспонированную матрицу перехода $\|S\|^{\mathrm{T}}$, то координатный столбец в «старом» базисе равен произведению матрицы перехода на координатный столбец в «новом» базисе.

Действительно, рассматривая столбцы $\|x\|_g$ и $\|x\|_g$ в формулах перехода как двухиндексные матрицы, получаем $\xi_{i1} = \sum\limits_{j=1}^{n} \sigma_{ij}\xi'_{j1}$, что по правилам умножения матриц записывается как $\|x\|_g = \|S\| \|x\|_{g'}$ (см. § 5.1).

Используя теорему 5.1.1 (о транспонировании произведения матриц), равенство (7.4.1) $\left\| g_1', g_2', \ldots, g_n' \right\|^{\mathrm{T}} = \|S\|^{\mathrm{T}} \left\| g_1, g_2, \ldots, g_n \right\|^{\mathrm{T}}$ можно также записать в следующем матричном виде:

$$\left\| g_1', g_2', \ldots, g_n' \right\| = \left\| g_1, g_2, \ldots, g_n \right\| \|S\| .$$

В заключение выясним, как операции с элементами линейного пространства выполняются в координатной форме.

Пусть в конкретном базисе в Λ^n имеем $x = \sum\limits_{i=1}^{n} \xi_i g_i$ и $y = \sum\limits_{i=1}^{n} \eta_i g_i$, тогда в силу определения базиса и аксиом линейного пространства справедливо следующее:

1°. Для *критерия сравнения*: два элемента в Λ^n равны (то есть $x = y$) тогда и только тогда, когда $\|x\|_g = \|y\|_g$.

2°. Для операции *сложения*: $\|x + y\|_g = \|x\|_g + \|y\|_g$.

3°. Для операции *умножения числа на элемент*: $\|\lambda x\|_g = \lambda \|x\|_g$.

Откуда следует, что элементы конечномерного линейного пространства не только могут представляться матрицами (столбцами), но и правила выполнения операций с этими элементами в координатах *совпадают с определением соответствующих матричных операций.*

Наконец, заключение о линейной зависимости или независимости некоторого набора элементов в Λ^n можно делать, применяя теорему 6.5.3 (о ранге матрицы) к матрице, столбцы которой суть координатные представления элементов этого набора.

7.5. Изоморфизм линейных пространств

Рассмотрим два линейных пространства: множество многочленов $P_2(\tau)$ степени не выше, чем 2, и множество векторов трехмерного геометрического пространства.

Операции сложения многочленов и их умножения числа на многочлен выглядят следующим образом:

$$\left(\xi_1 + \xi_2\tau + \xi_3\tau^2\right) + \left(\eta_1 + \eta_2\tau + \eta_3\tau^2\right) = (\xi_1 + \eta_1) + (\xi_2 + \eta_2)\,\tau + (\xi_3 + \eta_3)\,\tau^2,$$
$$\lambda\left(\xi_1 + \xi_2\tau + \xi_3\tau^2\right) = (\lambda\xi_1) + (\lambda\xi_2)\,\tau + (\lambda\xi_3)\,\tau^2.$$

Те же операции с трехмерными векторами в координатной форме в свою очередь записываются так:

$$\left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{matrix} \right\| + \left\| \begin{matrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{matrix} \right\| = \left\| \begin{matrix} \xi_1 + \eta_1 \\ \xi_2 + \eta_2 \\ \xi_3 + \eta_3 \end{matrix} \right\| , \qquad \lambda \left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{matrix} \right\| = \left\| \begin{matrix} \lambda\xi_1 \\ \lambda\xi_2 \\ \lambda\xi_3 \end{matrix} \right\| .$$

Сопоставляя эти записи, можно заключить, что природа данных множеств не играет роли, когда исследуются их характеристики, использующие только критерии равенства, операции сложения и умножения числа на элемент.

Отмеченное свойство линейных пространств носит название *изоморфизма.* Более точно его описывает

Определение 7.5.1	Два линейных пространства Λ_1 и Λ_2 называются *изоморфными*, если существует взаимно однозначное отображение $\hat{F}\,:\;\Lambda_1 \to \Lambda_2$, такое, что $\forall\lambda \in \mathbb{C}$ и $\forall x, y \in \Lambda_1$: $1^\circ.\ \hat{F}(x+y) = \hat{F}(x) + \hat{F}(y)$; $2^\circ.\ \hat{F}(\lambda x) = \lambda\hat{F}(x)$. Отображение $\hat{F}(x)$ называется *изоморфизмом* линейных пространств Λ_1 и Λ_2.

Напомним, что отображение $\hat{F}$ является взаимно однозначным (биективным), если разные элементы из Λ_1 имеют в Λ_2 разные образы (инъективность), а каждый элемент из Λ_2 является образом некоторого элемента из Λ_1 (сюръективность).

Теорема 7.5.1 (об изоморфизме) **Два линейных конечномерных пространства Λ_1 и Λ_2 изоморфны тогда и только тогда, когда их размерности равны.**

Доказательство.

1°. Пусть $\dim \Lambda_1 = \dim \Lambda_2$. Используя в качестве изоморфизма отображение, при котором каждому элементу из Λ_1 ставится в соответствие элемент из Λ_2, имеющий те же самые координаты, и используя правила операций с элементами в координатном представлении, приходим к заключению об изоморфности линейных пространств Λ_1 и Λ_2.

2°. Допустим теперь, что $n = \dim \Lambda_1 > \dim \Lambda_2 = m$, а пространства Λ_1 и Λ_2. изоморфны. Возьмем в Λ_1 некоторую линейную комбинацию n линейно независимых элементов, равную нулевому элементу. Эта линейная комбинация обязана быть тривиальной.

В пространстве Λ_2 эта же линейная комбинация образов выбранных элементов будет также равняться нулевому элементу, поскольку в силу определения 7.5.1 нулевой элемент переходит в нулевой элемент.

При этом образы выбранных элементов обязаны быть в Λ_2 линейно зависимыми (поскольку мы предположили, что $n > m$), и, следовательно, рассматриваемая линейная комбинация может быть нетривиальной. Это противоречит предположению о том, что $n > m$.

Аналогичные рассуждения в предположении, что $n < m$, также приводят к противоречию, и, следовательно, $n = m$.

Теорема доказана.

Пример 7.5.1 Изоморфизм одномерных пространств вещественных чисел $x \in \mathbb{R}$ и всех положительных чисел $y \in \mathbb{R}^+$ (с операциями, определенными в условии задачи 7.1.3) задается при помощи функций $y = e^x$ и $x = \ln y$.

Очевидным следствием теоремы 7.5.1 является изоморфизм любого линейного n-мерного пространства Λ^n и линейного пространства n-компонентных столбцов, позволяющий убедиться в наличии свойств элементов Λ^n, аналогичным свойствам столбцов, установленных в § 6.5–6.7.

Например, имеет место

Теорема 7.5.2 **Максимальное число линейно независимых элементов в любом конечном наборе элементов из Λ^n равно рангу матрицы, столбцы которой суть координатные представления элементов данного набора в некотором базисе.**

Доказательство.

Следует из изоморфности линейного пространства Λ^n и линейного пространства всех n-компонентных столбцов, а также из теоремы 6.5.3 (о ранге матрицы).

Теорема доказана.

Следствие 7.5.1 **Множество, состоящее из k элементов в Λ^n, линейно зависимо тогда и только тогда, когда ранг матрицы, столбцы которой суть координатные представления этих элементов в некотором базисе, меньше, чем $\min\{k, n\}$.**

Следствие 7.5.2 **Матрица перехода от некоторого базиса к другому невырожденная, то есть $\det\|S\| \neq 0$.**

Доказательство.

Предположим противное: $\det\|S\| = 0$, при этом $\mathrm{rg}\|S\| < n$ и столбцы матрицы перехода линейно зависимые.
Но тогда будут линейно зависимыми также и элементы $g'_1, g'_2, \ldots, g'_n$, а это противоречит предположению о том, что $\{ g'_1, g'_2, \ldots, g'_n \}$ — базис.

Следствие доказано.

Отметим также, что факт равенства или неравенства двух элементов в координатной форме можно проверять в любом базисе, поскольку в силу невырожденности матрицы $\|S\|$ и формул перехода оказываются справедливыми соотношения

$$\|x\|_g = \|y\|_g \quad \Longleftrightarrow \quad \|S\|\,\|x\|_{g'} = \|S\|\,\|y\|_{g'} \quad \Longleftrightarrow \quad \|x\|_{g'} = \|y\|_{g'} .$$

Следствие 7.5.3 **Существует матрица $\|T\| = \|S\|^{-1}$ перехода от базиса $\{ g'_1, g'_2, \ldots, g'_n \}$ к базису $\{ g_1, g_2, \ldots, g_n \}$, называемая** *матрицей обратного перехода*.

Для матрицы обратного перехода справедливы (проверьте это самостоятельно) соотношения

$$\left\| \begin{matrix} g_1 \\ g_2 \\ \ldots \\ g_n \end{matrix} \right\| = \|T\|^{\mathrm{T}} \left\| \begin{matrix} g'_1 \\ g'_2 \\ \ldots \\ g'_n \end{matrix} \right\| , \qquad \|x\|_{g'} = \|T\|\,\|x\|_g , \qquad \det\|T\| \cdot \det\|S\| = 1 .$$

Пусть в Λ^n задан базис $\{ g_1, g_2, \ldots, g_n \}$, в котором координатное разложение элементов имеет вид $x = \sum\limits_{i=1}^{n} \xi_i g_i$. Тогда имеет место

Следствие 7.5.4 **Каждая** *однородная* **система m линейных уравнений с n неизвестными**

$$\sum_{i=1}^{n} \alpha_{ji}\xi_i \quad \forall j = [1, m]$$

определяет некоторое подпространство Ω в Λ^n.

Доказательство.

Следует из того факта, что подпространство Ω в силу теоремы 6.7.2 является линейной оболочкой нормальной фундаментальной системы решений данной системы линейных уравнений, а Λ^n изоморфно линейному пространству n-компонентных столбцов $\| \xi_1 \, \xi_2 \, \ldots \, \xi_n \|^{\mathrm{T}}$.

Следствие доказано.

Таким образом, каждое подпространство в Λ^n может быть задано либо однородной системой линейных уравнений, либо как линейная оболочка фундаментальной системы ее решений.

Также справедливо

Следствие 7.5.5 **Каждая совместная** *неоднородная* **система m линейных уравнений с n неизвестными**

$$\sum_{i=1}^{n} \alpha_{ji}\xi_i = \beta_j \quad \forall j = [1, m]$$

определяет некоторую гиперплоскость Γ в Λ^n.

Доказательство.

Аналогично доказательству следствия 7.5.4.

Следствие доказано.

Задача 7.5.1 *Проверить, что набор элементов $\{ g_1, g_2, g_3 \}$ является базисом в Λ^3, и найти координатное представление элемента x в этом базисе, если в некотором исходном базисе*

$$\|g_1\| = \left\| \begin{matrix} 1 \\ 1 \\ 1 \end{matrix} \right\|, \ \|g_2\| = \left\| \begin{matrix} 2 \\ 1 \\ 0 \end{matrix} \right\|, \ \|g_3\| = \left\| \begin{matrix} 3 \\ 0 \\ 1 \end{matrix} \right\| \ \text{и}$$

$$\|x\| = \left\| \begin{matrix} 1 \\ 3 \\ 1 \end{matrix} \right\|.$$

Решение. 1°. Для того чтобы из элементов g_1, g_2, g_3 можно было образовать в Λ^3 базис, необходимо и достаточно (определение 7.2.2), чтобы эти элементы были линейно независимыми. Например, по следствию 7.5.1, данное условие равносильно

$$\operatorname{rg}\left\|\begin{array}{ccc} 1 & 2 & 3 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{array}\right\| = 3 \quad \Longleftrightarrow$$

$$\Longleftrightarrow \quad \det\left\|\begin{array}{ccc} 1 & 2 & 3 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{array}\right\| = -4 \neq 0\,,$$

что верно. Следовательно, элементы g_1, g_2, g_3 образуют базис.

2°. Обозначим искомые координаты элемента x как ξ_1, ξ_2, ξ_3. Тогда $x = \xi_1 g_1 + \xi_2 g_2 + \xi_3 g_3$, или в координатной форме:

$$\left\|\begin{array}{c} 1 \\ 3 \\ 1 \end{array}\right\| = \xi_1 \left\|\begin{array}{c} 1 \\ 1 \\ 1 \end{array}\right\| + \xi_2 \left\|\begin{array}{c} 2 \\ 1 \\ 0 \end{array}\right\| + \xi_3 \left\|\begin{array}{c} 3 \\ 0 \\ 1 \end{array}\right\|.$$

3°. Использовав критерий равенства двух элементов и правила выполнения операций сложения и умножения на число в координатной форме, получим для неизвестных ξ_1, ξ_2, ξ_3 систему линейных уравнений

$$\left\{\begin{array}{ccccccc} \xi_1 & + & 2\xi_2 & + & 3\xi_3 & = & 1, \\ \xi_1 & + & \xi_2 & & & = & 3, \\ \xi_1 & & & + & \xi_3 & = & 1, \end{array}\right.$$

решив которую (например, по правилу Крамера — теорема 6.4.1 или методом Гаусса — § 6.8), получим $\xi_1 = 2, \xi_2 = 1, \xi_3 = -1$. Откуда следует, что элемент в базисе $\{g_1, g_2, g_3\}$ имеет координатное представление

Решение получено.

$$\|x\|_g = \left\|\begin{array}{r} 2 \\ 1 \\ -1 \end{array}\right\|.$$

Задача 7.5.2 *Найти в Λ_3 матрицу перехода от базиса, образованного элементами $\{g'_1, g'_2, g'_3\}$, к базису $\{g''_1, g''_2, g''_3\}$, если в некотором исходном базисе $\{g_1, g_2, g_3\}$:*

$$\|g'_1\|_g = \left\|\begin{matrix} 1 \\ 1 \\ 1 \end{matrix}\right\|, \quad \|g'_2\|_g = \left\|\begin{matrix} 2 \\ 1 \\ 0 \end{matrix}\right\|, \quad \|g'_3\|_g = \left\|\begin{matrix} 3 \\ 0 \\ 1 \end{matrix}\right\|,$$

$$\|g''_1\|_g = \left\|\begin{matrix} 7 \\ 3 \\ 3 \end{matrix}\right\|, \quad \|g''_2\|_g = \left\|\begin{matrix} 16 \\ 5 \\ 6 \end{matrix}\right\|, \quad \|g''_3\|_g = \left\|\begin{matrix} 22 \\ 7 \\ 8 \end{matrix}\right\|.$$

Решение. 1°. Пусть $\|x\|_g$, $\|x\|_{g'}$ и $\|x\|_{g''}$ обозначают координатные столбцы элемента x в трех базисах: исходном, $\{g'_1, g'_2, g'_3\}$ и $\{g''_1, g''_2, g''_3\}$ соответственно. Тогда (по определению 7.4.2 и в силу теоремы 7.4.1) имеют место равенства

$$\|x\|_g = \|G\|\|x\|_{g'} \qquad \text{и} \qquad \|x\|_g = \|F\|\|x\|_{g''}\,, \quad (7.5.1)$$

где матрицы $\|G\|$ и $\|F\|$ составлены из координатных столбцов базисных элементов $\{g'_1, g'_2, g'_3\}$ и $\{g''_1, g''_2, g''_3\}$, то есть

$$\|G\| = \left\|\begin{matrix} 1 & 2 & 3 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{matrix}\right\| \qquad \text{и} \qquad \|F\| = \left\|\begin{matrix} 7 & 16 & 22 \\ 3 & 5 & 7 \\ 3 & 6 & 8 \end{matrix}\right\|.$$

2°. Обозначим через $\|S\|$ искомую матрицу перехода от базиса $\{g'_1, g'_2, g'_3\}$ к базису $\{g''_1, g''_2, g''_3\}$, для которой $\|x\|_{g'} = \|S\|\|x\|_{g''}$. Но из условий (7.5.1) также следует, что

$$\|x\|_{g'} = \|G\|^{-1}\|F\ \|x\|_{g''}\,,$$

поскольку матрица $\|G\|$ (как матрица перехода) невырожденная. Тогда $\|S\|\|x\|_{g''} = \|G\|^{-1}\|F\|\|x\|_{g''}$ для любого элемента $\|x\|_{g''}$, а это в силу леммы 5.1.2 означает, что искомая матрица перехода

$$\|S\| = \|G\|^{-1}\|F\|\,.$$

3°. Подсчитав произведение

$$\|G\|^{-1}\|F\| = \left\|\begin{array}{ccc} 1 & 2 & 3 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{array}\right\|^{-1} \left\|\begin{array}{ccc} 7 & 16 & 22 \\ 3 & 5 & 7 \\ 3 & 6 & 8 \end{array}\right\| ,$$

используя, например, схему, описанную в § 6.8, для вычисления произведений вида $\|A\|^{-1}\|B\|$, получаем

$$\|S\| = \left\|\begin{array}{ccc} 2 & 3 & 4 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{array}\right\| .$$

Решение получено.

Задача 7.5.3 *В линейном пространстве многочленов степени не выше, чем 3, найти размерность и базис пересечения линейных оболочек следующих двух наборов элементов*:

$$\begin{array}{rcrcrcrcr} x_1(\tau) & = & 1 & + & 2\tau & + & \tau^2 & + & 3\tau^3 \, , \\ x_2(\tau) & = & -1 & + & 8\tau & - & 6\tau^2 & + & 5\tau^3 \, , \\ x_3(\tau) & = & & & 10\tau & - & 5\tau^2 & + & 8\tau^3 \end{array}$$

и

$$\begin{array}{rcrcrcrcr} y_1(\tau) & = & 1 & + & 4\tau & - & \tau^2 & + & 5\tau^3 \, , \\ y_2(\tau) & = & 3 & - & 2\tau & + & 6\tau^2 & + & 3\tau^3 \, , \\ y_3(\tau) & = & 4 & + & 2\tau & + & 5\tau^2 & + & 8\tau^3 \, . \end{array}$$

Решение. 1°. Линейное пространство алгебраических многочленов степени не выше, чем 3, четырехмерное. При этом в базисе

$$\{ \, g_1(\tau) = 1, \quad g_2(\tau) = \tau, \quad g_3(\tau) = \tau^2, \quad g_4(\tau) = \tau^3 \, \}$$

элемент $x(\tau) = \xi_1 g_1(\tau) + \xi_2 g_2(\tau) + \xi_3 g_3(\tau) + \xi_4 g_4(\tau)$ будет иметь координатное представление (столбец) вида

$$\|x\|_g = \left\|\begin{array}{c} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{array}\right\| .$$

Линейное пространство алгебраических многочленов степени не выше, чем 3, по теореме 7.5.1 изоморфно линейному пространству четырехкомпонентных столбцов, в котором мы и будем решать задачу.

2°. По теореме 7.4.1 каждая из линейных оболочек является подпространством. Первое из них Π_1 образовано элементами вида $x(\tau) = \lambda_1 x_1(\tau) + \lambda_2 x_2(\tau) + \lambda_3 x_3(\tau)$, а второе подпространство Π_2 соответственно состоит из элементов $y(\tau) = \mu_1 y_1(\tau) + \mu_2 y_2(\tau) + \mu_3 y_3(\tau)$.

Координатная форма описания, например, Π_1, согласно условию задачи, будет $\forall \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$:

$$\left\| \begin{array}{c} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{array} \right\| = \lambda_1 \left\| \begin{array}{r} 1 \\ 2 \\ 1 \\ 3 \end{array} \right\| + \lambda_2 \left\| \begin{array}{r} -1 \\ 8 \\ -6 \\ 5 \end{array} \right\| + \lambda_3 \left\| \begin{array}{r} 0 \\ 10 \\ -5 \\ 8 \end{array} \right\| . \quad (7.5.2)$$

В координатном виде подпространства (а значит, и линейные оболочки) можно также задавать *однородными системами линейных уравнений.* Такой способ описания подпространств в данной задаче более удобен, поскольку описание пересечения оболочек получается объединением систем, задающих эти оболочки, в одну систему.

3°. Составим однородные системы линейных уравнений, задающих подпространства Π_1 и Π_1[12].

Пусть каждое из уравнений этих систем имеет вид

$$\| \, \alpha_1 \; \alpha_2 \; \alpha_3 \; \alpha_4 \, \| \left\| \begin{array}{c} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{array} \right\| = 0 \, .$$

Тогда, используя (7.5.2), приходим к условию для Π_1:

$$\| \, \alpha_1 \; \alpha_2 \; \alpha_3 \; \alpha_4 \, \| \left(\lambda_1 \left\| \begin{array}{r} 1 \\ 2 \\ 1 \\ 3 \end{array} \right\| + \lambda_2 \left\| \begin{array}{r} -1 \\ 8 \\ -6 \\ 5 \end{array} \right\| + \lambda_3 \left\| \begin{array}{r} 0 \\ 10 \\ -5 \\ \xi_8 \end{array} \right\| \right) = 0 \, ,$$

[12] Для этой цели возможно также использовать теорему 6.6.1 (Кронекера – Капелли). См., например, решение задачи 8.4.1 (пункт 2°).

которое будет выполняться $\forall\lambda_k,\ k=1,2,3$, если числа $\alpha_i,\ i=1,2,3,4$, суть решение следующей системы линейных уравнений:

$$\left\{\begin{array}{rcrcrcrcl} \alpha_1 & + & 2\alpha_2 & + & \alpha_3 & + & 3\alpha_4 & = & 0, \\ -\alpha_1 & + & 8\alpha_2 & - & 6\alpha_3 & + & 5\alpha_4 & = & 0, \\ & & 10\alpha_2 & - & 5\alpha_3 & + & 8\alpha_4 & = & 0. \end{array}\right.$$

Решив эту систему, например, по схеме, описанной в § 6.8, получим общее решение в виде

$$\left\|\begin{array}{c} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{array}\right\| = \kappa_1 \left\|\begin{array}{r} -4 \\ 1 \\ 2 \\ 0 \end{array}\right\| + \kappa_2 \left\|\begin{array}{r} -7 \\ -4 \\ 0 \\ 5 \end{array}\right\| \qquad \forall\kappa_1,\kappa_2\,.$$

Откуда заключаем, что существует лишь два независимых набора искомых чисел α_1, α_2, α_3, α_4, с помощью которых однородная система линейных уравнений, задающая подпространство Π_1, может быть записана как

$$\left\{\begin{array}{rcrcrcrcl} -4\xi_1 & + & \xi_2 & + & 2\xi_3 & & & = & 0, \\ -7\xi_1 & - & 4\xi_2 & & & + & 5\xi_4 & = & 0. \end{array}\right.$$

Аналогично строим однородную систему линейных уравнений, задающую Π_2:

$$\left\{\begin{array}{rcrcrcrcl} -22\xi_1 & + & 9\xi_2 & + & 14\xi_3 & & & = & 0, \\ -11\xi_1 & - & 6\xi_2 & & & + & 7\xi_4 & = & 0. \end{array}\right.$$

4°. Наконец, пересечение подпространств Π_1 и Π_2 будет задаваться системой, получаемой объединением найденных систем, в одну

$$\left\{\begin{array}{rcrcrcrcl} -4\xi_1 & + & \xi_2 & + & 2\xi_3 & & & = & 0, \\ -7\xi_1 & - & 4\xi_2 & & & + & 5\xi_4 & = & 0, \\ -22\xi_1 & + & 9\xi_2 & + & 14\xi_3 & & & = & 0, \\ -11\xi_1 & - & 6\xi_2 & & & + & 7\xi_4 & = & 0, \end{array}\right.$$

общее решение которой есть

$$\left\| \begin{array}{c} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{array} \right\| = \sigma \left\| \begin{array}{r} 2 \\ -6 \\ 7 \\ -2 \end{array} \right\| \qquad \forall \sigma \in \mathbb{R}\,.$$

В итоге, используя изоморфизм «в обратную сторону», получаем, что $\dim(\Pi_1 \cap \Pi_2) = 1$, а базис пересечения линейных оболочек двух наборов многочленов, заданных в условии задачи, состоит из одного элемента, например,

`Решение получено.`

$$b(\tau) = 2 - 6\tau + 7\tau^2 - 2\tau^3.$$

8. Линейные зависимости в линейном пространстве

8.1. Линейные операторы

Определение 8.1.1	Пусть *каждому* элементу x линейного пространства Λ поставлен в соответствие *единственный* элемент y линейного пространства Λ^*. Тогда говорят, что задан *оператор* $\hat{A}$, действующий в Λ и имеющий значения в Λ^*, действие которого обозначается как $y = \hat{A}x$ или $y = \hat{A}(x)$. При этом элемент y называется *образом* элемента x, а элемент x — *прообразом* элемента y.

Как и в § 5.2, операторы подразделяются на *отображения*, если $\Lambda^* \not\subseteq \Lambda$, и *преобразования*, если $\Lambda^* \subseteq \Lambda$. В дальнейшем, за исключением особо оговоренных случаев, будем предполагать, что из контекста ясно, идет ли речь об отображении или о преобразовании.

Определение 8.1.2	Оператор называется *линейным*, если для любых $x, y \in \Lambda$ и любого числа λ имеют место равенства $1^\circ.\ \hat{A}(x+y) = \hat{A}x + \hat{A}y$, $2^\circ.\ \hat{A}(\lambda x) = \lambda \hat{A}x$.

Пример 8.1.1

1°. В пространстве двумерных столбцов линейным оператором является правило

$$\left\| \begin{matrix} \eta_1 \\ \eta_2 \end{matrix} \right\| = \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\| \left\| \begin{matrix} \xi_1 \\ \xi_2 \end{matrix} \right\| ,$$

связывающее столбец-прообраз $x = \left\| \begin{matrix} \xi_1 \\ \xi_2 \end{matrix} \right\|$ со столбцом-образом $y = \left\| \begin{matrix} \eta_1 \\ \eta_2 \end{matrix} \right\|$.

2°. В пространстве бесконечно дифференцируемых функций линейным оператором является операция дифференцирования, ставящая в соответствие каждому элементу этого пространства его производную функцию.

3°. В пространстве непрерывных функций $f(\tau)$ линейным оператором является операция умножения такой функции на независимую переменную τ.

Задача 8.1.1

Доказать, что операторы в примерах 1°, 2° и 3° являются линейными.

Задача 8.1.2

Является ли линейным оператор $\hat{A}$, ставящий каждому элементу x в соответствие фиксированный элемент a?

Решение. Если элемент $a = o$, то $\hat{A}$ — линейный оператор.

Действительно, если оператор линейный, то, с одной стороны,

$$\hat{A}(a + a) = \hat{A}a + \hat{A}a = 2\hat{A}a = 2a ,$$

но, с другой стороны, $\hat{A}(a + a) = a$. Значит,

$$2a = a \quad \Longrightarrow \quad a = o.$$

Решение получено.

8.2. Действия с линейными операторами

Определение 8.2.1

Линейные операторы $\hat{A}$ и $\hat{B}$ называются *равными* (что обозначается как $\hat{A} = \hat{B}$), если

$$\forall x \in \Lambda : \qquad \hat{A}x = \hat{B}x .$$

Суммой линейных операторов и $\hat{A}$ и $\hat{B}$ называется оператор $\hat{C}$ (что символически обозначается равенством $\hat{C} = \hat{A} + \hat{B}$), ставящий каждому элементу x линейного пространства Λ в соответствие элемент $\hat{A}x + \hat{B}x$.

Лемма 8.2.1 **Сумма двух линейных операторов является линейным оператором.**

Доказательство.

Пусть $x, y \in \Lambda$ и λ, μ суть числа, а $\hat{C} = \hat{A} + \hat{B}$, тогда, в силу определений 8.1.2 и 8.2.1,

$$\begin{aligned}
\hat{C}\big(\lambda x + \mu y\big) &= \hat{A}\big(\lambda x + \mu y\big) + \hat{B}\big(\lambda x + \mu y\big) = \\
&= \lambda \hat{A}x + \mu \hat{A}y + \lambda \hat{B}x + \mu \hat{B}y = \\
&= \lambda \hat{A}x + \lambda \hat{B}x + \mu \hat{A}y + \mu \hat{B}y = \\
&= \lambda\big(\hat{A}x + \hat{B}x\big) + \mu\big(\hat{A}y + \hat{B}y\big) = \\
&= \lambda \hat{C}x + \mu \hat{C}y\,.
\end{aligned}$$

Сравнивая начало и конец полученной цепочки равенств, приходим к заключению о линейности оператора $\hat{C}$.

Лемма доказана.

Определение 8.2.2 *Нулевым* оператором $\hat{O}$ называется оператор, ставящий каждому элементу x линейного пространства Λ в соответствие нулевой элемент этого пространства.

Определение 8.2.3 *Оператором, противоположным оператору* $\hat{A}$, называется оператор, обозначаемый $\tilde{\hat{A}}$, ставящий каждому элементу x пространства Λ в соответствие элемент $\tilde{x}$ (см. определение 7.1.1).

Из решения задачи 8.1.2 следует, что нулевой оператор линейный. Покажите самостоятельно, что оператор, противоположный любому линейному оператору, также линейный.

Лемма 8.2.2 **Для любых линейных операторов $\hat{A}$, $\hat{B}$ и $\hat{C}$ выполняются соотношения**

$$\begin{aligned}
&\hat{A}+\hat{B}=\hat{B}+\hat{A}\,,\\
&\big(\hat{A}+\hat{B}\big)+\hat{C}=\hat{A}+\big(\hat{B}+\hat{C}\big)\,,\\
&\hat{A}+\hat{O}=\hat{A}\,,\quad \hat{A}+\tilde{\hat{A}}=\hat{O}\,.
\end{aligned}$$

Доказательство.

Справедливость утверждения леммы непосредственно вытекает из определений 8.2.1, 8.2.3 и аксиоматики линейного пространства.

Лемма доказана.

Определение 8.2.4 *Произведением числа λ на линейный оператор $\hat{A}$* называется оператор (обозначаемый $\lambda\hat{A}$), ставящий каждому элементу x линейного пространства Λ в соответствие элемент $\lambda\big(\hat{A}x\big)$.

Лемма 8.2.3 **Для произведения числа на линейный оператор справедливы соотношения**

$$\begin{aligned}
&\big(\alpha\beta\big)\hat{A}=\alpha\big(\beta\big)\hat{A}\,,\quad 1\hat{A}=\hat{A}\,,\\
&\big(\alpha+\beta\big)\hat{A}=\alpha\hat{A}+\beta\hat{A}\,,\\
&\alpha\big(\hat{A}+\hat{B}\big)=\alpha\hat{A}+\alpha\hat{B}\,.
\end{aligned}$$

Доказательство.

Утверждение леммы проверяется непосредственно. Например, для третьего равенства имеем

$$\forall x\in\Lambda\quad\text{и}\quad\forall\alpha,\beta\in\mathbb{R}\ :$$

$$\hat{A}\big(\big(\alpha+\beta\big)x\big)=\hat{A}\big(\alpha x+\beta x\big)=\big(\alpha+\beta\big)\hat{A}x=\alpha\hat{A}x+\beta\hat{A}x\,.$$

Лемма доказана.

Теорема 8.2.1 **Множество** *всех* **линейных операторов, действующих в линейном пространстве Λ, является** *линейным пространством***.**

Доказательство.

Следует из определений 7.1.1, 8.2.1–8.2.4 и лемм 8.2.1, 8.2.2.

Теорема доказана.

Определение 8.2.5 *Произведением* (иногда *композицией* или *суперпозицией*) линейных операторов $\hat{A}$ и $\hat{B}$ называется оператор (обозначаемый как $\hat{A}\hat{B}$), ставящий каждому элементу x линейного пространства Λ в соответствие элемент $\hat{A}\big(\hat{B}x\big)$.

Теорема 8.2.2 **Произведение линейных операторов является** *линейным* **оператором, для которого справедливы соотношения**

$$\big(\hat{A}\hat{B}\big)\hat{C} = \hat{A}\big(\hat{B}\hat{C}\big), \quad \big(\hat{A}+\hat{B}\big)\hat{C} = \hat{A}\hat{C} + \hat{B}\hat{C},$$
$$\hat{A}\big(\hat{B}+\hat{C}\big) = \hat{A}\hat{B} + \hat{A}\hat{C}.$$

Доказательство.

Докажем вначале линейность произведения линейных операторов. Действительно, $\forall x, y \in \Lambda$ и любых чисел α, β:

$$\hat{A}\hat{B}\big(\alpha x + \beta y\big) = \hat{A}\Big(\hat{B}\big(\alpha x + \beta y\big)\Big) = \hat{A}\Big(\alpha\hat{B}x + \beta\hat{B}y\Big) = \\ = \alpha\hat{A}\big(\hat{B}x\big) + \beta\hat{A}\big(\hat{B}y\big) = \alpha\big(\hat{A}\hat{B}\big)x + \beta\big(\hat{A}\hat{B}\big)y.$$

Проверим теперь сочетательный закон для произведения линейных операторов. Имеем

$$\big(\hat{A}\big(\hat{B}\hat{C}\big)\big)x = \hat{A}\big(\hat{B}\hat{C}x\big) = \hat{A}\big(\hat{B}\big(\hat{C}x\big)\big),$$

но, с другой стороны,

$$\big(\big(\hat{A}\hat{B}\big)\hat{C}\big)x = \big(\hat{A}\hat{B}\big)\hat{C}x = \hat{A}\big(\hat{B}\big(\hat{C}x\big)\big),$$

что и требовалось показать. Остальные утверждения теоремы проверяются аналогично.

Теорема доказана.

Замечание 8.2.1. В общем случае произведение линейных операторов не обладает перестановочным свойством (или, иначе говоря, операторы *не коммутируют*), то есть $\hat{A}\hat{B} \neq \hat{B}\hat{A}$.

Определение 8.2.6	Оператор $\hat{A}\hat{B}-\hat{B}\hat{A}$ называется *коммутатором* операторов $\hat{A}$ и $\hat{B}$.

Коммутатор коммутирующих операторов есть нулевой оператор.

Задача 8.2.1 *В линейном пространстве алгебраических многочленов вида $P_n(\tau) = \sum\limits_{k=0}^{n} \alpha_k \tau^k$ найти коммутатор для операторов:*

- *$\hat{A}$, ставящего в соответствие многочлену его производную функцию, и*
- *$\hat{B}$ –– оператора умножения независимой переменной на многочлен.*

Решение. Построим оператор $\hat{A}\hat{B}-\hat{B}\hat{A}$. Для любого $P_n(\tau)$ имеем

$$\hat{A}P_n(\tau) = \frac{d}{d\tau}P_n(\tau) = \frac{d}{d\tau}\left(\sum_{k=0}^{n} \alpha_k \tau^k\right) = \sum_{k=1}^{n} k\alpha_k \tau^{k-1} ,$$

$$\hat{B}P_n(\tau) = \tau P_n(\tau) = \sum_{k=0}^{n} \alpha_k \tau^{k+1} .$$

Откуда получаем

$$\hat{A}\hat{B}\,P_n(\tau) = \hat{A}\big(\hat{B}\,P_n(\tau)\big) = \frac{d}{d\tau}\sum_{k=0}^{n} \alpha_k \tau^{k+1} = \sum_{k=0}^{n}(k+1)\alpha_k \tau^k ,$$

$$\hat{B}\hat{A}\,P_n(\tau) = \hat{B}\big(\hat{A}\,P_n(\tau)\big) = \tau\sum_{k=1}^{n} k\alpha_k \tau^{k-1} = \sum_{k=0}^{n} k\alpha_k \tau^k ,$$

$$\big(\hat{A}\hat{B}-\hat{B}\hat{A}\big)P_n(\tau) = \sum_{k=0}^{n}(k+1)\alpha_k \tau^k - \sum_{k=0}^{n} k\alpha_k \tau^k = \sum_{k=0}^{n} \alpha_k \tau^k = \\ = P_n(\tau),$$

Решение получено. и, следовательно, линейные операторы $\hat{A}$ и $\hat{B}$ не коммутируют.

В рассмотренной выше задаче 8.2.1 оказалось, что действие коммутатора $\hat{A}\hat{B} - \hat{B}\hat{A}$ на любой элемент линейного пространства многочленов не приводит к изменению этого элемента. Для операторов, обладающих таким свойством, используют специальное наименование.

Определение 8.2.7

Оператор $\hat{E}$ называется *единичным* (или *тождественным*) оператором, если *каждому* элементу x линейного пространства Λ он ставит в соответствие *тот же самый* элемент, то есть

$$\hat{E}x = x \qquad \forall x \in \Lambda\,.$$

Докажите самостоятельно, что $\forall \hat{A}$: $\hat{A}\hat{E} = \hat{E}\hat{A} = \hat{A}$, а также линейность и единственность $\hat{E}$.

Определение 8.2.8

Оператор $\hat{B}$ называется *обратным* для линейного оператора $\hat{A}$ (обозначается $\hat{A}^{-1}$), если

$$\hat{A}\hat{B} = \hat{B}\hat{A} = \hat{E}\,.$$

Пример 8.2.1

В линейном пространстве функций $f(\tau)$, имеющих на $[\alpha, \beta]$ производную любого порядка и удовлетворяющих условиям $f^{(k)}(\alpha) = 0 \quad \forall k = 0, 1, 2, \ldots$, оператор дифференцирования $\hat{A}f = \dfrac{df}{d\tau}$ и оператор интегрирования с переменным верхним пределом $\hat{B}f = \int\limits_{\alpha}^{\tau} f(u)\,du$ являются взаимно обратными.

Действительно,

$$\hat{A}\hat{B}\,f = \frac{d}{d\tau}\int\limits_{\alpha}^{\tau} f(u)\,du = f(\tau) = \hat{E}f \qquad \text{и}$$

$$\hat{B}\hat{A}\,f = \int\limits_{\alpha}^{\tau} \frac{df}{d\tau}du = F(\tau) - f(\alpha) = f(\tau) = \hat{E}f\,.$$

Замечание 8.2.2. 1°. Не для всякого линейного оператора существует обратный оператор. Например, нулевой оператор $\hat{O}$ не имеет обратного. Действительно, пусть $\hat{O}x = o \quad \forall x \in \Lambda$, тогда для любого $\hat{A}$ имеет место $\hat{A}\hat{O}\,x = \hat{A}\big(\hat{O}x\big) = o$, и, следовательно, равенство $\hat{A}\hat{O} = \hat{E}$ не выполняется ни при каком $\hat{A}$.

2°. Обратный оператор, если существует, то он единственный. (Покажите это самостоятельно, использовав идею доказательства леммы 5.1.1.)

3°. Из условия $\hat{A}\hat{B} = \hat{E}$ может не следовать выполнение равенства $\hat{B}\hat{A} = \hat{E}$. Это имеет место, например, в пространстве многочленов вида $P_n(\tau) = \sum\limits_{k=0}^{n} \alpha_k \tau^k$ для пары операторов $\hat{A}$ и $\hat{B}$, где $\hat{B}$ есть оператор умножения многочлена на независимую переменную, а оператор $\hat{A}$ многочлену $P_n(\tau)$ ставит в соответствие многочлен $\sum\limits_{k=1}^{n} \alpha_k \tau^{k-1}$.

8.3. Координатное представление линейных операторов

Пусть в Λ^n заданы базис $\{\, g_1,\, g_2,\, \ldots,\, g_n \}$ и линейный оператор $\hat{A}$, имеющий образы в Λ^m с базисом $\{\, f_1,\, f_2,\, \ldots,\, f_m \}$. Иначе говоря, $\hat{A}$ является *отображением* вида $\hat{A}:\ \Lambda^n \longrightarrow \Lambda^m$.

В § 7.2 показано, что $\forall x \in \Lambda^n$ существует единственное разложение

$$x = \sum_{i=1}^{n} \xi_i g_i\,, \qquad \text{то есть} \qquad \|x\|_g = \|\, \xi_1\, \xi_2\, \ldots\, \xi_n \,\|^{\mathrm{T}}\,.$$

Аналогично в Λ^m существует единственное разложение образа отображения $y = \hat{A}x$, для которого в силу линейности $\hat{A}$ справедливо представление вида

$$y = \hat{A}x = \hat{A}\left(\sum_{i=1}^{n} \xi_i g_i\right) = \sum_{i=1}^{n} \xi_i \hat{A} g_i \,.$$

Приняв во внимание возможность и единственность в Λ^m разложения $\hat{A}g_i = \sum\limits_{k=1}^{m} \alpha_{ki} f_k \quad \forall i = [1, n]$, с одной стороны, получаем, что

$$y = \sum_{k=1}^{m}\left(\sum_{i=1}^{n} \alpha_{ki}\xi_i\right) f_k \,.$$

С другой стороны, если $\|y\|_f = \| \; \eta_1 \, \eta_2 \, \ldots \, \eta_m \; \|^{\mathrm{T}}$ — координатное представление элемента y в базисе $\{ f_1, f_2, \ldots, f_m \}$, то имеет место равенство $y = \sum\limits_{k=1}^{m} \eta_k f_k$.

Наконец, в силу единственности разложения элемента конечномерного пространства по базису, получаем

$$\eta_k = \sum_{i=1}^{n} \alpha_{ki}\xi_i \qquad \forall k = [1, m] \,. \tag{8.3.1}$$

Данные соотношения позволяют находить координатное представление *образов* элементов линейного пространства по координатному представлению их *прообразов*. При этом отметим, что каждый линейный оператор вида $\hat{A} : \; \Lambda^n \longrightarrow \Lambda^m$ в паре конкретных базисов $\{ g_1, g_2, \ldots, g_n \}$ и $\{ f_1, f_2, \ldots, f_m \}$ полностью и однозначно описывается матрицей размера $m \times n$ с элементами α_{ki}.

Определение 8.3.1	Матрица $\|\hat{A}\|_{fg}$ размера $m \times n$, i-м столбцом которой является координатное представление $\|\hat{A}g_i\|_f$, называется *матрицей линейного оператора* $\hat{A}$ в базисах $\{ g_1, g_2, \ldots, g_n \}$ и $\{ f_1, f_2, \ldots, f_m \}$.

В использованных обозначениях

$$\left\|\hat{A}\right\|_{fg} = \left\| \begin{array}{ccccc} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \cdots & \alpha_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{m1} & \alpha_{m2} & \alpha_{m3} & \cdots & \alpha_{mn} \end{array} \right\| .$$

В матричной форме *уравнения связи* (8.3.1) координатных представлений образов и прообразов будут иметь вид

$$\|y\|_f = \left\|\hat{A}\right\|_{fg} \|x\|_g \,, \tag{8.3.2}$$

в чем легко убедиться, использовав в (8.3.1) для столбцов *двухиндексную* форму записи: $\eta_{k1} = \sum\limits_{i=1}^{n} \alpha_{ki}\xi_{i1} \quad \forall k = [1, m]$ [13] .

Заметим, что (8.3.1) структурно совпадает с символической формулой $y = \hat{A}x$, однако, в отличие от последней, равенство (8.3.1) полностью и однозначно описывает также и «механизм действия» линейного оператора $\hat{A}$.

Также очевидно, что определение 8.3.1 и формула (8.3.1) могут использоваться только в конечномерном случае, то есть когда существует базис.

По определению 8.3.1 столбцы матрицы $\left\|\hat{A}\right\|_{fg}$ суть координатные представления в Λ^m *образов* базисных элементов пространства Λ^n. В этом случае оказывается справедливой

Теорема 8.3.1 **Между множеством всех линейных операторов вида $\hat{A} : \Lambda^n \longrightarrow \Lambda^m$ и множеством всех матриц размера $m \times n$ имеется** *взаимно однозначное соответствие***.**

Доказательство.

По определению 8.3.1 каждому линейному оператору вида $\hat{A} : \Lambda^n \longrightarrow \Lambda^m$, для конкретной пары базисов $\{ g_1, g_2, \ldots, g_n\} \in \Lambda^n$ и $\{ f_1, f_2, \ldots, f_m\} \in \Lambda^m$, можно, и притом единственным способом, сопоставить матрицу размера $m \times n$.

С другой стороны, для любой матрицы размера $m \times n$ равенство (8.3.2) может быть принято за определение некоторого оператора вида $\hat{A} : \Lambda^n \longrightarrow \Lambda^m$, однозначность и линейность которого следует из свойств операций сложения матриц и умножения числа на матрицу.

Теорема доказана.

[13] См. пояснение к определению 1.1.3.

Пример 8.3.1 В трехмерном векторном пространстве с ортонормированным базисом $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ рассмотрим оператор $\hat{A}$, ортогонально проектирующий радиусы-векторы точек на плоскость Oxy — пространство с базисом $\{\vec{e}_1, \vec{e}_2\}$. Линейность такого оператора показана в § 2.1.

Поскольку в этом случае для $\hat{A}: \Lambda^3 \to \Lambda^2$

$$\begin{cases} \hat{A}\vec{e}_1 = 1\vec{e}_1 + 0\vec{e}_2 + 0\vec{e}_3\,, \\ \hat{A}\vec{e}_2 = 0\vec{e}_1 + 1\vec{e}_2 + 0\vec{e}_3\,, \end{cases} \qquad \text{то} \qquad \left\|\hat{A}\right\|_{ee} = \left\| \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{matrix} \right\|.$$

Действия с линейными операторами в матричной форме

Будем рассматривать далее оператор вида $\hat{A}: \Lambda^n \to \Lambda^n$, то есть *линейное преобразование*, действующее в Λ^n с базисом $\{g_1, g_2, \ldots, g_n\} \in \Lambda^n$, матрица которого квадратная, порядка n. Введенные в § 1.1 и § 5.1 операции с матрицами позволяют выполнять в конкретном базисе действия с линейными операторами в следующей форме.

1°. **Критерий равенства операторов:** $\hat{A} = \hat{B} \quad \Leftrightarrow \quad \|\hat{A}\|_g = \|\hat{B}\|_g$.

Действительно, согласно определению 8.2.1 условие $\hat{A} = \hat{B}$ означает, что $\hat{A}x = \hat{B}x \quad \forall x \in \Lambda^n$ или же в координатной форме в силу (8.3.2) $\|\hat{A}\|_g\|x\|_g = \|\hat{B}\|_g\|x\|_g \quad \forall x \in \Lambda^n$.

Но тогда по лемме 5.1.2 матрица $\|\hat{A}\|_g - \|\hat{B}\|_g$ нулевая и, следовательно, условие $\hat{A} = \hat{B}$ равносильно $\|\hat{A}\|_g = \|\hat{B}\|_g$.

2°. **Сложение операторов:** $\|\hat{A} + \hat{B}\,\|_g = \|\hat{A}\|_g + \|\hat{B}\|_g$.

Действительно, из разложений $\hat{A}g_i = \sum\limits_{k=1}^{n} \alpha_{ki}g_k \quad \forall i = [1, n]$ и $\hat{B}g_i = \sum\limits_{k=1}^{n} \beta_{ki}g_k \quad \forall i = [1, n]$ следует, что

$$(\hat{A} + \hat{B})x = \hat{A}x + \hat{B}x = \sum_{k=1}^{n} \alpha_{ki}g_k + \sum_{k=1}^{n} \beta_{ki}g_k = \sum_{k=1}^{n} (\alpha_{ki} + \beta_{ki})g_k\,.$$

3°. `Умножение числа на оператор:` $\|\lambda\hat{A}\|_g = \lambda\|\hat{A}\|_g$.

Из $\hat{A}g_i = \sum\limits_{k=1}^{n}\alpha_{ki}g_k \quad \forall i=[1,n]$ для любого числа λ находим, что

$$(\lambda\hat{A})g_i = \hat{A}(\lambda g_i) = \hat{A}\left(\lambda\sum_{k=1}^{n}\alpha_{ki}g_k\right) = \sum_{k=1}^{n}(\lambda\alpha_{ki})g_k\,.$$

4°. `Произведение операторов:` $\|\hat{A}\hat{B}\,\|_g = \|\hat{A}\|_g\|\hat{B}\|_g$.

По определению матрицы линейного оператора имеем

$$(\hat{A}\hat{B})g_i = \hat{A}(\hat{B}g_i) = \hat{A}\sum_{k=1}^{n}\beta_{ki}g_k = \sum_{k=1}^{n}\beta_{ki}\hat{A}\,(g_k) =$$
$$= \sum_{k=1}^{n}\beta_{ki}\sum_{j=1}^{n}\alpha_{jk}g_j = \sum_{j=1}^{n}\left(\sum_{k=1}^{n}\alpha_{jk}\beta_{ki}\right)g_j = \sum_{j=1}^{n}\kappa_{ji}g_j\,,$$

где $\kappa_{ji} = \sum\limits_{k=1}^{n}\alpha_{jk}\beta_{ki}$, что совпадает с определением 5.1.1 для произведения матриц $\|\hat{A}\|_g$ и $\|\hat{B}\|_g$.

5°. `Обращение операторов:` $\|\hat{A}^{-1}\|_g = \|\hat{A}\|_g^{-1}$.

Будем предполагать, что обратный оператор существует. Поскольку из определения 8.2.8 следует, что $\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \hat{E}$, и, принимая во внимание результат пункта 4°, получаем, что искомое матричное представление $\|\hat{A}^{-1}\|_g$ оператора $\hat{A}^{-1}$ должно удовлетворять соотношениям

$$\|\hat{A}^{-1}\|_g\|\hat{A}\|_g = \|\hat{A}\|_g\|\hat{A}^{-1}\|_g = \|\hat{E}\|_g,$$

то есть являться обратной матрицей к матрице $\|\hat{A}\|_g$.

`Следствие 8.3.1` **Размерность линейного пространства линейных отображений вида $\hat{A}:\ \Lambda^n \longrightarrow \Lambda^m$ равна $n\times m$.**

`Доказательство.`

Из теоремы 8.3.1 и правил действий с линейными операторами в матричной форме следует изоморфизм линейного пространства линейных операторов $\hat{A}:\ \Lambda^n \longrightarrow \Lambda^m$ и линейного пространства всех матриц размера $m\times n$. Но тогда по теореме 7.5.1 (об изоморфизме) их размерности равны.

`Следствие доказано.`

Изменение матрицы линейного оператора при замене базиса

Выясним, как меняется $\left\|\hat{A}\right\|_{fg}$ — матрица линейного отображения $\hat{A}:\ \Lambda^n \longrightarrow \Lambda^m$ при замене базисов.

Пусть в пространстве Λ^n даны два базиса $\{\,g_1,\,g_2,\,\ldots,\,g_n\}$ и $\{\,g'_1,\,g'_2,\,\ldots,\,g'_n\}$, связанные матрицей перехода $\|G\|$, а в Λ^m — два базиса $\{\,f_1,\,f_2,\,\ldots,\,f_m\}$ и $\{\,f'_1,\,f'_2,\,\ldots,\,f'_m\}$ с матрицей перехода $\|F\|$. Найдем соотношение, связывающее $\left\|\hat{A}\right\|_{f'g'}$ и $\left\|\hat{A}\right\|_{fg}$. В этом случае справедлива

Теорема 8.3.2

Матрица линейного оператора $\left\|\hat{A}\right\|_{f'g'}$ в базисах $\{\,g'_1,\,g'_2,\,\ldots,\,g'_n\}$ и $\{\,f'_1,\,f'_2,\,\ldots,\,f'_m\}$ связана с матрицей этого же оператора $\left\|\hat{A}\right\|_{fg}$ в базисах $\{\,g_1,\,g_2,\,\ldots,\,g_n\}$ и $\{\,f_1,\,f_2,\,\ldots,\,f_m\}$ соотношением

$$\left\|\hat{A}\right\|_{f'g'} = \|F\|^{-1}\left\|\hat{A}\right\|_{fg}\|G\|\,.$$

Доказательство.

По теореме 7.3.1 при переходе от базиса $\{\,g_1,\,g_2,\,\ldots,\,g_n\}$ к базису $\{\,g'_1,\,g'_2,\,\ldots,\,g'_n\}$ компоненты элементов $x\in\Lambda^n$ — прообраза и $y\in\Lambda^m$ — образа при действии оператора $\hat{A}$ в этих базисах связаны формулами перехода, то есть равенствами

$$\|x\|_g = \|G\|\|x\|_{g'} \qquad \text{и} \qquad \|y\|_f = \|F\|\|y\|_{f'}\,,$$

где

$$\|x\|_g = \|\,\xi_1,\,\xi_2,\,\ldots,\,\xi_n\,\|^{\mathrm{T}} \qquad \text{и} \qquad \|x\|_{g'} = \|\,\xi'_1,\,\xi'_2,\,\ldots,\,\xi'_n\,\|^{\mathrm{T}}$$

и аналогично

$$\|y\|_f = \|\,\eta_1,\,\eta_2,\,\ldots,\,\eta_m\,\|^{\mathrm{T}} \qquad \text{и} \qquad \|y\|_{f'} = \|\,\eta'_1,\,\eta'_2,\,\ldots,\,\eta'_m\,\|^{\mathrm{T}}\,.$$

При этом в рассматриваемых базисах образы и прообразы элементов связаны соотношениями

$$\|y\|_f = \left\|\hat{A}\right\|_{fg}\|x\|_g \qquad \text{и} \qquad \|y\|_{f'} = \left\|\hat{A}\right\|_{f'g'}\|x\|_{g'}\,,$$

и поскольку матрица перехода имеет обратную, то из выписанных соотношений последовательно получаем

$$\|y\|_{f'} = \|F\|^{-1}\|y\|_f = \|F\|^{-1}\left\|\hat{A}\right\|_{fg}\|x\|_g = \|F\|^{-1}\left\|\hat{A}\right\|_{fg}\|G\|\|x\|_{g'}.$$

Но, с другой стороны, $\|y\|_{f'} = \left\|\hat{A}\right\|_{f'g'}\|x\|_{g'}$, и мы приходим к равенству

$$\left(\left\|\hat{A}\right\|_{f'g'} - \|F\|^{-1}\left\|\hat{A}\right\|_{fg}\|G\|\right)\|x\|_{g'} = \|o\| \qquad \forall x \in E^n\,,$$

из которого в силу произвольности столбца $\|x\|_{g'}$ и леммы 5.1.2 матрица, стоящая в круглых скобках, *нулевая*. Откуда следует справедливость утверждения теоремы.

Теорема доказана.

Следствие 8.3.2 **Матрица линейного** *преобразования* $\hat{A}$ **при переходе от базиса** $\{g_1, g_2, \ldots, g_n\}$ **к другому базису** $\{g'_1, g'_2, \ldots, g'_n\}$ **с матрицей перехода** $\|S\|$ **в** Λ^n **изменяется по правилу**

$$\left\|\hat{A}\right\|_{g'} = \|S\|^{-1}\left\|\hat{A}\right\|_g\|S\|\,.$$

Доказательство.

Следует непосредственно из утверждения теоремы 8.3.2.

Следствие доказано.

Следствие 8.3.3 **Значение** *определителя* **матрицы линейного преобразования** *не зависит* **от выбора базиса в** Λ^n**.**

Доказательство.

По следствию 8.3.2 $\det\left\|\hat{A}\right\|_{g'} = \det\left(\|S\|^{-1}\left\|\hat{A}\right\|_g\|S\|\right)$, а в силу теоремы 6.2.4

$$\det\left(\|S\|^{-1}\left\|\hat{A}\right\|_g\|S\|\right) = \left(\det\|S\|^{-1}\right)\left(\det\left\|\hat{A}\right\|_g\right)\left(\det\|S\|\right)\,,$$

а $\det \|S\|^{-1} \cdot \det \|S\| = 1$, то окончательно получаем, что

$$\det \left\| \hat{A} \right\|_{g'} = \det \left\| \hat{A} \right\|_{g} .$$

Следствие доказано.

В заключение проверьте самостоятельно, что в силу теоремы 8.3.2 в любом базисе нулевой оператор будет иметь нулевую матрицу, а единичный оператор — единичную.

8.4. Область значений и ядро линейного оператора

Трактуя линейный оператор, действующий в линейном пространстве как некоторое обобщение понятия *функции*, естественно рассмотреть вопрос об области определения и области значений линейных операторов.

Под областью значений линейного оператора $\hat{A}$ будем понимать множество образов всех элементов $x \in \Lambda$, то есть элементов вида $\hat{A}x$. В этом случае очевидно, что для любого линейного оператора его область определения совпадает с Λ.

Ответ на вопрос: «Что представляет собой область значений линейного оператора?» дает

Теорема 8.4.1 **Пусть $\hat{A}$ – линейный оператор, действующий в линейном пространстве Λ. Тогда**

1°. Множество элементов $\mathrm{Im}\hat{A}$ есть *подпространство* **в Λ.**

2°. Если, кроме того, $\Lambda = \Lambda^n$ с базисом $\{ g_1, g_2, \ldots, g_n \}$, то $\dim \mathrm{Im}\hat{A} = \mathrm{rg} \left\| \hat{A} \right\|_g$**.**

Доказательство.

Пусть Λ^* есть множество элементов вида $\hat{A}x$ и пусть при этом $y_1, y_2 \in \Lambda^*$. Тогда существуют $x_1, x_2 \in \Lambda$, такие, что $y_1 = \hat{A}x_1$ и $y_2 = \hat{A}x_2$. В силу линейности оператора $\hat{A}$ имеем

$$\lambda_1 y_1 + \lambda_2 y_2 = \lambda_1 \hat{A}x_1 + \lambda_2 \hat{A}xy_2 = \hat{A}\big(\lambda_1 x_1 + \lambda_2 x_2\big) \in \Lambda^*,$$

и потому Λ^* есть подпространство.

Пусть $\Lambda = \Lambda^n$ с базисом $\{ g_1, g_2, \ldots, g_n \}$. Поскольку каждый элемент $x \in \Lambda$ есть некоторая линейная комбинация базисных элементов, то в силу линейности каждый элемент из области значений Λ^* есть та же самая линейная комбинация элементов $\hat{A}g_1, \hat{A}g_2, \ldots, \hat{A}g_n$, то есть Λ^* — линейная оболочка множества $\{ \hat{A}g_1, \hat{A}g_2, \ldots, \hat{A}g_n \}$.

Пусть максимальное число линейно независимых элементов в этой линейной оболочке оказалось равным k. Тогда, применяя теорему 7.4.1, приходим к заключению, что размерность Λ^* есть k, а из теоремы 7.5.2 следует, что и $\operatorname{rg} \left\| \hat{A} \right\|_g = k$.

Теорема доказана.

Определение 8.4.1 *Рангом* линейного оператора $\hat{A}$ в Λ^n называется размерность его области значений. Ранг линейного оператора $\hat{A}$ обозначается как $\operatorname{rg}\hat{A}$ (или $\operatorname{rank}\hat{A}$).

Следствие 8.4.1 **В Λ^n $\operatorname{rg}\hat{A} = \operatorname{rg} \left\| \hat{A} \right\|_g \leq n$ и не зависит от выбора базиса.**

Следствие 8.4.2 **Размерность области значений линейного оператора $\hat{A}$, действующего на некотором подпространстве Λ^* линейного пространства Λ, не превосходит $\dim \Lambda^*$.**

Доказательство.

Поскольку подпространство является линейным пространством, то к нему применима теорема 8.4.1.

Следствие доказано.

Теорема 8.4.2 **Ранг произведения линейных операторов $\hat{A}$ и $\hat{B}$ не превосходит ранга каждого из этих операторов.**

Доказательство.

Рассмотрим область значений линейного оператора $\hat{A}\hat{B}$. По следствию 8.4.2 это подпространство имеет размерность не большую, чем размерность области значений оператора $\hat{B}$.

С другой стороны, область значений оператора $\hat{A}\hat{B}$ содержится в области значений оператора $\hat{A}$, и, следовательно, размерность области значений $\hat{A}\hat{B}$ не превосходит размерности области значений $\hat{A}$.

Теорема доказана.

Теорема 8.4.3 **Если квадратная матрица $\|A\|$ невырожденная, то для любой квадратной матрицы $\|B\|$ того же размера** $\operatorname{rg}\big(\|A\|\|B\|\big) = \operatorname{rg}\big(\|B\|\|A\|\big) = \operatorname{rg}\|B\|$.

Доказательство.

Будем рассматривать матрицы $\|A\|$ и $\|B\|$ как координатные представления линейных операторов $\hat{A}$ и $\hat{B}$ в некотором базисе пространства Λ^n.

Если $\det\|A\| \neq 0$, то существует $\|A\|^{-1}$ и в силу теоремы 8.4.2 имеем, с одной стороны, $\operatorname{rg}\big(\|A\|\|B\|\big) \leq \operatorname{rg}\|B\|$, но с другой стороны,

$$\operatorname{rg}\|B\| = \operatorname{rg}\big(\|A\|^{-1}\|A\|\|B\|\big) \leq \operatorname{rg}\big(\|A\|\|B\|\big).$$

Поэтому $\operatorname{rg}\big(\|A\|\|B\|\big) = \operatorname{rg}\big(\|B\|\|A\|\big) = \operatorname{rg}\|B\|$.

Теорема доказана.

Замечание 8.4.1. 1°. Если матрица $\|B\|$ не квадратная, но существует одно из произведений $\|A\|\|B\|$ или $\|B\|\|A\|$, то при $\det\|A\| \neq 0$ также верны равенства $\operatorname{rg}\big(\|A\|\|B\|\big) = \operatorname{rg}\|B\|$ или соответственно $\operatorname{rg}\big(\|B\|\|A\|\big) = \operatorname{rg}\|B\|$.

В этом можно убедиться, заменив матрицу $\|B\|$ матрицей $\|B\|^+$, являющейся дополнением нулевыми столбцами или нулевыми строками $\|B\|$ до квадратной так, чтобы существовали $\|A\|\|B\|^+$ или $\|B\|^+\|A\|$, ибо очевидно, что $\operatorname{rg}\|B\|^+ = \operatorname{rg}\|B\|$.

2°. Ранг произведения матриц может быть меньше рангов каждого из сомножителей. Например,

$$\begin{Vmatrix} 1 & 0 \\ 0 & 0 \end{Vmatrix} \begin{Vmatrix} 0 & 0 \\ 0 & 1 \end{Vmatrix} = \begin{Vmatrix} 0 & 0 \\ 0 & 0 \end{Vmatrix}.$$

Другой (помимо области значений и ранга) важной характеристикой линейного оператора $\hat{A}$ является совокупность элементов x линейного пространства Λ, называемая *ядром* линейного оператора и обозначаемая $\ker \hat{A}$.

Определение 8.4.2	Ядро линейного оператора $\hat{A}$ состоит из элементов $x \in \Lambda$, таких, что $\hat{A}x = o$.

Теорема 8.4.4 **Если $\Lambda = \Lambda^n$ и $\operatorname{rg}\hat{A} = r$, то $\ker A$ есть подпространство в Λ^n и $\dim \ker \hat{A} = n - r$.**

Доказательство.

Пусть в базисе $\{ g_1, g_2, \ldots, g_n \}$ оператор $\hat{A}$ имеет матрицу $\left\| \hat{A} \right\|_g = \|\alpha_{ij}\|$. По следствию 8.4.1 $\operatorname{rg} \left\| \hat{A} \right\|_g = r$ для любого базиса.

В координатной форме равенство $\hat{A}x = o$, то есть условие принадлежности некоторого элемента $x \in \Lambda^n$ с координатным представлением $\|x\|_g = \| \, \xi_1, \xi_2, \ldots, \xi_n \, \|^{\mathrm{T}}$ ядру оператора $\hat{A}$, в силу (8.3.2), имеет вид $\left\| \hat{A} \right\|_g \|x\|_g = \|o\|$ или

$$\sum_{j=1}^{n} \alpha_{ij} \xi_j = 0 \qquad \forall i = [1, n] \,. \tag{8.4.1}$$

При этом решения однородной системы линейных уравнений (8.4.1) образуют в своей совокупности линейное пространство.

Наконец, поскольку размерность ядра есть максимальное число линейно независимых решений этой системы уравнений, то она, согласно теореме 6.7.1, равна $n - r$.

Теорема доказана.

Типы линейных отображений

Как было отмечено в § 8.1, в тех случаях, когда область значений оператора не принадлежит области определения, следует говорить об отображении. В § 7.5 было использовано понятие *взаимно однозначного отображения*, называемого иногда *биекцией*.

Для отображений также выделяются специальные случаи так называемых инъективных и сюръективных отображений. Рассмотрим эти случаи подробнее.

Определение 8.4.3	Отображение $y = \hat{A}x \quad x \in \Omega,\ y \in \Theta$ множества Ω в множество Θ называется *инъективным* (или *инъекцией*), если из условия $\hat{A}x_1 = \hat{A}x_2$ вытекает $x_1 = x_2$.

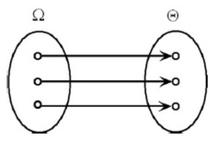
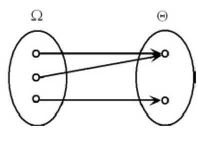
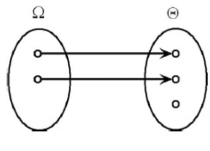
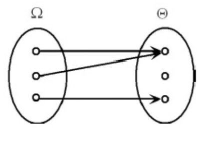
В случае инъекции множество всех значений оператора $\hat{A}$ может не совпадать с Θ.

Определение 8.4.4	Отображение $y = \hat{A}x \quad x \in \Omega,\ y \in \Theta$ множества Ω в множество Θ называется *сюръективным* (или *сюръекцией*), если каждый элемент из Θ имеет прообраз в Ω.

В случае сюръекции прообраз любого элемента из Θ всегда существует в Ω, но, вообще говоря, он не единственен.

В табл. 8.4.1 для сравнения приведены примеры отображений различных типов.

Таблица 8.4.1

Тип отображения	Инъективное	Неинъективное
Сюръективное		
Несюръективное		

Отметим также, что в конечномерном случае сюръективность отображения $\hat{A} : \Lambda^n \longrightarrow \Lambda^m$ означает выполнение условия $\Theta = \Lambda^m$, а инъективность — условия $\ker \hat{A} = \{o\}$.

Альтернативную форму условий инъективности и сюръективности в конечномерном случае дает

Теорема 8.4.5 **Ранг матрицы линейного оператора, являющегося сюръективным отображением, равен числу ее строк, а ранг матрицы инъективного отображения равен числу ее столбцов.**

Доказательство.

1°. Пусть в базисах $\{ g_1, g_2, \dots, g_n\}$ и $\{ f_1, f_2, \dots, f_m\}$ отображение $\hat{A} : \Lambda^n \longrightarrow \Lambda^m$ имеет матрицу $\left\|\hat{A}\right\|_{fg}$, причем $\operatorname{rg}\left\|\hat{A}\right\|_{fg} = m$.

Тогда система линейных уравнений $\left\|\hat{A}\right\|_{fg} \|x\|_g = \|y\|_f$ по теореме 6.6.1 (Кронекера – Капелли) имеет решение $\forall y \in \Lambda^m$, поскольку для расширенной матрицы этой системы ранг также равен m. Значит, для $\hat{A}$ каждый образ имеет хотя бы один прообраз, и данное отображение сюръективно.

2°. Пусть $\operatorname{rg}\left\|\hat{A}\right\|_{fg} = n$. Тогда, по теореме 6.4.1 (Крамера), система линейных уравнений вида

$$\left\|\hat{A}\right\|_{fg} (\|x_1\|_g - \|x_2\|_g) = \|o\|_f$$

имеет единственное решение, которое очевидно тривиальное. Поэтому равные образы имеют равные прообразы, и, следовательно, рассматриваемое отображение инъективно.

Теорема доказана.

Иными словами, инъективность отображения $\hat{A}$ равносильна выполнению равенств $\operatorname{rg}\hat{A} = \operatorname{rg}\left\|\hat{A}\right\|_{fg} = \dim \Lambda^n = n$, а сюръективность — $\operatorname{rg}\hat{A} = \operatorname{rg}\left\|\hat{A}\right\|_{fg} = \dim \Lambda^m = m$.

Наконец, отображение, являющееся одновременно и инъективным и сюръективным, будет *взаимно однозначным*, или *биекцией*.

В общем случае, исследование свойств линейного отображения может оказаться достаточно сложной задачей. Однако если область его значений принадлежит конечномерному линейному пространству, то, пользуясь теоремой 7.5.1 (об изоморфизме), можно попытаться свести исследование отображения к исследованию преобразования, установив изоморфизм между областью значений отображения и некоторым подпространством в области его определения.

Для этого в случае, когда линейный оператор $\hat{A}$ является преобразованием в Λ^n с базисом $\{ g_1, g_2, \ldots, g_n \}$, оказывается полезным дать уточняющее определению 8.3.1.

Определение 8.4.5	Квадратная, порядка n, матрица $\left\|\hat{A}\right\|_g$, столбцы которой есть координатные представления элементов $\hat{A}g_1, \hat{A}g_2, \ldots, \hat{A}g_n$ в базисе $\{ g_1, g_2, \ldots, g_n\}$, называется *матрицей линейного преобразования* $\hat{A}$ этом в базисе.

Использование описанного подхода иллюстрирует

Пример 8.4.1 1°. Оператор $\hat{Pr}$, ставящий в соответствие каждой точке трехмерного геометрического пространства ее ортогональную проекцию на некоторую фиксированную прямую, проходящую через начало координат, очевидно есть отображение $\Lambda^3 \longrightarrow \Lambda^1$, которое, однако, можно рассматривать и как преобразование трехмерного пространства в одномерное подпространство.

Отметим, что, хотя в данном случае и отображение, и преобразование реализуют геометрически одну и ту же операцию, вид задающих ее матриц может быть различным.

Например, пусть в ортонормированной системе координат $\{O, x, y, z\}$ прямая, на которую выполняется ортогональное проектирование, задана направляющим вектором $\|\vec{e}_1^{\,*}\|_e = \|\, 1\ 1\ 1 \,\|^{\mathrm{T}}$. В этом случае $\vec{r}^{\,*}$ — радиус-вектор ортогональной проекции на прямую точки $\vec{r}$ с $\|\vec{r}^{\,*}\|_e = \|\, x\ y\ z \,\|^{\mathrm{T}}$ равен

$$\vec{r}^{\,*} = \frac{(\vec{r}, \vec{e}^{\,*})}{|\vec{e}^{\,*}|^2}\,\vec{e}^{\,*}\,,$$

или $\left\| \begin{array}{c} x^* \\ y^* \\ z^* \end{array} \right\| = \left\| \begin{array}{c} \dfrac{x+y+z}{3} \\ \dfrac{x+y+z}{3} \\ \dfrac{x+y+z}{3} \end{array} \right\|$, и, следовательно, матрица

преобразования имеет вид $\left\| \hat{Pr} \right\|_e = \dfrac{1}{3} \left\| \begin{array}{ccc} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{array} \right\|$.

Но, с другой стороны, приняв вектор $\vec{e}_1^{\,*}$ за базисный в Λ^1, получим, согласно определению 8.4.5, матрицу отображения в виде $\left\| \hat{Pr} \right\|_{e^* e} = \dfrac{1}{3} \| \, 1 \; 1 \; 1 \, \|$ ..

2°. Пусть линейный оператор ставит в соответствие каждой матрице второго порядка $\left\| \begin{array}{cc} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{array} \right\|$ двумерный столбец вида $\left\| \begin{array}{c} \alpha_{11} + \alpha_{12} \\ \alpha_{21} + \alpha_{22} \end{array} \right\|$.

Исследование свойств данного отображения можно свести к исследованию свойств преобразования, ставящего в соответствие квадратным матрицам $\left\| \begin{array}{cc} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{array} \right\|$ квадратные матрицы вида $\left\| \begin{array}{cc} \alpha_{11} + \alpha_{12} & 0 \\ \alpha_{21} + \alpha_{22} & 0 \end{array} \right\|$.

Задача 8.4.1 *Линейный оператор $\hat{A}: \quad \Lambda^3 \longrightarrow \Lambda^3$ в некотором базисе задан матрицей $\left\| \begin{array}{ccc} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 5 & 7 \end{array} \right\|$. Найти его ядро и множество значений. Выяснить, является ли данный оператор инъективным или сюръективным.*

Решение. 1°. Пусть $\|x\|_g = \| \, \xi_1 \; \xi_2 \; \xi_3 \, \|^{\mathrm{T}}$ и $\|y\|_g = \| \, \eta_1 \; \eta_2 \; \eta_3 \, \|^{\mathrm{T}}$ — координатные представления соответственно прообраза и образа оператора $y = \hat{A}x$. Тогда ядро – множество элементов x, таких, что $\hat{A}x = 0$ задается в координатном представлении системой линейных уравнений

$$\left\|\hat{A}\right\|_g \|x\| = \|o\| \qquad \text{или} \qquad \left\{\begin{array}{rcrcrcr} \xi_1 & + & 2\xi_2 & + & 3\xi_3 & = & 0\,, \\ 2\xi_1 & + & 3\xi_2 & + & 4\xi_3 & = & 0\,, \\ 3\xi_1 & + & 5\xi_2 & + & 7\xi_3 & = & 0\,, \end{array}\right.$$

общее решение которой есть $\left\|\begin{array}{c} \xi_1 \\ \xi_2 \\ \xi_3 \end{array}\right\| = \lambda \left\|\begin{array}{r} 1 \\ -2 \\ 1 \end{array}\right\|$.

Отсюда заключаем, что ядро линейного оператора $\hat{A}$ есть линейная оболочка элемента $\|x\|_g = \|\ 1\ \ -2\ \ 1\ \|^{\mathrm{T}}$, и поскольку оно не состоит только из нулевого элемента, то данный оператор неинъективный.

К этому же заключению можно прийти, приняв во внимание, что

$$\operatorname{rg} \left\|\hat{A}\right\|_g = \operatorname{rg} \left\|\begin{array}{ccc} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 5 & 7 \end{array}\right\| = \operatorname{rg} \left\|\begin{array}{ccc} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right\| = 2 < 3$$

— числа столбцов матрицы оператора.

2°. Область значений линейного оператора $\hat{A}$ состоит из элементов $y \in \Theta$, таких, что $y = \hat{A}x \ \ \forall x \in \Omega$. В координатной форме принадлежность элемента y ко множеству значений означает совместность системы линейных уравнений

$$\left\|\begin{array}{ccc} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 5 & 7 \end{array}\right\| \left\|\begin{array}{c} \xi_1 \\ \xi_2 \\ \xi_3 \end{array}\right\| = \left\|\begin{array}{c} \eta_1 \\ \eta_2 \\ \eta_3 \end{array}\right\| .$$

Следовательно, нам необходимо выяснить, при каких значениях η_1, η_2, η_3 данная система линейных уравнений совместна. Это можно сделать, например, при помощи теоремы 6.6.1 (Кронекера – Капелли), сравнив ранги основной и расширенной матриц данной системы. Действительно, из равенств

$$2 = \operatorname{rg} \left\|\begin{array}{ccc} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 5 & 7 \end{array}\right\| = \operatorname{rg} \left\|\begin{array}{ccc|c} 1 & 2 & 3 & \eta_1 \\ 2 & 3 & 4 & \eta_2 \\ 3 & 5 & 7 & \eta_3 \end{array}\right\| =$$

$$= \operatorname{rg} \left\| \begin{array}{ccc|ccccc} 1 & 2 & 3 & \eta_1 & & & & \\ 0 & 1 & 2 & 2\eta_1 & - & \eta_2 & & \\ 0 & 0 & 0 & -\eta_1 & - & \eta_2 & + & \eta_3 \end{array} \right\| = 2$$

следует, что для совместности необходимо и достаточно, чтобы $-\eta_1 - \eta_2 + \eta_3 = 0$, а это, в свою очередь, означает, что множество значений оператора $\hat{A}$ состоит из элементов вида

$$\left\| \begin{matrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{matrix} \right\| = \mu_1 \left\| \begin{matrix} -1 \\ 1 \\ 0 \end{matrix} \right\| + \mu_2 \left\| \begin{matrix} 1 \\ 0 \\ 1 \end{matrix} \right\| \qquad \forall \mu_1,\, \mu_2\,,$$

являющихся решениями системы из одного уравнения с тремя неизвестными: $\eta_1 + \eta_2 - \eta_3 = 0$.

Решение получено.

В заключение заметим, что поскольку не каждый элемент $y \in \Theta \in \Lambda^3$ имеет прообраз в $\Omega = \Lambda^3$, то данное отображение не является и сюръективным.

8.5. Инвариантные подпространства и собственные векторы

Определение 8.5.1

Подпространство Λ^* линейного пространства Λ называется *инвариантным* подпространством линейного оператора[14] $\hat{A}$, если $\forall x \in \Lambda^* \quad \hat{A}x \in \Lambda^*$

Пример 8.5.1

1°. Множество радиусов-векторов точек некоторой прямой на плоскости Oxy, проходящей через начало координат, является инвариантным подпространством оператора поворота на угол π этих радиусов-векторов вокруг оси Oz (см. рис. 8.1).

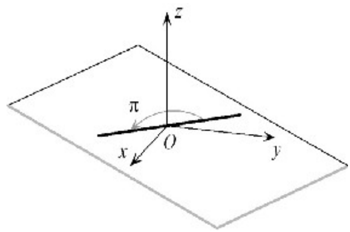


Рис. 8.1

[14] В этом параграфе речь пойдет только об операторах, являющихся преобразованиями.

2°. Для оператора дифференцирования в линейном пространстве функций $f(\tau)$, имеющих на (α, β) производную любого порядка, n-мерным инвариантным подпространством является линейная оболочка совокупности элементов вида $\{ e^{\lambda_1 \tau}, e^{\lambda_2 \tau}, \ldots, e^{\lambda_n \tau} \}$, где $\lambda_1, \lambda_2, \ldots \lambda_n$ — некоторые, попарно различные константы.

Зависимость вида матрицы линейного оператора $\hat{A}$ и инвариантности подпространства в Λ^n дает

Теорема 8.5.1 **Линейный оператор $\hat{A}$, заданный в линейном пространстве Λ^n с базисом $\{ g_1, g_2, \ldots, g_n \}$, имеет матрицу вида**

$$\left\| \begin{array}{ccc|ccc} \alpha_{11} & \ldots & \alpha_{1r} & \alpha_{1,r+1} & \ldots & \alpha_{1n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha_{r1} & \ldots & \alpha_{rr} & \alpha_{r,r+1} & \ldots & \alpha_{rn} \\ \hline 0 & \ldots & 0 & \alpha_{r+1,r+1} & \ldots & \alpha_{r+1,n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & \alpha_{n,r+1} & \ldots & \alpha_{nn} \end{array} \right\|$$

тогда и только тогда, когда линейная оболочка подмножества первых r базисных элементов $\{ g_1, g_2, \ldots, g_r \}$ есть Λ^* — инвариантное подпространство оператора $\hat{A}$.

Доказательство.

Докажем достаточность. Пусть матрица оператора $\hat{A}$ имеет указанный в формулировке теоремы вид. Тогда образ любой линейной комбинации элементов $\{ g_1, g_2, \ldots, g_r \}$ будет принадлежать их линейной оболочке, поскольку (в силу определения 8.3.1) j-й столбец матрицы линейного оператора есть образ j-го базисного элемента. Иначе говоря, если $\sum\limits_{k=1}^{r} \lambda_k g_k \in \Lambda^*$, то и

$$\hat{A} \left(\sum_{k=1}^{r} \lambda_k g_k \right) = \sum_{k=1}^{r} \lambda_k \hat{A} g_k = \sum_{k=1}^{r} \lambda_k \sum_{i=1}^{r} \alpha_{ik} g_i =$$

$$= \sum_{i=1}^{r} \left(\sum_{k=1}^{r} \alpha_{ik} \lambda_k \right) g_k = \sum_{i=1}^{r} \beta_i g_i \in \Lambda^* .$$

Наконец, из теоремы 7.4.1 следует, что Λ^* — подпространство. Достаточность доказана.

Докажем необходимость. Пусть Λ^* — инвариантное подпространство линейного оператора $\hat{A}$ — является линейной оболочкой подмножества базисных векторов $\{\, g_1,\, g_2,\, \ldots,\, g_r \}$.

Тогда образ любого, в том числе и базисного, элемента, принадлежащего Λ^*, также будет принадлежать Λ^*. Это, в свою очередь, означает, что

$$\hat{A}g_k = \sum_{i=1}^{r} \alpha_{ik} g_i + \sum_{i=r+1}^{n} 0 \cdot g_i \qquad \forall k = [1, r]$$

и, в сочетании с определением 8.4.5, доказывает необходимость.

Теорема доказана.

Задача 8.5.1 *Показать, что всякое инвариантное подпространство обратимого линейного оператора $\hat{A}$ является также инвариантным подпространством оператора $\hat{A}^{-1}$.*

Решение. Пусть $x \in \Lambda^*$, где Λ^* — инвариантное подпространство оператора $\hat{A}$, тогда по условию задачи $y = \hat{A}x \in \Lambda^*$. Если оператор обратимый, то для него существует обратный $\hat{A}^{-1}$ и связь элементов x и y можно записать также в виде $x = \hat{A}^{-1}y$.

Но поскольку как x, так и y принадлежат Λ^*, то это означает инвариантность подпространства Λ^* относительно оператора $\hat{A}^{-1}$.

Решение получено.

В приложениях важную роль играют определяемые ниже «собственные векторы» и «собственные значения» линейных преобразований.

Определение 8.5.2 *Ненулевой* элемент f называется *собственным вектором* линейного преобразования $\hat{A}$, если существует число λ, такое, что $\hat{A}f = \lambda f$.

Число λ в этом случае называется *собственным значением* $\hat{A}$, соответствующим (или отвечающему) собственному вектору f.

Задача 8.5.2 *Показать, что если линейное преобразование $\hat{A}$ имеет собственный вектор f с соответствующим ему собственным значением λ, то элемент f будет также являться собственным вектором линейного преобразования $\hat{A}^2 = \hat{A}\hat{A}$ с собственным значением λ^2.*

Решение. По условию $\hat{A}f = \lambda f$, но тогда в силу линейности $\hat{A}$ имеем

$$\hat{A}^2 f = \hat{A}(\hat{A}f) = \hat{A}(\lambda f) = \lambda \hat{A}f = \lambda^2 f\,.$$

Решение получено.

Заметим, что, по определению 8.5.2, f ненулевой элемент ядра линейного преобразования $\hat{A} - \lambda\hat{E}$, поскольку $\big(\hat{A} - \lambda\hat{E}\big)f = o$.

Замечание о важности собственных векторов

Допустим, что для некоторого линейного преобразования $\hat{A}$, заданного в Λ^n, удалось найти n *линейно независимых* собственных векторов $g_1, g_2, \ldots, g_n$, для которых выполнены равенства

$$\hat{A}g_1 = \lambda_1 g_1, \quad \hat{A}g_2 = \lambda_2 g_2, \quad \ldots, \quad \hat{A}g_n = \lambda_n g_n\,.$$

Если принять набор этих элементов за базис $\{g_1, g_2, \ldots, g_n\}$, то данные соотношения можно рассматривать как *координатные разложения* образов базисных элементов:

$$\hat{A}g_k = 0g_1 + 0g_2 + \ldots + \lambda_k g_k + \ldots + 0g_n \qquad \forall k = [1, n]\,.$$

Согласно теореме 7.2.1, эти разложения единственны, поэтому, исходя из определения 8.3.1, можно утверждать, что матрица линейного преобразования в этом базисе будет иметь *диагональный* вид:

$$\big\|\hat{A}\big\|_g = \left\|\begin{array}{cccccc} \lambda_1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \lambda_3 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_{n-1} & 0 \\ 0 & 0 & 0 & \ldots & 0 & \lambda_n \end{array}\right\|,$$

благодаря которому исследование свойств этого преобразования может существенно упроститься.

В том случае, когда удается найти базис, в котором матрица линейного преобразования имеет диагональный вид, данное преобразование принято называть *диагонализуемым.*

Вычисление собственных векторов и собственных значений линейного оператора в Λ^n

Выберем в Λ^n некоторый базис $\{ g_1, g_2, \ldots, g_n \}$, в котором координатное разложение f — собственного вектора линейного преобразования $\hat{A}$ будет $f = \sum\limits_{i=1}^{n} \xi_i g_i$, а линейное преобразование $\hat{A}$ имеет матрицу $\left\| \hat{A} \right\|_g = \| \alpha_{ij} \|$.

Пользуясь результатами, полученными в § 8.3, символическое равенство $\hat{A} f = \lambda f$ можно записать при помощи матричных операций в виде $\left\| \hat{A} \right\|_g \| f \|_g = \lambda \| f \|_g$, или же в координатной форме:

$$\left\{ \begin{array}{ccccccccc} \alpha_{11}\xi_1 & + & \alpha_{12}\xi_2 & + & \ldots & + & \alpha_{1n}\xi_n & = & \lambda\,\xi_1, \\ \alpha_{21}\xi_1 & + & \alpha_{22}\xi_2 & + & \ldots & + & \alpha_{2n}\xi_n & = & \lambda\,\xi_2, \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1}\xi_1 & + & \alpha_{n2}\xi_2 & + & \ldots & + & \alpha_{nn}\xi_n & = & \lambda\,\xi_n, \end{array} \right.$$

что по правилам действий с линейными операторами в координатах (см. §8.3) равносильно $\left\| \hat{A} - \lambda \hat{E} \right\|_g \| f \|_g = \| o \|_g$ или

$$\left\{ \begin{array}{ccccccccc} (\alpha_{11} - \lambda)\xi_1 & + & \alpha_{12}\xi_2 & + & \ldots & + & \alpha_{1n}\xi_n & = & 0, \\ \alpha_{21}\xi_1 & + & (\alpha_{22} - \lambda)\xi_2 & + & \ldots & + & \alpha_{2n}\xi_n & = & 0, \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1}\xi_1 & + & \alpha_{n2}\xi_2 & + & \ldots & + & (\alpha_{nn} - \lambda)\xi_n & = & 0. \end{array} \right. \tag{8.5.1}$$

Система уравнений (8.5.1) с неизвестными $\{ \lambda, \xi_1, \xi_2, \ldots, \xi_n \}$ нелинейная, но если принять λ за параметр, то относительно неизвестных $\{ \xi_1, \xi_2, \ldots, \xi_n \}$ она линейная и однородная.

Согласно определению 8.5.2 собственный вектор f должен быть *ненулевым.* Покажем, что этого можно добиться путем подбора специальных значений параметра λ.

Действительно, необходимым и достаточным условием существования *ненулевого* частного решения однородной системы n линейных уравнений с n неизвестными, согласно следствию 6.7.2[15], является *равенство нулю определителя ее основной матрицы.*

Поэтому условие, которому должны удовлетворять искомые значения λ, будет иметь вид

$$\det \left\| \, \alpha_{ij} - \lambda \, \delta_{ij} \right\| = 0, \tag{8.5.2}$$

[15] Но не по теореме Крамера!

где δ_{ij} — символ Кронекера, или же

$$\det \begin{Vmatrix} \alpha_{11}-\lambda & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22}-\lambda & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn}-\lambda \end{Vmatrix} = 0\,.$$

Определение 8.5.3	Уравнение (8.5.2) называется *характеристическим уравнением*, а функция от λ, равная $\det\left\|\,\hat{A}-\lambda\,\hat{E}\right\|_g$, — *характеристическим многочленом* преобразования $\hat{A}$, действующего в Λ^n .

Теорема 8.5.2 **Характеристический многочлен линейного преобразования не зависит от выбора базиса в Λ^n.**

Доказательство.

Заметим, что преобразование $\hat{A}-\lambda\,\hat{E}$, очевидно, линейное в силу линейности операторов $\hat{A}$ и $\hat{E}$. Тогда, согласно следствию 8.3.3, определитель его матрицы не меняется при замене базиса.

Поэтому при переходе от базиса $\{\,g_1,\,g_2,\,\dots,\,g_n\}$ к базису $\{\,g'_1,\,g'_2,\,\dots,\,g'_n\}$: $\quad \det\left\|\,\hat{A}-\lambda\,\hat{E}\right\|_g = \det\left\|\,\hat{A}-\lambda\,\hat{E}\right\|_{g'}$.

Теорема доказана.

Характеристическое уравнение является алгебраическим уравнением n-й степени относительно λ, что следует из определения детерминанта 6.1.2 и формулы (8.5.2).

В итоге мы получаем универсальный $\forall\Lambda^n$ алгоритм вычисления собственных значений и соответствующих им собственных векторов:

Решив характеристическое уравнение (8.5.2), из однородной системы уравнений (8.5.1) можно найти собственные векторы, соответствующие последовательно подставляемым в основную матрицу этой системы, найденным собственным значениям.

Примеры использования данного алгоритма в Λ^n иллюстрируют решения задач 8.6.1 и 8.6.2. В случае же линейных пространств, не имеющих базиса, задача отыскания собственных значений и построения собственных векторов может оказаться значительно сложнее.

Например, в линейном пространстве функций, имеющих на некотором интервале производную любого порядка, линейный оператор дифференцирования $\hat{A} = \dfrac{d}{d\tau}$ имеет бесконечно много собственных векторов (функций) $f(\tau) = \alpha e^{\lambda\tau}$ (где α – произвольная ненулевая константа) и собственных значений λ, удовлетворяющих дифференциальному уравнению вида $\dfrac{df}{d\tau} = \lambda f$.

8.6. Свойства собственных векторов и собственных значений

Теорема 8.6.1 **В комплексном линейном пространстве Λ^n всякое линейное преобразование имеет хотя бы один собственный вектор.**

Доказательство.

Поскольку характеристическое уравнение является алгебраическим уравнением n-й степени относительно λ, то к нему применима *основная теорема высшей алгебры*[16], утверждающая, что такое уравнение имеет хотя бы один комплексный корень.

Теорема доказана.

В случае вещественного линейного пространства теорема 8.6.1 неверна. Например, линейный оператор поворота в пространстве плоскости Oxy вокруг оси Oz на угол $\varphi \neq \pi k$ не имеет ни одного собственного вектора.

Действительно, характеристическое уравнение для этого оператора имеет вид (см. § 5.5):

$$\det \left\| \begin{matrix} \cos\varphi - \lambda & -\sin\varphi \\ \sin\varphi & \cos\varphi - \lambda \end{matrix} \right\| = 0 \qquad \text{или} \qquad \lambda^2 - 2\lambda\cos\varphi + 1 = 0\,,$$

то есть $\lambda_{1,2} = \cos\varphi \pm i\sin\varphi$. Откуда следует, что при $\varphi \neq \pi k$ вещественных решений данное характеристическое уравнение не имеет.

Для вещественного линейного конечномерного пространства оказывается справедливой

[16] Доказывается, например, в курсе ТФКП.

Теорема 8.6.2 **В вещественном линейном пространстве Λ^n всякое линейное преобразование имеет либо хотя бы один собственный вектор, либо двумерное инвариантное подпространство.**

Доказательство.

Если характеристическое уравнение имеет вещественный корень, то из системы (8.5.1) находим собственный вектор.

Пусть характеристическое уравнение имеет комплексный корень $\lambda = \alpha + i\beta$ с $\beta \neq 0$. Тогда, решив систему (8.5.1), получим соответствующий ему комплексный собственный вектор $f = u + iv$, где u и v — элементы Λ^n, представляемые вещественными n-компонентными столбцами.

Покажем теперь, что u и v линейно независимые. Допустим противное: $u = \kappa v$. Тогда из соотношения $\hat{A}f = \lambda f$ имеем $\hat{A}(\kappa + i)v = \lambda(\kappa + i)v$ или $\hat{A}v = \lambda v$, откуда следует вещественность λ, что противоречит предположению о невещественности собственного значения.

С другой стороны, опять-таки из $\hat{A}f = \lambda f$ имеем

$$\hat{A}(u + iv) = (\alpha + i\beta)(u + iv)$$

или, в силу линейности $\hat{A}$,

$$\hat{A}u + i\hat{A}v = (\alpha u - \beta v) + i(\beta u + \alpha v)\,.$$

Тогда по определению равенства комплексных чисел получаем, что

$$\begin{cases} \hat{A}u = \alpha u - \beta v, \\ \hat{A}v = \beta u + \alpha v. \end{cases}$$

Но это и означает, что $\hat{A}$ имеет двумерное инвариантное подпространство, совпадающее с двумерной линейной оболочкой элементов u и v, поскольку

$$\hat{A}(\xi u + \eta v) = \xi\hat{A}u + \eta\hat{A}v = \xi(\alpha u - \beta v) + \eta(\beta u + \alpha v) =$$

$$= (\xi\alpha + \eta\beta)u + (\eta\alpha - \xi\beta)v \qquad \forall \xi, \eta \in \mathbb{R}.$$

Теорема доказана.

`Задача 8.6.1` *Найти собственные значения и собственные векторы линейного преобразования $\hat{A}$, действующего в пространстве трехмерных столбцов и заданного матрицей*

$$\left\|\begin{array}{rrr} -1 & -2 & 2 \\ -2 & -1 & 2 \\ -3 & -2 & 3 \end{array}\right\| .$$

`Решение.`1°. Рассмотрим сначала случай, когда $\hat{A}$ действует в комплексном линейном пространстве. Будем искать собственные значения по формулам (8.5.1) – (8.5.2). Характеристическое уравнение имеет вид

$$\det\left\|\begin{array}{lll} -1-\lambda & -2 & 2 \\ -2 & -1-\lambda & 2 \\ -3 & -2 & 3-\lambda \end{array}\right\| = 0 \qquad \Longrightarrow$$

$$\Longrightarrow \qquad (\lambda-1)(\lambda^2+1) = 0 \qquad \Longrightarrow \qquad \lambda_1 = 1,\ \lambda_{2,3} = \pm i\,.$$

То есть из трех собственных значений одно $\lambda_1 = 1$ — вещественное, а два $\lambda_2 = i$ и $\lambda_3 = -i$ — комплексные.

2°. Найдем соответствующие им собственные векторы. Для $\lambda = 1$ по формулам (8.5.2) имеем

$$\left\|\begin{array}{rrr} -2 & -2 & 2 \\ -2 & -2 & 2 \\ -3 & -2 & 2 \end{array}\right\| \left\|\begin{array}{c} \varphi_{(1)1} \\ \varphi_{(1)2} \\ \varphi_{(1)3} \end{array}\right\| = \left\|\begin{array}{c} 0 \\ 0 \\ 0 \end{array}\right\| ,$$

где $\|\ \varphi_{(1)1}\ \varphi_{(1)2}\ \varphi_{(1)3}\ \|^{\mathrm{T}}$ — координатное представление собственного вектора $f_{(1)}$, отвечающего $\lambda = 1$.

Очевидным набором элементарных преобразований приводим полученную систему уравнений к упрощенному виду:

$$\left\|\begin{array}{rrr} 1 & 1 & -1 \\ 1 & 0 & 0 \end{array}\right\| \left\|\begin{array}{c} \varphi_{(1)1} \\ \varphi_{(1)2} \\ \varphi_{(1)3} \end{array}\right\| = \left\|\begin{array}{c} 0 \\ 0 \end{array}\right\| .$$

Откуда получаем, что собственный вектор, отвечающий $\lambda = 1$, представим как

$$\left\|f_{(1)}\right\| = \left\|\begin{array}{c} \varphi_{(1)1} \\ \varphi_{(1)2} \\ \varphi_{(1)3} \end{array}\right\| = \mu \left\|\begin{array}{c} 0 \\ 1 \\ 1 \end{array}\right\| \qquad \forall \mu \neq 0\,.$$

3°. Пусть теперь $\lambda = i$, тогда систему линейных уравнений (8.5.1)

$$\left\| \begin{array}{lll} -1-i & -2 & 2 \\ -2 & -1-i & 2 \\ -3 & -2 & 3-i \end{array} \right\| \left\| \begin{array}{c} \varphi_{(2)1} \\ \varphi_{(2)2} \\ \varphi_{(2)3} \end{array} \right\| = \left\| \begin{array}{c} 0 \\ 0 \\ 0 \end{array} \right\|$$

можно упростить, разделив[17] обе части первого уравнения на $1+i$. Заметим, что в полученной таким образом системе

$$\left\| \begin{array}{lll} -1 & -1+i & 1-i \\ -2 & -1-i & 2 \\ -3 & -2 & 3-i \end{array} \right\| \left\| \begin{array}{c} \varphi_{(2)1} \\ \varphi_{(2)2} \\ \varphi_{(2)3} \end{array} \right\| = \left\| \begin{array}{c} 0 \\ 0 \\ 0 \end{array} \right\|$$

третье уравнение оказывается суммой первых двух и его можно отбросить как линейно зависимое. Заменив затем второе уравнение разностью удвоенного первого и второго, получим

$$\left\| \begin{array}{rrr} -1 & -1+\ i & 1-\ i \\ 0 & -1+3i & -2i \end{array} \right\| \left\| \begin{array}{c} \varphi_{(2)1} \\ \varphi_{(2)2} \\ \varphi_{(2)3} \end{array} \right\| = \left\| \begin{array}{c} 0 \\ 0 \end{array} \right\| .$$

Полагая значение свободного неизвестного ξ_3 равным числу $-1-3i$, находим второй собственный вектор:

$$\left\| f_{(2)} \right\| = \left\| \begin{array}{c} \varphi_{(2)1} \\ \varphi_{(2)2} \\ \varphi_{(2)3} \end{array} \right\| = \mu \left\| \begin{array}{r} 2i \\ 2i \\ -1+3i \end{array} \right\| \quad \forall \mu \neq 0\,.$$

4°. Проведя аналогичные вычисления, найдем, что собственный вектор, отвечающий третьему собственному значению $\lambda_3 = -i$, имеет вид

$$\left\| f_{(3)} \right\| = \left\| \begin{array}{c} \varphi_{(3)1} \\ \varphi_{(3)2} \\ \varphi_{(3)3} \end{array} \right\| = \mu \left\| \begin{array}{r} -2i \\ -2i \\ -1-3i \end{array} \right\| \quad \forall \mu \neq 0\,.$$

[17]Правило деления (равно как и другие действия) для комплексных чисел описано в приложении 3.

(Покажите самостоятельно, что комплексная сопряженность f_2 и f_3 не случайна, то есть если λ_2 и λ_3 комплексно сопряжены, то будут комплексно сопряжены и отвечающие им собственные векторы f_2 и f_3.)

5°. Если оператор $\hat{A}$ действует в вещественном линейном пространстве, то согласно теореме 8.6.2 $\hat{A}$ имеет собственный вектор $\|f_1\| = \|\,0\,1\,1\,\|^{\mathrm{T}}$, отвечающий собственному значению $\lambda_1 = 1$, и инвариантное подпространство, являющееся линейной оболочкой элементов $u = \mathrm{Re} f_2$ и $v = \mathrm{Im} f_2$, то есть которое будет состоять из элементов вида

$$\left\|\begin{matrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{matrix}\right\| = \mu_1 \left\|\begin{matrix} 0 \\ 0 \\ -1 \end{matrix}\right\| + \mu_2 \left\|\begin{matrix} 2 \\ 2 \\ 3 \end{matrix}\right\| \qquad \forall \mu_1,\, \mu_2 \in \mathbb{R}\,.$$

Заметим, что при необходимости искомое инвариантное подпространство может быть задано и как однородная система линейных уравнений, которая в данном примере имеет вид

$$\xi_1 - \xi_2 = 0$$

(см., например, решение задачи 8.4.1).

Решение получено.

Теорема 8.6.3

Совокупность *всех* **собственных векторов, отвечающих некоторому собственному значению линейного преобразования $\hat{A}$, дополненная нулевым элементом линейного пространства Λ, является в Λ** *инвариантным подпространством* **преобразования $\hat{A}$.**

Доказательство.

Пусть $\hat{A} f_1 = \lambda f_1$ и $\hat{A} f_2 = \lambda f_2$. Тогда для любых α и β, не равных нулю одновременно, имеем

$$\hat{A}\big(\alpha f_1 + \beta f_2\big) = \alpha \hat{A} f_1 + \beta \hat{A} f_2 = \alpha\lambda f_1 + \beta\lambda f_2 = \lambda\big(\alpha f_1 + \beta f_2\big)\,,$$

что доказывает справедливость утверждения теоремы.

Теорема доказана.

Определение 8.6.1

Подпространство, состоящее из собственных векторов линейного преобразования $\hat{A}$, отвечающих некоторому собственному значению, дополненных нулевым элементом, называется *инвариантным собственным* (или просто *собственным*) подпространством линейного преобразования $\hat{A}$.

Теорема 8.6.4

Всякое *собственное* подпространство линейного преобразования $\hat{A}$ является также *инвариантным* подпространством линейного преобразования $\hat{B}$, если $\hat{A}$ и $\hat{B}$ *коммутируют*.

Доказательство.

Пусть Λ^* — инвариантное собственное подпространство $\hat{A}$, то есть $\hat{A}f = \lambda f \quad \forall f \in \Lambda^*$. Тогда будет справедливо равенство $\hat{B}\hat{A}f = \hat{B}\bigl(\lambda f\bigr)$, а в силу коммутируемости и линейности $\hat{A}$ и $\hat{B}$ будет также верно и $\hat{A}\bigl(\hat{B}f\bigr) = \lambda\bigl(\hat{B}f\bigr) \quad \forall f \in \Lambda^*$.

Последнее условие означает, что $\hat{B}f \in \Lambda^* \quad \forall f \in \Lambda^*$, то есть Λ^* — инвариантное подпространство оператора $\hat{B}$.

Теорема доказана.

Теорема 8.6.5

Собственные векторы линейного преобразования, отвечающие *различным* собственным значениям, линейно независимы.

Доказательство.

Один собственный вектор линейно независим как ненулевой.

Пусть имеются m линейно независимых собственных векторов $f_1, f_2, \ldots, f_m$ линейного преобразования $\hat{A}$, отвечающих *различным* собственным значениям $\lambda_1, \lambda_2, \ldots, \lambda_m$.

Покажем, что в этом случае будет линейно независим и $m+1$ собственный вектор $f_1, f_2, \ldots, f_m, f_{m+1}$, если эти элементы также отвечают различным собственным значениям.

Предположим противное: существует нетривиальная и равная нулевому элементу линейная комбинация собственных векторов $f_1, f_2, \ldots, f_m, f_{m+1}$, то есть такая, что

$$\kappa_1 f_1 + \kappa_2 f_2 + \ldots + \kappa_m f_m + \kappa_{m+1} f_{m+1} = o\,, \tag{8.6.1}$$

причем без ограничения общности можно считать, что $\kappa_{m+1} \neq 0$.

Подействуем оператором $\hat{A}$ на обе части равенства (8.6.1), получим

$$\hat{A}\left(\kappa_1 f_1 + \kappa_2 f_2 + \ldots + \kappa_m f_m + \kappa_{m+1} f_{m+1}\right) =$$

$$= \kappa_1 \lambda_1 f_1 + \kappa_2 \lambda_2 f_2 + \ldots + \kappa_m \lambda_m f_m + \kappa_{m+1} \lambda_{m+1} f_{m+1} = o\,. \tag{8.6.2}$$

С другой стороны, умножая обе части равенства (8.6.1) на λ_{m+1} и вычитая почленно результат этого умножения из равенства (8.6.2), получаем

$$\kappa_1\left(\lambda_1 - \lambda_{m+1}\right) f_1 + \kappa_2\left(\lambda_2 - \lambda_{m+1}\right) f_2 + \ldots$$

$$\ldots + \kappa_m\left(\lambda_m - \lambda_{m+1}\right) f_m = o\,.$$

Поскольку все собственные значения разные, а векторы $f_1, f_2, \ldots, f_m$ линейно независимые, то мы приходим к заключению, что $\kappa_1 = \kappa_2 = \ldots = \kappa_m = 0$.

Но тогда из (8.6.1) следует $\kappa_{m+1} = 0$, что противоречит сделанному выше предположению, и по принципу математической индукции из линейной независимости элементов $f_1, f_2, \ldots, f_m$ следует линейная независимость элементов $f_1, f_2, \ldots, f_m, f_{m+1}$.

Теорема доказана.

Следствие 8.6.1 **Линейное преобразование $\hat{A}$ в Λ^n может иметь (с точностью до произвольного ненулевого множителя) не более чем n собственных векторов, отвечающих различным собственным значениям.**

Теорема 8.6.6 **Если у линейного преобразования $\hat{A}$, действующего в Λ^n, есть n различных собственных значений, то существует базис, образованный собственными векторами, в котором матрица данного линейного оператора имеет диагональный вид, причем на ее главной диагонали расположены собственные числа оператора $\hat{A}$.**

Доказательство.

Следует из теоремы 8.6.5 и замечания о важности собственных векторов § 8.5.

Теорема доказана.

Теорема 8.6.7 **Пусть Λ^* — собственное подпространство в Λ^n линейного преобразования $\hat{A}$, отвечающее некоторому собственному значению λ_0** *кратности* k**. Тогда** $1 \leq \dim \Lambda^* \leq k$.

Доказательство.

Выберем в Λ^n базис $\{ g_1, g_2, \ldots, g_m, g_{m+1}, \ldots, g_{n-1}, g_n \}$ так, чтобы его первые $m = \dim \Lambda^*$ элементов принадлежали подпространству Λ^*.

В силу условия кратности собственного значения λ_0

$$\hat{A} g_i = \lambda_0 g_i \qquad \forall i = [1, m] .$$

Поэтому матрица $\left\| \hat{A} - \lambda \hat{E} \right\|_g$ в этом базисе согласно определению 8.3.1 (см. § 8.3) будет иметь вид

$$\left\| \hat{A} - \lambda \hat{E} \right\|_g =$$

$$= \left\| \begin{array}{ccccccc} \lambda_0 - \lambda & 0 & \ldots & 0 & \alpha_{1,m+1} & \ldots & \alpha_{1n} \\ 0 & \lambda_0 - \lambda & \ldots & 0 & \alpha_{2,m+1} & \ldots & \alpha_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_0 - \lambda & \alpha_{m,m+1} & \ldots & \alpha_{mn} \\ 0 & 0 & \ldots & 0 & \alpha_{m+1,m+1} - \lambda & \ldots & \alpha_{m+1,n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & \alpha_{n,m+1} & \ldots & \alpha_{nn} - \lambda \end{array} \right\| .$$

Откуда следует, что $\qquad \det \left\| \hat{A} - \lambda \hat{E} \right\|_g = (\lambda_0 - \lambda)^m P_{n-m}(\lambda).$

При этом множители вида $(\lambda_0 - \lambda)$ могут содержаться также и в многочлене $P_{n-m}(\lambda)$, значит, $m \leq k$, где k — кратность λ_0, корня характеристического многочлена $\det \left\| \hat{A} - \lambda \hat{E} \right\|_g$.

Условие $1 \leq \dim \Lambda^*$ очевидно верное, поскольку подпространство Λ^* содержит ненулевые собственные векторы.

Теорема доказана.

Таким образом, размерность собственного подпространства Λ^*, отвечающего собственному значению λ кратности k, может оказаться меньше k, что иллюстрирует

Задача 8.6.2 *Найти собственные значения и собственные векторы линейного преобразования, в пространстве двумерных столбцов, заданного матрицей* $\left\| \begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix} \right\|$.

Решение. Находим собственные значения:

$$\det \left\| \begin{matrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{matrix} \right\| = (1-\lambda)^2 = 0\,,$$

то есть $\quad \lambda_{1,2} = 1, \quad$ и кратность этого собственного значения 2.

Найдем теперь собственные векторы:

$$\left\| \begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix} \right\| \left\| \begin{matrix} \varphi_1 \\ \varphi_2 \end{matrix} \right\| = \left\| \begin{matrix} 0 \\ 0 \end{matrix} \right\| \quad \Longrightarrow \quad \|f\| = \mu \left\| \begin{matrix} 1 \\ 0 \end{matrix} \right\| \; \forall \mu \neq 0\,.$$

Таким образом, получаем, что данный линейный оператор имеет одномерное собственное подпространство, отвечающее собственному значению $\lambda = 1$ кратности 2.

Решение получено.

На основании теорем 8.6.2, 8.6.5 и 8.6.6 приходим к выводу, что базис в конечномерном вещественном линейном пространстве, образованный из собственных векторов действующего в нем линейного преобразования, может не существовать из-за *невещественности* и/или *кратности* его собственных значений.

Теорема 8.6.8 **Линейное преобразование $\hat{A}$ в Λ^n имеет нулевое собственное значение тогда и только тогда, когда оно не является взаимно однозначным.**

Доказательство.

Линейное преобразование $\hat{A}$ имеет в Λ^n собственное значение, равное нулю, тогда и только тогда, когда его матрица вырожденная, то есть в любом базисе $\det \left\| \hat{A} \right\|_g = 0$.

Пусть в Λ^n координатный столбец образа связан с координатным столбцом прообраза

$$\left\| \begin{array}{c} \eta_1 \\ \eta_2 \\ \dots \\ \eta_n \end{array} \right\| = \left\| \begin{array}{cccc} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{array} \right\| \left\| \begin{array}{c} \xi_1 \\ \xi_2 \\ \dots \\ \xi_n \end{array} \right\| .$$

Тогда из теоремы 6.4.1 (Крамера) следует, что для заданного координатного столбца элемента-образа $\left\| \eta_1\, \eta_2\, \dots\, \eta_n \right\|^{\mathrm{T}}$ эта система линейных уравнений, у которой неизвестными являются компоненты столбца элемента-прообраза $\left\| \xi_1\, \xi_2\, \dots\, \xi_n \right\|^{\mathrm{T}}$, либо оказывается *несовместной* (элемент-прообраз не существует), либо будет иметь согласно следствию 6.7.1 *неединственное* решение (элемент-прообраз определяется неоднозначно).

Теорема доказана.

Определение 8.6.2	Степенью порядка $k \geq 2$ квадратной матрицы $\|Q\|$, обозначаемую как $\|Q\|^k$, называется произведение следующего вида: $\underbrace{\|Q\| \cdot \|Q\| \cdot \ldots \cdot \|Q\|}_{k \text{ сомножителей}}$. Будем также считать, что $\|Q\|^1 = \|Q\|$ и $\|Q\|^0 = \|E\|$.

Теорема 8.6.9 (Гамильтона - Кэли) **Матрица линейного преобразования $\hat{A}$ в Λ^n является корнем характеристического уравнения этого преобразования.**

Доказательство.

Докажем теорему в предположении, что собственные векторы преобразования $\hat{A}$ образуют в Λ^n базис $\{ f_1, f_2, \dots, f_n \}$.

Пусть данное линейное преобразование в этом базисе имеет матрицу $\left\|\hat{A}\right\|_f$ и характеристическое уравнение вида $\sum\limits_{k=1}^{n} \alpha_k \lambda^k = 0$. Тогда в силу линейности $\hat{A}$ для собственного вектора f, отвечающего собственному значению λ, имеем (см. задачу 8.5.2):

$$\left(\sum_{k=1}^{n}\alpha_k\left\|\hat{A}\right\|_f^k\right)\|f\| = \sum_{k=1}^{n}\alpha_k\left(\left\|\hat{A}\right\|_f^k\|f\|\right) =$$

$$= \sum_{k=1}^{n}\alpha_k\left(\left\|\hat{A}\right\|_f\left(\left\|\hat{A}\right\|_f\ldots\left(\left\|\hat{A}\right\|_f\|f\|\right)\ldots\right)\right) =$$

$$= \sum_{k=1}^{n}\alpha_k\left(\lambda^k\|f\|\right) = \left(\sum_{k=1}^{n}\alpha_k\lambda^k\right)\|f\| = 0\cdot\|f\| = \|o\|\,.$$

Но поскольку это соотношение верно для всех базисных векторов, то оно будет верно и для каждого элемента $x \in \Lambda^n$. Тогда из леммы 5.1.2 следует, что

$$\sum_{k=1}^{n}\alpha_k\left\|\hat{A}\right\|_f^k = \left\|\hat{O}\right\|_f\,.$$

Наконец, выполнив переход (с матрицей перехода $\|S\|$) к произвольному базису $\{\,g_1,\,g_2,\,\ldots,\,g_n\,\}$, получим

$$\sum_{k=1}^{n}\alpha_k\left\|\hat{A}\right\|_g^k = \sum_{k=1}^{n}\alpha_k\left(\|S\|^{-1}\left\|\hat{A}\right\|_f\|S\|\right)^k =$$

$$= \sum_{k=1}^{n}\alpha_k\left(\|S\|^{-1}\left\|\hat{A}\right\|_f\|S\|\cdot\|S\|^{-1}\left\|\hat{A}\right\|_f\|S\|\ldots\|S\|^{-1}\left\|\hat{A}\right\|_f\|S\|\right) =$$

$$= \sum_{k=1}^{n}\alpha_k\left(\|S\|^{-1}\left\|\hat{A}\right\|_f^k\|S\|\right) = \|S\|^{-1}\left(\sum_{k=1}^{n}\alpha_k\left\|\hat{A}\right\|_f^k\right)\|S\| =$$

$$= \|S\|^{-1}\left\|\hat{O}\right\|_f\|S\| = \left\|\hat{O}\right\|_g\,.$$

Теорема доказана.

8.7. Линейные функционалы

Рассмотрим специальный случай линейного оператора, когда его область значений содержится в одномерном линейном пространстве, изоморфном множеству вещественных чисел.

Такого рода зависимости, следуя классификации, введенной в § 5.2, следует относить к функционалам. Напомним данное ранее

Определение 8.7.1 Пусть *каждому* элементу x линейного пространства Λ поставлено в соответствие *однозначно* определяемое число, обозначаемое $f(x)$. Тогда говорят, что в Λ задан функционал $f(x)$.

Пример 8.7.1

1°. В пространстве n-компонентных столбцов можно задать функционал, сопоставив столбцу $\| \, \xi_1 \, \xi_2 \, \ldots, \, \xi_n \, \|^{\mathrm{T}}$ число $\sum\limits_{i=1}^{n} \varphi_i \xi_i$, где $\varphi_i \quad i = [1, n]$ — некоторые фиксированные константы.

2°. В векторном геометрическом пространстве функционалом является длина вектора, то есть $f(\vec{r}\,) = |\vec{r}\,|$.

3°. В пространстве функций $x(\tau)$, определенных на $[-1, 1]$, функционалом является так называемая «дельта-функция Дирака», обозначаемая как $\delta(x(\tau))$, ставящая в соответствие каждой функции $x(\tau)$ ее значение в нуле, то есть число $x(0)$.

4°. В пространстве функций $x(\tau)$, непрерывных на (α, β), функционалом является

$$f\bigl(x(\tau)\bigr) = \int\limits_{\alpha}^{\beta} p(\tau) x(\tau) \, d\tau \, ,$$

где $p(\tau)$ — некоторая непрерывная на (α, β) функция.

5°. В линейном пространстве квадратных матриц вида $\|A\| = \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\|$ функционалом является

$$f(\|A\|) = \det \left\| \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{matrix} \right\| .$$

Определение 8.7.2 Функционал $f(x)$ называется *линейным* функционалом (или линейной формой), если $\forall x, y \in$ и $\forall \lambda, \mu \in \mathbb{R}$

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) \, .$$

Задача 8.7.1 *Доказать, что в примере 8.7.1 функционалы 1°, 3° и 4° являются линейными, а функционалы 2° и 5° – нет.*

Представление линейного функционала в Λ^n

Пусть в Λ^n дан базис $\{ g_1, g_2, \ldots, g_n \}$ и пусть элемент $x \in \Lambda^n$ имеет координатное разложение $x = \sum\limits_{i=1}^{n} \xi_i g_i$ и координатное представление $\left\| \, \xi_1 \, \xi_2 \, \ldots \, \xi_n \, \right\|^{\mathrm{T}}$. Тогда для каждого линейного функционала $f(x)$ справедливы соотношения

$$f(x) = f\left(\sum_{i=1}^{n} \xi_i g_i\right) = \sum_{i=1}^{n} \xi_i f(g_i) = \sum_{i=1}^{n} \varphi_i \xi_i \,.$$

Числа $\varphi_i = f(g_i) \quad \forall i = [1, n]$ принято называть *компонентами* (или *координатами*) линейного функционала $f(x)$ в данном базисе $\{ g_1, g_2, \ldots, g_n \}$.

Из этих равенств следует легко проверяемая

Теорема 8.7.1 **Каждый линейный функционал $f(x)$ в Λ^n в конкретном базисе $\{ g_1, g_2, \ldots, g_n \}$ имеет однозначно определяемую строку компонентов**

$$\|f\|_g = \left\| \, \varphi_1 \, \varphi_2 \, \ldots \, \varphi_n \, \right\| ,$$

а каждая строка $\left\| \, \varphi_1 \, \varphi_2 \, \ldots \, \varphi_n \, \right\|$ в этом базисе определяет линейный функционал $f(x) = \sum\limits_{i=1}^{n} \varphi_i \xi_i$, или в матричном виде $f(x) = \|f\|_g \|x\|_g$.

Запись координатного представления линейного функционала в Λ^n в виде строки (а не столбца!) обеспечивает соответствие этого представления определению 8.4.5, поскольку линейный функционал в Λ^n можно рассматривать как *линейное отображение* вида $\Lambda^n \longrightarrow \Lambda^1$, имеющего матрицу размета $1 \times n$.

Получим теперь правило изменения компонент линейного функционала в Λ^n при переходе от одного базиса к другому. Пусть в Λ^n даны два базиса $\{ g_1, g_2, \ldots, g_n \}$ и $\{ g'_1, g'_2, \ldots, g'_n \}$, связанные матрицей перехода $\|S\| = \| \, \sigma_{ij} \|$, где $g'_j = \sum\limits_{i=1}^{n} \sigma_{ij} g_i \quad \forall j = [1, n]$.

Координатные представления некоторого элемента будут иметь в рассматриваемых базисах вид $x = \sum\limits_{i=1}^{n} \xi_i g_i = \sum\limits_{j=1}^{n} \xi'_j g'_j$, а координатные представления линейного функционала $f(x)$ — соответственно

$$f(x) = \sum_{i=1}^{n} \varphi_i \xi_i = \sum_{j=1}^{n} \varphi'_j \xi'_j \,.$$

Найдем выражения для величин φ'_j через φ_i. Во введенных обозначениях в силу линейности функционала получаем

$$\varphi'_j = f(g'_j) = f\left(\sum_{i=1}^{n} \sigma_{ij} g_i\right) = \sum_{i=1}^{n} \sigma_{ij} f(g_i) = \sum_{i=1}^{n} \sigma_{ij} \varphi_i \qquad \forall j = [1, n]\,,$$

что доказывает справедливость следующего утверждения.

Теорема 8.7.2 **В Λ^n в базисах $\{\, g_1, g_2, \ldots, g_n \,\}$ и $\{\, g'_1, g'_2, \ldots, g'_n \,\}$ компоненты координатных представлений линейного функционала**

$$\|f\|_g = \left\| \, \varphi_1 \, \varphi_2 \, \ldots \, \varphi_n \, \right\| \quad \text{и} \quad \|f\|_{g'} = \left\| \, \varphi'_1 \, \varphi'_2 \, \ldots \, \varphi'_n \, \right\|$$

связаны соотношением $\varphi'_j = \sum\limits_{i=1}^{n} \sigma_{ij} \varphi_i \quad \forall j = [1, n]$, **где коэффициенты σ_{ij} суть элементы матрицы $\|S\|$ — перехода от первого базиса ко второму.**

В матричной форме это утверждение имеет вид $\|f\|_{g'} = \|f\|_g \|S\|$.

Проверьте самостоятельно, что компоненты линейного функционала в Λ^n преобразуются при замене базиса так же, как преобразуются столбцы, составленные из базисных элементов (см. § 7.3).

Сопряженное (двойственное) пространство. Взаимный (биортогональный) базис

Поскольку линейные функционалы являются частным случаем линейных операторов, то для них можно ввести операции сравнения, сложения и умножения на число.

Задача 8.7.2 *Показать, что в Λ^n с базисом $\{ g_1, g_2, \dots, g_n \}$ операции сложения и умножения на число для линейных функционалов $p(x)$ и $q(x)$ в координатном представлении имеют вид*

$$\| p+q \|_g = \| \varphi_1+\psi_1 \;\; \varphi_2+\psi_2 \;\; \dots \;\; \varphi_n+\psi_n \|$$

и

$$\| \lambda p \|_g = \| \lambda\varphi_1 \;\; \lambda\varphi_2 \;\; \dots \;\; \lambda\varphi_n \| ,$$

где $\|p\|_g = \Big\| \varphi_1 \, \varphi_2 \dots \varphi_n \Big\|$ *и* $\|q\|_g = \Big\| \psi_1 \, \psi_2 \dots \psi_n \Big\|$.

Нетрудно убедиться, что для множества линейных функционалов будут справедливы все утверждения § 8.2, в том числе и

Теорема 8.7.3 **Множество всех линейных функционалов, заданных в линейном пространстве Λ , является линейным пространством.**

Определение 8.7.3	Линейное пространство линейных функционалов, заданных в Λ, называется *сопряженным* (или *двойственным*) пространству Λ и обозначается Λ^+.

Теорема 8.7.1 устанавливает взаимно однозначное соответствие между множествами линейных функционалов в Λ^n и n-компонентных строк, последнее из которых является линейным n-мерным пространством.

Принимая во внимание, что операции с линейными функционалами в координатном представлении в Λ^n совпадают с аналогичными операциями для n-компонентных строк, можно прийти к заключению об *изоморфности* линейных пространств Λ^n и Λ^{n+}. Поэтому будет справедлива

Теорема 8.7.4 **Размерность пространства Λ^{n+}, сопряженного Λ^n, равна n.**

Как и во всяком n-мерном линейном пространстве, в Λ^{n+} существует базис. Пусть он состоит из элементов $\{ r_1, r_2, \dots, r_n \}$. Тогда каждый элемент $f \in \Lambda^{n+}$ может быть однозначно представлен в виде линейной комбинации базисных элементов, то есть $f = \sum\limits_{i=1}^{n} \rho_i r_i$, а стандартное для n-мерного пространства *столбцовое* координатное представление этого элемента будет иметь вид $\quad \|f\|_r = \left\| \, \rho_1 \, \rho_2 \, \ldots \, \rho_n \, \right\|^{\mathrm{T}}$.

Связь между координатными представлениями линейного функционала в базисах $\{ g_1, g_2, \ldots, g_n \} \subset \Lambda^n$ и $\{ r_1, r_2, \ldots, r_n \} \subset \Lambda^{n+}$ задается квадратной, порядка n, матрицей $\|\Gamma\|_{rg}$, составленной из чисел $\gamma_{ij} = r_i(g_j) \quad i, j = [1, n]$, являющихся значениями функционала $r_i(x)$ на элементах g_j.

Задача 8.7.3	*Доказать, что если* $\{ g_1, g_2, \ldots, g_n \}$ *— базис в* Λ^n*, а* $\{ r_1, r_2, \ldots, r_n \}$ *– базис в* Λ^{n+}*, то* $\quad \|f\|_g = \|f\|_r^{\mathrm{T}} \|\Gamma\|_{rg}$.

Определение 8.7.4	Если матрица $\|\Gamma\|_{rg}$ единичная, то есть $\|\Gamma\|_{rg} = \|E\|$, то базисы $\{ g_1, g_2, \ldots, g_n \}$ и $\{ r_1, r_2, \ldots, r_n \}$ называются *взаимными* (*биортогональными*).

Отметим, что во взаимной паре базисов для любого линейного функционала $f(x)$ его координатные представления в Λ^n и в Λ^{n+} связаны очевидным соотношением $\quad \|f\|_r = \|f\|_g^{\mathrm{T}}$.

Вторичное сопряженное (вторичное двойственное) пространство

Поскольку Λ^{n+} является n-мерным линейным пространством, то в нем так же, как и в Λ^n, возможно определять линейные функционалы и рассматривать их совокупность как новое линейное пространство Λ^{n++}, сопряженное к Λ^{n+}. Будем называть это пространство *вторичным сопряженным* для линейного пространства Λ^n.

Очевидно, что линейные пространства Λ^n, Λ^{n+} и Λ^{n++} n-мерные и, следовательно, изоморфны друг другу. Однако, как мы увидим, для пространств Λ^n и Λ^{n++} существует особый изоморфизм, позволяющий не делать различия между этими пространствами.

Пусть x — некоторый элемент из Λ^n, а $X(f)$ — действующий в Λ^{n+} функционал, такой, что $X(f) = f(x) \quad \forall f \in \Lambda^{n+}$.

Убедимся вначале, что $X(f)$ линейный на Λ^{n+}, то есть он будет некоторым элементом в Λ^{n++}.

Действительно,

$$X(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 f_1(x) + \lambda_2 f_2(x) = \lambda_1 X(f_1) + \lambda_2 X(f_2)$$

$$\forall \lambda_1, \lambda_2 \in \mathbb{R}, \qquad \forall f_1, f_2 \in \Lambda^{n+} .$$

Это означает, что $X(f) \in \Omega \quad \forall f \in \Lambda^{n+}$, где, согласно теореме 8.4.1, множество Ω есть подпространство линейного пространства Λ^{n++}.

Теперь рассмотрим отображение $\Lambda^n \longrightarrow \Omega \quad \forall x \in \Lambda^n, \ \forall y \in \Omega$ вида $y = X(f(x))$.

Оно будет линейным, как произведение (композиция) линейных отображений $f(x)$ и $X(f)$, и взаимно однозначным. Следовательно, это отображение устанавливает изоморфизм линейного пространства Λ^n и множества Ω, что в силу теоремы 7.5.1 дает $\dim \Omega = \dim \Lambda^n = n$.

Наконец, отметим, что сочетание условий

$$\dim \Lambda^{n++} = n = \dim \Omega \qquad \text{и} \qquad \Omega \subseteq \Lambda^{n++}$$

означает совпадение множества Ω и линейного пространства Λ^{n++}.

Таким образом, в итоге мы приходим к заключению, что отображение $y = X(f(x)) \quad \forall x \in \Lambda^n, \ \forall y \in \Lambda^{n++}$ устанавливает *тождественное* взаимно однозначное соответствие между элементами линейных пространств Λ^n и Λ^{n++}, позволяющее считать их *одним и тем же* пространством Λ^n.

Это также позволяет записывать связь между значениями линейных функционалов, действующих в Λ^n и Λ^{n+}, в следующем симметричном виде: $f(x) = x(f)$.

9. Нелинейные зависимости в линейном пространстве

9.1. Билинейные функционалы

Определение 9.1.1 Пусть в линейном пространстве Λ каждой упорядоченной паре элементов x и y поставлено в соответствие число $B(x, y)$ так, что

$$1)\quad B(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1 B(x_1, y) + \alpha_2 B(x_2, y)\,,$$

$$2)\quad B(x, \beta_1 y_1 + \beta_2 y_2) = \beta_1 B(x, y_1) + \beta_2 B(x, y_2)$$

$$\forall x, x_1, x_2, y, y_1, y_2 \in \Lambda, \qquad \forall \alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R},$$

тогда говорят, что в Λ задан *билинейный функционал* (или *билинейная форма*).

Пример 9.1.1

1°. Произведение двух линейных функционалов $F(x)$ и $G(y)$, определенных в Λ, $\quad B(x, y) = F(x)G(y) \quad$ есть билинейный функционал.

2°. Двойной интеграл

$$B(x, y) = \iint\limits_{\Omega} K(\sigma, \tau) x(\sigma) y(\tau)\, d\sigma d\tau =$$

$$= \int\limits_{\alpha}^{\beta} x(\sigma) \left(\int\limits_{\gamma}^{\delta} K(\sigma, \tau) y(\tau)\, d\tau \right) d\sigma\,,$$

где функция двух переменных $K(\sigma, \tau)$ непрерывна на прямоугольнике $\{\Omega: \ \alpha \leq \sigma \leq \beta, \ \gamma \leq \tau \leq \delta\}$, есть билинейный функционал в линейном пространстве непрерывных функций $x(\sigma)$ и $y(\tau)$.

3°. Билинейным функционалом является *скалярное произведение* векторов на плоскости или в пространстве.

Билинейные функционалы в Λ^n

Пусть в Λ^n заданы базис $\{g_1, g_2, \ldots, g_n\}$ и билинейный функционал $B(x, y)$. Найдем формулу, выражающую его значение через координаты аргументов.

Предположим, что в данном базисе $x = \sum\limits_{i=1}^{n} \xi_i g_i$ и $y = \sum\limits_{j=1}^{n} \eta_j g_j$, тогда, согласно определению 9.1.1, справедливы равенства

$$B(x,y) = B\left(\sum_{i=1}^{n} \xi_i g_i\,, \sum_{j=1}^{n} \eta_j g_j\right) = \sum_{i=1}^{n} \xi_i B\left(g_i\,, \sum_{j=1}^{n} \eta_j g_j\right) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \xi_i \eta_j B\left(g_i, g_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \beta_{ij} \xi_i \eta_j\,. \qquad (9.1.1)$$

Определение 9.1.2	Числа $\beta_{ij} = B(g_i, g_j) \quad \forall i, j = [1, n]$ называются *компонентами билинейного функционала* $B(x, y)$ в базисе $\{g_1, g_2, \ldots, g_n\}$, а матрица $\|B\|_g = \|\beta_{ij}\|$ — *матрицей билинейного функционала* в этом базисе.

В Λ^n с базисом $\{g_1, g_2, \ldots, g_n\}$ билинейный функционал может быть представлен в виде

$$B(x,y) = \sum_{i=1}^{n}\sum_{j=1}^{n} \beta_{ij}\xi_i\eta_j = \sum_{i=1}^{n}\sum_{j=1}^{n} \xi_{i1}\beta_{ij}\eta_{j1} = \sum_{i=1}^{n} \xi_{1i}^{\mathrm{T}} \sum_{j=1}^{n} \beta_{ij}\eta_{j1} =$$

$$= \|\,\xi_1\,\xi_2\,\ldots\,\xi_n\,\| \left\| \begin{matrix} \beta_{11} & \beta_{12} & \ldots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \ldots & \beta_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \beta_{n1} & \beta_{n2} & \ldots & \beta_{nn} \end{matrix} \right\| \left\| \begin{matrix} \eta_1 \\ \eta_2 \\ \ldots \\ \eta_n \end{matrix} \right\| =$$

$$= \|x\|_g^{\mathrm{T}} \|B\|_g \|y\|_g\,,$$

где столбцы $\|x\|_g$ и $\|y\|_g$ — координатные представления элементов x и y в данном базисе.

Матрица билинейного функционала зависит от выбора базиса. Правило изменения матрицы билинейного функционала при замене базиса дает

Теорема 9.1.1 **Пусть $\|S\|$ — матрица перехода от базиса $\{\, g_1,\, g_2,\, \ldots,\, g_n \,\}$ к базису $\{\, g'_1,\, g'_2,\, \ldots,\, g'_n \,\}$, тогда**

$$\|B\|_{g'} = \|S\|^{\mathrm{T}} \|B\|_g \|S\| \,.$$

Доказательство.

По определению матрицы перехода от одного базиса к другому в пространстве Λ^n (см. § 7.3) имеют место соотношения $g'_j = \sum\limits_{i=1}^{n} \sigma_{ij} g_i \quad \forall j = [1, n]\,,$ но тогда

$$\beta'_{pq} = B\big(g'_p, g'_q\big) = B\left(\sum_{i=1}^{n} \sigma_{ip} g_i, \sum_{j=1}^{n} \sigma_{jq} g_j\right) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{ip}\sigma_{jq} B\big(g_i, g_j\big) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{ip}\sigma_{jq}\beta_{ij} = \sum_{i=1}^{n} \sigma^{\mathrm{T}}_{pi} \sum_{j=1}^{n} \beta_{ij}\sigma_{jq} \quad \forall p, q = [1, n].$$

Теорема доказана.

Следствие 9.1.1 $\det \|B\|_{g'} = \det \|B\|_g \det^2 \|S\| \,.$

Доказательство.

Вытекает из теоремы 9.1.1, а также свойств детерминанта матрицы (см. теоремы 6.2.1 и 6.2.4).

Следствие доказано.

Отметим, что в силу невырожденности матрицы перехода *знак определителя* матрицы билинейного функционала не зависит от выбора базиса.

Следствие 9.1.2 *Ранг* **матрицы билинейного функционала не зависит от выбора базиса.**

Доказательство.

Следует из теоремы 8.4.3 и невырожденности матрицы перехода $\|S\|$.

Следствие доказано.

Определение 9.1.3	Билинейный функционал называется *симметричным*, если для любой упорядоченной пары элементов x и y линейного пространства Λ имеет место равенство $B(y,x) = B(x,y)$.

Теорема 9.1.2 **Для симметричности билинейного функционала в Λ^n необходимо и достаточно, чтобы его матрица была** *симметрической***.**

Доказательство.

Необходимость следует из соотношений

$$\beta_{ji} = B(g_j, g_i) = B(g_i, g_j) = \beta_{ij} \qquad \forall i, j = [1, n]\,.$$

Докажем достаточность.

Действительно, если $\beta_{ji} = \beta_{ij} \qquad \forall i, j = [1, n]\,,$ то

$$B(y,x) = \sum_{i=1}^{n}\sum_{j=1}^{n}\beta_{ji}\eta_j\xi_i = \sum_{i=1}^{n}\sum_{j=1}^{n}\beta_{ji}\xi_i\eta_j =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\beta_{ij}\xi_i\eta_j = B(x,y)\,.$$

Теорема доказана.

9.2. Квадратичные функционалы

Определение 9.2.1	Пусть в линейном пространстве Λ каждому элементу x поставлено в соответствие число $\Phi(x) = B(x,x)$, где $B(x,y)$ — некоторый билинейный функционал в Λ, тогда говорят, что в Λ задан *квадратичный функционал* (или *квадратичная форма*).

В общем случае в вещественном линейном пространстве по заданному квадратичному функционалу нельзя восстановить порождающий его билинейный функционал, однако это можно сделать в случае симметричного билинейного функционала.

Действительно, пусть квадратичный функционал $\Phi(x)$ порожден симметричным билинейным функционалом $B(x,y)$, тогда для любых x и y имеют место равенства

$$\Phi(x+y) = B(x+y, x+y) = B(x,x) + B(x,y) + B(y,x) + B(y,y) =$$

$$= B(x,x) + 2B(x,y) + B(y,y) = \Phi(x) + 2B(x,y) + \Phi(y)\,.$$

Откуда $\quad B(x,y) = \dfrac{\Phi(x+y) - \Phi(x) - \Phi(y)}{2}$.

Определение 9.2.2	В Λ^n симметрическая матрица билинейного функционала $\dfrac{\Phi(x+y) - \Phi(x) - \Phi(y)}{2}$ называется *матрицей квадратичного функционала* $\Phi(x)$.

Пусть матрица квадратичного функционала $\|\Phi\|_g = \|\varphi_{ij}\|$. Тогда из свойств билинейного функционала следует, что если в Λ^n задан базис $\{g_1,\, g_2,\, \ldots,\, g_n\}$, то квадратичный функционал может быть представлен в виде

$$\Phi(x) = \sum_{i=1}^{n}\sum_{j=1}^{n}\varphi_{ij}\xi_i\xi_j =$$

$$= \|\,\xi_1\,\xi_2\,\ldots\,\xi_n\,\| \left\| \begin{matrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1n} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_{n1} & \varphi_{n2} & \cdots & \varphi_{nn} \end{matrix} \right\| \left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \cdots \\ \xi_n \end{matrix} \right\| =$$

$$= \|x\|_g^{\mathrm{T}}\|\Phi\|_g\|x\|_g\,,$$

где столбец $\|x\|_g$ — координатное представление элемента $x = \sum\limits_{i=1}^{n} \xi_i g_i$ в данном базисе. Замена базиса, в свою очередь, приводит к изменению матрицы квадратичного функционала по формуле

$$\|\Phi\|_{g'} = \|S\|^{\mathrm{T}} \|\Phi\|_g \|S\| \,,$$

следующей из теоремы 9.1.1.

Отметим, что иногда целесообразно строить квадратичный функционал $\Phi(x)$ по порождающему билинейному функционалу, просимметрировав предварительно последний.

Действительно, для любого $B(x,y)$ можно указать симметричный билинейный функционал вида $\dfrac{B(x,y)+B(y,x)}{2}$, который будет порождать *тот же самый* квадратичный функционал $\Phi(x)$, что и $B(x,y)$. В этом случае очевидно, что

$$\varphi_{ij} = \frac{\beta_{ij}+\beta_{ji}}{2} = \frac{\beta_{ji}+\beta_{ij}}{2} = \varphi_{ji} \qquad \forall i,j = [1,n]$$

— элементы симметрической матрицы.

Пример 9.2.1 Пусть в Λ^3 задан билинейный функционал
$B_1(x,y) = \xi_1\eta_1 + 3\xi_2\eta_2 - \xi_2\eta_1 - 3\xi_1\eta_2 + 2\xi_3\eta_1 - \xi_2\eta_3 - \xi_3\eta_2 =$

$$= \|\, \xi_1\ \xi_2\ \xi_3 \,\| \left\| \begin{array}{rrr} 1 & -3 & 0 \\ -1 & 3 & -1 \\ 2 & -1 & 0 \end{array} \right\| \left\| \begin{array}{c} \eta_1 \\ \eta_2 \\ \eta_3 \end{array} \right\| ,$$

имеющий несимметрическую матрицу и, в силу теоремы 9.1.2, не являющийся симметричным. Порождаемый им в квадратичный функционал будет иметь вид

$$\Phi_1(x) = \xi_1^2 + 3\xi_2^2 - 4\xi_1\xi_2 + 2\xi_1\xi_3 - 2\xi_2\xi_3 \,.$$

В то же время просимметрованный билинейный функционал

$$B_2(x,y) = \frac{B_1(x,y)+B_1(y,x)}{2} =$$

$$= \xi_1\eta_1 + 3\xi_2\eta_2 - 2\xi_2\eta_1 - 2\xi_1\eta_2 + \xi_3\eta_1 + \xi_1\eta_3 - \xi_2\eta_3 - \xi_3\eta_2 =$$

$$= \|\, \xi_1\ \xi_2\ \xi_3 \,\| \left\| \begin{array}{rrr} 1 & -2 & 1 \\ -2 & 3 & -1 \\ 1 & -1 & 0 \end{array} \right\| \left\| \begin{array}{c} \eta_1 \\ \eta_2 \\ \eta_3 \end{array} \right\| ,$$

имеющий симметрическую матрицу и порождающий в Λ^3 квадратичный функционал вида

$$\Phi_2(x) = \xi_1^2 + 3\xi_2^2 - 4\xi_1\xi_2 + 2\xi_1\xi_3 - 2\xi_2\xi_3\,,$$

который совпадает с $\Phi_1(x)$ и имеет симметрическую матрицу

$$\|\Phi_2\|_g = \left\| \begin{array}{rrr} 1 & -2 & 1 \\ -2 & 3 & -1 \\ 1 & -1 & 0 \end{array} \right\| = \|\Phi_1\|_g\,.$$

На практике нередко оказывается необходимым отыскание базисов, в которых квадратичный функционал имеет наиболее простой и удобный для исследования вид. Понятие *простой* уточняет

Определение 9.2.3	Квадратичный функционал $\Phi(x)$ имеет *диагональный* вид в базисе $\{g_1,\, g_2,\, \ldots,\, g_n\} \subset \Lambda^n$, если он в этом базисе представим как $$\Phi(x) = \sum_{i=1}^{n} \lambda_i \xi_i^2\,,$$ где $\lambda_i \quad \forall i = [1,n]$ — некоторые числа. Если, кроме того, числа $\lambda_i \quad \forall i = [1,n]$ принимают лишь значения 0 или ± 1, то говорят, что квадратичный функционал в данном базисе имеет *канонический* вид.

Теорема 9.2.1 (Метод Лагранжа) **Для каждого квадратичного функционала в Λ^n существует базис, в котором функционал имеет канонический вид.**

Доказательство.

Воспользуемся методом математической индукции.

1°. При $n = 1$ в любом базисе $\Phi(x) = \varphi_{11}\xi_1^2$. Если $\varphi_{11} = 0$, то мы уже имеем канонический вид, если же $\varphi_{11} \neq 0$, то при помощи очевидно невырожденной замены переменных $\xi_1' = \sqrt{|\varphi_{11}|}\;\xi_1$ приходим к каноническому виду.

2°. Предположим, что утверждение теоремы верно для квадратичных функционалов, зависящих от $n-1$ переменной, и рассмотрим случай n переменных.

Будем считать, что $\varphi_{11} \neq 0$. Этого можно добиться изменением нумерации переменных в случае, когда хотя бы одно из чисел $\varphi_{ii} \quad \forall i = [2, n]$ не равно нулю. Если же $\varphi_{ii} = 0 \quad \forall i = [1, n]$, то без ограничения общности можно считать, что $\varphi_{12} \neq 0$. Тогда, выполнив невырожденную замену переменных

$$\begin{aligned} &\xi_1 = \xi_1' + \xi_2', \\ &\xi_2 = \xi_1' - \xi_2', \\ &\xi_3 = \xi_3', \qquad \xi_4 = \xi_4', \quad \ldots, \qquad \xi_n = \xi_n', \end{aligned}$$

получим координатное представление квадратичного функционала с ненулевым диагональным элементом

$$\Phi(x) = 2\xi_1' - 2\xi_2' + R\big(\,\xi_1',\, \xi_2',\, \ldots \xi_n'\,\big)\,,$$

где $R\big(\,\xi_1',\, \xi_2',\, \ldots \xi_n'\,\big)$ не содержит квадратов от ξ_1' и ξ_2'.

3°. В записи квадратичного функционала сгруппируем слагаемые, содержащие переменную ξ_1:

$$\Phi(x) = \sum_{i=1}^{n}\sum_{k=1}^{n} \varphi_{ik}\xi_i\xi_k =$$

$$= \varphi_{11}\left(\xi_1^2 + 2\sum_{i=2}^{n}\frac{\varphi_{1i}}{\varphi_{11}}\xi_1\xi_i\right) + \sum_{i=2}^{n}\sum_{k=2}^{n}\varphi_{ik}\xi_i\xi_k,$$

и выделим полный квадрат, воспользовавшись соотношениями

$$\sum_{i=1}^{n}\sum_{k=1}^{n}\alpha_i\alpha_k = \left(\sum_{i=1}^{n}\alpha_i\right)\left(\sum_{k=1}^{n}\alpha_k\right) = \left(\sum_{i=1}^{n}\alpha_i\right)^2 =$$

$$= \left(\alpha_1 + \sum_{i=2}^{n}\alpha_i\right)^2 = \alpha_1^2 + 2\alpha_1\sum_{i=2}^{n}\alpha_i + \left(\sum_{i=2}^{n}\alpha_i\right)^2 =$$

$$= \alpha_1^2 + 2\sum_{i=2}^{n}\alpha_1\alpha_i + \sum_{i=2}^{n}\sum_{k=2}^{n}\alpha_i\alpha_k\,.$$

Получим

$$\Phi(x) = \varphi_{11}\left(\xi_1^2 + 2\sum_{i=2}^{n}\frac{\varphi_{1i}}{\varphi_{11}}\xi_1\xi_i + \sum_{i=2}^{n}\sum_{k=2}^{n}\frac{\varphi_{1i}\varphi_{1k}}{\varphi_{11}^2}\xi_i\xi_k\right) +$$

$$+\sum_{i=2}^{n}\sum_{k=2}^{n}\left(\varphi_{ik} - \frac{\varphi_{1i}\varphi_{1k}}{\varphi_{11}^2}\right)\xi_i\xi_k$$

и, окончательно,

$$\Phi(x) = \varphi_{11}\left(\xi_1 + \sum_{i=2}^{n}\frac{\varphi_{1i}}{\varphi_{11}}\xi_1\right)^2 + \sum_{i=2}^{n}\sum_{k=2}^{n}\left(\varphi_{ik} - \frac{\varphi_{1i}\varphi_{1k}}{\varphi_{11}^2}\right)\xi_i\xi_k.$$

В последней формуле первое слагаемое есть полный квадрат, а второе — квадратичный функционал, не зависящий от ξ_1 и приводящийся, согласно предположению индукции, к каноническому виду некоторой невырожденной заменой переменных $\quad \xi'_k = \sum\limits_{i=2}^{n}\tau_{ki}\xi_i \quad \forall k = [2, n]$.

4°. Выполним замену переменных квадратичного функционала $\Phi(x)$ по формулам

$$\left\{\begin{array}{lcl} \xi'_1 & = & \sqrt{|\varphi_{11}|}\left(\xi_1 + \sum\limits_{i=2}^{n}\dfrac{\varphi_{1i}}{\varphi_{11}}\xi_1\right), \\ \xi'_k & = & \sum\limits_{i=2}^{n}\tau_{ki}\xi_i \qquad \forall k = [2, n], \end{array}\right. \tag{9.2.1}$$

которая приведет к представлению его в каноническом виде. Поскольку в силу $\varphi_{11} \neq 0$ матрица выполненной замены переменных

$$\|T\| = \left\|\begin{array}{cccc} \sqrt{|\varphi_{11}|} & \sqrt{|\varphi_{11}|}\dfrac{\varphi_{12}}{\varphi_{11}} & \dots & \sqrt{|\varphi_{11}|}\dfrac{\varphi_{1n}}{\varphi_{11}} \\ 0 & \tau_{22} & \dots & \tau_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & \tau_{n2} & \dots & \tau_{nn} \end{array}\right\|$$

имеет определитель, не равный нулю, то замена (9.2.1) невырожденная. Но тогда матрица $\|T\|$ имеет обратную: $\|S\| = \|T\|^{-1}$ (см. следствие 7.5.3), которая в свою очередь является матрицей перехода от исходного базиса к искомому *каноническому* базису.

Теорема доказана.

`Замечание 9.2.1.` Базис, в котором квадратичный функционал имеет канонический вид, не единственный, равно как не является единственным сам канонический вид квадратичного функционала в Λ^n.

Вариант метода Лагранжа, использованный в доказательстве теоремы 9.2.1, не всегда оказывается самой эффективной (с точки зрения затрат вычислительных усилий) процедурой. Иногда приведение матрицы квадратичного функционала к диагональному (или каноническому) виду более рационально выполнять путем использования некоторого набора элементарных преобразований.

Ранее (на с. 286) показано, что при переходе от исходного базиса $\{g_1, g_2, \ldots, g_n\}$ к новому $\{g'_1, g'_2, \ldots, g'_n\}$ с матрицей перехода $\|S\|$ матрица квадратичного функционала $\|\Phi\|_g$ меняется по правилу $\|\Phi\|_{g'} = \|S\|^{\mathrm{T}}\|\Phi\|_g\|S\|$.

Заметим, в этом случае можно (см. теоремы 6.8.1 и 6.8.2) рассматривать матрицу $\|S\|$ как матрицу некоторого *элементарного преобразования*, переводящего симметрическую матрицу $\|\Phi\|_g$ в симметрическую матрицу $\|\Phi\|_{g'}$.

Пусть матрица $\|S\|$ такова, что умножение ее справа на матрицу $\|\Phi\|_g$ приводит последнюю к нижнему треугольному виду, тогда в силу теоремы 6.8.2 матрица $\|\Phi\|_{g'}$ оказывается диагональной.

С другой стороны, в силу определения матрицы перехода, матрица $\|S\|$ представима как произведение матриц элементарных преобразований, последовательно примененных к столбцам единичной матрицы.

Поэтому, выполнив диагонализацию некоторым набором элементарных преобразований (выполняемых на каждом шаге процедуры как с ее строками, так и с ее столбцами) и применив затем тот же самый набор элементарных преобразований к столбцам единичной матрицы, мы получим одновременно как диагональный вид квадратичного функционала $\Phi(x) = \sum\limits_{i=1}^{n} \lambda_i \xi_i^2$, так и $\|x\|_g = \|S\|\|x\|_{g'}$ — формулы перехода от исходного базиса $\{g_1, g_2, \ldots, g_n\}$ к новому базису $\{g'_1, g'_2, \ldots, g'_n\}$, в котором матрица этого квадратичного функционала оказывается диагональной.

Описанный алгоритм иллюстрируется следующим примером.

`Задача 9.2.1` *Привести в Λ^3 к диагональному виду квадратичный функционал*

$$\Phi(x) = -2\xi_1^2 - \xi_2^2 - 4\xi_3^2 - 8\xi_1\xi_2 + 2\xi_1\xi_3 - 8\xi_2\xi_3 .$$

`Решение.` Предварительно в исходном базисе расширим матрицу данного функционала единичной матрицей

$$\left\| \begin{array}{rrr|rrr} -2 & -4 & 1 & 1 & 0 & 0 \\ -4 & -1 & -4 & 0 & 1 & 0 \\ 1 & -4 & -4 & 0 & 0 & 1 \end{array} \right\| .$$

1°. На первом шаге процедуры выполним следующие элементарные операции:

— заменим вторую строку исходной матрицы разностью второй и третьей ее строк;

$$\left\| \begin{array}{rrr|rrr} -2 & -4 & 1 & 1 & 0 & 0 \\ -5 & 3 & 0 & 0 & 1 & 0 \\ 1 & -4 & -4 & 0 & 0 & 1 \end{array} \right\| ,$$

— в полученной матрице заменим второй столбец разностью второго и третьего столбцов. Кроме того, в единичной матрице заменим (по тому же правилу) второй столбец разностью второго и третьего, получим

$$\left\| \begin{array}{rrr|rrr} -2 & -5 & 1 & 1 & 0 & 0 \\ -5 & 3 & 0 & 0 & 1 & 0 \\ 1 & 0 & -4 & 0 & -1 & 1 \end{array} \right\| .$$

2°. На втором шаге заменяем вначале в левой части первую строку утроенной первой, сложенную со второй, взятой с коэффициентом 5. Соответственно такое же преобразование выполняется со столбцами как в левой, так и в правой части. Получаем следующую расширенную матрицу:

$$\left\| \begin{array}{rrr|rrr} -93 & 0 & 3 & 3 & 0 & 0 \\ 0 & 3 & 0 & 5 & 1 & 0 \\ 3 & 0 & -4 & -5 & -1 & 1 \end{array} \right\| .$$

3°. На третьем шаге в левой части заменяем третью строку первой, сложенную с третьей, взятой с коэффициентом 31.

Наконец, выполнив такие же преобразования со столбцами в обеих частях, получаем матрицу

$$\left\|\begin{array}{rrr|rrr} -93 & 0 & 0 & 3 & 0 & 3 \\ 0 & 3 & 0 & 5 & 1 & 5 \\ 0 & 0 & -3751 & -5 & -1 & 26 \end{array}\right\| .$$

Откуда следует, что перейдя к базису

$$\left\{ \left\|\begin{array}{r} 3 \\ 5 \\ -5 \end{array}\right\| , \left\|\begin{array}{r} 0 \\ 1 \\ -1 \end{array}\right\| , \left\|\begin{array}{r} 3 \\ 5 \\ 26 \end{array}\right\| \right\}$$

и выполнив замену координат по формулам перехода

$$\begin{cases} \xi_1 = 3\xi'_1 + 3\xi'_3 , \\ \xi_2 = 5\xi'_1 + \xi'_2 + 5\xi'_3 , \\ \xi_3 = -5\xi'_1 - \xi'_2 + 26\xi'_3 , \end{cases}$$

мы получим следующий диагональный вид исходного квадратичного функционала:

Решение получено.

$$\Phi(x) = -93\,{\xi'_1}^2 + 3\,{\xi'_2}^2 - 3751\,{\xi'_3}^2 .$$

9.3. Исследование знака квадратичного функционала

Несмотря на неединственность диагонального или канонического представления, квадратичные функционалы в Λ^n обладают рядом важных свойств, инвариантных относительно (то есть не зависящих от) выбора базиса. Одной из таких характеристик является ранг квадратичного функционала, которую вводит

Определение 9.3.1	Максимальное число не равных нулю коэффициентов канонического вида квадратичного функционала $\Phi(x)$ в Λ^n называется его *рангом* и обозначается как $\operatorname{rg}\Phi$ (или $\operatorname{rank}\Phi$).

Теорема 9.3.1 **Ранг квадратичного функционала в Λ^n не зависит от выбора базиса.**

Доказательство.

По следствию 9.1.2 ранг матрицы билинейного функционала не зависит от выбора базиса. Поэтому не будет зависеть от выбора базиса и ранг матрицы порождаемого им квадратичного функционала.

С другой стороны, в силу теорем 8.4.3 и 9.1.1 ранг матрицы квадратичного функционала равен числу ненулевых коэффициентов в каноническом виде, а значит, и его рангу.

Теорема доказана.

При исследовании знака значений квадратичного функционала оказывается полезным использование следующих его характеристик.

Определение 9.3.2

1°. Максимальное число положительных коэффициентов диагонального (канонического) вида квадратичного функционала $\Phi(x)$ в Λ^n называется его *положительным индексом инерции* и обозначается $\mathrm{rg}_+ \Phi$.

2°. Максимальное число отрицательных коэффициентов диагонального (канонического) вида квадратичного функционала $\Phi(x)$ называется его *отрицательным индексом инерции* и обозначается $\mathrm{rg}_- \Phi$.

3°. Разность между положительным и отрицательным индексами инерции называется *сигнатурой* квадратичного функционала $\Phi(x)$ и обозначается $\operatorname{sgn} \Phi$.

Теорема 9.3.2 (инерции квадратичных функционалов) **Значения положительного и отрицательного индексов инерции, а также сигнатуры квадратичного функционала $\Phi(x)$ в Λ^n** *не зависят* **от выбора базиса, в котором этот функционал имеет диагональный (или канонический) вид.**

Доказательство.

1°. Пусть квадратичный функционал $\Phi(x)$ имеет в некотором базисе $\{g_1, g_2, \ldots, g_n\}$ координатное представление $\Phi(x) = \sum\limits_{i=1}^{n}\sum\limits_{k=1}^{n}\varphi_{ik}\xi_i\xi_k$ и пусть существуют два различных базиса $\{g'_1, g'_2, \ldots, g'_n\}$ и $\{g"_1, g"_2, \ldots, g"_n\}$, в которых данный функционал имеет следующий диагональный вид:

$$\Phi(x) = \sum_{i=1}^{k}\lambda_i\eta_i^2 - \sum_{i=k+1}^{m}\lambda_i\eta_i^2, \qquad m \le n, \;\; \lambda_i > 0 \;\; \forall i = [1, m]$$

и, соответственно,

$$\Phi(x) = \sum_{i=1}^{p}\mu_i\kappa_i^2 - \sum_{i=p+1}^{q}\mu_i\kappa_i^2, \qquad q \le n, \quad \mu_i > 0 \;\; \forall i = [1, q]\,,$$

где $\|x\|_{g'} = \|\, \eta_1\, \eta_2\, \ldots\, \eta_n \,\|^{\mathrm{T}}$ и $\|x\|_{g''} = \|\, \kappa_1\, \kappa_2\, \ldots\, \kappa_n \,\|^{\mathrm{T}}$.

В этом случае должны существовать формулы перехода от диагональных базисов к исходному, такие, что

$$\eta_s = \sum_{j=1}^{n}\omega_{sj}\xi_j \quad \text{и} \quad \kappa_s = \sum_{j=1}^{n}\theta_{sj}\xi_j \qquad \forall s = [1, n]\,. \quad (9.3.1)$$

2°. Для некоторого элемента

$$x = \sum_{j=1}^{n}\xi_j g_j = \sum_{j=1}^{n}\eta_j g'_j = \sum_{j=1}^{n}\kappa_j g''_j$$

приравняем значения функционала $\Phi(x)$ в диагональных базисах

$$\sum_{i=1}^{k}\lambda_i\eta_i^2 - \sum_{i=k+1}^{m}\lambda_i\eta_i^2 = \sum_{i=1}^{p}\mu_i\kappa_i^2 - \sum_{i=p+1}^{q}\mu_i\kappa_i^2$$

и преобразуем полученное равенство к виду

$$\sum_{i=1}^{k}\lambda_i\eta_i^2 + \sum_{i=p+1}^{q}\mu_i\kappa_i^2 = \sum_{i=k+1}^{m}\lambda_i\eta_i^2 + \sum_{i=1}^{p}\mu_i\kappa_i^2\,. \qquad (9.3.2)$$

3°. Исследуем полученное соотношение. Допустим, что $k < p$, и предположим, что элемент x имеет в диагональных базисах координаты

$$\eta_j = 0 \quad \forall j = [1, k] \qquad \text{и} \qquad \kappa_j = 0 \quad \forall j = [p+1, n]\,. \tag{9.3.3}$$

Этих условий меньше, чем n, поскольку $k < p$. Если их подставить в равенства (9.3.1), то мы получим однородную систему линейных уравнений относительно неизвестных $\xi_1, \xi_2, \ldots, \xi_n$. Поскольку число таких уравнений меньше числа неизвестных, то можно утверждать, что она в силу теоремы 6.7.1 имеет нетривиальные решения, и, следовательно, элемент может быть *ненулевым.*

С другой стороны, из равенства (9.3.2), положительности чисел $\lambda_j \;\; \forall j = [1, m]$ и $\mu_j \;\; \forall j = [1, q]$, а также условий (9.3,3) следует, что и все $\kappa_j = 0 \;\; \forall j = [1, n]$.

Тогда в силу (9.3.1) мы получаем однородную систему n линейных уравнений

$$\kappa_s = \sum_{j=1}^{n} \theta_{sj} \xi_j \qquad \forall s = [1, n]$$

с n неизвестными и невырожденной (поскольку она есть матрица перехода) основной матрицей, имеющую только тривиальное решение, то есть элемент x обязан быть *нулевым.* Полученное противоречие показывает ошибочность предположения о том, что $k < p$.

4°. Аналогичными рассуждениями показываем, что невозможно и соотношение $k > p$. Поэтому приходим к заключению, что $k = p$.

5°. По теореме 9.3.1 $m = q$ и потому $k - m = p - q$.

Теорема доказана.

При исследовании знака значений квадратичного функционала оказывается полезным понятие его *знаковой определенности.*

Определение 9.3.3	1°. Квадратичный функционал $\Phi(x)$ называется *положительно определенным* на подпространстве $\Omega \subseteq \Lambda$, если $\Phi(x) > 0$ для любого ненулевого $x \in \Omega$.

2°. Квадратичный функционал $\Phi(x)$ называется *отрицательно определенным* на подпространстве $\Omega \subseteq \Lambda$, если $\Phi(x) < 0$ для любого ненулевого $x \in \Omega$.

3°. Если же Ω совпадает с Λ, то говорят, что квадратичный функционал $\Phi(x)$ является *положительно* (или соответственно *отрицательно*) *определенным*.

4°. Если же $\Phi(x) \geq 0$ $\big(\Phi(x) \leq 0\big)$ для всех $x \in \Lambda$, то говорят, что квадратичный функционал является *положительно* (*отрицательно*) *полуопределенным*.

Теорема 9.3.3 **Максимальная размерность подпространства Ω в Λ^n, на котором квадратичный функционал положительно (отрицательно) определен, равняется положительному (отрицательному) индексу инерции этого функционала.**

Доказательство.

Следует из теоремы 9.3.2 и очевидного равенства числа положительных (отрицательных) элементов матрицы квадратичного функционала (в его диагональном представлении) размерности подпространства Ω.

Теорема доказана.

В ряде прикладных задач возникает потребность в исследовании знаковой определенности квадратичного функционала без приведения его к диагональному виду. Удобное для этого условие положительной определенности квадратичного функционала дает

Теорема 9.3.4 (критерий Сильвестра) **Для положительной определенности квадратичного функционала в Λ^n необходимо и достаточно, чтобы у его матрицы были** *положительными* **все главные миноры вида**

$$\det\left\|\begin{array}{cccc} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1k} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_{k1} & \varphi_{k2} & \cdots & \varphi_{kk} \end{array}\right\| \qquad \forall k = [1, n]\,.$$

`Доказательство достаточности.`

1°. Воспользуемся методом математической индукции.
Для $n = 1$ достаточность очевидна. Допустим, что из положительности всех главных миноров матрицы квадратичного функционала порядка $k[1, n-1]$ следует возможность приведения квадратичного функционала, зависящего от $n-1$ переменной, к виду $\Phi(x) = \sum\limits_{i=1}^{n-1} \xi_i^2$.

2°. Покажем, что в этом случае к такому же виду можно будет привести и квадратичный функционал, зависящий от n переменных.
Сначала в выражении для квадратичного функционала, зависящего от n переменных, выделим слагаемые, содержащие ξ_n:

$$\Phi(x) = \sum_{i=1}^{n} \sum_{k=1}^{n} \varphi_{ik} \xi_i \xi_k =$$

$$= \sum_{i=1}^{n-1} \sum_{k=1}^{n-1} \varphi_{ik} \xi_i \xi_k + 2 \sum_{i=1}^{n-1} \varphi_{in} \xi_i \xi_n + \varphi_{nn} \xi_n^2 .$$

Двойная сумма в правой части этой цепочки равенств есть некоторый квадратичный функционал $\Psi(x)$, зависящий от $n-1$ переменной, причем главные миноры его матрицы совпадают с главными минорами матрицы $\Phi(x)$ до порядка $n-1$ включительно, которые, по предположению индукции, положительны.
Откуда следует, что квадратичный функционал $\Psi(x)$ положительно определенный, и для него существует невырожденная замена переменных

$$\xi_k = \sum_{i=1}^{n-1} \sigma_{ki} \eta_i \qquad \forall k = [1, n-1] ,$$

приводящая его к каноническому виду $\Psi(x) = \sum\limits_{i=1}^{n-1} \eta_i^2$.
Выпишем представление квадратичного функционала $\Phi(x)$ в новых переменных:

$$\Phi(x) = \sum_{i=1}^{n-1} \eta_i^2 + 2 \sum_{i=1}^{n-1} \varphi'_{in} \eta_i \xi_n + \varphi_{nn} \xi_{nn}^2$$

и выделим в нем полные квадраты:

$$\Phi(x) = \sum_{i=1}^{n-1}\left(\eta_i^2+2\varphi'_{in}\eta_i\xi_n+\varphi'^2_{in}\xi_n^2\right)+\left(\varphi_{nn}-\sum_{i=1}^{n-1}\varphi'^2_{in}\right)\xi^2_{nn} =$$

$$= \sum_{i=1}^{n-1}\zeta_i^2+\varphi''_{nn}\xi_n^2\,, \tag{9.3.4}$$

где

$$\varphi''_{nn} = \varphi_{nn} - \sum_{i=1}^{n-1}\varphi'^2_{in}\ : \qquad \zeta_i = \eta_i+\varphi'_{in}\xi_n \quad \forall i = [1, n-1]\,.$$

В матричном виде эту замену переменных можно записать как

$$\left\|\begin{array}{l} \zeta_1 \\ \zeta_2 \\ \dots \\ \zeta_{n-1} \\ \xi_n \end{array}\right\| = \left\|\begin{array}{ccccc} 1 & 0 & \dots & 0 & \varphi'_{1n} \\ 0 & 1 & \dots & 0 & \varphi'_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & \varphi'_{n-1,n} \\ 0 & 0 & \dots & 0 & 1 \end{array}\right\| \left\|\begin{array}{l} \eta_1 \\ \eta_2 \\ \dots \\ \eta_{n-1} \\ \xi_n \end{array}\right\|,$$

и поскольку определитель ее матрицы отличен от нуля, то эта замена невырожденная.

3°. Для доказательства положительной определенности $\Phi(x)$ осталось показать, что $\varphi''_{nn} > 0$. Воспользуемся тем, что в силу следствия 9.1.1 определитель матрицы квадратичного функционала сохраняет знак при замене базиса. По условию теоремы знак определителя матрицы квадратичного функционала в исходном базисе *положительный*, поскольку этот определитель имеет вид

$$\det\left\|\begin{array}{cccc} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1k} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_{k1} & \varphi_{k2} & \cdots & \varphi_{kk} \end{array}\right\|$$

и является левым верхним главным минором порядка n.

Но тогда из (9.3.4) — выражения для $\Phi(x)$ в конечном базисе — мы получаем, что определитель матрицы квадратичного функционала равен φ''_{nn} — положительному числу.

Поэтому можно сделать невырожденную замену переменных $\zeta_n = \xi_n \sqrt{\varphi''_{nn}}$, приводящую к каноническому виду $\sum\limits_{i=1}^{n} \zeta_i^2$ функционал $\Phi(x)$.

Следовательно, квадратичный функционал положительно определен для числа переменных n, а значит, в силу математической индукции, и для любого числа переменных.

Достаточность доказана.

Доказательство необходимости для критерия Сильвестра положительной определенности квадратичного функционала будет приведено в § 10.3.

Исходя из критерия Сильвестра для положительной определенности квадратичного функционала, можно получить аналогичный критерий *отрицательной* определенности квадратичного функционала.

Следствие 9.3.1 **Для отрицательной определенности квадратичного функционала в Λ^n необходимо и достаточно, чтобы главные миноры матрицы функционала (указанные в формулировке теоремы 9.3.4) четного порядка были положительны, а нечетного порядка – отрицательны.**

Доказательство.

Пусть квадратичный функционал $\Phi(x)$ отрицательно определенный, тогда функционал $-\Phi(x)$ будет, очевидно, положительно определенным. Применив к нему критерий Сильвестра положительной определенности, получим, в силу линейного свойства определителя (см. следствие 6.2.3), для главного минора k-го порядка условие

$$\det \left\| \begin{matrix} -\varphi_{11} & -\varphi_{12} & \cdots & -\varphi_{1k} \\ -\varphi_{21} & -\varphi_{22} & \cdots & -\varphi_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ -\varphi_{k1} & -\varphi_{k2} & \cdots & -\varphi_{kk} \end{matrix} \right\| =$$

$$= (-1)^k \det \left\| \begin{matrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1k} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_{k1} & \varphi_{k2} & \cdots & \varphi_{kk} \end{matrix} \right\| > 0 \qquad \forall k = [1, n] \, .$$

Откуда и следует доказываемое утверждение.

`Следствие доказано.`

9.4. Инварианты линий второго порядка на плоскости

Независимость значений ранга и сигнатуры квадратичного функционала от выбора базиса позволяет выполнить классификацию линий второго порядка на плоскости способом, отличным от приведенного в теореме 4.4.1.

Рассмотрим линию второго порядка на плоскости Oxy в базисе $\{g_1, g_2\}$ и с началом координат в точке O. Эта линия в общем случае задается согласно определению 4.4.1 уравнением вида

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0\,,$$

где числа A, B, C, D, E и F произвольны с одним лишь ограничением, что A, B и C не равны нулю одновременно, т.е. $|A| + |B| + |C| > 0$.

Нетрудно проверить, что при замене начала координат коэффициенты A, B и C не меняются, а при смене базиса преобразуются как коэффициенты квадратичного функционала (см. теорему 9.1.1). Поэтому можно считать, что многочлен $Ax^2 + 2Bxy + Cy^2$ задает в исходном базисе $\{g_1, g_2\}$ квадратичный функционал

$$\Phi(x, y) = Ax^2 + 2Bxy + Cy^2$$

с матрицей $\left\| \begin{matrix} A & B \\ B & C \end{matrix} \right\|$.

Из теорем 9.2.2 и 9.3.1 следует, что $\operatorname{rg}\Phi$ — *ранг* и $|\operatorname{sgn}\Phi|$ — *модуль сигнатуры* квадратичного функционала не зависят от выбора системы координат и, следовательно, являются инвариантами линии второго порядка на плоскости.

Модуль сигнатуры здесь необходим, поскольку одновременное изменение знаков всех коэффициентов уравнения линии второго порядка меняет уравнение, но линия при этом останется той же.

В запись уравнения линии второго порядка на плоскости входят также и коэффициенты D, E и F. Поэтому необходимо выяснить, существуют ли другие инварианты, образованные при помощи коэффициентов A, B, C, D, E и F.

Для этого рассмотрим вспомогательный квадратичный функционал в Λ^3 с базисом $\{g_1, g_2, g_3\}$:

$$\Psi(x,y,z) = Ax^2 + 2Bxy + Cy^2 + 2Dxz + 2Eyz + Fz^2$$

с матрицей $\left\| \begin{array}{ccc} A & B & D \\ B & C & E \\ D & E & F \end{array} \right\|$. В этом случае совокупность всех точек в Λ^3, для которых $\Psi(x,y,1) = 0$, есть рассматриваемая нами линия второго порядка, расположенная в пространстве на плоскости $z = 1$.

Сделаем в Λ^3 такую замену координат, при которой плоскость $z = 1$ переходит сама в себя. Найдем для такой замены правило изменения коэффициентов квадратичного функционала $\Psi(x,y,z)$.

Лемма 9.4.1 **Матрица $\|S\|$ перехода от декартовой системы координат $\{O, g_1, g_2, g_3\}$ к декартовой системе координат $\{O', g_1', g_2', g_3'\}$, для которой плоскость $z = 1$ переходит сама в себя, имеет вид**

$$\|S\| = \left\| \begin{array}{ccc} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ 0 & 0 & 1 \end{array} \right\| .$$

Доказательство.

Воспользуемся формулами замены координат в плоскости Oxy:

$$\left\{ \begin{array}{lcl} x & = & \sigma_{11}x' + \sigma_{12}y' + \sigma_{13}\,, \\ y & = & \sigma_{21}x' + \sigma_{22}y' + \sigma_{23}. \end{array} \right.$$

В матричном виде эти соотношения для Λ^3 при $z = 1$ и $z' = 1$ принимают вид

$$\left\| \begin{array}{c} x \\ y \\ 1 \end{array} \right\| = \left\| \begin{array}{ccc} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ 0 & 0 & 1 \end{array} \right\| \left\| \begin{array}{c} x' \\ y' \\ 1 \end{array} \right\| .$$

Невырожденность матрицы $\|S\|$ следует из очевидного условия $\det \left\| \begin{array}{cc} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{array} \right\| \neq 0\,.$

Лемма доказана.

Поскольку ранг и сигнатура квадратичного функционала не меняются при любых заменах базиса, то это будет верным и для замен, переводящих плоскость $z = 1$ саму в себя. Поэтому $\operatorname{rg}\Psi$ и $\operatorname{sgn}\Psi$ сохраняются при таких заменах, а числа $\operatorname{rg}\Psi$ и $|\operatorname{sgn}\Psi|$ являются инвариантами уравнения линии второго порядка. Таким образом, доказана

`Теорема 9.4.1` **При любых заменах декартовой системы координат на плоскости числа** $\operatorname{rg}\Phi$, $|\operatorname{sgn}\Phi|$, $\operatorname{rg}\Psi$ **и** $|\operatorname{sgn}\Psi|$ **являются инвариантами линии второго порядка.**

Подсчитаем значения четырех параметров $\operatorname{rg}\Phi$, $|\operatorname{sgn}\Phi|$, $\operatorname{rg}\Psi$ и $|\operatorname{sgn}\Psi|$ для каждого из девяти канонических видов линий второго порядка на плоскости, приведенных в формулировке теоремы 4.4.1, в предположении, что система координат $\{O',\, g'_1,\, g'_2,\, g'_3\}$ оказалась канонической.

Например, для эллипса имеем

$$\Phi(x', y') = \frac{x'^2}{a^2} + \frac{y'^2}{b^2}\,, \qquad \Psi(x', y', z') = \frac{x'^2}{a^2} + \frac{y'^2}{b^2} - z'^2\,.$$

Это дает $\operatorname{rg}\Phi = 2$, $|\operatorname{sgn}\Phi| = 2$, $\operatorname{rg}\Psi = 3$ и $|\operatorname{sgn}\Psi| = 1$.

Такой метод дает нужный результат в восьми случаях из девяти. Однако для параболы (хотя с функционалом $\Phi(x', y') = y'^2$ проблем не возникает) подсчет значений ранга и сигнатуры по виду функционала в каноническом базисе $\Psi(x', y', z') = y'^2 - 2px'$ невозможен, поскольку его матрица имеет недиагональный вид:

$$\|\Psi\|_{g'} = \left\|\begin{array}{rrr} 0 & 0 & -p \\ 0 & 1 & 0 \\ -p & 0 & 0 \end{array}\right\|.$$

Для получения диагонального вида и подсчета ранга и сигнатуры здесь можно использовать матрицу перехода следующего вида:

$$\|S\| = \left\|\begin{array}{rrr} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{array}\right\|,$$

которая приводит матрицу $\|\Psi\|_{g'}$ к диагональному виду.

Такая замена координат хотя и не обеспечивает перехода плоскости $z = 1$ самой в себя, но она, как всякая линейная замена, сохраняет ранг и сигнатуру.

Таблица 9.4.1

Вид линии	**Уравнение**	rg Ψ	$\lvert\mathrm{sgn}\,\Psi\rvert$	rg Φ	$\lvert\mathrm{sgn}\,\Phi\rvert$
Эллипс	$\frac{x'^2}{a^2}+\frac{y'^2}{b^2}=1$	3	1	2	2
Мнимый эллипс	$\frac{x'^2}{a^2}+\frac{y'^2}{b^2}=-1$	3	3	2	2
Изолированная точка	$\frac{x'^2}{a^2}+\frac{y'^2}{b^2}=0$	2	2	2	2
Гипербола	$\frac{x'^2}{a^2}-\frac{y'^2}{b^2}=1$	3	1	2	0
Пересекающиеся прямые	$\frac{x'^2}{a^2}-\frac{y'^2}{b^2}=0$	2	0	2	0
Парабола	$y'^2 = 2px'$	3	1	1	1
Параллельные прямые	$y'^2 = a^2 \quad \forall x'$	2	0	1	1
Совпадающие прямые	$y'^2 = 0 \quad \forall x'$	1	1	1	1
Пара мнимых прямых	$y'^2 = -a^2 \quad \forall x'$	2	2	1	1

Действительно,

$$\|\Psi\|_{g''} = \|s\|^{\mathrm{T}}\|\Psi\|_{g'}\|s\| =$$

$$= \left\| \begin{array}{rrr} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{array} \right\| \left\| \begin{array}{rrr} 0 & 0 & -p \\ 0 & 1 & 0 \\ -p & 0 & 0 \end{array} \right\| \left\| \begin{array}{rrr} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{array} \right\| =$$

$$= \left\| \begin{array}{rrr} 2p & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2p \end{array} \right\|,$$

что в итоге дает $\quad \operatorname{rg}\Phi = 1,\ |\operatorname{sgn}\Phi| = 1,\ \operatorname{rg}\Psi = 3$ и $|\operatorname{sgn}\Psi| = 1.$

Полученные результаты сведем в таблицу 9.4.1, из которой следует, что каждый вид линии второго порядка на плоскости имеет свой уникальный набор значений инвариантов, который может быть принят за признак принадлежности некоторой линии второго порядка к конкретному виду.

В заключение отметим, что

1°. Для линий второго порядка на плоскости существуют и другие инварианты, например, ортогональными инвариантами являются числа $A + C$ и $\det\left\|\begin{matrix} A & B \\ B & C \end{matrix}\right\|$. Справедливость этого утверждения показана в § 4.4.

2°. Схема, аналогичная рассмотренной, может быть использована и для классификации поверхностей второго порядка в пространстве.

9.5. Экстремальные свойства квадратичных функционалов

Из теоремы 9.2.1 следует существование в Λ^n базиса, в котором квадратичный функционал $\Phi(x)$ имеет диагональный вид. Допустим, что этот базис $\{\, g'_1,\, g'_2,\, \ldots,\, g'_n\}$ построен так, что $\quad \Phi(x) = \sum\limits_{i=1}^{n} \lambda_i \xi_i'^2$ и $\lambda_1 \le \lambda_2 \le\ \ldots \le \lambda_n$. Тогда справедлива

Теорема 9.5.1 **Для квадратичного функционала $\Phi(x)$ в Λ^n верны оценки $\lambda_1 = \min\limits_{x\in\Omega}\Phi(x)$ и $\lambda_n = \max\limits_{x\in\Omega}\Phi(x)$, где $\Omega = \left\{\, x \;\middle|\; \sum\limits_{i=1}^{n} \xi_i'^2 = 1 \,\right\}$.**

Доказательство.

Если в рассматриваемом базисе $\Phi(x) = \sum\limits_{i=1}^{n} \lambda_i \xi_i'^2$, то в силу соотношений $\lambda_1 \le \lambda_2 \le\ \ldots \le \lambda_n$ будут иметь место неравенства

$$\lambda_1 \sum_{i=1}^{n} \xi_i'^2 \le \sum_{i=1}^{n} \lambda_i \xi_i'^2 \qquad \text{и} \qquad \sum_{i=1}^{n} \lambda_i \xi_i'^2 \le \lambda_n \sum_{i=1}^{n} \xi_i'^2 \,.$$

Но поскольку $\sum\limits_{i=1}^{n} \xi_i'^2 = 1$, то будут также справедливы и оценки

$$\lambda_1 \leq \sum_{i=1}^{n} \lambda_i \xi_i'^2 \qquad \text{и} \qquad \sum_{i=1}^{n} \lambda_i \xi_i'^2 \leq \lambda_n \,.$$

То есть при $x = \| \, 1 \; 0 \, \ldots \, 0 \, \|^{\mathrm{T}}$ достигается максимум, а при $x = \| \, 0 \; 0 \, \ldots \, 1 \, \|^{\mathrm{T}}$ — минимум значений функционала.

Теорема доказана.

9.6. Полилинейные функционалы

По аналогии с билинейными функционалами, зависящими от пары элементов линейного пространства, можно определить нелинейные функционалы (называемые *полилинейными*), обладающие аналогичными свойствами, но зависящие от большего числа аргументов.

Определение 9.6.1

Пусть в линейном пространстве Λ каждому упорядоченному набору из k элементов $\{ x_1, x_2, \ldots, x_n \}$ поставлено в соответствие число $\Phi\Big(x_1, x_2, \ldots, x_n \Big)$ так, что $\forall j = [1, k]$

$$\Phi\Big(x_1, x_2, \ldots, \alpha x_j' + \beta x_j'', \ldots, x_n \Big) =$$

$$= \alpha \Phi\Big(x_1, x_2, \ldots, x_j', \ldots, x_n \Big) +$$

$$+ \beta \Phi\Big(x_1, x_2, \ldots, x_j'', \ldots, x_n \Big)$$

$$\forall x_j', x_j'' \in \Lambda \qquad \forall \alpha, \beta \in \mathbb{R} \,,$$

тогда говорят, что в Λ задан *полилинейный* функционал, а именно *k-линейный* функционал.

Пример 9.6.1: 1°. Произведение k, определенных в Λ, линейных функционалов $F_1(x)$, $F_2(x)$, $\ldots$, $F_k(x)$, то есть

$$\Phi(x_1, x_2, \ldots, x_k) = F_1(x_1)\, F_2(x_2) \, \ldots F_k(x_k),$$

является k-линейным функционалом в Λ.

2°. Смешанное произведение трех векторов в трехмерном геометрическом пространстве является *трилинейным* функционалом.

3°. Определитель n-го порядка есть полилинейный функционал от n элементов в Λ^n в случае, когда координатные представления этих элементов являются столбцами данного определителя.

4°. Стоит также отметить, что к полилинейным функционалам относятся специальные математические объекты, называемые *тензорами*. Эти объекты позволяют единообразно описывать и использовать в Λ^n, введенные в нашем курсе независимо друг от друга, понятия элементов, операторов и функционалов. Тензоры, их определение, свойства и операции с ними, подробно рассматриваются в Приложении 4.

10. Евклидово пространство

10.1. Определение и основные свойства

В произвольном линейном пространстве отсутствуют понятия *длины, расстояния, величины угла* и других *метрических характеристик.* Однако их использование становится возможным, если в линейном пространстве дополнительно ввести специальную, определяемую ниже, операцию.

Определение 10.1.1 Пусть в вещественном линейном пространстве Λ каждой упорядоченной паре элементов x и y поставлено в соответствие единственное вещественное число (x, y), называемое *скалярным произведением*, так, что выполнены аксиомы:

$$
\begin{array}{ll}
1^\circ. & (x, y) = (y, x), \\
2^\circ. & (x, y + z) = (x, y) + (x, z) \quad \forall z \in \Lambda, \\
3^\circ. & (\lambda x, y) = \lambda (x, y) \quad \forall \lambda \in \mathbb{R}, \\
4^\circ. & (x, x) \geq 0 \quad \forall x \in \Lambda, \\
 & \text{причем} \qquad (x, x) = 0 \quad \Longleftrightarrow \quad x = o,
\end{array}
$$

тогда говорят, что задано *евклидово пространство* E.

Замечание 10.1.1. Аксиомы 1°–4° в совокупности означают, что скалярное произведение есть *билинейный* (следует из 2° и 3°) и *симметричный* (следует из 1°) функционал, который, кроме того, порождает *положительно определенный квадратичный* (следует из 4°) функционал.

Любой билинейный функционал, обладающий данными свойствами, может использоваться в качестве скалярного произведения.

Пример 10.1.1

1°. Трехмерное геометрическое пространство со скалярным произведением, введенным по правилам § 2.2, является евклидовым.

2°. Пространство n-мерных столбцов

$$x = \| \, \xi_1 \, \xi_2 \, \ldots \xi_n \, \|^{\mathrm{T}} \qquad \text{и} \qquad y = \| \, \eta_1 \, \eta_2 \, \ldots \eta_n \, \|^{\mathrm{T}}$$

со скалярным произведением, определяемым по формуле $(x, y) = \sum\limits_{i=1}^{n} \xi_i \eta_i$, есть евклидово пространство.

3°. Евклидовым будет пространство непрерывных на $\tau \in [\alpha, \beta]$ функций $x(\tau)$ и $y(\tau)$ со скалярным произведением $(x, y) = \int\limits_{\alpha}^{\beta} x(\tau) y(\tau) \, d\tau$.

Задача 10.1.1 *Можно ли в трехмерном пространстве скалярное произведение определить как произведение длин векторов на куб косинуса угла между ними?*

Решение. Нет, нельзя, так как не будет выполняться аксиома 3° в определении 10.1.1.

Определение 10.1.2 В евклидовом пространстве E назовем

1) *нормой* (или *длиной*) элемента $x \in E$ число $|x| = \sqrt{(x, x)}$;
2) *расстоянием* между элементами x и y число $|x - y|$.

Замечание 10.1.2 Использование для обозначения нормы элемента одинарных вертикальных ограничителей вида $| \ldots |$ не приводит к каким-либо конфликтам с введенными ранее обозначениями, поскольку для линейного пространства вещественных чисел норма числа, очевидно, совпадает с его абсолютной величиной, для комплексного числа норма совпадает с его модулем, а для линейного пространства геометрических векторов — с длиной вектора.

Теорема 10.1.1 (Неравенство Коши - Буняковского) **Для любых $x, y \in E$ имеет место неравенство**

$$\bigl|(x,y)\bigr| \leq |\,x\,|\,|\,y\,|\,.$$

Доказательство.

Для $\forall x, y \in E$ и любого числа $\tau \in \mathbb{R}$ элемент $x - \tau y \in E$.

Согласно четырем аксиомам из определения 10.1.1 и определению 10.1.2 имеем

$$0 \leq (x - \tau y, x - \tau y) = (x,x) - (x,y)\tau - (y,x)\tau + (y,y)\tau^2 =$$

$$= (x,x) - 2(x,y)\tau + (y,y)\tau^2 = |\,x\,|^2 - 2(x,y)\tau + |\,y\,|^2\tau^2.$$

Полученный квадратный трехчлен неотрицателен $\forall \tau$ тогда и только тогда, когда его дискриминант неположителен, то есть

$$(x,y)^2 - |\,x\,|^2|\,y\,|^2 \leq 0 \qquad \Longrightarrow \qquad \bigl|(x,y)\bigr| \leq |\,x\,|\,|\,y\,|\,.$$

Теорема доказана.

Обратите внимание, что вертикальные ограничители в левой части неравенства Коши – Буняковского обозначают абсолютную величину числа, а в правой — нормы элементов в E.

Задача 10.1.2 *Показать, что неравенство Коши – Буняковского превращается в равенство тогда и только тогда, когда элементы x и y линейно зависимы.*

Следствие 10.1.1 (неравенство треугольника) **Для любых $x, y \in E$ имеет место неравенство**

$$\bigl|\,x + y\,\bigr| \leq |\,x\,| + |\,y\,|\,.$$

Доказательство.

Из аксиом евклидова пространства и неравенства Коши – Буняковского имеем

$$\left| x+y \right|^2 = (x+y,x+y) = (x,x)+2(x,y)+(y,y) \le$$

$$\le |\,x\,|^2 + 2|\,x\,|\,|\,y\,| + |\,y\,|^2 = \Big(|\,x\,| + |\,y\,|\Big)^2 .$$

Откуда, в силу неотрицательности чисел $|\,x\,|$, $|\,y\,|$ и $|\,x\,|+|\,y\,|$, получаем неравенство треугольника.

Следствие доказано.

Отметим, что неравенства Коши – Буняковского и треугольника для евклидова пространства из примера 10.1.1 (2°) имеют вид

$$\left| \sum_{i=1}^{n} \xi_i \eta_i \right| \le \sqrt{\sum_{j=1}^{n} \xi_j^2}\,\sqrt{\sum_{k=1}^{n} \eta_k^2} \qquad \forall \xi_i, \eta_i \quad i=[1,n]\,,$$

$$\sqrt{\sum_{i=1}^{n} \big(\xi_i+\eta_i\big)^2} \le \sqrt{\sum_{j=1}^{n} \xi_j^2} + \sqrt{\sum_{k=1}^{n} \eta_k^2} \qquad \forall \xi_i, \eta_i \quad i=[1,n]\,,$$

в то время как для евклидова пространства из примера 10.1.1 (3°) соответственно:

$$\left| \int\limits_{\alpha}^{\beta} x(\tau) y(\tau)\, d\tau \right| \le \sqrt{\int\limits_{\alpha}^{\beta} x^2(\tau)\, d\tau}\,\sqrt{\int\limits_{\alpha}^{\beta} y^2(\tau)\, d\tau}\,,$$

$$\sqrt{\int\limits_{\alpha}^{\beta} \big(x(\tau)+y(\tau)\big)^2 d\tau} \le \sqrt{\int\limits_{\alpha}^{\beta} x^2(\tau)\, d\tau} + \sqrt{\int\limits_{\alpha}^{\beta} y^2(\tau)\, d\tau}\,.$$

Определение 10.1.3	В евклидовом пространстве E *величиной угла* между ненулевыми элементами x и y назовем число $\alpha \in [0,\pi]$, удовлетворяющее соотношению $$\cos\alpha = \frac{(x,y)}{|x||y|}\,.$$

Из неравенства Коши – Буняковского (теорема 10.1.1) следует, что величина угла существует для любой пары ненулевых элементов в E.

Определение 10.1.4	В евклидовом пространстве E элементы x и y называются *ортогональными*, если $(x, y) = 0$.

Заметим, что нулевой элемент евклидова пространства ортогонален любому другому элементу.

10.2. Ортонормированный базис. Ортогонализация базиса

Определение 10.2.1	В конечномерном евклидовом пространстве E^n базис $\{ e_1, e_2, \ldots, e_n \}$ называется *ортонормированным*, если $(e_i, e_j) = \delta_{ij} \quad \forall i, j = [1, n]$, где δ_{ij} — символ Кронекера.

Теорема 10.2.1 (Грама - Шмидта) **Во всяком евклидовом пространстве E^n существует ортонормированный базис.**

Доказательство.

1°. Пусть в E^n дан некоторый, вообще говоря, неортогональный базис $\{ g_1, g_2, \ldots, g_n \}$. Построим вначале базис $\{ e'_1, e'_2, \ldots, e'_n \}$ из попарно ортогональных (но не обязательно нормированных) элементов. Последовательное построение этих элементов будем называть процессом *ортогонализации базиса.*

Возьмем $e'_1 = g_1$. Элемент e'_2 будем искать в виде
$$e'_2 = g_2 + \alpha_{21} e'_1 \,,$$
где α_{21} — некоторая константа. Подберем α_{21} так, чтобы $(e'_1, e'_2) = 0$, для этого достаточно, чтобы

$$(e'_1, e'_2) = (e'_1, g_2 + \alpha_{21} e'_1) = (e'_1, g_2) + \alpha_{21}(e'_1, e'_1) = 0 \quad \Longrightarrow$$

$$\Longrightarrow \qquad \alpha_{21} = -\frac{(e'_1, g_2)}{(e'_1, e'_1)}.$$

Заметим, что $e'_2 \neq o$. Действительно, из

$$o = e'_2 = g_2 + \alpha_{21} e'_1 = g_2 + \alpha_{21} g_1$$

следует линейная зависимость g_1 и g_2, что противоречит условию принадлежности этих элементов базису (см. лемму 7.2.2).

2°. Допустим теперь, что нам удалось ортогонализовать $k-1$ элемент, и в качестве e'_k выберем элемент вида

$$e'_k = g_k + \sum_{j=1}^{k-1} \alpha_{kj} e'_j \,. \tag{10.2.1}$$

Потребуем, чтобы $(e'_k, e'_i) = 0 \quad \forall i = [1, k-1]$, но тогда, в силу $(e'_j, e'_i) = 0$ при $i \neq j$, имеем

$$0 = (\, g_k + \sum_{j=1}^{k-1} \alpha_{kj} e'_j,\ e'_i\,) = (g_k, e'_i) + \alpha_{ki}(e'_i, e'_i) \qquad \Longrightarrow$$

$$\Longrightarrow \qquad \alpha_{ki} = -\frac{(g_k, e'_i)}{(e'_i, e'_i)} \quad \forall i = [1, k-1]\,.$$

Покажем теперь, что в этом случае $e'_k \neq o$.

Допустим противное: $e'_k = o$. Поскольку все построенные ранее элементы $e'_i \quad \forall i = [1, k-1]$ суть некоторые линейные комбинации элементов $g_i \quad \forall i = [1, k-1]$, то в силу (10.2.1) мы приходим к линейной зависимости набора элементов $g_i \quad \forall i = [1, k]$, что противоречит условию теоремы. Следовательно, $e'_k \neq o$.

3°. Процесс ортогонализации продолжается до исчерпания множества элементов $g_i \quad \forall i = [1, n]$. Затем достаточно пронормировать элементы $e'_i \ \forall i = [1, n]$, чтобы получить искомый ортонормированный базис

$$\{\, e_1,\, e_2,\, \ldots,\, e_n \,\}, \text{ где } \ e_i = \frac{e'_i}{|e'_i|} \quad \forall i = [1, n] \ .$$

Теорема доказана.

Замечания: 1°. Методом Грама – Шмидта можно выполнять ортогонализацию наборов элементов, являющихся счетными множествами.

2°. Процесс ортогонализации Грама – Шмидта может быть применен к любой, в том числе и к линейно зависимой, системе элементов евклидова пространства.

Если ортогонализуемая система линейно зависима, то на некотором шаге мы получим нулевой элемент, после отбрасывания которого можно продолжить процесс ортогонализации.

10.3. Координатное представление скалярного произведения

Полезным инструментом исследования свойств некоторого набора элементов $\{ f_1, f_2, \ldots, f_k \}$ в евклидовом пространстве E является матрица Грама.

Определение 10.3.1	В евклидовом пространстве E *матрицей Грама* системы элементов $\{ f_1, f_2, \ldots, f_k \}$ называется квадратная, порядка k матрица вида $$\left\|\Gamma\right\|_f = \left\| \begin{matrix} (f_1,f_1) & (f_1,f_2) & \ldots & (f_1,f_k) \\ (f_2,f_1) & (f_2,f_2) & \ldots & (f_2,f_k) \\ \ldots & \ldots & \ldots & \ldots \\ (f_k,f_1) & (f_k,f_2) & \ldots & (f_k,f_k) \end{matrix} \right\| .$$

Матрица Грама симметрическая в силу коммутативности скалярного произведения.

Пусть в E^n дан базис $\{ g_1, g_2, \ldots, g_n \}$. Скалярное произведение элементов $x = \sum\limits_{i=1}^{n} \xi_i g_i$ и $y = \sum\limits_{j=1}^{n} \eta_j g_j$, в силу определения 10.1.1, представимо в виде

$$(x,y) = (\sum_{i=1}^{n} \xi_i g_i, \sum_{j=1}^{n} \eta_j g_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij}\xi_i\eta_j ,$$

где γ_{ij} — компоненты матрицы $\left\|\Gamma\right\|_g$, называемой *базисной матрицей Грама.*

Заметим, что эта матрица является матрицей симметричного билинейного функционала, задающего скалярное произведение. Тогда (принимая во внимание определение 9.1.2) координатное представление скалярного произведения может быть записано так:

$$(x,y) = \left\|x\right\|_g^{\mathrm{T}} \left\|\Gamma\right\|_g \left\|y\right\|_g =$$

$$= \left\| \, \xi_1 \, \xi_2 \, \ldots \xi_n \, \right\| \left\| \begin{matrix} (g_1, g_1) & (g_1, g_2) & \ldots & (g_1, g_n) \\ (g_2, g_1) & (g_2, g_2) & \ldots & (g_2, g_n) \\ \ldots & \ldots & \ldots & \ldots \\ (g_n, g_1) & (g_n, g_2) & \ldots & (g_n, g_n) \end{matrix} \right\| \left\| \begin{matrix} \eta_1 \\ \eta_2 \\ \ldots \\ \eta_n \end{matrix} \right\| ,$$

где $\left\|x\right\|_g$ и $\left\|y\right\|_g$ — координатные представления (столбцы) элементов x и y в базисе $\{ g_1, g_2, \ldots, g_n \}$. Очевидно, что эта формула согласуется с определениями в § 2.3 и § 9.2.

Заметим, наконец, что $\left\|\Gamma\right\|_e = \left\|E\right\|$ в ортонормированном базисе $\{ e_1, e_2, \ldots, e_n \}$, и потому формула для скалярного произведения принимает в этом случае упрощенный вид:

$$(x, y) = \left\|x\right\|_g^{\mathrm{T}} \left\|y\right\|_g = \sum_{i=1}^{n} \xi_i \eta_i \, .$$

Теорема 10.3.1 **Для базисной матрицы Грама $\left\|\Gamma\right\|_g$ в любом базисе $\det\left\|\Gamma\right\|_g > 0$.**

Доказательство.

Из определения 10.1.1 следует, что скалярное произведение есть билинейный, симметричный функционал, поэтому при переходе от базиса $\{ g_1, g_2, \ldots, g_n \}$ к базису $\{ g'_1, g'_2, \ldots, g'_n \}$ (с матрицей перехода $\|S\|$) по теореме 9.1.1 для матрицы Грама имеют место равенства

$$\|\Gamma\|_{g'} = \|S\|^{\mathrm{T}} \|\Gamma\|_g \|S\|, \qquad \det \|\Gamma\|_{g'} = \det \|\Gamma\|_g \det^2 \|S\|,$$

где $\det \|S\| \neq 0$.
Откуда следует, что значение $\operatorname{sgn} \det \|\Gamma\|_g$ инвариантно, то есть не изменяется при замене базиса. А, приняв во внимание, что в ортонормированном базисе $\det \|\Gamma\|_e = 1$, приходим к заключению, что в любом базисе $\det \|\Gamma\|_g > 0$.

Теорема доказана.

Следствие 10.3.1 **Система элементов $\{ f_1, f_2, \ldots, f_k \}$ в E *линейно независима* тогда и только тогда, когда определитель матрицы Грама этой системы положителен.**

Доказательство.

Покажем, что, если элементы $\{ f_1, f_2, \ldots, f_k \}$ линейно зависимы, то определитель их матрицы Грама равен нулю. Действительно, пусть существуют не равные нулю одновременно числа $\lambda_1, \lambda_2, \ldots, \lambda_k,$ такие что $\sum\limits_{i=1}^{k} \lambda_i f_i = o\,.$

Умножив последовательно $\forall j = [1, k]$ это равенство скалярно слева на f_j, получим

$$\sum_{i=1}^{k} \lambda_i (f_j, f_i) = o \qquad \forall j = [1, k]$$

или, в матричном виде,

$$\sum_{i=1}^{k} \lambda_i \left\| \begin{array}{c} (q_1, g_i) \\ (g_2, g_i) \\ \ldots \\ (g_k, g_i) \end{array} \right\| = \left\| \begin{array}{c} 0 \\ 0 \\ \ldots \\ 0 \end{array} \right\| .$$

То есть нетривиальная линейная комбинация столбцов матрицы Грама, имеющая коэффициентами числа λ_i, будет равна нулевому столбцу. Значит, эти столбцы линейно зависимые и тогда будет равен нулю определитель матрицы Грама (см. лемму 6.5.2 и теорему 6.5.2).
С другой стороны, если элементы $\{ f_1, f_2, \ldots, f_k \}$ линейно независимы, то они образуют базис в своей линейной оболочке, и для них справедливо утверждение теоремы 10.3.1.

Следствие доказано.

Теперь можно доказать необходимость в теореме 9.3.4.

Теорема 9.3.4 (критерий Сильвестра) **Для положительной определенности квадратичного функционала в Λ^n необходимо и достаточно, чтобы у его матрицы были** *положительными* **все главные миноры вида**

$$\det \left\| \begin{array}{cccc} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1k} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_{k1} & \varphi_{k2} & \cdots & \varphi_{kk} \end{array} \right\| \qquad \forall k = [1, n]\,.$$

Доказательство необходимости.

1°. В § 10.1 было отмечено, что введение скалярного произведения в линейном пространстве равносильно заданию некоторого симметричного билинейного функционала, порождающего положительно определенный квадратичный функционал.

Обратно, по положительно определенному квадратичному функционалу однозначно восстанавливается породивший его симметричный билинейный функционал, который можно принять за скалярное произведение.

2°. Покажем, что у положительно определенного квадратичного функционала все главные миноры (указанного в условии теоремы вида) его матрицы положительны. Действительно, если ввести в Λ^n скалярное произведение при помощи порождающего его билинейного функционала, то матрица этого квадратичного функционала в базисе $\{ g_1, g_2, \ldots, g_n \}$ есть матрица Грама этого базиса.

Рассмотрим последовательно линейные оболочки систем элементов вида

$$\{ g_1, g_2, \ldots, g_k \} \qquad \forall k = [1, n] .$$

Все эти системы линейно независимые (как подмножества базиса) и по теореме 10.3.1 соответствующие им матрицы Грама имеют положительные определители, то есть

$$\det \begin{Vmatrix} \beta_{11} & \beta_{12} & \ldots & \beta_{1k} \\ \beta_{21} & \beta_{22} & \ldots & \beta_{2k} \\ \ldots & \ldots & \ldots & \ldots \\ \beta_{k1} & \beta_{k2} & \ldots & \beta_{kk} \end{Vmatrix} =$$

$$= \det \begin{Vmatrix} (f_1, f_1) & (f_1, f_2) & \ldots & (f_1, f_k) \\ (f_2, f_1) & (f_2, f_2) & \ldots & (f_2, f_k) \\ \ldots & \ldots & \ldots & \ldots \\ (f_k, f_1) & (f_k, f_2) & \ldots & (f_k, f_k) \end{Vmatrix} > 0 \quad \forall k = [1, n] .$$

Теорема доказана.

Теорема 10.3.2 **Координатный столбец элемента x евклидова пространства E^n в любом базисе $\{ g_1, g_2, \ldots, g_n \}$ может быть представлен в виде**

$$\|x\|_g = \|\Gamma\|_g^{-1} \|b\|_g ,$$

где $\|\Gamma\|_g$ — базисная матрица Грама, а

$$\|b\|_g = \left\| \begin{matrix} (x, g_1) \\ (x, g_2) \\ \ldots \\ (x, g_n) \end{matrix} \right\| .$$

Доказательство.

Умножим последовательно обе части равенства $x = \sum\limits_{i=1}^{n} \xi_i g_i$ скалярно на $g_k \ \ \forall k = [1, n]$. В результате получим систему уравнений

$$\sum_{i=1}^{n} (g_i, g_k)\xi_i = (x, g_k) \qquad \forall k = [1, n] ,$$

основная матрица которой есть базисная матрица Грама. Поскольку в силу теоремы 10.3.1 эта матрица невырожденная, приходим к доказываемой формуле

$$\|x\|_g = \|\Gamma\|_g^{-1} \|b\|_g .$$

Теорема доказана.

Следствие 10.3.2 **В ортонормированном базисе $\{ e_1, e_2, \ldots, e_n \}$ евклидова пространства E^n для любого элемента $x = \sum\limits_{i=1}^{n} \xi_i e_i$ имеют место равенства**

$$\xi_k = (x, e_k) \qquad \forall k = [1, n].$$

Замечание 10.3.1. Хотя данная формула справедлива для E^n, она может быть использована для обобщения понятия координат на случай евклидовых пространств с неограниченным числом линейно независимых элементов (см. § 12.3).

10.4. Ортогональные матрицы в евклидовом пространстве

Согласно определению 5.1.4 квадратная матрица $\|Q\|$, удовлетворяющая соотношению $\|Q\|^{-1} = \|Q\|^{\mathrm{T}}$, называется *ортогональной.*

Для любой ортогональной матрицы справедливы равенства

$$\|Q\|\|Q\|^{\mathrm{T}} = \|Q\|^{\mathrm{T}}\|Q\| = \|E\| \qquad \text{и} \qquad \det\|Q\| = \pm 1\,.$$

Кроме того, в евклидовом пространстве будут справедливы следующие теоремы.

Теорема 10.4.1 **Ортогональные матрицы (и только они!) могут служить в E^n матрицами перехода от одного ортонормированного базиса к другому ортонормированному.**

Доказательство.

Рассмотрим два различных ортонормированных базиса $\{\, e_1, e_2, \ldots, e_n \,\}$ и $\{\, e'_1, e'_2, \ldots, e'_n \,\}$ в E^n с матрицей перехода $\|S\|$ от первого базиса ко второму.

Поскольку в этих базисах матрица Грама единичная, то из соотношения $\|\Gamma\|_{e'} = \|S\|^{\mathrm{T}}\|\Gamma\|_e\|S\|$ следует равенство $\|E\| = \|S\|^{\mathrm{T}}\|E\|\|S\|$, или $\|E\| = \|S\|^{\mathrm{T}}\|S\|$.

Поскольку матрица перехода $\|S\|$ невырожденная, то окончательно имеем $\|S\|^{-1} = \|S\|^{\mathrm{T}}$.

Теорема доказана.

В координатной форме равенство $\|S\|^{\mathrm{T}}\|S\| = \|E\|$ принимает вид

$$\sum_{i=1}^{n} \sigma_{ki}\sigma_{ij} = \delta_{kj} \qquad \forall k, j = [1, n]\,,$$

частный случай которого для $n = 3$ был получен в § 2.9.

Теорема 10.4.2 **Собственные значения линейного оператора, имеющего в E^n ортогональную матрицу, равны по модулю единице.**

Доказательство.

Из равенства $\|\hat{A}\|_g\|f\|_g = \lambda\|f\|$ по правилу транспонирования произведений матриц следует, что $\|f\|_g^{\mathrm{T}}\|\hat{A}\|_g^{\mathrm{T}} = \lambda\|f\|_g^{\mathrm{T}}$.

Перемножив эти равенства почленно, получим

$$\|f\|_g^{\mathrm{T}}\|\hat{A}\|_g^{\mathrm{T}}\|\hat{A}\|_g\|f\|_g = \lambda^2\|f\|_g^{\mathrm{T}}\|f\|_g\,.$$

В силу ортогональности $\|\hat{A}\|_g$ имеем $\|\hat{A}\|_g^{\mathrm{T}}\|\hat{A}\|_g = \|E\|$, а потому $\|f\|_g^{\mathrm{T}}\|f\|_g = \lambda^2\|f\|_g^{\mathrm{T}}\|f\|_g$ и, наконец, $\lambda^2 = 1$, поскольку собственные векторы ненулевые.

Теорема доказана.

Ортогональные матрицы также играют важную роль в вычислительных методах алгебры, что, например, иллюстрирует

Теорема 10.4.3 **Если матрица $\|A\|$ невырожденная, то ее разложение вида $\|A\| = \|Q\|\|R\|$, где $\|Q\|$ — ортогональная матрица, а $\|R\|$ — верхняя треугольная матрица с положительными диагональными элементами, единственно.**

Доказательство.

Предположим, что имеется два разложения

$$\|A\| = \|Q\|_1\|R\|_1 = \|Q\|_2\|R\|_2\,.$$

Из невырожденности $\|A\|$ следует невырожденность $\|R\|_1$ и $\|R\|_2$, поскольку $\|Q\|_1$ и $\|Q\|_2$ ортогональные и, очевидно, невырожденные. Тогда последнее равенство можно переписать в виде $\|Q\|_2^{\mathrm{T}}\|Q\|_1 = \|R\|_2\|R\|_1^{-1}$, где $R\|_1^{-1}$ — также верхняя треугольная матрица.

Заметим, что $\|R\|_2\|R\|_1^{-1}$ есть диагональная матрица. Действительно, с одной стороны, она верхняя треугольная матрица как произведение верхних треугольных. С другой стороны, $\|R\|_2\|R\|_1^{-1}$ должна быть и нижней треугольной, поскольку она ортогональная (как произведение двух ортогональных матриц $\|Q\|_2^{\mathrm{T}}\|Q\|_1$) и ее обратная матрица совпадает с транспонированной.

Очевидно, что диагональная ортогональная матрица может иметь на диагонали лишь элементы, равные по модулю единице. Но диагональные элементы $\|R\|_1$ и $\|R\|_2$ положительны по условию, поэтому остается возможным лишь случай $\|R\|_1\|R\|_2^{-1} = \|E\|$, откуда и следует единственность разложения.

Теорема доказана.

Отметим, что в силу теоремы 10.4.3 решение неоднородной системы линейных уравнений $\|A\|\|x\| = \|b\|$ может быть сведено к разложению невырожденной матрицы $\|A\|$ — на произведение верхней треугольной $\|R\|$ и ортогональной $\|Q\|$, поскольку в этом случае система преобразуется к легко решаемому виду $\|R\|\|x\| = \|Q\|^{\mathrm{T}}\|b\|$.

10.5. Ортогональные дополнения и ортогональные проекции в евклидовом пространстве

Пусть задано множество $\Omega \subseteq E$. Рассмотрим множество $\Omega^{\perp} \subseteq E$, состоящее из *всевозможных* элементов $x \in \Omega^{\perp}$, ортогональных *всем* элементам из Ω.

Определение 10.5.1	В евклидовом пространстве E множество $\Omega^{\perp}$ элементов x, таких, что $(x, y) = 0 \quad \forall y \in \Omega$, называется *ортогональным дополнением* множества Ω.

Теорема 10.5.1 **$\Omega^{\perp}$ — ортогональное дополнение k-мерного подпространства $\Omega \subset E^n$ является подпространством размерности $n - k$.**

Доказательство.

Пусть в E^n со стандартным скалярным произведением дан ортонормированный базис и пусть $\Omega^{\perp} \subset E^n$ — ортогональное дополнение к $\Omega \subset E^n$. Выберем некоторый базис $\{ g_1, g_2, \ldots, g_k \}$ в Ω. Тогда из условия ортогональности произвольного элемента $x \in \Omega^{\perp}$ каждому элементу Ω следует, что

$$(g_j, x) = 0 \qquad \forall j = [1, k]$$

или же в координатной форме:

$$\begin{cases} \varepsilon_{11}\xi_1 + \varepsilon_{12}\xi_2 + \ldots + \varepsilon_{1k}\xi_k = 0, \\ \varepsilon_{21}\xi_1 + \varepsilon_{22}\xi_2 + \ldots + \varepsilon_{2k}\xi_k = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \varepsilon_{k1}\xi_1 + \varepsilon_{k2}\xi_2 + \ldots + \varepsilon_{kk}\xi_k = 0, \end{cases}$$

где $\|g_j\|_e = \|\, \varepsilon_{1j}\, \varepsilon_{1j}\, \ldots \varepsilon_{1j}\,\|^{\mathrm{T}} \quad \forall j = [1,k]$ и соответственно $\|x\|_e = \|\, \xi_1\, \xi_2\, \ldots \xi_k\,\|^{\mathrm{T}}$.

Эта однородная система линейных уравнений (неизвестными в которой являются компоненты элемента x), определяющая ортогональное дополнение $\Omega^{\perp}$, имеет ранг k в силу линейной независимости элементов $\{\, g_1,\, g_2,\, \ldots,\, g_k \,\}$. Тогда по теореме 6.7.1 у нее есть $n - k$ линейно независимых решений, образующих базис подпространства $\Omega^{\perp}$.

Теорема доказана.

Убедимся теперь в справедливости следующего утверждения.

Теорема 10.5.2 **Если $\Omega^{\perp}$ — ортогональное дополнение подпространства $\Omega \subseteq E$, то Ω является в E ортогональным дополнением к $\Omega^{\perp}$.**

Доказательство.

Для каждого элемента $x \in \Omega^{\perp}$ по определению 10.5.1 имеет место равенство $(y, x) = 0 \quad \forall y \in \Omega$.

Но это означает, что $\forall y \in \Omega$ также будет справедливым $(x, y) = 0 \quad \forall y \in \Omega^{\perp}$, то есть Ω является ортогональным дополнением к $\Omega^{\perp}$ в E.

Теорема доказана.

Определение 10.5.2	В евклидовом пространстве E элемент y называется ортогональной проекцией элемента x на подпространство E^*, если $1^\circ. \quad y \in E^*;$ $2^\circ. \quad (x - y, z) = 0 \quad \forall z \in E^*.$

Теорема 10.5.3 **Если $\Omega \subseteq E$ является k-мерным подпространством E, то элемент y — ортогональная проекция x на Ω — существует и единственен.**

Доказательство.

Если в Ω существует базис $\{\, g_1,\, g_2,\, \ldots,\, g_k \,\}$, то элемент $y \in \Omega$ может быть представлен в виде $\quad y = \sum\limits_{j=1}^{k} \xi_j g_j$.

Условие $(x-y, z) = 0 \quad \forall z \in \Omega$ равносильно ортогональности $x - y$ каждому из базисных элементов подпространства Ω, то есть $(x - y, g_j) = 0 \quad \forall j = [1, k]$, и, следовательно, числа $\xi_j \quad \forall j = [1, k]$ могут быть найдены из системы линейных уравнений $\quad (x - \sum\limits_{j=1}^{k} \xi_j g_j,\, g_i) = 0 \quad \forall i = [1, k] \quad$ или

$$\sum_{j=1}^{k} (g_i, g_j)\, \xi_j = (g_i, x) \qquad \forall i = [1, k]\,.$$

Поскольку основная матрица этой системы, как базисная матрица Грама (см. следствие 10.3.1), невырожденная, то по теореме 6.4.1 (Крамера) решение данной системы существует и единственно.

Теорема доказана.

Отметим, что если базис $\{\, e_1,\, e_2,\, \ldots,\, e_k \,\}$ в подпространстве Ω ортонормированный, то ортогональная проекция элемента x на Ω есть элемент вида $\quad y = \sum\limits_{j=1}^{k} (x, e_j) e_j$.

Задача 10.5.1 *В евклидовом пространстве E^4 со стандартным скалярным произведением в некотором ортонормированном базисе система линейных уравнений*

$$\begin{cases} \ \ \xi_1 + \xi_2 - \xi_3 - \xi_4 = 0, \\ 2\xi_1 + \xi_2 \qquad\qquad\ \ = 0 \end{cases}$$

задает подпространство Ω. Найти в этом базисе матрицу оператора ортогонального проектирования элементов E^4 на Ω.

Решение.1°. За базис в Ω примем пару элементов g_1 и g_2, координатные представления которых в исходном базисе $\{\, e_1,\, e_2,\, e_3,\, e_4 \,\}$ суть линейно независимые решения системы линейных уравнений, задающей Ω, например, $\|g_1\|_e = \|\; -1\ 2\ 1\ 0\; \|^{\mathrm{T}}$ и $\|g_2\|_e = \|\; -1\ 2\ 0\ 1\; \|^{\mathrm{T}}$.

2°. Поскольку $\dim \Omega = 2$, то размерность ортогонального дополнения к Ω согласно теореме 10.5.1 также равна 2. За базис в этом ортогональном дополнении удобно принять элементы g_3 и g_4, такие, что

$$\|g_3\|_e = \| \ 1 \ \ 1 \ -1 \ -1 \ \|^{\mathrm{T}}, \qquad \|g_4\|_e = \| \ 2 \ -1 \ \ 0 \ \ 0 \ \|^{\mathrm{T}},$$

поскольку они линейно независимы и ортогональны каждому элементу из подпространства Ω, как образованные из коэффициентов заданной в условии задачи системы линейных уравнений.

3°. Элементы g_1, g_2, g_3 и g_4 линейно независимые по построению и образуют базис в E^4. Значит, каждый элемент из E^4 может быть представлен и притом единственным образом как линейная комбинация элементов этого базиса. Искомый оператор $\hat{A}$ ортогонального проектирования элементов E^4 на Ω должен, очевидно, удовлетворять соотношениям

$$\hat{A}g_1 = g_1, \quad \hat{A}g_2 = g_2, \quad \hat{A}g_3 = o, \quad \hat{A}g_4 = o\,,$$

в силу которых его матрица в базисе $\{\, g_1,\ g_2,\ g_3,\ g_4 \,\}$ будет иметь следующий вид:

$$\left\|\hat{A}\right\|_g = \left\| \begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} \right\| .$$

4°. С другой стороны, матрица перехода $\|S\|$ от базиса $\{\, e_1,\ e_2,\ e_3,\ e_4 \,\}$ к базису $\{\, g_1,\ g_2,\ g_3,\ g_4 \,\}$ есть

$$\|S\| = \left\| \begin{matrix} -1 & -1 & 1 & 2 \\ 2 & 2 & 1 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{matrix} \right\| .$$

А поскольку $\left\|\hat{A}\right\|_g = \|S\|^{-1} \left\|\hat{A}\right\|_e \|S\|$ и, следовательно, $\left\|\hat{A}\right\|_e = \|S\| \left\|\hat{A}\right\|_g \|S\|^{-1}$, то, воспользовавшись определением 5.1.1 и методом, описанным в § 6.8 (с. 207), найдем, что

$$\left\|\hat{A}\right\|_e =$$

$$= \left\| \begin{array}{rrrr} -1 & -1 & 1 & 2 \\ 2 & 2 & 1 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{array} \right\| \left\| \begin{array}{rrrr} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right\| \left\| \begin{array}{rrrr} -1 & -1 & 1 & 2 \\ 2 & 2 & 1 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{array} \right\|^{-1} =$$

$$= \frac{1}{11} \left\| \begin{array}{rrrr} 2 & -4 & -1 & -1 \\ -4 & 8 & 2 & 2 \\ -1 & 2 & 6 & -5 \\ -1 & 2 & -5 & 6 \end{array} \right\| .$$

`Решение получено.`

`Замечание 10.5.1.` Геометрическая интерпретация ортогонального проектирования вполне очевидна, однако эта операция используется и в других приложениях. Например, если E есть евклидово пространство непрерывных на $[\alpha, \beta]$ функций $x(\tau)$ со скалярным произведением

$$(x, y) = \int\limits_{\alpha}^{\beta} x(\tau) y(\tau)\, d\tau\, ,$$

а $\Omega \subset E$ — подпространство алгебраических многочленов степени не выше, чем n, то ортогональная проекция $x(\tau)$ — элемента E на Ω может рассматриваться как наилучшее приближение на $[\alpha, \beta]$ непрерывной функции линейной комбинацией степенных многочленов.

10.6. Сопряженные операторы в евклидовом пространстве

Поскольку евклидово пространство является частным случаем линейного пространства, то все изложенные в главе 8 утверждения справедливы и для линейных операторов, действующих в евклидовом пространстве.

Однако операция скалярного произведения позволяет выделять в евклидовых пространствах специфические классы линейных операторов, обладающих рядом полезных свойств.

Определение 10.6.1	Линейный оператор $\hat{A}^{+}$, заданный в евклидовом пространстве E, называется *сопряженным* линейному оператору $\hat{A}$, если $\forall x, y \in E$ имеет место равенство $(\hat{A}x, y) = (x, \hat{A}^{+}y)$.

Пример 10.6.1 В евклидовом пространстве, образованном бесконечно дифференцируемыми функциями, равными нулю вне некоторого конечного интервала, со скалярным произведением

$$(x, y) = \int\limits_{-\infty}^{+\infty} x(\tau)y(\tau)\,d\tau\,,$$

для линейного оператора дифференцирования $\hat{A} = \dfrac{d}{d\tau}$ сопряженным ему будет оператор $\hat{A}^{+} = -\dfrac{d}{d\tau}$.

Действительно, согласно правилу интегрирования несобственных интегралов по частям это следует из определения 10.6.1 и равенств

$$(\hat{A}x, y) = \int\limits_{-\infty}^{+\infty} \frac{dx(\tau)}{d\tau} y(\tau)\,d\tau =$$

$$= x(\tau)y(\tau)\Bigg|_{-\infty}^{+\infty} + \int\limits_{-\infty}^{+\infty} x(\tau)\left(-\frac{dy(\tau)}{d\tau}\right)\,d\tau = (x, \hat{A}^{+}y)\,,$$

поскольку проинтегрированная часть равна нулю $\forall x(\tau), y(\tau) \in E$.

Рассмотрим теперь конечномерное евклидово пространство E^n с базисом $\{\, g_1,\, g_2,\, \ldots,\, g_n \,\}$ и выясним связь матриц линейных операторов $\hat{A}^{+}$ и $\hat{A}$ в этом базисе, предположив, что сопряженный оператор существует.

Пусть в данном базисе матрицы операторов $\hat{A}^{+}$ и $\hat{A}$ имеют соответственно вид $\left\|\hat{A}^{+}\right\|_g$ и $\left\|\hat{A}\right\|_g$, а координатные представления элементов

x и y суть $\|x\|_g = \|\,\xi_1\,\xi_2\,\ldots\xi_k\,\|^{\mathrm{T}}$. и $\|y\|_g = \|\,\eta_1\,\eta_2\,\ldots\eta_k\,\|^{\mathrm{T}}$, тогда равенство $(\hat{A}x, y) = (x, \hat{A}^{+}y)$ можно записать как

$$\left(\left\|\hat{A}\right\|_g \|x\|_g\right)^{\mathrm{T}} \|\Gamma\|_g \|\|y\|_g = \|x\|_g^{\mathrm{T}} \|\Gamma\|_g \left\|\hat{A}^{+}\right\|_g \|y\|_g\,, \tag{10.6.1}$$

где $\|\Gamma\|_g$ — матрица Грама выбранного в E^n базиса.

В силу соотношения $(\|A\|\|B\|)^{\mathrm{T}} = \|B\|^{\mathrm{T}}\|A\|^{\mathrm{T}}$ и свойства дистрибутивности умножения матриц равенство (10.6.1) преобразуется к виду

$$\|x\|_g^{\mathrm{T}} \left(\left\|\hat{A}\right\|_g^{\mathrm{T}} \|\Gamma\|_g - \|\Gamma\|_g \left\|\hat{A}^{+}\right\|_g\right) \|y\|_g = 0\,.$$

Поскольку это равенство справедливо при любых x и y, то, приняв во внимание невырожденность матрицы Грама и проведя рассуждения, аналогичные использованным при доказательстве леммы 5.1.2, заключаем, что матрица, стоящая в круглых скобках, нулевая. Тогда из соотношения $\left\|\hat{A}\right\|_g^{\mathrm{T}} \|\Gamma\|_g - \|\Gamma\|_g \left\|\hat{A}^{+}\right\|_g = \|O\|$ следует

$$\left\|\hat{A}^{+}\right\|_g = \|\Gamma\|_g^{-1} \left\|\hat{A}\right\|_g^{\mathrm{T}} \|\Gamma\|_g\,, \tag{10.6.2}$$

которое, в частности, для ортонормированного базиса $\{\, e_1,\, e_2,\, \ldots,\, e_n \,\}$ имеет вид $\left\|\hat{A}^{+}\right\|_e = \left\|\hat{A}\right\|_e^{\mathrm{T}}$.

Лемма 10.6.1 **Если $(x, \hat{A}y) = 0 \quad \forall x, y \in E$, то оператор $\hat{A}$ нулевой.**

Доказательство.

Пусть справедливо равенство $(x, \hat{A}y) = 0 \quad \forall x, y \in E$. Тогда оно будет верным и для $x = \hat{A}y$.

Но из равенства $(\hat{A}y, \hat{A}y) = 0 \quad \forall y \in E$ согласно определению 10.1.1 следует, что $\hat{A}y = o$.

Наконец, в силу произвольности элемента y и определения 8.2.2 приходим к заключению, что $\hat{A} = \hat{O}$.

Лемма доказана.

Теорема 10.6.1 **Каждый линейный оператор в евклидовом пространстве E^n имеет единственный сопряженный оператор.**

Доказательство.

Существование в E^n оператора $\hat{A}^+$, сопряженного оператору $\hat{A}$, следует из возможности построения для любого линейного оператора с матрицей $\|\hat{A}\|$ линейного оператора с матрицей, определяемой формулой (10.6.2).

Покажем теперь, что, если оператор $\hat{A}^+$ существует в E, то он единственный. Предположим, что $\hat{A}$ имеет два сопряженных оператора $\hat{A}^+$ и $\hat{A}^\times$. Это означает, что $\forall x, y \in E^n$ одновременно выполнены равенства

$$(\hat{A}x, y) = (x, \hat{A}^+ y) \qquad \text{и} \qquad (\hat{A}x, y) = (x, \hat{A}^\times y)\,.$$

Вычитая их почленно, получаем $\quad 0 = (x, \left(\hat{A}^+ - \hat{A}^\times\right) y)$, но тогда по лемме 10.6.1

$$\hat{A}^+ - \hat{A}^\times = \hat{O} \qquad \Longrightarrow \qquad \hat{A}^+ = \hat{A}^\times\,.$$

Теорема доказана.

Теорема 10.6.2 **Для любых линейных операторов $\hat{A}$ и $\hat{B}$ в E имеет место равенство** $\quad \left(\hat{A}\hat{B}\right)^+ = \hat{B}^+\hat{A}^+\,.$

Доказательство.

Имеет место $\forall x, y \in E$:

$$(\left(\hat{A}\hat{B}\right)^+ x, y) = (x, \hat{A}\hat{B}y) = (\hat{A}^+x, \hat{B}y) = (\hat{B}^+\hat{A}^+x, y)\,.$$

Откуда следует $\quad (\left(\left(\hat{A}\hat{B}\right)^+ - \hat{B}^+\hat{A}^+\right) x, y) = 0 \quad \forall x, y \in E\,,$ но тогда по лемме 10.6.1

$$\left(\hat{A}\hat{B}\right)^+ - \hat{B}^+\hat{A}^+ = \hat{O} \qquad \Longrightarrow \qquad \left(\hat{A}\hat{B}\right)^+ = \hat{B}^+\hat{A}^+\,.$$

Теорема доказана.

Теорема 10.6.3 **Для линейного оператора $\hat{A}$ в E имеет место равенство** $\quad \left(\hat{A}^+\right)^+ = \hat{A}\,.$

Доказательство.

Имеет место $\forall x, y \in E$:

$$((\hat{A}^{+})^{+} x, y) = (x, \hat{A}^{+} y) = (\hat{A} x, y)\,.$$

Откуда следует $\quad (\left((\hat{A}^{+})^{+} - \hat{A}\right) x, y) = 0 \quad \forall x, y \in E\,,\quad$ но тогда по лемме 10.6.1

$$(\hat{A}^{+})^{+} - \hat{A} = \hat{O} \qquad \Longrightarrow \qquad (\hat{A}^{+})^{+} = \hat{A}\,.$$

Теорема доказана.

Теорема 10.6.4 **Ортогональное дополнение в E^n области значений оператора $\hat{A}$ является ядром оператора $\hat{A}^{+}$.**

Доказательство.

1°. Покажем вначале, что ядро оператора $\hat{A}$, обозначаемое как $\ker \hat{A}$, содержится во множестве $\left(\mathrm{Im}\hat{A}\,\right)^{\perp}$ — ортогональном дополнении области значений оператора $\hat{A}$. Действительно, любой элемент $y \in \ker \hat{A}^{+}$, то есть такой, что $\hat{A}^{+} y = o$, ортогонален элементу $\quad b = \hat{A} x \quad \forall x \in E^n$, поскольку $\quad (b, y) = (\hat{A} x, y) = (x, \hat{A}^{+} y) = (x, o) = 0\,.$

2°. Теперь сравним размерности $\ker \hat{A}$ и $\left(\mathrm{Im}\hat{A}\,\right)^{\perp}$. С одной стороны, в силу невырожденности базисной матрицы Грама, формулы 10.6.2 и теоремы 8.4.3

$$\dim\ker \hat{A}^{+} = n - \mathrm{rg}\left\|\hat{A}^{+}\right\| = n - \mathrm{rg}\left(\|S\|^{-1}\left\|\hat{A}\right\|^{\mathrm{T}}\|S\|\right) =$$

$$= n - \mathrm{rg}\left\|\hat{A}\right\|^{\mathrm{T}} = n - \mathrm{rg}\left\|\hat{A}\right\|\,.$$

Но, с другой стороны, по теореме 8.4.1 размерность области значений $\hat{A}$ равна $\mathrm{rg}\left\|\hat{A}\right\|$, и тогда по теореме 10.5.1 $\left(\mathrm{Im}\hat{A}\,\right)^{\perp} = n - \mathrm{rg}\left\|\hat{A}\right\|$.

Наконец, из равенства $\dim\ker \hat{A}^{+} = \dim\left(\mathrm{Im}\hat{A}\,\right)^{\perp}$ и соотношения $\ker \hat{A}^{+} \subseteq \left(\mathrm{Im}\hat{A}\,\right)^{\perp}$ следует совпадение множеств $\quad \ker \hat{A}^{+}$ и $\left(\mathrm{Im}\hat{A}\,\right)^{\perp}$.

Теорема доказана.

Замечание 10.6.1. 1) В использованных обозначениях теорема 10.6.4 допускает формулировку, совпадающую с формулировкой теоремы 6.7.3 (Фредгольма), поскольку равенство $\hat{A}x = b$ означает, что элемент b принадлежит области значений линейного оператора $\hat{A}$.

2) В предположении, что столбцы $\|y\|$ и $\|b\|$ суть координатные представления элементов E^m в ортонормированном базисе, теорема допускает также и нижеследующую формулировку.

Теорема 10.6.5 (Фредгольма) **Система линейных уравнений $\|A\|\|x\| = \|b\|$ совместна тогда и только тогда, когда каждое решение однородной системы вида $\|A\|^{\mathrm{T}}\|y\| = \|o\|$ ортогонально столбцу свободных членов $\|b\|$.**

10.7. Самосопряженные операторы

Определение 10.7.1 Линейный оператор $\hat{R}$ в евклидовом пространстве E называется *самосопряженным*, если $\forall x, y \in E$ имеет место равенство $(\hat{R}x, y) = (x, \hat{R}y)$.

Пример 10.7.1 В евклидовом пространстве E операторы вида $\hat{A} + \hat{A}^+$, $\hat{A}\hat{A}^+$ и $\hat{A}^+\hat{A}$ будут самосопряженными для каждого линейного оператора $\hat{A}$.

Действительно, для оператора $\hat{A}^+\hat{A}$, например, имеем, что $\forall x, y \in E \quad (\hat{A}^+\hat{A}x, y) = (\hat{A}x, \hat{A}y) = (x, \hat{A}^+\hat{A}y)$, откуда и следует его самосопряженность.

Свойства самосопряженных операторов сформулируем в виде следующих утверждений.

Лемма 10.7.1 **Линейный оператор $\hat{R}$ в E^n является самосопряженным тогда и только тогда, когда его матрица в каждом *ортонормированном* базисе *симметрическая*.**

Доказательство.

Из определения 10.7.1 и формулы

$$\left\|\hat{R}^+\right\|_g = \|\Gamma\|_g^{-1} \left\|\hat{R}\right\|_g^{\mathrm{T}} \|\Gamma\|_g$$

для некоторого ортонормированного базиса $\{ e_1, e_2, \ldots, e_n \}$ в силу самосопряженности $\hat{R}$ имеем $\left\|\hat{R}^+\right\|_e = \left\|\hat{R}\right\|_e^{\mathrm{T}}$.

Перейдем теперь к другому ортонормированному базису $\{ e'_1, e'_2, \ldots, e'_n \}$. Матрица перехода $\|S\|$, как было показано в § 10.4, при такой замене ортогональная, то есть для нее $\|S\|^{\mathrm{T}} = \|S\|^{-1}$. Поэтому

$$\left\|\hat{R}^+\right\|_{e'}^{\mathrm{T}} = \left(\|S\|^{-1} \left\|\hat{R}^+\right\|_e \|S\|\right)^{\mathrm{T}} = \left(\|S\|^{\mathrm{T}} \left\|\hat{R}^+\right\|_e \|S\|\right)^{\mathrm{T}} =$$

$$= \|S\|^{\mathrm{T}} \left\|\hat{R}^+\right\|_e^{\mathrm{T}} \left(\|S\|^{\mathrm{T}}\right)^{\mathrm{T}} = \|S\|^{\mathrm{T}} \left\|\hat{R}^+\right\|_e \|S\| =$$

$$= \|S\|^{-1} \left\|\hat{R}^+\right\|_e \|S\| = \left\|\hat{R}^+\right\|_{e'} .$$

Верно и обратное: если $\left\|\hat{R}^+\right\|_e = \left\|\hat{R}\right\|_e^{\mathrm{T}}$, то $\forall x, y \in E^n$

$$(\hat{R}x, y) = \left\|\hat{R}x\right\|_e^{\mathrm{T}} \|y\|_e = \left(\left\|\hat{R}\right\|_e \|x\|_e\right)^{\mathrm{T}} \|y\|_e =$$

$$= \|x\|_e^{\mathrm{T}} \left\|\hat{R}\right\|_e^{\mathrm{T}} \|y\|_e = \|x\|_e^{\mathrm{T}} \left\|\hat{R}\right\|_e \|y\|_e =$$

$$= \|x\|_e^{\mathrm{T}} \left\|\hat{R}y\right\|_e = (x, \hat{R}y) ,$$

то есть оператор $\hat{R}$ — самосопряженный.

Лемма доказана.

Признак самосопряженности может быть сформулирован как

Следствие 10.7.1 **Если линейный оператор в E^n имеет** *симметрическую* **матрицу в некотором ортонормированном базисе, то он** *самосопряженный*.

Лемма 10.7.2 **Все собственные значения самосопряженного оператора $\hat{R}$ в E^n** *вещественные* **числа.**

Доказательство.

Допустим противное: пусть характеристическое уравнение самосопряженного оператора $\hat{R}$ имеет комплексный корень $\lambda = \alpha + i\beta$, где $\beta \neq 0$.

По теореме 8.6.2 оператор $\hat{R}$ в этом случае имеет двумерное инвариантное подпространство. Было показано, что в этом случае существует пара линейно независимых элементов x и x таких, что

$$\begin{cases} \hat{R}x = \alpha x - \beta y\,, \\ \hat{R}y = \alpha y + \beta x\,. \end{cases}$$

Умножая эти равенства скалярно: первое — справа на y, второе — слева на x, получим

$$\begin{cases} (\hat{R}x, y) = \alpha(x, y) - \beta(y, y)\,, \\ (x, \hat{R}y) = \alpha(x, y) + \beta(y, y)\,. \end{cases}$$

Вычитая почленно второе равенство из первого и принимая во внимание самосопряженность $\hat{R}$, приходим к заключению, что $\beta = 0$. Однако это противоречит предположению о том, что собственное значение λ невещественное.

Лемма доказана.

Лемма 10.7.3 **Собственные векторы самосопряженного оператора действующего в E, отвечающие** *различным* **собственным значениям,** *ортогональны***.**

Доказательство.

Пусть для самосопряженного оператора $\hat{R}$ имеют место равенства

$$\hat{R}f_1 = \lambda_1 f_1\,, \qquad \hat{R}f_2 = \lambda_2 f_2\,,$$

где ненулевые элементы f_1 и f_2 — собственные векторы оператора $\hat{R}$ и $\lambda_1 \neq \lambda_2$ — соответствующие им собственные значения.

Умножив эти равенства соответственно: первое — скалярно справа на f_2, второе — скалярно слева на f_1, получим

$$\begin{cases} (\hat{R}f_1, f_2) = \lambda_1(f_1, f_2)\,, \\ (f_1, \hat{R}f_2) = \lambda_2(f_1, f_2)\,. \end{cases}$$

Вычитая эти равенства почленно и учитывая, что $\hat{R}$ — самосопряженный оператор, приходим к скалярному равенству $(\lambda_1 - \lambda_2)(f_1, f_2) = 0,$ откуда, в силу $\lambda_1 \neq \lambda_2$, получаем, что $(f_1, f_2) = 0$.

Лемма доказана.

Лемма 10.7.4 **Пусть Ω — инвариантное подпространство самосопряженного оператора $\hat{R}$, действующего в E , и пусть $\Omega^{\perp}$ — ортогональное дополнение к Ω в E. Тогда $\Omega^{\perp}$ — также инвариантное подпространство оператора $\hat{R}$.**

Доказательство.

Ω инвариантно для оператора $\hat{R}$, то есть $\forall x \in \Omega$ также и $\hat{R}x \in \Omega$. Если $\Omega^{\perp}$ — ортогональное дополнение Ω, то $\forall x \in \Omega$ и $\forall y \in \Omega^{\perp}$ будет $(x, y) = 0$.

Поскольку Ω — инвариантное подпространство $\hat{R}$, то будет также иметь место $(\hat{R}x, y) = 0$. Но в силу самосопряженности $\hat{R}$ тогда верно и $(x, \hat{R}y) = 0$.

Последнее равенство означает, что $\forall y \in \Omega^{\perp}$ имеет место $\hat{R}y \in \Omega^{\perp}$, то есть подпространство $\Omega^{\perp}$ инвариантно для оператора $\hat{R}$.

Лемма доказана.

Теорема 10.7.1 **Для любого самосопряженного оператора $\hat{R}$ в E^n существует ортонормированный базис, состоящий из собственных векторов $\hat{R}$.**

Доказательство.

Для самосопряженного оператора $\hat{R}$ в E^n по лемме 10.7.2 существует, по крайней мере, одно вещественное собственное значение λ_1.

Из системы уравнений (8.5.1) находим отвечающий λ_1 собственный вектор e_1. Без ограничения общности можно считать, что $|e_1| = 1$. Если $n = 1$, то доказательство завершено.

Рассмотрим E^1 — линейную оболочку элемента e_1, являющуюся одномерным инвариантным собственным подпространством $\hat{R}$. Пусть E^{n-1} — ортогональное дополнение к E^1. Тогда по лемме 10.7.4 E^{n-1} — также инвариантное подпространство оператора $\hat{R}$.

Инвариантность E^{n-1} относительно $\hat{R}$ позволяет рассматривать его как оператор, действующий только в E^{n-1}. При этом очевидно, что $\hat{R}$ — самосопряженный оператор, поскольку из условия $(\hat{R}x, y) = (x, \hat{R}y) \quad \forall x, y \in E^n$ следует, что $(\hat{R}x, y) = (x, \hat{R}y) \quad \forall x, y \in E^{n-1}$.

Применяя в E^{n-1} изложенные выше рассуждения, найдем λ_2 — новое собственное значение $\hat{R}$ и соответствующий ему собственный вектор e_2. Без ограничения общности можно считать, что $|e_2| = 1$. При этом λ_2 может случайно совпасть с λ_1, однако из построения ясно, что $(e_2, e_1) = 0$.

Если $n = 2$, то построение базиса завершено. Иначе рассмотрим E^2 — линейную оболочку $\{e_1, e_2\}$ и ее ортогональное дополнение E^{n-2}, найдем новое собственное значение λ_3 и соответствующий ему собственный вектор e_3 и т.д.

Аналогичные рассуждения проводим до исчерпания E^n.

Теорема доказана.

Следствие 10.7.2 **В базисе, построенном в теореме 10.7.1, самосопряженный оператор $\hat{R}$ имеет диагональную матрицу в E^n.**

Доказательство.

Вытекает из замечания о важности собственных векторов, приведенного в § 8.5.

Следствие доказано.

Следствие 10.7.3 **Размерность собственного инвариантного подпространства, отвечающего некоторому собственному значению самосопряженного оператора, равна кратности этого собственного значения.**

Доказательство.

Из доказательства теоремы 10.7.1 следует, что число построенных в нем линейно независимых элементов не зависит от совпадения или несовпадения значений собственных чисел оператора $\hat{R}$.

Следствие доказано.

Следствие 10.7.4 **Если линейный оператор $\hat{A}$ в E^n имеет n попарно ортогональных собственных векторов, то он самосопряженный.**

Доказательство.

Пронормируем собственные векторы оператора $\hat{A}$ и примем их за ортонормированный базис, в котором матрица этого линейного оператора $\|\hat{A}\|_e$ диагональная и, следовательно, симметрическая. Тогда в силу леммы 10.7.1 линейный оператор $\hat{A}$ самосопряженный.

Следствие доказано.

Следствие 10.7.5 **Если $\|R\|$ — симметрическая матрица, то существует $\|Q\|$ — ортогональная матрица, такая, что матрица**

$$\|D\| = \|Q\|^{\mathrm{T}}\|R\|\|Q\| = \|Q\|^{-1}\|R\|\|Q\|$$

диагональна.

Доказательство.

В ортонормированном базисе симметрическая матрица $\|R\|$ определяет самосопряженный оператор в E^n, поэтому в качестве искомой матрицы $\|Q\|$ можно выбрать матрицу перехода от данного ортонормированного базиса к ортонормированному базису, образованному собственными векторами этого оператора по схеме, использованной в доказательстве теоремы 10.7.1.

Следствие доказано.

Теорема 10.7.2 **Два самосопряженных оператора $\hat{A}$ и $\hat{B}$ имеют общую систему собственных векторов в E^n тогда и только тогда, когда $\hat{A}\hat{B} = \hat{B}\hat{A}$.**

Доказательство.

Докажем необходимость.

Пусть $\hat{A}a = \lambda a$ и $\hat{B}a = \mu a$, тогда

$$\hat{B}\hat{A}a = \hat{B}\lambda a = \lambda\hat{B}a = \lambda\mu a \qquad \hat{A}\hat{B}Aa = \hat{A}\mu a = \mu\hat{A}a = \mu\lambda a$$

и, вычитая почленно, получим, что $\left(\hat{A}\hat{B} - \hat{B}\hat{A}\right)a = o$.

Поскольку элемент a — произвольный собственный вектор, то данное соотношение верно и для всей совокупности собственных векторов, а значит, и для любого элемента в E^n, так как из собственных векторов можно образовать базис. Поэтому $\hat{A}\hat{B} - \hat{B}\hat{A} = \hat{O}$. Докажем достаточность.

Пусть самосопряженные операторы $\hat{A}$ и $\hat{B}$ коммутируют и пусть, кроме того, $\hat{A}a = \lambda a$.

Рассмотрим здесь лишь случай, когда все собственные значения оператора $\hat{A}$ различны.

Покажем, что элемент евклидова пространства $b = \hat{B}a$ является собственным вектором оператора $\hat{A}$. Действительно, в силу $\hat{A}\hat{B} = \hat{B}\hat{A}$ имеем

$$\hat{A}b = \hat{A}\hat{B}a = \hat{B}\hat{A}a = \hat{B}\lambda a = \lambda\hat{B}a = \lambda b\,.$$

С другой стороны, если все собственные значения кратности единица, то λ — это собственное значение, которое отвечает a и b одновременно. Поэтому $b = \kappa a$, и из $b = \hat{B}a$ следует $\hat{B}a = \kappa a$. Значит, a — собственный вектор оператора $\hat{B}$.

Теорема доказана.

10.8. Ортогональные операторы

Определение 10.8.1	Линейный оператор $\hat{Q}$, действующий в евклидовом пространстве E, называется *ортогональным* (или *изометрическим*), если $\forall x, y \in E$ имеет место равенство $(\hat{Q}x, \hat{Q}y) = (x, y)$.

Из определения 10.8.1 следует, что ортогональный оператор сохраняет нормы и величины углов между элементами.

Действительно,

$$\left|\hat{Q}x\right| = \sqrt{(\hat{Q}x, \hat{Q}x)} = \sqrt{(x, x)} = |x| \,,$$

$$\cos\psi = \frac{(\hat{Q}x, \hat{Q}y)}{\left|\hat{Q}x\right| \left|\hat{Q}y\right|} = \frac{(x, y)}{|x|\,|y|} = \cos\varphi \,,$$

где ψ — величина угла между элементами $\hat{Q}x$ и $\hat{Q}y$, а φ — величина угла между ненулевыми элементами x и y.

Теорема 10.8.1 **Если ортогональный оператор $\hat{Q}$ имеет сопряженный оператор $\hat{Q}^{+}$, то он имеет и обратный оператор $\hat{Q}^{-1}$, причем $\hat{Q}^{-1} = \hat{Q}^{+}$.**

Доказательство.

По определению 10.8.1

$$(\hat{Q}x, \hat{Q}y) = (x, y) \qquad \forall x, y \in E \,.$$

Откуда следует, что $(\hat{Q}^{+}\hat{Q}x, y) = (x, y)$ или, что то же самое, $(\left(\hat{Q}^{+}\hat{Q} - \hat{E}\right) x, y) = 0$. Это, в силу леммы 10.6.1, дает $\hat{Q}^{+}\hat{Q} = \hat{E}$.

Из последнего равенства вытекает, что $\hat{Q}^{+}\hat{Q}\hat{Q}^{+} = \hat{E}\hat{Q}^{+}$, а в силу того, что единичный оператор коммутирует с любым другим, получаем $\hat{Q}^{+}\hat{Q}\hat{Q}^{+} = \hat{Q}^{+}\hat{E}$ или $\hat{Q}\hat{Q}^{+} = \hat{E}$.

Наконец, по определению 8.2.8 в силу $\hat{Q}^{+}\hat{Q} = \hat{Q}\hat{Q}^{+} = \hat{E}$ приходим к $\hat{Q}^{-1} = \hat{Q}^{+}$.

Теорема доказана.

Следствие 10.8.1 **Операторы $\hat{Q}^{-1}$ и $\hat{Q}^{+}$ также ортогональные.**

Теорема 10.8.2 **Матрица ортогонального оператора в E^n ортогональная в каждом ортонормированном базисе.**

Доказательство.

Пусть оператор $\hat{Q}$ ортогональный. Тогда по теореме 10.8.1 $\hat{Q}^{-1} = \hat{Q}^{+}$ и в силу § 8.3 (4°) в любом ортонормированном базисе справедливы равенства

$$\left\|\hat{Q}\right\|_e^{-1} = \left\|\hat{Q}^{-1}\right\|_e = \left\|\hat{Q}^{+}\right\|_e = \left\|\hat{Q}\right\|_e^{\mathrm{T}} .$$

Но тогда $\left\|\hat{Q}\right\|_e^{-1} = \left\|\hat{Q}\right\|_e^{\mathrm{T}}$, что и означает, согласно определению 5.1.4, ортогональность матрицы $\left\|\hat{Q}\right\|_e$.

Теорема доказана.

Признак ортогональности линейного оператора в E^n дает

Теорема 10.8.3 **Для того чтобы линейный оператор в E^n был ортогональным, достаточно, чтобы его матрица была ортогональной в** *некотором* **ортонормированном базисе.**

Доказательство.

1°. Пусть у линейного оператора $\hat{Q}$ его матрица ортогональная в некотором ортонормированном базисе $\{ e_1, e_2, \ldots, e_n \}$, то есть $\left\|\hat{Q}\right\|_e^{-1} = \left\|\hat{Q}\right\|_e^{\mathrm{T}}$. Тогда

$$(\hat{Q}x, \hat{Q}y) = \left\|\hat{Q}x\right\|_e^{\mathrm{T}} \left\|\hat{Q}y\right\|_e = \left(\left\|\hat{Q}\right\|_e \|x\|_e\right)^{\mathrm{T}} \left\|\hat{Q}\right\|_e \|y\|_e =$$

$$= \|x\|_e^{\mathrm{T}} \left\|\hat{Q}\right\|_e^{\mathrm{T}} \left\|\hat{Q}\right\|_e \|y\|_e = \|x\|_e^{\mathrm{T}} \left\|\hat{Q}\right\|_e^{-1} \left\|\hat{Q}\right\|_e \|y\|_e =$$

$$= \|x\|_e^{\mathrm{T}} \|E\| \, \|y\|_e = \|x\|_e^{\mathrm{T}} \|y\|_e = (x, y) \qquad \forall x, y \in E .$$

Значит, условие ортогональности $\hat{Q}$ выполнено в данном ортонормированном базисе.

2°. Перейдем теперь к $\{ e'_1, e'_2, \ldots, e'_n \}$ — некоторому другому ортонормированному базису и убедимся, что условие ортогональности при этом переходе не нарушится.

Действительно, в силу ортогональности матрицы перехода $\|S\|$, связывающей два ортонормированных базиса, имеем

$$\|\hat{Q}\|_{e'}^{-1} = \Big(\|S\|^{-1}\|\hat{Q}\|_e\|S\|\Big)^{-1} = \|S\|^{-1}\|\hat{Q}\|_e^{-1}\left(\|S\|^{-1}\right)^{-1} =$$

$$= \|S\|^{-1}\|\hat{Q}\|_e^{-1}\|S\| = \|S\|^{\mathrm{T}}\|\hat{Q}\|_e^{\mathrm{T}}\|S\| = \|S\|^{\mathrm{T}}\|\hat{Q}\|_e^{\mathrm{T}}\big(\|S\|^{\mathrm{T}}\big)^{\mathrm{T}} =$$

$$= \Big(\|S\|^{\mathrm{T}}\|\hat{Q}\|_e\|S\|\Big)^{\mathrm{T}} = \Big(\|S\|^{-1}\|\hat{Q}\|_e\|S\|\Big)^{\mathrm{T}} = \|\hat{Q}\|_{e'}^{\mathrm{T}} .$$

Теорема доказана.

В ряде приложений оказывается полезной

Теорема 10.8.4 (О полярном разложении) **Любой линейный оператор $\hat{A}$ в E^n с $\det\|\hat{A}\| \neq 0$ может быть единственным образом представлен в виде $\hat{A} = \hat{Q}\hat{R}$, где оператор $\hat{Q}$ — ортогональный, а оператор $\hat{R}$ — самосопряженный и имеющий положительные собственные значения.**

Доказательство.

1°. Покажем вначале, что самосопряженный оператор $\hat{A}^+\hat{A}$ (см. пример 10.7.1) имеет только положительные собственные значения. Действительно, пусть $\hat{A}^+\hat{A}f = \lambda f$, тогда, с одной стороны,

$$(\hat{A}^+\hat{A}f, f) = (\hat{A}f, \hat{A}f) > 0$$

при $f \neq o$, а с другой:

$$(\hat{A}^+\hat{A}f, f) = (\hat{A}f, \hat{A}f) = \lambda(f, f) > 0 .$$

Но тогда $\lambda > 0$ в силу аксиоматики евклидова пространства, поскольку из предположения, что $\hat{A}f = o$ при $f \neq o$, следует

$$\hat{A}f = 0f \qquad \Longrightarrow \qquad \det\|\hat{A}\| = 0 ,$$

но это противоречит условию теоремы.

2°. Пусть $\{ e_1, e_2, \ldots, e_n \}$ — ортонормированный базис, состоящий из собственных векторов оператора $\hat{A}^+\hat{A}$. Рассмотрим множество элементов $\hat{A}e_i \quad \forall i = [1, n]$, для которых имеем

$$(\hat{A}e_i, \hat{A}e_j) = (\hat{A}^+\hat{A}e_i, e_j) = (\lambda_i e_i, e_j) = \lambda_i \delta_{ij} \quad \forall i, j = [1, n]\,.$$

Но это означает, что $\left\{ e_i' = \dfrac{1}{\sqrt{\lambda_i}} \hat{A}e_i \quad \forall i = [1, n] \right\}$ — также базис и притом ортонормированный.

3°. Если за искомый оператор $\hat{Q}$ мы принимаем ортогональный оператор, переводящий ортонормированный базис $\{ e_1, e_2, \ldots, e_n \}$ в ортонормированный базис $\{ e_1', e_2', \ldots, e_n' \}$, то в качестве $\hat{R}$ можно будет взять оператор $\hat{Q}^{-1}\hat{A}$.

Действительно, во-первых, имеет место очевидное равенство $\hat{A} = \hat{Q}\hat{R}$. Во-вторых, из соотношений

$$\hat{R}e_i = \hat{Q}^{-1}\hat{A}e_i = \hat{Q}^{-1}\sqrt{\lambda_i}e_i' = \sqrt{\lambda_i}e_i$$

(поскольку $e_i' = \hat{Q}e_i \quad \forall i = [1, n]$) следует, что базисные элементы e_i суть собственные векторы оператора $\hat{R}$, отвечающие положительным собственным значениям $\sqrt{\lambda_i}$. Но тогда матрица $\|\hat{R}\|_e$ в базисе $\{ e_1, e_2, \ldots, e_n \}$ диагональная и потому симметрическая. Значит, в силу леммы 10.7.1, оператор $\hat{R}$ самосопряженный.

4°. Покажем, наконец, единственность построенного разложения.

Во введенных обозначениях справедливо равенство $\hat{A}^+\hat{A} = \hat{R}^2$, поскольку из $\hat{A} = \hat{Q}\hat{R}$ и $\hat{A}^+ = \hat{R}^+\hat{Q}^+$ следует, что

$$\hat{A}^+\hat{A} = \hat{R}^+\hat{Q}^+\hat{Q}\hat{R} = \hat{R}^+\hat{Q}^{-1}\hat{Q}\hat{R} = \hat{R}^+\hat{R}\,,$$

то в силу самосопряженности $\hat{R} \qquad \hat{R}^+\hat{R} = \hat{R}^2\,.$

Предположим, что существуют два различных самосопряженных оператора $\hat{R}_1$ и $\hat{R}_2$ с положительными собственными значениями, такие, что

$$\hat{A}^+\hat{A} = \hat{R}_1^2 \quad \text{и} \quad \hat{A}^+\hat{A} = \hat{R}_2^2 \qquad \Longrightarrow \qquad \hat{R}_1^2 - \hat{R}_2^2 = \hat{O}.$$

Заметим, что $\hat{R}_1$ и $\hat{R}_2$ по построению (см. 2°) имеют общую систему собственных векторов, а потому они в силу теоремы 10.7.2 коммутируют. Но тогда, согласно определениям § 8.2, справедливы равенства

$$\hat{O} = \hat{R}_1^2 - \hat{R}_2^2 = \hat{R}_1^2 - \hat{R}_1\hat{R}_2 + \hat{R}_2\hat{R}_1 - \hat{R}_2^2 = \left(\hat{R}_1 - \hat{R}_2\right)\left(\hat{R}_1 + \hat{R}_2\right).$$

Из невырожденности и линейности $\hat{R}_1$ и $\hat{R}_2$, в силу теоремы 8.6.8, оператор $\hat{R}_1 + \hat{R}_2$ также невырожденный и поэтому из равенства $\hat{R}_1^2 - \hat{R}_2^2$ следует $\hat{R}_1 - \hat{R}_2 = \hat{O}$.

Таким образом, $\hat{R}$ — самосопряженный оператор, определяемый по $\hat{A}$ однозначно. При этом $\hat{Q} = \hat{A}\hat{R}^{-1}$ и, значит, также определяется однозначно по $\hat{A}$.

Теорема доказана.

Замечание 10.8.1. 1°. Теорема о полярном разложении является обобщением теоремы 5.5.2 о возможности представления аффинного преобразования плоскости в виде произведения двух операторов, первый из которых ортогональный, а второй – сжатие по двум взаимно перпендикулярным направлениям, матрица которого диагональная.

2°. В случае вырожденного оператора $\hat{A}$ разложение, аналогичное указанному в теореме 10.8.2, с неотрицательными собственными значениями самосопряженного оператора существует, но не единственно.

Задача 10.8.1 *В некотором ортонормированном базисе E^2 линейный оператор $\hat{A}$ имеет матрицу* $\|\hat{A}\|_{e^0} = \begin{Vmatrix} \sqrt{2} & -1 \\ 0 & \sqrt{2} \end{Vmatrix}$. *Найти его полярное разложение.*

Решение. 1°. Выполним искомое разложение по схеме, использованной в доказательстве теоремы 10.8.2.

Матрица оператора $\hat{A}^{+}\hat{A}$ в исходном (стандартном) ортонормированном базисе $\{\, e_1^o\, ;\ \ e_2^o \,\}$ имеет вид

$$\|\hat{A}^{+}\hat{A}\|_{e^o} = \|\hat{A}^{+}\|_{e^o}\|\hat{A}\|_{e^o} = \|\hat{A}\|^{\mathrm{T}}_{e^o}\|\hat{A}\|_{e^o} =$$

$$= \left\| \begin{matrix} \sqrt{2} & 0 \\ -1 & \sqrt{2} \end{matrix} \right\| \left\| \begin{matrix} \sqrt{2} & -1 \\ 0 & \sqrt{2} \end{matrix} \right\| = \left\| \begin{matrix} 2 & -\sqrt{2} \\ -\sqrt{2} & 3 \end{matrix} \right\| .$$

Собственные значения и собственные векторы оператора с этой матрицей равны соответственно

$$\lambda_1 = 1\,,\ \lambda_2 = 4\,, \quad \|f_1\|_{e^o} = \left\| \begin{matrix} \sqrt{2} \\ 1 \end{matrix} \right\| ,\ \|f_2\|_{e^o} = \left\| \begin{matrix} -1 \\ \sqrt{2} \end{matrix} \right\| .$$

поэтому (сохраняя обозначения, использованные в доказательстве теоремы 10.8.4) получим координатные представления в $\{\, e_1^o\, ;\ e_2^o \,\}$ для элементов, образующих ортонормированные базисы $\{\, e_1\, ;\ e_2 \,\}$ и $\{\, e_1'\, ;\ e_2' \,\}$ вида

$$e_1 = \frac{f_1}{|f_1|} = \frac{1}{\sqrt{3}}\left\| \begin{matrix} \sqrt{2} \\ 1 \end{matrix} \right\| , \qquad e_2 = \frac{f_2}{|f_2|} = \frac{1}{\sqrt{3}}\left\| \begin{matrix} -1 \\ \sqrt{2} \end{matrix} \right\| ,$$

$$e_1' = \frac{1}{\sqrt{\lambda_1}}\hat{A}e_1 = \frac{1}{\sqrt{3}}\left\| \begin{matrix} \sqrt{2} & -1 \\ 0 & \sqrt{2} \end{matrix} \right\| \left\| \begin{matrix} \sqrt{2} \\ 1 \end{matrix} \right\| = \frac{1}{\sqrt{3}}\left\| \begin{matrix} 1 \\ \sqrt{2} \end{matrix} \right\| ,$$

$$e_2' = \frac{1}{\sqrt{\lambda_2}}\hat{A}e_2 = \frac{1}{2\sqrt{3}}\left\| \begin{matrix} \sqrt{2} & -1 \\ 0 & \sqrt{2} \end{matrix} \right\| \left\| \begin{matrix} -1 \\ \sqrt{2} \end{matrix} \right\| = \frac{1}{\sqrt{3}}\left\| \begin{matrix} -\sqrt{2} \\ 1 \end{matrix} \right\| .$$

2°. Обозначая матрицы перехода

$$\|G\| = \frac{1}{\sqrt{3}}\left\| \begin{matrix} \sqrt{2} & -1 \\ 1 & \sqrt{2} \end{matrix} \right\| \quad \text{и} \quad \|F\| = \frac{1}{\sqrt{3}}\left\| \begin{matrix} 1 & -\sqrt{2} \\ \sqrt{2} & 1 \end{matrix} \right\|$$

соответственно от исходного базиса $\{\, e_1^o\, ;\ e_2^o \,\}$ к базисам $\{\, e_1\, ;\ e_2 \,\}$ и $\{\, e_1'\, ;\ e_2' \,\}$ и рассуждая так же, как при решении задачи 7.5.2, получаем для искомого ортогонального оператора $\hat{Q}$ координатное представление $\left\|\hat{Q}\right\|_{e^o} = \|F\|\|G\|^{-1}$.

Действительно, в рассматриваемом случае преобразование

$$\{\,e_1\,;\,e_2\,\}\xrightarrow{\hat{Q}}\{\,e'_1\,;\,e'_2\,\}$$

может быть представлено как произведение (последовательное выполнение) преобразований

$$\{\,e_1\,;\,e_2\,\}\xrightarrow{\hat{G}^{-1}}\{\,e^o_1\,;\,e^o_2\,\}\quad\text{и}\quad\{\,e^o_1\,;\,e^o_2\,\}\xrightarrow{\hat{F}}\{\,e'_1\,;\,e'_2\,\}.$$

Следовательно, $\left\|\hat{Q}\right\|_{e^o}=\left\|\hat{F}\right\|_{e^o}\left\|\hat{G}^{-1}\right\|_{e^o}$. Откуда, в силу определений (8.3.1) и (7.4.2), а также равенства $\left\|\hat{G}^{-1}\right\|_{e^o}=\left\|\hat{G}\right\|^{-1}_{e^o}$ получаем, что $\left\|\hat{Q}\right\|_{e^o}=\|F\|\|G\|^{-1}$.

Учитывая, что матрица $\|\hat{G}\|$ ортогональная (как матрица перехода, связывающая два ортонормированных базиса), находим матрицу

$$\left\|\hat{Q}\right\|_{e^o}=\|F\|\|G\|^{-1}=\|F\|\|G\|^{\mathrm{T}}=$$

$$=\frac{1}{\sqrt{3}}\left\|\begin{matrix}1&-\sqrt{2}\\\sqrt{2}&1\end{matrix}\right\|\cdot\frac{1}{\sqrt{3}}\left\|\begin{matrix}\sqrt{2}&1\\-1&\sqrt{2}\end{matrix}\right\|=\frac{1}{3}\left\|\begin{matrix}2\sqrt{2}&-1\\1&2\sqrt{2}\end{matrix}\right\|,$$

которая в исходном ортонормированном базисе также ортогональная.

3°. Поскольку $\hat{R}=\hat{Q}^{-1}\hat{A}$, то

$$\left\|\hat{R}\right\|_{e^o}=\left\|\hat{Q}^{-1}\right\|_{e^o}\left\|\hat{A}\right\|_{e^o}=\left\|\hat{Q}\right\|^{-1}_{e^o}\left\|\hat{A}\right\|_{e^o}=\left\|\hat{Q}\right\|^{\mathrm{T}}_{e^o}\left\|\hat{A}\right\|_{e^o}=$$

$$=\frac{1}{3}\left\|\begin{matrix}2\sqrt{2}&1\\-1&2\sqrt{2}\end{matrix}\right\|\left\|\begin{matrix}\sqrt{2}&-1\\0&\sqrt{2}\end{matrix}\right\|=\frac{1}{3}\left\|\begin{matrix}4&-\sqrt{2}\\-\sqrt{2}&5\end{matrix}\right\|,$$

и, следовательно, искомое полярное разложение имеет вид

$$\left\|\hat{A}\right\|_{e^o}=\left\|\hat{Q}\right\|_{e^o}\left\|\hat{R}\right\|_{e^o}=\left\|\begin{matrix}\frac{2\sqrt{2}}{3}&-\frac{1}{3}\\\frac{1}{3}&\frac{2\sqrt{2}}{3}\end{matrix}\right\|\left\|\begin{matrix}\frac{4}{3}&-\frac{\sqrt{2}}{3}\\-\frac{\sqrt{2}}{3}&\frac{5}{3}\end{matrix}\right\|.$$

Решение получено.

11. Унитарное пространство

11.1. Определение унитарного пространства

Определение 11.1.1 Пусть в комплексном линейном пространстве U каждой упорядоченной паре элементов a и b поставлено в соответствие комплексное[18] число $\langle a|b\rangle$, называемое их скалярным произведением, так, что выполнены аксиомы:

1°. $\langle a|b\rangle = \overline{\langle b|a\rangle}$;

2°. $\langle \lambda a|b\rangle = \overline{\lambda}\,\langle a|b\rangle$;

3°. $\langle a_1 + a_2|b\rangle = \langle a_1|b\rangle + \langle a_2|b\rangle$;

4°. $\langle a|a\rangle$ — вещественное неотрицательное число, причем
$$\langle a|a\rangle = 0 \quad \Longleftrightarrow \quad a = o \ ,$$

тогда говорят, что задано *унитарное пространство.*

Для обозначения скалярного произведения в унитарном пространстве используются не круглые скобки, принятые в евклидовом пространстве, а скобки типа «брэкет»: $\langle\ ,\ \rangle$.

`Замечание 11.1.1.` Вид аксиомы 1° позволяет избежать проблемы, которая возникает в случае использования евклидовского правила коммутативности (т.е. пункта 1

[18] Определение и основные свойства комплексных чисел приводятся в приложении 3.

определения 10.1.1) скалярного произведения для комплексных линейных пространств.

Действительно, если принять, что $\langle a|b\rangle = \langle b|a\rangle$, то $\langle a|\lambda b\rangle = \overline{\lambda}\,\langle a|b\rangle$, и очевидно, что при некотором ненулевом a и $\lambda = i$:

$$\langle ia|ia\rangle = \overline{i}\,\overline{i}\,\langle a|a\rangle = (-i)\,(-i)\,\langle a|a\rangle = -\,\langle a|a\rangle\ ,$$

но тогда число либо $\langle ia|ia\rangle$, либо $\langle a|a\rangle$ не положительно, что противоречит аксиоме 4°.

В случае же равенства $\langle a|b\rangle = \overline{\langle b|a\rangle}$ вынос λ из второго сомножителя скалярного произведения выполняется иначе:

$$\langle a|\lambda b\rangle = \overline{\langle\lambda b|a\rangle} = \overline{\overline{\lambda}\,\langle b|a\rangle} = \overline{\overline{\lambda}}\,\overline{\langle b|a\rangle} = \lambda\,\langle a|b\rangle\,,$$

поскольку $\overline{\overline{\lambda}} = \lambda$. Это в рассматриваемом примере приводит к равенству $\langle ia|ia\rangle = \overline{i}\,i\,\langle a|a\rangle = \langle a|a\rangle$, которое согласуется с аксиомой 4°.

Пример 11.1.1

1°. Пространство n-мерных столбцов

$$a = \|\ \xi_1,\, \xi_2,\, \ldots,\, \xi_n\ \|^{\mathrm{T}} \qquad b = \|\ \eta_1,\, \eta_2,\, \ldots,\, \eta_n\ \|^{\mathrm{T}}\,,$$

где $\xi_j,\, \eta_j \quad \forall j = [1, n]$ — комплексные числа, со скалярным произведением, определяемым по формуле

$$\langle a|b\rangle = \sum_{j=1}^{n} \overline{\xi}_j \eta_j,$$

является унитарным.

2°. Унитарным будет пространство непрерывных на $[\alpha, \beta]$ комплекснозначных функций вещественного аргумента со скалярным произведением:

$$\langle a|b\rangle = \int\limits_{\alpha}^{\beta} \overline{a}(\tau) b(\tau)\, d\tau\,.$$

В унитарных пространствах, как правило, существуют аналоги определений и теорем, справедливых для евклидова пространства. Например, *неравенство Коши – Буняковского* будет иметь следующий вид $\langle a|a\rangle\,\langle b|b\rangle \geq \langle a|b\rangle\,\langle b|a\rangle$. Действительно,

$$\langle a|a\rangle\,\langle b|b\rangle \geq |\langle a|b\rangle|^2 = \langle a|b\rangle\,\overline{\langle a|b\rangle} = \langle a|b\rangle\,\langle b|a\rangle\ .$$

В конечномерном унитарном пространстве U^n произвольный базис $\{\, g_1,\, g_2,\, \ldots,\, g_n \,\}$ при необходимости может быть ортогонализирован по схеме Грама – Шмидта.

Выражение для скалярного произведения в координатах аналогично соответствующей формуле в евклидовом пространстве:

$$\langle x|y\rangle = \overline{\|x\|}_g^{\mathrm{T}}\,\|\Gamma\|_g \|y\|_g =$$

$$= \left\|\, \overline{\xi}_1\, \overline{\xi}_2\, \ldots \overline{\xi}_n \,\right\| \left\| \begin{matrix} \langle g_1|g_1\rangle & \langle g_1|g_2\rangle & \ldots & \langle g_1|g_n\rangle \\ \langle g_2|g_1\rangle & \langle g_2|g_2\rangle & \ldots & \langle g_2|g_n\rangle \\ \ldots & \ldots & \ldots & \ldots \\ \langle g_n|g_1\rangle & \langle g_n|g_2\rangle & \ldots & \langle g_n|g_n\rangle \end{matrix} \right\| \left\| \begin{matrix} \eta_1 \\ \eta_2 \\ \ldots \\ \eta_n \end{matrix} \right\| ,$$

где $\|x\|_g$ и $\|y\|_g$ — координатные представления (столбцы) элементов x и y в базисе $\{\, g_1,\, g_2,\, \ldots,\, g_n \,\}$ и где $\|\Gamma\|_g$ — базисная матрица Грама в унитарном пространстве U^n. Заметим, что поскольку $\langle g_i|g_j\rangle = \overline{\langle g_j|g_i\rangle}$, то имеет место равенство $\|\Gamma\|_g^{\mathrm{T}} = \overline{\|\Gamma\|}_g$.

Определение 11.1.2	Матрица $\|R\|$, удовлетворяющая соотношению $\|R\|^{\mathrm{T}} = \overline{\|R\|}$, называется *эрмитовой*. Матрица $\|Q\|$, удовлетворяющая соотношениям $\|Q\|^{\mathrm{T}}\overline{\|Q\|} = \|E\|$ и $\overline{\|Q\|}\|Q\|^{\mathrm{T}} = \|E\|$, называется *унитарной*.

Определитель унитарной матрицы есть комплексное число, модуль которого равен единице. Действительно, согласно определению 11.1.2,

$$\det\left(\|Q\|^{\mathrm{T}}\overline{\|Q\|}\right) = \det\|Q\|^{\mathrm{T}} \det\overline{\|Q\|} = \det\|Q\|^{\mathrm{T}}\,\overline{\det\|Q\|} =$$

$$= \left|\det\|Q\|\right|^2 = \det\|E\| = 1\,.$$

11.2. Линейные операторы в унитарном пространстве

Для унитарного пространства имеются определения, аналогичные введенным для линейных операторов в главе 10.

В данном параграфе будут рассмотрены лишь специфические особенности линейных операторов, действующих в унитарном пространстве.

Определение 11.2.1 Линейный оператор $\hat{Q}$, действующий в унитарном пространстве U, называется *унитарным* (или *изометрическим*), если $\forall a, b \in U$ имеет место равенство $\left\langle \hat{Q}a|\hat{Q}b \right\rangle = \langle a|b \rangle$.

Замечание 11.2.1. Унитарный линейный оператор, действующий в конечномерном унитарном пространстве , в ортонормированном базисе имеет унитарную матрицу.

Определение 11.2.2 Линейный оператор $\hat{A}^+$, действующий в унитарном пространстве U, называется эрмитово сопряженным линейному оператору $\hat{A}$, если $\forall a, b \in U$ имеет место равенство $\left\langle \hat{A}a|\hat{b} \right\rangle = \left\langle a|\hat{A}^+ b \right\rangle$.

Теорема 11.2.1 **Для линейных операторов $\hat{A}$ и $\hat{B}$, действующих в унитарном пространстве U, справедливы соотношения: $\left(\hat{A}\hat{B}\right)^+ = \hat{B}^+\hat{A}^+$ и $\left(\lambda\hat{A}\right)^+ = \overline{\lambda}\hat{A}^+$.**

Доказательство.

Докажем первое соотношение. $\forall a, b \in U$ имеет место равенство

$$\left\langle \left(\hat{A}\hat{B}\right) a|\hat{b} \right\rangle = \left\langle \hat{B}a|\hat{A}^+ b \right\rangle = \left\langle a|\hat{B}^+\hat{A}^+ b \right\rangle .$$

Откуда по определению 11.2.2 имеем, что $\left(\hat{A}\hat{B}\right)^+ = \hat{B}^+\hat{A}^+$.

Аналогично, $\forall a, b \in U$

$$\left\langle \left(\lambda\hat{A}\right) a|\hat{b} \right\rangle = \overline{\lambda} \left\langle \hat{A}a|b \right\rangle = \overline{\lambda} \left\langle a|\hat{A}^+ b \right\rangle = \left\langle a|\overline{\lambda}\hat{A}^+ b \right\rangle .$$

для любого комплексного числа λ.

Теорема доказана.

Для эрмитово сопряженных операторов, действующих в конечномерном пространстве U^n, имеет место

Теорема 11.2.2 **Матрица оператора $\hat{A}^+$, эрмитово сопряженного оператору $\hat{A}$ в U^n, в базисе $\{ g_1, g_2, \ldots, g_n \}$ определяется соотношением** $\left\| \hat{A}^+ \right\|_g = \|\Gamma\|_g^{-1} \overline{\left\| \hat{A} \right\|_g}^{\mathrm{T}} \|\Gamma\|_g$**.**

Доказательство этой теоремы аналогично выводу формулы (10.6.1) для евклидова пространства.

11.3. Эрмитовы операторы

Определение 11.3.1 Линейный оператор $\hat{R}$, действующий в унитарном пространстве U, называется *эрмитово самосопряженным* (или просто *эрмитовым*), если $\hat{R}^+ = \hat{R}$.

Эрмитов оператор, действующий в унитарном пространстве, обладает свойствами, аналогичными свойствам самосопряженного оператора в евклидовом пространстве. В частности:

1°. Собственные значения эрмитова оператора вещественны.

2°. Собственные векторы, отвечающие различным собственным значениям эрмитова оператора, ортогональны.

3°. Для каждого эрмитова оператора в U^n существует ортонормированный базис, состоящий из его собственных векторов.

4°. В любом ортонормированном базисе унитарного пространства U^n эрмитов оператор имеет эрмитову матрицу.

Определение 11.3.2 Собственное значение λ линейного оператора $\hat{A}$ называется *вырожденным*, если отвечающее ему собственное подпространство имеет размерность больше единицы.

Приведем формулировки и обоснование наиболее важных свойств эрмитовых операторов.

Теорема 11.3.1 **Два эрмитовых оператора $\hat{R}$ и $\hat{S}$ имеют общую систему собственных векторов в U^n тогда и только тогда, когда эти операторы коммутируют, то есть когда $\hat{R}\hat{S} = \hat{S}\hat{R}$.**

Доказательство.

Докажем необходимость.

Пусть $\hat{R}a = \lambda a$ и $\hat{S}a = \mu a$, тогда

$$\hat{S}\hat{R}a = \hat{S}\lambda a = \lambda\hat{S}a = \lambda\mu a, \qquad \hat{R}\hat{S}a = \hat{R}\mu a = \mu\hat{R}a = \mu\lambda a$$

и, вычитая почленно, получим, что $\left(\hat{R}\hat{S} - \hat{S}\hat{R}\right) a = o$.

Поскольку элемент a — произвольный собственный вектор, то данное соотношение верно и для всей совокупности собственных векторов, а значит, и для любого элемента в U^n, так как из собственных векторов можно образовать базис. Поэтому $\hat{R}\hat{S} - \hat{S}\hat{R} = \hat{O}$.

Докажем достаточность.

Пусть самосопряженные операторы $\hat{R}$ и $\hat{S}$ коммутируют и пусть, кроме того, $\hat{R}a = \lambda a$.

Рассмотрим здесь лишь случай, когда все собственные значения оператора $\hat{R}$ различны.

Покажем, что элемент евклидова пространства $b = \hat{S}a$ является собственным вектором оператора $\hat{R}$. Действительно, в силу $\hat{R}\hat{S} = \hat{S}\hat{R}$ имеем

$$\hat{R}b = \hat{R}\hat{S}a = \hat{S}\hat{R}a = \hat{S}\lambda a = \lambda\hat{S}a = \lambda b\,.$$

С другой стороны, если все собственные значения кратности единица, то λ — это собственное значение, которое отвечает a и b одновременно. Поэтому $b = \kappa a$, и из $b = \hat{S}a$ следует $\hat{S}a = \kappa a$. Значит, a — собственный вектор оператора $\hat{S}$.

Теорема доказана.

Теорема 11.3.2 (о вырождении) **Если эрмитов оператор $\hat{A}$ коммутирует с каждым из двух некоммутирующих между собой эрмитовых операторов $\hat{B}$ и $\hat{C}$, то все собственные значения оператора $\hat{A}$ вырожденные.**

Доказательство.

Пусть Ω — линейная оболочка элемента f — является одномерным собственным подпространством оператора $\hat{A}$, отвечающим его собственному значению λ кратности единица. То есть предположим, что $\dim\Omega = 1$.
Из коммутируемости операторов $\hat{A}$ и $\hat{B}$ (по теореме 11.3.1) имеем, что $\hat{B}f = \mu f$, а из коммутируемости $\hat{A}$ и $\hat{C}$ следует, что $\hat{C}f = \kappa f$. Но тогда в силу $\hat{A}f = \lambda f$ справедливы равенства

$$\hat{A}\hat{B}f = \hat{B}\hat{A}f = \lambda\mu f \qquad \Longrightarrow \qquad \hat{C}\hat{B}\hat{A}f = \hat{C}\hat{B}\big(\lambda f\big) = \lambda\mu\kappa f$$

$$\hat{A}\hat{C}f = \hat{C}\hat{A}f = \lambda\kappa f \qquad \Longrightarrow \qquad \hat{B}\hat{C}\hat{A}f = \hat{B}\hat{C}\big(\lambda f\big) = \lambda\mu\kappa f$$

Из этих равенств получаем, что

$$\left(\hat{C}\hat{B} - \hat{B}\hat{C}\right)\lambda f = o \qquad \forall f \in \Omega\,,$$

то есть $\hat{C}\hat{B} - \hat{B}\hat{C} = \hat{O}$, и, значит, операторы $\hat{B}$ и $\hat{C}$ коммутируют.

Но последнее утверждение противоречит условию теоремы, и, следовательно, необходимо допустить существование более чем одного линейно независимого элемента в Ω.

Теорема доказана.

Таблица 11.3.1a

Евклидово пространство	**Унитарное пространство**
Правило выноса константы из первого сомножителя в скалярном произведении: $(\lambda a, b) = \lambda(a, b)$	Правило выноса константы из первого сомножителя в скалярном произведении: $\langle\lambda a\vert b\rangle = \overline{\lambda}\,\langle a\vert b\rangle$
Ортогональный оператор $\hat{Q}$: $(\hat{Q}a, \hat{Q}b) = (a, b) \quad \forall a, b \in E$	Унитарный оператор $\hat{Q}$: $\left\langle \hat{Q}a\vert\hat{Q}b\right\rangle = \langle a\vert b\rangle \quad \forall a, b \in U$
Ортогональная матрица $\Vert A\Vert$: $\Vert A\Vert^{\mathrm{T}}\Vert A\Vert = \Vert E\Vert$	Унитарная матрица $\Vert A\Vert$: $\overline{\Vert A\Vert}^{\mathrm{T}}\Vert A\Vert = \Vert E\Vert$

Таблица 11.3.1b

Евклидово пространство	**Унитарное пространство**
В ортонормированном базисе в E^n ортогональный оператор имеет ортогональную матрицу	В ортонормированном базисе в U^n унитарный оператор имеет унитарную матрицу
Сопряженный оператор $\hat{A}^+$: $$(\hat{A}a, b) = (a, \hat{A}^+ b) \quad \forall a, b \in E$$	Эрмитово сопряженный оператор $\hat{A}^+$: $$\left\langle \hat{A}a \middle\vert b \right\rangle = \left\langle a \middle\vert \hat{A}^+ b \right\rangle \quad \forall a, b \in U$$
В E^n сопряженный оператор имеет матрицу $$\left\Vert \hat{A}^+ \right\Vert_g = \Vert\Gamma\Vert_g^{-1} \left\Vert \hat{A} \right\Vert_g^{\mathrm{T}} \Vert\Gamma\Vert_g$$	В U^n эрмитово сопряженный оператор имеет матрицу $$\left\Vert \hat{A}^+ \right\Vert_g = \Vert\Gamma\Vert_g^{-1} \overline{\left\Vert \hat{A} \right\Vert_g}^{\mathrm{T}} \Vert\Gamma\Vert_g$$
Самосопряженный оператор $\hat{R}$: $$(\hat{R}a, b) = (a, \hat{R}b) \quad \forall a, b \in E$$	Эрмитово самосопряженный оператор $\hat{R}$: $$\left\langle \hat{R}a \middle\vert b \right\rangle = \left\langle a \middle\vert \hat{R}b \right\rangle \quad \forall a, b \in U$$
Симметрическая матрица $\Vert A\Vert$: $$\Vert A\Vert^{\mathrm{T}} = \Vert A\Vert$$	Эрмитова матрица $\Vert A\Vert$: $$\overline{\Vert A\Vert}^{\mathrm{T}} = \Vert A\Vert$$
В ортонормированном базисе в E^n самосопряженный оператор имеет симметрическую матрицу	В ортонормированном базисе в U^n эрмитов оператор имеет эрмитову матрицу
Из собственных векторов самосопряженного оператора в E^n можно образовать ортонормированный базис	Из собственных векторов эрмитова оператора в U^n можно образовать ортонормированный базис

В таблицах 11.3.1a и 11.3.1b для сравнения приведены сопоставимые понятия и свойства для евклидова и унитарного пространств.

11.4. Эрмитовы формы. Среднее значение и дисперсия эрмитова оператора

Как и в любом линейном пространстве, в унитарном пространстве можно ввести билинейные и квадратичные формы (функционалы). Например, в унитарном пространстве непрерывных комплекснозначных функций $\varphi(\tau)$ вещественного аргумента квадратичным функционалом является интеграл

$$\Phi\big(\varphi(\tau)\big) = \iint\limits_{\Omega} \overline{\varphi(\omega)} K(\omega, \tau) \varphi(\tau)\, d\omega d\tau\,, \qquad \Omega \subseteq E^2.$$

Определение 11.4.1 Квадратичный функционал $\Phi(x) = \left\langle x|\hat{R}x \right\rangle$, где $x \in U$, а линейный оператор $\hat{R}$ — эрмитов, называется *эрмитовым функционалом* (или *эрмитовой формой*) в унитарном пространстве U.

Определение 11.4.2 Число $\underset{-a}{\hat{R}} = \left\langle a|\hat{R}a \right\rangle$ называется *средним значением эрмитова оператора* $\hat{R}$ *по* a — нормированному элементу из унитарного пространства U.

Замечания 11.4.1

1°. Если a — нормированный (то есть для которого верно $|a| = \sqrt{\langle a|a\rangle} = 1$) собственный вектор эрмитова оператора $\hat{R}$ с соответствующим собственным значением λ, то $\underset{-a}{\hat{R}} = \lambda$, поскольку в этом случае

$$\underset{-a}{\hat{R}} = \left\langle a|\hat{R}a \right\rangle = \langle a|\lambda a\rangle = \lambda \langle a|a\rangle = \lambda\,.$$

2°. Среднее значение эрмитова оператора вещественно. По определению эрмитовости имеем $\hat{A}^{+} = \hat{A}$, тогда

$$\underset{-a}{\hat{R}} = \left\langle a|\hat{R}a \right\rangle = \left\langle a|\hat{R}^{+}a \right\rangle = \left\langle \hat{R}a|a \right\rangle = \overline{\left\langle a|\hat{R}a \right\rangle}.$$

Но если некоторое число равно своему комплексному сопряжению, то оно вещественно.

3°. Если оператор умножения на константу κ определить как $\hat{\kappa} = \kappa \hat{E}$, где $\hat{E}$ — единичный оператор, то будет верно равенство $\underline{\hat{R} - \underset{-a}{\hat{R}}}_{\,a} = 0$. Действительно,

$$\underline{\hat{R} - \underset{-a}{\hat{R}}}_{\,a} = \left\langle a \middle| \left(\hat{R} - \underset{-a}{\hat{R}}\, \hat{E} \right) a \right\rangle = \left\langle a \middle| \hat{R} a \right\rangle - \left\langle a \middle| \underset{-a}{\hat{R}}\, a \right\rangle =$$

$$= \underset{-a}{\hat{R}} - \underset{-a}{\hat{R}} \langle a | a \rangle = \underset{-a}{\hat{R}} - \underset{-a}{\hat{R}} = 0\,.$$

Определение 11.4.3 Число $\underset{=a}{\hat{R}} = \underline{\left(\hat{R} - \underset{-a}{\hat{R}} \right)^2}_{\,a}$ называется *дисперсией* эрмитова оператора $\hat{R}$ по a — нормированному элементу унитарного пространства.

Важные свойства дисперсии описывают следующие теоремы.

Теорема 11.4.1 **Дисперсия эрмитова оператора $\hat{R}$, действующего в унитарном пространстве, есть** *вещественное неотрицательное число*, **для которого справедливо равенство** $\underset{=a}{\hat{R}} = \underline{\hat{R}^2}_{\,a} - \left(\underset{-a}{\hat{R}} \right)^2$ **.**

Доказательство.

Покажем вначале, что число $\underset{=a}{\hat{R}}$ вещественное и неотрицательное. Оператор $\hat{R} - \underset{-a}{\hat{R}}$, очевидно, эрмитов, поскольку эрмитовыми являются операторы $\hat{R}$ (по условию теоремы) и $\underset{-a}{\hat{R}}$ (как оператор умножения на константу). Тогда

$$\underset{=a}{\hat{R}} = \left\langle a \middle| \left(\hat{R} - \underset{-a}{\hat{R}} \right)^2 a \right\rangle = \left\langle a \middle| \left(\hat{R} - \underset{-a}{\hat{R}} \right) \left(\hat{R} - \underset{-a}{\hat{R}} \right) a \right\rangle =$$

$$= \left\langle \left(\hat{R} - \underset{-a}{\hat{R}} \right)^+ a \middle| \left(\hat{R} - \underset{-a}{\hat{R}} \right) a \right\rangle =$$

$$= \left\langle \left(\hat{R} - \underset{-a}{\hat{R}} \right) a \middle| \left(\hat{R} - \underset{-a}{\hat{R}} \right) a \right\rangle \geq 0\,.$$

Наконец, исходя из определения 11.4.2, получаем

$$\underset{=a}{\hat{R}} = \underline{\left(\hat{R} - \underset{-a}{\hat{R}}\right)^2}_{\,a} = \left\langle a \middle| \left(\hat{R} - \underset{-a}{\hat{R}}\right)^2 a \right\rangle =$$

$$= \left\langle a \middle| \left(\hat{R}^2 - 2\hat{R}\underset{-a}{\hat{R}} + \left(\underset{-a}{\hat{R}}\right)^2\right) a \right\rangle =$$

$$= \left\langle a \middle| \hat{R}^2 a \right\rangle - 2 \left\langle a \middle| \hat{R} a \right\rangle \underset{-a}{\hat{R}} + \left(\underset{-a}{\hat{R}}\right)^2 \left\langle a \middle| a \right\rangle =$$

$$= \underline{\hat{R}^2}_{\,a} - 2\left(\underset{-a}{\hat{R}}\right)^2 + \left(\underset{-a}{\hat{R}}\right)^2 = \underline{\hat{R}^2}_{\,a} - \left(\underset{-a}{\hat{R}}\right)^2 .$$

Теорема доказана.

Теорема 11.4.2 **Для эрмитова оператора $\hat{R}$, действующего в унитарном пространстве, дисперсия, взятая по его нормированному собственному вектору, равняется нулю.**

Доказательство.

Пусть $\hat{R}a = \lambda a$, тогда

$$\underset{=a}{\hat{R}} = \underline{\hat{R}^2}_{\,a} - \left(\underset{-a}{\hat{R}}\right)^2 = \left\langle a \middle| \hat{R}^2 a \right\rangle - \left\langle a \middle| \hat{R} a \right\rangle^2 = \left\langle a \middle| \hat{R}\hat{R} a \right\rangle - \left\langle a \middle| \lambda a \right\rangle^2 =$$

$$= \left\langle a \middle| \hat{R}\lambda a \right\rangle - \lambda^2 \left\langle a \middle| a \right\rangle^2 = \lambda \left\langle a \middle| \lambda a \right\rangle - \lambda^2 \left\langle a \middle| a \right\rangle^2 = \lambda^2 - \lambda^2 = 0 ,$$

поскольку $\left\langle a \middle| a \right\rangle = 1$.

Теорема доказана.

11.5. Соотношение неопределенностей

Для эрмитовых операторов, действующих в унитарном пространстве, справедлива

`Теорема 11.5.1 (соотношение неопределенностей)` **Для двух эрмитовых операторов $\hat{A}$ и $\hat{B}$, заданных в унитарном пространстве, имеет место соотношение**

$$\underset{=a}{\hat{A}}\ \underset{=a}{\hat{B}} \geq \frac{1}{4}\left|\ \underline{\hat{A}\hat{B}-\hat{B}\hat{A}}_{\,a}\right|^2 .$$

`Доказательство.`

1°. Рассмотрим оператор $\hat{Q} = \bigl(\hat{A} - \underset{-a}{\hat{A}}\bigr) + \tau\bigl(\hat{B} - \underset{-a}{\hat{B}}\bigr)\, i\,,$ где τ — вещественный параметр. Для $\hat{Q}$ эрмитово сопряженным будет оператор $\hat{Q}^{+} = \bigl(\hat{A} - \underset{-a}{\hat{A}}\bigr) - \tau\bigl(\hat{B} - \underset{-a}{\hat{B}}\bigr)\, i\,,$ поскольку эрмитовыми являются следующие четыре оператора: $\hat{A}$, $\underset{-a}{\hat{A}}$, $\hat{B}$, $\underset{-a}{\hat{B}}$.

Заметим также, что оператор $\hat{Q}^{+}\hat{Q}$ — эрмитов и что из пункта 4° определения 11.1.1 следует

$$\left\langle a \,\middle|\, \hat{Q}^{+}\hat{Q}a\right\rangle = \left\langle \hat{Q}a \,\middle|\, \hat{Q}a\right\rangle \geq 0\,..$$

2°. Выражая $\hat{Q}^{+}\hat{Q}$ через операторы $\hat{A}$, $\underset{-a}{\hat{A}}$, $\hat{B}$, $\underset{-a}{\hat{B}}$, получаем, что $\hat{Q}^{+}\hat{Q} =$

$$= \Bigl(\bigl(\hat{A} - \underset{-a}{\hat{A}}\bigr) + \tau\bigl(\hat{B} - \underset{-a}{\hat{B}}\bigr)\, i\Bigr)\Bigl(\bigl(\hat{A} - \underset{-a}{\hat{A}}\bigr) + \tau\bigl(\hat{B} - \underset{-a}{\hat{B}}\bigr)\Bigr)\ i =$$

$$= \bigl(\hat{A} - \underset{-a}{\hat{A}}\bigr)^2 + \tau\bigl(\hat{A}\hat{B} - \hat{B}\hat{A}\bigr)i + \tau^2\bigl(\hat{B} - \underset{-a}{\hat{B}}\bigr)^2.$$

3°. Обозначим $\hat{C} = -\bigl(\hat{A}\hat{B} - \hat{B}\hat{A}\bigr)i$ и подсчитаем среднее значение эрмитова оператора $\hat{Q}^{+}\hat{Q}$ на элементе a :

$$\underline{\hat{Q}^{+}\hat{Q}}_{\,a} = \underline{\bigl(\hat{A} - \underset{-a}{\hat{A}}\bigr)^2}_{\,a} + \tau\, \underline{\bigl(\hat{A}\hat{B} - \hat{B}\hat{A}\bigr)i}_{\,a} + \tau^2\, \underline{\bigl(\hat{B} - \underset{-a}{\hat{B}}\bigr)^2}_{\,a} =$$

$$= \underset{=a}{\hat{A}} - \tau\, \underset{-a}{\hat{C}} + \tau^2\, \underset{=a}{\hat{B}} \geq 0 \qquad \forall \tau\,.$$

Полученное значение $\underline{\hat{Q}^+\hat{Q}}_{\,a}$ есть вещественный квадратный трехчлен относительно τ, который должен быть неотрицательным при любом τ. Откуда следует, что его дискриминант не положителен: $\Big(\underset{-a}{\hat{C}}\Big)^2 - 4\,\underset{=a}{\hat{A}}\,\underset{=a}{\hat{B}} \leq 0$, или окончательно

$$\underset{=a}{\hat{A}}\,\underset{=a}{\hat{B}} \geq \frac{1}{4}\left|\,\underline{\hat{A}\hat{B} - \hat{B}\hat{A}}_{\,a}\right|^2 .$$

Теорема доказана.

Замечание 11.5.1. Теорема 11.5.1 верна как в конечномерном унитарном пространстве, так и в случае, когда унитарное пространство не имеет базиса.

Заметим также, что в курсах теоретической физики и квантовой механики неравенство указанное в формулировке этой теоремы имеет название *соотношение неопределенности Гайзенберга*.

12. Прикладные задачи линейной алгебры

В данной главе рассматриваются некоторые классы задач, имеющие важное значение в прикладных разделах математики, таких, как математическая физика, теория оптимального управления, математическая экономика, вычислительная математика и т.д., причем общим для этих задач является совместное использование в процессе их решения понятий и методов из различных разделов линейной алгебры.

12.1. Приведение в E^n квадратичных форм к диагональному виду

Задача отыскания базиса, в котором квадратичный функционал имеет диагональный или канонический вид, достаточно часто встречается в различных приложениях механики, физики, теории оптимального управления.

Приведение к диагональному виду квадратичной формы, заданной в ортонормированном базисе

Пусть в ортонормированном базисе $\{ e_1, e_2, \ldots, e_n \}$ евклидова пространства E^n задана некоторая квадратичная форма (квадратичный функционал) $\Phi(x)$. Рассмотрим задачу отыскания в E^n ортонормированного базиса $\{ e'_1, e'_2, \ldots, e'_n \}$, в котором координатное представление $\Phi(x)$ (то есть его матрица) имеет диагональный вид.

Принципиальная разрешимость подобной задачи для неортонормированного базиса следует из теоремы 9.2.1. Очевидно, что такой базис не единственный, и возникает вопрос о возможности построения в E^n ортонормированного базиса, в котором данный квадратичный функционал имеет диагональный вид.

Напомним предварительно (см. § 9.2), что квадратичный функционал в E^n может быть задан в координатной форме:

$$\Phi(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} \varphi_{ij} \xi_i \xi_j = \|x\|_g^{\mathrm{T}} \|\Phi\|_g \|x\|_g \,,$$

где столбец $\|x\|_g$ — координатное представление элемента $x = \sum\limits_{i=1}^{n} \xi_i g_i$ в базисе $\{ e_1, e_2, \ldots, e_n \}$, а $\|\Phi\|_g$ — матрица функционала $\Phi(x)$, описываемая определением 9.2.2.

Замена этого базиса при помощи матрицы перехода $\|S\|$ на базис $\{ e'_1, e'_2, \ldots, e'_n \}$ приводит к изменению матрицы квадратичного функционала по формуле $\|\Phi\|_{g'} = \|S\|^{\mathrm{T}} \|\Phi\|_g \|S\|$.

В Λ^n для построении базиса, в котором квадратичный функционал имеет диагональный вид, использовался (см. § 9.3) *метод выделения полных квадратов* (называемый иногда *методом Лагранжа*), применение которого может потребовать значительных затрат вычислительных ресурсов. В E^n для решения этой задачи более эффективным оказывается алгоритм [19], основой которого является

Теорема 12.1.1 **Для всякого квадратичного функционала, заданного в ортонормированном базисе, существует ортонормированный базис, в котором этот функционал имеет диагональный вид.**

Доказательство.

1°. Матрица квадратичного функционала изменяется по правилу $\|\Phi\|_{e'} = \|S\|^{\mathrm{T}} \|\Phi\|_e \|S\|$, где $\|S\| = \|\sigma_{ij}\|$ — матрица перехода от базиса без штрихов к базису со штрихами, то есть $e'_j = \sum\limits_{i=1}^{n} \sigma_{ij} e_i \quad \forall j = [1, n]$.

[19] Иногда метод построения ортонормированного базиса, в котором квадратичный функционал имеет диагональный вид, называют «приведением квадратичного функционала к диагональному виду при помощи ортогонального преобразования базиса».

2°. Поскольку матрица перехода от одного ортонормированного базиса к другому ортогональная (§ 10.4), то для нее справедливо равенство $\|S\|^{-1} = \|S\|^{\mathrm{T}}$. Откуда вытекает, что в данном случае $\|\Phi\|_{e'} = \|S\|^{-1}\|\Phi\|_e\|S\|$.

3°. Матрица $\|\Phi\|_e$ симметрическая, поэтому в ортонормированном базисе $\{e_1, e_2, \ldots, e_n\}$ она определяет самосопряженный оператор $\hat{\Phi}$ (лемма 10.7.1), матрица которого в базисе $\{e'_1, e'_2, \ldots, e'_n\}$ находится по формуле (теорема 8.3.2): $\|\Phi\|_{e'} = \|S\|^{-1}\|\Phi\|_e\|S\| = \|S\|^{\mathrm{T}}\|\Phi\|_e\|S\|$.

4°. Совпадение формул *изменения матриц* квадратичного функционала и самосопряженного оператора при переходе от одного *ортонормированного* базиса к другому позволяет решить нашу задачу, использовав в качестве базиса $\{e'_1, e'_2, \ldots, e'_n\}$ ортонормированный базис из собственных векторов оператора $\hat{\Phi}$.

Этот базис существует (см. теорему 10.7.1) и в нем матрица оператора $\hat{\Phi}$ (а значит, и матрица квадратичного функционала $\Phi(x)$) имеет диагональный вид, причем на главной диагонали расположены собственные значения самосопряженного оператора $\hat{\Phi}$.

Теорема доказана.

Заметим, что утверждение теоремы 12.1.1 согласуется с утверждением следствия 10.7.4.

Определение 12.1.1	Линейный самосопряженный оператор $\hat{\Phi}$ называется *присоединенным* к квадратичному функционалу $\Phi(x)$ в E^n.

При этом очевидно выполнение равенства $\Phi(x) = (x, \hat{\Phi}x) \quad \forall x \in E^n$.

Определение 12.1.2	Функционал $\rho(x) = \dfrac{(x, \hat{R}x)}{(x, x)}$, заданный в E^n для некоторого самосопряженного оператора $\hat{R}$, называется *отношением Релея*.

Используя теорему 12.1.1, можно упростить процедуру оценки экстремальных значений квадратичного функционала. В качестве примера рассмотрим задачу нахождения максимума и минимума.

Следствие 12.1.1 **В ортонормированном базисе в E^n максимальное (минимальное) значение $\rho(x)$ равно максимальному (минимальному) собственному значению оператора $\hat{R}$, и это значение достигается на соответствующем собственном векторе этого оператора.**

Доказательство.

Поскольку при переходе к ортонормированному базису, образованному из собственных векторов самосопряженного оператора $\hat{R}$ (в силу теоремы 12.1.1), справедливы соотношения

$$\rho(x) = \frac{(x, \hat{R}x)}{(x, x)} = \frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \alpha_{ij}\xi_i\xi_j}{\sum\limits_{i=1}^{n} \xi_i^2} = \frac{\sum\limits_{i=1}^{n} \lambda_i \xi_i'^2}{\sum\limits_{i=1}^{n} \xi_i'^2},$$

то, проводя рассуждения, аналогичные доказательству теоремы 9.5.1, получаем, что $\lambda_{\min} \le \rho(x) \le \lambda_{\max}$.

Следствие доказано.

Проиллюстрируем применение теоремы 12.1.1 на примере решения следующей задачи.

Задача 12.1.1 *При помощи ортогональной замены координат привести к диагональному виду в E^3 квадратичный функционал*

$$\Phi(x) = 2\xi_1\xi_2 + 2\xi_1\xi_3 - 2\xi_2\xi_3\,.$$

Решение. 1°. Пусть исходный ортонормированный базис стандартный, то есть состоит из элементов $\{e_1;\, e_2; e_3\}$ с координатными представлениями:

$$\|\, e_1\| = \|1\ 0\ 0\,\|^{\mathrm{T}}, \quad \|\, e_2\| = \|\,0\ 1\ 0\,\|^{\mathrm{T}}, \quad \|\, e_3\| = \|\,0\ 0\ 1\,\|^{\mathrm{T}}.$$

Восстановим по квадратичному функционалу $\Phi(x)$ порождающий его симметричный билинейный функционал $B(x, y)$, использовав формулу

$$B(x,y) = \frac{\Phi(x+y) - \Phi(x) - \Phi(y)}{2}$$

(см. § 9.2). Проверьте самостоятельно, что в данном случае

$$B(x,y) = \xi_1\eta_2 + \xi_1\eta_3 - \eta_2\xi_3 + \eta_1\xi_2 + \eta_1\xi_3 - \eta_2\xi_3\,,$$

где $\|\, x\|_e = \|\, \xi_1\ \xi_2\ \xi_3\,\|^{\mathrm{T}}$ и $\|\, y\|_e = \|\, \eta_1\ \eta_2\ \eta_3\,\|^{\mathrm{T}}$.

Следовательно, матрица функционала $\Phi(x)$ имеет вид

$$\|\Phi\|_e = \left\|\begin{array}{rrr} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{array}\right\|.$$

2°. Рассмотрим в ортонормированном базисе в E^3 самосопряженный оператор $\hat{\Phi}$, имеющий ту же самую матрицу $\|\hat{\Phi}\|_e = \|\Phi\|_e$.

Найдем его собственные значения, составив характеристическое уравнение (8.5.2):

$$\det\left\|\begin{array}{rrr} -\lambda & 1 & 1 \\ 1 & -\lambda & -1 \\ 1 & -1 & -\lambda \end{array}\right\| = 0 \qquad \text{или} \qquad \lambda^3 - 3\lambda + 2 = 0\,.$$

Оно имеет корни: $\lambda_1 = -2$, $\lambda_{2,3} = 1$, которые и являются искомыми собственными значениями.

Заметим, что если нас интересует только диагональный вид квадратичного функционала, то его уже можно написать, основываясь на следствии 10.7.1 и теореме 12.1.1 :

$$\Phi(x) = -2\xi_1^{\prime 2} + \xi_2^{\prime 2} + \xi_3^{\prime 2},$$

и закончить на этом решение задачи.

3°. В случае, когда требуется также построить и новый «диагональный» ортонормированный базис (или, что равносильно, найти матрицу $\|S\|$ — матрицу перехода от исходного ортонормированного базиса к «диагональному»), нам потребуются еще и собственные векторы оператора $\hat{\Phi}$.

Для этого последовательно подставим найденные собственные значения в систему (8.4.1) и найдем ее общие решения.

Пусть собственный вектор f имеет координатное представление $\|f\|_e = \| \ \varphi_1 \ \varphi_2 \ \varphi_3 \ \|^{\mathrm{T}}$. При $\lambda = -2$ система (8.4.1) такова:

$$\left\| \begin{array}{rrr} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{array} \right\| \left\| \begin{array}{c} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{array} \right\| = \left\| \begin{array}{c} 0 \\ 0 \\ 0 \end{array} \right\| .$$

Заметим, что ранг основной матрицы этой системы равен 2, поскольку третье уравнение есть разность первых двух и имеется ненулевой минор порядка 2.

Далее, действуя по схеме, описанной в § 6.8 (*метод Гаусса*), получаем для компонентов собственного вектора систему условий

$$\left\{ \begin{array}{rcrcr} 2\varphi_1 & + & \varphi_2 & = & -\varphi_3\,, \\ \varphi_1 & + & 2\varphi_2 & = & \varphi_3\,. \end{array} \right.$$

Принимая φ_3 за свободное неизвестное, получим собственный вектор f_1 с координатным представлением $\|f_1\|_e = \| -1 \ 1 \ 1 \ \|^{\mathrm{T}}$.

Кратность собственного значения $\lambda = 1$ равна 2, и в силу следствия 10.7.2 ему должны отвечать два линейно независимых (но не обязательно ортогональных) собственных вектора. Конкретно, компоненты собственного вектора должны удовлетворять следующей системе уравнений:

$$\left\| \begin{array}{rrr} -1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \end{array} \right\| \left\| \begin{array}{c} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{array} \right\| = \left\| \begin{array}{c} 0 \\ 0 \\ 0 \end{array} \right\| ,$$

из которых независимое только одно. Общее решение этой системы будет иметь вид

$$\left\| \begin{array}{c} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{array} \right\| = \alpha \left\| \begin{array}{c} 1 \\ 1 \\ 0 \end{array} \right\| + \beta \left\| \begin{array}{c} 1 \\ 0 \\ 1 \end{array} \right\| \qquad \forall \alpha, \beta \in \mathbb{R}\,.$$

Каждый столбец такого вида ортогонален f_1, но выбранные нами конкретные фундаментальные решения $\| \; 1 \; 1 \; 0 \; \|^{\mathrm{T}}$ и $\| \; 1 \; 0 \; 1 \; \|^{\mathrm{T}}$ не ортогональны друг другу.

Поэтому пару ортогональных собственных векторов, отвечающих $\lambda = 1$, сформируем из первого фундаментального решения и ортогональной ему линейной комбинации первого и второго.

Условие ортогональности столбцов $\left\| \begin{matrix} 1 \\ 1 \\ 0 \end{matrix} \right\|$ и $\left\| \begin{matrix} \alpha + \beta \\ \alpha \\ \beta \end{matrix} \right\|$, очевидно, есть $2\alpha + \beta = 0$. Откуда, например, выбрав $\alpha = 1$ и $\beta = -2$, получим f_2 и f_3 соответственно с координатными представлениями $\|f_2\|_e = \| \; 1 \; 1 \; 0 \; \|^{\mathrm{T}}$ и $\|f_3\|_e = \| \; -1 \;\; 1 \;\; -2 \; \|^{\mathrm{T}}$.

4°. Набор элементов $\{f_1;\, f_2; f_3\}$ является в E^3 ортогональным, но не нормированным базисом. Чтобы построить ортонормированный базис, выполним нормировку каждого из элементов базиса $\{f_1;\, f_2; f_3\}$. В результате получим искомый базис $\{e'_1;\, e'_2; e'_3\}$ с координатными столбцам:

$$\|e'_1\| = \left\| \begin{matrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{matrix} \right\| , \quad \|e'_2\| = \left\| \begin{matrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{matrix} \right\| , \quad \|e'_3\| = \left\| \begin{matrix} -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \end{matrix} \right\| .$$

Матрица перехода от исходного базиса $\{e_1;\, e_2; e_3\}$ к «диагональному» $\{e'_1;\, e'_2; e'_3\}$ базису имеет вид

$$\|S\| = \left\| \begin{matrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{1}{\sqrt{6}} \end{matrix} \right\| .$$

Ее столбцами являются коэффициенты координатных разложений элементов базиса $\{e'_1;\, e'_2; e'_3\}$ по базису $\{e_1;\, e_2; e_3\}$.

Матрица $\|S\|$ ортогональная, то есть для нее верно равенство $\|S\|^{-1} = \|S\|^{\mathrm{T}}$, то есть можно не только записывать формулы перехода, но также легко получать равенства, выражающие «новые» координаты через «старые».

Действительно (см. § 7.4), из стандартной формулы перехода $\|x\|_e = \|S\|\|x\|_{e'}$ следует

$$\|x\|_{e'} = \|S\|^{-1}\|x\|_e \qquad \Longrightarrow \qquad \|x\|_{e'} = \|S\|^{\mathrm{T}}\|x\|_e \, ,$$

или в координатах

$$\left\| \begin{matrix} \xi'_1 \\ \xi'_2 \\ \xi'_3 \end{matrix} \right\| = \left\| \begin{matrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \end{matrix} \right\| \left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{matrix} \right\| .$$

Решение получено.

Построение базиса, в котором два квадратичных функционала (один из которых знакоопределенный) имеют диагональный вид

Пусть в некотором базисе $\{ g_1, g_2, \ldots, g_n \}$ линейного пространства Λ^n задана пара квадратичных функционалов $\Phi(x)$ и $\Psi(x)$, первый из которых знакоопределенный (например, положительно). Рассмотрим задачу отыскания базиса, в котором функционал $\Phi(x)$ имеет канонический, а функционал $\Psi(x)$ — диагональный вид.

Отметим, что условие знаковой определенности одного из приводимых квадратичных функционалов существенно, поскольку в общем случае два различных квадратичных функционала одним линейным преобразованием к диагональному виду не приводятся.

Действительно, известно (см. доказательство теоремы 4.4.1), что квадратичный функционал $\Phi(x) = A\xi_1^2 + 2B\xi_1\xi_2 + C\xi_2^2$ в Λ^2 можно привести к диагональному виду при помощи линейного оператора, сводящегося к повороту плоскости радиусов-векторов на угол α, который удовлетворяет уравнению $(A - C)\sin 2\alpha = 2B\cos 2\alpha$.

Однако для пары квадратичных функционалов $\Phi_1(x) = \xi_1^2 - \xi_2^2$ и $\Phi_2(x) = \xi_1\xi_2$ угла α, удовлетворяющего одновременно условиям $2\sin 2\alpha = 0$ и $0 = \cos 2\alpha$, очевидно, не существует.

Опишем теперь алгоритм приведения заданных в исходном базисе $\{ g_1, g_2, \ldots, g_n \}$ в Λ^n пары квадратичных функционалов $\Phi(x)$ и $\Psi(x)$, первый из которых положительно определенный, соответственно к каноническому и диагональному виду.

1°. Поскольку квадратичный функционал $\Phi(x)$ положительно определенный, то для него в Λ^n найдется базис $\{ g'_1, g'_2, \ldots, g'_n \}$, в котором он имеет канонический вид, причем все его коэффициенты равны единице (см. теорему 9.2.1).

   Приведем этот функционал к каноническому виду каким-либо методом, например, выделив полные квадраты с нормировкой, если она потребуется. Одновременно *тем же самым методом* преобразуем и квадратичный функционал $\Psi(x)$.

2°. Введем в Λ^n скалярное произведение с единичной матрицей Грама, то есть по формуле $(x, y) = \sum\limits_{i=1}^{n} \xi'_i \eta'_i$, превратив тем самым данное линейное пространство Λ^n в евклидово E^n. Отметим, что в этом случае базис

$$\{ g'_1, g'_2, \ldots, g'_n \} = \{ e'_1, e'_2, \ldots, e'_n \},$$

   в котором $\Phi(x)$ имеет канонический вид, *ортонормированный.*

3°. Наконец, строим третий, также ортонормированный, базис $\{ e"_1, e"_2, \ldots, e"_n \}$, приводя квадратичный функционал $\Psi(x)$ к *диагональному* виду по схеме, описанной в доказательстве теоремы 12.1.1. Пусть переход от второго к третьему базису выполняется при помощи *ортогональной* матрицы перехода $\|S\|$.

   При этом переходе квадратичный функционал $\Phi(x)$ не потеряет канонического вида, поскольку из условия $\|\Phi\|_{e'}$ и ортогональности $\|S\|$ следует, что

$$\|\Phi\|_{e''} = \|S\|^{\mathrm{T}} \|\Phi\|_{e'} \|S\| = \|S\|^{-1} \|E\| \|S\| = \|S\|^{-1} \|S\| = \|E\| .$$

   Таким образом, построен базис, в котором квадратичный функционал $\Phi(x)$ имеет канонический вид, а функционал $\Psi(x)$ — диагональный.

В заключение отметим, что матрица перехода от исходного базиса к искомому ортонормированному базису равна *произведению* матрицы перехода, при котором знакоопределенный квадратичный функционал приводится к каноническому виду, и ортогональной матрицы $\|S\|$.

Задача 12.1.2 *Найти замену переменных, приводящую квадратичные функционалы* $\Phi(x) = \xi_1^2 + 2\xi_1\xi_2 + 3\xi_2^2$ *и* $\Psi(x) = 4\xi_1^2 + 16\xi_1\xi_2 + 6\xi_2^2$ *соответственно к каноническому и диагональному виду.*

Решение. 1°. Исследуем квадратичные функционалы $\Phi(x)$ и $\Psi(x)$ на знаковую определенность. Из критерия Сильвестра (теорема 9.3.2, следствие 9.3.1) и неравенств

$$\det \left\| \begin{array}{cc} 1 & 1 \\ 1 & 3 \end{array} \right\| > 0 \qquad \text{и} \qquad \det \left\| \begin{array}{cc} 4 & 8 \\ 8 & 6 \end{array} \right\| < 0$$

заключаем, что $\Phi(x)$ — положительно определенный квадратичный функционал, в то время как функционал $\Psi(x)$ не является знакоопределенным.

2°. Приведем положительно определенный квадратичный функционал к каноническому виду методом Лагранжа. Поскольку

$$\Phi(x) = \xi_1^2 + 2\xi_1\xi_2 + 3\xi_2^2 = \left(\xi_1 + \xi_2\right)^2 + 2\xi_2^2\,,$$

то, выполнив замену переменных

$$\left\{ \begin{array}{ccc} \xi_1' & = & \xi_1 + \xi_2\,, \\ \xi_2' & = & \sqrt{2}\,\xi_2 \end{array} \right. \qquad \text{или} \qquad \left\{ \begin{array}{ccr} \xi_1 & = & \xi_1' - \frac{1}{\sqrt{2}}\xi_2'\,, \\ \xi_2 & = & \frac{1}{\sqrt{2}}\xi_2'\,, \end{array} \right.$$

мы получим $\Phi(x) = \xi_1'^2 + \xi_2'^2$ и соответственно $\Psi(x) = 4\xi_1'^2 + 4\sqrt{2}\,\xi_1'\xi_2' - 3\xi_2'^2$.

3°. Введение в Λ^2 скалярного произведения с единичной матрицей Грама означает, что координаты $\{\ \xi_1';\ \xi_2'\ \}$ являются координатами евклидова пространства E^2 с базисом $\{\ e_1';\ e_2'\ \}$, где $\|\, e_1' \| = \|\, 1\ \ 0\, \|^{\mathrm{T}}$ и соответственно $\|\, e_2' \| = \|\, 0\ 1\, \|^{\mathrm{T}}$.

Матрица квадратичного функционала $\Psi(x)$ в этом базисе $\left\| \begin{array}{cc} 4 & 2\sqrt{2} \\ 2\sqrt{2} & -3 \end{array} \right\|$.

.

Она задает присоединенный самосопряженный оператор, имеющий собственные значения $\lambda_1 = 5$ и $\lambda_2 = -4$, а также ортогональные и нормированные собственные векторы f_1 и f_2 с координатными столбцами

$$\|f_1\|_{e'} = \left\| \begin{matrix} \frac{2\sqrt{2}}{3} \\ \frac{1}{3} \end{matrix} \right\|, \qquad \|f_2\|_{e'} = \left\| \begin{matrix} -\frac{1}{3} \\ \frac{2\sqrt{2}}{3} \end{matrix} \right\|,$$

которые примем за третий базис $\{\, e''_1;\ e''_2 \,\}$.

4°. Матрица перехода от ортонормированного базиса $\{\, e'_1;\ e'_2 \,\}$ к ортонормированному базису $\{\, e''_1;\ e''_2 \,\}$, в котором $\Phi(x) = {\xi''_1}^2 + {\xi''_2}^2$ и $\Psi(x) = 5{\xi''_1}^2 - 4{\xi''_2}^2$, ортогональная и имеет вид $\|S\| = \left\| \begin{matrix} \frac{2\sqrt{2}}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2\sqrt{2}}{3} \end{matrix} \right\|$.

Откуда окончательно получаем, что

$$\left\{ \begin{array}{rcrcr} \xi''_1 & = & \frac{2\sqrt{2}}{3}\xi'_1 & + & \frac{1}{3}\xi'_2\,, \\ \xi''_2 & = & -\frac{1}{3}\xi'_1 & + & \frac{2\sqrt{2}}{3}\,\xi'_2 \end{array} \right. \qquad \Longrightarrow$$

Решение получено.

$$\Longrightarrow \qquad \left\{ \begin{array}{rcrcr} \xi''_1 & = & \frac{2\sqrt{2}}{3}\xi_1 & + & \sqrt{2}\xi_2\,, \\ \xi''_2 & = & -\frac{1}{3}\xi_1 & + & \xi_2\,. \end{array} \right.$$

Если в задаче одновременного приведения пары квадратичных функционалов, один из которых знакоопределенный соответственно к каноническому и диагональному виду, требуется найти лишь эти виды (а не формулы замены переменных), то можно воспользоваться более простой схемой расчетов.

Допустим, что положительно определенный квадратичный функционал $\Phi(x)$ приведен при помощи некоторой матрицы перехода $\|W\|$ к каноническому виду, то есть $\|W\|^{\mathrm{T}}\|\Phi\|\|W\| = \|E\|$. После того же преобразования матрица квадратичного функционала $\Psi(x)$ будет иметь вид $\|\Psi^*\| = \|W\|^{\mathrm{T}}\|\Psi\|\|W\|$.

Согласно теореме 12.1.1 в ортонормированном базисе для построения диагонального вида квадратичного функционала $\Psi(x)$ достаточно найти собственные числа самосопряженного оператора, матрица которого есть $\|\Psi^*\|$.

Найдем выражение для этой матрицы, учитывающее связь между матрицами $\|\Phi\|$ и $\|W\|$.

Из равенства $\quad \|W\|^{\mathrm{T}}\|\Phi\|\|W\| = \|E\| \quad$ очевидно получается, что $\|W\| = \left(\|W\|^{\mathrm{T}}\|\Phi\|\right)^{-1}$. Тогда, используя правила обращения и транспонирования произведения матриц, перестановочность обращения и транспонирования, а также симметричность и невырожденность матрицы $\|\Phi\|$, получаем

$$\|\Psi^*\| = \|W\|^{\mathrm{T}}\|\Psi\|\|W\| = \Big(\left(\|W\|^{\mathrm{T}}\|\Phi\|\right)^{-1}\Big)^{\mathrm{T}}\|\Psi\|\|W\| =$$

$$= \Big(\left(\|W\|^{\mathrm{T}}\|\Phi\|\right)^{\mathrm{T}}\Big)^{-1}\|\Psi\|\|W\| = \Big(\|\Phi\|\|W\|\Big)^{-1}\|\Psi\|\|W\| =$$

$$= \|W\|^{-1}\Big(\|\Phi\|^{-1}\|\Psi\|\Big)\|W\|\,.$$

Это означает, что матрица $\|\Psi^*\|$ может рассматриваться как результат изменения матрицы линейного оператора $\Phi^{-1}\Psi$ при замене базиса с матрицей перехода $\|W\|$.

Поскольку собственные значения линейного оператора не зависят от выбора базиса, то решение задачи сводится к определению собственных значений оператора, имеющего матрицу $\|\Phi\|^{-1}\|\Psi\|$.

Собственные векторы и собственные значения этого оператора находятся согласно § 8.5 из системы линейных уравнений

$$\|\Phi\|^{-1}\|\Psi\|\|f\| = \lambda\|f\| \qquad \Longrightarrow \qquad \Big(\|\Psi\| - \lambda\|\Phi\|\Big)\|f\| = \|o\|\,.$$

Условие существования ненулевых $\|f\|:\quad \det\Big(\|\Psi\| - \lambda\|\Phi\|\Big) = 0$ — алгебраическое уравнение относительно λ, корни которого и являются искомыми коэффициентами диагонального представления квадратичного функционала $\Psi(x)$.

Проиллюстрируем применение данного метода для нахождения диагонального вида квадратичных форм в задаче 12.1.2. В этом случае $\|\Phi\| = \left\|\begin{matrix} 1 & 1 \\ 1 & 3 \end{matrix}\right\|$ и $\|\Psi\| = \left\|\begin{matrix} 4 & 8 \\ 8 & 6 \end{matrix}\right\|$, то есть для определения коэффициентов диагонального представления квадратичного функционала Ψ необходимо решить уравнение

$$\det\left(\left\|\begin{matrix} 4 & 8 \\ 8 & 6 \end{matrix}\right\| - \lambda\left\|\begin{matrix} 1 & 1 \\ 1 & 3 \end{matrix}\right\|\right) = 0 \qquad \text{или} \qquad \det\left\|\begin{matrix} 4-\lambda & 8-\lambda \\ 8-\lambda & 6-3\lambda \end{matrix}\right\| = 0.$$

Проверьте, что данное уравнение имеет корни $\lambda_1 = 5$ и $\lambda_2 = -4$. Значит, искомые канонический и диагональный виды будут такими: $\Phi(x) = \xi_1'^2 + \xi_2'^2$ и $\Psi(x) = 5\xi_1'^2 - 4\xi_2'^2$.

12.2. Классификация поверхностей второго порядка

Операция приведения квадратичного функционала к диагональному виду позволяет выполнять классификацию поверхностей второго порядка методом, отличающимся от изложенного в § 4.5.

Пусть в евклидовом пространстве E^3 со стандартным ортонормированным базисом $\{\,e_1,\ e_2,\ e_3\,\}$ вида

$$\|e_1\| = \|\ 1\ 0\ 0\ \|^{\mathrm{T}}\,, \quad \|e_2\| = \|\ 0\ 1\ 0\ \|^{\mathrm{T}}\,, \quad \|e_3\| = \|\ 0\ 0\ 1\ \|^{\mathrm{T}}$$

поверхность S задана уравнением второго порядка

$$\alpha_{11}\xi_1^2 + \alpha_{22}\xi_2^2 + \alpha_{33}\xi_3^2 + 2\alpha_{12}\xi_1\xi_2 + 2\alpha_{13}\xi_1\xi_3 + 2\alpha_{23}\xi_2\xi_3 +$$
$$+2\alpha_{14}\xi_1 + 2\alpha_{24}\xi_2 + 2\alpha_{34}\xi_3 + \alpha_{44} = 0\,, \quad \text{где} \qquad \sum_{k=1}^{3}\sum_{i=1}^{k}|\alpha_{ki}| > 0\,.$$

Слагаемые второго порядка в левой части данного уравнения можно рассматривать как квадратичный функционал в E^3. Приведем его к диагональному виду ортогональным преобразованием по схеме, описанной в § 12.1. Получим уравнение

$$\lambda_1\xi_1'^2 + \lambda_2\xi_2'^2 + \lambda_3\xi_3'^2 + 2\alpha_{14}'\xi_1' + 2\alpha_{24}'\xi_2' + 2\alpha_{34}'\xi_3' + \alpha_{44}' = 0\,,$$
$$\text{где} \qquad |\lambda_1| + |\lambda_2| + |\lambda_3| > 0\,.$$

для которого рассмотрим три следующих случая.

1. Центральный случай: здесь $\lambda_1\lambda_2\lambda_3 \neq 0$ или (проверьте это самостоятельно), что в силу теоремы 8.6.8 то же самое,

$$\det\left\|\begin{matrix}\alpha_{11} & \alpha_{12} & \alpha_{13}\\ \alpha_{12} & \alpha_{22} & \alpha_{23}\\ \alpha_{13} & \alpha_{23} & \alpha_{33}\end{matrix}\right\| \neq 0\,.$$

После переноса начала координат, устраняющего линейные слагаемые, получаем уравнение

$$\lambda_1 {\xi_1''}^2 + \lambda_{21} {\xi_2''}^2 + \lambda_3 {\xi_3''}^2 + \alpha_{44}'' = 0\,,$$

для которого можно выделить следующие варианты:

если $\alpha_{44}'' \neq 0$, то имеем

1) *мнимый эллипсоид* при
$\operatorname{sgn}\lambda_i = \operatorname{sgn}\alpha_{44}'' \quad \forall i = 1, 2, 3;$

2) *эллипсоид* при
$\operatorname{sgn}\lambda_i = -\operatorname{sgn}\alpha_{44}'' \quad \forall i = 1, 2, 3;$

3) *однополостный гиперболоид* при
$\operatorname{sgn}\lambda_1 = \operatorname{sgn}\lambda_2 = -\operatorname{sgn}\lambda_3 = -\operatorname{sgn}\alpha_{44}'';$

4) *двуполостный гиперболоид* при
$\operatorname{sgn}\lambda_1 = -\operatorname{sgn}\lambda_2 = -\operatorname{sgn}\lambda_3 = -\operatorname{sgn}\alpha_{44}'',$

если $\alpha_{44}'' = 0$, то имеем

5) *изолированная точка* при
$\operatorname{sgn}\lambda_1 = \operatorname{sgn}\lambda_2 = \operatorname{sgn}\lambda_3;$

6) *Конус* при
$\operatorname{sgn}\lambda_1 = \operatorname{sgn}\lambda_2 = -\operatorname{sgn}\lambda_3.$

II. Первый нецентральный случай, где $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, $\lambda_3 = 0$. После переноса начала координат приходим к уравнению

$$\lambda_1 {\xi_1''}^2 + \lambda_2 {\xi_2''}^2 + 2\alpha_{34}''\xi_3'' + \alpha_{44}'' = 0\,,$$

для которого выделяем варианты:

если $\alpha_{34}'' \neq 0$, то уравнение приводится к

$$\lambda_1 {\xi_1'''}^2 + \lambda_2 {\xi_2'''}^2 + 2\alpha_{34}'''\xi_3''' = 0,$$

и тогда имеем

7) *эллиптический параболоид* при
$\operatorname{sgn}\lambda_1 = \operatorname{sgn}\lambda_2;$

8) *гиперболический параболоид* при
$\operatorname{sgn}\lambda_1 = -\operatorname{sgn}\lambda_2;$

если $\alpha''_{34} = 0,\ \alpha''_{44} \neq 0$, то имеем

9) *мнимый эллиптический цилиндр* при
$\operatorname{sgn}\lambda_i = \operatorname{sgn}\alpha''_{44} \quad \forall i = 1, 2;$

10) *эллиптический цилиндр* при
$\operatorname{sgn}\lambda_i = -\operatorname{sgn}\alpha''_{44} \quad \forall i = 1, 2;$

11) *гиперболический цилиндр* при
$\operatorname{sgn}\lambda_1 = -\operatorname{sgn}\lambda_2;,$

если $\alpha''_{34} = 0,\ \alpha''_{44} = 0$, то получаем

12) *изолированную прямую* при
$\operatorname{sgn}\lambda_1 = \operatorname{sgn}\lambda_2;$

13) *пару пересекающихся плоскостей* при
$\operatorname{sgn}\lambda_1 = -\operatorname{sgn}\lambda_2.$

III. Второй нецентральный случай, где $\lambda_1 \neq 0,\ \lambda_2 = \lambda_3 = 0$. После переноса начала координат приходим к уравнению

$$\lambda_1 \xi_1^{''2} + 2\alpha''_{24}\xi'_2 + 2\alpha''_{34}\xi''_3 + \alpha''_{44} = 0\,,$$

для анализа которого целесообразно перейти в новую ортонормированную систему координат по формулам

$$\xi'''_1 = \xi''_1, \qquad \xi'''_2 = \frac{\alpha''_{24}\xi''_2 + \alpha''_{34}\xi''_3}{\sqrt{\alpha_{24}^{''2} + \alpha_{34}^{''2}}}, \qquad \xi'''_3 = \frac{\alpha''_{34}\xi''_2 - \alpha''_{24}\xi''_3}{\sqrt{\alpha_{24}^{''2} + \alpha_{34}^{''2}}},$$

что, очевидно, является поворотом в плоскости $O''\xi''_2\xi''_3$. В итоге получаем уравнение

$$\lambda_1 \xi_1^{'''2} + 2\xi'''_2\sqrt{\alpha_{24}^{''2} + \alpha_{34}^{''2}} + \alpha'''_{44} = 0$$

и соответствующие ему варианты:

если $\alpha''_{24} \neq 0$ или $\alpha''_{34} \neq 0$, то после переноса начала координат имеем

14) *параболический цилиндр*;

если $\alpha''_{24} = \alpha''_{34} = 0$, то имеем

15) *пару мнимых параллельных плоскостей* при
$\operatorname{sgn}\lambda_1 = \operatorname{sgn}\alpha'''_{44};$

16) *пару параллельных плоскостей* при $\operatorname{sgn}\lambda_1 = -\operatorname{sgn}\alpha'''_{44}$;

если $\alpha''_{24} = \alpha''_{34} = \alpha''_{44} = 0$, то мы имеем

17) *пару совпадающих плоскостей.*

`Замечание 12.2.1.` 1) Для классификации конкретной поверхности второго порядка необходимо привести квадратичную часть уравнения к диагональному виду и выполнить переносы начала координат.

2) Для классификации линий второго порядка на плоскости можно использовать схему, аналогичную рассмотренной для поверхностей второго порядка.

12.3. Использование координатного метода в линейных пространствах без базиса

Рассмотрим возможные методы координатного описания элементов в линейных пространствах, не имеющих базиса.

Согласно определению 7.2.2 базисом в линейном пространстве является упорядоченный набор $n \in \mathbb{N}$ линейно независимых элементов $\{ g_1,\ g_2,\ \dots,\ g_n \}$ такой, что добавление к нему любого элемента пространства делает этот набор линейно зависимым. В этом случае, по определению 7.2.3, линейное пространство называется n-мерным.

Это означает, что линейное пространство не имеет базиса, если либо в нем вообще нет линейно независимых наборов элементов (пример — линейное пространство, состоящее из одного нулевого элемента), либо для любого линейно независимого набора, состоящего из $n \in \mathbb{N}$ элементов, существует $n+1$-й элемент, добавление которого в набор сохраняет линейную независимость.

Примером линейного пространства, в котором есть линейно независимые множества элементов, но нет базиса, может служить линейное пространство функций $f(\tau)$, непрерывных на отрезке $[-1, 1]$. Действительно, $\forall k \in \mathbb{N}$ набор его элементов вида $1,\ \tau,\ \tau^2,\ \dots,\tau^k$ является линейно независимым. При этом существует непрерывная функция

τ^{k+1}, добавление которой в набор сохранит его линейную независимость.

Суть координатного метода состоит (теорема 7.2.1) в том, что в n-мерном линейном пространстве Λ^n каждый элемент $x \in \Lambda^n$ может быть представлен единственным образом как линейная комбинация элементов, образующих базис

$$x = \sum_{j=1}^{n} \xi_j g_j \,, \tag{12.3.1}$$

где упорядоченный набор чисел $\{\,\xi_1,\ \xi_2,\ \ldots\ ,\ \xi_n\,\}$ (называемых координатами элемента x в базисе $\{\,g_1,\ g_2,\ \ldots\ ,\ g_n\,\}$) дает полное описание этого элемента и позволяет оперировать с ним, не принимая во внимание его «природу» (свойство изоморфизма (см. §7.5)).

Полезность и удобство координатного описания элементов линейного пространства вполне очевидны. Поэтому представляется целесообразной попытка обобщения понятия координат на случай линейных пространств с неограниченным числом линейно независимых элементов.

Возможный подход к решению этой проблемы состоит в следующем.

Вначале напомним теорему 10.3.2, утверждающую, что в n-мерном евклидовом пространстве E^n с базисом $\{\,g_1,\ g_2,\ \ldots\ ,\ g_n\,\}$ равенство (12.3.1) $\forall x \in E^n$ равносильно системе линейных уравнений

$$\sum_{j=1}^{n} \gamma_{ij}\xi_j = \beta_i \qquad \forall i = [1, n]\,, \tag{12.3.2}$$

где $\quad \gamma_{ij} = (g_i, g_j), \quad \beta_i = (x, g_i) \quad$ и $\quad \|x\|_g = \|\,\xi_1\ \xi_2\ \ldots\ \xi_n\,\|^{\mathrm{T}}$.

Основная матрица системы (12.3.2) есть базисная матрица Грама, которая невырождена. В силу чего система (12.3.2) всегда имеет решение и притом единственное.

Правые же части (12.3.2) суть скалярные произведения, первый сомножитель в которых x. Значит, они линейные функции от тех же $\{\,\xi_1,\ \xi_2,\ \ldots\ ,\ \xi_n\,\}$ — координат элемента x. Поэтому, на первый взгляд, от этой теоремы большой пользы нет.

Однако если базис $\{\,e_1,\ e_2,\ \ldots\ ,\ e_n\,\}$ ортонормированный, то базисная матрица Грама единичная и для нее $\gamma_{ij} = (e_i, e_j) = \delta_{ij}$. Система (12.3.2) в этом случае распадается на n линейных уравнений, с одним неизвестным каждое:

$$\xi_j = (x, e_j) \qquad \forall j = [1, n]\,. \tag{12.3.3}$$

И если набор линейно независимых, нормированных и попарно ортогональных элементов $\{e_1, \ e_2, \ \dots \ , \ e_n \, , \dots\}$ не является конечным, то числа $\xi_j \quad \forall j \in \mathbb{N}$ можно принять (*по определению*!) за обобщение евклидовых координат.

Рассмотрим эту идею подробнее.

Предварительно вспомним, что в любом евклидовом пространстве имеется возможность количественно оценивать «степень близости» элементов x и y величиной *нормы* их разности (см. определение 10.1.2), то есть числом $|x-y| = \sqrt{(x-y, x-y)}$.

Это, в свою очередь, позволяет ввести (по аналогии с числовыми последовательностями) понятия *сходимости* и *фундаментальности* для последовательностей элементов в евклидовом пространстве.

Тогда использование понятий *ряда* (как суммы с неограниченным числом слагаемых), а также *сходящегося ряда* (то есть ряда, у которого последовательность частичных сумм имеет предел) оказывается корректным и в E.

Тот факт, что скалярное произведение аксиоматически существует в *любом* евклидовом пространстве (независимо от того, имеется в нем базис или нет), позволяет рассуждать следующим образом.

Пусть $\Upsilon \subset E$ есть множество *линейно независимых*, попарно *ортогональных* и *нормированных* элементов $\{e_1, \ e_2, \ \dots \ , \ e_k, \ \dots\}$ таких, что $\xi_j = (x, e_j)$ и $x = \sum\limits_j \xi_j e_j$.

В случае конечного числа элементов множества Υ данное пространство E имеет базис и выполняется (12.3.1).

Если же число элементов множества Υ неограниченно, а ряд $\sum\limits_{j=1}^{+\infty} \xi_j e_j$ сходится к элементу x, то числа $\xi_j = (x, e_j)$ можно принять за координаты элемента x в евклидовом пространстве E с неограниченным числом линейно независимых элементов и считать верным равенство

$$x = \sum_{j=1}^{+\infty} \xi_j e_j \; . \tag{12.3.4}$$

В этом случае аналогом базиса в E становится упорядоченный набор линейно независимых, нормированных и попарно ортогональных элементов $\{e_1, \ e_2, \ \dots \ , \ e_k, \ \dots\}$, а координатным представлением элемента x в этом базисе $\|x\|_e$ — числовая последовательность $\{\xi_j\}$.

Дадим удобное для дальнейшего

Определение 12.3.1	Набор $\{ e_1,\ e_2,\ \dots,\ e_k,\ \dots \}$ линейно независимых элементов в E таких, что $(e_i, e_j) = 0 \quad \forall i \neq j$, будем называть *ортогональной системой.*

Исследование методов построения ортогональных систем и условий сходимости, образованных из них рядов, выходит за рамки курса линейной алгебры. Здесь же мы ограничимся рассмотрением лишь двух иллюстративных примеров.

В первом из них в евклидовом пространстве функций $x(\tau)$, непрерывных при $\tau \in [-1, 1]$, со скалярным произведением

$$(x, y) = \int\limits_{-1}^{1} x(\tau) y(\tau)\, d\tau$$

в качестве «строительного материала» возьмем $\{ g_k(\tau) \}$ — последовательность линейно независимых элементов $\{ \tau^k \} \quad \forall k = 0, 1, 2, \dots$ Элементы такой последовательности не являются ни нормированными, ни попарно ортогональными. Действительно, например,

$$|g_1| = \sqrt{(g_1, g_1)} = \sqrt{\int\limits_{-1}^{1} \tau^2\, d\tau} = \sqrt{\frac{2}{3}} \neq 1\,, \qquad (g_1, g_3) = \int\limits_{-1}^{1} \tau \cdot \tau^3\, d\tau = \frac{2}{5} \neq 0\,.$$

Однако при помощи процедуры ортогонализации Грама – Шмидта (теорема 10.2.1, которая применима и к наборам с неограниченным числом элементов), из последовательности $\{ g_k(\tau) \}$ можно построить ортогональную систему, а после нормировки и ортонормированную.

В результате (это доказывается в курсе *гармонического анализа*) получается последовательность $\{ e_k(\tau) \}$ ортонормированных степенных многочленов вида

$$e_0(\tau) = 1\,, \qquad e_1(\tau) = \tau\,, \qquad e_2(\tau) = \frac{1}{2}\big(3\tau^2 - 1\big)\,, \quad \dots,$$

$$\dots, \quad e_k(\tau) = \frac{1}{2^k k!} \frac{d^k}{d\tau^k} \big(\tau^2 - 1\big)^k\,, \quad \dots,$$

называемых *полиномами Лежандра.* Этот новый набор многочленов можно использовать для представления на отрезке $[-1, 1]$ любой непрерывной функции $x(\tau)$ при помощи ряда

$$x(\tau) = \sum_{k=0}^{+\infty} \xi_k e_k(\tau)\,, \qquad \text{где} \quad \xi_k = \int\limits_{-1}^{1} x(\tau) e_k(\tau)\, d\tau\,. \qquad (12.3.5)$$

Ряды вида (12.3.5) принято называть *рядами Фурье* и их частичные суммы являются аппроксимациями функции $x(\tau)$ на *всем* промежутке $[-1, 1]$, в то время как *степенные* ряды (типа ряда Тейлора) используются как *локальные* аппроксимации для малых окрестностей фиксированных точек.

Покажем теперь, что ортогональные системы элементов в бесконечномерных евклидовых пространствах можно также строить по схемам, принципиально отличающимся от метода Грама – Шмидта.

В качестве второго примера рассмотрим схему, основанную на том факте, что собственные векторы самосопряженного линейного преобразования, отвечающие различным собственным значениям, образуют ортогональную систему.

Напомним предварительно некоторые рассмотренные нами ранее определения и теоретические факты.

1°. `Опр. 10.6.1.` Линейное преобразование $\hat{A}^{+}$ называется сопряженным линейному преобразованию $\hat{A}$, если
$$\forall x, y \in E \quad (\hat{A}x, y) = (x, \hat{A}^{+}y)\,.$$

2°. `Опр. 10.7.1.`Линейное преобразование $\hat{R}$ называется самосопряженным линейным преобразованием, если
$$\forall x, y \in E \quad (\hat{R}x, y) = (x, \hat{R}y)\,.$$

3°. `Опр. 8.5.2.` Ненулевой элемент f называется собственным вектором линейного преобразования $\hat{A}$, отвечающему собственному значению λ, если $\hat{A}f = \lambda f$.

4°. `Лемма 10.7.3.` Собственные векторы самосопряженного преобразования, отвечающие различным собственным значениям, попарно ортогональны. При этом в E^n из собственных векторов любого самосопряженного преобразования можно образовать ортонормированный базис (`теорема 10.7.1`).

5°. `Пример 10.7.1.` Для любого линейного преобразования $\hat{A}$ линейные преобразования $\hat{A}^{+}\hat{A}$ и $\hat{A}\,\hat{A}^{+}$ самосопряженные и имеют неотрицательные собственные значения.

Рассмотрим множество всевозможных функций $x(\tau)$, имеющих производную любого порядка на отрезке $[-1, 1]$, у которых
$$x^{(k)}(-1) = x^{(k)}(1) \quad \forall k = 0, 1, 2, \ldots ,$$
где k — порядок производной.

Нетрудно проверить, что это множество есть линейное пространство. Превратим его в евклидово пространство, введя скалярное произведение элементов $x(\tau)$ и $y(\tau)$, опять-таки по формуле

$$(x,y) = \int\limits_{-1}^{1} x(\tau)y(\tau)\,d\tau\,.$$

При этом оператор дифференцирования $\hat{D} = \dfrac{d}{d\tau}$ будет линейным преобразованием, действующим в данном пространстве.

Найдем для него сопряженное преобразование. Интегрирование по частям дает равенства

$$(\hat{D}x,y) = \int\limits_{-1}^{1} \frac{dx(\tau)}{d\tau}y(\tau)\,d\tau = x(\tau)y(\tau)\Bigg|_{-1}^{1} - \int\limits_{-1}^{1} \frac{(\tau)}{d\tau}y(\tau)\,d\tau =$$

$$= \int\limits_{-1}^{1} x(\tau)\left(-\frac{dy(\tau)}{d\tau}\right)\,d\tau = (x,\hat{D}^{+}y)\,.$$

Откуда $\hat{D}^{+} = -\dfrac{d}{d\tau}$.

Тогда оператор $\hat{D}\hat{D}^{+} = \dfrac{d}{d\tau}\left(-\dfrac{d}{d\tau}\right) = -\dfrac{d^2}{d\tau^2}$ будет самосопряженным, с неотрицательными собственными значениями.

Найдем его собственные векторы.

Условие $\hat{D}\hat{D}^{+}f = \lambda f$ в данном случае является (если положить в нем $\lambda = \omega^2 \geq 0$) дифференциальным уравнением вида

$$-\frac{d^2}{d\tau^2}f = \lambda f \qquad \Longrightarrow \qquad \frac{d^2 f}{d\tau^2} + \omega^2 f = 0\,,$$

общее вещественное решение которого описывается формулой

$$f(\tau) = A\cos\omega\tau + B\sin\omega\tau\,,$$

причем $A^2 + B^2 \neq 0$, поскольку собственный вектор — ненулевой по своему определению.

В данном случае характеристического уравнения у преобразования нет. Поэтому его собственные значения найдем другим способом, используя условие $f(-1) = f(1)$. Конкретно получаем, что из

$$A\cos(-\omega) + B\sin(-\omega) = A\cos\omega + B\sin\omega \qquad \Longrightarrow \qquad \sin\omega = 0\,,$$

то есть $\omega = \pi k$, где k — любое целое число.

Нетрудно также проверить, что любая производная функции

$$f(\tau) = A\cos\pi k\tau + B\sin\pi k\tau$$

имеет на концах отрезка $[-1,1]$ равные значения.

Таким образом, система линейно независимых, попарно ортогональных функций, являющихся собственными векторами самосопряженного оператора $\hat{D}\hat{D}^{+}$, имеет вид

$$f_k(\tau) = A\cos\pi k\tau + B\sin\pi k\tau \qquad \forall k = 0,\, 1, 2,\, \ldots$$

Эта система, называемая обычно *тригонометрической* (так же как и система полиномов Лежандра), может использоваться для аппроксимации интегрируемой на $[-1,1]$ функции $x(\tau)$ в виде ряда

$$\sum_{k=0}^{\infty} \xi_k f_k(\tau) \qquad \text{или} \qquad \frac{A_0}{2} + \sum_{k=1}^{\infty}\Big(A_k\cos\pi k\tau + B_k\sin\pi k\tau\Big)\,,$$

где коэффициенты A_k и B_k определяются $\forall k = 0,\, 1,\, 2,\, \ldots$ формулами, аналогичными (12.3.5), например,

$$A_k = \int\limits_{-1}^{1} x(v)\cos\pi kv\, dv \qquad \text{и} \qquad B_k = \int\limits_{-1}^{1} x(v)\sin\pi kv\, dv\,.$$

Приложение 1.
Свойства линий второго порядка на плоскости

В § 4.4 были перечислены конкретные типы линий второго порядка, различие между которыми сохраняется при переходе из одной декартовой системы координат в другую. В данном приложении будут рассмотрены характерные свойства линий этих типов.

Приложение 1.1. Вырожденные линии второго порядка

К вырожденным линиям второго порядка будем относить все типы, перечисленные во всех строках таблицы теоремы 4.4.1, кроме последней. Кратко опишем их свойства [20].

1°. Тип линии «Несовпадающие прямые»

Уравнение $\dfrac{x'^2}{a^2} - \dfrac{y'^2}{b^2} = 0$ определяет пару пересекающихся прямых в системе координат $\left\{O',\, \vec{e'}_1,\, \vec{e'}_2\right\}$. В свою очередь уравнение вида $y'^2 = a^2 \quad \forall x'$ при $a > 0$ определяет пару параллельных прямых.

[20] Напомним, что свойства линий и поверхностей второго порядка по умолчанию рассматриваются в прямоугольной (ортонормированной) системе координат.

Пример Прил. 1.1.1. Пусть на плоскости Oxy задана линия второго порядка $3x^2 + 4xy + y^2 = 0$.

Преобразовав ее уравнение выделением полных квадратов к виду $(2x + y)^2 - x^2 = 0$ (*метод Лагранжа*), получим $(3x + y)(x + y) = 0$, то есть две прямые $y = -3x$ и $y = -x$ (см. рис. Прил. 1.1.1).

Проверьте самостоятельно, что в данном случае параметр $\Delta = -1 < 0$, а угол поворота осей системы координат $\alpha = \dfrac{1}{2}\operatorname{arctg} 2$.

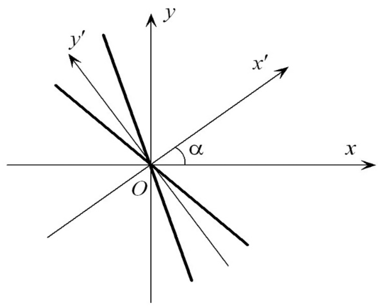


Рис. Прил. 1.1.1

2°. Тип линии «Совпадающие прямые»

Уравнение $y'^2 = 0 \quad \forall x'$ определяет прямую $y' = 0$ в системе координат $\left\{O', \vec{e'}_1, \vec{e'}_2\right\}$. Оно получается из уравнения линии типа 1° предельным переходом при $b \to +0$.

3°. Тип линии «Изолированная точка»

Уравнение $\dfrac{x'^2}{a^2} + \dfrac{y'^2}{b^2} = 0$ определяет единственную точку O' — начало координат системы $\left\{O', \vec{e'}_1, \vec{e'}_2\right\}$.

4°. Тип линии «Пустые множества»

На плоскости не существует точек, координаты которых удовлетворяют уравнениям $\dfrac{x'^2}{a^2} + \dfrac{y'^2}{b^2} = -1$ или $y'^2 = -a^2 \quad \forall x'$ при $a \geq 0$. Однако эти случаи иногда называют *мнимыми линиями*.

Приложение 1.2. Эллипс и его свойства

Определение Прил. 1.2.1	Линия, уравнение которой в некоторой прямоугольной декартовой системе координат есть $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$, где $a \geq b > 0$, называется *эллипсом.*

Определение Прил. 1.2.2	Число $\varepsilon = \dfrac{\sqrt{a^2-b^2}}{a}$ называется *эксцентриситетом* эллипса. Точки $\left\Vert \begin{matrix} \pm\varepsilon a \\ 0 \end{matrix} \right\Vert$ называются *фокусами* эллипса. Прямые $x = \pm\dfrac{a}{\varepsilon}$ называются *директрисами* эллипса. Число $p = \dfrac{b^2}{a}$ называется *фокальным параметром* эллипса.

Свойства эллипса

1. Эллипс — *ограниченная* линия: $|x| \leq a$ и $|y| \leq b$, что следует из записи его канонического уравнения в виде $y = \pm\dfrac{b}{a}\sqrt{a^2-x^2}$.

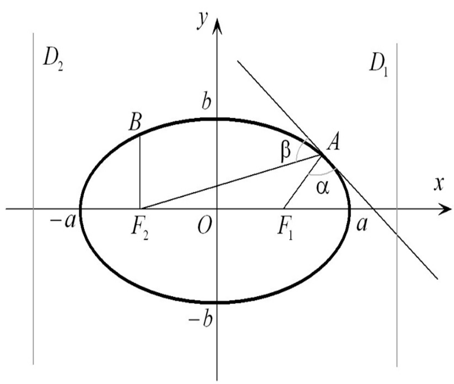


Рис. Прил. 1.2.1

2. Эллипс обладает *осевой симметрией* относительно осей Ox и Oy, а также *центральной симметрией* относительно начала координат O. Это вытекает из отношений

$$\left\| \begin{array}{r} -x \\ y \end{array} \right\| \in L \;\Leftrightarrow\; \left\| \begin{array}{r} x \\ y \end{array} \right\| \in L \;\Leftrightarrow\; \left\| \begin{array}{r} x \\ -y \end{array} \right\| \in L \;\Leftrightarrow\; \left\| \begin{array}{r} -x \\ -y \end{array} \right\| \in L\,,$$

очевидных для канонического уравнения эллипса L.

Обозначим как $\rho(P, Q)$ расстояние между геометрическими объектами P и Q (например, между точкой и прямой), а через α и β обозначим углы между касательной к эллипсу в точке A и фокальными радиусами — отрезками F_1A и F_2A.

Теорема Прил. 1.2.1 **Пусть $A = \|\, x\, y \,\|^{\mathrm{T}}$ есть некоторая точка, принадлежащая эллипсу L, заданному каноническим уравнением, тогда имеют место следующие соотношения:**

1°. $r_1 = |\vec{F_1A}| - \varepsilon a,$

$r_2 = |\vec{F_2A}| + \varepsilon a$**;**

2°. $r_1 + r_2 = |\vec{F_1A}| + |\vec{F_2A}| = 2a$**;**

3°. $\dfrac{\rho(A, F_1)}{\rho(A, D_1)} = \dfrac{\rho(A, F_2)}{\rho(A, D_2)} = \varepsilon$**;**

4°. $\dfrac{\rho(M, F_1)}{\rho(M, D_1)} = \varepsilon \quad \forall M \in L$**;**

5°. $|\vec{F_2B}| = p$**, где** $\vec{F_2B}$ **ортогонален оси** Ox**;**

6°. $\alpha = \beta$**.**

Доказательство.

1°. Имеем (см. рис. Прил. 1.2.1):

$$r_1 = \sqrt{(x - a\varepsilon)^2 + y^2}\,, \qquad r_2 = \sqrt{(x + a\varepsilon)^2 + y^2}\,.$$

Тогда, учитывая каноническое уравнение и определение эксцентриситета, получаем для $i = 1, 2$:

$$r_i = \sqrt{(x \pm a\varepsilon)^2 + y^2} = \sqrt{(x \pm a\varepsilon)^2 + \frac{b^2}{a^2}(a^2 - x^2)} =$$

$$= \sqrt{(x \pm a\varepsilon)^2 + (1 - \varepsilon^2)(a^2 - x^2)} =$$

$$= \sqrt{x^2 \pm 2xa\varepsilon + a^2\varepsilon^2 + a^2 - a^2\varepsilon^2 - x^2 + x^2\varepsilon^2} =$$

$$= \sqrt{a^2 \pm 2xa\varepsilon + x^2\varepsilon^2} = |\, a \pm x\varepsilon|\,.$$

Но поскольку $|x| \leq a$ и $0 \leq \varepsilon < 1$, то $x \pm a\varepsilon \geq 0$ и, следовательно, $r_1 = |\vec{F_1 A}| = a - \varepsilon x$, $\qquad r_2 = |\vec{F_2 A}| = a + \varepsilon x$.

2°. Утверждение 2° очевидно в силу 1°.

3°. Из рис. Прил. 1.2.1 следует

$$\frac{\rho(A, F_1)}{\rho(A, D_1)} = \frac{x - a\varepsilon}{\frac{a}{\varepsilon} - x} = \varepsilon, \qquad \frac{\rho(A, F_2)}{\rho(A, D_2)} = \frac{x + a\varepsilon}{\frac{a}{\varepsilon} + x} = \varepsilon\,.$$

4°. Справедливость 4° докажите самостоятельно.

5°. Наконец, координаты F_2 суть $\{-\varepsilon a;\, 0\}$, поэтому

$$\left|\vec{F_2 B}\right| = \frac{b}{a}\sqrt{a^2 - (\varepsilon a)^2} = \frac{b}{a}\sqrt{a^2 - \frac{a^2 - b^2}{a^2}a^2} = \frac{b^2}{a} = p\,.$$

.

6°. Доказательство приводится после доказательства теоремы Прил. 1.2.2.

Теорема доказана.

Проведение касательных к эллипсу

Теорема Прил. 1.2.2 **Пусть A есть точка с координатами $\{x_0\, y_0\}$, принадлежащая эллипсу L, заданному каноническим уравнением, тогда уравнение касательной к этому эллипсу, проходящей через точку А, имеет вид** $\dfrac{x_0 x}{a^2} + \dfrac{y_0 y}{b^2} = 1\,.$

`Доказательство.`

Уравнение касательной, проведенной к эллипсу в точке A, имеет вид $\quad y - y_0 = y'(x_0)\big(x - x_0\big)$.
Для эллипса из канонического уравнения по правилу дифференцирования функции, заданной неявно, находим

$$\frac{2x}{a^2} + \frac{2yy'}{b^2} = 0\,,$$

то есть при $y_0 \neq 0 \qquad y'(x_0) = -\dfrac{b^2}{a^2}\dfrac{x_0}{y_0}.$

Но тогда $\qquad y - y_0 = -\dfrac{b^2}{a^2}\dfrac{x_0}{y_0}\big(x - x_0\big)\,, \qquad$ и, принимая во внимание, что $\quad \dfrac{x_0^2}{a^2} + \dfrac{y_0^2}{b^2} = 1\,,$ окончательно получаем

$$\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} = 1\,.$$

Наконец, непосредственно проверяем утверждение теоремы для точек $\{\pm a,\, 0\}$, в которых уравнения касательных имеют вид $\; x = \pm a\,.$

`Теорема доказана.`

`Доказательство пункта 6° теоремы Прил. 1.2.1.`

Пусть касательная к эллипсу проведена через точку A, имеющую координаты $\;\{x_0,\, y_0\}$. Тогда d_2 — расстояние от фокуса F_2 с координатами $\;\{-\varepsilon a,\, 0\}$ до касательной с уравнением $\dfrac{x_0}{a^2}x + \dfrac{y_0}{b^2}y - 1 = 0$ — равно (см. задачу 3.2.1):

$$d_2 = \frac{1}{\Delta}\left|\frac{x_0}{a^2}(-\varepsilon a) + \frac{y_0}{b^2}0 - 1\right| = \frac{1}{\Delta a}|\varepsilon x_0 + a| = \frac{r_2}{\Delta a}\,,$$

где $\quad \Delta = \sqrt{\dfrac{x_0^2}{a^2} + \dfrac{y_0^2}{b^2}}\,.$

Аналогично находим d_1 — расстояние от фокуса F_1 с координатами $\;\{-\varepsilon a,\, 0\}$ до касательной: $\qquad d_1 = \dfrac{r_1}{a\Delta}.$

Тогда имеем

$$\sin\alpha = \frac{d_1}{r_1} = \frac{1}{a\Delta} = \frac{d_2}{r_2} = \sin\beta\,.$$

Поскольку углы α и β острые, то из $\sin\alpha = \sin\beta$ следует равенство $\alpha = \beta$.

`Пункт 6° теоремы Прил. 1.2.1 доказан.`

Утверждения теорем Прил. 1.2.1 и Прил. 1.2.2 допускают следующие геометрические интерпретации:

> *Фокальное свойство эллипса*: эллипс есть геометрическое место точек, сумма расстояний от которых до двух фокусов постоянна и равна $2a$.
>
> *Директрисальное свойство эллипса*: эллипс (исключая случай окружности) есть геометрическое место точек, отношение расстояния от которых до данной точки (фокуса) к расстоянию до данной прямой (директрисе) постоянно и меньше единицы.
>
> *Оптическое свойство эллипса*: любой луч света, исходящий из одного фокуса, после отражения в эллипсе (по правилу: «угол падения равен углу отражения») проходит через другой фокус.

Уравнение эллипса в полярной системе координат

Поместим полюс полярной системы координат в левый фокус эллипса, а полярную ось направим по линии, соединяющей его фокусы. Для произвольной точки A, лежащей на эллипсе (рис. Прил. 1.2.2), имеем

$$\rho = r_2 = a + \varepsilon x = a + \varepsilon(\rho\cos\varphi - a\varepsilon) = a + \varepsilon\rho\cos\varphi - a\varepsilon^2\,.$$

Откуда $\rho(1-\varepsilon\cos\varphi) = a - a\varepsilon^2$ и окончательно

$$\rho = \frac{p}{1-\varepsilon\cos\varphi}.$$

(Сравните эти формулы с выкладками в § 4.6.)

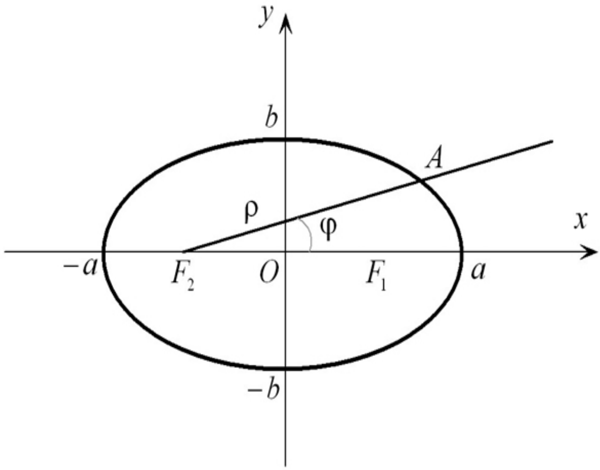


Рис. Прил. 1.2.2

Приложение 1.3. Гипербола и ее свойства

Определение Прил. 1.3.1	Линия, уравнение которой в некоторой прямоугольной декартовой системе координат есть $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$, где $a > 0,\ b > 0,$ называется *гиперболой*.

Определение Прил. 1.3.2	Число $\varepsilon = \dfrac{\sqrt{a^2+b^2}}{a}$ называется *эксцентриситетом* гиперболы. Точки $\left\lVert\begin{matrix}\pm\varepsilon a\\ 0\end{matrix}\right\rVert$ называются *фокусами* гиперболы. Прямые $x = \pm\dfrac{a}{\varepsilon}$ называются *директрисами* гиперболы. Число $p = \dfrac{b^2}{a}$ называется *фокальным параметром* гиперболы.

Свойства гиперболы

1°. Гипербола — неограниченная линия, существующая для $|x| \geq a$, что следует из записи ее канонического уравнения в следующем виде: $\quad y = \pm\dfrac{b}{a}\sqrt{x^2 - a^2}$.

2°. Гипербола обладает осевой симметрией относительно осей Ox и Ox, а также центральной симметрией относительно начала координат. Это вытекает из отношений

$$\left\|\begin{matrix} -x \\ y \end{matrix}\right\| \in L \;\Leftrightarrow\; \left\|\begin{matrix} x \\ y \end{matrix}\right\| \in L \;\Leftrightarrow\; \left\|\begin{matrix} x \\ -y \end{matrix}\right\| \in L \;\Leftrightarrow\; \left\|\begin{matrix} -x \\ -y \end{matrix}\right\| \in L\,,$$

очевидных для канонического уравнения гиперболы L.

Через α и β обозначим углы между касательной и фокальными радиусами (рис. Прил. 1.3.1). Напомним также введенное в курсе математического анализа

Определение Прил. 1.3.3	Прямая $y = ux + v$ называется *асимптотой* для графика функции $y = f(x)$ при $x \to \infty$, если $u = \lim\limits_{x\to\infty} \dfrac{f(x)}{x}$ и $v = \lim\limits_{x\to\infty} (f(x) - ux)$.

3°. Гипербола обладает асимптотами вида $\quad y = \pm\dfrac{b}{a}x$.

Действительно, для функции $\quad y = \dfrac{b}{a}\sqrt{x^2 - a^2} \quad$ имеем

$$u = \lim_{x\to\infty} \frac{b}{ax}\sqrt{x^2 - a^2} = \frac{b}{a}\lim_{x\to\infty} \frac{|x|}{x}\sqrt{1 - \frac{a^2}{x^2}} =$$

$$= \frac{b}{a}\lim_{x\to\infty} \operatorname{sgn}(x)\sqrt{1 - \frac{a^2}{x^2}} = \begin{cases} \;\;b/a & \text{при } x \to +\infty, \\ -b/a & \text{при } x \to -\infty. \end{cases}$$

Кроме того,

$$v = \lim_{x\to\pm\infty} \left(\frac{b}{a}\sqrt{x^2 - a^2} \mp \frac{b}{a}x\right) = \frac{b}{a}\lim_{x\to\pm\infty} \left(\sqrt{x^2 - a^2} \mp x\right) =$$

$$= \frac{b}{a}\lim_{x\to\pm\infty} \frac{x^2 - a^2 - x^2}{\sqrt{x^2 - a^2} \pm x} = -ab \lim_{x\to\pm\infty} \frac{1}{\sqrt{x^2 - a^2} \pm x} = 0\,.$$

Свойства гиперболы иллюстрирует рис. Прил. 1.3.1.

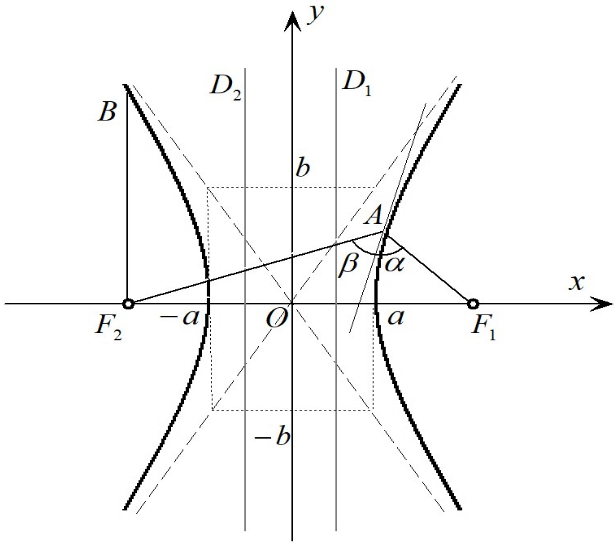


Рис. Прил. 1.3.1

Теорема Прил. 1.3.1

Пусть $A = \| \, x \, y \, \|^{\mathrm{T}}$ есть точка, принадлежащая гиперболе L, заданной каноническим уравнением, тогда имеют место следующие соотношения:

1°. Для правой ветви $(x > a)$**:**

$$r_1 = |\vec{F_1A}| = -a + \varepsilon x,$$
$$r_2 = |\vec{F_2A}| = \quad a + \varepsilon x.$$

Для левой ветви $(x < -a)$**:**

$$r_1 = |\vec{F_1A}| = \quad a - \varepsilon x,$$
$$r_2 = |\vec{F_2A}| = -a - \varepsilon x.$$

2°. $\quad |\, r_1 - r_2| = 2a$**;**

3°. $\quad \dfrac{\rho(A, F_1)}{\rho(A, D_1)} = \dfrac{\rho(A, F_2)}{\rho(A, D_2)} = \varepsilon$**;**

4°. $\quad \dfrac{\rho(M, F_1)}{\rho(M, D_1)} = \varepsilon \quad \forall M \in L$**;**

5°. $\quad |\vec{F_2B}| = p$**, где** $\vec{F_2B}$ **ортогонален оси** Ox**;**

6°. $\quad \alpha = \beta$**.**

Доказательство.

1°. Доказательство аналогично доказательству теоремы Прил. 1.2.1, поэтому ограничимся здесь лишь нахождением величин

$$r_1 = \sqrt{(x - a\varepsilon)^2 + y^2}\,,$$

$$r_2 = \sqrt{(x + a\varepsilon)^2 + y^2}\,.$$

Учитывая каноническое уравнение и определение эксцентриситета, получаем для $i = 1,\,2$

$$r_i = \sqrt{(x \pm a\varepsilon)^2 + y^2} = \sqrt{(x \pm a\varepsilon)^2 + \frac{b^2}{a^2}(a^2 - x^2)} =$$

$$= \sqrt{(x \pm a\varepsilon)^2 + (1 - \varepsilon^2)(a^2 - x^2)} =$$

$$= \sqrt{a^2 \pm 2xa\varepsilon + x^2\varepsilon^2} = |\,a \pm x\varepsilon|\,.$$

Но поскольку для гиперболы $|x| \geq a$ и $\varepsilon > 1$, то для правой ветви имеем

$$r_1 = |\overrightarrow{F_1A}| = -a + \varepsilon x, \qquad r_2 = |\overrightarrow{F_2A}| = \quad a + \varepsilon x,$$

а для левой соответственно

$$r_1 = |\overrightarrow{F_1A}| = \quad a - \varepsilon x, \qquad r_2 = |\overrightarrow{F_2A}| = -a - \varepsilon x.$$

Пункт 1° доказан. Откуда следует, что пункты 2° и 3° также верны.

4°. Справедливость этого условия докажите самостоятельно.

5°. Поскольку координаты F_2 суть $\{-\varepsilon a;\, 0\}$, то имеем

$$\left|\overrightarrow{F_2B}\right| = \frac{b}{a}\sqrt{(\varepsilon a)^2 - a^2} = \frac{b}{a}\sqrt{\frac{a^2+b^2}{a^2}a^2 - a^2} = \frac{b^2}{a} = p\,.$$

6°. Этот пункт докажите самостоятельно по аналогии с доказательством свойства 6° теоремы Прил. 1.2.1, и используя также утверждение теоремы Прил. 1.3.2.

Теорема доказана.

Замечания о свойствах гиперболы

Каноническое уравнение, изучаемой в курсе элементарной математики, гиперболы $y = \dfrac{1}{x}$ в прямоугольной системе координат находится путем следующей замены координат:

$$\begin{cases} x = \dfrac{1}{\sqrt{2}}x' - \dfrac{1}{\sqrt{2}}y' , \\ y = \dfrac{1}{\sqrt{2}}x' + \dfrac{1}{\sqrt{2}}y' , \end{cases}$$

реализующей поворот плоскости Oxy на угол $\dfrac{\pi}{4}$ против часовой стрелки вокруг начала координат.

Из теоремы Прил. 1.3.1 следуют альтернативные формулировки геометрических свойств гиперболы:

> *Фокальное свойство гиперболы*: гипербола есть геометрическое место точек, абсолютная величина разности расстояний от которых до двух фокусов постоянна и равна $2a$.
>
> *Директрисальное свойство гиперболы*: гипербола есть геометрическое место точек, отношение расстояния от которых до данной точки (фокуса) к расстоянию до данной прямой (директрисы) постоянно и больше единицы.
>
> *Оптическое свойство гиперболы*: касательная в любой точке гиперболы образует с фокальными радиусами точки касания равные углы. (Изображение точечного источника света, расположенного в одном из фокусов, есть мнимое и находится в другом фокусе гиперболы.)

Проведение касательных к гиперболе

Теорема Прил. 1.3.2 **Пусть $A = \| \, x_0 \, y_0 \, \|^{\mathrm{T}}$ есть точка, принадлежащая гиперболе L, заданной каноническим уравнением. Тогда уравнение касательной к L, проходящей через точку A, имеет вид** $\dfrac{x_0 x}{a^2} - \dfrac{y_0 y}{b^2} = 1$ **.**

Доказательство.

Уравнение касательной к графику функции $y(x)$ в точке x_0 имеет вид $y - y_0 = y'(x_0)\big(x - x_0\big)$.

Для гиперболы из канонического уравнения по правилу дифференцирования неявно заданной функции получаем

$$\frac{2x_0}{a^2} - \frac{2y_0 y'}{b^2} = 0\,,$$

то есть при $y_0 \neq 0$ производная равна $y'(x_0) = \dfrac{b^2}{a^2}\dfrac{x_0}{y_0}$, откуда

$$y - y_0 = \frac{b^2}{a^2}\frac{x_0}{y_0}\big(x - x_0\big)\,.$$

Принимая во внимание, что из $A \in L \implies \dfrac{x_0^2}{a^2} - \dfrac{y_0^2}{b^2} = 1\,,$

окончательно получаем $\dfrac{x_0 x}{a^2} - \dfrac{y_0 y}{b^2} = 1\,.$

Наконец, непосредственно проверяем утверждение теоремы для точек с $y_0 = 0$, где уравнения касательных имеют вид $x = \pm a$.

Теорема доказана.

Уравнение гиперболы в полярной системе координат

Поместим полюс полярной системы координат в правый фокус гиперболы, а полярную ось направим по положительной полуоси Ox (см. рис. Прил. 1.3.2).

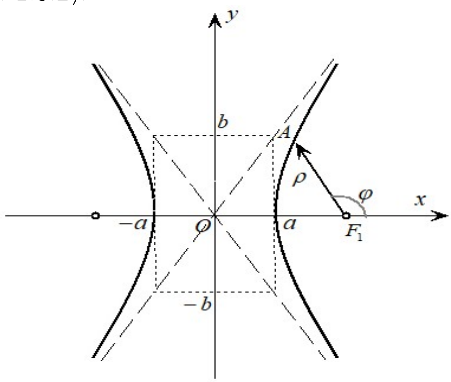


Рис. Прил. 1.3.2

Тогда для произвольной точки A, лежащей на правой ветви гиперболы:

$$\rho = r_1 = -a + \varepsilon x = -a + \varepsilon(\rho\cos\varphi + a\varepsilon) = -a + \varepsilon\rho\cos\varphi + a\varepsilon^2 \,.$$

Откуда

$$\rho(1 - \varepsilon\cos\varphi) = -a + a\varepsilon^2 = -a + a\,\frac{a^2 + b^2}{a^2} = \frac{b^2}{a} = p$$

и окончательно $\rho = \dfrac{p}{1 - \varepsilon\cos\varphi}$. (Сравните эти формулы с выкладками в § 4.6.)

Приложение 1.4. Парабола и ее свойства

Определение Прил. 1.4.1	Линия, уравнение которой в некоторой прямоугольной декартовой системе координат есть $y^2 = 2px$, где $p > 0$, называется *параболой*.

Определение Прил. 1.4.2	Точка $\left\Vert \begin{matrix} \frac{p}{2} \\ 0 \end{matrix} \right\Vert$ называется *фокусом* параболы. Прямая $x = -\dfrac{p}{2}$ называется *директрисой* параболы. Число p называется *фокальным параметром* параболы.

Свойства параболы иллюстрирует рис. Прил. 1.4.1, на котором через α обозначен угол между касательной и фокальным радиусом, а через β — угол между касательной и положительным направлением оси Ox.

Свойства параболы

1°. Парабола — неограниченная кривая, существующая $\forall x \geq 0$.

2°. Парабола L обладает осевой симметрией относительно оси Ox, что вытекает из отношения

$$\left\| \begin{matrix} x \\ y \end{matrix} \right\| \in L \qquad \Longleftrightarrow \qquad \left\| \begin{matrix} x \\ -y \end{matrix} \right\| \in L\,,$$

очевидного для канонического уравнения параболы.

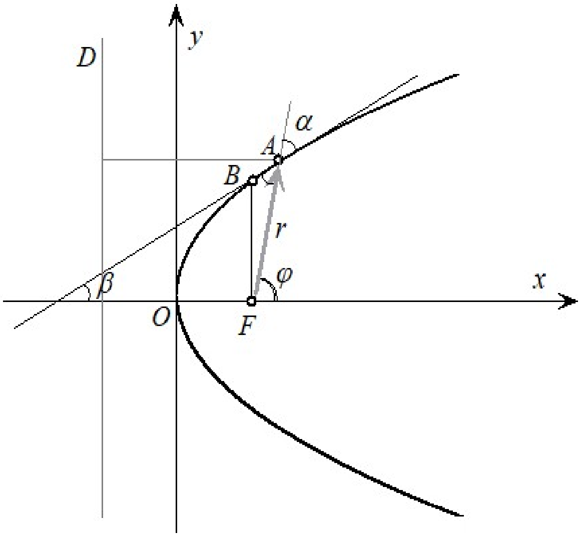


Рис. Прил. 1.4.1

3°. Для параболы имеет место монотонное возрастание абсолютной величины y при возрастании x, причем в нуле касательная к параболе вертикальна.

Теорема Прил. 1.4.1 **Пусть $A = \| \, x \; y \, \|^{\mathrm{T}}$ есть точка, принадлежащая параболе L, заданной каноническим уравнением, тогда имеют место следующие соотношения:**

1°. $\quad r = |\vec{FA}| = x + \dfrac{p}{2}$**;**

2°. $\quad \dfrac{\rho(A,F)}{\rho(A,D)} = 1 \quad \forall A \in L\,,$ $\qquad$ **то есть** $\varepsilon = 1$**;**

3°. $\quad |\vec{FB}| = p\,$**;**

4°. $\quad \alpha = \beta$**.**

Доказательство.

1°. Используя каноническое уравнение, получаем

$$r = \sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} = \sqrt{\left(x - \frac{p}{2}\right)^2 + 2px} =$$

$$= \sqrt{\left(x + \frac{p}{2}\right)^2} = \left|x + \frac{p}{2}\right| = x + \frac{p}{2},$$

поскольку $x \geq -\dfrac{p}{2}$.

2°. Из рис. Прил. 1.4.1 с учетом 1° очевидны равенства

$$\rho(F, A) = r = \frac{p}{2} + x = \rho(F, D) \quad \forall A \in L \qquad \Longrightarrow \qquad \varepsilon = 1\,.$$

3°. Имеем $\quad |\vec{FB}| = \sqrt{2\,p\,\dfrac{p}{2}} = p\,.$

4°. Доказательство приводится после доказательства теоремы Прил. 1.4.2.

Теорема доказана.

Замечания о свойствах параболы

Каноническое уравнение параболы вида $y = ax^2$, изучаемой в курсе элементарной математики, получается путем взаимного переименования координатных переменных x и y.

Из теоремы Прил. 1.4.1 следует возможность альтернативных формулировок свойств параболы.

> *Директориальное свойство параболы*: парабола есть геометрическое место точек, отношение расстояния от которых до данной точки (фокуса) к расстоянию до данной прямой (директрисы) постоянно и равно единице.
>
> *Оптическое свойство параболы*: касательная в любой точке гиперболы образует равные углы с фокальным радиусом точки касания и положительным направлением оси абсцисс.

(Каждый луч света, выходящий из фокуса параболы, после отражения от параболы распространяется параллельно ее оси.)

Проведение касательных к параболе

Теорема Прил. 1.4.2 **Пусть $A = \| \, x_0 \, y_0 \, \|^{\mathrm{T}}$ есть точка, принадлежащая параболе, заданной каноническим уравнением, тогда уравнение касательной к этой параболе, проходящей через точку A, будет $\quad y_0 y = p(x+x_0)$.**

Доказательство.

Уравнение касательной к графику функции $y(x)$ в точке A имеет вид $\quad y - y_0 = y'(x_0)\big(x - x_0\big)$.
Для параболы из канонического уравнения находим, что $2yy' = 2p, \quad$ то есть $\quad y'(x_0) = \dfrac{p}{y_0}, \ \ y_0 \neq 0.$

Но тогда $\quad y - y_0 = \dfrac{p}{y_0}\big(x - x_0\big), \quad$ и, принимая во внимание, что $y_0^2 = 2px_0$, окончательно получаем $\quad y_0 y = p(x + x_0).$

Наконец, непосредственно проверяем утверждение теоремы для точки O , где уравнение касательной $x = 0$.

Теорема доказана.

Доказательство свойства 5° теоремы Прил.1.4.1

Направляющий вектор касательной к параболе в точке ее касания есть $\left\| \begin{matrix} y_0 \\ p \end{matrix} \right\|$, а направляющий вектор фокального радиуса — $\left\| \begin{matrix} x_0 - \dfrac{p}{2} \\ y_0 \end{matrix} \right\|$. Поэтому в силу 1° теоремы 1.4.1

$$\cos\alpha = \frac{y_0\left(x_0 - \dfrac{p}{2}\right) + py_0}{\sqrt{y_0^2 + p^2}\sqrt{\left(x_0 - \dfrac{p}{2}\right)^2 + y_0^2}} = \frac{y_0}{\sqrt{y_0^2 + p^2}}.$$

Но, с другой стороны, косинус угла β — угла между векторами $\left\|\begin{matrix} y_0 \\ p \end{matrix}\right\|$ и $\left\|\begin{matrix} 1 \\ 0 \end{matrix}\right\|$ — выражается той же формулой. Поскольку углы α и β острые, то они равны.

Пункт 5° теоремы Прил.1.4.1 доказан.

Уравнение параболы в полярной системе координат

Поместим полюс полярной системы координат в фокус параболы, а полярную ось направим по линии, перпендикулярной директрисе и проходящей через ее фокус (рис. Прил. 1.4.1). Для произвольной точки A, лежащей на параболе:

$$r = \frac{p}{2} + x = \frac{p}{2} + \frac{p}{2} + r\cos\varphi = p + r\cos\varphi\,.$$

Откуда $r(1 - \cos\varphi) = p$ и окончательно $r = \dfrac{p}{1 - \cos\varphi}$. (Сравните эти формулы с выкладками в § 4.6.)

Приложение 2. Свойства поверхностей второго порядка

В теореме 4.5.1 были перечислены конкретные типы поверхностей второго порядка, различие между которыми сохраняется при переходе из одной декартовой системы координат в другую. В данном приложении будут рассмотрены основные свойства поверхностей этих типов.

Приложение 2.1. Вырожденные поверхности второго порядка

К вырожденным поверхностям второго порядка относятся типы, указанные в табл. 4.5.1.

В первых двух столбцах этой таблицы перечислены типы пустых множеств, а также объекты точечного и линейного типа, исследование которых полностью аналогично случаям, рассмотренным в приложении 1, в ортонормированной, канонической системе координат $\{ O, \vec{e}_1, \vec{e}_2, \vec{e}_3 \}$.

Первые три типа поверхностей, содержащиеся в третьей колонке таблицы, являются частными случаями цилиндрической поверхности, образующая которых параллельна оси Oz, а направляющими служат плоские кривые — эллипс, гипербола и парабола, соответственно расположенные в плоскости Oxy.

Описание свойств невырожденных поверхностей второго порядка (табл. 4.5.2) будет (как и в случае линий 2-го порядка) выполняться в ортонормированной системе координат.

Нетрудно убедиться, что в общем случае в сечении поверхности второго порядка плоскостью получается линия второго порядка. При этом для описания основных свойств поверхностей второго порядка достаточно рассмотреть сечения, параллельные координатным плоскостям.

Приложение 2.2. Эллипсоид

Определение Прил. 2.2.1	Поверхность, задаваемая в некоторой ортонормированной системе координат уравнением вида $$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1, \quad \text{где } a>0,\ b>0,\ c>0\,,$$ называется *эллипсоидом.*

Свойства эллипсоида

1°. Эллипсоид — ограниченная поверхность, поскольку из его канонического уравнения следует, что $\quad |x|\le a, \quad |y|\le b, \quad |z|\le c\,.$

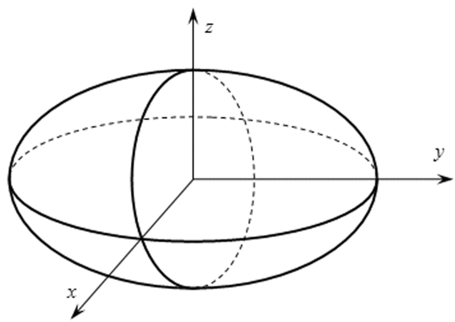


Рис. Прил. 2.2.1

2°. Эллипсоид обладает:

- центральной симметрией относительно начала координат;
- осевой симметрией относительно координатных осей;
- плоскостной симметрией относительно каждой из координатных плоскостей.

3°. В сечении эллипсоида плоскостью, ортогональной любой из осей координат, получается эллипс, точка или пустое множество. Например, используя секущую плоскость $z = z_0$, где $|z_0| < c$, получаем следующую систему уравнений, задающую линию сечения:

$$\begin{cases} \dfrac{x^2}{\left(a\sqrt{1-\dfrac{z_0^2}{c^2}}\right)^2} + \dfrac{y^2}{\left(b\sqrt{1-\dfrac{z_0^2}{c^2}}\right)^2} = 1\,, \\ z = z_0\,, \end{cases}$$

являющейся эллипсом (см. рис. Прил. 2.2.1).

Приложение 2.3. Эллиптический параболоид

Определение Прил. 2.3.1	Поверхность, задаваемая в некоторой ортонормированной системе координат уравнением вида $$\frac{x^2}{a^2}+\frac{y^2}{b^2}=2z, \quad \text{где } a>0,\ b>0\,,$$ называется *эллиптическим параболоидом.*

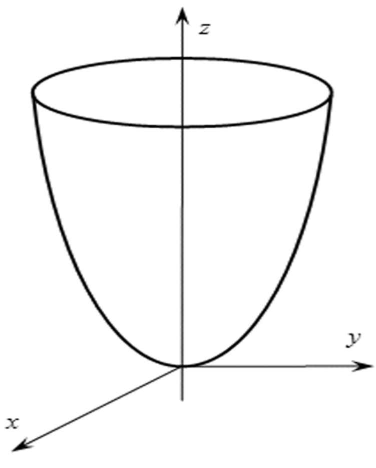


Рис. Прил. 2.3.1

Свойства эллиптического параболоида

1°. Эллиптический параболоид – *неограниченная* поверхность, поскольку из его канонического уравнения следует, что $z \geq 0$ и принимает сколь угодно большие значения.

2°. Эллиптический параболоид обладает
- *осевой симметрией* относительно оси Oz;
- *плоскостной симметрией* относительно координатных плоскостей Oxz и Oyz.

3°. В сечении эллиптического параболоида плоскостью, ортогональной оси Oz, получается *эллипс, точка или пустое множество*, а плоскостями, ортогональными осям Ox или Oy, — *параболы.* Например, рассматривая сечение плоскостью $z = z_0 > 0$, получаем плоскую линию

$$\left\{ \begin{array}{rcl} \dfrac{x^2}{\left(a\sqrt{2z_0}\right)^2} + \dfrac{y^2}{\left(b\sqrt{2z_0}\right)^2} & = & 1\,, \\ z & = & z_0\,, \end{array} \right.$$

являющуюся эллипсом. С другой стороны, сечение плоскостью $y = y_0$ дает линию

$$\left\{ \begin{array}{rcl} x^2 & = & 2a^2\left(z - \dfrac{y_0^2}{2b^2}\right)\,, \\ y & = & y_0\,, \end{array} \right.$$

которая есть парабола. Для случая плоскости $x = x_0$ система уравнений, задающая линию сечения, имеет аналогичный вид

$$\left\{ \begin{array}{rcl} y^2 & = & 2b^2\left(z - \dfrac{x_0^2}{2a^2}\right)\,, \\ x & = & x_0\,. \end{array} \right.$$

Приложение 2.4. Гиперболический параболоид

Определение Прил. 2.4.1	Поверхность, задаваемая в некоторой ортонормированной системе координат уравнением вида $$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 2z, \quad \text{где } a > 0,\ b > 0\,,$$ называется *гиперболическим параболоидом.*

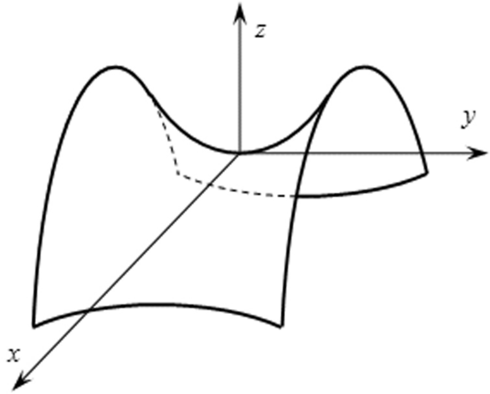


Рис. Прил. 2.4.1

Свойства гиперболического параболоида

1°. Гиперболический параболоид – *неограниченная* поверхность, поскольку из его канонического уравнения следует, что z — любое.

2°. Гиперболический параболоид обладает
- *осевой симметрией* относительно оси Oz;
- *плоскостной симметрией* относительно координатных плоскостей Oxz и Oyz.

3°. В сечении гиперболического параболоида плоскостью, ортогональной оси Oz, получается *гипербола*, а плоскостями, ортогональными осям Ox или Oy, — *параболы*. Например, используя секущую плоскость $z = z_0 > 0$, получаем плоскую линию

$$\left\{ \begin{array}{rcl} \dfrac{x^2}{\left(a\sqrt{2z_0}\,\right)^2} - \dfrac{y^2}{\left(b\sqrt{2z_0}\,\right)^2} & = & 1\,, \\ z & = & z_0\,, \end{array} \right.$$

являющуюся гиперболой. При $z = z_0 < 0$ в сечении будет также получаться гипербола

$$\left\{ \begin{array}{rcl} \dfrac{x^2}{\left(a\sqrt{-2z_0}\,\right)^2} - \dfrac{y^2}{\left(b\sqrt{-2z_0}\,\right)^2} & = & -1\,, \\ z & = & z_0\,. \end{array} \right.$$

С другой стороны, сечение плоскостью $y = y_0$ дает линию

$$\begin{cases} x^2 & = & 2a^2\left(z + \dfrac{y_0^2}{2b^2}\right), \\ y & = & y_0\,, \end{cases}$$

которая есть парабола. Для случая сечения плоскостью $x = x_0$ система уравнений имеет аналогичный вид

$$\begin{cases} y^2 & = & 2b^2\left(z + \dfrac{x_0^2}{2a^2}\right), \\ x & = & x_0\,. \end{cases}$$

Из полученных уравнений следует, что гиперболический параболоид может быть получен поступательным перемещением в пространстве параболы так, что ее вершина перемещается вдоль другой параболы, ось которой параллельна оси первой параболы, ветви направлены противоположно, а плоскости парабол взаимно перпендикулярны.

4°. Гиперболический параболоид имеет два семейства *прямолинейных образующих.*

Действительно, если записать уравнение данной поверхности в виде

$$\left(\frac{x}{a} + \frac{y}{b}\right)\left(\frac{x}{a} - \frac{y}{b}\right) = 2z\,,$$

то можно прийти к заключению, что при любых значениях параметра α точки, лежащие на прямых (проверьте самостоятельно, что это — прямые!)

$$\begin{cases} \dfrac{x}{a} + \dfrac{y}{b} & = & 2\alpha\,, \\ \alpha\left(\dfrac{x}{a} - \dfrac{y}{b}\right) & = & z \end{cases} \qquad \text{и} \qquad \begin{cases} \dfrac{x}{a} - \dfrac{y}{b} & = & 2\alpha\,, \\ \alpha\left(\dfrac{x}{a} + \dfrac{y}{b}\right) & = & z\,, \end{cases}$$

также принадлежат и гиперболическому параболоиду, поскольку почленное перемножение уравнений плоскостей, задающих эти прямые, дает уравнение гиперболического параболоида.

Заметим, что для каждой точки гиперболического параболоида существует *пара прямых*, проходящих через эту точку и целиком лежащих на гиперболическом параболоиде. Уравнения этих прямых могут быть получены (с точностью до некоторого общего ненулевого множителя) путем подбора конкретных значений параметра α.

Приложение 2.5. Однополостный гиперболоид

Определение Прил. 2.5.1	Поверхность, задаваемая в некоторой ортонормированной системе координат уравнением вида $$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1, \quad \text{где } a>0,\ b>0,\ c>0\,,$$ называется *однополостным гиперболоидом.*

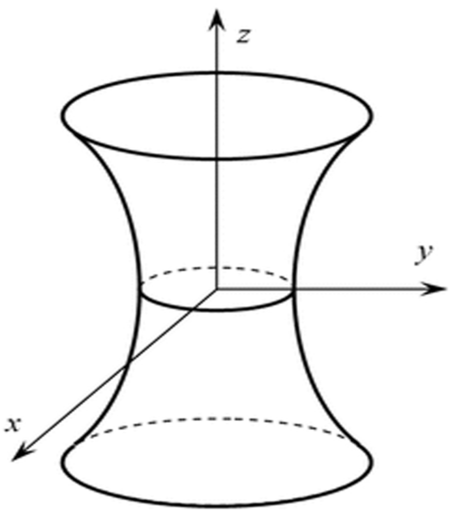


Рис. Прил. 2.5.1

Свойства однополостного гиперболоида

1°. Однополостный гиперболоид — это *неограниченная* поверхность, поскольку из его канонического уравнения (см. опр. Прил. 2.5.1) следует, что $z \in (-\infty, +\infty)$.

2°. Однополостный гиперболоид обладает

- *центральной симметрией* относительно начала координат;
- *осевой симметрией* относительно всех координатных осей.
- *плоскостной симметрией* относительно всех координатных плоскостей.

3°. В сечении однополостного гиперболоида плоскостью, ортогональной оси Oz, получается *эллипс*, а плоскостями, ортогональными осям Ox или Oy, — *гиперболы*. Вывод уравнений для линий сечения аналогичен рассмотренным ранее случаям.

4°. Однополостный гиперболоид имеет два семейства *прямолинейных образующих*.

Действительно, если записать уравнение данной поверхности в виде

$$\left(\frac{x}{a}+\frac{z}{c}\right)\left(\frac{x}{a}-\frac{z}{c}\right)=\left(1+\frac{y}{b}\right)\left(1-\frac{y}{b}\right),$$

то можно прийти к заключению, что при любых, не равных нулю одновременно, значениях параметров α и β точки, лежащие на прямых

$$\begin{cases} \alpha\left(\dfrac{x}{a}+\dfrac{z}{c}\right) = \beta\left(1-\dfrac{y}{b}\right), \\ \beta\left(\dfrac{x}{a}-\dfrac{z}{c}\right) = \alpha\left(1+\dfrac{y}{b}\right) \end{cases}$$

и

$$\begin{cases} \alpha\left(\dfrac{x}{a}+\dfrac{z}{c}\right) = \beta\left(1+\dfrac{y}{b}\right), \\ \beta\left(\dfrac{x}{a}-\dfrac{z}{c}\right) = \alpha\left(1-\dfrac{y}{b}\right), \end{cases}$$

также принадлежат и однополостному гиперболоиду, поскольку почленное перемножение уравнений плоскостей, задающих эти прямые, дает уравнение однополостного гиперболоида.

Заметим, что для каждой точки однополостного гиперболоида существует *пара прямых*, проходящих через эту точку и целиком лежащих на нем. Уравнения этих прямых могут быть получены (с точностью до некоторого общего ненулевого множителя) путем подбора значений параметров α и β.

Проверьте самостоятельно, что нормальные векторы плоскостей в обеих системах неколлинеарны.

Приложение 2.6. Двуполостный гиперболоид

Определение Прил. 2.6.1	Поверхность, задаваемая в некоторой ортонормированной системе координат уравнением вида $$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1, \quad \text{где } a > 0,\ b > 0,\ c > 0\,,$$ называется *двуполостным гиперболоидом.*

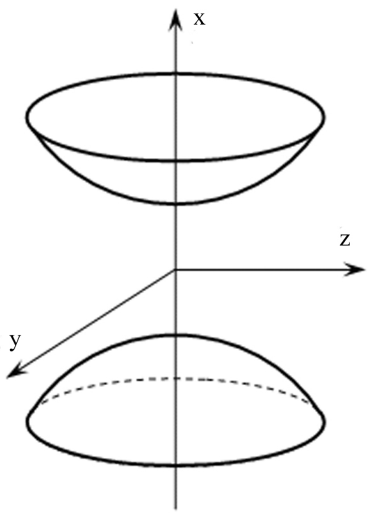


Рис. Прил. 2.6.1

Свойства двуполостного гиперболоида

1°. Двуполостный гиперболоид — это *неограниченная* поверхность, поскольку из его канонического уравнения следует неравенство $a \leq |x| < +\infty$.

2°. Двуполостный гиперболоид обладает

- *центральной симметрией* относительно начала координат;
- *осевой симметрией* относительно всех координатных осей.
- *плоскостной симметрией* относительно всех координатных плоскостей.

3°. В сечении двуполостного гиперболоида плоскостью, ортогональной оси Ox, при $|x| > a$ получается *эллипс*, а плоскостями, ортогональными осям Oy или Oz, — *гиперболы*.

Приложение 2.7. Поверхности вращения

Пусть некоторая линия, расположенная в плоскости Oxz, имеет уравнение $F(x, z) = 0$. Если вращать эту линию вокруг оси Oz, то каждая ее точка будет описывать окружность.

Определение Прил. 2.7.1	Совокупность точек, координаты которых удовлетворяют уравнению $F\left(\pm\sqrt{x^2+y^2}, z\right) = 0$, называется *поверхностью вращения*.

Пример Прил. 2.7.1 К поверхностям вращения, например, относятся:

1°. Эллипсоид вращения
$$\frac{x^2+y^2}{a^2} + \frac{z^2}{c^2} = 1\,.$$

2°. Конус вращения $\quad k^2z^2 = x^2 + y^2\,.$

Замечание Прил. 2.7.1. Поверхности вращения линии второго порядка не всегда задаются уравнениями второго порядка.

Например, если квадратную параболу $z^2 = 2px$ вращать вокруг оси Ox, то получается эллиптический параболоид вращения, однако при вращении этой же кривой вокруг оси Oz получится поверхность, задаваемая уравнением вида
$$z^2 = \pm 2p\sqrt{x^2+y^2} \quad \text{или} \quad z^4 = 4p^2(x^2+y^2).$$

Задача Прил. 2.7.1 *Составить уравнение поверхности вращения, получаемой при вращении линии $z^2 = 2px$ вокруг оси Ox.*

Решение. Зафиксируем на вращаемой линии точку с координатами $\| x_0 \ 0 \ z_0 \|^{\mathrm{T}}$. Линия, получаемая при вращении этой точки вокруг оси Ox в плоскости $x = x_0$, есть окружность, задаваемая уравнением $y^2 + z^2 = z_0^2$.

С другой стороны, $z_0^2 = 2px_0$, поэтому $y^2 + z^2 = 2px_0$.

Наконец, в силу произвольности точки $\| x_0 \ 0 \ z_0 \|^{\mathrm{T}}$, выбранной на линии вращения, получаем, что уравнение поверхности вращения

$$y^2 + z^2 = 2px$$

Решение получено. есть уравнение эллиптического параболоида.

Приложение 3.
Комплексные числа

Рассмотрим двумерное линейное пространство Ω, изоморфное (см. §7.5) линейному пространству радиусов-векторов на плоскости, с ортонормированной системой координат $\{ O, e_1, e_2 \}$.

Каждый элемент z пространства Ω в некотором базисе однозначно задается двухкомпонентным столбцом $\begin{Vmatrix} \alpha \\ \beta \end{Vmatrix}$. Если базисные элементы пространства Ω суть $e_1 = \begin{Vmatrix} 1 \\ 0 \end{Vmatrix}$ и $e_2 = \begin{Vmatrix} 0 \\ 1 \end{Vmatrix}$, то произвольный элемент этого пространства z представляется в виде

$$z = \begin{Vmatrix} \alpha \\ \beta \end{Vmatrix} = \alpha \begin{Vmatrix} 1 \\ 0 \end{Vmatrix} + \beta \begin{Vmatrix} 0 \\ 1 \end{Vmatrix} = \alpha e_1 + \beta e_2 \,.$$

Введем новую операцию — *операцию умножения* элементов рассматриваемого линейного пространства.

Определение Прил. 3.1.1	Результатом операции умножения *двух* элементов $z_1 = \begin{Vmatrix} \alpha_1 \\ \beta_1 \end{Vmatrix}$ и $z_2 = \begin{Vmatrix} \alpha_2 \\ \beta_2 \end{Vmatrix}$ пространства Ω является элемент этого же пространства: $$z_1 \cdot z_2 = \begin{Vmatrix} \alpha_1\alpha_2 - \beta_1\beta_2 \\ \alpha_1\beta_2 + \alpha_2\beta_1 \end{Vmatrix}.$$
Определение Прил. 3.1.2	Двумерное линейное пространство с базисом $$\left\{ e_1 = \begin{Vmatrix} 1 \\ 0 \end{Vmatrix} ;\ \ e_2 = \begin{Vmatrix} 0 \\ 1 \end{Vmatrix} \right\},$$ в котором введена операция умножения элементов согласно определению Прил. 3.1.1, называется *множеством комплексных чисел*, а каждый элемент этого множества z — *комплексным числом.*

`Замечания Прил. 3.1.1.`

1°. Операция умножения комплексных чисел коммутативна и обладает распределительным свойством относительно операции сложения, что следует непосредственно из ее определения.

2°. Операция умножения комплексных чисел позволяет ввести операцию *деления*: частным от деления комплексного числа z_1 на ненулевое z_2 называется комплексное число z^*, такое, что $z_1 = z_2 z^*$.

3°. Нетрудно убедиться, что подмножество комплексных чисел вида $\left\| \begin{matrix} \alpha \\ 0 \end{matrix} \right\|$, где α — произвольное вещественное число, в силу определения Прил. 3.1.2 обладает всеми свойствами вещественных чисел, и потому можно говорить, что вещественные числа есть подмножество комплексных чисел.

На практике более употребительна специальная, упрощающая выкладки, форма записи комплексных чисел: в представлении

$$z = \alpha \left\| \begin{matrix} 1 \\ 0 \end{matrix} \right\| + \beta \left\| \begin{matrix} 0 \\ 1 \end{matrix} \right\| = \alpha e_1 + \beta e_2$$

символ e_1 опускается (заменяется не записываемым явно множителем «единица»), а символ e_2 заменяется символом i (называемым иногда «мнимой единицей»). Тогда произвольное комплексное число z представимо как $\alpha + \beta i$, а записи операций с комплексными числами принимают следующий вид:

$$z_1 + z_2 = (\alpha_1 + \beta_1 i) + (\alpha_2 + \beta_2 i) = (\alpha_1 + \alpha_2) + (\beta_1 + \beta_2) i\,,$$

$$\lambda z = \lambda(\alpha + \beta i) = \lambda\alpha + \lambda\beta i\,, \quad \text{где } \lambda \text{ — вещественное число}\,,$$

$$z_1 \cdot z_2 = (\alpha_1 + \beta_1 i) \cdot (\alpha_2 + \beta_2 i) = (\alpha_1\alpha_2 - \beta_1\beta_2) + (\alpha_1\beta_2 + \beta_1\alpha_2) i\,. \quad (3.1.1)$$

Данная форма записи удобна тем, что с комплексными числами можно оперировать как с обычными алгебраическими двучленами, если принять во внимание, что $i^2 = -1$, поскольку

$$i^2 = i \cdot i = (0 + 1i)(0 + 1i) = \left\| \begin{matrix} 0 \\ 1 \end{matrix} \right\| \cdot \left\| \begin{matrix} 0 \\ 1 \end{matrix} \right\| = \left\| \begin{matrix} -1 \\ 0 \end{matrix} \right\| = (-1) + 0i = -1.$$

Действительно, например, перемножая комплексные числа как двучлены и заменяя повсюду i^2 на число (-1), мы формально приходим к соотношению (3.1.1):

$$z_1 \cdot z_2 = (\alpha_1 + \beta_1 i) \cdot (\alpha_2 + \beta_2 i) = \alpha_1\alpha_2 + \alpha_1\beta_2 i + \alpha_2\beta_1 i + \beta_1 i\beta_2 i =$$

$$= \alpha_1\alpha_2 + \alpha_1\beta_2 i + \alpha_2\beta_1 i + \beta_1\beta_2 i^2 = (\alpha_1\alpha_2 - \beta_1\beta_2) + (\alpha_1\beta_2 + \beta_1\alpha_2) i\,,$$

которое согласуется с введенным выше определением Прил. 3.1.1.

Достаточно просто может выполняться также и операция *деления*:

$$\frac{z_1}{z_2} = \frac{\alpha_1 + \beta_1 i}{\alpha_2 + \beta_2 i} = \frac{(\alpha_1 + \beta_1 i)(\alpha_2 - \beta_2 i)}{(\alpha_2 + \beta_2 i)(\alpha_2 - \beta_2 i)} =$$

$$= \frac{(\alpha_1\alpha_2 + \beta_1\beta_2) + (\alpha_2\beta_1 - \alpha_1\beta_2) i}{\alpha_2^2 + \beta_2^2} = \frac{\alpha_1\alpha_2 + \beta_1\beta_2}{\alpha_2^2 + \beta_2^2} + \frac{\alpha_2\beta_1 - \alpha_1\beta_2}{\alpha_2^2 + \beta_2^2} i\,.$$

`Определение Прил. 3.1.3`

Для комплексного числа $z = \alpha + \beta i$:

1°. Вещественное число α называется *вещественной частью* z и обозначается $\operatorname{Re} z$.

2°. Вещественное число β называется *мнимой частью* z и обозначается $\operatorname{Im} z$.

3°. Вещественное число $\rho = \sqrt{\alpha^2 + \beta^2}$ называется *модулем* z и обозначается $|z|$.

4°. Вещественное число φ, такое, что

$$\begin{cases} \cos\varphi = \dfrac{\alpha}{\sqrt{\alpha^2 + \beta^2}}, \\ \sin\varphi = \dfrac{\beta}{\sqrt{\alpha^2 + \beta^2}}, \end{cases}$$

называется *аргументом* z и обозначается $\arg z$ при условии, что $z \neq 0$.

5°. Число $\alpha - \beta i$ называется *комплексно-сопряженным числу* z и обозначается $\overline{z}$.

`Замечания Прил. 3.1.2.`

1°. Определения, аналогичные пунктам 1°, 2° и 5°, могут быть сделаны и для матриц, элементами которых являются комплексные числа.

2°. Поскольку существует взаимно однозначное соответствие множества радиусов-векторов на плоскости и множества комплексных чисел, то комплексные числа можно изображать точками на плоскости $O\alpha\beta$.

3°. В силу данных выше определений подмножество комплексных чисел вида $z = \alpha + 0 \cdot i \quad \forall \alpha \in \mathbb{R}$ и множество вещественных чисел обладают одинаковыми свойствами. Поэтому их можно считать одним и тем же множеством.

Свойства комплексного сопряжения

Имеют место следующие, легко проверяемые, свойства $\forall z \in \mathbb{C}$:

1°. $\overline{(\overline{z})} = z$.

2°. Число z будет вещественным тогда и только тогда, когда $\overline{z} = z$.

3°. Число $\overline{z}z = \alpha^2 + \beta^2$ всегда вещественное и неотрицательное.

4°. $\overline{(z_1 + z_2)} = \overline{z}_1 + \overline{z}_2\,, \qquad \overline{(z_1 \cdot z_2)} = \overline{z}_1 \cdot \overline{z}_2\,.$

5°. Если $P_n(z) = \sum\limits_{k=0}^{n} \alpha_k z^k$ — многочлен с *вещественными* коэффициентами, имеющий корень λ, то этот многочлен также будет иметь и корень $\overline{\lambda}$. Действительно, пусть $\sum\limits_{k=0}^{n} \alpha_k \lambda^k = 0$,

тогда $\quad 0 = \overline{0} = \overline{\left(\sum\limits_{k=0}^{n} \alpha_k \lambda^k \right)} = \sum\limits_{k=0}^{n} \alpha_k \overline{\lambda}^k.$

`Замечание Прил 3.1.3.` Если алгебраическое уравнение с вещественными коэффициентами имеет комплексные корни, то они попарно сопряжены, а алгебраическое уравнение с вещественными коэффициентами нечетной степени имеет, по крайней мере, один вещественный корень.

`Задача Прил. 3.1.1` *На множестве комплексных чисел решить уравнение* $z^2 + 1 = 0$.

`Решение.` Переписывая это уравнение, приняв, что

$$z = \alpha + \beta i = \left\| \begin{matrix} \alpha \\ \beta \end{matrix} \right\|, \text{ получаем } \left\| \begin{matrix} \alpha \\ \beta \end{matrix} \right\| \cdot \left\| \begin{matrix} \alpha \\ \beta \end{matrix} \right\| + \left\| \begin{matrix} 1 \\ 0 \end{matrix} \right\| = \left\| \begin{matrix} 0 \\ 0 \end{matrix} \right\|.$$

Заметим, что здесь мы воспользовались развернутыми представлениями чисел

$$1 = 1 + 0i = \left\| \begin{array}{c} 1 \\ 0 \end{array} \right\| \qquad \text{и} \qquad 0 = 0 + 0i = \left\| \begin{array}{c} 0 \\ 0 \end{array} \right\| .$$

Выполняя умножение (по определению Прил. 3.1.1) и сложение в левой части уравнения, приходим к равенству

$$\left\| \begin{array}{c} \alpha^2 - \beta^2 + 1 \\ 2\alpha\beta \end{array} \right\| = \left\| \begin{array}{c} 0 \\ 0 \end{array} \right\| .$$

Но поскольку два комплексных числа равны тогда и только тогда, когда одновременно равны их вещественные и мнимые части, то мы получаем следующую систему нелинейных уравнений относительно вещественных неизвестных α и β :

$$\begin{cases} \alpha^2 - \beta^2 + 1 = 0 \,, \\ \qquad\quad 2\alpha\beta = 0 \,, \end{cases}$$

которая имеет два решения $\alpha = 0 \;\; \beta = \pm 1$. Поэтому исходное уравнение также имеет два решения:

Решение получено.

$$z_1 = \left\| \begin{array}{c} 0 \\ 1 \end{array} \right\| = 1 + 0i = i \quad \text{и} \quad z_2 = \left\| \begin{array}{c} 0 \\ -1 \end{array} \right\| = 0 - +0i = -i \,.$$

Тригонометрическая и экспоненциальная формы записи комплексных чисел

Используя определение Прил. 3.1.3, можно получить специальную форму записи ненулевых комплексных чисел, называемую *тригонометрической*:

$$z = \alpha + \beta i = \sqrt{\alpha^2 + \beta^2} \left(\frac{\alpha}{\sqrt{\alpha^2 + \beta^2}} + \frac{\beta}{\sqrt{\alpha^2 + \beta^2}} i \right) = \rho(\cos\varphi + i \sin\varphi).$$

Тригонометрическая форма записи комплексных чисел аналогична описанию точки, изображающей комплексное число, в полярной системе координат.

Пусть направляющим вектором полярной оси служит $e_1 = \left\| \begin{matrix} 1 \\ 0 \end{matrix} \right\|$, а полюс совпадает с началом ортонормированной декартовой системы координат $\{O, e_1, e_2\}$.

При этом

— модуль комплексного числа $|z|$ равен ρ — расстоянию от начала координат до точки, изображающей данное число;

— значение аргумента $\arg z$ совпадает с величиной полярного угла φ, отсчитываемого против часовой стрелки от полярной оси (см. рис. Прил. 3.1.1).

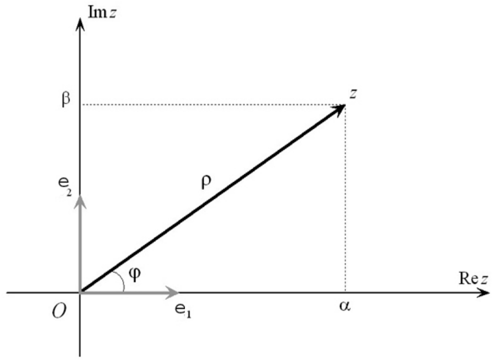


Рис. Прил. 3.1.1

Другой часто используемой формой представления комплексных чисел является их *экспоненциальная форма*, которая получается преобразованием тригонометрической формы по формуле Эйлера:

$$\forall z: \qquad e^{iz} = \cos z + i \sin z\,.$$

В этом случае из $z = |z|(\cos\varphi + i\sin\varphi)$ следует, что $z = \rho e^{i\varphi}$.

Использование экспоненциальной формы записи комплексных чисел может упростить решение некоторых задач, поскольку при перемножении комплексных чисел их модули перемножаются, а аргументы складываются: $z_1 z_2 = \rho_1 e^{i\varphi_1} \rho_2 e^{i\varphi_2} = \rho_1 \rho_2\, e^{i(\varphi_1 + \varphi_2)}$.

Пусть, например, требуется найти какое-нибудь значение $i^{\,i}$. Приняв во внимание, что $i = 0 + 1 \cdot i = \cos\frac{\pi}{2} + i \sin\frac{\pi}{2} = e^{i\frac{\pi}{2}}$, получим $i^{\,i} = \left(e^{i\frac{\pi}{2}}\right)^{i} = e^{i^2\frac{\pi}{2}} = e^{-\frac{\pi}{2}}$.

Задача Прил. 3.1.2 *Найти вещественные решения уравнения*

$$\cos\sqrt{x} = 5\,.$$

Решение. Из формулы Эйлера следует, что $\cos z = \dfrac{e^{iz}+e^{-iz}}{2}$, поэтому данное уравнение можно записать в виде

$$\frac{e^{i\sqrt{x}}+e^{-i\sqrt{x}}}{2} = 5$$

или $\quad y + \dfrac{1}{y} - 10 = 0\,,\quad$ где $\quad y = e^{i\sqrt{x}}\,.$

Откуда находим, что $e^{i\sqrt{x}} = y = 5 \pm 2\sqrt{6}\,,$ то есть $i\sqrt{x} = \ln(5\pm 2\sqrt{6})\,.$

Поскольку $5 - 2\sqrt{6} = \dfrac{1}{5+2\sqrt{6}}$, то решение единственное и оно вещественное

Решение получено.

$$x = -\ln^2(5+2\sqrt{6})\,.$$

Приложение 4. Элементы тензорного исчисления

Приложение 4.1. Замечания об определении объектов в линейном пространстве

В предыдущих разделах курса линейной алгебры исследовались наиболее часто встречающиеся в приложениях виды объектов в линейном пространстве, такие как: элемент линейного пространства, линейный функционал (форма), линейный оператор, билинейный функционал (форма) и т.д., хотя вполне очевидно, что в линейном пространстве могут быть определены и иные, быть может, более сложные объекты, представляющие практический интерес.

Определение всех рассмотренных ранее объектов давалось *вне зависимости* от наличия или отсутствия базиса линейного пространства, причем в случае существования базиса для каждого из объектов приводился альтернативный, координатный способ его описания. А поскольку замена базиса меняет, вообще говоря, данное описание, то специально исследовался вопрос о характере этого изменения.

Однако естественно допустить, что в линейном пространстве Λ^n существуют объекты, которые можно определить, используя лишь значения их координат. Такой подход привлекателен тем, что, во-первых, в этом случае не требуется объяснять, что представляет собой данный объект безотносительно к базису, и, во-вторых, определения объектов разной природы могут быть выполнены единообразно.

С другой стороны, недостатком такой схемы является очевидная зависимость описания объекта от выбора базиса, то есть необходимо указывать (в самом определении объекта!), что происходит с его компонентами при переходе от одного базиса к другому.

Для оценки целесообразности использования определения объектов в Λ^n через их компоненты приведем в табл. Прил. 4.1.1 основные, рассмотренные нами ранее, типы объектов, формы их представления в базисе $\{ g_1, g_2, \ldots, g_n \}$ и правила изменения этого представления при переходе от базиса $\{ g_1, g_2, \ldots, g_n \}$ к базису $\{ g'_1, g'_2, \ldots, g'_n \}$.

Как и ранее, будем предполагать, что матрица перехода $\|S\|$ имеет компоненты σ_{ij}, где $g'_j = \sum\limits_{i=1}^{n} \sigma_{ij} g_i \quad \forall j \in [1, n]$, а матрица обратного перехода $\|T\|$ имеет компоненты τ_{ij}, то есть $g_j = \sum\limits_{i=1}^{n} \tau_{ij} g'_i \quad \forall j \in [1, n]$.

Таблица Прил. 4.1.1

Тип объекта в Λ^n	**Координатное представление в базисе** $\{ g_1, g_2, \ldots, g_n \}$	**Правило изменения координатного представления при переходе к базису** $\{ g'_1, g'_2, \ldots, g'_n \}$
Элемент x	Столбец $\|x\|_g = \left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \cdots \\ \xi_n \end{matrix} \right\|$	$\|x\|_{g'} = \|T\| \|x\|_g$ или $\xi'_j = \sum\limits_{i=1}^{n} \tau_{ij} \xi_j \quad \forall j \in [1, n]$
Линейный функционал $f(x)$	Строка $\|f\|_g = \| \varphi_1, \varphi_2, \ldots \varphi_n \|$	$\|f\|_{g'} = \|f\|_g \|S\|_g$ или $\varphi'_j = \sum\limits_{i=1}^{n} \varphi_i \sigma_{ij} \quad \forall j \in [1, n]$
Линейный оператор (линейное преобразование) $\hat{A}$	Матрица $\|\hat{A}\|_g = \left\| \begin{matrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \alpha_{31} & \alpha_{32} & \cdots & \alpha_{3n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{matrix} \right\|$, где $\hat{A} g_j = \sum\limits_{i=1}^{n} \alpha_{ij} g_i \quad \forall j = [1, n]$	$\|\hat{A}\|_{g'} = \|S\|^{-1} \|\hat{A}\|_g \|S\|$ или $\alpha'_{ki} = \sum\limits_{j=1}^{n} \sum\limits_{m=1}^{n} \tau_{kj} \alpha_{jm} \sigma_{mi}$ $\forall k, j = [1, n]$

Билинейный функционал (билинейная форма) $B(x,y)$	Матрица $|B|_g = $ $= \left\| \begin{matrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2n} \\ \beta_{31} & \beta_{32} & \cdots & \beta_{3n} \\ \cdots & \cdots & \cdots & \cdots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nn} \end{matrix} \right\|,$ где $\beta_{ij} = B(g_i, g_j) \quad \forall i, j = [1, n]$	$\|B\|_{g'} = \|S\|^{\mathrm{T}} \|B\|_g \|S\|$ или $\beta'_{ki} = \sum\limits_{j=1}^{n} \sum\limits_{m=1}^{n} \sigma^{\mathrm{T}}_{kj} \beta_{jm} \sigma_{mi}$ $\forall k, j = [1, n]$
Квадратичный функционал (квадратичная форма) $\Phi(x)$	Матрица $|\Phi|_g = $ $= \left\| \begin{matrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1n} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2n} \\ \varphi_{31} & \varphi_{32} & \cdots & \varphi_{3n} \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_{n1} & \varphi_{n2} & \cdots & \varphi_{nn} \end{matrix} \right\|,$ где $\varphi_{ij} = \dfrac{\beta_{ij} + \beta_{ji}}{2} \quad \forall i, j = [1, n]$	$\|\Phi\|_{g'} = \|S\|^{\mathrm{T}} \|\Phi\|_g \|S\|$ или $\varphi'_{ki} = \sum\limits_{j=1}^{n} \sum\limits_{m=1}^{n} \sigma^{\mathrm{T}}_{kj} \varphi_{jm} \sigma_{mi}$ $\forall k, j = [1, n]$

Сопоставление формул третьей колонки таблицы позволяет заметить, что для данных объектов:

1°. Значения их компонентов в базисе $\{ g'_1, g'_2, \ldots, g'_n \}$ линейны по значениям компонентов в базисе $\{ g_1, g_2, \ldots, g_n \}$.

2°. Коэффициентами в этих формулах служат либо компоненты матриц $\|S\|$ или $\|T\| = \|S\|^{-1}$, либо и той и другой одновременно.

В курсе линейной алгебры нами были рассмотрены далеко не все виды объектов, которые обладают подобными трансформационными свойствами.

Например, для $\forall x, y \in \Lambda^n$ можно ввести произведение $x \otimes y$, поставив в каждом базисе упорядоченной паре элементов

$$\|x\|_g = \| \, \xi_1 \, \xi_2 \, \ldots \, \xi_n \, \|^{\mathrm{T}} \qquad \text{и} \qquad \|y\|_g = \| \, \eta_1 \, \eta_2 \, \ldots \, \eta_n \, \|^{\mathrm{T}}$$

в соответствие матрицу $\|x \otimes y\|_g$ размера $n \times n$, вида

$$\|x \otimes y\|_g = \|x\|_g \|y\|_g^{\mathrm{T}} = \left\| \begin{array}{cccc} \xi_1\eta_1 & \xi_1\eta_2 & \cdots & \xi_1\eta_n \\ \xi_2\eta_1 & \xi_2\eta_2 & \cdots & \xi_2\eta_n \\ \cdots & \cdots & \cdots & \cdots \\ \xi_n\eta_1 & \xi_n\eta_2 & \cdots & \xi_n\eta_n \end{array} \right\| .$$

Убедимся, что матричный объект $x \otimes y$ при переходе от базиса $\{ g_1, g_2, \ldots, g_n \}$ к базису $\{ g'_1, g'_2, \ldots, g'_n \}$ меняется в соответствии с правилами 1° и 2°.

Действительно, по формулам перехода

$$\xi'_k = \sum_{i=1}^{n} \tau_{ki}\xi_i \quad \forall k \in [1,n] \qquad \text{и} \qquad \eta'_m = \sum_{j=1}^{n} \tau_{mj}\xi_j \quad \forall m \in [1,n] \,,$$

имеющих матричный вид

$$\|x\|_{g'} = \|T\|\|x\|_g \quad \text{и} \quad \|y\|_{g'} = \|T\|\|y\|_g \quad \Longrightarrow \quad \|y\|_{g'}^{\mathrm{T}} = \|y\|_g^{\mathrm{T}}\|T\|^{\mathrm{T}} ,$$

получим, если использовать для столбцов двухиндексное описание,

$$\xi'_{k1}{\eta'}^{\mathrm{T}}_{1m} = \sum_{i=1}^{n}\sum_{j=1}^{n} \tau_{ki}\,\xi_{i1}\eta_{1j}^{\mathrm{T}}\tau_{jm}^{\mathrm{T}} \qquad \forall k,m = [1,n]$$

или то же самое, но в матричной форме:

$$\|x \otimes y\|_{g'} = \|x\|_{g'}\|y\|_{g'}^{\mathrm{T}} = \|T\|\|x\|_g\|y\|_g^{\mathrm{T}}\|T\|^{\mathrm{T}} = \|T\|\|x \otimes y\|_g\|T\|^{\mathrm{T}} .$$

Последнее равенство означает, что введенное нами произведение элементов $x, y \in \Lambda^n$ обладает свойствами 1° и 2°.

Рассмотрим другой пример, демонстрирующий существование более сложных объектов, обладающих данными свойствами.

Достаточно часто в приложениях линейное преобразование используется для описания зависимости одного вектора, характеризующего некоторое свойство точки пространства, от другого вектора, являющегося другой количественной характеристикой этой же точки.

Например, закон Гука связывает вектор силы $\vec{F} = \|\, F_x\, F_y\, F_z \,\|^{\mathrm{T}}$, возникающей в результате упругой деформации, с вектором деформации $\Delta\vec{r} = \|\, \Delta x\, \Delta y\, \Delta z \,\|^{\mathrm{T}}$ соотношением

$$\left\| \begin{array}{c} F_x \\ F_y \\ F_z \end{array} \right\| = \left\| \begin{array}{ccc} \kappa_{xx} & \kappa_{xy} & \kappa_{xz} \\ \kappa_{yx} & \kappa_{yy} & \kappa_{yz} \\ \kappa_{zx} & \kappa_{zy} & \kappa_{zz} \end{array} \right\| \left\| \begin{array}{c} \Delta x \\ \Delta y \\ \Delta z \end{array} \right\| .$$

Или же индукция электрического поля $\vec{D} = \| \, D_x \, D_y \, D_z \, \|^{\mathrm{T}}$ связана с напряженностью электрического поля $\vec{E} = \| \, E_x \, E_y \, E_z \, \|^{\mathrm{T}}$ формулой

$$\left\| \begin{array}{c} D_x \\ D_y \\ D_z \end{array} \right\| = \left\| \begin{array}{ccc} \theta_{xx} & \theta_{xy} & \theta_{xz} \\ \theta_{yx} & \theta_{yy} & \theta_{yz} \\ \theta_{zx} & \theta_{zy} & \theta_{zz} \end{array} \right\| \left\| \begin{array}{c} E_x \\ E_y \\ E_z \end{array} \right\| .$$

Если среда однородная, то коэффициенты матриц этих операторов константы. Однако если исследуемые свойства среды меняются от точки к точке, то соответствующие операторы уже не будут линейными, и может возникнуть вопрос о характере их зависимости от координат.

В этом случае полезно ввести в рассмотрение объект, компоненты которого являются частными производными компонентов матрицы оператора по переменным x, y и z. Для рассматриваемого примера таких частных производных будет 27, и их удобно представить в виде трехмерной таблицы (или, как иногда говорят, «куб-матрицы»).

Например, для закона Гука этот объект состоит из трех матриц:

$$\left\| \begin{array}{ccc} \dfrac{\partial \kappa_{xx}}{\partial x} & \dfrac{\partial \kappa_{xy}}{\partial x} & \dfrac{\partial \kappa_{xz}}{\partial x} \\ \dfrac{\partial \kappa_{yx}}{\partial x} & \dfrac{\partial \kappa_{yy}}{\partial x} & \dfrac{\partial \kappa_{yz}}{\partial x} \\ \dfrac{\partial \kappa_{zx}}{\partial x} & \dfrac{\partial \kappa_{zy}}{\partial x} & \dfrac{\partial \kappa_{zz}}{\partial x} \end{array} \right\| , \left\| \begin{array}{ccc} \dfrac{\partial \kappa_{xx}}{\partial y} & \dfrac{\partial \kappa_{xy}}{\partial y} & \dfrac{\partial \kappa_{xz}}{\partial y} \\ \dfrac{\partial \kappa_{yx}}{\partial y} & \dfrac{\partial \kappa_{yy}}{\partial y} & \dfrac{\partial \kappa_{yz}}{\partial y} \\ \dfrac{\partial \kappa_{zx}}{\partial y} & \dfrac{\partial \kappa_{zy}}{\partial y} & \dfrac{\partial \kappa_{zz}}{\partial y} \end{array} \right\| , \left\| \begin{array}{ccc} \dfrac{\partial \kappa_{xx}}{\partial z} & \dfrac{\partial \kappa_{xy}}{\partial z} & \dfrac{\partial \kappa_{xz}}{\partial z} \\ \dfrac{\partial \kappa_{yx}}{\partial z} & \dfrac{\partial \kappa_{yy}}{\partial z} & \dfrac{\partial \kappa_{yz}}{\partial z} \\ \dfrac{\partial \kappa_{zx}}{\partial z} & \dfrac{\partial \kappa_{zy}}{\partial z} & \dfrac{\partial \kappa_{zz}}{\partial z} \end{array} \right\| .$$

В общем случае n-мерного линейного пространства можно ввести объект, называемый *производной оператора*, обозначаемый $\left\| \dfrac{\partial \hat{A}}{\partial r} \right\|_g$ и задаваемый в конкретном базисе $\{ \, g_1, \, g_2, \, \ldots, \, g_n \, \}$ упорядоченным набором из n^3 чисел.

Найдем закон преобразования компонентов этого объекта при переходе от базиса $\{ \, g_1, \, g_2, \, \ldots, \, g_n \, \}$ к базису $\{ \, g'_1, \, g'_2, \, \ldots, \, g'_n \, \}$. Поскольку правило изменения компонентов матрицы оператора $\hat{A}$ при смене базиса в Λ^n имеет вид $\|\hat{A}\|_{g'} = \|S\|^{-1}\|\hat{A}\|_g\|S\|$ (или, в координатах, $\alpha'_{ki} = \sum\limits_{j=1}^{n}\sum\limits_{m=1}^{n} \tau_{kj}\alpha_{jm}\sigma_{mi} \quad \forall k, i = [1, n]$), то из правила дифференцирования сложной функции, а также формул перехода $\xi_p = \sum\limits_{l=1}^{n} \sigma_{pl}\xi'_l \quad \forall p = [1, n]$ следует, что $\forall k, i, t = [1, n]$:

$$\frac{\partial \alpha'_{ki}}{\partial \xi_t} = \sum_{j=1}^{n}\sum_{m=1}^{n} \tau_{kj}\frac{\partial \alpha_{jm}}{\partial \xi_t}\sigma_{mi} = \sum_{j=1}^{n}\sum_{m=1}^{n} \tau_{kj}\sum_{l=1}^{n}\frac{\partial \alpha_{jm}}{\partial \xi_l}\frac{\partial \xi_l}{\partial \xi'_t}\sigma_{mi} =$$

$$= \sum_{j=1}^{n} \sum_{m=1}^{n} \sum_{l=1}^{n} \tau_{kj} \sigma_{tl}^{\mathrm{T}} \frac{\partial \alpha_{jm}}{\partial \xi_l} \sigma_{mi} \, .$$

Откуда делаем заключение, что введенный нами новый объект — производная линейного оператора, также обладает свойствами 1° и 2°.

С другой стороны, отметим, что не всякий однозначно определяемый своими компонентами объект будет обладать подобными трансформационными свойствами.

Таков, например, однокомпонентный объект ω, значение которого для каждого элемента $\|x\|_g = \| \, \xi_1, \, \xi_2, \, \dots \, \xi_n \, \|^{\mathrm{T}}$ пространства Λ^n есть сумма компонентов x. Для него в базисе $\{ \, g_1, \, g_2, \, \dots, \, g_n \, \}$ имеем $\omega = \sum\limits_{i=1}^{n} \xi_i$. Заметим что, хотя значение ω' и определяется с помощью формул перехода однозначно в базисе $\{ \, g'_1, \, g'_2, \, \dots, \, g'_n \, \}$, оно не выражается линейно через ω, так как $\omega' = \sum\limits_{i=1}^{n} \xi'_i = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} \tau_{ij} \xi_j$.

Резюмируя вышесказанное, мы приходим к заключению, что в конечномерном линейном пространстве существует достаточно широкий класс объектов:

- полностью и однозначно описываемых совокупностью значений своих координат в любом базисе;
- обладающих свойствами 1° и 2°, однозначно определяющими изменения этих координат при переходе от одного базиса к другому.

Объекты, обладающие перечисленными свойствами, называют *тензорами*, уточняя это название в случае присутствия матриц $\|S\|$ или $\|S\|^{\mathrm{T}}$ в формулах пересчета компонентов тензора при замене базиса термином *ковариантный* (то есть преобразующийся так же, как и базисные элементы) или же в случае матриц $\|T\|$ или $\|T\|^{\mathrm{T}}$ — термином *контравариантный*.

Приложение 4.2. Определение и обозначение тензоров

Определение тензора, исходя из вышеизложенных соображений, можно было бы дать, например, в следующей форме.

Будем говорить, что в вещественном линейном пространстве Λ^n определен *тензор типа* (или *ранга*) (q,p) — q раз *контравариантный* и p раз *ковариантный* (или $(q+p)$-*валентный*), если в Λ^n задан объект, который в каждом базисе характеризуется упорядоченным набором n^{q+p} чисел $\xi_{j_1 j_2 \ldots j_q i_1 i_2 \ldots i_p}$ (где $j_1, j_2, \ldots, j_q = [1,n]$ — контравариантные индексы и $i_1, i_2, \ldots, i_p = [1,n]$ — ковариантные), преобразующихся при переходе от базиса $\{\, g_1,\, g_2,\, \ldots,\, g_n \,\}$ к базису $\{\, g'_1,\, g'_2,\, \ldots,\, g'_n \,\}$ по закону

$$\xi'_{j'_1 j'_2 \ldots j'_q i'_1 i'_2 \ldots i'_p} = \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_q=1}^{n} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_p=1}^{n} \tau_{j'_1 j_1} \tau_{j'_2 j_2} \ldots \tau_{j'_q j_q} \times$$

$$\times \xi_{j_1 j_2 \ldots j_q i_1 i_2 \ldots i_p} \sigma_{i_1 i'_1} \sigma_{i_2 i'_2} \ldots \sigma_{i_p i'_p}\,,$$

где $i'_k = [1,n],\ k = [1,p]$ и $j'_k = [1,n],\ k = [1,q]$, а σ_{ij} и τ_{ji} суть соответственно компоненты матрицы перехода $\|S\|$ и ей обратной $\|T\| = \|S\|^{-1}$.

Громоздкость записи и неудобочитаемость тензоров при использовании такой схемы обозначений очевидны уже на примере этого определения.

Поэтому в тензорном исчислении используется специальная, более компактная форма описания тензорных объектов и операций с ними, основу которой составляют следующие правила.

Запись тензоров

1°. Упорядоченный набор вещественных чисел, являющихся компонентами $(q+p)$-валентного тензора, образует $(q+p)$-мерную таблицу (называемую также $(q+p)$-мерным массивом, или $(q+p)$-мерной матрицей), каждый элемент которой однозначно определен набором значений контравариантных индексов $j_1,\, j_2,\, \ldots\, j_q$ и ковариантных индексов $i_1,\, i_2,\, \ldots\, i_p$.

Если какой-либо из индексов принимает значения от 1 до n, то в записи тензора перечень значений индекса не указывается и предполагается, что компоненты тензора для всех этих значений заданы.

Пример Прил.4.2.1 Запись $\xi_i = \eta_i$ означает, что $\xi_i = \eta_i \quad \forall i = [1,n]$.

2°. Порядок следования индексов в покомпонентной (координатной) записи тензоров *существен.*

Для того чтобы избежать возможной неоднозначности при последовательной записи координат тензора, применяется следующий алгоритм:

у первой координаты в списке все индексы полагаем равными единице, и если $n = 1$, то список координат тензора построен. Иначе, переходим к формированию индексного списка для второй координаты.

$\mathbb{A}$: Для рассматриваемой координаты берем список индексов предыдущей, в котором ищем *самый правый* индекс, значение которого можно увеличить (то есть индекс со значением, меньшим, чем n), и этот индекс *увеличиваем на единицу.*

Если значение увеличенного индекса оказывается меньше n, то список индексов для рассматриваемой координаты построен и переходим к $\mathbb{A}$ (то есть к построению индексного списка следующей координаты).

Если же значение увеличенного индекса равно n и он не первый в индексном списке, то значениям всех индексов, стоящих справа от него, присваивается единица и делается переход к $\mathbb{A}$. Иначе, построение списка координат тензора завершено.

`Пример Прил.4.2.2` Для тензора ξ_{ijk} в Λ^2 порядок записи координат таков: $\xi_{111},\ \xi_{112},\ \xi_{121},\ \xi_{122},\ \xi_{211},\ \xi_{212},\ \xi_{221},\ \xi_{222}\,.$

3°. В тензорных записях для отличия контравариантных индексов от ковариантных принято первые обозначать *верхними* индексами, а вторые – *нижними.* При этом, чтобы сохранить общий порядок следования индексов в списке, в запись каждого индекса добавляется символ «точка» под каждым верхним индексом и над каждым нижним.

`Пример Прил.4.2.3` $\xi_{i\,\bullet\bullet\,ml}^{\bullet\,jk\,\bullet\bullet}\ .$

4°. Если точки не использованы в записи тензора, то предполагается, что нижние индексы следуют в списке после верхних.

Пример Прил.4.2.4 Линейное преобразование (линейный оператор) $\hat{A}$, переводящее базис $\{ g_1, g_2, \ldots, g_n \}$ в базис $\{ g'_1, g'_2, \ldots, g'_n \}$, является двухвалентным (или второго ранга) тензором типа $(1,1)$ (один раз контравариантным и один раз ковариантным), поскольку

$$\|\hat{A}\|_{g'} = \|T\| \|\hat{A}\|_g \|S\| \,, \qquad \text{или в силу } 3^\circ$$

$$\alpha'^{j'}_{i'} = \sum_{i=1}^{n} \sum_{j=1}^{n} \tau^{j'}_{j} \alpha^{j}_{i} \sigma^{i}_{i'} \qquad \forall i', j' = [1, n] \,.$$

Причем его компоненты α^j_i совпадают с компонентами матрицы перехода σ_{ij} — как следствие совпадения определения *матрицы перехода* 7.3.2 от нештрихованного базиса к штрихованному и определения *матрицы линейного оператора* 8.3.1 в исходном базисе.

Соглашение о суммировании

Пусть имеется выражение, являющееся произведением сомножителей, имеющих как верхние, так и нижние индексы, причем некоторый индекс встречается в записи выражения дважды: один раз как верхний, а второй раз как нижний. Тогда такое выражение означает *сумму выражений данного вида*, выписанных для всех значений повторяющегося индекса.

В случае присутствия в выражении нескольких пар совпадающих индексов имеет место многократное суммирование.

Пример Прил.4.2.5

1°. Квадратичный функционал (квадратичная форма) записывается теперь в виде $\Phi(x) = \varphi_{ij}\xi^i\xi^j$.

2°. Система линейных уравнений вида

$$\left\{ \begin{array}{ccccccccc} \alpha^1_1\xi^1 & + & \alpha^1_2\xi^2 & + & \cdots & + & \alpha^1_n\xi^n & = & \beta^1 \,, \\ \alpha^2_1\xi^1 & + & \alpha^2_2\xi^2 & + & \cdots & + & \alpha^2_n\xi^n & = & \beta^2 \,, \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \,, \\ \alpha^n_1\xi^1 & + & \alpha^n_2 1\xi^2 & + & \cdots & + & \alpha^n_n\xi^n & = & \beta^n \end{array} \right.$$

с учетом соглашений о тензорных обозначениях записывается просто как $\alpha^k_i\xi^i = \beta^k$.

Используя соглашения о тензорных обозначениях и принимая во внимание, что числа σ_i^j и τ_i^j (компоненты матриц прямого и обратного перехода между базисами $\{ g_1, g_2, \ldots, g_n \}$ и $\{ g'_1, g'_2, \ldots, g'_n \}$, то есть коэффициентами в формулах

$$\xi^j = \sum_{i=1}^{n} \sigma_i^j \xi'^i \qquad \text{и} \qquad \xi'^j = \sum_{i=1}^{n} \tau_i^j \xi^i \qquad \forall j = [1, n])$$

являются также компонентами тензоров типа (1,1), сформулируем

`Определение Прил. 4.2.1`	Будем говорить, что в вещественном линейном пространстве Λ^n определен тензор типа (q, p) — q раз контравариантный и p раз ковариантный, если в Λ^n задан объект, который в каждом базисе характеризуется упорядоченным набором из n^{q+p} чисел вида $\xi^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$, где $j_m \quad \forall m = [1, q]$ — контравариантные индексы и $i_k \quad \forall k = [1, p]$ — ковариантные, которые изменяются при переходе от базиса $\{ g_1, g_2, \ldots, g_n \}$ к $\{ g'_1, g'_2, \ldots, g'_n \}$ по закону $$\xi'^{j'_1 j'_2 \ldots j'_q}_{i'_1 i'_2 \ldots i'_p} = \tau^{j'_1}_{j_1} \tau^{j'_2}_{j_2} \cdots \tau^{j'_q}_{j_q} \; \sigma^{i_1}_{i'_1} \sigma^{i_2}_{i'_2} \cdots \sigma^{i_p}_{i'_p} \; \xi^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p} \,.$$

`Определение Прил. 4.2.2`	Число $(q + p)$ называется *валентностью* (или *рангом*) тензора $\xi^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$.

`Определение Прил. 4.2.3`	Два тензора называются *равными*, если они одного и того же типа и во всех базисах имеют равные компоненты (координаты).

`Замечание Прил. 4.2.1.` 1°. Для равенства тензоров одного типа достаточно, чтобы их компоненты были равны лишь в некотором базисе, так как из формул пересчета компонентов следует, что эти тензоры будут иметь равные компоненты и в любом другом базисе.

2°. Если объект характеризуется одним числом, причем не зависящим от выбора базиса, то такой инвариант принято считать тензором типа (0,0).

Таблица Прил. 4.2.1

Тип объекта в Λ^n	**Тип тензора и его координатное представление в базисе** $\{\, g_1,\, g_2,\, \ldots,\, g_n \,\}$	**Тензорное правило изменения координат тензора при переходе к базису** $\{\, g'_1,\, g'_2,\, \ldots,\, g'_n \,\}$
Элемент x	Одновалентный тензор (один раз контравариантный) типа $(1,0)$ ξ^j	$\xi'^j = \tau^j_i \xi^i$
Линейный функционал $f(x)$	Одновалентный тензор (один раз ковариантный) типа $(0,1)$ φ_i	$\phi'_i = \sigma^j_i \varphi_j$
Линейный оператор $\hat{A}$	Двухвалентный тензор (один раз контравариантный и один раз ковариантный) типа $(1,1)$ α^j_i	$\alpha'^k_m = \tau^k_i \sigma^j_m \alpha^i_j$
Билинейный функционал (билинейная форма) $B(x,y)$	Двухвалентный тензор (дважды ковариантный) типа $(0,2)$ β_{ji}	$\beta'_{km} = \sigma^j_k \sigma^i_m \beta_{ji}$
Квадратичный функционал (квадратичная форма) $\Phi(x)$	Двухвалентный тензор (дважды ковариантный) типа $(0,2)$ φ_{ji}	$\Phi_{km} = \sigma^j_k \sigma^i_m \varphi_{ji}$
Символ Кронекера $\delta^j_i = \begin{cases} 1, \text{при } j = i, \\ 0, \text{при } j \neq i \end{cases}$	Двухвалентный тензор (один раз контравариантный и один раз ковариантный) типа $(1,1)$ δ^j_i	$\delta'^k_m = \tau^k_i \sigma^j_m \delta^i_j$

Таблица Прил. 4.2.1 содержит примеры простейших тензорных объектов и правил пересчета их координат при замене базиса.

Отметим, что последний из приведенных в таблице Прил. 4.2.1 тензоров — символ Кронекера — во всех базисах имеет компоненты, совпадающие с компонентами единичной матрицы, если считать, что верхний индекс этого тензора есть номер строки, а нижний – столбца.

Действительно, по определению Прил. 4.2.1 справедливы соотношения

$$\delta'^{k}_{m} = \tau^k_i \sigma^j_m \delta^i_j = \tau^k_i \sigma^i_m = \begin{cases} 1 \text{ при } & m = k, \\ 0 \text{ при } & m \neq k. \end{cases}$$

Последнее равенство имеет место, поскольку выражение $\tau^k_i \sigma^i_m$ есть результат произведения двух невырожденных, взаимно обратных матриц, компоненты которых совпадают с компонентами тензоров τ^k_i и σ^i_m.

Замечания о матричной записи тензоров

В ряде случаев тензоры удобно представлять в виде блочных матриц, то есть матриц, элементами которых являются обычные матрицы с числовыми элементами. При этом используются такие правила:

1°. Тензор типа $(1, 0)$ записывается матрицей-столбцом. Тензор типа $(0, 1)$ записывается матрицей-строкой.

2°. Элементы матриц, используемых для записи тензоров, нумеруются нижними индексами, порядок следования индексов определен выше, в правиле 2° «Запись тензоров». Обратите внимание, что при этом такая запись тензоров валентности большей, чем 1, не будет отражать тип тензора.

3°. Первый индекс определяет номер строки в числовой матрице, второй индекс – номер столбца. Третий индекс определяет номер строки в блочной матрице, состоящей из числовых матриц, четвертый индекс соответственно — номер столбца в блочной матрице.

Приведем для иллюстрации общий вид матричной записи тензора четвертой валентности в двумерном пространстве:

$$\left\| \begin{array}{cc|cc} \alpha_{1111} & \alpha_{1211} & \alpha_{1112} & \alpha_{1212} \\ \alpha_{2111} & \alpha_{2211} & \alpha_{2112} & \alpha_{2212} \\ \hline \alpha_{1121} & \alpha_{1221} & \alpha_{1122} & \alpha_{1222} \\ \alpha_{2121} & \alpha_{2221} & \alpha_{2122} & \alpha_{2222} \end{array} \right\| .$$

Задача Прил.4.2.1 *Каждой паре элементов x и y четырехмерного пространства сопоставляется число $f(x, y)$, определяемое в стандартном базисе $\{ g_1, g_2, g_3, g_4 \}$ через компоненты этих элементов $\xi_1, \xi_2, \xi_3, \xi_4$ и $\eta_1, \eta_2, \eta_3, \eta_4$ по формуле $f(x, y) = \xi_1\eta_3 + \xi_2\eta_4$.*

Показать, что данное сопоставление определяет тензор, найти его тип, выписать его компоненты в данном базисе.

`Решение.` 1°. Очевидно, что данное сопоставление линейно по каждому из аргументов. Рассуждая так же, как при выводе формулы (9.1.1), приходим к заключению, что данное сопоставление полностью и однозначно описывается упорядоченным набором чисел $f(g_j, g_i)\ \forall i, j$.

2°. Найдем закон изменения элементов этого набора при замене базиса. Пусть $g'_i = \sum\limits_{k=1}^{4} \sigma_i^k g_k \quad \forall i = [1,4]$ при переходе от базиса $\{\, g_1,\, g_2,\, g_3,\, g_4 \,\}$ к базису $\{\, g'_1,\, g'_2,\, g'_3,\, g'_4 \,\}$. Тогда, в силу линейности сопоставления, имеем

$$f(g'_i, g'_j) = f\left(\sum_{k=1}^{4} \sigma_i^k g_k, \sum_{l=1}^{4} \sigma_j^l g_l\right) = \sum_{k=1}^{4}\sum_{l=1}^{4} \sigma_i^k \sigma_j^l f(g_k, g_l).$$

Поскольку компоненты исследуемого объекта в новом базисе выражаются линейно через компоненты в старом, а коэффициентами этого выражения служат попарные произведения элементов матрицы перехода $\|S\|$, то по определению Прил. 4.2.1 данный объект является тензором типа $(0, 2)$.

3°. В исходном базисе

$$\left\{ \|g_1\| = \left\| \begin{array}{c} 1 \\ 0 \\ 0 \\ 0 \end{array} \right\|,\ \|g_2\| = \left\| \begin{array}{c} 0 \\ 1 \\ 0 \\ 0 \end{array} \right\|,\ \|g_3\| = \left\| \begin{array}{c} 0 \\ 0 \\ 1 \\ 0 \end{array} \right\|,\ \|g_4\| = \left\| \begin{array}{c} 0 \\ 0 \\ 0 \\ 1 \end{array} \right\| \right\}$$

По условию задачи ненулевыми будут лишь значения $f(g_1, g_3) = 1,\ f(g_2, g_4) = 3$. Таким образом, искомая матрица тензора имеет вид

$$\left\| \begin{array}{cccc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right\| .$$

`Решение получено.`

Приложение 4.3. Операции с тензорами

Вводимые ниже операции с тензорами во всех случаях требуют обоснования того, что результатом каждой из них является также тензор. В рамках данного курса эти утверждения предлагаются в качестве упражнений.

Напомним также, что условия равенства тензоров были сформулированы в определении Прил. 4.2.3.

Сложение тензоров

Определение Прил. 4.3.1

Пусть заданы два тензора типа (q,p) $\alpha^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$ и $\beta^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$.
Тензор типа (q,p) $\gamma^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$ называется *суммой* тензоров $\alpha^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$ и $\beta^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$, если в каждом базисе имеет место равенство

$$\gamma^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p} = \alpha^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p} + \beta^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p} \;.$$

Пример Прил. 4.3.1

Сумма двух линейных операторов в Λ^n α^j_i и β^j_i, являющихся тензорами типа (1,1), есть также линейный оператор γ^j_i и, следовательно, тензор типа (1,1), для компонент (координат) которого в каждом базисе справедливо равенство $\gamma^j_i = \alpha^j_i + \beta^j_i$.
Напомним, что, согласно правилам записи тензоров, последняя формула состоит из n^2 скалярных равенств.

Произведение тензора и числа

Определение Прил. 4.3.2

Пусть задан тензор типа (q,p) $\alpha^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$ и число λ.
Тензор типа (q,p) $\gamma^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$ называется *произведением* тензора $\alpha^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$ и числа λ, если в каждом базисе имеет место равенство

$$\gamma^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p} = \lambda\, \alpha^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p} \;.$$

Замечание Прил. 4.3.1. Нетрудно показать, что множество тензоров типа $(q + p)$ с операциями сложения и умножения на число является линейным пространством размерности n^{q+p} .

Тензорное произведение

Определение Прил. 4.3.3 Пусть заданы два тензора типа (q,p) $\alpha^{j_1j_2\ldots j_q}_{i_1i_2\ldots i_p}$ и типа (r,s) $\beta^{k_1k_2\ldots k_r}_{t_1t_2\ldots t_s}$.
Тензор типа $(q+r, p+s)$ $\gamma^{\,j_1j_2\ldots j_qk_1k_2\ldots k_r}_{\,i_1i_2\ldots i_pt_1t_2\ldots t_s}$ называется *произведением* тензоров $\alpha^{j_1j_2\ldots j_q}_{i_1i_2\ldots i_p}$ и $\beta^{\,k_1k_2\ldots k_r}_{\,t_1t_2\ldots t_s}$, если в каждом базисе имеет место равенство

$$\gamma^{\,j_1j_2\ldots j_qk_1k_2\ldots k_r}_{\,i_1i_2\ldots i_pt_1t_2\ldots t_s} = \alpha^{j_1j_2\ldots j_q}_{i_1i_2\ldots i_p} \cdot \beta^{\,k_1k_2\ldots k_r}_{\,t_1t_2\ldots t_s} \ .$$

Иногда тензорное произведение обозначают символом $\otimes$.

Пример Прил. 4.3.2 Мы видели, что элементы линейного пространства Λ^n являются один раз контравариантными тензорами. Найдем произведение их упорядоченной пары по определению Прил. 4.3.3. Получаем, что $x\otimes y = \xi^k\eta^i$ есть дважды контравариантный тензор.
Заметим, $x\otimes y \neq y\otimes x$. Дело в том, что хотя и $\xi^k\eta^i = \eta^i\xi^k$, но *порядок следования* компонент в записи тензоров *различный*. Следовательно, тензорное произведение *не коммутативно*.

Задача Прил. 4.3.1 *Определить тип и матрицу тензора $c = a\otimes b$, если a — тензор типа $(0,3)$ с матрицей* $\left\| \begin{array}{cc} 1 & 2 \\ 3 & 4 \\ \hline 5 & 6 \\ 7 & 8 \end{array} \right\|$ *и b — тензор типа $(0,1)$ с матрицей* $\| \, 9 \ 10 \, \|$.

Решение. По определению тензорного произведения c есть тензор типа $(0,4)$ с матрицей, составленной (с учетом соглашения о порядке индексов) из поэлементных произведений вида $\alpha_{ijk}\beta_t$, где α_{ijk} и β_t — компоненты тензоров a и b соответственно.

Таким образом, матрица тензора c имеет вид

$$\left\|\begin{array}{cc|cc} 1\cdot 9 & 2\cdot 9 & 1\cdot 10 & 2\cdot 10 \\ 3\cdot 9 & 4\cdot 9 & 3\cdot 10 & 4\cdot 10 \\ \hline 5\cdot 9 & 6\cdot 9 & 5\cdot 10 & 6\cdot 10 \\ 7\cdot 9 & 8\cdot 9 & 7\cdot 10 & 8\cdot 10 \end{array}\right\| = \left\|\begin{array}{cc|cc} 9 & 18 & 10 & 20 \\ 27 & 36 & 30 & 40 \\ \hline 45 & 54 & 50 & 60 \\ 63 & 72 & 70 & 80 \end{array}\right\| .$$

Решение получено.

Свертывание тензоров

Определение Прил. 4.3.4

Пусть дан тензор типа (q,p), причем $q \geq 1$ и $p \geq 1$ $\alpha^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$.

Выберем один верхний (например, j_r) и один нижний (например, i_s) индексы и в записи тензора заменим их обозначения одним и тем же символом (например, m).

Тензор типа $(q-1, p-1)$ $\beta^{j_1 j_2 \ldots j_{q-1}}_{i_1 i_2 \ldots i_{p-1}}$ называется *сверткой тензора* $\alpha^{j_1 j_2 \ldots j_r \ldots j_q}_{i_1 i_2 \ldots i_s \ldots i_p}$ *по индексам* j_r *и* i_s, если в каждом базисе имеет место равенство

$$\beta^{j_1 j_2 \ldots j_{q-1}}_{i_1 i_2 \ldots i_{p-1}} = \alpha^{j_1 j_2 \ldots m \ldots j_q}_{i_1 i_2 \ldots m \ldots i_p} .$$

Заметим, что в последнем равенстве правая часть — это сумма n слагаемых, где m — индекс, по которому выполняется суммирование, а само данное тензорное равенство равносильно $(q-1)\cdot(p-1)$ скалярным равенствам.

Пример Прил. 4.3.3

Свертка тензора типа $(1,1)$, являющегося линейным оператором α^j_i, есть тензор типа $(0,0)$, то есть инвариант относительно замены базиса, имеющий единственный компонент, равный

$$\alpha^m_m = \alpha^1_1 + \alpha^2_2 + \ldots + \alpha^n_n .$$

Данное выражение есть сумма диагональных элементов матрицы линейного оператора, которая не меняется при замене базиса.

Заметим, что данным свойством не обладает, например, матрица билинейного функционала.

Операция свертки часто комбинируется с операцией умножения тензоров.

Например, результатом произведения один раз ковариантного тензора φ_i на один раз контравариантный ξ^j, с последующей сверткой, является инвариант, представляющий значение линейного функционала в Λ^n. Действительно, $f(x) = \varphi_m \xi^m$. В этом случае говорят, что тензор φ_i *свертывается* с тензором ξ^j.

Задача Прил. 4.3.2

Даны тензоры:

– a *типа* $(1,1)$, *с элементами* α_i^j *и матрицей* $\left\| \begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{matrix} \right\|$;

– b *типа* $(1,0)$, *с элементами* β^j *и матрицей* $\| \, 2 \; -3 \;\; 4 \, \|^{\mathrm{T}}$;

– c *типа* $(0,1)$, *с элементами* γ_i *и матрицей* $\| \, 2 \; -3 \;\; 4 \, \|$;

Найти свертки $\alpha_i^j \beta^i$ *и* $\alpha_i^j \gamma_j$.

Решение. 1°. По определению операции свертывания, $\alpha_i^j \beta^i$ — тензор d типа $(1,0)$ с компонентами $\delta^j = \sum\limits_{m=1}^{3} \alpha_m^j \beta^m$. Поэтому

$$\begin{array}{lcl} \delta^1 & = & \alpha_1^1\beta^1 + \alpha_2^1\beta^2 + \alpha_3^1\beta^3 = 1\cdot 2 + 2\cdot(-3) + 3\cdot 4 = 8, \\ \delta^2 & = & \alpha_1^2\beta^1 + \alpha_2^2\beta^2 + \alpha_3^2\beta^3 = 4\cdot 2 + 5\cdot(-3) + 6\cdot 4 = 17, \\ \delta^3 & = & \alpha_1^3\beta^1 + \alpha_2^3\beta^2 + \alpha_3^3\beta^3 = 7\cdot 2 + 8\cdot(-3) + 9\cdot 4 = 26. \end{array}$$

Матрица тензора d есть $\| \, 8 \; 17 \; 26 \, \|^{\mathrm{T}}$.

2°. По определению операции свертывания, $\alpha_i^j \gamma_j$ — тензор f типа $(0,1)$ с компонентами $\varphi_i = \sum\limits_{m=1}^{3} \alpha_i^m \gamma_m$. Поэтому

$$\begin{array}{lcl} \varphi_1 & = & \alpha_1^1\gamma_1 + \alpha_1^2\gamma_2 + \alpha_1^3\gamma_3 = 1\cdot 2 + 4\cdot(-3) + 7\cdot 4 = 18, \\ \varphi_2 & = & \alpha_2^1\gamma_1 + \alpha_2^2\gamma_2 + \alpha_2^3\gamma_3 = 2\cdot 2 + 5\cdot(-3) + 8\cdot 4 = 21, \\ \varphi_3 & = & \alpha_3^1\gamma_1 + \alpha_3^2\gamma_2 + \alpha_3^3\gamma_3 = 3\cdot 2 + 6\cdot(-3) + 9\cdot 4 = 24. \end{array}$$

Решение получено.

Матрица тензора f есть $\| \, 18 \; 21 \; 24 \, \|$.

Транспонирование тензоров

Определение Прил. 4.3.5	Тензор, получаемый из исходного тензора путем перестановки местами некоторой пары ковариантных (или же пары контравариантных) индексов, называется тензором, *транспонированным к исходному по данной паре индексов.*

Транспонирование по любой паре ковариантных (или пары контравариантных) индексов, вообще говоря, приводит к его изменению, поскольку в определении тензора говорится об *упорядоченной* системе индексов. Однако при этом новый тензор будет того же типа, что и исходный.

В общем случае для группы, состоящей из N верхних (или нижних) индексов, существует $N!$ различных способов перестановок. Это означает, что транспонирование по такой группе индексов приводит к построению $N!$ новых тензоров.

Задача Прил. 4.3.3 *Дан тензор типа* $(2,1)$ *с элементами* α_i^{jk} *и матрицей* $\left\| \begin{array}{cc} 1 & 2 \\ 3 & 4 \\ \hline 5 & 6 \\ 7 & 8 \end{array} \right\|$. *Выполнить его транспонирование.*

Решение. Данный тензор можно транспонировать только по паре контравариантных индексов j и k. Поскольку $N = 2$, то вариантов перестановок, включая исходную, всего две. После перестановки соответствующих компонентов получаем тензор с матрицей $\left\| \begin{array}{cc} 1 & 3 \\ 2 & 4 \\ \hline 5 & 7 \\ 6 & 8 \end{array} \right\|$.

В данном случае транспонирование тензора свелось к обычному транспонированию двумерных слоев кубматрицы.

Решение получено.

Симметрирование и альтернирование тензоров

Определение Прил. 4.3.6	Тензор называется *симметричным* относительно группы (верхних или нижних) индексов, если он не меняется при перестановке любых двух индексов, принадлежащих данной группе.

Определение Прил. 4.3.7	Тензор называется *антисимметричным* или *кососимметричным* относительно группы (верхних или нижних) индексов, если он при перестановке любых двух индексов, принадлежащих данной группе, знаки всех координат тензора меняются на противоположные.

Выделим у некоторого тензора группу, состоящую из N индексов (либо верхних, либо нижних), построим путем всевозможных перестановок индексов данной группы $N!$ новых тензоров и возьмем их среднее арифметическое. В результате мы получим тензор, симметричный по выбранной группе индексов.

Данная операция называется *симметрированием тензора по группе индексов*. Группа индексов, по которой выполняется симметрирование тензора, выделяется круглыми скобками.

Пример Прил. 4.3.4

$$\begin{aligned}
N=1 \quad & \xi_{(j_1)} &=& \ \xi_{j_1},\\
N=2 \quad & \xi_{(j_1,j_2)} &=& \ \tfrac{1}{2!}\Big(\xi_{j_1,j_2}+\xi_{j_2,j_1}\Big),\\
N=3 \quad & \xi_{(j_1,j_2,j3)} &=& \ \tfrac{1}{3!}\Big(\xi_{j_1,j_2,j_3}+\xi_{j_3,j_1,j_2}+\xi_{j_2,j_3,j_1}+\\
& & & \quad +\xi_{j_1,j_3,j_2}+\xi_{j_2,j_1,j_3}+\xi_{j_3,j_2,j_1}\Big).
\end{aligned}$$

Операция симметрирования часто комбинируется с умножением, причем имеет место следующий порядок действий: сначала умножение, а потом симметрирование.

Пример Прил. 4.3.5

$$\xi^{(i}\eta^{j)}.$$

Выделим у тензора группу, состоящую из N индексов (либо верхних, либо нижних). Построим путем всевозможных перестановок индексов данной группы $N!$ новых тензоров, приписав каждому из них знак $(-1)^{\text{Б}(k_1,k_2,\ldots,k_N)}$. Здесь $\text{Б}(k_1,k_2,\ldots,k_N)$ — число беспорядков в перестановке чисел $\{k_1,k_2,\ldots,k_N\}$, образованных из $1,2,\ldots,N$. Наконец, возьмем среднее арифметическое этих новых тензоров.

В результате мы получим тензор, антисимметричный по выбранной группе индексов. Данная операция называется *альтернированием* *тензора по группе индексов*. Группа индексов, по которой выполняется альтернирование тензора, выделяется квадратными скобками.

Пример Прил. 4.3.6

$$N=1 \quad \xi_{[j_1]} = \xi_{j_1},$$

$$N=2 \quad \xi_{[j_1,j_2]} = \frac{1}{2!}\Big(\xi_{j_1,j_2} - \xi_{j_2,j_1}\Big),$$

$$N=3 \quad \xi_{[j_1,j_2,j3]} = \frac{1}{3!}\Big(\xi_{j_1,j_2,j_3} + \xi_{j_3,j_1,j_2} + \xi_{j_2,j_3,j_1} - {} - \xi_{j_1,j_3,j_2} - \xi_{j_2,j_1,j_3} - \xi_{j_3,j_2,j_1}\Big).$$

Операция альтернирования нередко комбинируется с умножением, причем имеет место тот же порядок действий, что и для симмертирования: сначала умножение, а потом альтернирование.

Пример Прил. 4.3.7 $\qquad \xi^{[i}\eta^{j]}.$

Отметим, что как симметрирование кососимметричного тензора, так и альтернирование симметричного дают нулевой тензор.

Задача Прил. 4.3.4 $\qquad$ *Тензор* α_{ijk} *имеет матрицу* $\left\|\begin{array}{cc} 1 & 2 \\ 3 & 4 \\ \hline 5 & 6 \\ 7 & 8 \end{array}\right\|$. *Найти матрицы тензоров* $\alpha_{(ij)k}$, $\alpha_{i(jk)}$, $\alpha_{i[jk]}$.

Решение. $\quad 1^\circ$. Тензор $\beta_{ijk} = \alpha_{jik}$, транспонированный к данному по паре индексов i и j, имеет матрицу $\left\|\begin{array}{cc} 1 & 3 \\ 2 & 4 \\ \hline 5 & 7 \\ 6 & 8 \end{array}\right\|$ (см. задачу Прил. 4.3.3).

Тензор $\gamma_{ijk} = \alpha_{ikj}$, транспонированный к исходному по паре индексов j и k, будет иметь матрицу $\left\|\begin{array}{cc} 1 & 5 \\ 3 & 7 \\ \hline 2 & 6 \\ 4 & 8 \end{array}\right\|$, в которой элементы первых столбцов блочных матриц исходного тензора записаны в первой блочной строке, а элементы вторых столбцов — во второй блочной строке.

2°. Тогда тензор $\alpha_{(ij)k} = \frac{1}{2}\Big(\alpha_{ijk} + \beta_{ijk}\Big)$ имеет матрицу

$$\left\|\begin{array}{cc} \frac{1+1}{2} & \frac{2+3}{2} \\ \frac{3+2}{2} & \frac{4+4}{2} \\ \hline \frac{5+5}{2} & \frac{6+7}{2} \\ \frac{7+6}{2} & \frac{8+8}{2} \end{array}\right\| = \left\|\begin{array}{cc} 1 & \frac{5}{2} \\ \frac{5}{2} & 8 \\ \hline 5 & \frac{13}{2} \\ \frac{13}{2} & 8 \end{array}\right\|,$$

тензор $\alpha_{i(jk)} = \frac{1}{2}\Big(\alpha_{ijk} + \gamma_{ijk}\Big)$ — матрицу

$$\left\|\begin{array}{cc} \frac{1+1}{2} & \frac{2+5}{2} \\ \frac{3+3}{2} & \frac{4+7}{2} \\ \hline \frac{5+2}{2} & \frac{6+6}{2} \\ \frac{7+4}{2} & \frac{8+8}{2} \end{array}\right\| = \left\|\begin{array}{cc} 1 & \frac{7}{2} \\ 3 & \frac{11}{2} \\ \hline \frac{2}{2} & 6 \\ \frac{11}{2} & 8 \end{array}\right\|,$$

а тензор $\alpha_{i[jk]} = \frac{1}{2}\Big(\alpha_{ijk} - \gamma_{ijk}\Big)$ — матрицу

$$\left\|\begin{array}{cc} \frac{1-1}{2} & \frac{2-5}{2} \\ \frac{3-3}{2} & \frac{4-7}{2} \\ \hline \frac{5-2}{2} & \frac{6-6}{2} \\ \frac{7-4}{2} & \frac{8-8}{2} \end{array}\right\| = \left\|\begin{array}{cc} 0 & -\frac{3}{2} \\ 0 & -\frac{3}{2} \\ \hline \frac{3}{2} & 0 \\ \frac{3}{2} & 0 \end{array}\right\|.$$

`Решение получено.`

Приложение 4.4. Тензоры в евклидовом пространстве

В случае евклидова пространства тензоры обладают дополнительными свойствами, следующими из того, что скалярное произведение есть билинейный функционал, а потому является дважды ковариантным тензором, компоненты которого в любом базисе совпадают с компонентами матрицы Грама.

Этот дважды ковариантный тензор иногда называют *фундаментальным метрическим тензором.*

Проиллюстрируем эти свойства следующим примером.

Пусть дан базис $\{ g_1, g_2, \ldots, g_n \}$ в E^n и его некоторый элемент x, являющийся одновалентным, один раз контравариантным тензором ξ^j. Свернем фундаментальный метрический тензор $\gamma_{ji} = (g_j, g_i)$ с тензором ξ^j, получим

$$\xi_i = \gamma_{ji}\xi^j = (g_j, g_i)\xi^j = (\xi^j g_j, g_i) = (x, g_i)\,.$$

Данное равенство означает, что элемент x однозначно характеризуется в каждом базисе также и компонентами один раз ковариантного тензора ξ_i, то есть строкой.

Числа ξ_i называются *ковариантными компонентами* (*координатами*) элемента x в базисе $\{ g_1, g_2, \ldots, g_n \}$, и они взаимно однозначно определяются обычными контравариантными компонентами элемента x в силу невырожденности матрицы Грама и системы уравнений $\gamma_{ji}\xi^j = \xi_i$.

Таким образом, в евклидовом пространстве *исчезает принципиальная разница* между ковариантными и контравариантными индексами тензоров. Более того, в ортонормированном базисе ковариантные и контравариантные компоненты элемента просто совпадают (см. теорему 10.3.2).

Операция опускания индекса

Определение Прил. 4.4.1	Пусть в E^n задан тензор типа (q, p) вида $\alpha^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$, где $q \geq 1$. Тензор типа $(q-1, p+1)$ вида $\beta^{\bullet\, j_2 j_3 \ldots j_q}_{i_0 i_1 i_2 \ldots i_p}$ называется результатом операции *опускания* контравариантного индекса j_1 у тензора $\alpha^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$, если в каждом базисе имеет место равенство $$\beta^{\bullet\, j_2 j_3 \ldots j_q}_{i_0 i_1 i_2 \ldots i_p} = \gamma_{i_0 j_1} \alpha^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}\,.$$

Заметим, что использование точек для указания порядка следования индексов в этой операции оказывается необходимым, чтобы сделать ее однозначной. Иначе непонятно, куда следует опустить индекс.

Операция поднятия индекса

Определение Прил. 4.4.2	Дважды контравариантный тензор, компоненты которого в любом базисе евклидова пространства E^n совпадают с матрицей, *обратной* матрице Грама, называется *контравариантным метрическим тензором.*

Убедимся вначале, что матрица, обратная матрице Грама, задает в каждом базисе тензор типа $(2,0)$. Имеем $\|\Gamma\|_{g'} = \|S\|^{\mathrm{T}}\|\Gamma\|_g\|S\|$.

Исходя из этого соотношения получаем следующее правило преобразования обратной матрицы Грама при замене базиса:

$$\|\Gamma\|_{g'}^{-1} =$$

$$= \left(\|S\|^{\mathrm{T}}\|\Gamma\|_g\|S\|\right)^{-1} = \|S\|^{-1}\|\Gamma\|_g^{-1}\left(\|S\|^{\mathrm{T}}\right)^{-1} = \|S\|^{-1}\|\Gamma\|_g^{-1}\left(\|S\|^{-1}\right)^{\mathrm{T}},$$

поскольку (проверьте это самостоятельно) для невырожденной матрицы $\|S\|$ справедливо равенство $\left(\|S\|^{\mathrm{T}}\right)^{-1} = \left(\|S\|^{-1}\right)^{\mathrm{T}}$. А это и означает, что обратная матрица Грама определяет во всех базисах дважды контравариантный тензор γ^{ij}.

По аналогии с операцией опускания индекса дадим

Определение Прил. 4.4.3	Пусть в E^n задан тензор типа (q,p) вида $\alpha^{j_1j_2\ldots j_q}_{i_1i_2\ldots i_p}$, где $p \geq 1$. Тензор типа $(q+1,p-1)$ вида $\beta^{\,j_0j_1j_2\ldots j_q}_{\bullet\, i_2\, i_3\ldots i_p}$ называется результатом операции *подъема* ковариантного индекса i_1 у тензора $\alpha^{j_1j_2\ldots j_q}_{i_1i_2\ldots i_p}$, если в каждом базисе имеет место равенство $$\beta^{\,j_0j_1j_2\ldots j_q}_{\bullet\, i_2\, i_3\ldots i_p} = \gamma^{j_0i_1}\alpha^{j_1j_2\ldots j_q}_{i_1i_2\ldots i_p}\,.$$

Задача Прил. 4.4.1 *В E^2 с фундаментальным метрическим тензором вида* $\gamma_{ij} = \begin{Vmatrix} 2 & 3 \\ 3 & 5 \end{Vmatrix}$ *тензор $\alpha^{i}_{\bullet jk}$ задан матрицей* $\begin{Vmatrix} 3 & 4 \\ 5 & 7 \\ \hline 2 & 5 \\ 1 & 3 \end{Vmatrix}$.

Найти матрицы тензоров α_{ijk} и $\alpha^{ij\bullet}_{\ \ \ k}$.

Решение. 1°. Для опускания первого индекса воспользуемся формулой $\alpha_{ijk} = \gamma_{im}\alpha^{m}_{\bullet jk}$. Получаем

$$\begin{array}{lclclcr}
\alpha_{111} & = & \gamma_{11}\alpha^1_{11} + \gamma_{12}\alpha^2_{11} & = & 2\cdot 3 + 3\cdot 5 & = & 21\,, \\
\alpha_{112} & = & \gamma_{11}\alpha^1_{12} + \gamma_{12}\alpha^2_{12} & = & 2\cdot 2 + 3\cdot 1 & = & 7\,, \\
\alpha_{121} & = & \gamma_{11}\alpha^1_{21} + \gamma_{12}\alpha^2_{21} & = & 2\cdot 4 + 3\cdot 7 & = & 29\,, \\
\alpha_{122} & = & \gamma_{11}\alpha^1_{22} + \gamma_{12}\alpha^2_{22} & = & 2\cdot 5 + 3\cdot 3 & = & 19\,, \\
\alpha_{211} & = & \gamma_{21}\alpha^1_{11} + \gamma_{22}\alpha^2_{11} & = & 3\cdot 3 + 5\cdot 5 & = & 34\,, \\
\alpha_{212} & = & \gamma_{21}\alpha^1_{12} + \gamma_{22}\alpha^2_{12} & = & 3\cdot 2 + 5\cdot 1 & = & 11\,, \\
\alpha_{221} & = & \gamma_{21}\alpha^1_{21} + \gamma_{22}\alpha^2_{21} & = & 3\cdot 4 + 5\cdot 7 & = & 47\,, \\
\alpha_{222} & = & \gamma_{21}\alpha^1_{22} + \gamma_{22}\alpha^2_{22} & = & 3\cdot 5 + 5\cdot 3 & = & 30\,.
\end{array}$$

Следовательно, матрица тензора α_{ijk} имеет вид

$$\left\| \begin{array}{rr} 21 & 29 \\ 34 & 47 \\ \hline 7 & 19 \\ 11 & 30 \end{array} \right\| .$$

2°. Для поднятия второго нижнего индекса следует применить формулу $\alpha^{ij}_{\bullet\bullet k} = \alpha^{i}_{\bullet mk}\gamma^{mj}$, где γ^{ij} — контравариантный метрический тензор, матрица которого обратна матрице тензора γ_{ij} и имеет вид

$$\gamma^{ij} = \left\| \begin{array}{cc} 2 & 3 \\ 3 & 5 \end{array} \right\|^{-1} = \left\| \begin{array}{rr} 5 & -3 \\ -3 & 2 \end{array} \right\|, \text{ тогда}$$

$$\begin{array}{lclcrclcr}
\alpha^{11}_{\bullet\bullet 1} & = & \alpha^1_{\bullet 11}\gamma^{11} + \alpha^1_{\bullet 21}\gamma^{21} & = & 3\cdot 5 + 4\cdot(-3) & = & 3\,, \\
\alpha^{12}_{\bullet\bullet 1} & = & \alpha^1_{\bullet 11}\gamma^{12} + \alpha^1_{\bullet 21}\gamma^{22} & = & -3\cdot 3 + \quad 4\cdot 2 & = & -1\,, \\
\alpha^{21}_{\bullet\bullet 1} & = & \alpha^2_{\bullet 11}\gamma^{11} + \alpha^2_{\bullet 21}\gamma^{21} & = & 5\cdot 5 + 7\cdot(-3) & = & 4\,, \\
\alpha^{12}_{\bullet\bullet 1} & = & \alpha^2_{\bullet 11}\gamma^{12} + \alpha^2_{\bullet 21}\gamma^{22} & = & 5\cdot(-3) + \quad 7\cdot 2 & = & -1\,, \\
\alpha^{11}_{\bullet\bullet 2} & = & \alpha^1_{\bullet 12}\gamma^{11} + \alpha^1_{\bullet 22}\gamma^{21} & = & 2\cdot 5 + (-3)\cdot 5 & = & -5\,, \\
\alpha^{12}_{\bullet\bullet 2} & = & \alpha^1_{\bullet 12}\gamma^{12} + \alpha^1_{\bullet 22}\gamma^{22} & = & 2\cdot(-3) + \quad 5\cdot 2 & = & 4\,, \\
\alpha^{21}_{\bullet\bullet 2} & = & \alpha^2_{\bullet 12}\gamma^{11} + \alpha^2_{\bullet 22}\gamma^{21} & = & 1\cdot 5 + 3\cdot(-3) & = & -4\,, \\
\alpha^{12}_{\bullet\bullet 2} & = & \alpha^2_{\bullet 12}\gamma^{12} + \alpha^2_{\bullet 22}\gamma^{22} & = & 1\cdot(-3) + \quad 3\cdot 2 & = & 3\,.
\end{array}$$

Таким образом, тензор $\alpha^{ij}_{\bullet\bullet k}$ имеет матрицу

$$\left\| \begin{array}{rr} 3 & -1 \\ 4 & -1 \\ \hline -5 & 4 \\ -4 & 3 \end{array} \right\| .$$

Решение получено.

В ортонормированном базисе очевидно, что $\gamma_{ij} = \gamma^{ij} = \delta_i^j$, то есть между ковариантными и контравариантными индексами нет никакой разницы, что также следует из равенства $\|T\| = \|S\|^{-1} = \|S\|^{\mathrm{T}}$, верного для ортонормированных базисов, и из определения тензоров.

Приложение 4.5. Тензоры в ортонормированном базисе

Совпадение ковариантных и контравариантных индексов в ортонормированных базисах евклидова пространства позволяет ввести в рассмотрение упрощенный класс тензоров, определенных только в таких базисах и называемых *евклидовыми тензорами.*

Два евклидовых тензора считаются одинаковыми, если один из них может быть преобразован во второй операциями опускания или поднятия индексов. Поэтому можно в дальнейшем рассматривать евклидовы тензоры как имеющие лишь нижние индексы. При этом, правда, придется допустить суммирование по паре совпадающих ковариантных индексов.

Евклидовы тензоры позволяют наглядно продемонстрировать связь методов тензорного исчисления и аппарата векторной алгебры в обычном трехмерном векторном пространстве E^3.

Введем предварительно в рассмотрение трехвалентный, так называемый *дискриминантный* тензор ε_{ijk}, определяемый во всех ортонормированных базисах по правилу:

$$\varepsilon_{ijk} = \begin{cases} (-1)^{\text{Б}(i,j,k)}, & \text{если среди индексов } i,j,k \text{ нет равных,} \\ \qquad 0, & \text{в остальных случаях.} \end{cases}$$

Здесь Б(i,j,k), как и раньше (см. § 6.1), обозначает число беспорядков в перестановке чисел $\{i,j,k\}$. Всего у тензора ε_{ijk}, антисимметричного по любой паре индексов, 27 компонентов, из которых только шесть ненулевых: три, равные 1, и три, равные -1.

Убедимся вначале, что объект ε_{ijk} преобразуется при переходе от одного ортонормированного базиса в E^3 к другому как трижды ковариантный тензор.

Для этого запишем выражения компонентов объекта в новом базисе через его компоненты в старом, использовав правила преобразования трижды ковариантного тензора. Имеем $\varepsilon'_{lmn} = \varepsilon_{ijk}\sigma_{li}\sigma_{mj}\sigma_{nk}$.

Это выражение распишем в явном виде, согласно соглашению о суммировании, учтя значения координат дискриминантного тензора в исходном базисе:

$$\varepsilon'_{lmn} = \sigma_{l1}\sigma_{m2}\sigma_{n3} + \sigma_{l3}\sigma_{m1}\sigma_{n2} + \sigma_{l2}\sigma_{m3}\sigma_{n1} -$$

$$-\sigma_{l2}\sigma_{m1}\sigma_{n3} - \sigma_{l3}\sigma_{m2}\sigma_{n1} - \sigma_{l1}\sigma_{m3}\sigma_{n2} =$$

$$= \det \begin{Vmatrix} \sigma_{l1} & \sigma_{l2} & \sigma_{l3} \\ \sigma_{m1} & \sigma_{m2} & \sigma_{m3} \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} \end{Vmatrix},$$

что, в свою очередь, по свойствам определителя дает

$$\varepsilon'_{lmn} = \begin{cases} (-1)^{\text{Б}(l,m,n)}, & \text{если среди индексов } l, m, n \text{ нет равных}, \\ 0, & \text{в остальных случаях}, \end{cases}$$

поскольку матрица $\|S\| = \begin{Vmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{Vmatrix}$ ортогональная (как матрица перехода от одного ортонормированного базиса к другому) и ее определитель равен ± 1.

Но если объект ε_{ijk} в новом произвольном ортонормированном базисе имеет (при использовании тензорных правил преобразования) те же компоненты, что и в исходном, то мы приходим к заключению (по определению тензора), что это трехвалентный евклидов тензор.

Тензоры и произведения векторов

Покажем теперь связь тензорного произведения элементов пространства E^3 и произведений векторов, ранее введенных в данном пособии (см. § 2.2 и § 2.4). Все базисы по-прежнему ортонормированные.

Рассмотрим два одновалентных ковариантных тензора ξ_i и η_k, которые в аналитической геометрии (что было показано ранее) интерпретируются как обычные геометрические векторы $\vec{a}$ и $\vec{b}$.

Их тензорное произведение $\xi_i\,\eta_k$ есть дважды ковариантный евклидов тензор, имеющий 9 компонентов, которые можно записать в виде матрицы:

$$\begin{Vmatrix} \xi_1\eta_1 & \xi_1\eta_2 & \xi_1\eta_3 \\ \xi_2\eta_1 & \xi_2\eta_2 & \xi_2\eta_3 \\ \xi_3\eta_1 & \xi_3\eta_2 & \xi_3\eta_3 \end{Vmatrix}.$$

Согласно правилам сложения тензоров и умножения их на число, данный тензор можно представить как сумму симметричного и антисимметричного тензоров:

$$\xi_i\,\eta_k = \frac{1}{2}\Big(\xi_i\,\eta_k + \xi_k\,\eta_i\Big) + \frac{1}{2}\Big(\xi_i\,\eta_k - \xi_k\,\eta_i\Big)\,,$$

или в матричном виде:

$$\left\| \begin{array}{ccc} \xi_1\eta_1 & \xi_1\eta_2 & \xi_1\eta_3 \\ \xi_2\eta_1 & \xi_2\eta_2 & \xi_2\eta_3 \\ \xi_3\eta_1 & \xi_3\eta_2 & \xi_3\eta_3 \end{array} \right\| =$$

$$= \frac{1}{2}\left\| \begin{array}{ccc} \xi_1\eta_1 + \xi_1\eta_1 & \xi_1\eta_2 + \xi_2\eta_1 & \xi_1\eta_3 + \xi_3\eta_1 \\ \xi_2\eta_1 + \xi_1\eta_2 & \xi_2\eta_2 + \xi_2\eta_2 & \xi_2\eta_3 + \xi_3\eta_2 \\ \xi_3\eta_1 + \xi_1\eta_3 & \xi_3\eta_2 + \xi_2\eta_3 & \xi_3\eta_3 + \xi_3\eta_3 \end{array} \right\| +$$

$$+ \frac{1}{2}\left\| \begin{array}{ccc} 0 & \xi_1\eta_2 - \xi_2\eta_1 & \xi_1\eta_3 - \xi_3\eta_1 \\ \xi_2\eta_1 - \xi_1\eta_2 & 0 & \xi_2\eta_3 - \xi_3\eta_2 \\ \xi_3\eta_1 - \xi_1\eta_3 & \xi_3\eta_2 - \xi_2\eta_3 & 0 \end{array} \right\| .$$

Рассмотрим каждое слагаемое в правой части по отдельности.

Во-первых, отметим, что из симметричности матричного представления для первого слагаемого следует существование ортонормированного базиса, в котором эта матрица диагональна.

Теперь покажем, что свертка этого слагаемого есть инвариант, то есть она не зависит от выбора базиса.

Действительно, учитывая, что ξ_i и η_k суть одновалентные ковариантные тензоры, и используя свойства матрицы перехода между двумя ортонормированными базисами $\|S\|$, получаем следующее правило преобразования их свертки:

$$\xi'_k\eta'_k = \sigma_{ki}\xi_i\sigma_{kj}\eta_j = \sigma^{\mathrm{T}}_{ik}\sigma_{kj}\xi_i\eta_j = \delta_{ij}\xi_i\eta_j = \xi_i\eta_i\,,$$

что и означает инвариантность этой свертки относительно замены базиса.

Откуда следует вывод: любой паре элементов (векторов) $\vec{a}$ и $\vec{b}$, имеющих соответственно компоненты ξ_i и η_k в E^3, можно поставить в соответствие не зависящее от выбора ортонормированного базиса число $\quad \xi_i\eta_i = \xi_1\eta_1 + \xi_2\eta_2 + \xi_3\eta_3 \quad$ (см. также § 2.9).

Выясним геометрический смысл этого инварианта, обозначаемого $(\vec{a},\vec{b})$.

Если среди векторов $\vec{a}$ и $\vec{b}$ есть нулевой, то значение инварианта, очевидно, равно нулю. Для любых же ненулевых векторов $\vec{a}$ и $\vec{b}$ всегда найдется ортонормированный базис, в котором их координатные представления соответственно имеют вид

$$\left\|\begin{array}{c} |\vec{a}| \\ 0 \\ 0 \end{array}\right\| \quad \text{и} \quad \left\|\begin{array}{c} |\vec{b}|\cos\varphi \\ |\vec{b}|\sin\varphi \\ 0 \end{array}\right\| ,$$

где φ — угол между $\vec{a}$ и $\vec{b}$. Тогда значение инварианта будет равно $(\vec{a},\vec{b}) = |\vec{a}||\vec{b}|\cos\varphi$, и мы приходим к формуле *скалярного произведения* векторов, которая обычно принимается за его определение.

Рассмотрим теперь второе слагаемое. Как нетрудно видеть, матрица

$$\left\|\begin{array}{ccc} 0 & \xi_1\eta_2-\xi_2\eta_1 & \xi_1\eta_3-\xi_3\eta_1 \\ \xi_2\eta_1-\xi_1\eta_2 & 0 & \xi_2\eta_3-\xi_3\eta_2 \\ \xi_3\eta_1-\xi_1\eta_3 & \xi_3\eta_2-\xi_2\eta_3 & 0 \end{array}\right\|$$

имеет только три независимых элемента, из чего следует, что паре векторов $\vec{a}$ и $\vec{b}$ в E^3 может быть поставлен в однозначное соответствие третий вектор, обозначаемый как $[\vec{a},\vec{b}]$, с координатным представлением

$$\left\|\begin{array}{c} \xi_2\eta_3-\xi_3\eta_2 \\ \xi_3\eta_1-\xi_1\eta_3 \\ \xi_1\eta_2-\xi_2\eta_1 \end{array}\right\| .$$

Исследуем его свойства.

Во-первых, заметим, что число независимых компонентов у кососимметричной части тензорного произведения элементов в случае пространства размерности n равно $\dfrac{(n-1)n}{2}$, поскольку это есть число компонентов, стоящих в матрице над ее главной диагональю. Поэтому только в E^3 это число совпадает с размерностью пространства, и только в E^3 произведению двух элементов можно подобным образом ставить в однозначное соответствие третий элемент.

Во-вторых, убедимся, что имеют место равенства $[\vec{a},\vec{b}]_i = \varepsilon_{ijk}\xi_j\eta_k$. Действительно, например, для $i=1$:

$$\begin{aligned} \varepsilon_{1jk}\xi_j\eta_k \quad = \quad & \varepsilon_{111}\xi_1\eta_1+\varepsilon_{112}\xi_1\eta_2+\varepsilon_{113}\xi_1\eta_3+ \\ & +\varepsilon_{121}\xi_2\eta_1+\varepsilon_{122}\xi_2\eta_2+\varepsilon_{123}\xi_2\eta_3+ \\ & +\varepsilon_{131}\xi_3\eta_1+\varepsilon_{132}\xi_3\eta_2+\varepsilon_{133}\xi_3\eta_3= \\ & =\xi_2\eta_3-\xi_3\eta_2\,. \end{aligned}$$

В-третьих, покажем инвариантность тензора $\kappa_i = \varepsilon_{ijk}\xi_j\eta_k$ при переходе от одного ортонормированного базиса к другому в E^3.

Пусть соотношение $\kappa'_i = \varepsilon'_{ijk}\xi'_j\eta'_k$ верно в новом ортонормированном базисе, тогда в исходном базисе будут справедливы равенства $\sigma_{is}\kappa_s = \varepsilon'_{ijk}\sigma_{jm}\xi_m\sigma_{kl}\eta_l$.

Умножив обе части последнего равенства на тензор σ_{qi} и свернув произведения по индексу i, получим

$$\sigma_{qi}\sigma_{is}\kappa_s = \sigma_{qi}\sigma_{jm}\sigma_{kl}\,\varepsilon'_{ijk}\xi_m\eta_l\ ,$$

но $\sigma_{qi}\sigma_{is}\kappa_s = \delta_{qs}\kappa_s = \kappa_q$, а $\sigma_{qi}\sigma_{jm}\sigma_{kl}\,\varepsilon'_{ijk} = \varepsilon_{qml}$, поскольку тензор ε_{ijk} инвариантен при переходе от одного ортонормированного базиса к другому.

Следовательно, $\kappa_i = \varepsilon_{iml}\xi_m\eta_l$, что и означает инвариантность этого элемента относительно замены базиса.

Выясним, наконец, геометрический смысл вектора $[\vec{a},\vec{b}]$. Сначала самостоятельно убедитесь, что в случае коллинеарных $\vec{a}$ и $\vec{b}$ тензор $\kappa_i = \varepsilon_{iml}\xi_m\eta_l$ (типа $(0,1)$) нулевой.

Заметим далее, что для любых неколлинеарных векторов $\vec{a}$ и $\vec{b}$ можно выбрать ортонормированный базис в E^3, в котором их координатные представления имеют вид соответственно

$$\left\|\begin{matrix} 0 \\ |\vec{a}| \\ 0 \end{matrix}\right\| \quad \text{и} \quad \left\|\begin{matrix} 0 \\ |\vec{b}|\cos\varphi \\ |\vec{b}|\sin\varphi \end{matrix}\right\| ,$$

где φ — угол между $\vec{a}$ и $\vec{b}$.

Тогда значение первого компонента $[\vec{a},\vec{b}]$ есть $|\vec{a}||\vec{b}|\sin\varphi$, в то время как остальные компоненты нулевые, и получилась формула векторного произведения, принимаемая обычно за его определение.

Таким образом, можно заключить, что введенные в курсе векторной алгебры операции скалярного и векторного произведений базируются не только на «их полезности для приложений», но и отражают инвариантные свойства тензорного произведения элементов евклидова пространства при переходах между ортонормированными базисами.

В заключение покажем, что тензоры могут быть эффективно использованы и в более сложных конструкциях векторной алгебры. Например:

1°. Смешанное произведение трех векторов (см. § 2.6) представимо в виде

$$(\vec{a},\vec{b},\vec{c}) = ([\vec{a},\vec{b}],\vec{c}) = [\vec{a},\vec{b}]_k\kappa_k = \varepsilon_{ijk}\xi_i\eta_j\kappa_k.$$

2°. Выражение для двойного векторного произведения трех векторов (см. § 2.8) может быть получено следующим образом:

$$[\vec{a},[\vec{b},\vec{c}]]_i =$$

$$= \varepsilon_{ijk}\xi_j[\vec{b},\vec{c}]_k = \varepsilon_{ijk}\xi_j\varepsilon_{klm}\eta_l\kappa_m = \varepsilon_{ijk}\varepsilon_{klm}\xi_j\eta_l\kappa_m\,.$$

Принимая во внимание, что $\varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$ (проверьте это самостоятельно), приходим к равенству

$$[\vec{a},[\vec{b},\vec{c}]]_i = \big(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}\big)\xi_j\eta_l\kappa_m =$$

$$= \eta_i\xi_m\kappa_m - \kappa_i\xi_j\eta_j = \eta_i(\vec{a},\vec{c}) - \kappa_i(\vec{a},\vec{b}),$$

или окончательно $\quad [\vec{a},[\vec{b},\vec{c}]] = \vec{b}(\vec{a},\vec{c}) - \vec{c}(\vec{a},\vec{b})\,.$

Список литературы

1. *Беклемишев Д.В.* Курс аналитической геометрии и линейной алгебры. 10-е изд., испр. Москва : Физматлит, 2005.
2. *Чехлов В.И.* Лекции по аналитической геометрии и линейной алгебре. Москва : МФТИ, 2005.
3. *Мальцев А.М.* Основы линейной алгебры. Москва : Наука, 1976.
4. *Постников М.М.* Лекции по геометрии. Москва : Наука, 1979.
5. *Беклемишев Д.В.* Дополнительные главы линейной алгебры. Москва : Наука, 1983.
6. *Шилов Г.Е.* Введение в теорию линейных пространств. Москва : Наука, 1976.
7. *Волков Т.Ф.* Тензоры и векторы : учебное пособие. Москва : МФТИ, 1976.
8. *Беклемишева Л.А., Петрович А.Ю., Чубаров И.А.* Сборник задач по аналитической геометрии и линейной алгебре. Москва : Физматлит, 2001.

Предметный указатель

Г

Д

Е

З

И

К

Л

М

Н

О

П

Р

С

Т

У

Ф

Х

Ц

Ч

Э

Я

Учебное издание

Умнов Александр Евгеньевич

Умнов Егор Александрович

АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ И ЛИНЕЙНАЯ АЛГЕБРА

Издание четвертое, исправленное и дополненное

Редактор *И. А. Волкова*
Корректор *О. П. Котова*
Дизайн обложки *Е. А. Казённовой*

Подписано в печать xx.xx.2025. Формат 60×84 $^1/_{16}$.
Усл. печ. л. 28,75. Уч.-изд. л. 26,45. Тираж 50 экз. Заказ №

Федеральное государственное автономное образовательное учреждение высшего образования «Московский физико-технический институт (национальный исследовательский университет)»
141700, Московская обл., г. Долгопрудный, Институтский пер., 9
Тел. (495) 408-58-22, e-mail: rio@mipt.ru

Отпечатано в полном соответствии с предоставленным оригиналом-макетом
ООО «Печатный салон ШАНС», г. Москва, ул. Ижорская, д. 13, стр. 2
Тел. (495) 484-26-55