МИНИСТЕРСТВО ОБЩЕГО И ПРОФЕССИОНАЛЬНОГО
ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

БАШКИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ШАРИПОВ Р. А.

КУРС ЛИНЕЙНОЙ АЛГЕБРЫ И МНОГОМЕРНОЙ ГЕОМЕТРИИ

УФА 1996

УДК 517.9

Шарипов Р. А. **Курс линейной алгебры и многомерной геометрии**: учебное пособие для вузов / Изд-е Башкирского ун-та. — Уфа, 1996. — 146 с. ISBN 5-7477-0099-5

Книга рассчитана как учебное пособие по основному курсу многомерной геометрии и линейной алгебры. На математическом факультете Башкирского Государственного университета этот предмет изучается на первом курсе во втором семестре. Он входит в программу базового математического образования для физико-математических факультетов и изучается во всех университетах России.

Подготовка книги к изданию выполнена методом компьютерной верстки на базе пакета $\mathcal{AMS}$-TEX от Американского Математического Общества. При этом были использованы кириллические шрифты семейства Lh, распространяемые Ассоциацией *CyrTUG* пользователей кириллического TEX'а.

Рецензенты: Кафедра Вычислительной Математики и Кибернетики УГАТУ,
д. ф.-м. н., проф. Пинчук С.И. (Челябинский Государственный Технологический Университет и Индианский Университет, США).

Контактная информация для связи с автором.

Место работы:	Математический факультет, Башкирский Государственный Университет, ул. Фрунзе 32, Уфа 450074, Россия
Тел.:	7-(3472)-23-67-18
Факс:	7-(3472)-23-67-74
Домашний адрес:	ул. Рабочая 5, Уфа 450003, Россия
Тел.:	7-(917)-75-55-786
E-mail:	R_Sharipov@ic.bashedu.ru, r-sharipov@mail.ru, ra_sharipov@lycos.com, ra_sharipov@hotmail.com
URL:	http://www.geocities.com/r-sharipov

ISBN 5-7477-0099-5

ОГЛАВЛЕНИЕ.

ПРЕДИСЛОВИЕ

Существует два подхода к изложению линейной алгебры и многомерной геометрии. Первый можно охарактеризовать как «координатно-матричный подход», второй — «инвариантно-геометрический подход».

В большинстве учебников используется координатно-матричный подход. Он начинается с рассмотрения систем линейных уравнений. Затем развивается теория детерминантов, матричная алгебра и геометрия пространства $\mathbb{R}^n$. Этот подход удобен для первоначального знакомства с предметом. В основе его лежат простые понятия — числа, наборы чисел, матрицы с числовыми элементами, линейные функции и линейные уравнения. Доказательства в идейном плане просты и носят, по существу, вычислительный характер. Однако, в определенный момент простота координатно-матричного подхода перестает быть преимуществом. Вычислительный характер доказательств делает их громоздкими, а желание ограничиться числовыми объектами препятствует введению и использованию новых понятий.

Инвариантно-геометрический подход, которого мы придерживаемся в данной книге, стартует с определения абстрактного линейного векторного пространства. При этом координатное представление векторов перестает играть первостепенную роль. На первый план выходят теоретико-множественные методы, принятые в современной алгебре и геометрии. Линейные векторные пространства оказываются тем объектом, где эти методы проявляются наиболее просто и эффективно. Доказательство многих фактов удается сделать более коротким и изящным.

Принятый в книге инвариантно-геометрический подход к изложению материала позволяет подготовить читателя к изучению более продвинутых разделов математики, таких, как дифференциальная геометрия, коммутативная алгебра, алгебраическая геометрия и алгебраическая топология. Изложение материала в книге является замкнутым. От читателя требуются лишь некоторые минимальные знания из матричной алгебры и теории детерминантов. Эти вопросы обычно излагаются в курсах общей алгебры и аналитической геометрии.

Под числовым полем в этой книге мы понимаем одно из трех полей: поле рациональных чисел $\mathbb{Q}$, поле вещественных чисел $\mathbb{R}$ и поле комплексных чисел $\mathbb{C}$. Поэтому знакомства с общей теорией числовых полей не требуется.

Автор благодарен Руденко Е. Б. за прочтение и редактирование рукописи книги.

Май, 1996 г. Р.А. Шарипов.

ГЛАВА I

ЛИНЕЙНЫЕ ВЕКТОРНЫЕ ПРОСТРАНСТВА И ЛИНЕЙНЫЕ ОТОБРАЖЕНИЯ.

§ 1. Множества и отображения.

Понятие *множества* является базовым понятием современной математики. Им принято обозначать всякую совокупность объектов, которые по каким-либо причинам необходимо выделить из числа остальных и объединить друг с другом. Объекты, составляющие данное множество, называются *элементами* этого множества. Множествам и их элементам обычно присваивают буквенные имена (идентификаторы). Пусть множество A состоит из трех объектов m, n и q. Это записывают следующим образом:

$$A = \{m,\ n,\ q\}.$$

Тот факт, что m является элементом множества A, обозначают посредством значка принадлежности $m \in A$. Запись $p \notin A$ означает, что объект p не является элементом множества A.

Имея некоторое количество множеств, можно собрать все их элементы в одно множество, которое называется *объединением* исходных множеств. Для обозначения операции объединения используется значок $\cup$. Выделив общую часть всех рассматриваемых множеств, мы получим новое множество, называемое *пересечением* этих множеств. Для обозначения операции пересечения используется значок $\cap$.

Если одно из множеств A является частью другого множества B, то это обозначается так: $A \subset B$ или $A \subseteq B$. При этом говорят, что A есть *подмножество* множества B. Значки $\subset$ и $\subseteq$ совершенно равнозначны. Просто, при втором варианте записи мы подчеркиваем, что условие $A \subset B$ не исключает совпадения $A = B$. Если $A \subsetneq B$, то говорят, что A есть *собственное подмножество* в множестве B.

Термином *пустое множество* обозначается множество $\varnothing$, не содержащее ни одного элемента. Пустое множество считается частью любого множества $\varnothing \subset A$.

Определение 1.1. Отображением $f\colon X \to Y$ из множества X в множество Y называется правило f, которое по каждому заданному $x \in X$ однозначно определяет некоторый элемент $y = f(x)$ из множества Y.

Множество X в определении 1.1 принято называть *областью определения* отображения f. Множество Y называется *областью значений* отображения f. Запись $f(x)$ означает применение правила f к элементу x из множества X. Элемент $y = f(x)$, полученный в результате применения f к элементу x, называют образом элемента x при отображении f.

Пусть $A \subset B$ — некоторое подмножество в X. Множество $f(A)$, составленное из образов всех элементов $x \in A$, называют *образом множества* A при отображении f. Можно дать формальное определение $f(A)$:

$$f(A) = \{y \in Y : \exists\, x\ ((x \in A)\ \ \&\ \ (f(x) = y))\}.$$

Если $A = X$, то образ $f(X)$ принято называть *образом отображения* f. Для него имеется специальное обозначение $f(X) = \operatorname{Im} f$. Образ отображения $f(X) = \operatorname{Im} f$ часто называют *множеством значений*. Не следует путать множество значений и область значений, это разные понятия.

Пусть y — некоторый элемент в Y. Рассмотрим множество $f^{-1}(y)$, состоящее из всех элементов $x \in X$, которые отображаются в y. Оно называется *полным прообразом* элемента y при отображении f:

$$f^{-1}(y) = \{x \in X : f(x) = y\}.$$

Пусть $B \subset Y$ — некоторое подмножество в Y. Объединив полные прообразы всех элементов множества B, мы получим *полный прообраз* этого множества:

$$f^{-1}(B) = \{x \in X : f(x) \in B\}.$$

Нетрудно видеть, что для $B = Y$ полный прообраз $f^{-1}(Y)$ совпадает с X. Поэтому никакого специального обозначения для $f^{-1}(Y)$ не вводят.

Определение 1.2. Отображение $f : X \to Y$ называется *инъективным*, если образы любых двух различных элементов $x_1 \neq x_2$ различны, т. е. $x_1 \neq x_2$ влечет $f(x_1) \neq f(x_2)$.

Определение 1.3. Отображение $f : X \to Y$ называется *сюръективным*, если полный прообраз $f^{-1}(y)$ любого элемента $y \in Y$ непуст.

Определение 1.4. Отображение $f : X \to Y$ называется *биективным* или *взаимно-однозначным*, если полный прообраз $f^{-1}(y)$ любого элемента $y \in Y$ состоит равно из одного элемента множества X.

Теорема 1.1. *Отображение $f : X \to Y$ биективно тогда и только тогда, когда оно инъективно и сюръективно одновременно.*

Док-во. Согласно утверждению теоремы 1.1 одновременная инъективность и сюръективность является необходимым и достаточным условием биективности отображения $f : X \to Y$. Начнем с доказательства необходимости.

Пусть отображение $f : X \to Y$ биективно. Тогда для любого $y \in Y$ полный прообраз $f^{-1}(y)$ состоит ровно из одного элемента. Поэтому он непуст. Это доказывает сюръективность отображения $f : X \to Y$. Допустим, что условие инъективности отображения $f : X \to Y$ нарушено. Тогда найдутся два различных элемента $x_1 \neq x_2$ в множестве X, для которых $f(x_1) = f(x_2)$. Обозначим $y = f(x_1) = f(x_2)$ и рассмотрим полный прообраз $f^{-1}(y)$. Из $f(x_1) = y$ выводим $x_1 \in f^{-1}(y)$, а из $f(x_2) = y$ получаем $x_2 \in f^{-1}(y)$. Значит, множество $f^{-1}(y)$ содержит, как минимум, два элемента x_1 и x_2. Это противоречит биективности отображения $f : X \to Y$. Полученное противоречие доказывает инъективность отображения $f : X \to Y$.

Перейдем к доказательству достаточности. Пусть отображение $f\colon X \to Y$ инъективно и сюръективно одновременно. В силу сюръективности множества $f^{-1}(y)$ непусты для всех $y \in Y$. Допустим, что какое-то из них $f^{-1}(y)$ содержит более одного элемента. Пусть $x_1 \neq x_2$ два различных элемента из $f^{-1}(y)$. Тогда $f(x_1) = y = f(x_2)$, что противоречит инъективности отображения $f\colon X \to Y$. Значит, все множества $f^{-1}(y)$ непусты и каждое состоит ровно из одного элемента. Это доказывает биективность отображения f. $\square$

Теорема 1.2. *Отображение $f\colon X \to Y$ сюръективно тогда и только тогда, когда $\operatorname{Im} f = Y$.*

Док-во. Пусть $f\colon X \to Y$ сюръективно. Тогда для любого элемента $y \in Y$ полный прообраз $f^{-1}(y)$ непуст. Выбрав некоторый элемент $x \in f^{-1}(y)$, мы получим $y = f(x)$. Значит, всякий элемент $y \in Y$ есть образ некоторого элемента x при отображении f. Следовательно, $\operatorname{Im} f = Y$.

Наоборот, пусть $\operatorname{Im} f = Y$. Тогда всякий элемент $y \in Y$ есть образ некоторого x, т. е. $y = f(x)$. Значит, для всякого $y \in Y$ прообраз $f^{-1}(y)$ непуст, что означает сюръективность отображения f. $\square$

Рассмотрим два отображения $f : X \to Y$ и $g : Y \to Z$. Выбрав произвольный элемент $x \in X$, мы можем применить к нему отображение f. В результате получится элемент $f(x) \in Y$, к которому можно применить отображение g. Последовательное применение двух отображений $g(f(x))$ задает правило, которое каждому $x \in X$ сопоставляет однозначно определенный элемент $z = g(f(x)) \in Z$. Полученное отображение $\varphi : X \to Z$, называется *композицией отображений* g и f и обозначается $\varphi = g \circ f$.

Теорема 1.3. *Композиция $g \circ f$ двух инъективных отображений $f\colon X \to Y$ и $g\colon Y \to Z$ инъективна.*

Док-во. Рассмотрим два вектора x_1 и x_2 из пространства X. Положим $y_1 = f(x_1)$ и $y_2 = f(x_2)$. Отсюда $g \circ f(x_1) = g(y_1)$ и $g \circ f(x_2) = g(y_2)$. В силу инъективности f из $x_1 \neq x_2$ вытекает $y_1 \neq y_2$. Из инъективности g выводим $g(y_1) \neq g(y_2)$, что означает $g \circ f(x_1) \neq g \circ f(x_2)$. Инъективность $g \circ f$ доказана. $\square$

Теорема 1.4. *Композиция $g \circ f$ любых двух сюръективных отображений $f\colon X \to Y$ и $g\colon Y \to Z$ сюръективна.*

Док-во. Рассмотрим некоторый произвольный вектор z из пространства Z. В силу сюръективности g прообраз $g^{-1}(z)$ непуст. Выберем некоторый вектор $y \in g^{-1}(z)$ и рассмотрим его прообраз $f^{-1}(y)$. В силу сюръективности f он непуст. Выбрав некоторый вектор $x \in f^{-1}(y)$, получаем $g \circ f(x) = z$, то есть $x \in (g \circ f)^{-1}(z)$. Значит, прообраз $(g \circ f)^{-1}$ непуст. Сюръективность композиции $g \circ f$ доказана. $\square$

В качестве прямого следствия из доказанных двух теорем получаем теорему о композиции биекций.

Теорема 1.5. *Композиция $g \circ f$ двух биективных отображений $f\colon X \to Y$ и $g\colon Y \to Z$ биективна.*

Пусть заданы три отображения $f\colon X \to Y$, $g\colon Y \to Z$ и $h\colon Z \to U$. Тогда мы можем образовать композиции этих отображений двумя способами:

$$\varphi = h \circ (g \circ f), \qquad\qquad \psi = (h \circ g) \circ f. \tag{1.1}$$

Факт совпадения этих двух отображений составляет содержание следующей теоремы об ассоциативности.

Теорема 1.6. *Операция взятия композиции отображений ассоциативна, т. е.* $h \circ (g \circ f) = (h \circ g) \circ f$.

Док-во. Согласно определению 1.1 проверка совпадения двух отображений $\varphi\colon X \to U$ и $\psi\colon X \to U$ из (1.1) сводится к проверке равенства $\varphi(x) = \psi(x)$ для произвольного элемента $x \in X$. Положим $\alpha = h \circ g$ и $\beta = g \circ f$. Тогда

$$\begin{aligned} &\varphi(x) = h \circ \beta(x) = h(\beta(x)) = h(g(f(x))),\\ &\psi(x) = \alpha \circ f(x) = \alpha(f(x)) = h(g(f(x))). \end{aligned} \tag{1.2}$$

Сравнивая правые части в равенствах (1.2), выводим требуемое совпадение $\varphi(x) = \psi(x)$ для отображений (1.1). Значит, $h \circ (g \circ f) = (h \circ g) \circ f$. Теорема доказана. $\square$

Рассмотрим отображение $f\colon X \to Y$, а также пару единичных отображений $\mathrm{id}_X\colon X \to X$ и $\mathrm{id}_Y\colon Y \to Y$. Последние два отображения определяются так:

$$\mathrm{id}_X(x) = x, \qquad\qquad \mathrm{id}_Y(y) = y.$$

Определение 1.5. Отображение $l\colon Y \to X$ называют *левым обратным* для отображения $f\colon X \to Y$, если $l \circ f = \mathrm{id}_X$.

Определение 1.6. Отображение $r\colon Y \to X$ называют *правым обратным* для отображения $f\colon X \to Y$, если $f \circ r = \mathrm{id}_Y$.

Вопрос о существовании левого и правого обратного отображений решается следующими двумя теоремами.

Теорема 1.7. *Отображение $f\colon X \to Y$ обладает левым обратным отображением l тогда и только тогда, когда оно инъективно.*

Теорема 1.8. *Отображение $f\colon X \to Y$ обладает правым обратным отображением r тогда и только тогда, когда оно сюръективно.*

Док-во теоремы 1.7. Пусть отображение $f\colon X \to Y$ обладает левым обратным отображением l. Выберем два вектора x_1 и x_2 из пространства X и положим $y_1 = f(x_1)$ и $y_2 = f(x_2)$. Тогда равенство $l \circ f = \mathrm{id}_X$ дает $x_1 = l(y_1)$ и $x_2 = l(y_2)$. Следовательно, равенство $y_1 = y_2$ влечет $x_1 = x_2$, а из $x_1 \neq x_2$ вытекает $y_1 \neq y_2$. Таким образом, из существования левого обратного отображения l мы вывели инъективность f.

Теперь в обратную сторону. Пусть отображение f инъективно. Выберем и фиксируем некоторый элемент $x_0 \in X$. Для каждого элемента $y \in \operatorname{Im} f$ прообраз $f^{-1}(y)$ непуст. Для всех таких y выберем и фиксируем элементы $x_y \in f^{-1}(y)$. Определим отображение $l\colon Y \to X$ следующим образом:

$$l(y) = \begin{cases} x_y & \text{при} \ \ y \in \operatorname{Im} f,\\ x_0 & \text{при} \ \ y \notin \operatorname{Im} f. \end{cases}$$

Рассмотрим композицию $l \circ f$. Нетрудно видеть, что для произвольного элемента $x \in X$ выполнено соотношение $l \circ f(x) = x_y$, где $y = f(x)$. Но $f(x_y) = y = f(x)$. Используя инъективность отображения f, получаем $x_y = x$. Значит, $l \circ f(x) = x$ для произвольного $x \in X$. Равенство $l \circ f = \mathrm{id}_X$ для отображения l доказано. Это отображение и есть требуемое левое обратное отображение для f. Теорема доказана. $\square$

Док-во теоремы 1.8. Пусть отображение $f\colon X \to Y$ обладает правым обратным отображением r. Для произвольного элемента $y \in Y$ из равенства $f \circ r = \mathrm{id}_Y$ выводим $y = f(r(y))$. Значит, $r(y) \in f^{-1}(y)$ и прообраз $f^{-1}(y)$ непуст. Сюръективность f установлена.

Теперь наоборот. Пусть отображение f сюръективно. Для каждого $y \in Y$ прообраз $f^{-1}(y)$ непуст. В каждом из множеств $f^{-1}(y)$ выберем и фиксируем ровно по одному элементу $x_y \in f^{-1}(y)$. Это позволяет определить отображение $r\colon Y \to X$, полагая $r(y) = x_y$. Но $f(x_y) = y$, поэтому $f(r(y)) = y$ и $f \circ r = \mathrm{id}_Y$. Существование правого обратного отображения r установлено. $\square$

Отметим, что отображения $l\colon Y \to X$ и $r\colon Y \to X$, построенные при доказательстве теорем 1.7 и 1.8, в общем случае не являются единственно возможными. Даже способ их построения содержит определенный элемент произвола.

Определение 1.7. Отображение $f^{-1}\colon Y \to X$ называют *двухсторонним обратным* отображением или просто *обратным* отображением для $f\colon X \to Y$, если выполнены следующие соотношения:

$$f^{-1} \circ f = \mathrm{id}_X, \qquad f \circ f^{-1} = \mathrm{id}_Y\,. \tag{1.3}$$

Теорема 1.9. *Отображение $f\colon X \to Y$ обладает левым и правым обратными отображениями l и r тогда и только тогда, когда оно биективно. В этом случае отображения l и r однозначно определены, совпадают друг с другом и определяют единственное двустороннее обратное отображение $l = r = f^{-1}$.*

Док-во. Первое утверждение теоремы 1.9 вытекает из теорем 1.7 и 1.8 и теоремы 1.1. Докажем оставшиеся утверждения теоремы 1.9. Совпадение $l = r$ вытекает из следующей цепочки равенств:

$$l = l \circ \mathrm{id}_Y = l \circ (f \circ r) = (l \circ f) \circ r = \mathrm{id}_X \circ r = r.$$

Отсюда же вытекает единственность левого обратного отображения. Если допустить существование еще одного левого обратного отображения l', то из $l = r$ и $l' = r$ вытекает $l = l'$.

Аналогичным образом, допустив существование еще одного правого обратного отображения r', получаем $l = r$ и $l = r'$. Откуда $r = r'$. Совпадающие друг с другом левое и правое обратные отображения задают единственное двухстороннее обратное отображение $f^{-1} = l = r$, удовлетворяющее соотношениям (1.3). $\square$

§ 2. Линейные векторные пространства.

Пусть M — некоторое множество. *Бинарной алгебраической операцией* на M называют некоторое правило, которое каждой упорядоченной паре элементов x, y из множества M сопоставляет некоторый однозначно определенный третий элемент $z \in M$. Это правило можно обозначать в форме $z = f(x, y)$. Такая форма записи называется *префиксной* формой записи алгебраической операции: в ней знак операции f предшествует элементам x и y, к которым он применяется. Имеется и другая — *инфиксная* форма записи алгебраической операции, когда знак операции ставится между элементами x и y. Примером могут служить бинарные операции сложения и умножения чисел: $z = x + y$, $z = x \cdot y$. Иногда роль знака алгебраической операции играют специальные скобки, а разделителем служит обычная запятая. Примером такого обозначения служит векторное произведение трехмерных векторов: $z = [x, y]$.

Пусть $\mathbb{K}$ — числовое поле. Под числовым полем в этой книге мы будем понимать одно из трех полей: поле рациональных чисел $\mathbb{K} = \mathbb{Q}$, поле вещественных чисел $\mathbb{K} = \mathbb{R}$ или поле комплексных чисел $\mathbb{K} = \mathbb{C}$. Скажем, что на множестве M задана операция *умножения на числа из поля* $\mathbb{K}$, если задано правило, которое каждой паре α, x, состоящей из числа $\alpha \in \mathbb{K}$ и элемента $x \in M$, ставит в соответствие некоторый элемент $y \in M$. Операция умножения на число записывается в инфиксной форме: $y = \alpha \cdot x$. Знак умножения в этой записи часто не ставится: $y = \alpha\, x$.

Определение 2.1. Множество V, оснащенное бинарной операцией сложения и операцией умножения на числа из поля $\mathbb{K}$, называется *линейным векторным пространством над полем* $\mathbb{K}$, если выполнены условия

(1) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ для всех $\mathbf{u}, \mathbf{v} \in V$;
(2) $(\mathbf{u} + v) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ для всех $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$;
(3) существует *нулевой элемент* $\mathbf{0} \in V$, такой, что $\mathbf{v} + \mathbf{0} = \mathbf{v}$ для всякого элемента $\mathbf{v} \in V$;
(4) для всякого $\mathbf{v} \in V$ существует *противоположный элемент* $\mathbf{v}' \in V$, такой, что $\mathbf{v} + \mathbf{v}' = \mathbf{0}$;
(5) $\alpha \cdot (\mathbf{u} + \mathbf{v}) = \alpha \cdot \mathbf{u} + \alpha \cdot \mathbf{v}$ для всякого числа $\alpha \in \mathbb{K}$ и для всяких элементов $\mathbf{u}, \mathbf{v} \in V$;
(6) $(\alpha + \beta) \cdot \mathbf{v} = \alpha \cdot \mathbf{v} + \beta \cdot \mathbf{v}$ для любых двух чисел $\alpha, \beta \in \mathbb{K}$ и для всякого элемента $\mathbf{v} \in V$;
(7) $\alpha \cdot (\beta \cdot \mathbf{v}) = (\alpha\beta) \cdot \mathbf{v}$ для любых двух чисел $\alpha, \beta \in \mathbb{K}$ и для всякого элемента $\mathbf{v} \in V$;
(8) $1 \cdot \mathbf{v} = \mathbf{v}$ для числа $1 \in \mathbb{K}$ и любого элемента $\mathbf{v} \in V$.

Элементы линейного векторного пространства принято называть *векторами*, а условия (1)-(8) называют аксиомами линейного векторного пространства. В зависимости от рассматриваемого случая $\mathbb{K} = \mathbb{Q}$, $\mathbb{K} = \mathbb{R}$ или $\mathbb{K} = \mathbb{C}$ мы будем различать *рациональные*, *вещественные* и *комплексные* линейные векторные пространства. Большинство приводимых в этой книге результатов справедливо для случая произвольного числового поля $\mathbb{K}$. При формулировке таких результатов мы не будем специально оговаривать тип линейного векторного пространства.

Аксиомы (1) и (2) — это аксиома *коммутативности*[1] и аксиома *ассоциативности* для операции сложения векторов. Аксиомы (5) и (6) выражают свойство *дистрибутивности*.

---

[1] Система аксиом (1)-(8) избыточна: аксиому (1) можно вывести из остальных. Автор признателен А. В. Муфтахову, сообщившему этот любопытный факт.

Теорема 2.1. *Алгебраические операции с векторами произвольного линейного векторного пространства V обладают следующими свойствами:*

(9) *нулевой вектор $\mathbf{0} \in V$ единственен;*
(10) *для любого вектора $\mathbf{v}$ из пространства V имеется ровно один противоположный вектор $\mathbf{v}'$;*
(11) *произведение числа $0 \in \mathbb{K}$ и любого вектора $\mathbf{v}$ из пространства V есть нулевой вектор: $0 \cdot \mathbf{v} = \mathbf{0}$;*
(12) *произведение любого числа $\alpha \in K$ и нулевого вектора есть нулевой вектор: $\alpha \cdot \mathbf{0} = \mathbf{0}$;*
(13) *умножение числа $-1 \in \mathbb{K}$ и вектора $\mathbf{v} \in V$ дает противоположный вектор: $(-1) \cdot \mathbf{v} = \mathbf{v}'$.*

Док-во. Свойства (9)-(13) вытекают непосредственно из аксиом (1)-(8). В силу этого их нумерация продолжает нумерацию аксиом линейного векторного пространства.

Предположим, что в линейном векторном пространстве имеются два элемента $\mathbf{0}$ и $\mathbf{0}'$ со свойствами нулевого вектора. Тогда для любого вектора $\mathbf{v} \in V$ из аксиомы (3) имеем $\mathbf{v} = \mathbf{v} + \mathbf{0}$ и $\mathbf{v} + \mathbf{0}' = \mathbf{v}$. Подставим $\mathbf{v} = \mathbf{0}'$ в первое равенство и $\mathbf{v} = \mathbf{0}$ во второе равенство. С учетом аксиомы (1) это дает

$$\mathbf{0}' = \mathbf{0}' + \mathbf{0} = \mathbf{0} + \mathbf{0}' = \mathbf{0}.$$

Значит, вектора $\mathbf{0}$ и $\mathbf{0}'$ совпадают: $\mathbf{0} = \mathbf{0}'$. Единственность нулевого вектора доказана.

Пусть $\mathbf{v}$ — некоторый произвольный вектор из V. Предположим, что имеются два вектора $\mathbf{v}'$ и $\mathbf{v}''$, противоположных вектору $\mathbf{v}$. Тогда

$$\mathbf{v} + \mathbf{v}' = \mathbf{0}, \qquad\qquad \mathbf{v} + \mathbf{v}'' = \mathbf{0}.$$

Следующая цепочка вычислений доказывают единственность противоположного вектора в линейном векторном пространстве V:

$$\begin{aligned}\mathbf{v}'' = \mathbf{v}'' + \mathbf{0} = \mathbf{v}'' + (\mathbf{v} + \mathbf{v}') = (\mathbf{v}'' + \mathbf{v}) + \mathbf{v}' = \\ = (\mathbf{v} + \mathbf{v}'') + \mathbf{v}' = \mathbf{0} + \mathbf{v}' = \mathbf{v}' + \mathbf{0} = \mathbf{v}'.\end{aligned}$$

При выводе $\mathbf{v}'' = \mathbf{v}'$ выше мы использовали аксиому (4), аксиомы ассоциативности (2) и дважды использовали аксиому коммутативности (1).

Пусть вновь $\mathbf{v}$ — некоторый произвольный вектор из V. Положим $\mathbf{x} = 0 \cdot \mathbf{v}$. Сложим вектор $\mathbf{x}$ с $\mathbf{x}$ и применим аксиому дистрибутивности (6). Это дает

$$\mathbf{x} + \mathbf{x} = 0 \cdot \mathbf{v} + 0 \cdot \mathbf{v} = (0 + 0) \cdot \mathbf{v} = 0 \cdot \mathbf{v} = \mathbf{x}.$$

Из доказанного равенства $\mathbf{x} + \mathbf{x} = \mathbf{x}$ легко выводим $\mathbf{x} = \mathbf{0}$:

$$\mathbf{x} = \mathbf{x} + \mathbf{0} = \mathbf{x} + (\mathbf{x} + \mathbf{x}') = (\mathbf{x} + \mathbf{x}) + \mathbf{x}' = \mathbf{x} + \mathbf{x}' = \mathbf{0}.$$

Здесь мы вновь воспользовались аксиомой ассоциативности (2). Свойство (11) доказано.

Пусть α — некоторое число из числового поля $\mathbb{K}$. Положим $\mathbf{x} = \alpha \cdot \mathbf{0}$, где $\mathbf{0}$ — нулевой вектор из V. Тогда

$$\mathbf{x} + \mathbf{x} = \alpha \cdot \mathbf{0} + \alpha \cdot \mathbf{0} = \alpha \cdot (\mathbf{0} + \mathbf{0}) = \alpha \cdot \mathbf{0} = \mathbf{x}.$$

Здесь мы использовали аксиому (5) и использовали свойство нулевого вектора из аксиомы (3). Из равенства $\mathbf{x} + \mathbf{x} = \mathbf{x}$ вытекает $\mathbf{x} = \mathbf{0}$ (см. выше). Свойство (12) доказано.

Пусть $\mathbf{v}$ — некоторый произвольный вектор из V. Положим $\mathbf{x} = (-1) \cdot \mathbf{v}$. Для вектора $\mathbf{x}$ выводим

$$\mathbf{v} + \mathbf{x} = 1 \cdot \mathbf{v} + \mathbf{x} = 1 \cdot \mathbf{v} + (-1) \cdot \mathbf{v} = (1 + (-1)) \cdot \mathbf{v} = 0 \cdot \mathbf{v} = \mathbf{0}.$$

Здесь мы использовали аксиому (8) и аксиому дистрибутивности (6). Из выведенного равенства $\mathbf{v} + \mathbf{x} = \mathbf{0}$ видим, что $\mathbf{x}$ есть вектор, противоположный вектору $\mathbf{v}$ в смысле аксиомы (4). Из доказанного выше свойства единственности противоположного вектора (10) находим $\mathbf{x} = \mathbf{v}'$, то есть $(-1) \cdot \mathbf{v} = \mathbf{v}'$. Теорема полностью доказана. $\square$

Аксиомы коммутативности и ассоциативности позволяют не заботиться о расстановке скобок и о порядке слагаемых при записи сумм векторов. Свойство (8) и аксиомы (7) и (8) дают

$$(-1) \cdot \mathbf{v}' = (-1) \cdot ((-1) \cdot \mathbf{v}) = ((-1)(-1)) \cdot \mathbf{v} = 1 \cdot \mathbf{v} = \mathbf{v}.$$

Это равенство делает естественным обозначение $\mathbf{v}' = -\mathbf{v}$ для противоположного вектора. При этом

$$-\alpha \cdot \mathbf{v} = -(\alpha \cdot \mathbf{v}) = (-1) \cdot (\alpha \cdot \mathbf{v}) = (-\alpha) \cdot \mathbf{v}.$$

Операция *вычитания* векторов противоположна операции сложения. Она определяется как сложение с противоположным вектором: $\mathbf{x} - \mathbf{y} = \mathbf{x} + (-\mathbf{y})$. Следующие свойства операции вычитания

$$\begin{aligned}
&(\mathbf{a} + \mathbf{b}) - \mathbf{c} = \mathbf{a} + (\mathbf{b} - \mathbf{c}),\\
&(\mathbf{a} - \mathbf{b}) + \mathbf{c} = \mathbf{a} - (\mathbf{b} - \mathbf{c}),\\
&(\mathbf{a} - \mathbf{b}) - \mathbf{c} = \mathbf{a} - (\mathbf{b} + \mathbf{c}),\\
&\alpha \cdot (\mathbf{x} - \mathbf{y}) = \alpha \cdot \mathbf{x} - \alpha \cdot \mathbf{y}
\end{aligned}$$

делают вычисления с векторами естественными, простыми и очень похожими на вычисления с числами. Доказательство перечисленных свойств мы оставляем читателю.

Рассмотрим некоторые примеры линейных векторных пространств. Вещественное *арифметическое линейное векторное пространство* $\mathbb{R}^n$ определяется как множество всевозможных упорядоченных наборов из n вещественных чисел $x^1, \ldots, x^n$. Такие наборы удобно изображать в форме *вектор-столбцов*. Алгебраические операции с вектор-столбцами выполняются покомпонентно:

$$\left\|\begin{matrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{matrix}\right\| + \left\|\begin{matrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{matrix}\right\| = \left\|\begin{matrix} x^1 + y^1 \\ x^2 + y^2 \\ \vdots \\ x^n + y^n \end{matrix}\right\| \qquad \alpha \cdot \left\|\begin{matrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{matrix}\right\| = \left\|\begin{matrix} \alpha \cdot x^1 \\ \alpha \cdot x^2 \\ \vdots \\ \alpha \cdot x^n \end{matrix}\right\| \tag{2.1}$$

Мы представляем читателю возможность самостоятельно убедиться в том, что множество $\mathbb{R}^n$ с алгебраическими операциями (2.1) есть линейное векторное пространство над полем вещественных чисел. Аналогичным образом определяется рациональное арифметическое векторное пространство $\mathbb{Q}^n$ над полем $\mathbb{Q}$ и комплексное арифметическое векторное пространство $\mathbb{C}^n$ над полем $\mathbb{C}$.

Рассмотрим множество m-кратно непрерывно дифференцируемых вещественнозначных функций на отрезке $[-1, 1]$. Это множество обычно обозначают так: $C^m([-1, 1])$. Операции сложения функций и умножения функции на число в $C^m([-1, 1])$ определяются поточечно. Значение функции $f + g$ в точке a есть сумма значений функций f и g в этой точке, а значение функции $\alpha \cdot f$ в точке a есть произведение чисел α и $f(a)$. Нетрудно убедиться в том, что множество функций $C^m([-1, 1])$ с поточечным сложением и умножением на число есть линейное векторное пространство над полем вещественных чисел $\mathbb{R}$. Проверку этого мы также оставляем читателю.

Определение 2.2. Непустое подмножество $U \subset V$ линейного векторного пространства V над полем $\mathbb{K}$ называется подпространством в V, если:

(1) из $\mathbf{u}_1, \mathbf{u}_2 \in U$ вытекает $\mathbf{u}_1 + \mathbf{u}_2 \in U$;
(2) из $\mathbf{u} \in U$ вытекает $\alpha \cdot \mathbf{u} \in U$ для любого числа $\alpha \in \mathbb{K}$.

Пусть U — подпространство линейного векторного пространства V. Рассмотрим U как отдельное изолированное множество. В силу условий (1) и (2) это множество замкнуто относительно операций сложения и умножения на число. Нетрудно показать, что нулевой вектор $\mathbf{0}$ принадлежит U и для всякого $\mathbf{u} \in U$ противоположный вектор $\mathbf{u}'$ также принадлежит U. Это вытекает из формул $\mathbf{0} = 0 \cdot \mathbf{u}$ и $\mathbf{u}' = (-1) \cdot \mathbf{u}$, где $\mathbf{u} \in U$. Основываясь на этих фактах, нетрудно доказать, что любое подпространство $U \subset V$, рассматриваемое изолированно, само является линейным векторным пространством над полем $\mathbb{K}$. Действительно, выполнение аксиом (3) и (4) мы уже установили. Проверка аксиом (1), (2) и (5)-(8) состоит в проверке некоторых равенств, записанных в терминах операций сложения и умножения на число. Будучи выполненными для произвольных векторов из V, эти равенства выполнены и для векторов из подмножества $U \subset V$. Замкнутость U относительно алгебраических операций гарантирует нам, что все определяемые этими равенствами вычисления не выводят за пределы множества U.

В качестве примеров отметим следующие подпространства в функциональном пространстве $C^m([-1, 1])$:

- подпространство четных функций;
- подпространство нечетных функций;
- подпространство полиномов.

§ 3. Линейная зависимость и независимость.

Пусть $\mathbf{v}_1, \ldots, \mathbf{v}_n$ — некоторая система векторов из линейного векторного пространства V. Операции сложения и умножения на числа позволяют образовать из этих векторов выражения вида

$$\mathbf{y} = \alpha_1 \cdot \mathbf{v}_1 + \ldots + \alpha_n \cdot \mathbf{v}_n. \tag{3.1}$$

Выражение вида (3.1) называется *линейной комбинацией* векторов $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Числа $\alpha_1, \ldots, \alpha_n$ из поля $\mathbb{K}$ называются *коэффициентами* линейной комбинации, а вектор $\mathbf{y}$ называется *значением* линейной комбинации. Будем говорить, что линейная комбинация *равна нулю*, если равно нулю ее значение.

Линейная комбинация называется *тривиальной*, если все ее коэффициенты равны нулю: $\alpha_1 = \ldots = \alpha_n = 0$. В противном случае линейная комбинация называется *нетривиальной*.

Определение 3.1. Система векторов $\mathbf{v}_1, \ldots, \mathbf{v}_n$ из линейного векторного пространства V называется *линейно зависимой*, если существует нетривиальная линейная комбинация из этих векторов, равная нулю.

Определение 3.2. Система векторов $\mathbf{v}_1, \ldots, \mathbf{v}_n$ из линейного векторного пространства V называется *линейно независимой*, если из равенства нулю линейной комбинации этих векторов вытекает ее тривиальность.

Понятие линейной независимости получается прямым логическим отрицанием понятия линейной зависимости. Читатель может может дать несколько эквивалентных формулировок для определения этого понятия. Мы привели лишь одну из таких формулировок, которая, на наш взгляд, будет наиболее удобна в дальнейшем.

Введем еще одно понятие, связанное с линейными комбинациями. Скажем, что вектор $\mathbf{y}$ *линейно выражается* через вектора $\mathbf{v}_1, \ldots, \mathbf{v}_n$, если $\mathbf{y}$ есть значение некоторой линейной комбинации, составленной из векторов $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

Теорема 3.1. *Отношение линейной зависимости векторов обладает следующими основными свойствами:*

(1) *система векторов, содержащая нулевой вектор, линейно зависима;*
(2) *система векторов, содержащая линейно зависимую подсистему, линейно зависима;*
(3) *если система векторов линейно зависима, то один из этих векторов линейно выражается через остальные;*
(4) *если система векторов $\mathbf{v}_1, \ldots, \mathbf{v}_n$ линейно независима, а добавление вектора $\mathbf{v}_{n+1}$ делает ее линейно зависимой, то вектор $\mathbf{v}_{n+1}$ линейно выражается через $\mathbf{v}_1, \ldots, \mathbf{v}_n$;*
(5) *если вектор $\mathbf{x}$ линейно выражается через $\mathbf{y}_1, \ldots, \mathbf{y}_m$, а каждый из векторов $\mathbf{y}_1, \ldots, \mathbf{y}_m$ линейно выражается через $\mathbf{z}_1, \ldots, \mathbf{z}_n$, то $\mathbf{x}$ линейно выражается через вектора $\mathbf{z}_1, \ldots, \mathbf{z}_n$.*

Док-во. Пусть система векторов $\mathbf{v}_1, \ldots, \mathbf{v}_n$ содержит нулевой вектор. Для определенности можно считать $\mathbf{v}_k = \mathbf{0}$. Составим из векторов $\mathbf{v}_1, \ldots, \mathbf{v}_n$ следующую линейную комбинацию:

$$0 \cdot \mathbf{v}_1 + \ldots + 0 \cdot \mathbf{v}_{k-1} + 1 \cdot \mathbf{v}_k + 0 \cdot \mathbf{v}_{k+1} + \ldots + 0 \cdot \mathbf{v}_n = \mathbf{0}.$$

Эта линейная комбинация нетривиальна, ибо коэффициент при $\mathbf{v}_k$ отличен от нуля. А значение этой комбинации равно нулю. Свойство (1) доказано.

Пусть система векторов $\mathbf{v}_1, \ldots, \mathbf{v}_n$ содержит линейно зависимую подсистему. Поскольку понятие линейной зависимости нечувствительно к порядку нумерации векторов, можно считать, что линейно зависимая подсистема состоит из первых k векторов $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Тогда существует нетривиальная линейная комбинация этих k векторов, равная нулю:

$$\alpha_1 \cdot \mathbf{v}_1 + \ldots + \alpha_k \cdot \mathbf{v}_k = \mathbf{0}.$$

Эту линейную комбинацию можно превратить в линейную комбинацию из полной системы векторов $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Надо лишь добавить недостающие вектора с нулевыми коэффициентами:

$$\alpha_1 \cdot \mathbf{v}_1 + \ldots + \alpha_k \cdot \mathbf{v}_k + 0 \cdot \mathbf{v}_{k+1} + \ldots + 0 \cdot \mathbf{v}_n = \mathbf{0}.$$

Полученная линейная комбинация нетривиальна, а ее значение равно нулю. Свойство (2) доказано.

Пусть вектора $\mathbf{v}_1, \ldots, \mathbf{v}_n$ линейно зависимы. Тогда существует нетривиальная линейная комбинация этих векторов, равная нулю:

$$\alpha_1 \cdot \mathbf{v}_1 + \ldots + \alpha_n \cdot \mathbf{v}_n = \mathbf{0}. \tag{3.2}$$

Нетривиальность линейной комбинации векторов (3.2) означает, что один из ее коэффициентов отличен от нуля. Пусть для определенности $\alpha_k \neq 0$. Запишем линейную комбинацию (3.2) более подробно:

$$\alpha_1 \cdot \mathbf{v}_1 + \ldots + \alpha_k \cdot \mathbf{v}_k + \ldots + \alpha_n \cdot \mathbf{v}_n = \mathbf{0}.$$

Перенесем слагаемое $\alpha_k \cdot v_k$ в правую часть равенства и поделим полученное в результате этого выражение на $-\alpha_k$:

$$\mathbf{v}_k = -\frac{\alpha_1}{\alpha_k} \cdot \mathbf{v}_1 - \ldots - \frac{\alpha_{k-1}}{\alpha_k} \cdot \mathbf{v}_{k-1} - \frac{\alpha_{k+1}}{\alpha_k} \cdot \mathbf{v}_{k+1} - \ldots - \frac{\alpha_n}{\alpha_k} \cdot \mathbf{v}_n.$$

Отсюда видим, что один из векторов $\mathbf{v}_k$ линейно выражается через остальные векторы системы. Свойство (3) доказано.

Пусть теперь система векторов $\mathbf{v}_1, \ldots, \mathbf{v}_n$ линейно независима, а добавление вектора $\mathbf{v}_{n+1}$ делает ее линейно зависимой. Тогда имеется нетривиальная линейная комбинация векторов $\mathbf{v}_1, \ldots, \mathbf{v}_{n+1}$, равная нулю:

$$\alpha_1 \cdot \mathbf{v}_1 + \ldots + \alpha_n \cdot \mathbf{v}_n + \alpha_{n+1} \cdot \mathbf{v}_{n+1} = \mathbf{0}.$$

Докажем, что $\alpha_{n+1} \neq 0$. Допустив противоположное $\alpha_{n+1} = 0$, мы получим нетривиальную линейную комбинацию векторов $\mathbf{v}_1, \ldots, \mathbf{v}_n$, равную нулю:

$$\alpha_1 \cdot \mathbf{v}_1 + \ldots + \alpha_n \cdot \mathbf{v}_n = \mathbf{0}.$$

А это противоречит линейной независимости первых n векторов $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Значит, $\alpha_{n+1} \neq 0$, что позволяет повторить прием, использованный выше:

$$\mathbf{v}_{n+1} = -\frac{\alpha_1}{\alpha_{n+1}} \cdot \mathbf{v}_1 - \ldots - \frac{\alpha_n}{\alpha_{n+1}} \cdot \mathbf{v}_n.$$

Полученное выражение для вектора $\mathbf{v}_{n+1}$ доказывает свойство (4).

Пусть вектор $\mathbf{x}$ линейно выражается через $\mathbf{y}_1, \ldots, \mathbf{y}_m$, а каждый из векторов $\mathbf{y}_1, \ldots, \mathbf{y}_m$ выражается через $\mathbf{z}_1, \ldots, \mathbf{z}_n$. Этот факт запишем так:

$$\mathbf{x} = \sum_{i=1}^{m} \alpha_i \cdot \mathbf{y}_i, \qquad\qquad \mathbf{y}_i = \sum_{j=1}^{n} \beta_{ij} \cdot \mathbf{z}_j.$$

Подставив вторую формулу в первую, для $\mathbf{x}$ получаем:

$$\mathbf{x} = \sum_{i=1}^{m} \alpha_i \cdot \left(\sum_{j=1}^{n} \beta_{ij} \cdot \mathbf{z}_j \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} \alpha_i \beta_{ij} \right) \cdot \mathbf{z}_j$$

Полученное выражение для вектора $\mathbf{x}$ показывает, что он линейно выражается через вектора $\mathbf{z}_1, \ldots, \mathbf{z}_n$. Свойство (5) доказано. Это завершает доказательство всей теоремы 3.1. $\square$

Отметим важное следствие из свойства (2) в теореме 3.1.

Следствие. *Любая подсистема в линейно независимой системе векторов линейно независима.*

Следующее свойство отношения линейной зависимости известно как теорема Штейница. Она описывает некоторый количественный аспект этого понятия.

Теорема 3.2 (Штейниц). *Если вектора $\mathbf{x}_1, \ldots, \mathbf{x}_n$ линейно независимы и если каждый вектор этой системы векторов линейно выражается через вектора $\mathbf{y}_1, \ldots, \mathbf{y}_m$, то $m \geqslant n$.*

Док-во. Доказательство проведем индукцией по числу n векторов в системе $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Начнем со случая $n = 1$. Линейная независимость системы, состоящей из одного вектора, эквивалентна условию $\mathbf{x}_1 \neq 0$. Для выражения ненулевого вектора $\mathbf{x}_1$ через вектора системы $\mathbf{y}_1, \ldots, \mathbf{y}_m$ эта система должна содержать, как минимум, один вектор. Значит, $m \geqslant 1$. База индукции доказана.

Предположим, что теорема справедлива для случая $n = k$. В этом предположении докажем справедливость теоремы при $n = k + 1$. В этом случае мы имеем систему линейно независимых векторов $\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}$, каждый из которых выражается через $\mathbf{y}_1, \ldots, \mathbf{y}_m$. Изобразим это обстоятельство формулами

$$\begin{aligned} &\mathbf{x}_1 = \alpha_{11} \cdot \mathbf{y}_1 + \ldots + \alpha_{1m} \cdot \mathbf{y}_m, \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ &\mathbf{x}_k = \alpha_{k1} \cdot \mathbf{y}_1 + \ldots + \alpha_{km} \cdot \mathbf{y}_m. \end{aligned} \tag{3.3}$$

Аналогичную формулу для выражения погледнего $(k + 1)$-вектора $\mathbf{x}_{k+1}$ через вектора $\mathbf{y}_1, \ldots, \mathbf{y}_m$ запишем в несколько ином виде:

$$\mathbf{x}_{k+1} = \beta_1 \cdot \mathbf{y}_1 + \ldots + \beta_m \cdot \mathbf{y}_m.$$

В силу линейной независимости системы векторов $\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}$ вектор $\mathbf{x}_{k+1}$ отличен от нуля. Поэтому среди чисел $\beta_1, \ldots, \beta_m$ есть, по меньшей мере, одно ненулевое. Выполнив, если это необходимо, перенумерацию векторов $\mathbf{y}_1, \ldots, \mathbf{y}_m$, можем считать, что $\beta_m \neq 0$. Тогда

$$\mathbf{y}_m = \frac{1}{\beta_m} \cdot \mathbf{x}_{k+1} - \frac{\beta_1}{\beta_m} \cdot \mathbf{y}_1 - \ldots - \frac{\beta_{m-1}}{\beta_m} \cdot \mathbf{y}_{m-1}. \tag{3.4}$$

Подставим (3.4) в соотношения (3.3) и произведем в них приведение подобных членов. После этого соотношения (3.3) можно записать в виде

$$\mathbf{x}_i - \frac{\alpha_{im}}{\beta_m} \cdot \mathbf{x}_{k+1} = \sum_{j=1}^{m-1} \left(\alpha_{ij} - \beta_j \frac{\alpha_{im}}{\beta_m} \right) \cdot \mathbf{y}_j, \tag{3.5}$$

где $i = 1, \ldots, k$. Для того, чтобы упростить запись полученных формул (3.5) введем следующие обозначения:

$$\mathbf{x}_i^* = \mathbf{x}_i - \frac{\alpha_{im}}{\beta_m} \cdot \mathbf{x}_{k+1}, \qquad \alpha_{ij}^* = \alpha_{ij} - \beta_j \frac{\alpha_{im}}{\beta_m}. \tag{3.6}$$

В этих обозначениях соотношения (3.5) запишутся в виде

$$\begin{aligned} &\mathbf{x}_1^* = \alpha_{11}^* \cdot \mathbf{y}_1 + \ldots + \alpha_{1\,m-1}^* \cdot \mathbf{y}_{m-1}, \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ &\mathbf{x}_k^* = \alpha_{k1}^* \cdot \mathbf{y}_1 + \ldots + \alpha_{k\,m-1}^* \cdot \mathbf{y}_{m-1}. \end{aligned} \tag{3.7}$$

Согласно полученным формулам k векторов $\mathbf{x}_1^*, \ldots, \mathbf{x}_k^*$ линейно выражаются через $\mathbf{y}_1, \ldots, \mathbf{y}_{m-1}$. Для того, чтобы применить предположение индукции, необходимо установить линейную независимость векторов $\mathbf{x}_1^*, \ldots, \mathbf{x}_k^*$. Рассмотрим линейную комбинацию этих векторов, равную нулю:

$$\gamma_1 \cdot \mathbf{x}_1^* + \ldots + \gamma_k \cdot \mathbf{x}_k^* = \mathbf{0}. \tag{3.8}$$

Подставив в нее выражения для $\mathbf{x}_i^*$ из (3.6), после приведения подобных слагаемых в формуле (3.8) получим

$$\gamma_1 \cdot \mathbf{x}_1 + \ldots + \gamma_k \cdot \mathbf{x}_k - \left(\sum_{i=1}^{k} \gamma_i \frac{\alpha_{im}}{\beta_m} \right) \cdot \mathbf{x}_{k+1} = \mathbf{0}.$$

В силу линейной независимости исходной системы векторов $\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}$ выводим $\gamma_1 = \ldots = \gamma_k = 0$. Значит, линейная комбинация (3.8) тривиальна, что доказывает линейную независимость векторов $\mathbf{x}_1^*, \ldots, \mathbf{x}_k^*$. Теперь, применив предположение индукции к соотношениям (3.7), получим $m - 1 \geqslant k$. Отсюда вытекает требуемое неравенство $m \geqslant k + 1$, которое доказывает справедливость теоремы для случая $n = k + 1$. Индукционный переход завершен и теорема доказана. $\square$

§ 4. Порождающие системы и базисы.

Пусть $S \subset V$ некоторое непустое подмножество в линейном векторном пространстве V. Множество S может состоять как из конечного, так и из бесконечного числа векторов. Обозначим через $\langle S \rangle$, множество всех векторов, каждый из которых линейно выражается через некоторое конечное число векторов, взятых из множества S:

$$\langle S \rangle = \{\mathbf{v} \in V\colon\, \exists\, n\ (\mathbf{v} = \alpha_1 \cdot \mathbf{s}_1 + \ldots + \alpha_n \cdot \mathbf{s}_n,\ \text{где } \mathbf{s}_i \in S)\}.$$

Множество $\langle S \rangle$ принято называть *линейной оболочкой* подмножества $S \subset V$.

Теорема 4.1. *Линейная оболочка любого подмножества $S \subset V$ есть подпространство в линейном векторном пространстве V.*

Док-во. Для доказательства того, что $\langle S \rangle$ есть подпространство в V, надо проверить два условия из определения 2.2. Пусть $\mathbf{u}_1, \mathbf{u}_2 \in \langle S \rangle$. Тогда

$$\begin{aligned} \mathbf{u}_1 &= \alpha_1 \cdot \mathbf{s}_1 + \ldots + \alpha_n \cdot \mathbf{s}_n, \\ \mathbf{u}_2 &= \beta_1 \cdot \mathbf{s}_1^* + \ldots + \beta_m \cdot \mathbf{s}_m^*. \end{aligned}$$

Сложив эти два выражения, видим, что вектор $\mathbf{u}_1 + \mathbf{u}_2$ также линейно выражается через конечное число векторов, взятых из S. Поэтому $\mathbf{u}_1 + \mathbf{u}_2 \in \langle S \rangle$.

Пусть $\mathbf{u} \in \langle S \rangle$ и $\mathbf{u} = \alpha_1 \cdot \mathbf{s}_1 + \ldots + \alpha_n \cdot \mathbf{s}_n$. Для вектора $\alpha \cdot \mathbf{u}$ из этого получаем

$$\alpha \cdot \mathbf{u} = (\alpha\, \alpha_1) \cdot \mathbf{s}_1 + \ldots + (\alpha\, \alpha_n) \cdot \mathbf{s}_n,$$

откуда $\alpha \cdot \mathbf{u} \in \langle S \rangle$. Условия (1) и (2) для $\langle S \rangle$ выполнены. Теорема доказана. $\square$

Теорема 4.2. *Операция взятия линейной оболочки в линейном векторном пространстве V обладает следующими свойствами:*

(1) *если $S \subset U$ и U есть подпространство в V, то $\langle S \rangle \subset U$;*
(2) *линейная оболочка множества $S \subset V$ есть пересечение всех подпространств, содержащих в себе множество S.*

Док-во. Пусть $\mathbf{u} \in \langle S \rangle$ и $S \subset U$, где U — подпространство в V. Тогда $\mathbf{u} = \alpha_1 \cdot \mathbf{s}_1 + \ldots + \alpha_n \cdot \mathbf{s}_n$, причем из $\mathbf{s}_i \in S$ вытекает $\mathbf{s}_i \in U$. Но значение любой линейной комбинации, составленной из векторов подпространства U, принадлежит этому подпространству. Следовательно, $\mathbf{u} \in U$. Это доказывает включение $\langle S \rangle \subset U$.

Обозначим через W пересечение всех подпространств в V, содержащих в себе множество S. В силу уже доказанного свойства (1) множество $\langle S \rangle$ содержится в каждом из таких подпространств. Следовательно $\langle S \rangle \subset W$. С другой стороны, множество $\langle S \rangle$ само является подпространством, содержащим в себе множество S. Значит, оно входит в число тех подпространств, пересечением которых является W. Отсюда $W \subset \langle S \rangle$. Из полученных двух включений вытекает $\langle S \rangle = W$. Теорема доказана. $\square$

Пусть $\langle S \rangle = U$. При этом говорят, что множество $S \subset V$ *порождает* подпространство U посредством линейных комбинаций. Для случая $U = V$ эту терминологию закрепим в следующем определении.

Определение 4.1. Множество $S \subset V$ называется *порождающим множеством* или *порождающей системой векторов* в линейном векторном пространстве V, если $\langle S \rangle = V$.

В линейном векторном пространстве может существовать несколько порождающих систем векторов: $\langle S \rangle = \langle R \rangle = V$. В связи с этим возникает вопрос о выборе некоторой минимальной такой системы.

Определение 4.2. Порождающая система векторов $S \subset V$ называется *минимальной порождающей системой*, если никакая меньшая подсистема $S' \subsetneq S$ не является порождающей системой: $\langle S' \rangle \neq V$.

Определение 4.3. Система векторов $S \subset V$ называется *линейно независимой*, если любая конечная подсистема векторов $\mathbf{s}_1, \ldots, \mathbf{s}_n$, взятая из S, линейно независима.

В случае, когда система векторов S конечна, определение 4.3 сводится к определению 3.2. В общем случае связь между свойствами минимальности и линейной независимости для порождающих систем векторов дает следующая теорема.

Теорема 4.3. *Порождающая система векторов минимальна тогда и только тогда, когда она линейно независима.*

Док-во. Пусть порождающая система векторов $S \subset V$ минимальна. Докажем ее линейную независимость. Допустим, что система векторов S зависима. Тогда в ней имеется некоторая линейно зависимая конечная подсистема векторов $\mathbf{s}_1, \ldots, \mathbf{s}_n$. В силу пункта (3) теоремы 3.1 один из векторов $\mathbf{s}_k$ в этой подсистеме линейно выражается через остальные. Рассмотрим подсистему $S' = S \setminus \{\mathbf{s}_k\}$ в S, которая получается удалением вектора $\mathbf{s}_k$ из S. Очевидно, $S' \subsetneq S$. Докажем, что $\langle S' \rangle = V$. В силу $\langle S \rangle = V$ произвольный вектор $\mathbf{v} \in V$ линейно выражается через некоторое конечное число векторов из S. Если вектор $\mathbf{s}_k$ не участвует в этом выражении, то $\mathbf{v}$ линейно выражается через вектора системы S'. Если же $\mathbf{s}_k$ входит в выражение для вектора $\mathbf{v}$, мы можем исключить его из этого выражения в силу пункта (5) теоремы 3.1, пользуясь тем, что сам вектор $\mathbf{s}_k$ линейно выражается через вектора системы S'. Доказанное равенство $\langle S' \rangle = V$ противоречит минимальности порождающей системы векторов S. Полученное противоречие доказывает линейную независимость системы векторов S.

Пусть, наоборот, порождающая система векторов S линейно независима. Если она не является минимальной, то в ней найдется подсистема $S' \subsetneq S$, которая также порождает пространство V. Выберем некоторый вектор $\mathbf{s} \in S$, не принадлежащий S'. Этот вектор линейно выражается через некоторое конечное число векторов из подсистемы S':

$$\mathbf{s} = \alpha_1 \cdot \mathbf{s}_1 + \ldots + \alpha_n \cdot \mathbf{s}_n.$$

Переписав это в форме $\alpha_1 \cdot \mathbf{s}_1 + \ldots + \alpha_n \cdot \mathbf{s}_n + (-1) \cdot \mathbf{s} = \mathbf{0}$, мы видим, что вектора $\mathbf{s}_1, \ldots, \mathbf{s}_n, \mathbf{s}$ составляют линейно зависимую конечную подсистему векторов в S. Это противоречит линейной независимости системы векторов S. Полученное противоречие доказывает минимальность S. Теорема доказана. $\square$

Определение 4.4. Линейное векторное пространство V называется *конечномерным*, если в нем существует порождающая система $S = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, состоящая из конечного числа векторов.

В любом линейном векторном пространстве существует по меньшей мере одна порождающая система векторов, например, $S = V$. Однако, вопрос о существовании минимальных порождающих систем векторов нетривиален. В случае произвольных линейных векторных пространств он решается положительно, но доказательство этого факта неэлементарно и неконструктивно. Оно основывается на *аксиоме выбора* (см. [1]). Класс конечномерных линейных векторных пространств выделен тем, что для него доказательство существования минимальных порождающих систем векторов осуществляется элементарными средствами.

Теорема 4.4. *Пусть V — конечномерное линейное векторное пространство. Тогда в нем существует, по меньшей мере, одна минимальная порождающая система векторов. Число векторов в любых двух минимальных порождающих системах векторов $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ и $\{\mathbf{y}_1, \ldots, \mathbf{y}_m\}$ одинаково. Оно называется размерностью пространства V и обозначается $m = \dim V$.*

Док-во. Пусть $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ — некоторая конечная порождающая система векторов в конечномерном линейном векторном пространстве V. Если эта система не минимальна, то она линейно зависима, и один из векторов линейно выражается через остальные. Исключив этот вектор из системы, мы получим конечную порождающую систему векторов, число векторов в которой на единицу меньше, чем в исходной. Если полученная система также не минимальна, повторим процедуру исключения вектора. В силу конечности общего числа векторов в исходной системе после некоторого числа исключений, мы получим линейно независимую систему векторов, которая и будет минимальной.

Пусть $\{\mathbf{y}_1, \ldots, \mathbf{y}_m\}$ — минимальная порождающая система векторов в V и пусть S — некоторая другая минимальная порождающая система пространства V. Система векторов S линейно независима и каждый вектор из S линейно выражается через вектора $\mathbf{y}_1, \ldots, \mathbf{y}_m$. В силу теоремы Штейница число векторов в S конечно и $n = |S| \leqslant m$. Но любой из векторов $\mathbf{y}_i$ также линейно выражается через вектора системы $S = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$. Из той же теоремы Штейница заключаем: $m \leqslant n$. Из полученных двух неравенств вытекает равенство $m = n$. Теорема доказана. □

Размерность $\dim V$ является целочисленным числовым инвариантом конечномерного пространства V. Конечномерное пространство размерности m называется m-мерным. Возвращаясь к примерам линейных векторных пространств из параграфа 2, отметим, что $\dim \mathbb{R}^n = n$, а функциональное пространство $C^m([-1,1])$ не является конечномерным.

Теорема 4.5. *Пусть V — конечномерное линейное векторное пространство. Тогда имеют место следующие факты:*

1. *количество векторов во всякой линейно независимой системе векторов $\mathbf{x}_1, \ldots, \mathbf{x}_k$ пространства V не превосходит его размерности;*
2. *всякое подпространство U конечномерного пространства V конечномерно, причем $\dim U \leqslant \dim V$;*
3. *из совпадения размерностей $\dim U = \dim V$ для подпространства $U \subset V$ вытекает совпадение $U = V$;*
4. *любая линейно независимая система векторов, число векторов в которой совпадает с $\dim V$, является порождающей системой в V.*

Док-во. Пусть $\dim V = m$. Фиксируем некоторую минимальную порождающую систему векторов $\mathbf{y}_1, \ldots, \mathbf{y}_m$ в V. Всякий вектор линейно независимой системы векторов $\mathbf{x}_1, \ldots, \mathbf{x}_k$ линейно выражается через вектора $\mathbf{y}_1, \ldots, \mathbf{y}_m$. Применяя в этой ситуации теорему Штейница, заключаем, что $k \leqslant m$. Первый пункт теоремы доказан.

Пусть U — подпространство в V. Рассмотрим всевозможные линейно независимые системы $\mathbf{u}_1, \ldots, \mathbf{u}_k$, составленные из векторов подпространства U. В силу уже доказанного пункта (1) теоремы число векторов в таких системах ограничено, оно не превосходит $\dim V$. Поэтому можно подобрать систему $\mathbf{u}_1, \ldots, \mathbf{u}_k$, число векторов в которой является максимально возможным: $k = k_{\max} \leqslant \dim V$. Пусть $\mathbf{u}$ — некоторый произвольный вектор из подпространства U. Добавление его к системе $\mathbf{u}_1, \ldots, \mathbf{u}_k$ делает ее линейно зависимой (это следует из максимальности k). Применив свойство (4) из теоремы 3.1, заключаем, что вектор u линейно выражается через вектора $\mathbf{u}_1, \ldots, \mathbf{u}_k$. Значит, $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ есть порождающая система векторов в U, состоящая из конечного числа векторов. Ее минимальность вытекает из линейной независимости. Конечномерность подпространства U доказана. Оценка для его размерности вытекает из неравенства $\dim U = k \leqslant m = \dim V$.

Пусть вновь U — подпространство в V. Рассмотрим случай совпадения размерностей $\dim U = \dim V$. Выберем некоторую минимальную порождающую систему векторов $\mathbf{u}_1, \ldots, \mathbf{u}_m$ в U. Она линейно независима. Добавление произвольного вектора $\mathbf{v} \in V$ к ней делает ее линейно зависимой, ибо в V не может быть линейно независимых систем из $(m+1)$-го вектора. Применив свойство (4) из теоремы 3.1, заключаем, что вектор $\mathbf{v}$ линейно выражается через систему векторов $\mathbf{u}_1, \ldots, \mathbf{u}_m$:

$$\mathbf{v} = \alpha_1 \cdot \mathbf{u}_1 + \ldots + \alpha_m \cdot \mathbf{u}_m.$$

Из этой формулы немедленно получаем $\mathbf{v} \in U$. Ввиду произвольности выбора $\mathbf{v} \in V$ мы получаем $U = V$. Третий пункт теоремы доказан.

Пусть $\mathbf{x}_1, \ldots, \mathbf{x}_m$ — линейно независимая система векторов из V, число векторов m в которой совпадает с размерностью пространства V. Обозначим через U линейную оболочку системы векторов $\mathbf{x}_1, \ldots, \mathbf{x}_m$. Эти вектора порождают U, ввиду линейной независимости они образуют минимальную порождающую систему в U. Значит, $\dim U = m = \dim V$. Применяя уже доказанный пункт (3) теоремы, получаем

$$\langle\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}\rangle = U = V.$$

Таким образом, вектора $\mathbf{x}_1, \ldots, \mathbf{x}_m$ образуют порождающую систему в V. Теорема доказана. $\square$

Определение 4.5. Упорядоченная минимальная порождающая система векторов $\mathbf{e}_1, \ldots, \mathbf{e}_n$, т. е. та, в которой зафиксирован порядок следования векторов, называется *базисом* конечномерного векторного пространства V.

Теорема 4.6 (критерий базиса). *Упорядоченная система из n векторов $\mathbf{e}_1, \ldots, \mathbf{e}_n$ является базисом в пространства V тогда и только тогда, когда*

(1) *вектора $\mathbf{e}_1, \ldots, \mathbf{e}_n$ линейно независимы;*
(2) *любой вектор пространства V линейно выражается через $\mathbf{e}_1, \ldots, \mathbf{e}_n$.*

Доказательство теоремы очевидно. Условие (2) теоремы 4.6 означает, что система векторов $\mathbf{e}_1, \ldots, \mathbf{e}_n$ является порождающей. А условие (1) эквивалентно ее минимальности.

Теорема 4.6 является, по существу, переформулировкой определения 4.5. Мы сформулировали ее исключительно для упрощения терминологии. Термины «порождающая система» и «минимальная порождающая система» достаточно громоздки для частого использования.

Теорема 4.7. *Пусть* $\mathbf{e}_1, \ldots, \mathbf{e}_s$ *— базис в подпространстве* $U \subset V$ *и пусть* $\mathbf{v} \in V$ *— некоторый вектор, не принадлежащий подпространству* U*. Тогда система векторов* $\mathbf{e}_1, \ldots, \mathbf{e}_s, \mathbf{v}$ *линейно независима.*

Док-во. Действительно, если система векторов $\mathbf{e}_1, \ldots, \mathbf{e}_s$, $\mathbf{v}$, полученная добавлением вектора $\mathbf{v}$ к $\mathbf{e}_1, \ldots, \mathbf{e}_s$, линейно зависима, то мы оказываемся в условиях пункта (4) теоремы 3.1. Тогда $\mathbf{v}$ линейно выражается через вектора $\mathbf{e}_1, \ldots, \mathbf{e}_s$, что противоречит условию $\mathbf{v} \notin U$. Полученное противоречие доказывает теорему. $\square$

Теорема 4.8 (о дополнении базиса). *Пусть* U *— подпространство в конечномерном линейном векторном пространстве* V*. Тогда произвольный базис* $\mathbf{e}_1, \ldots, \mathbf{e}_s$ *из* U *может быть дополнен до базиса* $\mathbf{e}_1, \ldots, \mathbf{e}_s, \mathbf{e}_{s+1}, \ldots, \mathbf{e}_n$ *в* V*.*

Док-во. Обозначим $U = U_0$. Если $U_0 = V$, то дополнять базис нет необходимости — базис $\mathbf{e}_1, \ldots, \mathbf{e}_s$ является базисом в V. Если же $U_0 \neq V$, то обозначим через $\mathbf{e}_{s+1}$ некоторый произвольный вектор из V, не содержащийся в U_0. По теореме 4.7 вектора $\mathbf{e}_1, \ldots, \mathbf{e}_s, \mathbf{e}_{s+1}$ линейно независимы.

Обозначим через U_1 линейную оболочку векторов $\mathbf{e}_1, \ldots, \mathbf{e}_s, \mathbf{e}_{s+1}$. Для подпространства U_1 также возникает альтернатива: $U_1 = V$ либо $U_1 \neq V$. В первом случае процесс дополнения базиса завершен. Во втором случае он продолжается и приводит к построению подпространства U_2 с базисом $\mathbf{e}_1, \ldots, \mathbf{e}_s, \mathbf{e}_{s+1}, \mathbf{e}_{s+2}$. В результате продолжения этого процесса мы получим цепочку вложенных друг в друга подпространств

$$U_0 \subsetneq U_1 \subsetneq U_2 \subsetneq \ldots .$$

Такая цепочка подпространств не может быть бесконечной, ибо каждое следующее подпространство имеет на единицу большую размерность, чем предыдущее. А их размерность ограничена сверху размерностью V. На $(n-s)$-ом шаге этот процесс завершится совпадением $U_{n-s} = V$ и мы получим базис $\mathbf{e}_1, \ldots, \mathbf{e}_s, \mathbf{e}_{s+1}, \ldots, \mathbf{e}_n$ в V, дополняющий базис $\mathbf{e}_1, \ldots, \mathbf{e}_s$ из U. $\square$

§ 5. Координаты. Преобразование координат векторов при замене базиса.

Пусть V — некоторое конечномерное линейное векторное пространство над полем $\mathbb{K}$ и пусть $\dim V = n$. На протяжении этого параграфа мы рассматриваем только конечномерные пространства. Выберем некоторый базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в V. Система векторов $\mathbf{e}_1, \ldots, \mathbf{e}_n$ порождает V, поэтому любой вектор $\mathbf{x}$ из пространства V может быть представлен в виде линейной комбинации базисных векторов $\mathbf{e}_1, \ldots, \mathbf{e}_n$:

$$\mathbf{x} = x^1 \cdot \mathbf{e}_1 + \ldots + x^n \cdot \mathbf{e}_n. \tag{5.1}$$

Линейная комбинация (5.1) называется *разложением* вектора $\mathbf{x}$ по базису $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Ее коэффициенты $\mathbf{x}^1, \ldots, \mathbf{x}^n$ — это числа из поля $\mathbb{K}$. Они называются *компонентами* или *координатами* вектора $\mathbf{x}$ в этом базисе.

Буквенные обозначения для координат вектора $\mathbf{x}$ в (5.1) мы маркируем верхними индексами. Такое обозначение для координат векторов обусловлено специальным соглашением, известным как *тензорная нотация*. Тензорная нотация была введена для упрощения громоздких вычислений в дифференциальной геометрии и в общей теории относительности (см. [2] и [3]). С другими правилами тензорной нотации можно познакомиться при рассмотрении координатной теории тензоров (см. [7][1]).

Теорема 5.1. *Для всякого вектора $\mathbf{x} \in V$ его разложение по базису пространства V однозначно определено.*

Док-во. Существование разложения (5.1) для вектора $\mathbf{x}$ вытекает из пункта (2) теоремы 4.7. Допустим, что существует некоторое другое разложение

$$\mathbf{x} = {x'}^1 \cdot \mathbf{e}_1 + \ldots + {x'}^n \cdot \mathbf{e}_n. \tag{5.2}$$

Вычитая (5.1) из этого равенства, получим соотношение

$$({x'}^1 - x^1) \cdot \mathbf{e}_1 + \ldots + ({x'}^n - x^n) \cdot \mathbf{e}_n = \mathbf{0}. \tag{5.3}$$

В силу линейной независимости базисных векторов $\mathbf{e}_1, \ldots, \mathbf{e}_n$ из равенства нулю линейной комбинации (5.3) вытекает ее тривиальность, т. е. имеют место равенства ${x'}^i - x^i = 0$. Тогда из них мы выводим

$${x'}^1 = x^1, \ \ldots, \ {x'}^n = x^n.$$

Значит, разложения (5.1) и (5.2) совпадают. Однозначность разложения векторов по базису доказана. $\square$

Фиксировав некоторый базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в пространстве V и разложив вектор $\mathbf{x}$ по этому базису, мы можем записать его координаты в вектор-столбцы. В силу теоремы 5.1 это определяет взаимно-однозначное отображение $\psi\colon V \to \mathbb{K}^n$. При этом нетрудно проверить, что

$$\psi(\mathbf{x}+\mathbf{y}) = \begin{Vmatrix} x^1 + y^1 \\ \vdots \\ x^n + y^n \end{Vmatrix}, \qquad \psi(\alpha \cdot \mathbf{x}) = \begin{Vmatrix} \alpha \cdot x^1 \\ \vdots \\ \alpha \cdot x^n \end{Vmatrix}. \tag{5.4}$$

Формулы (5.4) показывают, что базис — это очень удобное средство работы с векторами. Алгебраические операции с векторами заменяются алгебраическими операциями с их координатами. Однако, у координатного подхода есть один недостаток. Отображение ψ существенно зависит от выбора базиса. А канонического выбора базиса нет: в общем случае ни один из базисов не является предпочтительным перед другим. Это приводит к необходимости рассмотрения различных базисов и пересчету координат векторов при замене базиса.

[1] Ссылка [7] добавлена в 2004 году при переводе книги на английский язык.

Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ — два базиса в линейном векторном пространстве V. Эти базисы мы будем называть *«волнистым»* и *«неволнистым»*. Первый неволнистый базис будем называть исходным или *старым* базисом, а второй — *новым* базисом. Выбрав i-ый вектор нового (волнистого) базиса, разложим его по векторам старого базиса:

$$\tilde{\mathbf{e}}_i = S_i^1 \cdot \mathbf{e}_1 + \ldots + S_i^n \cdot \mathbf{e}_n. \tag{5.5}$$

В соответствии с тензорной нотацией координаты вектора $\tilde{\mathbf{e}}_i$ в разложении (5.5) маркируются верхним индексом. Нижний индекс i есть просто номер раскладываемого вектора $\tilde{\mathbf{e}}_i$. Всего разложения (5.5) однозначно определяют n^2 чисел S_i^j, которые принято записывать в виде матрицы:

$$S = \begin{Vmatrix} S_1^1 & \ldots & S_n^1 \\ \vdots & \ddots & \vdots \\ S_1^n & \ldots & S_n^n \end{Vmatrix}. \tag{5.6}$$

Верхний индекс элемента S_i^j — это номер строки, нижний индекс — номер столбца. Матрица S в (5.6) называется *матрицей перехода* из старого базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в новый базис $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$.

Поменяв базисы $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ ролями, мы можем записать разложение вектора $\mathbf{e}_j$ по волнистому базису:

$$\mathbf{e}_j = T_j^1 \cdot \tilde{\mathbf{e}}_1 + \ldots + T_j^n \cdot \tilde{\mathbf{e}}_n. \tag{5.7}$$

Коэффициенты разложений (5.7) определяют матрицу T, которую принято называть *матрицей обратного перехода.* Разумеется, понятие прямого и обратного перехода относительны: они зависят от того, какой из базисов считать старым или исходным, а какой — новым.

Теорема 5.2. *Матрица прямого перехода S и матрица обратного перехода T, определяемые двумя разложениями* (5.5) *и* (5.7), *являются обратными матрицами друг для друга.*

Напомним, что две квадратные матрицы являются обратными друг для друга, если их произведение есть единичная матрица: $S\,T = 1$. Операцию умножения матриц мы здесь не определяем, считая ее известной из курса общей алгебры.

Док-во. Начнем доказательство теоремы с того, что запишем соотношения (5.5) и (5.7) в сокращенном виде:

$$\tilde{\mathbf{e}}_i = \sum_{k=1}^{n} S_i^k \cdot \mathbf{e}_k, \qquad\qquad \mathbf{e}_j = \sum_{i=1}^{n} T_j^i \cdot \tilde{\mathbf{e}}_i. \tag{5.8}$$

Подставим первое соотношение (5.8) во второе. Это дает

$$\mathbf{e}_j = \sum_{i=1}^{n} T_j^i \cdot \left(\sum_{k=1}^{n} S_i^k \cdot \mathbf{e}_k \right) = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} S_i^k\, T_j^i \right) \cdot \mathbf{e}_k. \tag{5.9}$$

Символ δ^k_j, называемый *символом Кронекера*, используется для обозначения следующего набора чисел:

$$\delta^k_j = \begin{cases} 1 & \text{при } k = j, \\ 0 & \text{при } k \neq j. \end{cases} \tag{5.10}$$

Используем символ Кронекера, определенный согласно (5.10), для преобразования левой части равенства (5.9):

$$\mathbf{e}_j = \sum_{k=1}^{n} \delta^k_j \cdot \mathbf{e}_k. \tag{5.11}$$

Соотношение (5.11) и правая часть соотношения (5.9) определяют разложение одного и того же вектора e_j по базису $\mathbf{e}_1, \ldots, \mathbf{e}_n$. В силу теоремы 5.1 об однозначности разложения вектора по базису имеем

$$\sum_{i=1}^{n} S^k_i \, T^i_j = \delta^k_j.$$

Легко видеть, что полученные соотношения в точности эквивалентны матричному равенству $S\,T = 1$. Теорема доказана. $\square$

Следствие. *Матрица прямого перехода S и матрица обратного перехода T невырождены, причем* $\det S \, \det T = 1$.

Док-во. Соотношение $\det S \, \det T = 1$ вытекает из доказанного выше соотношения $S\,T = 1$. Этот факт известен из курса общей алгебры. Но если произведение двух чисел равно 1, то ни одно из них не может быть нулем:

$$\det S \neq 0, \qquad\qquad \det T \neq 0.$$

Это и означает невырожденность матриц S и T. Следствие доказано. $\square$

Теорема 5.3. *Всякая невырожденная матрица S порядка $n \times n$ может быть получена как матрица перехода из некоторого базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в некоторый другой базис $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$.*

Док-во. Выберем базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в V произвольным образом и фиксируем его, после чего определим векторы $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ из соотношений (5.5). Докажем линейную независимость полученных векторов. Для этого рассмотрим линейную комбинацию этих векторов, равную нулю:

$$\alpha^1 \cdot \tilde{\mathbf{e}}_1 + \ldots + \alpha^n \cdot \tilde{\mathbf{e}}_n = \mathbf{0}. \tag{5.12}$$

Подставив (5.5) в это соотношение, можно преобразовать его к виду

$$\left(\sum_{i=1}^{n} S^1_i \, \alpha^i\right) \cdot \mathbf{e}_1 + \ldots + \left(\sum_{i=1}^{n} S^n_i \, \alpha^i\right) \cdot \mathbf{e}_n = \mathbf{0}.$$

Из линейной независимости базисных векторов $\mathbf{e}_1, \ldots, \mathbf{e}_n$ вытекает равенство нулю всех выписанных выше сумм. Расписав эти суммы, мы получим однородную систему линейных уравнений относительно $\alpha^1, \ldots, \alpha^n$:

$$\begin{array}{ccccccc} S^1_1\,\alpha^1 & + & \ldots & + & S^1_n\,\alpha^n & = & 0, \\ \multicolumn{7}{c}{\dotfill} \\ S^n_1\,\alpha^1 & + & \ldots & + & S^n_n\,\alpha^n & = & 0. \end{array}$$

Матрица этой системы уравнений совпадает с S. Из курса алгебры известно, что однородная система линейных уравнений с невырожденной квадратной матрицей S имеет единственное чисто нулевое решение $\alpha^1 = \ldots = \alpha^n = 0$. Значит, равная нулю линейная комбинация (5.12) тривиальна. Пользуясь доказанной линейной независимостью векторов $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$, применим к ним пункт (4) теоремы 4.5. Из него вытекает, что вектора $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ образуют базис, а матрица S служит матрицей перехода из $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в этот базис. Теорема доказана. $\square$

Рассмотрим два базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ в пространстве V, связанных матрицей перехода S. Пусть $\mathbf{x}$ — некоторый произвольный вектор пространства V. Он может быть разложен в каждом из этих базисов

$$\mathbf{x} = \sum_{k=1}^{n} x^k \cdot \mathbf{e}_k, \qquad\qquad \mathbf{x} = \sum_{i=1}^{n} \tilde{x}^i \cdot \tilde{\mathbf{e}}_i. \tag{5.13}$$

При этом задание координат вектора в одном базисе определяет этот вектор, а значит, определяет и его координаты в другом базисе.

Теорем 5.4. *Связь между координатами вектора* $\mathbf{x}$ *в базисах* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *и* $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$, *связанных матрицей перехода* S, *дается формулами*

$$x^k = \sum_{i=1}^{n} S^k_i\,\tilde{x}^i, \qquad\qquad \tilde{x}^i = \sum_{k=1}^{n} T^i_k\,x^k. \tag{5.14}$$

Соотношения (5.14) известны как *формулы преобразования координат вектора при замене базиса.*

Док-во. Для доказательства первого из соотношений (5.14) подставим во второе соотношение (5.13) разложение вектора $\tilde{\mathbf{e}}_i$ по базису без волны в форме (5.8). В результате такой подстановки получим

$$\mathbf{x} = \sum_{i=1}^{n} \tilde{x}^i \cdot \left(\sum_{k=1}^{n} S^k_i \cdot \mathbf{e}_k\right) = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} S^k_i\,\tilde{x}^i\right) \cdot \mathbf{e}_k.$$

Сравнив полученное разложение вектора $\mathbf{x}$ с первым разложением (5.13), в силу теоремы об однозначности разложения вектора по базису выводим:

$$x^k = \sum_{i=1}^{n} S^k_i\,\tilde{x}^i.$$

Это в точности первая из формул преобразования (5.14). Вторая формула (5.14) доказывается аналогично. $\square$

§ 6. Пересечения и суммы подпространств.

Пусть задано некоторое количество подпространств в линейном векторном пространстве V. Для обозначения этого факта запишем $U_i \subset V$, где $i \in I$. Число подпространств может быть конечным или счетным, тогда их можно пронумеровать натуральными числами. Однако, в общем случае подпространства приходится индексировать элементами некоторого индексного множества I, которое может быть конечным, счетным или даже несчетным. Обозначим через U и S пересечение и объединение всех подпространств:

$$U = \bigcap_{i \in I} U_i, \qquad\qquad S = \bigcup_{i \in I} U_i. \tag{6.1}$$

Теорема 6.1. *Пересечение любого числа подпространств линейного векторного пространства V есть подпространство в пространстве V.*

Док-во. Множество U в (6.1) непусто, ибо нулевой вектор содержится в каждом из U_i. Проверим для U условия (1) и (2) из определения 2.2.

Пусть вектора $\mathbf{u}_1$, $\mathbf{u}_2$ и $\mathbf{u}$ принадлежат U. Тогда они принадлежат U_i для каждого $i \in I$. Но U_i есть подпространство, поэтому $\mathbf{u}_1 + \mathbf{u}_2 \in U_i$ и $\alpha \cdot \mathbf{u} \in U_i$ для всех $i \in I$. Отсюда $\mathbf{u}_1 + \mathbf{u}_2 \in U$ и $\alpha \cdot \mathbf{u} \in U$. Теорема доказана. $\square$

Множество S в (6.1), вообще говоря, подпространством не является. Поэтому на базе S вводится следующее понятие.

Определение 6.1. *Суммой* подпространств U_i, $i \in I$, называется линейная оболочка объединения этих подпространств.

Для обозначения суммы подпространств $W = \langle S \rangle$ используем знак суммы:

$$W = \left\langle \bigcup_{i \in I} U_i \right\rangle = \sum_{i \in I} U_i.$$

Теорема 6.2. *Вектор $\mathbf{w}$ принадлежит сумме семейства подпространств $U_i \subset V$, $i \in I$ тогда и только тогда, когда он представляется в виде суммы конечного числа векторов, взятых из подпространств U_i этого семейства:*

$$\mathbf{w} = \mathbf{u}_{i_1} + \ldots + \mathbf{u}_{i_k}, \text{ где } \mathbf{u}_i \in U_i. \tag{6.2}$$

Док-во. Пусть S — объединение подпространств семейства $U_i \subset V$, $i \in I$. Тогда их сумма — это линейная оболочка $W = \langle S \rangle$. Пусть $\mathbf{w} \in W$, тогда $\mathbf{w}$ есть линейная комбинация некоторого конечного числа векторов из S:

$$\mathbf{w} = \alpha_1 \cdot \mathbf{s}_1 + \ldots + \alpha_k \cdot \mathbf{s}_k.$$

Но S есть объединение подпространств U_i, поэтому $\mathbf{s}_m$ принадлежит U_{i_m} и $\alpha_m \cdot \mathbf{s}_m = \mathbf{u}_{i_m} \in U_{i_m}$, где $m = 1, \ldots, k$. Это приводит к соотношению (6.2) для нашего вектора $\mathbf{w}$.

Теперь, наоборот, пусть вектор $\mathbf{w}$ задан соотношением (6.2). Тогда для $\mathbf{u}_{i_m}$ мы имеем $\mathbf{u}_{i_m} \in U_{i_m} \subset S$. Поэтому вектор $\mathbf{w}$ принадлежит линейной оболочке множества S. Теорема доказана. $\square$

Определение 6.2. Сумма W семейства подпространств U_i, $i \in I$, называется *прямой суммой*, если для каждого вектора $\mathbf{w} \in W$ разложение (6.2) единственно. Для прямой суммы используется специальное обозначение:

$$W = \bigoplus_{i \in I} U_i.$$

Теорема 6.3. *Пусть $W = U_1 + \ldots + U_k$ есть сумма конечного числа конечномерных подпространств. Размерность W равна сумме размерностей подпространств U_i тогда и только тогда, когда эта сумма прямая: $W = U_1 \oplus \ldots \oplus U_k$.*

Док-во. Выберем базис в каждом из подпространств. Пусть $\dim U_i = s_i$ и пусть $\mathbf{e}_{i\,1}, \ldots, \mathbf{e}_{i\,s_i}$ — базис в U_i. Объединим все базисные векторы $\mathbf{e}_{i\,m}$ в одну систему векторов, упорядоченную в алфавитном порядке:

$$\mathbf{e}_{1\,1}, \ldots, \mathbf{e}_{1\,s_1}, \quad \ldots, \quad \mathbf{e}_{k\,1}, \ldots, \mathbf{e}_{k\,s_k}. \tag{6.3}$$

Система векторов (6.3) является порождающей в W. В силу $W = U_1 + \ldots + U_k$ для произвольного вектора $\mathbf{w} \in W$ имеется разложение вида (6.2):

$$\mathbf{w} = \mathbf{u}_1 + \ldots + \mathbf{u}_k, \text{ где } \mathbf{u}_i \in U_i. \tag{6.4}$$

Разложив каждый из векторов $\mathbf{u}_i$ в (6.4) по базису в соответствующем подпространстве U_i, получим разложение вектора $\mathbf{w}$ по векторам системы (6.3). Значит, вектора (6.3) порождают подпространство W.

Если $\dim W = \dim U_1 + \ldots + \dim U_k$, то число векторов в порождающей системе (6.3) является минимальным. Поэтому (6.3) — это базис в W. Из всякого разложения (6.4) выводится разложение вектора w по базису (6.3):

$$\mathbf{w} = \left(\sum_{j=1}^{s_1} \alpha_{1\,j} \cdot \mathbf{e}_{1\,j}\right) + \ldots + \left(\sum_{j=1}^{s_k} \alpha_{k\,j} \cdot \mathbf{e}_{k\,j}\right). \tag{6.5}$$

Суммы в круглых скобках при этом определяются разложениями векторов $\mathbf{u}_1, \ldots, \mathbf{u}_k$ по базисам в соответствующих подпространствах $U_1, \ldots, U_k$:

$$\mathbf{u}_i = \sum_{j=1}^{s_i} \alpha_{i\,j} \cdot \mathbf{e}_{i\,j}. \tag{6.6}$$

В силу (6.6) существование двух различных разложений (6.4) для некоторого вектора w означало бы существование двух различных разложений (6.5) этого вектора в базисе (6.3). Значит, разложение (6.4) единственно и сумма подпространств $W = U_1 + \ldots + U_k$ является прямой.

Пусть, наоборот, $W = U_1 \oplus \ldots \oplus U_k$. Вектора (6.3) порождают W. Докажем линейную независимость этих векторов. Для этого рассмотрим их линейную комбинацию и приравняем ее к нулю:

$$\mathbf{0} = \left(\sum_{j=1}^{s_1} \alpha_{1\,j} \cdot \mathbf{e}_{1\,j}\right) + \ldots + \left(\sum_{j=1}^{s_k} \alpha_{k\,j} \cdot \mathbf{e}_{k\,j}\right). \tag{6.7}$$

Обозначим через $\tilde{\mathbf{u}}_1, \ldots, \tilde{\mathbf{u}}_k$ значения сумм, заключенных в круглые скобки в (6.7). Легко видеть, что $\tilde{\mathbf{u}}_i \in U_i$, поэтому (6.7) есть разложение (6.4) для $\mathbf{w} = 0$. Но $\mathbf{0} = \mathbf{0} + \ldots + \mathbf{0}$ и $\mathbf{0} \in U_i$. Сумма подпространств $U_1, \ldots, U_k$ прямая, следовательно, это разложение нуля должно совпасть с (6.7). Отсюда

$$\mathbf{0} = \sum_{j=1}^{s_i} \alpha_{i\,j} \cdot \mathbf{e}_{i\,j} \quad \text{для всех} \quad i = 1, \ldots, k.$$

Полученные соотношения — это разложения нулевого вектора по базисам в подпространствах U_i. Значит, $\alpha_{i\,j} = 0$. Линейная независимость векторов (6.3) доказана. Для определения размерности W посчитаем число векторов $\mathbf{e}_{i\,j}$ в системе (6.3): $\dim W = s_1 + \ldots + s_k = \dim U_1 + \ldots + \dim U_k$. $\square$

Замечание. Если сумма подпространств $U_1, \ldots, U_k$ не является прямой, вектора (6.3) по прежнему порождают их сумму W, но они уже не являются линейно независимыми. Поэтому получается неравенство:

$$\dim W < \dim U_1 + \ldots + \dim U_k. \tag{6.8}$$

Уточнение этого неравенства в общем случае достаточно сложно. Мы уточним его для случая суммы двух подпространств.

Теорема 6.4. *Размерность суммы двух произвольных конечномерных подпространств U_1 и U_2 в линейном векторном пространстве V равна сумме их размерностей за вычетом размерности их пересечения:*

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2). \tag{6.9}$$

Док-во. Из включения $U_1 \cap U_2 \subset U_1$ и из неравенства (6.8) заключаем, что все рассматриваемые в теореме подпространства конечномерны. Положим $\dim(U_1 \cap U_2) = s$ и выберем базис $\mathbf{e}_1, \ldots, \mathbf{e}_s$ в пересечении $U_1 \cap U_2$.

Используя включение $U_1 \cap U_2 \subset U_1$ применим теорему 4.8 о дополнении базиса. Это позволяет дополнить базис $\mathbf{e}_1, \ldots, \mathbf{e}_s$ в пересечении $U_1 \cap U_2$ до базиса $\mathbf{e}_1, \ldots, \mathbf{e}_s, \mathbf{e}_{s+1}, \ldots, \mathbf{e}_{s+p}$ в U_1. Тогда $\dim U_1 = s + p$. Аналогичным образом, в силу включения $U_1 \cap U_2 \subset U_2$ определяется базис $\mathbf{e}_1, \ldots, \mathbf{e}_s, \mathbf{e}_{s+p+1}, \ldots, \mathbf{e}_{s+p+q}$ в U_2, который дополняет базис из пересечения $U_1 \cap U_2$. Для размерности U_2 это дает $\dim U_2 = s + q$.

Объединим два построенных базиса и рассмотрим вектора

$$\mathbf{e}_1, \ldots, \mathbf{e}_s, \mathbf{e}_{s+1}, \ldots, \mathbf{e}_{s+p}, \mathbf{e}_{s+p+1}, \ldots, \mathbf{e}_{s+p+q}. \tag{6.10}$$

Докажем, что вектора (6.10) образуют базис в сумме подпространств $U_1 + U_2$. Пусть $\mathbf{w}$ некоторый произвольный вектор в $U_1 + U_2$. Соотношение (6.2) для этого вектора имеет вид: $\mathbf{w} = \mathbf{u}_1 + \mathbf{u}_2$. Разложим вектора $\mathbf{u}_1$ и $\mathbf{u}_2$ по построенным базисам в пространствах U_1 и U_2:

$$\mathbf{u}_1 = \sum_{i=1}^{s} \alpha_i \cdot \mathbf{e}_i + \sum_{j=1}^{p} \beta_{s+j} \cdot \mathbf{e}_{s+j},$$
$$\mathbf{u}_2 = \sum_{i=1}^{s} \tilde{\alpha}_i \cdot \mathbf{e}_i + \sum_{j=1}^{q} \gamma_{s+p+j} \cdot \mathbf{e}_{s+p+j}.$$

Сложив эти два разложения, мы получим, что $\mathbf{w}$ линейно выражается через вектора (6.10). Значит, вектора (6.10) порождают подпространство $U_1 + U_2$.

Для доказательства линейной независимости системы векторов (6.10) рассмотрим их линейную комбинацию, равную нулю:

$$\sum_{i=1}^{s+p} \alpha_i \cdot \mathbf{e}_i + \sum_{i=1}^{q} \alpha_{s+p+i} \cdot \mathbf{e}_{s+p+i} = \mathbf{0}. \tag{6.11}$$

Преобразуем это выражение, перенеся вторую сумму в правую часть:

$$\sum_{i=1}^{s+p} \alpha_i \cdot \mathbf{e}_i = -\sum_{i=1}^{q} \alpha_{s+p+i} \cdot \mathbf{e}_{s+p+i}.$$

Обозначим через $\mathbf{u}$ значение левой и правой частей этого равенства. Тогда для вектора $\mathbf{u}$ получаем следующие формулы:

$$\mathbf{u} = \sum_{i=1}^{s+p} \alpha_i \cdot \mathbf{e}_i, \qquad\qquad \mathbf{u} = -\sum_{i=1}^{q} \alpha_{s+p+i} \cdot \mathbf{e}_{s+p+i}. \tag{6.12}$$

В силу первого из этих равенств $\mathbf{u} \in U_1$, в силу второго — этот вектор принадлежит U_2. Отсюда $\mathbf{u} \in U_1 \cap U_2$. Значит, он раскладывается по базису $\mathbf{e}_1, \ldots, \mathbf{e}_s$ в пересечении двух подпространств U_1 и U_2:

$$\mathbf{u} = \sum_{i=1}^{s} \beta_i \cdot \mathbf{e}_i. \tag{6.13}$$

Из сравнения этого разложения со вторым разложением из (6.12) получаем:

$$\sum_{i=1}^{s} \beta_i \cdot \mathbf{e}_i + \sum_{i=1}^{q} \alpha_{s+p+i} \cdot \mathbf{e}_{s+p+i} = \mathbf{0}. \tag{6.14}$$

Но вектора $\mathbf{e}_1, \ldots, \mathbf{e}_s, \mathbf{e}_{s+p+1}, \ldots, \mathbf{e}_{s+p+q}$ составляют базис U_2. Они линейно независимы. Поэтому все коэффициенты в (6.14) равны нулю. В частности:

$$\alpha_{s+p+1} = \ldots = \alpha_{s+p+q} = 0. \tag{6.15}$$

Подставив $\beta_1 = \ldots = \beta_s = 0$ в (6.13) получаем $\mathbf{u} = 0$. Теперь из первого разложения (6.12) выводим следующую формулу:

$$\sum_{i=1}^{s+p} \alpha_i \cdot \mathbf{e}_i = \mathbf{0}.$$

Из линейной независимости $\mathbf{e}_1, \ldots, \mathbf{e}_s, \mathbf{e}_{s+1}, \ldots, \mathbf{e}_{s+p}$ вытекает зануление всех коэффициентов выписанном выше разложении:

$$\alpha_1 = \ldots = \alpha_s = \alpha_{s+1} = \ldots = \alpha_{s+p} = 0. \tag{6.16}$$

Объединив (6.15) и (6.16), получаем тривиальность линейной комбинации (6.11). Значит, вектора (6.10) линейно независимы и образуют базис в $U_1 + U_2$. Для размерности этого подпространства имеем

$$\begin{aligned}\dim(U_1 + U_2) &= s + p + q = (s + p) + (s + q) - s = \\ &= \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2).\end{aligned}$$

Соотношение (6.9) и вся теорема 6.4 в целом доказаны. □

§ 7. Смежные классы по подпространству. Понятие факторпространства.

Пусть V — некоторое линейное векторное пространство и пусть U — подпространство в нем. *Смежным классом* вектора $\mathbf{v} \in V$ по подпространству U называется следующее множество:

$$\mathrm{Cl}_U(\mathbf{v}) = \{\mathbf{w} \in V\colon\ \mathbf{w} - \mathbf{v} \in U\}. \tag{7.1}$$

Вектор $\mathbf{v}$ называется *определяющим вектором* или *представителем* класса (7.1). Класс $\mathrm{Cl}_U(\mathbf{v})$ устроен очень просто. Он получается добавлением вектора $\mathbf{v}$ к каждому вектору $\mathbf{u}$ из подпространства U. Особенно просто устроен класс нулевого вектора: $\mathrm{Cl}_U(\mathbf{0}) = U$. Он называется *нулевым* классом.

Теорема 7.1. *Смежные классы по подпространству $U \subset V$ обладают следующими свойствами:*

(1) $\mathbf{a} \in \mathrm{Cl}_U(\mathbf{a})$;
(2) *если* $\mathbf{a} \in \mathrm{Cl}_U(\mathbf{b})$, *то* $\mathbf{b} \in \mathrm{Cl}_U(\mathbf{a})$;
(3) *если* $\mathbf{a} \in \mathrm{Cl}_U(\mathbf{b})$ *и* $\mathbf{b} \in \mathrm{Cl}_U(\mathbf{c})$, *то* $\mathbf{a} \in \mathrm{Cl}_U(\mathbf{c})$.

Док-во. Первый пункт теоремы очевиден. Действительно, разность $\mathbf{a} - \mathbf{a} = \mathbf{0}$ есть вектор из подпространства U. Отсюда в силу определения (7.1) имеем $\mathbf{a} \in \mathrm{Cl}_U(\mathbf{a})$.

Пусть $\mathbf{a} \in \mathrm{Cl}_U(\mathbf{b})$. Тогда $\mathbf{a} - \mathbf{b} \in U$. Но $\mathbf{b} - \mathbf{a} = (-1) \cdot (\mathbf{a} - \mathbf{b})$, следовательно, $\mathbf{b} - \mathbf{a} \in U$ и $\mathbf{b} \in \mathrm{Cl}_U(\mathbf{a})$. Второй пункт теоремы доказан.

Пусть $\mathbf{a} \in \mathrm{Cl}_U(\mathbf{b})$ и $\mathbf{b} \in \mathrm{Cl}_U(\mathbf{c})$. Тогда $\mathbf{a} - \mathbf{b} \in U$ и $\mathbf{b} - \mathbf{c} \in U$. Но $\mathbf{a} - \mathbf{c} = (\mathbf{a} - \mathbf{b}) + (\mathbf{b} - \mathbf{c})$, значит, $\mathbf{a} - \mathbf{c} \in U$ и $\mathbf{a} \in \mathrm{Cl}_U(\mathbf{c})$. Это завершает доказательство теоремы. □

Пусть $\mathbf{a} \in \mathrm{Cl}_U(b)$. Это условие определяет некоторую взаимосвязь или зависимость между векторами $\mathbf{a}$ и $\mathbf{b}$. Такая взаимосвязь не является жесткой: условие $\mathbf{a} \in \mathrm{Cl}_U(\mathbf{b})$ не исключает $\mathbf{a}' \in \mathrm{Cl}_U(\mathbf{b})$ для некоторого другого вектора $\mathbf{a}'$. Подобные нежесткие связи в математике описываются понятием *бинарного отношения*. Подробнее о бинарных отношениях можно прочесть в книгах [1] и [4]. Запишем $\mathbf{a} \sim \mathbf{b}$ в качестве сокращения записи $\mathbf{a} \in \mathrm{Cl}_U(\mathbf{b})$. Теорема 7.1 определяет следующие свойства введенного таким способом бинарного отношения:

(1) *рефлексивность*: $\mathbf{a} \sim \mathbf{a}$;
(2) *симметричность*: $\mathbf{a} \sim \mathbf{b}$ влечет $\mathbf{b} \sim \mathbf{a}$;
(3) *транзитивность*: $\mathbf{a} \sim \mathbf{b}$ и $\mathbf{b} \sim \mathbf{c}$ влечет $\mathbf{a} \sim \mathbf{c}$.

Бинарные отношения, обладающие свойствами рефлексивности, симметричности и транзитивности, называются *отношениями эквивалентности*. Отношения эквивалентности определяют разбиение множества V, на котором они заданы, в объединение попарно не пересекающихся подмножеств, называемых *классами эквивалентности*:

$$\mathrm{Cl}(\mathbf{v}) = \{\mathbf{w} \in V\colon\ \mathbf{w} \sim \mathbf{v}\}. \tag{7.2}$$

В рассматриваемом нами конкретном случае определение (7.2) в точности совпадает с определением (7.1). Для соблюдения замкнутости изложения мы не будем более использовать запись $\mathbf{a} \sim \mathbf{b}$ вместо $\mathbf{a} \in \mathrm{Cl}_U(\mathbf{b})$ и не будем ссылаться на теорию бинарных отношений (хотя она и проста и широко известна). Вместо этого выведем результат о разбиении V на попарно не пересекающиеся смежные классы из следующей теоремы.

Теорема 7.2. *Если два смежных класса* $\mathrm{Cl}_U(\mathbf{a})$ *и* $\mathrm{Cl}_U(\mathbf{b})$ *пересекаются, то они совпадают.*

Док-во. Действительно, пусть пересечение указанных в теореме классов непусто. Тогда найдется элемент $\mathbf{c}$, принадлежащий им обоим: $\mathbf{c} \in \mathrm{Cl}_U(\mathbf{a})$ и $\mathbf{c} \in \mathrm{Cl}_U(b)$. Из $\mathbf{c} \in \mathrm{Cl}_U(\mathbf{b})$ в силу пункта (2) теоремы 7.1 выводим $\mathbf{b} \in \mathrm{Cl}_U(\mathbf{c})$. Соединив $\mathbf{b} \in \mathrm{Cl}_U(\mathbf{c})$ и $\mathbf{c} \in \mathrm{Cl}_U(\mathbf{a})$, в силу пункта (3) теоремы 7.1 получаем $\mathbf{b} \in \mathrm{Cl}_U(\mathbf{a})$. Имеется и обратное включение $\mathbf{a} \in \mathrm{Cl}_U(\mathbf{b})$, которое вытекает в силу свойства (2).

Докажем совпадением множеств $\mathrm{Cl}_U(\mathbf{a})$ и $\mathrm{Cl}_U(\mathbf{b})$. Выберем произвольный элемент $\mathbf{x} \in \mathrm{Cl}_U(\mathbf{a})$. Из $\mathbf{x} \in \mathrm{Cl}_U(\mathbf{a})$ и $\mathbf{a} \in \mathrm{Cl}_U(\mathbf{b})$ выводим $\mathbf{x} \in \mathrm{Cl}_U(\mathbf{b})$. Следовательно $\mathrm{Cl}_U(\mathbf{a}) \subseteq \mathrm{Cl}_U(\mathbf{b})$. Аналогичным образом выводится включение $\mathrm{Cl}_U(\mathbf{a}) \subseteq \mathrm{Cl}_U(\mathbf{b})$. Из этих двух включений получаем $\mathrm{Cl}_U(\mathbf{a}) = \mathrm{Cl}_U(\mathbf{b})$. $\square$

Множество всех смежных классов пространства V по подпространству U называется *фактормножеством* V/U. В силу доказанной выше теоремы различные классы из фактормножества имеют пустое пересечение, а объединение всех таких классов совпадает с пространством V:

$$V = \bigcup_{Q \in V/U} Q.$$

Это равенство вытекает из того, что всякий вектор $\mathbf{v} \in V$ содержится в некотором смежном классе $v \in Q$. Класс Q определяется любым своим представителем $\mathbf{v}$ по формуле $Q = \mathrm{Cl}_U(\mathbf{v})$. В силу сказанного следующая теорема получается простой переформулировкой определения смежного класса.

Теорема 7.3. *Два вектора* $\mathbf{v}$ *и* $\mathbf{w}$ *принадлежат одному смежному классу по подпространству* U *тогда и только тогда, когда их разность* $\mathbf{v}-\mathbf{w}$ *принадлежит подпространству* U.

Определение 7.1. Пусть Q_1 и Q_2 — два смежных класса по подпространству U. Суммой классов Q_1 и Q_2 называется смежный класс Q, определяемый соотношением $Q = \mathrm{Cl}_U(\mathbf{v}_1 + \mathbf{v}_2)$, где $\mathbf{v}_1 \in Q_1$ и $\mathbf{v}_2 \in Q_2$.

Определение 7.2. Пусть Q — смежный класс по подпространству U. Произведением класса Q на число $\alpha \in \mathbb{K}$ называется смежный класс P по подпространству U, определяемый соотношением $P = \mathrm{Cl}_U(\alpha \cdot \mathbf{v})$, где $\mathbf{v} \in Q$.

Для сложения классов и умножения классов на число принято использовать те же знаки алгебраических операций, что и в случае векторов: то есть $Q = Q_1 + Q_2$ и $P = \alpha \cdot Q$. Определения 7.1 и 7.2 можно изобразить так:

$$\begin{aligned} &\mathrm{Cl}_U(\mathbf{v}_1) + \mathrm{Cl}_U(\mathbf{v}_2) = \mathrm{Cl}_U(\mathbf{v}_1 + \mathbf{v}_2), \\ &\alpha \cdot \mathrm{Cl}_U(\mathbf{v}) = \mathrm{Cl}_U(\alpha \cdot \mathbf{v}). \end{aligned} \tag{7.3}$$

Заметим, что определения 7.1 и 7.2 требуют некоторых комментариев. Действительно, класс $Q = Q_1 + Q_2$ в определении 7.1 и класс $P = \alpha \cdot Q$ в определении 7.2 определяются выбором векторов $\mathbf{v}_1 \in Q_1$, $\mathbf{v}_2 \in Q_2$ и $\mathbf{v} \in Q$. Выбор представителя в классе неоднозначен, поэтому однозначность результата операции в определениях 7.1 и 7.2 требует специального обоснования. Это же относиться и к формулам (7.3). Такое обоснование называется доказательством корректности определения 7.1 и 7.2.

Теорема 7.4. *Определения 7.1 и 7.2 корректны и определяемые ими операции сложения классов и умножения класса на число не зависят от выбора представителей в этих классах.*

Док-во. Начнем с операции сложения классов. Рассмотрим два выбора представителей в смежных классах Q_1 и Q_2. Пусть $\mathbf{v}_1, \tilde{\mathbf{v}}_1 \in Q_1$ и $\mathbf{v}_2, \tilde{\mathbf{v}}_2 \in Q_2$. Тогда в силу теоремы 7.3 находим:

$$\tilde{\mathbf{v}}_1 - \mathbf{v}_1 \in U, \qquad\qquad \tilde{\mathbf{v}}_2 - \mathbf{v}_2 \in U.$$

Отсюда: $(\tilde{\mathbf{v}}_1 + \tilde{\mathbf{v}}_2) - (\mathbf{v}_1 + \mathbf{v}_2) = (\tilde{\mathbf{v}}_1 - \mathbf{v}_1) + (\tilde{\mathbf{v}}_2 - \mathbf{v}_2) \in U$. Значит, классы векторов $\tilde{\mathbf{v}}_1 - \mathbf{v}_1$ и $\tilde{\mathbf{v}}_2 - \mathbf{v}_2$ совпадают

$$\mathrm{Cl}_U(\tilde{\mathbf{v}}_1 + \tilde{\mathbf{v}}_2) = \mathrm{Cl}_U(\mathbf{v}_1 + \mathbf{v}_2).$$

Это доказывает корректность определения 7.1 для сложения классов.

Теперь два выбора представителей в классе Q. Положим $\mathbf{v}, \tilde{\mathbf{v}} \in Q$. Тогда $\tilde{\mathbf{v}} - \mathbf{v} \in U$. Отсюда выводим $\alpha \cdot \tilde{\mathbf{v}} - \alpha \cdot \mathbf{v} = \alpha \cdot (\tilde{\mathbf{v}} - \mathbf{v}) \in U$. Это дает $\mathrm{Cl}_U(\alpha \cdot \tilde{\mathbf{v}}) = \mathrm{Cl}_U(\alpha \cdot \mathbf{v})$, что доказывает корректность определения 7.2.

Определения 7.1 и 7.2 оснащают фактормножество V/U двумя алгебраическими операциями — сложения классов и умножения классов на число.

Теорема 7.5. *Фактормножество V/U пространства V по подпространству U, оснащенное алгебраическими операциями* (7.3), *является линейным векторным пространством. Оно называется факторпространством пространства V по подпространству U.*

Док-во. Доказательство теоремы сводится к проверке выполнения аксиом (1)-(8) линейного векторного пространства для V/U. Коммутативность и ассоциативность сложения классов вытекают из следующих вычислений:

$$\begin{aligned}&\mathrm{Cl}_U(\mathbf{v}_1) + \mathrm{Cl}_U(\mathbf{v}_2) = \mathrm{Cl}_U(\mathbf{v}_1 + \mathbf{v}_2) = \\ &\qquad = \mathrm{Cl}_U(\mathbf{v}_2 + \mathbf{v}_1) = \mathrm{Cl}_U(\mathbf{v}_2) + \mathrm{Cl}_U(\mathbf{v}_1),\end{aligned}$$

$$\begin{gathered}(\mathrm{Cl}_U(\mathbf{v}_1) + \mathrm{Cl}_U(\mathbf{v}_2)) + \mathrm{Cl}_U(\mathbf{v}_3) = \mathrm{Cl}_U(\mathbf{v}_1 + \mathbf{v}_2) + \mathrm{Cl}_U(\mathbf{v}_3) = \\ = \mathrm{Cl}_U((\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{v}_3) = \mathrm{Cl}_U(\mathbf{v}_1 + (\mathbf{v}_2 + \mathbf{v}_3)) = \\ \mathrm{Cl}_U(\mathbf{v}_1) + \mathrm{Cl}_U(\mathbf{v}_2 + \mathbf{v}_3) = \mathrm{Cl}_U(\mathbf{v}_1) + (\mathrm{Cl}_U(\mathbf{v}_2) + \mathrm{Cl}_U(\mathbf{v}_3)).\end{gathered}$$

Они являются следствием соответствующих свойств сложения векторов.

Для проверки выполнения аксиомы (3) мы должны указать нулевой элемент в V/U. Выберем в качестве нуля нулевой класс $\mathbf{0} = \mathrm{Cl}_U(\mathbf{0})$. Такой выбор нуля удовлетворяет необходимому условию из аксиомы (3):

$$\mathrm{Cl}_U(\mathbf{v}) + \mathrm{Cl}_U(\mathbf{0}) = \mathrm{Cl}_U(\mathbf{v} + \mathbf{0}) = \mathrm{Cl}_U(\mathbf{v}).$$

Для проверки выполнения аксиомы (4) мы должны указать противоположный класс Q' для $Q = \mathrm{Cl}_U(\mathbf{v})$. Определим его так: $Q' = \mathrm{Cl}_U(\mathbf{v}')$. Тогда:

$$Q + Q' = \mathrm{Cl}_U(\mathbf{v}) + \mathrm{Cl}_U(\mathbf{v}') = \mathrm{Cl}_U(\mathbf{v} + \mathbf{v}') = \mathrm{Cl}_U(\mathbf{0}) = \mathbf{0}.$$

Проверка оставшихся аксиом (5)-(8) сводится просто к прямым вычислениям на базе формул (7.3), корректность которых установлена:

$$\begin{aligned}&\alpha\cdot(\mathrm{Cl}_U(\mathbf{v}_1) + \mathrm{Cl}_U(\mathbf{v}_2)) = \alpha\cdot\mathrm{Cl}_U(\mathbf{v}_1 + \mathbf{v}_2) =\\ &= \mathrm{Cl}_U(\alpha\cdot(\mathbf{v}_1 + \mathbf{v}_2)) = \mathrm{Cl}_U(\alpha\cdot\mathbf{v}_1 + \alpha\cdot\mathbf{v}_2) =\\ &= \mathrm{Cl}_U(\alpha\cdot\mathbf{v}_1) + \mathrm{Cl}_U(\alpha\cdot\mathbf{v}_2) = \alpha\cdot\mathrm{Cl}_U(\mathbf{v}_1) + \alpha\cdot\mathrm{Cl}_U(\mathbf{v}_2),\end{aligned}$$

$$\begin{aligned}&(\alpha+\beta)\cdot\mathrm{Cl}_U(\mathbf{v}) = \mathrm{Cl}_U((\alpha+\beta)\cdot\mathbf{v}) = \mathrm{Cl}_U(\alpha\cdot\mathbf{v} + \beta\cdot\mathbf{v}) =\\ &= \mathrm{Cl}_U(\alpha\cdot\mathbf{v}) + \mathrm{Cl}_U(\beta\cdot\mathbf{v}) = \alpha\cdot\mathrm{Cl}_U(\mathbf{v}) + \beta\cdot\mathrm{Cl}_U(\mathbf{v}),\end{aligned}$$

$$\begin{aligned}&\alpha\cdot(\beta\cdot\mathrm{Cl}_U(\mathbf{v})) = \alpha\cdot\mathrm{Cl}_U(\beta\cdot\mathbf{v}) = \mathrm{Cl}_U(\alpha\cdot(\beta\cdot\mathbf{v})) =\\ &= \mathrm{Cl}_U((\alpha\beta)\cdot\mathbf{v}) = (\alpha\beta)\cdot\mathrm{Cl}_U(\mathbf{v}),\end{aligned}$$

$$1\cdot\mathrm{Cl}_U(\mathbf{v}) = \mathrm{Cl}_U(1\cdot\mathbf{v}) = \mathrm{Cl}_U(\mathbf{v}).$$

Выписанные равенства завершают проверку того, что фактормножество V/U наделено структурой линейного векторного пространства. □

Отметим, что при проверке аксиомы (4) мы определили противоположный класс Q' для класса $Q = \mathrm{Cl}_U(\mathbf{v})$ соотношением $Q' = \mathrm{Cl}_U(\mathbf{v}')$. Можно установить корректность такого определения. Однако, в этом нет настоятельной необходимости, ибо в силу свойства (10) из теоремы 2.1 противоположный класс Q' для Q единственен.

Понятие факторпространства в равной мере применимо и к конечномерным и к бесконечномерным пространствам V. Конечномерность или бесконечномерность подпространства U также не играет роли. Единственное упрощение в конечномерном случае состоит в том, что мы можем вычислить размерность факторпространства V/U.

Теорема 7.6. *Если пространство V конечномерно, то для любого его подпространства U соответствующее факторпространство V/U также конечномерно и его размерность удовлетворяет соотношению*

$$\dim U + \dim(V/U) = \dim V. \tag{7.4}$$

Док-во. Если $U = V$, то факторпространство V/U состоит из единственного нулевого класса $V/U = \{0\}$. Размерность такого факторпространства равна нулю и равенство (7.4) выполнено тривиальным образом.

Рассмотрим нетривиальный случай $U \subsetneq V$. В силу теоремы 4.5 подпространство U конечномерно. Пусть $\dim V = n$ и $\dim U = s$, тогда $s < n$. Выберем базис $\mathbf{e}_1, \ldots, \mathbf{e}_s$ в подпространстве U и в соответствии с теоремой 4.8 рассмотрим его дополнение $\mathbf{e}_{s+1}, \ldots, \mathbf{e}_n$ до базиса в V. Для каждого из дополнительных векторов рассмотрим его класс:

$$\mathbf{E}_1 = \mathrm{Cl}_U(\mathbf{e}_{s+1}),\ \ldots,\ \mathbf{E}_{n-s} = \mathrm{Cl}_U(\mathbf{e}_n). \tag{7.5}$$

Покажем, что классы (7.5) порождают факторпространство V/U. Действительно, пусть Q — произвольный класс из V/U и пусть $v \in Q$ — некоторый представитель этого класса. Разложим вектор $\mathbf{v}$ по базису в V:

$$\mathbf{v} = (\alpha_1 \cdot \mathbf{e}_1 + \ldots + \alpha_s \cdot \mathbf{e}_s) + \beta_1 \cdot \mathbf{e}_{s+1} + \ldots + \beta_{n-s} \cdot \mathbf{e}_n.$$

Обозначим через $\mathbf{u}$ начальный фрагмент этого разложения: $\mathbf{u} = \alpha_1 \cdot \mathbf{e}_1 + \ldots + \alpha_s \cdot \mathbf{e}_s$. Ясно, что $\mathbf{u} \in U$. Тогда

$$\mathbf{v} = \mathbf{u} + \beta_1 \cdot \mathbf{e}_{s+1} + \ldots + \beta_{n-s} \cdot \mathbf{e}_n.$$

Но $Cl_U(\mathbf{u}) = 0$ по причине $\mathbf{u} \in U$. Для класса $Q = \mathrm{Cl}_U(\mathbf{v})$ это дает $Q = \mathrm{Cl}_U(\mathbf{u}) + \beta_1 \cdot \mathrm{Cl}_U(\mathbf{e}_{s+1}) + \ldots + \beta_{n-s} \cdot \mathrm{Cl}_U(\mathbf{e}_n)$. Отсюда

$$Q = \beta_1 \cdot \mathbf{E}_1 + \ldots + \beta_{n-s} \cdot \mathbf{E}_{n-s}.$$

Значит, $\mathbf{E}_1, \ldots, \mathbf{E}_{n-s}$ есть конечная порождающая система в V/U. Это доказывает конечномерность факторпространства V/U. Для определения его размерности докажем линейную независимость классов (7.5). Для этого рассмотрим линейную комбинацию этих классов, равную нулю:

$$\gamma_1 \cdot \mathbf{E}_1 + \ldots + \gamma_{n-s} \cdot \mathbf{E}_{n-s} = \mathbf{0}. \tag{7.6}$$

Переходя от классов к векторам, из этого соотношения получаем:

$$\begin{aligned}\gamma_1 \cdot \mathrm{Cl}_U(\mathbf{e}_{s+1}) + \ldots + \gamma_{n-s} \cdot \mathrm{Cl}_U(\mathbf{e}_n) =\\ = \mathrm{Cl}_U(\gamma_1 \cdot \mathbf{e}_{s+1} + \ldots + \gamma_{n-s} \cdot \mathbf{e}_n) = \mathrm{Cl}_U(\mathbf{0}).\end{aligned}$$

Обозначим $\mathbf{u} = \gamma_1 \cdot \mathbf{e}_{s+1} + \ldots + \gamma_{n-s} \cdot \mathbf{e}_n$. Из выведенного соотношения для этого вектора имеем $\mathrm{Cl}_U(\mathbf{u}) = \mathrm{Cl}_U(\mathbf{0})$, что дает $\mathbf{u} \in U$. Разложим вектор $\mathbf{u}$ по базису в подпространстве: $\mathbf{u} = \alpha_1 \cdot \mathbf{e}_1 + \ldots + \alpha_s \cdot \mathbf{e}_s$. Тогда приравняв два выражения для вектора $\mathbf{u}$, получим

$$-\alpha_1 \cdot \mathbf{e}_1 - \ldots - \alpha_s \cdot \mathbf{e}_s + \gamma_1 \cdot \mathbf{e}_{s+1} + \ldots + \gamma_{n-s} \cdot \mathbf{e}_n = \mathbf{0}.$$

Это линейная комбинация базисных векторов пространства V, равная нулю. Ввиду линейной независимости векторов $\mathbf{e}_1, \ldots, \mathbf{e}_n$ она тривиальна. Значит, $\gamma_1 = \ldots = \gamma_{n-s} = 0$. Это доказывает тривиальность линейной комбинации (7.6) и линейную независимость классов (7.5). Для размерности факторпространства получаем $\dim(V/U) = n - s$, что доказывает соотношение (7.4). $\square$

§ 8. Линейные отображения.

Определение 8.1. Пусть V и W — два линейных векторных пространства над числовым полем $\mathbb{K}$. Отображение $f\colon V \to W$ из пространства в V в пространство W называется *линейным отображением*, если:

(1) $f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2)$ для любых $\mathbf{v}_1, \mathbf{v}_2 \in V$;
(2) $f(\alpha \cdot \mathbf{v}) = \alpha \cdot f(\mathbf{v})$ для любого $\mathbf{v} \in V$ и любого числа α из поля $\mathbb{K}$.

Простейшим и немедленным следствием свойств (1) и (2) линейного отображения является соотношение $f(\mathbf{0}) = \mathbf{0}$. Действительно:

$$f(\mathbf{0}) = f(\mathbf{0} + (-1) \cdot \mathbf{0}) = f(\mathbf{0}) + (-1) \cdot f(\mathbf{0}) = \mathbf{0}. \tag{8.1}$$

Теорема 8.1. *Линейные отображения обладают следующими свойствами:*

(1) *тождественное отображение* $\mathrm{id}_V : V \to V$ *линейного векторного пространства* V *на себя линейно;*
(2) *композиция двух линейных отображений* $f: V \to W$ *и* $g: W \to U$ *есть линейное отображение* $g \circ f : V \to U$*;*
(3) *если линейное отображение* $f: V \to W$ *биективно, то обратное отображение* $f^{-1} : W \to V$ *также линейно.*

Док-во. Линейность тождественного отображения очевидна. Проверка условий (1) и (2) из определения 8.1 для него имеет следующий вид:

$$\begin{aligned}
&\mathrm{id}_V(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{v}_1 + \mathbf{v}_2 = \mathrm{id}_V(\mathbf{v}_1) + \mathrm{id}_V(\mathbf{v}_2),\\
&\mathrm{id}_V(\alpha \cdot \mathbf{v}) = \alpha \cdot \mathbf{v} = \alpha \cdot \mathrm{id}_V(\mathbf{v}).
\end{aligned}$$

Докажем второй пункт теоремы. Рассмотрим композицию отображений $g \circ f$. Для такой композиции условия (1) и (2) из определения линейности 8.1 проверяются следующим образом:

$$\begin{aligned}
g \circ f(\mathbf{v}_1 + \mathbf{v}_2) = g(f(\mathbf{v}_1 + \mathbf{v}_2) = g(f(\mathbf{v}_1) + f(\mathbf{v}_2)) = \\
= g(f(\mathbf{v}_1)) + g(f(\mathbf{v}_2)) = g \circ f(\mathbf{v}_1) + g \circ f(\mathbf{v}_2),\\
g \circ f(\alpha \cdot \mathbf{v}) = g(f(\alpha \cdot \mathbf{v})) = g(\alpha \cdot f(\mathbf{v})) = \alpha \cdot g(f(\mathbf{v})) = \alpha \cdot g \circ f(\mathbf{v}).
\end{aligned}$$

Теперь докажем третий пункт теоремы. Для этого рассмотрим биективное отображение $f: V \to W$. Оно имеет однозначно определенное обратное отображение $f^{-1} : W \to V$ (см. теорему 1.9). Сделаем следующие обозначения:

$$\begin{aligned}
&\mathbf{z}_1 = f^{-1}(\mathbf{w}_1 + \mathbf{w}_2) - f^{-1}(\mathbf{w}_1) - f^{-1}(\mathbf{w}_2),\\
&\mathbf{z}_2 = f^{-1}(\alpha \cdot \mathbf{w}) - \alpha \cdot f^{-1}(\mathbf{w}).
\end{aligned}$$

Линейность обратного отображения, очевидно, эквивалентна занулению векторов $\mathbf{z}_1$ и $\mathbf{z}_2$. Вычислим вектора $f(\mathbf{z}_1)$ и $f(\mathbf{z}_2)$:

$$\begin{aligned}
f(\mathbf{z}_1) &= f(f^{-1}(\mathbf{w}_1 + \mathbf{w}_2) - f^{-1}(\mathbf{w}_1) - f^{-1}(\mathbf{w}_2)) = \\
&= f(f^{-1}(\mathbf{w}_1 + \mathbf{w}_2)) - f(f^{-1}(\mathbf{w}_1)) - f(f^{-1}(\mathbf{w}_2)) = \\
&= (\mathbf{w}_1 + \mathbf{w}_2) - \mathbf{w}_1 - \mathbf{w}_2 = \mathbf{0},\\
f(\mathbf{z}_2) &= f(f^{-1}(\alpha \cdot \mathbf{w}) - \alpha \cdot f^{-1}(\mathbf{w})) = f(f^{-1}(\alpha \cdot \mathbf{w}))- \\
&- \alpha \cdot f(f^{-1}(\mathbf{w})) = \alpha \cdot \mathbf{w} - \alpha \cdot \mathbf{w} = \mathbf{0}.
\end{aligned}$$

Биективное отображение инъективно. Поэтому из полученных равенств $f(\mathbf{z}_1) = \mathbf{0}$, $f(\mathbf{z}_2) = \mathbf{0}$ и из равенства $f(\mathbf{0}) = \mathbf{0}$, выведенного в (8.1), вытекает $\mathbf{z}_1 = \mathbf{z}_2 = \mathbf{0}$. Теорема доказана. $\square$

С каждым линейным отображением $f\colon V \to W$ связаны два подмножества: *ядро* $\operatorname{Ker} f \subset V$ и *образ* $\operatorname{Im} f \subset W$. Образ $\operatorname{Im} f = f(V)$ для линейного отображения определяется так же, как это было сделано в случае произвольного отображения выше в параграфе 1:

$$\operatorname{Im} f = \{\mathbf{w} \in W\colon \exists\, \mathbf{v}\ ((\mathbf{v} \in V)\ \ \&\ \ (f(\mathbf{v}) = \mathbf{w}))\}.$$

Ядром линейного отображения $f\colon V \to W$ называется множество элементов из пространства V, отображающихся в ноль:

$$\operatorname{Ker} f = \{\mathbf{v} \in V\colon\ f(\mathbf{v}) = \mathbf{0}\}$$

Теорема 8.2. *Ядро и образ для линейного отображения $f\colon V \to W$ являются подпространствами в пространствах V и W соответственно.*

Док-во. Для доказательства теоремы проверим условия (1) и (2) из определения 2.2 применительно к подмножествам $\operatorname{Ker} f \subset V$ и $\operatorname{Im} f \subset W$.

Пусть $\mathbf{v}_1, \mathbf{v}_2 \in \operatorname{Ker} f$. Тогда $f(\mathbf{v}_1) = \mathbf{0}$ и $f(\mathbf{v}_2) = \mathbf{0}$. Пусть также $\mathbf{v} \in \operatorname{Ker} f$. Тогда $f(\mathbf{v}) = \mathbf{0}$. Отсюда выводим

$$\begin{aligned}
&f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2) = \mathbf{0} + \mathbf{0} = \mathbf{0},\\
&f(\alpha \cdot \mathbf{v}) = \alpha \cdot f(\mathbf{v}) = \alpha \cdot \mathbf{0} = \mathbf{0}.
\end{aligned}$$

Значит, $\mathbf{v}_1 + \mathbf{v}_2 \in \operatorname{Ker} f$ и $\alpha \cdot \mathbf{v} \in \operatorname{Ker} f$. Это доказывает утверждение теоремы относительно $\operatorname{Ker} f$.

Пусть $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w} \in \operatorname{Im} f$. Тогда существуют векторы $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}$ из V, такие, что $f(\mathbf{v}_1) = \mathbf{w}_1$, $f(\mathbf{v}_2) = \mathbf{w}_2$ и $f(\mathbf{v}) = \mathbf{w}$. Отсюда

$$\begin{aligned}
&\mathbf{w}_1 + \mathbf{w}_2 = f(\mathbf{v}_1) + f(\mathbf{v}_2) = f(\mathbf{v}_1 + \mathbf{v}_2),\\
&\alpha \cdot \mathbf{w} = \alpha \cdot f(\mathbf{v}) = f(\alpha \cdot \mathbf{v}).
\end{aligned}$$

Значит, $\mathbf{w}_1 + \mathbf{w}_2 \in \operatorname{Im} f$ и $\alpha \cdot \mathbf{w} \in \operatorname{Im} f$. Теорема доказана. □

Напомним, что в силу теоремы 1.2 линейное отображение $f\colon V \to W$ сюръективно тогда и только тогда, когда $\operatorname{Im} f = W$. Аналогичное утверждение имеется для $\operatorname{Ker} f$.

Теорема 8.3. *Линейное отображение $f\colon V \to W$ инъективно тогда и только тогда, когда* $\operatorname{Ker} f = \{\mathbf{0}\}$.

Док-во. Пусть f инъективно и пусть $\mathbf{v} \in \operatorname{Ker} f$. Тогда $f(\mathbf{0}) = \mathbf{0}$ и $f(\mathbf{v}) = \mathbf{0}$. Но из $\mathbf{v} \neq 0$, вытекало бы $f(\mathbf{v}) \neq f(\mathbf{0})$. Следовательно, $\mathbf{v} = \mathbf{0}$ и ядро отображения f состоит только из нулевого вектора: $\operatorname{Ker} f = \{\mathbf{0}\}$.

Теперь наоборот, пусть $\operatorname{Ker} f = \{\mathbf{0}\}$. Рассмотрим пару не совпадающих векторов $\mathbf{v}_1 \neq \mathbf{v}_2$ из V. Значит, $\mathbf{v}_1 - \mathbf{v}_2 \neq \mathbf{0}$ и $\mathbf{v}_1 - \mathbf{v}_2 \notin \operatorname{Ker} f$. Поэтому $f(\mathbf{v}_1 - \mathbf{v}_2) \neq \mathbf{0}$. Используя линейность отображения f, из этого выводим: $f(\mathbf{v}_1) - f(\mathbf{v}_2) \neq \mathbf{0}$ и $f(\mathbf{v}_1) \neq f(\mathbf{v}_2)$. Следовательно отображение f инъективно. Теорема доказана. □

Следующая теорема известна как теорема о линейной независимости прообразов для линейно независимой системы векторов.

ТЕОРЕМА 8.4. *Пусть $f\colon V \to W$ — некоторое линейное отображение и пусть вектора $\mathbf{v}_1, \ldots, \mathbf{v}_s$ из пространства V таковы, что соответствующие им вектора $f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)$ в W линейно независимы. Тогда сами вектора $\mathbf{v}_1, \ldots, \mathbf{v}_s$ также линейно независимы.*

ДОК-ВО. Для доказательства теоремы рассмотрим линейную комбинацию векторов $\mathbf{v}_1, \ldots, \mathbf{v}_s$, равную нулю:

$$\alpha_1 \cdot \mathbf{v}_1 + \ldots + \alpha_s \cdot \mathbf{v}_s = \mathbf{0}.$$

Применив к обоим частям этого равенства отображение f и воспользовавшись его линейностью, получим соотношение

$$\alpha_1 \cdot f(\mathbf{v}_1) + \ldots + \alpha_s \cdot f(\mathbf{v}_s) = \mathbf{0}.$$

Теперь из линейной независимости $f(\mathbf{v}_1), \ldots, f(\mathbf{v}_s)$ выводим равенство нулю всех коэффициентов: $\alpha_1 = \ldots = \alpha_s = 0$. Значит, и исходная линейная комбинация тривиальна, что доказывает линейную независимость исходной системы векторов $\mathbf{v}_1, \ldots, \mathbf{v}_s$. $\square$

Линейное векторное пространство — это некоторое множество. Но это не просто множество, это структурированное множество. На нем определены алгебраические операции и выполнены аксиомы (1)-(8). Линейные отображения — это отображения, согласованные со структурами линейных векторных пространств. На алгебраическом языке отображения, согласованные с алгебраическими структурами в связываемых ими множествах, называются *морфизмами.* На этом языке линейные отображения — это морфизмы линейных векторных пространств. Биективные линейные отображения называются *изоморфизмами* линейных векторных пространств.

ОПРЕДЕЛЕНИЕ 8.2. Два линейных векторных пространства V и W называются *изоморфными*, если существует биективное линейное отображение $f\colon V \to W$, связывающее эти два пространства.

Первым примером изоморфизма линейных векторных пространств было отображение $\psi\colon\ V \to \mathbb{K}^n$ из (5.4). Из существования такого отображения вытекает следующая теорема.

ТЕОРЕМА 8.5. *Всякое n-мерное линейное векторное пространство V изоморфно арифметическому линейному векторному пространству $\mathbb{K}^n$.*

Изоморфные линейные векторные пространства имеют много общего. Часто их вообще можно считать неразличимыми. В частности, имеет место следующий факт.

ТЕОРЕМА 8.6. *Если линейное пространство V изоморфно конечномерному пространству W, то V конечномерно и размерности этих двух пространств совпадают:* $\dim V = \dim W$.

ДОК-ВО. Пусть $f\colon V \to W$ — изоморфизм пространств V и W. Положим для определенности $\dim W = n$ и выберем базис $\mathbf{h}_1, \ldots, \mathbf{h}_n$ в W. При помощи обратного отображения $f^{-1}\colon W \to V$ определим вектора $\mathbf{e}_i = f^{-1}(\mathbf{h}_i)$,

$i = 1, \ldots, n$. Пусть $\mathbf{v}$ — некоторый произвольный вектор из V. Отобразим его с помощью f в пространство W и разложим по базису:

$$f(\mathbf{v}) = \alpha_1 \cdot \mathbf{h}_1 + \ldots + \alpha_n \cdot \mathbf{h}_n.$$

Применим отображение f^{-1} к обеим частям этого равенства. В результате, используя линейность f^{-1}, получим разложение

$$\mathbf{v} = \alpha_1 \cdot \mathbf{e}_1 + \ldots + \alpha_n \cdot \mathbf{e}_n.$$

Из этого разложения видим, что $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ есть порождающая система в V, состоящая из конечного числа векторов. Конечномерность пространства V доказана. Линейная независимость $\mathbf{e}_1, \ldots, \mathbf{e}_n$ вытекает из теоремы 8.4 о линейной независимости прообразов. Значит, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ есть базис в V и $\dim V = n = \dim W$. Теорема доказана. $\square$

§ 9. Матрица линейного отображения.

Пусть $f\colon V \to W$ — отображение из n-мерного векторного пространства V в m-мерное векторное пространство W. Выберем базис $\mathbf{e}_1 \ldots, \mathbf{e}_n$ в пространстве V и базис $\mathbf{h}_1, \ldots, \mathbf{h}_m$ в пространстве W. Рассмотрим образы векторов $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и разложим их по базису $\mathbf{h}_1, \ldots, \mathbf{h}_m$:

$$\begin{array}{ccccccc} f(\mathbf{e}_1) & = & F_1^1 \cdot \mathbf{h}_1 & + & \ldots & + & F_1^m \cdot \mathbf{h}_m, \\ \multicolumn{7}{c}{\dotfill} \\ f(\mathbf{e}_n) & = & F_n^1 \cdot \mathbf{h}_1 & + & \ldots & + & F_n^m \cdot \mathbf{h}_m. \end{array} \tag{9.1}$$

В (9.1) мы имеем n разложений, из которых определяется $n\,m$ чисел F_j^i. Они записываются в прямоугольную матрицу F размера $m\times n$, которая называется *матрицей линейного отображения* f в паре базисов $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\mathbf{h}_1, \ldots, \mathbf{h}_m$:

$$F = \left\| \begin{matrix} F_1^1 & \ldots & F_n^1 \\ \vdots & \ldots & \vdots \\ F_1^m & \ldots & F_n^m \end{matrix} \right\|. \tag{9.2}$$

При записи элемента F_j^i в матрицу (9.2) верхний индекс определяет номер строки, нижний — номер столбца. Иными словами, матрица F составляется из вектор-столбцов, образованных координатами векторов $f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)$ в базисе $\mathbf{h}_1, \ldots, \mathbf{h}_m$. Разложения (9.1), определяющие компоненты этой матрицы, удобно записать так:

$$f(\mathbf{e}_j) = \sum_{i=1}^{m} F_j^i \cdot \mathbf{h}_i. \tag{9.3}$$

Пусть $\mathbf{x}$ — некоторый произвольный вектор из V и пусть $\mathbf{y} = f(\mathbf{x})$ — его образ при отображении f. Разложим $\mathbf{x}$ по базису: $\mathbf{x} = x^1 \cdot \mathbf{e}_1 + \ldots + x^n \cdot \mathbf{e}_n$. Тогда с учетом (9.3) для вектора $\mathbf{y}$ получим:

$$f(\mathbf{x}) = \sum_{j=1}^{n} x^j \cdot f(\mathbf{e}_j) = \sum_{j=1}^{n} x^j \cdot \left(\sum_{i=1}^{m} F_j^i \cdot \mathbf{h}_i \right).$$

Поменяв порядок суммирования в полученном выражении, мы получим разложение вектора $\mathbf{y}$ по базису $\mathbf{h}_1, \ldots, \mathbf{h}_m$:

$$\mathbf{y} = f(\mathbf{x}) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} F^i_j \, x^j \right) \cdot \mathbf{h}_i.$$

В силу однозначности такого разложения для координат вектора $\mathbf{y}$ в базисе $\mathbf{h}_1, \ldots, \mathbf{h}_m$ получится следующее выражение:

$$y^i = \sum_{j=1}^{n} F^i_j \, x^j. \tag{9.4}$$

Формула (9.4) — это основное применение матрицы линейного отображения. Она позволяет вычислить координаты вектора $f(\mathbf{x})$ непосредственно по координатам вектора $\mathbf{x}$. В матричной форме эта формула имеет вид

$$\begin{Vmatrix} y^1 \\ \vdots \\ y^m \end{Vmatrix} = \begin{Vmatrix} F^1_1 & \ldots & F^1_n \\ \vdots & \ldots & \vdots \\ F^m_1 & \ldots & F^m_n \end{Vmatrix} \cdot \begin{Vmatrix} x^1 \\ \vdots \\ x^n \end{Vmatrix}. \tag{9.5}$$

Вспомним, что сопоставление вектору x вектор-столбца из его координат мы договорились оформлять в виде линейного отображения $\psi\colon V \to \mathbb{K}^n$ (см. (5.8)). Аналогичное отображение для векторов y из W обозначим через $\tilde{\psi}: W \to \mathbb{K}^m$. Матричное соотношение (9.5) можно трактовать как отображение $F: \mathbb{K}^n \to \mathbb{K}^m$. Перечисленные три отображения вместе с отображением f можно изобразить в виде диаграммы:

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\psi"'] & W \arrow[d, "\tilde{\psi}"] \\ \mathbb{K}^n \arrow[r, "F"'] & \mathbb{K}^m \end{tikzcd} \tag{9.6}$$

Подобные диаграммы называются *коммутативными*, если композиции отображений, соответствующие «прохождению по стрелкам диаграммы», не зависят от «маршрута» прохождения. Применительно к диаграмме (9.6) коммутативность означает $\tilde{\psi} \circ f = F \circ \psi$. Ввиду биективности линейных отображений ψ и $\tilde{\psi}$ условие коммутативности диаграммы (9.6) можно изобразить двумя эквивалентными формулами следующего вида:

$$F = \tilde{\psi} \circ f \circ \psi^{-1}, \qquad\qquad f = \tilde{\psi}^{-1} \circ F \circ \psi. \tag{9.7}$$

Читатель легко проверит, что соотношения (9.7) выполнены в силу способа построения матрицы F. Следовательно, диаграмма (9.6) коммутативна.

Посмотрим на соотношения (9.7) с несколько иной точки зрения. Пусть заданы два пространства V и W размерности n и m соответственно. Пусть также задана некоторая произвольная матрица F размера $m \times n$. Тогда соотношение (9.5) определит линейное отображение $F: \mathbb{K}^n \to \mathbb{K}^m$. Выбрав базисы $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\mathbf{h}_1, \ldots, \mathbf{h}_m$ в пространствах V и W и, пользуясь вторым соотношением (9.7), определим отображение $f: V \to W$. Матрица этого отображения f в базисах $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\mathbf{h}_1, \ldots, \mathbf{h}_m$ в точности совпадает с F. В результате доказана следующая теорема.

Теорема 9.1. *Всякая прямоугольная матрица F размера $m \times n$ может быть реализована как матрица некоторого линейного отображения $f\colon V \to W$ из n-мерного пространства V в m-мерное пространство W в некоторой паре базисов в этих пространствах.*

Более прямой способ доказательства теоремы 9.1 может быть основан на следующей теореме.

Теорема 9.2. *Для всякого базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в n-мерном линейном векторном пространстве V и для любого набора $\mathbf{w}_1, \ldots, \mathbf{w}_n$ из n-векторов в другом пространстве W существует линейное отображение $f: V \to W$, такое, что $f(\mathbf{e}_i) = \mathbf{w}_i$, где $i = 1, \ldots, n$.*

Док-во. Задание базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в V определяет линейное отображение $\psi: V \to \mathbb{K}^n$ (см. (5.8)). Для построения необходимого отображения f определим отображение $\varphi\colon \mathbb{K}^n \to W$ следующим соотношением:

$$\varphi: \begin{Vmatrix} x^1 \\ \vdots \\ x^n \end{Vmatrix} \mapsto x^1 \cdot \mathbf{w}_1 + \ldots + x^n \cdot \mathbf{w}_n.$$

Теперь нетрудно убедиться, что искомое отображение есть композиция двух построенных отображений $f = \varphi \circ \psi$. $\square$

Вернемся к исходной ситуации. Пусть задано отображение $f: V \to W$, которое определяет матрицу F после выбора базисов $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\mathbf{h}_1, \ldots, \mathbf{h}_m$ в пространствах V и W. Зависимость матрицы F от выбора базисов существенна. Для того, чтобы описать такую зависимость, рассмотрим четыре базиса: два базиса в V и два в W. Пусть S и P — матрицы перехода для указанных пар базисов в V и W соответственно:

$$\tilde{\mathbf{e}}_k = \sum_{j=1}^{n} S^j_k \cdot \mathbf{e}_j, \qquad\qquad \tilde{\mathbf{h}}_r = \sum_{i=1}^{m} P^i_r \cdot \mathbf{h}_i.$$

Пусть $T = S^{-1}$ и $Q = P^{-1}$ — соответствующие матрицы обратного перехода:

$$\mathbf{e}_j = \sum_{k=1}^{n} T^k_j \cdot \tilde{\mathbf{e}}_k, \qquad\qquad \mathbf{h}_i = \sum_{r=1}^{m} Q^r_i \cdot \tilde{\mathbf{h}}_r.$$

Используем эти соотношения вместе с соотношением (9.3) для того, чтобы проделать следующие вычисления:

$$\begin{aligned} f(\tilde{\mathbf{e}}_k) = \sum_{j=1}^{n} S^j_k \cdot f(\mathbf{e}_j) = \sum_{j=1}^{n} S^j_k \cdot \left(\sum_{i=1}^{m} F^i_j \cdot \mathbf{h}_i\right) = \\ = \sum_{j=1}^{n} S^j_k \cdot \left(\sum_{i=1}^{m} F^i_j \cdot \left(\sum_{r=1}^{m} Q^r_i \cdot \tilde{\mathbf{h}}_r\right)\right). \end{aligned}$$

После перемены порядка суммирования результат этого вычисления можно записать в форме следующего соотношения:

$$f(\tilde{\mathbf{e}}_k) = \sum_{r=1}^{m} \left(\sum_{i=1}^{m} \sum_{j=1}^{n} Q^r_i \, F^i_j \, S^j_k \right) \cdot \tilde{\mathbf{h}}_r .$$

Двойные суммы в скобках являются коэффициентами разложения векторов $f(\tilde{\mathbf{e}}_k)$ по базису $\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_m$. Они определяют матрицу линейного отображения f в волнистых базисах $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ и $\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_m$:

$$\tilde{F}^r_k = \sum_{i=1}^{m} \sum_{j=1}^{n} Q^r_i \, F^i_j \, S^j_k . \tag{9.8}$$

Аналогичным образом можно вывести противоположное соотношение, выражающее компоненты матрицы F через компоненты $\tilde{F}$:

$$F^i_j = \sum_{r=1}^{m} \sum_{k=1}^{n} P^i_r \, \tilde{F}^r_k \, T^k_j . \tag{9.9}$$

Соотношения (9.8) и (9.9) называются *формулами преобразования матрицы линейного отображения при замене базисов.* Эти соотношения допускают матричную форму записи. В этом случае они выглядят так:

$$\tilde{F} = P^{-1} \, F \, S, \qquad\qquad F = P \, \tilde{F} \, S^{-1}. \tag{9.10}$$

Формулы преобразования типа (9.10) подводят нас к очень широкому кругу задач «о приведении к каноническому виду». В данном случае изменение базисов в пространствах V и W изменяет вид матрицы линейного отображения $f\colon V \to W$. Задача о приведении к каноническому виду состоит в нахождении оптимального выбора базисов, в которых матрица F имеет максимально простой (канонический) вид. Следующая теорема, дающая решение данной задачи, известна как теорема о приведении к *почти диагональному* виду.

Теорема 9.3. *Пусть $f\colon V \to W$ — некоторое ненулевое линейное отображение из n-мерного пространства V в m-мерное пространство W. Тогда существует такой выбор базисов в V и W, при котором матрица F этого отображения имеет следующий почти диагональный вид:*

$$F = \begin{array}{c} \overbrace{\qquad\qquad\qquad}^{s} \qquad\qquad\qquad \\ \left\| \begin{array}{ccccccc} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{array} \right\| \end{array} \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\1\\\vdots\\1\end{array}}\right\} s \\ \vphantom{\begin{array}{c}1\\\vdots\\1\end{array}} \end{array} \tag{9.11}$$

Док-во. Чисто нулевое отображение $0\colon V \to W$ отображает любой вектор в нулевой. Матрица такого отображения состоит из одних нулей при любом выборе базисов в V и W. Задача о приведении к каноническому виду для такой матрицы просто не ставится.

Пусть $f\colon V \to W$ — ненулевое линейное отображение. Число $s = \dim(\operatorname{Im} f)$ называется *рангом* отображения f. Ранг ненулевого отображения отличен от нуля. Построение канонического базиса в W начнем с выбора базиса $\mathbf{h}_1, \ldots, \mathbf{h}_s$ в образе $\operatorname{Im} f$. Для каждого базисного вектора $\mathbf{h}_i \in \operatorname{Im} f$ существует вектор $\mathbf{e}_i \in V$, такой, что $f(\mathbf{e}_i) = \mathbf{h}_i$, $i = 1, \ldots, s$. Вектора $\mathbf{e}_1, \ldots, \mathbf{e}_s$ линейно независимы в силу теоремы 8.4. Пусть $r = \dim(\operatorname{Ker} f)$. Выберем базис в $\operatorname{Ker} f$ и обозначим базисные вектора через $\mathbf{e}_{s+1}, \ldots, \mathbf{e}_{s+r}$. Далее рассмотрим систему векторов

$$\mathbf{e}_1, \ldots, \mathbf{e}_s, \mathbf{e}_{s+1}, \ldots, \mathbf{e}_{s+r} \tag{9.12}$$

и докажем, что она является базисом в V. Для этого используем теорему 4.6.

Начнем с условия (1) теоремы 4.6. Для доказательства линейной независимости векторов (9.12) рассмотрим их линейную комбинацию, равную нулю:

$$\begin{aligned}\alpha_1 \cdot \mathbf{e}_1 + \ldots + \alpha_s \cdot \mathbf{e}_s &+ \\ + \alpha_{s+1} \cdot \mathbf{e}_{s+1} &+ \ldots + \alpha_{s+r} \cdot \mathbf{e}_{s+r} = \mathbf{0}.\end{aligned} \tag{9.13}$$

Применим к обеим частям равенства (9.13) отображение f и при этом учтем, что $f(\mathbf{e}_i) = \mathbf{h}_i$ при $i = 1, \ldots, s$. Остальные векторы принадлежат ядру отображения, поэтому $f(\mathbf{e}_{s+i}) = \mathbf{0}$. С учетом этого из (9.13) выводим:

$$\alpha_1 \cdot \mathbf{h}_1 + \ldots + \alpha_s \cdot \mathbf{h}_s = \mathbf{0}.$$

Вектора $\mathbf{h}_1, \ldots, \mathbf{h}_s$ составляют базис в $\operatorname{Im} f$. Они линейно независимы, поэтому $\alpha_1 = \ldots = \alpha_s = 0$. Учет этого сводит (9.13) к соотношению:

$$\alpha_{s+1} \cdot \mathbf{e}_{s+1} + \ldots + \alpha_{s+r} \cdot \mathbf{e}_{s+r} = \mathbf{0}.$$

Вектора $\mathbf{e}_{s+1}, \ldots, \mathbf{e}_{s+r}$ составляют базис в $\operatorname{Ker} f$. Они линейно независимы, поэтому $\alpha_{s+1} = \ldots = \alpha_{s+r} = 0$. В итоге мы доказали зануление всех коэффициентов в линейной комбинации (9.13). Следовательно, вектора (9.12) линейно независимы.

Пусть $\mathbf{v}$ — некоторый произвольный вектор из V. Тогда $f(\mathbf{v})$ принадлежит $\operatorname{Im} f$. Разложим $f(\mathbf{v})$ по базису $\mathbf{h}_1, \ldots, \mathbf{h}_s$:

$$f(\mathbf{v}) = \beta_1 \cdot \mathbf{h}_1 + \ldots + \beta_s \cdot \mathbf{h}_s. \tag{9.14}$$

Вспомним, что $f(\mathbf{e}_i) = \mathbf{h}_i$ при $i = 1, \ldots, s$. Тогда из (9.14) выводим

$$\begin{aligned}\mathbf{0} = f(\mathbf{v}) - \beta_1 \cdot f(\mathbf{e}_1) - \ldots - \beta_s \cdot f(\mathbf{e}_s) &= \\ = f(\mathbf{v} - \beta_1 \cdot \mathbf{e}_1 - \ldots - \beta_s \cdot \mathbf{e}_s)&\end{aligned} \tag{9.15}$$

Обозначим $\tilde{\mathbf{v}} = \mathbf{v} - \beta_1 \cdot \mathbf{e}_1 - \ldots - \beta_s \cdot \mathbf{e}_s$. Из (9.15) для него имеем: $f(\tilde{\mathbf{v}}) = \mathbf{0}$. Значит, $\tilde{\mathbf{v}} \in \operatorname{Ker} f$. Разложим вектор $\tilde{\mathbf{v}}$ по базису в ядре отображения f:

$$\tilde{\mathbf{v}} = \beta_{s+1} \cdot \mathbf{e}_{s+1} + \ldots + \beta_{s+r} \cdot \mathbf{e}_{s+r}.$$

Объединив $\tilde{\mathbf{v}} = \mathbf{v} - \beta_1 \cdot e_1 - \ldots - \beta_s \cdot e_s$ с выписанным разложением, получаем:

$$\mathbf{v} = \beta_1 \cdot \mathbf{e}_1 + \ldots + \beta_{s+r} \cdot \mathbf{e}_{s+1}.$$

Вывод: вектора (9.12) порождают пространство V. Условие (2) теоремы 4.6 также доказано. Значит, вектора (9.12) составляют базис в V. Это дает

$$\dim V = s + r. \tag{9.16}$$

Для завершения доказательства теоремы осталось дополнить базис $\mathbf{h}_1, \ldots, \mathbf{h}_s$ из $\operatorname{Im} f$ до базиса $\mathbf{h}_1, \ldots, \mathbf{h}_s, \mathbf{h}_{s+1}, \ldots, \mathbf{h}_m$ в W. При $j = 1, \ldots, s$ для векторов $f(\mathbf{e}_j)$ имеем разложение

$$f(\mathbf{e}_j) = \mathbf{h}_j = \sum_{i=1}^{s} \delta^i_j \cdot \mathbf{h}_i + \sum_{i=s+1}^{m} 0 \cdot \mathbf{h}_i.$$

При $j = s + 1, \ldots, s + r$ для $f(\mathbf{e}_j)$ разложение чисто нулевое:

$$f(\mathbf{e}_j) = \mathbf{0} = \sum_{i=1}^{s} 0 \cdot \mathbf{h}_i + \sum_{i=s+1}^{m} 0 \cdot \mathbf{h}_i.$$

В силу полученных разложений матрица отображения f в построенных базисах имеет требуемый вид (9.10). □

В процессе доказательства этой теоремы мы одновременно доказали следующую теорему.

Теорема 9.4. *Пусть $f: V \to W$ — линейное отображение из n-мерного пространства V в пространство W. Тогда*

$$\dim(\operatorname{Ker} f) + \dim(\operatorname{Im} f) = \dim V. \tag{9.17}$$

Теорема 9.4 известна как теорема *о сумме размерностей ядра и образа* линейного отображения. Утверждение теоремы в форме соотношения (9.17) немедленно вытекает из (9.16).

§ 10. Алгебраические операции с отображениями. Пространство гомоморфизмов $\operatorname{Hom}(V, W)$.

Определение 10.1. Пусть V и W — два линейных векторных пространства и пусть $f\colon V \to W$ и $g\colon V \to W$ — два отображения из V в W. Суммой отображений f и h называется отображение $h\colon V \to W$, определяемое соотношением $h(\mathbf{v}) = f(\mathbf{v}) + g(\mathbf{v})$ для всех $\mathbf{v} \in V$.

Определение 10.2. Пусть V и W — два линейных векторных пространства над числовым полем $\mathbb{K}$ и пусть $f : V \to W$ — отображение из V в W. Произведением числа $\alpha \in \mathbb{K}$ на отображение f называется отображение $h\colon V \to W$, определяемое соотношением $h(\mathbf{v}) = \alpha \cdot f(\mathbf{v})$ для всех $\mathbf{v} \in V$.

Алгебраические операции, вводимые определениями 10.1 и 10.2, принято называть *поточечным сложением* отображений и *поточечным умножением* отображений на число. Действительно, они выполняются «поточечно» путем сложения значений исходных отображений и умножения этих значений на число отдельно для каждого значения аргумента $\mathbf{v} \in V$. Для обозначения этих операций используются те же значки, что и для соответствующих операций с векторами: $h = f + g$ и $h = \alpha \cdot f$. Запись $(f + g)(\mathbf{v})$ понимается как применение суммы отображений $f + g$ к вектору $\mathbf{v}$. Запись же $f(\mathbf{v}) + g(\mathbf{v})$ означает сложение результатов применения отображений f и g к вектору $\mathbf{v}$. Хотя результаты этих действий совпадают, их смысл различен. Аналогичным образом следует различать левую и правую части следующего равенства

$$(\alpha \cdot f)(\mathbf{v}) = \alpha \cdot f(\mathbf{v}).$$

Обозначим через $\mathrm{Map}(V, W)$ множество всевозможных отображений из пространства V в пространство W. Иногда это множество обозначают так: W^V.

Теорема 10.1. *Множество* $\mathrm{Map}(V, W)$*, оснащенное операциями поточечного сложения и поточечного умножения отображений на число, является линейным векторным пространством.*

Док-во. Проверим выполнение аксиом линейного векторного пространства для $\mathrm{Map}(V, W)$. В случае первой аксиомы мы должны установить совпадение отображений $f + g$ и $g + f$. Совпадение отображений эквивалентно совпадению результатов их применения к произвольному вектору $\mathbf{v} \in V$. Следующие вычисления как раз и устанавливают такое совпадение:

$$(f + g)(\mathbf{v}) = f(\mathbf{v}) + g(\mathbf{v}) = g(\mathbf{v}) + f(\mathbf{v}) = (g + f)(\mathbf{v}).$$

Как мы видим из проделанных вычислений, равенство $f+g = g+f$ вытекает из аксиомы коммутативности для пространства W в силу поточечного характера операции сложения отображений. Аналогичное обстоятельство проявляется и при проверке аксиом (2), (5) и (6):

$$\begin{aligned}
&((f + g) + h)(\mathbf{v}) = (f + g)(\mathbf{v}) + h(\mathbf{v}) = (f(\mathbf{v}) + g(\mathbf{v})) + h(\mathbf{v}) = \\
&\quad = f(\mathbf{v}) + (g(\mathbf{v}) + h(\mathbf{v})) = f(\mathbf{v}) + (g + h)(\mathbf{v}) = (f + (g + h))(\mathbf{v})
\end{aligned}$$

$$\begin{aligned}
&(\alpha \cdot (f + g))(\mathbf{v}) = \alpha \cdot (f + g)(\mathbf{v}) = \alpha \cdot (f(\mathbf{v}) + g(\mathbf{v})) = \\
&\quad = \alpha \cdot f(\mathbf{v}) + \alpha \cdot g(\mathbf{v}) = (\alpha \cdot f)(\mathbf{v}) + (\alpha \cdot g)(\mathbf{v}) = (\alpha \cdot f + \alpha \cdot g)(\mathbf{v})
\end{aligned}$$

$$\begin{aligned}
&((\alpha + \beta) \cdot f)(\mathbf{v}) = (\alpha + \beta) \cdot f(\mathbf{v}) = \alpha \cdot f(\mathbf{v}) + \beta \cdot f(\mathbf{v}) = \\
&\qquad\qquad = (\alpha \cdot f)(\mathbf{v}) + (\beta \cdot f)(\mathbf{v}) = (\alpha \cdot f + \beta \cdot f)(\mathbf{v})
\end{aligned}$$

В случае аксиомы (7) соответствующие выкладки имеют вид $(\alpha \cdot (\beta \cdot f))(\mathbf{v}) = \alpha \cdot (\beta \cdot f)(\mathbf{v}) = \alpha \cdot (\beta \cdot f(\mathbf{v})) = (\alpha\beta) \cdot f(\mathbf{v}) = ((\alpha\beta) \cdot f)(\mathbf{v})$, а в случае аксиомы (8) эти выкладки еще проще: $(1 \cdot f)(\mathbf{v}) = 1 \cdot f(\mathbf{v}) = f(\mathbf{v})$.

Остается проверить аксиомы (3) и (4). В качестве нулевого элемента в пространстве $\mathrm{Map}(V, W)$ выберем нулевое отображение $0\colon V \to W$, которое отображает все вектора из V в нулевой вектор из W. Тогда $(f + 0)(\mathbf{v}) = f(\mathbf{v}) + 0(\mathbf{v}) = f(\mathbf{v}) + \mathbf{0} = f(\mathbf{v})$. Аксиома (3) выполнена.

Для отображения $f \in \operatorname{Map}(V, W)$ противоположное отображение f' определим так $f' = (-1) \cdot f$. Тогда

$$\begin{aligned}(f + f')(\mathbf{v}) = (f + (-1) \cdot f)(\mathbf{v}) = f(\mathbf{v}) +\\ + ((-1) \cdot f)(\mathbf{v}) = f(\mathbf{v}) + (-1) \cdot f(\mathbf{v}) = \mathbf{0} = 0(\mathbf{v})\end{aligned}$$

Аксиома (3) также выполнена. Это завершает доказательство теоремы. $\square$

В типичной ситуации пространство $\operatorname{Map}(V, W)$ очень велико. Даже для конечномерных пространств V и W оно обычно бывает бесконечномерным. В линейной алгебре принято изучать гораздо меньшее подмножество в пространстве $\operatorname{Map}(V, W)$, состоящее из всех линейных отображений из V в W. Оно обозначается $\operatorname{Hom}(V, W)$ и называется *множеством гомоморфизмов*. Следующие две теоремы показывают замкнутость $\operatorname{Hom}(V, W)$ относительно алгебраических операций в $\operatorname{Map}(V, W)$. Поэтому мы можем говорить о $\operatorname{Hom}(V, W)$ как о *пространстве гомоморфизмов*.

Теорема 10.2. *Поточечная сумма двух линейных отображений $f\colon V \to W$ и $g\colon V \to W$ есть линейное отображение из пространства V в пространство W.*

Теорема 10.3. *Поточечное произведение всякого линейного отображения $f : V \to W$ на число $\alpha \in \mathbb{K}$ есть линейное отображение из пространства V в пространство W.*

Док-во. Пусть $h = f + g$ есть сумма отображений f и g. Следующие выкладки доказывают линейность отображения h:

$$\begin{aligned}h(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1 + \mathbf{v}_2) + g(\mathbf{v}_1 + \mathbf{v}_2) = (f(\mathbf{v}_1) +\\ + f(\mathbf{v}_2)) + (g(\mathbf{v}_1) + g(\mathbf{v}_2)) = (f(\mathbf{v}_1) +\\ + g(\mathbf{v}_1)) + (f(\mathbf{v}_2) + g(\mathbf{v}_2)) = h(\mathbf{v}_1) + h(\mathbf{v}_2),\end{aligned}$$

$$\begin{aligned}h(\beta \cdot \mathbf{v}) = f(\beta \cdot \mathbf{v}) + g(\beta \cdot \mathbf{v}) = \beta \cdot f(\mathbf{v}) +\\ + \beta \cdot g(\mathbf{v}) = \beta \cdot (f(\mathbf{v}) + g(\mathbf{v})) = \beta \cdot h(\mathbf{v}).\end{aligned}$$

Теперь рассмотрим произведение отображения f на число $\alpha \in \mathbb{K}$. Положим $h = \alpha \cdot f$ и проделаем аналогичные выкладки:

$$\begin{aligned}h(\mathbf{v}_1 + \mathbf{v}_2) = \alpha \cdot f(\mathbf{v}_1 + \mathbf{v}_2) = \alpha \cdot (f(\mathbf{v}_1) +\\ + f(\mathbf{v}_2)) = \alpha \cdot f(\mathbf{v}_1) + \alpha \cdot f(\mathbf{v}_2) = h(\mathbf{v}_1) + h(\mathbf{v}_2)\end{aligned}$$

$$\begin{aligned}h(\beta \cdot \mathbf{v}) = \alpha \cdot f(\beta \cdot \mathbf{v}) = \alpha \cdot (\beta \cdot f(\mathbf{v})) =\\ = (\alpha\beta) \cdot f(\mathbf{v}) = (\beta\alpha) \cdot f(\mathbf{v}) = \beta \cdot (\alpha \cdot f(\mathbf{v})) = \beta \cdot h(\mathbf{v})\end{aligned}$$

Проделанные выкладки завершают доказательство теорем 10.2 и 10.3. $\square$

Пространство гомоморфизмов $\operatorname{Hom}(V, W)$ является подпространством в пространстве всех отображений $\operatorname{Map}(V, W)$. Оно гораздо меньше по своим «размерам» и состоит из объектов, изучение которых входит в предмет ведения линейной алгебры. Для конечномерных пространств V и W пространство $\operatorname{Hom}(V, W)$ также конечномерно. Это вытекает из следующей теоремы.

Теорема 10.4. *Для двух конечномерных пространств V и W пространство* $\operatorname{Hom}(V,W)$ *конечномерно и его размерность определяется формулой:*

$$\dim(\operatorname{Hom}(V,W)) = \dim(V)\dim(W). \tag{10.1}$$

Док-во. Пусть $\dim V = n$ и $\dim W = m$. Выберем базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в пространстве V и базис $\mathbf{h}_1, \ldots, \mathbf{h}_m$ в пространстве W. Пусть $1 \leqslant i \leqslant n$ и $1 \leqslant j \leqslant m$. Для каждой фиксированной пары индексов i, j из указанных диапазонов рассмотрим следующий набор из n векторов пространства W:

$$\mathbf{w}_1 = \mathbf{0},\ \ldots,\ \mathbf{w}_{i-1} = \mathbf{0},\ \mathbf{w}_i = \mathbf{h}_j,\ \mathbf{w}_{i+1} = \mathbf{0},\ \ldots,\ \mathbf{w}_n = \mathbf{0}.$$

Все вектора в этом наборе, за исключением вектора $\mathbf{w}_i$, равны нулю, а вектор $\mathbf{w}_i$ совпадает с базисным вектором $\mathbf{h}_j$. Применим теорему 9.2 к базису $\mathbf{e}_1, \ldots, \mathbf{e}_n$ из V и набору векторов $\mathbf{w}_1, \ldots, \mathbf{w}_n$. Это определит нам линейное отображение $E^i_j\colon V \to W$, для которого $E^i_j(\mathbf{e}_s) = \mathbf{w}_s$ при всех $s = 1, \ldots, n$. Запишем это обстоятельство так:

$$E^i_j(\mathbf{e}_s) = \delta^i_s \cdot \mathbf{h}_j, \tag{10.2}$$

где δ^i_s — символ Кронекера. Всего мы имеем $n\,m$ отображений E^i_j, удовлетворяющих соотношениям (10.2):

$$E^i_j\colon V \to W,\ 1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant m. \tag{10.3}$$

Покажем, что отображения (10.3) порождают пространство $\operatorname{Hom}(V,W)$. Рассмотрим некоторое произвольное отображение $f \in \operatorname{Hom}(V,W)$. Пусть F — его матрица в базисах $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\mathbf{h}_1, \ldots, \mathbf{h}_m$ и пусть F^j_i — элементы этой матрицы. Тогда действие f на произвольный вектор $\mathbf{v} \in V$ определяется координатами этого вектора по формуле:

$$f(\mathbf{v}) = \sum_{i=1}^{n} v^i \cdot f(\mathbf{e}_i) = \sum_{i=1}^{n}\sum_{j=1}^{m} F^j_i\, v^i\, \mathbf{h}_j. \tag{10.4}$$

Для действия E^i_j на этот же вектор $\mathbf{v}$ из (10.2) выводим:

$$E^i_j(\mathbf{v}) = \sum_{s=1}^{n} v^s \cdot E^i_j(\mathbf{e}_s) = \sum_{s=1}^{n} (v^s\,\delta^i_s) \cdot \mathbf{h}_j = v^i \cdot \mathbf{h}_j. \tag{10.5}$$

Теперь из сравнения соотношений (10.4) и (10.5) находим

$$f(\mathbf{v}) = \sum_{i=1}^{n}\sum_{j=1}^{m} F^j_i \cdot E^i_j(\mathbf{v}).$$

Ввиду произвольности вектора $\mathbf{v} \in V$ это можно записать в форме разложения отображения f в линейную комбинацию отображений (10.3):

$$f = \sum_{i=1}^{n}\sum_{j=1}^{m} F^j_i \cdot E^i_j.$$

Значит, построенные выше отображения (10.3) порождают пространство гомоморфизмов $\mathrm{Hom}(V, W)$, откуда вытекает конечномерность этого пространства.

Для подсчета размерности пространства $\mathrm{Hom}(V, W)$ докажем линейную независимость отображений (10.3). Рассмотрим их линейную комбинацию, которую приравняем к нулю:

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\gamma_i^j\cdot E_j^i = 0. \tag{10.6}$$

Левая и правая части этого равенства представляют собой нулевое отображение $0\colon V \to W$. Подействуем им на базисный вектор $\mathbf{e}_s$. Это дает

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\gamma_i^j\cdot E_j^i(\mathbf{e}_s) = \sum_{i=1}^{n}\sum_{j=1}^{m}(\gamma_i^j\,\delta_s^i)\cdot\mathbf{h}_j = \mathbf{0}.$$

Суммирование по i можно выполнить явно. В результате этого получим линейные комбинации базисных векторов пространства W, равные нулю:

$$\sum_{j=1}^{m}\gamma_i^j\cdot\mathbf{h}_j = \mathbf{0}.$$

Используя линейную независимость $\mathbf{h}_1, \ldots, \mathbf{h}_m$, находим $\gamma_i^j = 0$. Это доказывает тривиальность линейной комбинации (10.6) и линейную независимость отображений (10.3). Значит, они образуют базис в пространстве $\mathrm{Hom}(V, W)$. Подсчет числа таких отображений доказывает равенство (10.2). $\square$

Смысл доказанной теоремы становится прозрачным в терминах матриц линейных отображений. После выбора базисов в V и W, линейные отображения из $\mathrm{Hom}(V, W)$ изображаются прямоугольными матрицами $m \times n$. При этом сумме отображений соответствует сумма их матриц, а при умножении отображения на число его матрица также умножается на это число.

Прямоугольные матрицы размера $m\times n$ сами образуют линейное векторное пространство $\mathbb{K}^{m\times n}$, изоморфное арифметическому пространству $\mathbb{K}^{mn}$. Таким образом, выбор базисов в пространствах V и W определяет некоторый изоморфизм пространства $\mathrm{Hom}(V, W)$ и пространства матриц $\mathbb{K}^{m\times n}$.

ГЛАВА II

ЛИНЕЙНЫЕ ОПЕРАТОРЫ.

§ 1. Линейные операторы. Алгебра эндоморфизмов $\operatorname{End}(V)$ и группа автоморфизмов $\operatorname{Aut}(V)$.

Линейное отображение $f\colon V \to V$, действующее из линейного векторного пространства V в то же самое пространство V, принято называть *линейным оператором.* Линейные операторы являются частной разновидностью линейных отображений. Поэтому к ним применимы все те результаты, которые мы рассмотрели в предыдущей главе. Однако уменьшение общности всегда добавляет конкретики. Поэтому теория линейных операторов оказывается богаче и сложнее теории линейных отображений. Она содержит не только усиление доказанных ранее теорем для данного более специального случая, но и некоторый круг задач, постановка которых для случая общих линейных отображений просто невозможна.

Рассмотрим пространство гомоморфизмов $\operatorname{Hom}(V, W)$. В случае совпадения $W = V$ это пространство принято называть *пространством эндоморфизмов* $\operatorname{End}(V) = \operatorname{Hom}(V, V)$. Оно состоит из линейных операторов $f\colon V \to V$, которые иногда называют *эндоморфизмами* пространства V. В отличие от пространства гомоморфизмов $\operatorname{Hom}(V, W)$ общего вида, пространство эндоморфизмов $\operatorname{End}(V)$ оснащено дополнительной бинарной операцией. Действительно, имея два линейных оператора $f, g \in \operatorname{End}(V)$, мы не только можем складывать эти операторы и умножать их на числа, но можем также строить композиции $f \circ g \in \operatorname{End}(V)$ и $g \circ f \in \operatorname{End}(V)$.

Теорема 1.1. *На множество эндоморфизмов* $\operatorname{End}(V)$, *оснащенном операцией композиции, помимо аксиом линейного векторного пространства* (1)-(8) *выполнены соотношения:*

(9) $(f + g) \circ h = f \circ h + g \circ h$;
(10) $(\alpha \cdot f) \circ h = \alpha \cdot (f \circ h)$;
(11) $f \circ (g + h) = f \circ g + f \circ h$;
(12) $f \circ (\alpha \cdot g) = \alpha \cdot (f \circ g)$;

Док-во. Каждое из равенств (9)-(12) есть операторное равенство. А равенство двух операторов состоит в совпадении результатов действия этих операторов на произвольный вектор $\mathbf{v}$ из пространства V:

$$((f + g) \circ h)(\mathbf{v}) = (f + g)(h(\mathbf{v})) = f(h(\mathbf{v})) + g(h(\mathbf{v})) = \\ = (f \circ h)(\mathbf{v}) + (g \circ h)(\mathbf{v}) = (f \circ h + g \circ h)(\mathbf{v})$$

$$((\alpha \cdot f) \circ h)(\mathbf{v}) = (\alpha \cdot f)(h(\mathbf{v})) = \alpha \cdot f(h(\mathbf{v})) = \\ = \alpha \cdot (f \circ h)(\mathbf{v}) = (\alpha \cdot (f \circ h))(\mathbf{v})$$

$$\begin{aligned}&(f\circ(g+h))(\mathbf{v})=f((g+h)(\mathbf{v}))=f(g(\mathbf{v})+h(\mathbf{v}))=\\&=f(g(\mathbf{v}))+f(h(\mathbf{v}))=(f\circ g)(\mathbf{v})+(f\circ h)(\mathbf{v})=(f\circ g+f\circ h)(\mathbf{v})\end{aligned}$$

$$\begin{aligned}&(f\circ(\alpha\cdot g))(\mathbf{v})=f((\alpha\cdot g)(\mathbf{v}))=f(\alpha\cdot g(\mathbf{v}))=\\&\qquad\qquad=\alpha\cdot f(g(\mathbf{v}))=\alpha\cdot(f\circ g)(\mathbf{v})=(\alpha\cdot(f\circ g))(\mathbf{v})\end{aligned}$$

Проделанные выкладки доказывают свойства (9)-(12) для композиции линейных операторов. $\square$

Фиксируем оператор $h\in\operatorname{End}(V)$ и рассмотрим композицию $f\circ h$ как правило, сопоставляющее каждому оператору f некоторый другой оператор $g=f\circ h$. Это определяет отображение

$$R_h\colon \operatorname{End}(V)\to\operatorname{End}(V).$$

Первые два свойства (9) и (10) из теоремы 1.1 в точности означают линейность отображения R_h. Отображение R_h называют *правым сдвигом* на h, поскольку оно действует как взятие композиции с h, где h стоит справа. Аналогичным образом можно определить *левый сдвиг* на h:

$$L_h\colon \operatorname{End}(V)\to\operatorname{End}(V),$$

действующий по правилу $L_h(f)=h\circ f$. Это отображение линейно в силу свойств (9) и (10) из теоремы 1.1.

Операция взятия композиции операторов есть дополнительная бинарная операция на пространстве $\operatorname{End}(V)$. Свойство линейности отображения R_h принято называть *линейностью по первому аргументу*, а свойство линейности L_h называют *линейностью по второму аргументу.* Бинарная операция, линейная по каждому из аргументов, называется *билинейной.* Ситуация, когда на некотором линейном векторном пространстве имеется дополнительная билинейная бинарная операция, встречается довольно часто.

Определение 1.1. Линейное векторное пространство A над числовым полем $\mathbb{K}$, оснащенное билинейной бинарной операцией умножения векторов, называется *алгеброй над полем* $\mathbb{K}$ или просто $\mathbb{K}$-*алгеброй.*

Операция умножения в алгебрах обозначается каким-либо значком типа $\bullet$ или $\circ$, хотя часто знак операции не ставится вовсе. Алгебра A называется *коммутативной алгеброй*, если операция умножения в ней коммутативна: $a\,b=b\,a$. Алгебра A называется *ассоциативной алгеброй*, если операция умножения ассоциативна: $(a\,b)\,\mathrm{c}=a\,(b\,c)$.

Из определения 1.1 и теоремы 1.1 видим, что линейное пространство $\operatorname{End}(V)$ с операцией композиции в качестве умножения является алгеброй над тем же полем, что и исходное пространство V. Эта алгебра называется *алгеброй эндоморфизмов* пространства V. Она ассоциативна в силу теоремы 1.6 из первой главы, однако, коммутативной она, вообще говоря, не является.

Композиция является умножением в алгебре эндоморфизмов $\operatorname{End}(V)$, а знак умножения в записи принято опускать. Умножение операторов является приоритетной операцией по сравнению со сложением. Приоритет же операции умножения операторов по сравнению с умножением операторов на числа и с умножением чисел между собой значения не имеет. Это следует из аксиомы (7) для пространства $\operatorname{End}(V)$ и из свойств (10) и (12) операции умножения. Такое соглашение позволяет ввести в рассмотрение степени операторов:

$$f^2 = f\,f, \qquad f^3 = f^2\,f, \qquad f^{n+1} = f^n\,f.$$

Если оператор f биективен и имеет обратный оператор f^{-1}, то можно рассмотреть и отрицательные степени этого оператора:

$$f^{-2} = f^{-1}\,f^{-1}, \qquad f^n\,f^{-n} = \mathrm{id}_V, \qquad f^{n+m} = f^n\,f^m,$$

причем последнее равенство выполнено как для положительных, так и для отрицательных значений n и m.

Определение 1.2. Алгебра A над полем $\mathbb{K}$ называется *алгеброй с единичным элементом* или *алгеброй с единицей*, если в ней найдется элемент 1, такой, что $1\,a = a$ и $a\,1 = a$ для всех элементов $a \in A$.

Алгебра эндоморфизмов $\operatorname{End}(V)$ является алгеброй с единицей. Роль единичного элемента в ней играет тождественный оператор $1 = \mathrm{id}_V$. Он называется также *единичным оператором* или *операторной единицей*.

Определение 1.3. Линейный оператор $f\colon V \to V$ называется *скалярным оператором*, если он получается из единичного оператора 1 умножением на некоторое число $\lambda \in \mathbb{K}$, то есть $f = \lambda \cdot 1$.

Основным предназначением операторов из $\operatorname{End}(V)$ является их действие на векторы из V. Пусть $a, b \in \operatorname{End}(V)$ и пусть $\mathbf{x}, \mathbf{y} \in V$. Тогда:

(1) $(a + b)(\mathbf{x}) = a(\mathbf{x}) + b(\mathbf{x})$;
(2) $a(\mathbf{x} + \mathbf{y}) = a(\mathbf{x}) + b(\mathbf{y})$.

Эти соотношения нам хорошо известны: первое вытекает из определения суммы двух операторов, а второе следует из линейности оператора a. Вопрос в необходимости скобок вокруг вектора $\mathbf{x}$ и вектора $\mathbf{y}$ в этих формулах. Они являются следствием «функциональной» формы записи действия оператора на вектор: значок оператора ставится слева, а значок вектора окружается скобками подобно аргументу в записи функции: $\mathbf{w} = f(\mathbf{v})$. В алгебре принят более «вольный» стиль записи: $\mathbf{w} = f\,\mathbf{v}$, значок оператора всегда стоит слева от значка вектора, к которому он применен. Если известно, что $f \in \operatorname{End}(V)$ и $\mathbf{v} \in V$, то такая запись не может вызвать разночтений. В более сложной ситуации запись $\mathbf{w} = \alpha\,f\,g\,\mathbf{v}$ допускает различные толкования даже если известно, что $\alpha \in \mathbb{K}$ есть число, $f, g \in \operatorname{End}(V)$ и $\mathbf{v} \in V$:

$$\begin{aligned} &\mathbf{w} = \alpha \cdot f(g(\mathbf{v})), && \mathbf{w} = (\alpha \cdot f)(g(\mathbf{v})),\\ &\mathbf{w} = (\alpha \cdot (f \circ g))(\mathbf{v}), && \mathbf{w} = ((\alpha \cdot f) \circ g)(\mathbf{v}). \end{aligned}$$

Но при любом из этих толкований значение вектора $\mathbf{w}$ будет одним и тем же. Поэтому в дальнейшем мы часто будем использовать алгебраическую форму записи для действия оператора на вектор, особенно при записи громоздких вычислений.

Пусть $f\colon V \to V$ — линейный оператор в конечномерном векторном пространстве V. В соответствии с общей схемой для построения его матрицы мы должны выбрать два базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\mathbf{h}_1, \ldots, \mathbf{h}_n$ в пространстве V и построить разложения (9.1) из первой главы. Такой подход абсолютно верен, в некоторых случаях он очень полезен. Однако, два базиса в одном пространстве — это некоторое излишество. Поэтому при построении матрицы линейного оператора принято выбирать второй базис совпадающим с первым. Тогда матрица F для оператора f определится из разложений:

$$\begin{array}{ccccccc} f(\mathbf{e}_1) & = & F^1_1\cdot\mathbf{e}_1 & + & \ldots & + & F^n_1\cdot\mathbf{e}_n, \\ \multicolumn{7}{c}{\dotfill} \\ f(\mathbf{e}_n) & = & F^1_n\cdot\mathbf{e}_1 & + & \ldots & + & F^n_n\cdot\mathbf{e}_n, \end{array} \tag{1.1}$$

которые кратко можно изобразить следующей формулой:

$$f(\mathbf{e}_j) = \sum_{i=1}^{n} F^i_j \cdot \mathbf{e}_i. \tag{1.2}$$

Матрица F, определяемая из разложений (1.1) или (1.2), называется *матрицей линейного оператора* f в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Это квадратная матрица размера $n \times n$, где $n = \dim V$.

Теорема 1.2. *Сопоставление оператору $f \in \operatorname{End}(V)$ его матрицы в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$ обладает следующими свойствами:*

(1) *сумме операторов соответствует сумма их матриц;*
(2) *произведению оператора на число соответствует произведение его матрицы на это число;*
(3) *композиции двух операторов соответствует произведение их матриц.*

Док-во. Рассмотрим операторы f, g и h из $\operatorname{End}(V)$. Пусть F, G и H их матрицы в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$. При доказательстве первого пункта положим: $h = f + g$. Тогда для $h(\mathbf{e}_j)$ запишем

$$\begin{aligned} h(\mathbf{e}_j) &= (f+g)\,\mathbf{e}_j = f(\mathbf{e}_j) + h(\mathbf{e}_j) = \\ &= \sum_{i=1}^{n} F^i_j\cdot\mathbf{e}_i + \sum_{i=1}^{n} G^i_j\cdot\mathbf{e}_i = \sum_{i=1}^{n}(F^i_j + G^i_j)\cdot\mathbf{e}_i = \sum_{i=1}^{n} H^i_j\cdot\mathbf{e}_i. \end{aligned}$$

Отсюда в силу однозначности разложения векторов по базису $H^i_j = F^i_j + G^i_j$ и $H = F + G$. Первый пункт теоремы доказан.

Второй пункт теоремы доказывается аналогично. Для этого положим $h = \alpha\cdot f$. Тогда аналогично предыдущему случаю для $h(\mathbf{e}_j)$ запишем

$$\begin{aligned} h(\mathbf{e}_j) &= (\alpha\cdot f)\,\mathbf{e}_j = \alpha\cdot f(\mathbf{e}_j) = \\ &= \alpha\cdot\left(\sum_{i=1}^{n} F^i_j\cdot\mathbf{e}_i\right) = \sum_{i=1}^{n}(\alpha\,F^i_j)\cdot\mathbf{e}_i = \sum_{i=1}^{n} H^i_j\cdot\mathbf{e}_i. \end{aligned}$$

Отсюда $H^i_j = \alpha\,F^i_j$ и $H = \alpha\,F$. Доказательство третьего пункта теоремы требует несколько больших вычислений. Положим $h = f \circ g$. Тогда

$$h(\mathbf{e}_j) = (f \circ g)\,\mathbf{e}_j = f(g(\mathbf{e}_g)) = f\left(\sum_{i=1}^{n} G^i_j \cdot \mathbf{e}_i\right) = \sum_{i=1}^{n} G^i_j \cdot f(\mathbf{e}_i) =$$
$$= \sum_{i=1}^{n} G^i_j \cdot \left(\sum_{s=1}^{n} F^s_i \cdot \mathbf{e}_s\right) = \sum_{s=1}^{n}\left(\sum_{i=1}^{n} F^s_i\, G^i_j\right) \cdot \mathbf{e}_s = \sum_{s=1}^{n} H^s_i \cdot \mathbf{e}_s.$$

Отсюда, в силу однозначности разложения векторов по базису, выводим

$$H^s_i = \sum_{i=1}^{n} F^s_i\, G^i_j.$$

Это соотношение легко интерпретируется как произведение квадратных матриц $H = F\,G$. Теорема доказана. $\square$

Из доказанной теоремы заключаем, что сопоставление оператору f его матрицы, задает изоморфизм алгебры эндоморфизмов $\operatorname{End}(V)$ и алгебры квадратных матриц $\mathbb{K}^{n\times n}$ со стандартным матричным умножением.

Теперь рассмотрим вопрос о преобразовании матрицы линейного оператора $f\colon V \to V$ при замене базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ на некоторый новый базис $\tilde{\mathbf{e}}_1, \ldots, \tilde{e}_n$. Пусть S — соответствующая матрица прямого перехода, а T — матрица обратного перехода. Нам нет необходимости выводить формулы преобразования по новой. Достаточно приспособить к нашему случаю формулы (9.10) из первой главы. Из совпадения базиса $\mathbf{h}_1, \ldots, \mathbf{h}_n$ с $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и из совпадения $\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_n$ с $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ получаем $P = S$. Тогда

$$\tilde{F} = S^{-1}\,F\,S, \qquad\qquad F = S\,\tilde{F}\,S^{-1}. \tag{1.3}$$

Формулы (1.3) и есть искомые формулы преобразования матрицы линейного оператора при замене базиса. Учитывая равенство $T = S^{-1}$, запишем формулы (1.3) в индексной форме:

$$\tilde{F}^q_p = \sum_{i=1}^{n}\sum_{j=1}^{n} T^q_i\, S^j_p\, F^i_j, \qquad\qquad F^i_j = \sum_{q=1}^{n}\sum_{p=1}^{n} S^i_q\, T^p_j\, \tilde{F}^q_p. \tag{1.4}$$

Формулы (1.3) имеют важное следствие для детерминантов матриц F и $\tilde{F}$. Действительно, из них выводим

$$\det\tilde{F} = \det(S^{-1})\,\det F\,\det S = (\det S)^{-1}\,\det F\,\det S = \det F.$$

Совпадение детерминантов матриц линейного оператора, вычисленных в различных базисах, означает, что это число не зависит от выбора базиса вовсе.

Определение 1.4. Детерминантом $\det f$ линейного оператора f, действующего в пространстве V, называется число, равное определителю его матрицы при произвольном выборе базиса в V.

Числовым инвариантом геометрического объекта называется числовая характеристика этого объекта, не зависящая от выбора каких-либо вспомогательных конструкций. Детерминант линейного оператора $\det f$ есть его числовой инвариант. Координаты вектора или компоненты матрицы линейного оператора числовыми инвариантами не являются. Другим числовым инвариантом линейного оператора является его ранг:

$$\operatorname{rank} f = \dim(\operatorname{Im} f).$$

Вскоре мы определим целый ряд других числовых инвариантов линейного оператора.

Из третьего пункта теоремы 1.2 вытекает следующая формула для детерминанта линейных операторов:

$$\det(f \circ g) = \det(f)\,\det(g). \tag{1.5}$$

Теорема 1.3. *Линейный оператор $f\colon V \to V$ в конечномерном пространстве V инъективен тогда и только тогда, когда он сюръективен.*

Док-во. Для доказательства этой теоремы применим теорему 1.2, теорему 8.3 и теорему 9.4 из первой главы. Инъективность оператора f эквивалентна условию $\operatorname{Ker} f = \{\mathbf{0}\}$, сюръективность f эквивалентны условию $\operatorname{Im} f = V$, а теорема 9.4 из главы I дает связь между размерностями этих пространств:

$$\dim(\operatorname{Ker} f) + \dim(\operatorname{Im} f) = \dim(V).$$

Если оператор оператор f инъективен, то $\operatorname{Ker} f = \{\mathbf{0}\}$ и $\dim(\operatorname{Ker} f) = 0$. Тогда $\dim(\operatorname{Im} f) = \dim(V)$. Применив третий пункт теоремы 4.5 из главы I, получаем $\operatorname{Im} f = V$, что доказывает сюръективность оператора f.

Пусть, наоборот, оператор f сюръективен. Тогда $\operatorname{Im} f = V$ и для размерности образа имеем $\dim(\operatorname{Im} f) = \dim(V)$. Отсюда $\dim(\operatorname{Ker} f) = 0$ и $\operatorname{Ker} f = \{\mathbf{0}\}$, что доказывает инъективность оператора f. □

Теорема 1.4. *Линейный оператор $f\colon V \to V$ в конечномерном пространстве V биективен тогда и только тогда, когда* $\det f \neq 0$.

Док-во. Пусть $\mathbf{x}$ — вектор из V и пусть $\mathbf{y} = f(\mathbf{x})$. Разложив оба вектора $\mathbf{x}$ и $\mathbf{y}$ по некоторому базису $\mathbf{e}_1, \ldots, \mathbf{e}_n$, мы получим следующую формулу, связывающую координаты этих векторов:

$$\begin{Vmatrix} y^1 \\ \vdots \\ y^n \end{Vmatrix} = \begin{Vmatrix} F^1_1 & \ldots & F^1_n \\ \vdots & \ddots & \vdots \\ F^n_1 & \ldots & F^n_n \end{Vmatrix} \cdot \begin{Vmatrix} x^1 \\ \vdots \\ x^n \end{Vmatrix}. \tag{1.6}$$

Формула (1.6) может быть выведена непосредственно или же ее можно получить из формулы (9.5) из первой главы. Из этой формулы видим, что вектор $\mathbf{x}$ принадлежит ядру оператора f тогда и только тогда, когда его координаты $x^1, \ldots, x^n$ удовлетворяют однородной системе линейных уравнений

$$\begin{Vmatrix} F^1_1 & \ldots & F^1_n \\ \vdots & \ddots & \vdots \\ F^n_1 & \ldots & F^n_n \end{Vmatrix} \cdot \begin{Vmatrix} x^1 \\ \vdots \\ x^n \end{Vmatrix} = \begin{Vmatrix} 0 \\ \vdots \\ 0 \end{Vmatrix}, \tag{1.7}$$

матрица которой совпадает с матрицей оператора f в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Поэтому ядро оператора f отлично от нуля тогда и только тогда, когда система уравнений (1.7) имеет нетривиальное решение. Здесь мы воспользуемся известным результатом из теории детерминантов: система уравнений (1.7) с квадратной матрицей F имеет нетривиальное решение тогда и только тогда, когда $\det F = 0$. Доказательство его можно найти в книге [5]. Из этого результата немедленно получаем, что условие $\operatorname{Ker} f \neq \{\mathbf{0}\}$ эквивалентно условию $\det f = 0$. Тогда условие $\det f \neq 0$ эквивалентно $\operatorname{Ker} f = \{\mathbf{0}\}$. Последнее в силу предыдущей теоремы и теоремы 1.1 из первой главы эквивалентно биективности оператора $f\colon V \to V$. Теорема доказана. $\square$

Оператор f с нулевым детерминантом $det f = 0$ называется *вырожденным*. Пользуясь этой терминологией сформулируем следствие из доказанной теоремы 1.4.

Следствие. *Линейный оператор $f : V \to V$ в конечномерном пространстве V имеет нетривиальное ядро $\operatorname{Ker} f \neq \{0\}$ тогда и только тогда, когда он вырожден. Всякий невырожденный оператор биективен.*

Вспомним, что биективное линейное отображение f из V в W называется изоморфизмом. В случае $W = V$ оно устанавливает изоморфизм пространства V с самим собой. Такое отображение называется *автоморфизмом* пространства V. Множество всех автоморфизмов обозначается через $\operatorname{Aut}(V)$, оно состоит из биективных линейных операторов $f : V \to V$. Это множество, очевидно, обладает следующими свойствами:

(1) если $f, g \in \operatorname{Aut}(V)$, то $f \circ g \in \operatorname{Aut}(V)$;
(2) если $f \in \operatorname{Aut}(V)$, то $f^{-1} \in \operatorname{Aut}(V)$;
(3) $1 \in \operatorname{Aut}(V)$, где 1 — тождественный оператор.

Нетрудно видеть, что перечисленные свойства наделяют множество $\operatorname{Aut}(V)$ структурой группы. *Группа автоморфизмов* $\operatorname{Aut}(V)$ есть подмножество в алгебре $\operatorname{End}(V)$, однако ни структуру алгебры, ни даже структуру линейного пространства она не наследует. Это видно, например из того, что нулевой оператор не принадлежит $\operatorname{Aut}(V)$. В случае конечномерного пространства V группа автоморфизмов состоит из всех невырожденных операторов.

§ 2. Операторы проектирования.

Пусть V — некоторое линейное векторное пространство, разложенное в прямую сумму двух своих подпространств U_1 и U_2:

$$V = U_1 \oplus U_2. \tag{2.1}$$

В силу разложения (2.1) каждый вектор $\mathbf{v} \in V$ имеет разложение в сумму

$$\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2, \text{ где } \mathbf{u}_1 \in U_1 \text{ и } \mathbf{u}_2 \in U_2, \tag{2.2}$$

причем компоненты $\mathbf{u}_1$ и $\mathbf{u}_2$ в разложении (2.2) определяются вектором $\mathbf{v}$ однозначно.

Определение 2.1. Оператор $P\colon V \to V$, сопоставляющий каждому вектору $\mathbf{v} \in V$ его первую компоненту $P(\mathbf{v}) = \mathbf{u}_1$ в разложении (2.2), называется *оператором проектирования* на подпространство U_1 параллельно второму подпространству U_2.

Теорема 2.1. *Для всякого разложения* (2.1) *оператор проектирования* P *на подпространство* U_1 *параллельно подпространству* U_2 *линеен.*

Док-во. Рассмотрим пару векторов $\mathbf{v}_1, \mathbf{v}_2$ из пространства V и для каждого из них рассмотрим разложения (2.2):

$$\begin{aligned}\mathbf{v}_1 &= \mathbf{u}_1 + \mathbf{u}_2,\\ \mathbf{v}_2 &= \tilde{\mathbf{u}}_1 + \tilde{\mathbf{u}}_2.\end{aligned}$$

Тогда $P(\mathbf{v}_1) = \mathbf{u}_1$ и $P(\mathbf{v}_2) = \tilde{\mathbf{u}}_1$. Сложим выписанные выше разложения для $\mathbf{v}_1$ и $\mathbf{v}_2$ и запишем полученный результат в виде

$$\mathbf{v}_1 + \mathbf{v}_2 = (\mathbf{u}_1 + \tilde{\mathbf{u}}_1) + (\mathbf{u}_2 + \tilde{\mathbf{u}}_2). \tag{2.3}$$

Тогда из $\mathbf{u}_1, \tilde{\mathbf{u}}_1 \in U_1$ и из $\mathbf{u}_2, \tilde{\mathbf{u}}_2 \in U_2$ выводим $\mathbf{u}_1 + \tilde{\mathbf{u}}_1 \in U_1$ и $\mathbf{u}_2 + \tilde{\mathbf{u}}_2 \in U_2$. Поэтому (2.3) есть разложение вида (2.2) для вектора $\mathbf{v}_1 + \mathbf{v}_2$. Это дает

$$P(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{u}_1 + \tilde{\mathbf{u}}_1 = P(\mathbf{v}_1) + P(\mathbf{v}_2). \tag{2.4}$$

Теперь рассмотрим разложение (2.2) для произвольного вектора $\mathbf{v} \in V$ и умножим его на число $\alpha \in \mathbb{K}$:

$$\alpha \cdot \mathbf{v} = (\alpha \cdot \mathbf{u}_1) + (\alpha \cdot \mathbf{u}_2).$$

Тогда $\alpha \cdot \mathbf{u}_1 \in U_1$ и $\alpha \cdot \mathbf{u}_2 \in U_2$, поэтому из определения оператора P получаем

$$P(\alpha \cdot \mathbf{v}) = \alpha \cdot \mathbf{u}_1 = \alpha \cdot P(\mathbf{v}). \tag{2.5}$$

Соотношения (2.4) и (2.5), как раз, и означают линейность оператора P. $\square$

Пусть вектор $\mathbf{u}$ в разложении выбран принадлежащим подпространству U_1. Тогда разложение (2.2) для него имеет вид $\mathbf{u} = \mathbf{u} + \mathbf{0}$, поэтому для таких векторов $P(\mathbf{u}) = \mathbf{u}$. Вектора из U_1 проектируются оператором P в самих себя. Это имеет важное следствие: $P^2 = P$. Действительно, для всякого $\mathbf{v} \in V$ имеем: $P(\mathbf{v}) \in U_1$, поэтому $P(P(\mathbf{v})) = P(\mathbf{v})$.

Наряду с оператором P, из (2.2) мы можем определить еще один оператор Q, действующий по правилу $Q(\mathbf{v}) = \mathbf{u}_2$. Он также является оператором проектирования и проектирует на U_2 параллельно U_1. Поэтому $Q^2 = Q$. Для суммы этих операторов имеем: $P + Q = 1$. Действительно, для любого $\mathbf{v} \in V$:

$$P(\mathbf{v}) + Q(\mathbf{v}) = \mathbf{u}_1 + \mathbf{u}_2 = \mathbf{v} = 1(\mathbf{v}).$$

Если $\mathbf{u} \in U_1$, то разложение (2.2) для него имеет вид $\mathbf{u} = \mathbf{u} + \mathbf{0}$, поэтому $Q(\mathbf{u}) = \mathbf{0}$. Точно так же $P(\mathbf{u}) = \mathbf{0}$ для векторов $\mathbf{u} \in U_2$. Отсюда выводим: $Q(P(\mathbf{v})) = \mathbf{0}$ и $P(Q(\mathbf{v})) = \mathbf{0}$ для любого $\mathbf{v} \in V$. Суммируя все это, запишем

$$\begin{aligned}&P^2 = P, &&P + Q = 1,\\ &Q^2 = Q, &&P\,Q = Q\,P = 0.\end{aligned} \tag{2.6}$$

Пара операторов проектирования, удовлетворяющая соотношениям (2.6), называется *согласованной парой проекторов.*

Для получения согласованной пары проекторов достаточно задать лишь один из них, скажем, P. Второй определится из соотношения $Q = 1 - P$. При этом соотношения (2.6) будут автоматически выполнены. Действительно:

$$\begin{aligned} PQ &= P \circ (1 - P) = P - P^2 = P - P = 0, \\ QP &= (1 - P) \circ P = P - P^2 = P - P = 0. \end{aligned}$$

Соотношение $Q^2 = Q$ для Q выводится аналогичным образом:

$$Q^2 = (1 - P) \circ (1 - P) = 1 - 2P + P = 1 - P = Q.$$

Теорема 2.2. *Оператор $P\colon V \to V$ является оператором проектирования на одно подпространство параллельно другому подпространству тогда и только тогда, когда $P^2 = P$.*

Док-во. В одну сторону теорема уже нами доказана: если P — оператор проектирования, то $P^2 = P$. Докажем ее в обратную сторону.

Пусть $P^2 = P$. Положим $Q = 1 - P$. Тогда для операторов P и Q выполнены все соотношения (2.6). Рассмотрим два подпространства

$$U_1 = \operatorname{Im} P, \qquad\qquad U_2 = \operatorname{Ker} P.$$

Для произвольного вектора $\mathbf{v} \in V$ получаем разложение

$$\mathbf{v} = 1(\mathbf{v}) = (P + Q)\mathbf{v} = P(\mathbf{v}) + Q(\mathbf{v}), \tag{2.7}$$

где, очевидно, $\mathbf{u}_1 = P(\mathbf{v}) \in \operatorname{Im} P$. Из соотношения $PQ = 0$ для второго вектора $\mathbf{u}_2 = Q(\mathbf{v})$ получаем

$$P(\mathbf{u}_2) = P(Q(\mathbf{v})) = \mathbf{0}.$$

То есть $\mathbf{u}_2 \in \operatorname{Ker} P$. Отсюда $V = \operatorname{Im} P + \operatorname{Ker} P$. Докажем, что эта сумма подпространств является прямой. Для этого покажем однозначность разложения

$$\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2, \tag{2.8}$$

где $\mathbf{u}_1 \in \operatorname{Im} P$ и $\mathbf{u}_2 \in \operatorname{Ker} P$. Из $\mathbf{u}_1 \in \operatorname{Im} P$ заключаем, что $\mathbf{u}_1 = P(\mathbf{v}_1)$ для некоторого $\mathbf{v}_1 \in V$. А из условия $\mathbf{u}_2 \in \operatorname{Ker} P$ получаем, что $P(\mathbf{u}_2) = \mathbf{0}$. Тогда из (2.8) выводим следующие соотношения:

$$\begin{aligned} &P(\mathbf{v}) = P(\mathbf{u}_1) + P(\mathbf{u}_2) = P(P(\mathbf{v}_1)) = P^2(\mathbf{v}_1) = P(\mathbf{v}_1) = \mathbf{u}_1, \\ &Q(\mathbf{v}) = (\mathbf{1} - P)\,\mathbf{v} = \mathbf{v} - P(\mathbf{v}) = \mathbf{v} - \mathbf{u}_1 = \mathbf{u}_2. \end{aligned}$$

Полученные соотношения означают, что всякое разложение (2.8) совпадает с разложением (2.7). Значит, оно единственно и

$$V = \operatorname{Im} P \oplus \operatorname{Ker} P.$$

Оператор P отображает произвольный вектор $\mathbf{v} \in V$ в первую компоненту разложения (2.8). Следовательно, P есть оператор проектирования на $\operatorname{Im} P$ параллельно $\operatorname{Ker} P$. $\square$

Пусть теперь пространство V разложено в прямую сумму нескольких подпространств. Запишем это следующим образом:

$$V = U_1 \oplus \ldots \oplus U_s. \tag{2.9}$$

Такому разложению пространства V соответствует однозначное разложение произвольного вектора $\mathbf{v} \in V$

$$\mathbf{v} = \mathbf{u}_1 + \ldots + \mathbf{u}_s \tag{2.10}$$

в сумму векторов $\mathbf{u}_i \in U_i$.

Определение 2.2. Оператор $P_i\colon V \to V$, сопоставляющий каждому вектору $\mathbf{v} \in V$ его i-ую компоненту $P_i(\mathbf{v}) = \mathbf{u}_i$ в разложении (2.10), называется *оператором проектирования* на подпространство U_i параллельно остальным подпространствам.

Доказательство линейности операторов P_i практически ничем не отличается от случая, рассмотренного в теореме 2.1. Оно основано на однозначности разложения (2.10).

Выберем вектор $\mathbf{u} \in U_i$. Тогда его разложение (2.10) имеет следующий вид:

$$\mathbf{u} = \mathbf{0} + \ldots + \mathbf{0} + \mathbf{u} + \mathbf{0} + \ldots + \mathbf{0}.$$

Поэтому для такого вектора $P_i(\mathbf{u}) = \mathbf{u}$ и $P_j(\mathbf{u}) = \mathbf{0}$ при $j \neq i$. Для операторов проектирования из определения 2.2 это дает

$$(P_i)^2 = P_i, \qquad\qquad P_i \circ P_j = 0 \text{ при } i \neq j. \tag{2.11}$$

Кроме того, из самого определения операторов P_i получаем

$$P_1 + \ldots + P_s = 1. \tag{2.12}$$

В силу первого из соотношений (2.11) теория отдельных операторов P_i ничем не отличается от теории проекторов, заданных разложением на две компоненты. При многокомпонентных разложениях основной интерес представляет «коллективное поведение» этих операторов. Семейство операторов проектирования $P_1, \ldots, P_s$ называется *согласованным семейством проекторов*, если эти операторы удовлетворяет соотношениям (2.11) и (2.12).

Теорема 2.3. *Семейство операторов $P_1, \ldots, P_s$ является семейством проекторов, задаваемых некоторым разложением* (2.9)*, тогда и только тогда, когда для них выполнены соотношения* (2.11) *и* (2.12).

Док-во. В одну сторону теорема уже доказана. Докажем ее в обратную сторону. Пусть соотношения (2.11) и (2.12) для операторов $P_1, \ldots, P_s$ выполнены. Определим подпространства U_i так: $U_i = \operatorname{Im} P_i$. Тогда из (2.12) для произвольного вектора $\mathbf{v} \in V$ получаем

$$\mathbf{v} = P_1(\mathbf{v}) + \ldots + P_s(\mathbf{v}), \tag{2.13}$$

где $P_i(\mathbf{v}) \in \operatorname{Im} P_i$. Следовательно, для пространства V имеет место разложение в сумму подпространств

$$V = \operatorname{Im} P_1 + \ldots + \operatorname{Im} P_s. \tag{2.14}$$

Докажем, что сумма (2.14) прямая. Рассмотрим некоторое разложение произвольного вектора $\mathbf{v} \in V$, отвечающее разложению (2.14):

$$\mathbf{v} = \mathbf{u}_1 + \ldots + \mathbf{u}_s, \tag{2.15}$$

причем $\mathbf{u}_i \in \operatorname{Im} P_i$. Из $\mathbf{u}_i \in \operatorname{Im} P_i$ получаем $\mathbf{u}_i = P(\mathbf{v}_i)$ для некоторых векторов $\mathbf{v}_i \in V$. Тогда из (2.15) выводим

$$P_i(\mathbf{v}) = P_i(\mathbf{u}_1 + \ldots + \mathbf{u}_s) = \sum_{j=1}^{s} P_i(P_j(\mathbf{v}_j)).$$

В силу соотношений (2.11) только одно слагаемое в полученной сумме отлично от нуля, поэтому $P_i(\mathbf{v}) = (P_i)^2\,\mathbf{v}_i = P_i(\mathbf{v}_i) = \mathbf{u}_i$. Полученное равенство показывает, что произвольное разложение (2.15) обязано совпадать с (2.13). То есть сумма (2.14) прямая и P_i есть оператор проектирования на i-ую компоненту этой суммы параллельно остальным компонентам. Теорема доказана. $\square$

Сейчас на примере оператора проектирования P мы рассмотрим вопрос о приведении его матрицы к каноническому виду.

Теорема 2.4. *Для всякого ненулевого оператора проектирования в конечномерном пространстве V существует базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$, в котором матрица оператора проектирования P имеет вид*

$$\mathcal{P} = \begin{Vmatrix} \overbrace{1 & 0 & \ldots & 0}^{s} & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{Vmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix}1\\0\\\vdots\\0\end{matrix}}\right\} s \\ \vphantom{\begin{matrix}0\\\vdots\\0\end{matrix}} \end{array} \tag{2.16}$$

Док-во. Рассмотрим подпространства $\operatorname{Im} P$ и $\operatorname{Ker} P$. Из условия $P \neq 0$ заключаем, что $s = \dim(\operatorname{Im} P) \neq 0$. Выберем базис $\mathbf{e}_1, \ldots, \mathbf{e}_s$ в $U_1 = \operatorname{Im} P$, и в случае если $U_1 \neq V$ дополним его базисом $\mathbf{e}_{s+1}, \ldots, \mathbf{e}_n$ из $U_2 = \operatorname{Ker} P$. Сумма этих двух подпространств прямая $V = U_1 \oplus U_2$, поэтому, соединив выбранные базисы, мы получим базис в V (см. доказательство теоремы 6.3 из главы I).

Применим оператор P к векторам построенного базиса. Он проектирует на U_1 параллельно U_2, поэтому мы можем записать

$$P(\mathbf{e}_i) = \begin{cases} \mathbf{e}_i & \text{при } \ i = 1, \ldots, s, \\ \mathbf{0} & \text{при } \ i = s+1, \ldots, n. \end{cases}$$

Отсюда выводим вид (2.16) для матрицы оператора P в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_s$. $\square$

§ 3. Инвариантные подпространства. Сужение и факторизация операторов.

Пусть $f\colon V \to V$ — линейный оператор и пусть U — некоторое подпространство в V. Сузим область определения f до подпространства U. При этом образ $\operatorname{Im} f$ сузится до $f(U)$. Однако, никакой гарантии, что подпространство $f(U)$ окажется включенным в подпространство U, нет. Поэтому в общем случае *сужение* необходимо трактовать как линейное отображение $f_U\colon U \to V$, действующее из U в V.

Определение 3.1. Подпространство U называется *инвариантным подпространством* линейного оператора $f\colon V \to V$, если $f(U) \subseteq U$, то есть если из $\mathbf{u} \in U$ вытекает $f(\mathbf{u}) \in U$.

В случае инвариантности подпространства U относительно f, сужение f_U уже можно трактовать как линейный оператор, действующий в пространстве U. Его действие на векторы $\mathbf{u} \in U$ совпадает с действием f на $\mathbf{u}$. На векторы, не принадлежащие U, сужение f_U не действует по определению.

Теорема 3.1. *Ядро и образ линейного оператора $f : V \to V$ являются инвариантными подпространствами для f.*

Док-во. Начнем с доказательства инвариантности ядра f. Пусть $\mathbf{u}$ принадлежит $\operatorname{Ker} f$, тогда $f(\mathbf{u}) = \mathbf{0}$. Значит, $f(\mathbf{u}) \in \operatorname{Ker} f$, ибо нулевой вектор принадлежит любому подпространству. Инвариантность ядра установлена.

Пусть теперь $\mathbf{u} \in \operatorname{Im} f$. Положим $\mathbf{w} = f(\mathbf{u})$, тогда $\mathbf{w}$ есть образ вектора $\mathbf{u}$, в силу чего $\mathbf{w} = f(\mathbf{u}) \in \operatorname{Im} f$. Инвариантность образа также установлена. □

Теорема 3.2. *Пересечение любого числа инвариантных подпространств и сумма любого числа инвариантных подпространств оператора $f\colon V \to V$ являются инвариантными подпространствами этого оператора.*

Док-во. Пусть U_i, $i \in I$ — некоторое семейство инвариантных подпространств оператора $f\colon V \to V$. Рассмотрим их пересечение и их сумму:

$$U = \bigcap_{i\in I} U_i, \qquad\qquad W = \sum_{i\in I} U_i.$$

Тот факт, что U и W являются подпространствами, мы установили в § 6 первой главы. Начнем с доказательства инвариантности U. Пусть вектор $\mathbf{u}$ принадлежит U. Тогда этот вектор принадлежит всем подпространствам U_i, которые инвариантны относительно f. Поэтому $f(\mathbf{u})$ также принадлежит всем U_i, а значит, принадлежит и их пересечению. Инвариантность U доказана.

Теперь рассмотрим вектор $\mathbf{w} \in W$. По определению суммы подпространств он допускает разложение в сумму

$$\mathbf{w} = \mathbf{u}_{i_1} + \ldots + \mathbf{u}_{i_s}, , \text{ где } \mathbf{u}_{i_r} \in U_{i_r}.$$

Применив оператор f к обеим частям этого равенства, получаем аналогичное разложение для вектора $f(\mathbf{w})$:

$$f(\mathbf{w}) = f(\mathbf{u}_{i_1}) + \ldots + f(\mathbf{u}_{i_s}),$$

причем из инвариантности U_i имеем $f(\mathbf{u}_{i_r}) \in U_{i_r}$. Отсюда $f(\mathbf{w}) \in W$, что доказывает инвариантность суммы инвариантных подпространств. $\square$

Пусть U — инвариантное подпространство оператора $f\colon V \to V$. Рассмотрим факторпространство V/U и определим оператор f_U, действующий в этом факторпространстве по следующей формуле:

$$f_{V/U}(Q) = \mathrm{Cl}_U(f(\mathbf{v})), \text{ где } Q = \mathrm{Cl}_U(\mathbf{v}). \tag{3.1}$$

Оператор $f_{V/U}\colon V/U \to V/U$, действующий по правилу (3.1), называется *фактороператором* оператора f по инвариантному подпространству U. Более кратко формулу (3.1) можно переписать так:

$$f_{V/U}(\mathrm{Cl}_U(\mathbf{v})) = \mathrm{Cl}_U(f(\mathbf{v})). \tag{3.2}$$

Формулы (3.1) и (3.2), подобно формулам (7.3) из первой главы, содержат элемент неопределенности, связанный с произволом в выборе представителя $\mathbf{v}$ в классе Q. Поэтому требуется доказать их корректность.

Теорема 3.3. *Формула* (3.1) *и эквивалентная ей формула* (3.2) *корректны и определяют линейный оператор $f_{V/U}$ в факторпространстве V/U.*

Док-во. Рассмотрим два возможных выбора представителя в смежном классе Q, то есть пусть $\mathbf{v}, \tilde{\mathbf{v}} \in Q$. Тогда $\tilde{\mathbf{v}} - \mathbf{v} \in U$. Следуя формуле (3.1), определим два возможных результата применения оператора $f_{V/U}$ к классу Q, исходя из выбранных предтавителей этого класса:

$$f_{V/U}(Q) = \mathrm{Cl}_U(f(\mathbf{v})), \qquad\qquad f_{V/U}(Q) = \mathrm{Cl}_U(f(\tilde{\mathbf{v}})).$$

Вычислим разницу этих двух возможных результатов:

$$\mathrm{Cl}_U(f(\tilde{\mathbf{v}})) - \mathrm{Cl}_U(f(\mathbf{v})) = \mathrm{Cl}_U(f(\tilde{\mathbf{v}}) - f(\mathbf{v})) = \mathrm{Cl}_U(f(\tilde{\mathbf{v}} - \mathbf{v})).$$

Но разность $\tilde{\mathbf{v}} - \mathbf{v} = \mathbf{u}$ принадлежит подпространству U. В силу инвариантности этого подпространства $\tilde{\mathbf{u}} = f(\mathbf{u}) \in U$. Поэтому

$$\mathrm{Cl}_U(f(\tilde{\mathbf{v}})) - \mathrm{Cl}_U(f(\mathbf{v})) = \mathrm{Cl}_U(\tilde{\mathbf{u}}) = \mathbf{0}.$$

Доказанное совпадение $\mathrm{Cl}_U(f(\tilde{\mathbf{v}})) = \mathrm{Cl}_U(f(\mathbf{v}))$ устанавливает корректность формулы (3.1), а значит, и формулы (3.2).

Теперь докажем линейность оператора $f_{V/U}\colon V/U \to V/U$. Соответствующие выкладки проделаем на базе формулы (3.1):

$$\begin{aligned}&f_{V/U}(Q_1 + Q_2) = f_{V/U}(\mathrm{Cl}_U(\mathbf{v}_1) + \mathrm{Cl}_U(\mathbf{v}_2)) = f_{V/U}(\mathrm{Cl}_U(\mathbf{v}_1 + \mathbf{v}_2)) =\\ &\qquad = \mathrm{Cl}_U(f(\mathbf{v}_1 + \mathbf{v}_2)) = \mathrm{Cl}_U(f(\mathbf{v}_1)) + \mathrm{Cl}_U(f(\mathbf{v}_2)) = f_{V/U}(Q_1) + f_{V/U}(Q_2),\end{aligned}$$

$$\begin{aligned}&f_{V/U}(\alpha \cdot Q) = f_{V/U}(\alpha \cdot \mathrm{Cl}_U(\mathbf{v})) = f_{V/U}(\mathrm{Cl}_U(\alpha \cdot \mathbf{v})) =\\ &\qquad = \mathrm{Cl}_U(f(\alpha \cdot \mathbf{v})) = \mathrm{Cl}_U(\alpha \cdot f(\mathbf{v})) = \alpha \cdot \mathrm{Cl}_U(f(\mathbf{v})) = \alpha \cdot f_{V/U}(Q).\end{aligned}$$

Проделанные выкладки доказывают линейность фактороператора $\mathrm{f}_{V/U}$. □

Теорема 3.4. *Пусть U — общее инвариантное подпространство для операторов $f, g \in \operatorname{End}(V)$. Тогда U является инвариантным подпространством и для операторов $f+g$, $\alpha \cdot f$ и $f \circ g$, причем для сужений и факторoператоров выполнены следующие соотношения:*

$$\begin{aligned}
&(f+g)_U = f_U + g_U\,; && (f+g)_{V/U} = f_{V/U} + g_{V/U}\,;\\
&(\alpha \cdot f)_U = \alpha \cdot f_U\,; && (\alpha \cdot f)_{V/U} = \alpha \cdot f_{V/U}\,;\\
&(f \circ g)_U = f_U \circ g_U\,; && (f \circ g)_{V/U} = f_{V/U} \circ g_{V/U}\,.
\end{aligned}$$

Док-во. Начнем с первого случая. Пусть $h = f + g$ и пусть $\mathbf{u} \in U$. Тогда $f(\mathbf{u}) \in U$ и $g(\mathbf{u}) \in U$ в силу инвариантности U относительно f и g. Поэтому $h(\mathbf{u}) = f(\mathbf{u}) + g(\mathbf{u}) \in U$. Это доказывает инвариантность U относительно h. Соотношение $h_U = f_U + g_U$ вытекает из $h = f + g$, поскольку действие сужений на вектор $\mathbf{u}$ из U не отличается от действия соответствующих операторов на этот вектор. Соотношение для факторoператоров вытекает из

$$\begin{gathered}
h_{V/U}\,\mathrm{Cl}_U(\mathbf{v}) = \mathrm{Cl}_U(h(\mathbf{v})) = \mathrm{Cl}_U(f(\mathbf{v}) + h(\mathbf{v})) =\\
= \mathrm{Cl}_U(f(\mathbf{v})) + \mathrm{Cl}_U(g(\mathbf{v})) = f_{V/U}\,\mathrm{Cl}_U(\mathbf{v}) + g_{V/U}\,\mathrm{Cl}_U(\mathbf{v}).
\end{gathered}$$

Случай $h = \alpha \cdot f$ мало чем отличается от предыдущего. Из $\mathbf{u} \in U$ вытекает $f(\mathbf{u}) \in U$, откуда $h(\mathbf{u}) = \alpha \cdot f(\mathbf{u}) \in U$. Соотношение $h_U = \alpha \cdot f_U$ очевидно в силу соображений, изложенных выше. Для факторoператора имеем:

$$\begin{gathered}
h_{V/U}\,\mathrm{Cl}_U(\mathbf{v}) = \mathrm{Cl}_U(h(\mathbf{v})) = \mathrm{Cl}_U(\alpha \cdot f(\mathbf{v}) =\\
= \alpha \cdot \mathrm{Cl}_U(f(\mathbf{v})) = \alpha \cdot f_{V/U}\,\mathrm{Cl}_U(\mathbf{v}) = (\alpha \cdot f_{V/U})\,\mathrm{Cl}_U(\mathbf{v}).
\end{gathered}$$

Теперь рассмотрим композицию $h = f \circ g$. Из $\mathbf{u} \in U$ выводим $\mathbf{w} = g(\mathbf{u}) \in U$, затем выводим $f(\mathbf{w}) \in U$, что означает инвариантность U относительно h: $h(\mathbf{u}) = f(g(\mathbf{u})) = f(\mathbf{w}) \in U$. Для сужений это дает:

$$h_U(\mathbf{u}) = h(\mathbf{u}) = f(g(\mathbf{u})) = f_U(g_U(\mathbf{u})),$$

что доказывает $h_U = f_U \circ g_U$. Для факторoператоров имеем:

$$\begin{gathered}
h_{V/U}\,\mathrm{Cl}_U(\mathbf{v}) = \mathrm{Cl}_U(h(\mathbf{v})) = \mathrm{Cl}_U(f(g(\mathbf{v})) = f_{V/U}\,\mathrm{Cl}(g(\mathbf{v})) =\\
= f_{V/U}(g_{V/U}\,\mathrm{Cl}_U(\mathbf{v})) = (f_{V/U} \circ g_{V/U})\,\mathrm{Cl}_U(\mathbf{v}).
\end{gathered}$$

Проделанные выкладки доказывают последнее соотношение теоремы 3.4. □

Теорема 3.5. *Пусть $V = U_1 \oplus \ldots \oplus U_s$ — разложение пространства V в прямую сумму своих подпространств. Подпространства $U_1, \ldots, U_s$ инвариантны относительно оператора $f\colon V \to V$ тогда и только тогда, когда соответствующие этому разложению проекторы $P_1, \ldots, P_s$ перестановочны с оператором f, то есть $f\,P_i = P_i\,f$ или, что эквивалентно, $f \circ P_i = P_i \circ f$.*

Док-во. Пусть все подпространства U_i инвариантны относительно f. Для произвольного $\mathbf{v} \in V$ у нас имеется разложение

$$\mathbf{v} = \mathbf{u}_1 + \ldots + \mathbf{u}_s,$$

где $\mathbf{u}_i \in U_i$ и $\mathbf{u}_i = P_i(\mathbf{v})$. Из этого разложения выводим

$$P_i(f(\mathbf{v})) = P_i(f(\mathbf{u}_1) + \ldots + f(\mathbf{u}_s)) = f(\mathbf{u}_i) = f(P_i(\mathbf{v})).$$

Здесь мы использовали $\mathbf{w}_j = f(\mathbf{u}_j) \in U_j$, что вытекает из инвариантности U_j относительно f, и свойство проекторов:

$$P_i(\mathbf{w}_j) = \begin{cases} \mathbf{w}_i & \text{при } j = i, \\ \mathbf{0} & \text{при } j \neq i. \end{cases}$$

Из произвольности вектора $\mathbf{v}$ в полученном равенстве $P_i(f(\mathbf{v})) = f(P_i(\mathbf{v}))$ выводим перестановочность $f\,P_i = P_i\,f$.

Пусть, наоборот, известна перестановочность оператора f со всеми проекторами $P_1, \ldots, P_s$, отвечающими разложению пространства V в прямую сумму подпространств $U_1, \ldots, U_s$. Пусть $\mathbf{u} \in U_i$ и пусть $\mathbf{w} = f(\mathbf{u})$. Тогда

$$P_i(\mathbf{w}) = P_i(f(\mathbf{v})) = f(P_i(\mathbf{u})) = f(\mathbf{u}) = \mathbf{w}.$$

Но $U_i = \operatorname{Im} P_i$, поэтому из $\mathbf{w} = P_i(\mathbf{w})$ заключаем, что $\mathbf{w} = f(\mathbf{u}) \in U_i$. Это доказывает инвариантность U_i относительно оператора f. $\square$

Рассмотрим оператор f, действующий в конечномерном пространстве V и обладающий инвариантным подпространством U. Пусть $\dim V = n$ и $\dim U = s$. Выберем базис $\mathbf{e}_1, \ldots, \mathbf{e}_s$ в U и дополним его до базиса в V векторами $\mathbf{e}_{s+1}, \ldots, \mathbf{e}_n$. При $j \leq s$ из инвариантности пространства U относительно оператора f имеем $f(\mathbf{e}_j) \in U$. Поэтому в разложениях

$$f(\mathbf{e}_j) = \sum_{i=1}^{s} F^i_j \cdot \mathbf{e}_i$$

для $j \leq s$ суммирование идет от единицы до s, а не до n. Это означает, что матрица оператора f в выбранном базисе имеет блочную структуру, причем левый нижний блок равен нулю:

$$F = \begin{array}{c} \overbrace{\qquad\qquad\qquad\qquad}^{s}\qquad\qquad\qquad\qquad \\ \left\| \begin{array}{ccccccc} F^1_1 & F^1_2 & \ldots & F^1_s & F^1_{s+1} & \ldots & F^1_n \\ F^2_1 & F^2_2 & \ldots & F^2_s & F^2_{s+1} & \ldots & F^2_n \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ F^s_1 & F^s_2 & \ldots & F^4_s & F^s_{s+1} & \ldots & F^s_n \\ 0 & 0 & \ldots & 0 & F^{s+1}_{s+1} & \ldots & F^{s+1}_n \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & F^n_{s+1} & \ldots & F^n_n \end{array} \right\| \end{array} \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \\ \end{array}\right\} s \\ \\ \\ \\ \end{array} \tag{3.3}$$

Матрицы такого вида называются *блочно-треугольными*. Левый верхний диагональный блок в матрице (3.3) совпадает с матрицей сужения $f_U\colon U \to U$.

Выясним смысл правого нижнего блока в этой матрице. Рассмотрим факторпространство V/U и классы дополнительных векторов в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$:

$$\mathbf{E}_1 = \mathrm{Cl}_U(\mathbf{e}_{s+1}),\ \ldots,\ \mathbf{E}_{n-s} = \mathrm{Cl}_U(\mathbf{e}_n). \tag{3.4}$$

При доказательстве теоремы 7.6 в первой главе мы установили, что эти классы образуют базис в факторпространстве V/U. Подействовав на них фактороператором $f_{V/U}$, получаем

$$\begin{aligned} f_{V/U}(\mathbf{E}_j) &= f_{V/U}\,\mathrm{Cl}_U(\mathbf{e}_{s+j}) = \mathrm{Cl}_U(f(\mathbf{e}_{s+j})) = \\ &= \sum_{i=1}^{s} F^i_{s+j}\cdot \mathrm{Cl}_U(\mathbf{e}_i) + \sum_{i=s+1}^{n} F^i_{s+j}\cdot \mathrm{Cl}_U(\mathbf{e}_i) \end{aligned}$$

Первая из сумм в этом выражении зануляется в силу того, что $\mathbf{e}_1, \ldots, \mathbf{e}_s$ принадлежат U. Тогда, сделав сдвиг индекса во второй сумме, находим:

$$f_{V/U}(\mathbf{E}_j) = \sum_{i=1}^{n-s} F^{s+i}_{s+j}\cdot \mathbf{E}_i.$$

Из этой формулы видим, что матрица факторопepатора в базисе (3.4) в точности совпадает с правым нижним диагональным блоком в матрице (3.3).

Теорема 3.6. *Пусть $f\colon V \to V$ линейный оператор в конечномерном пространстве, обладающий инвариантным подпространством U. Тогда его детерминант равен произведению детерминантов сужения и факторопepатора:*

$$\det f = (\det f_U)\,(\det f_{V/U}).$$

Доказательство этой теоремы вытекает из следующего известного факта из теории детерминантов: определитель блочно-треугольной матрицы равен произведению определителей всех диагональных блоков в ней.

§ 4. Собственные числа и собственные векторы.

Пусть $f : V \to V$ линейный оператор. Вектор $\mathbf{v} \neq \mathbf{0}$ из пространства V называется *собственным вектором* оператора f, если $f\,\mathbf{v} = \lambda\cdot\mathbf{v}$, где $\lambda \in \mathbb{K}$. Число λ называется *собственным числом* оператора f, отвечающим собственному вектору $\mathbf{v}$.

Одному собственному числу λ оператора f может соответствовать несколько или даже бесконечно много собственных векторов. Однако, по заданному собственному вектору соответствующее собственное число λ определяется однозначно. Действительно, из равенства $f\,\mathbf{v} = \lambda\cdot\mathbf{v} = \lambda'\cdot\mathbf{v}$ и из условия $\mathbf{v} \neq \mathbf{0}$ вытекает равенство $\lambda = \lambda'$.

Пусть $\mathbf{v}$ собственный вектор оператора $f\colon V \to V$. Рассмотрим оператор $h_\lambda = f - \lambda\cdot 1$. Уравнение $f\,\mathbf{v} = \lambda\cdot\mathbf{v}$, которому удовлетворяет собственный вектор $\mathbf{v}$, можно теперь переписать в форме

$$(f - \lambda\cdot 1)\,\mathbf{v} = \mathbf{0}. \tag{4.1}$$

Значит, $\mathbf{v} \in \operatorname{Ker}(f - \lambda \cdot 1)$. Условие $\mathbf{v} \neq \mathbf{0}$ означает, что ядро этого оператора отлично от нуля: $\operatorname{Ker}(f - \lambda \cdot 1) \neq \{\mathbf{0}\}$.

Определение 4.1. Число $\lambda \in \mathbb{K}$ называется собственным числом оператора $f\colon V \to V$, если подпространство $V_\lambda = \operatorname{Ker}(f - \lambda \cdot 1)$ отлично от нуля. Это подпространство $V_\lambda = \operatorname{Ker}(f - \lambda \cdot 1) \neq \{\mathbf{0}\}$ называется *собственным подпространством*, отвечающим собственному числу λ, а любой ненулевой вектор из V_λ называется собственным вектором оператора f, отвечающим λ.

Совокупность всех собственных чисел оператора f иногда называют *спектром* оператора f, а раздел математики, изучающий спектры операторов, называется *спектральной теорией операторов*. Наиболее просто устроена спектральная теория линейных операторов в конечномерных пространствах. Именно ее принято изучать в курсе линейной алгебры.

Пусть $f\colon V \to V$ линейный оператор в конечномерном пространстве. Для нахождения спектра такого оператора применим следствие из теоремы 1.4. Из него сразу получаем, что число $\lambda \in \mathbb{K}$ является собственным числом оператора f тогда и только тогда, когда оно удовлетворяет уравнению

$$\det(f - \lambda \cdot 1) = 0. \tag{4.2}$$

Уравнение (4.2) называется *характеристическим уравнением* оператора f, а его корни называются *характеристическими числами* оператора f.

Пусть $\dim V = n$. Вычисление детерминанта в формуле (4.2), по сути, сводится к вычислению детерминанта квадратной матрицы размера $n \times n$. Матрица оператора $h_\lambda = f - \lambda \cdot 1$ получается из матрицы оператора f путем вычитания числа λ из каждого элемента на главной диагонали:

$$H_\lambda = \begin{Vmatrix} F_1^1 - \lambda & F_2^1 & \ldots & F_n^1 \\ F_1^2 & F_2^2 - \lambda & \ldots & F_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ F_1^n & F_2^n & \ldots & F_n^n - \lambda \end{Vmatrix}. \tag{4.3}$$

Зависимость детерминанта матрицы (4.3) от λ носит полиномиальный характер. Поэтому для $\det(f - \lambda \cdot 1)$ мы можем записать

$$\det(f - \lambda \cdot 1) = (-\lambda)^n + F_1(-\lambda)^{n-1} + \ldots + F_n. \tag{4.4}$$

Полином в правой части (4.4) называется *характеристическим полиномом* оператора f. Коэффициенты $F_1, \ldots, F_n$ в характеристическом полиноме (4.4) находятся по матрице оператора f, вычисленной в каком-либо базисе. Левая часть (4.4) не зависит от выбора базиса, поэтому коэффициенты $F_1, \ldots, F_n$ также не зависят от выбора базиса. Они являются скалярными инвариантами оператора f. Наиболее известны первый и последний инварианты в (4.4):

$$F_1 = \operatorname{tr} f, \qquad\qquad F_n = \det f.$$

Инвариант F_1 называется *следом* оператора f. Он вычисляется по матрице F этого оператора в соответствии со следующей формулой:

$$\operatorname{tr} f = \sum_{i=1}^{n} F_i^i. \tag{4.5}$$

Мы не будем доказывать формулы (4.5), поскольку это известный факт из теории детерминантов. Мы лишь выведем инвариантность следа непосредственным образом на базе формулы (1.4) для преобразования матрицы линейного оператора при замене базиса:

$$\sum_{p=1}^{n} \tilde{F}_p^p = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\sum_{p=1}^{n} T_i^p\, S_p^j \right) F_j^i = \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_i^j\, F_j^i = \sum_{i=1}^{n} F_i^i.$$

После подстановки полинома (4.4) в (4.2) мы видим, что характеристическое уравнение (4.2) оператора f является полиномиальным уравнением n-ой степени относительно переменной λ:

$$(-\lambda)^n + F_1(-\lambda)^{n-1} + \ldots + F_n = 0. \tag{4.6}$$

Это позволяет делать выводы о количестве собственных чисел оператора f. Всякое собственное число $\lambda \in \mathbb{K}$ является корнем характеристического уравнения (4.6), однако не всякий корень уравнения (4.6) есть собственное число. Для того, чтобы характеристическое число λ оператора f было собственным числом, оно должно принадлежать полю $\mathbb{K}$. Из курса общей алгебры известно, что полное число корней уравнения (4.6) с учетом их кратностей равно n (см. [4]). Однако, некоторые из них, а может быть и все они, могут не принадлежать полю $\mathbb{K}$.

Теорема 4.1. *Количество собственных чисел оператора $f\colon V \to V$ не превосходит размерности пространства V.*

Рассмотрим случай $\mathbb{K} = \mathbb{Q}$. Корни полиномиального уравнения с рациональными коэффициентами не обязаны быть рациональными: в качестве примера достаточно рассмотреть уравнение $\lambda^2 - 3 = 0$. В вещественном случае $\mathbb{K} = \mathbb{R}$ полиномиальное уравнение с вещественными коэффициентами также может не иметь вещественных корней, например, уравнение $\lambda^2 + \sqrt{3} = 0$. Исключением в этом ряду является поле комплексных чисел.

Теорема 4.2. *Всякое полиномиальное уравнение n-ой степени с комплексными коэффициентами имеет ровно n корней в поле комплексных чисел $\mathbb{C}$ с учетом их кратностей.*

Мы не будем приводить здесь доказательство этой теоремы ссылаясь на курс общей алгебры (см. [4]). Теорема 4.2 известна как «основная теорема алгебры», а соответсвующее свойство поля $\mathbb{C}$ называется *алгебраической замкнутостью.*

Определение 4.2. Поле $\mathbb{K}$ называется алгебраически замкнутым, если корни всякого полиномиального уравнения с коэффициентами из $\mathbb{K}$ принадлежат полю $\mathbb{K}$.

Разумеется, поле $\mathbb{C}$ не является единственным алгебраически замкнутым полем. Однако, в списке полей $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, которые мы рассматриваем в этой книге, только поле комплексных чисел является алгебраически замкнутым.

Пусть λ — собственное число оператора f. Тогда λ есть корень уравнения (4.6). Кратность корня λ в уравнении (4.6) называется *кратностью* собственного числа λ.

Теорема 4.3. *Для линейного оператора $f\colon V \to V$ в комплексном векторном пространстве V количество собственных чисел с с учетом их кратностей в точности равно размерности пространства V.*

Это усиление теоремы 4.1 является простым следствием алгебраической замкнутости поля $\mathbb{C}$. В случае $\mathbb{K} = \mathbb{C}$ характеристический полином (4.4) раскладывается в произведение линейных множителей:

$$\det(f - \lambda \cdot \mathbf{1}) = \prod_{i=1}^{n} (\lambda_i - \lambda). \tag{4.7}$$

Для некоторых операторов это разложение может иметь место и в случае $\mathbb{K} = \mathbb{Q}$ или $\mathbb{K} = \mathbb{R}$, однако, здесь такая ситуация уже не является типичной. Если под $\lambda_1, \ldots, \lambda_n$ понимать не собственные а характеристические числа оператора f, разложение (4.7) имеется всегда.

Разложение (4.7) позволяет представить числовые инварианты $F_1, \ldots, F_n$ оператора f в виде элементарных симметрических многочленов от его характеристических чисел $F_i = \sigma_i(\lambda_1, \ldots, \lambda_n)$. В частности, для следа и для детерминанта оператора f это дает:

$$\operatorname{tr} f = \sum_{i=1}^{n} \lambda_i, \qquad\qquad \det f = \prod_{i=1}^{n} \lambda_i. \tag{4.8}$$

Теория симметрических многочленов излагается в курсе общей алгебры (см., например, книгу [4]).

Теорема 4.4. *Для всякого собственного числа λ оператора $f : V \to V$ собственное подпространство V_λ инвариантно относительно f.*

Док-во. Определение 4.1 собственного подпространства V_λ оператора f можно переформулировать следующим образом:

$$V_\lambda = \{\mathbf{v} \in V \colon f(\mathbf{v}) = \lambda \cdot \mathbf{v}\}.$$

Поэтому из $\mathbf{v} \in V_\lambda$ вытекает $f(\mathbf{v}) = \lambda \cdot \mathbf{v} \in V_\lambda$, что доказывает инвариантность собственного подпространства V_λ. $\square$

Мы знаем, что множество всех линейных операторов, действующих в пространстве V, образует алгебру $\operatorname{End}(V)$ над полем $\mathbb{K}$. Однако, эта алгебра слишком обширна. Рассмотрим один оператор $f \in \operatorname{End}(V)$ и дополним его единичным оператором 1. Оставаясь в пределах $\operatorname{End}(V)$, мы можем рассмотреть целые положительные степени оператора f, можем умножать их на числа из $\mathbb{K}$, можем складывать результаты таких умножений и можем добавлять к ним скалярные операторы, полученные умножением единичного оператора 1 на числа из $\mathbb{K}$. В результате всех этих действий мы получим разнообразные операторы, каждый из которых имеет следующий вид:

$$P(f) = \alpha_p \cdot f^p + \ldots + \alpha_1 \cdot f + \alpha_0 \cdot 1. \tag{4.9}$$

Множество всех операторов вида (4.9) называется *полиномиальной оболочкой* оператора f и обозначается $\mathbb{K}[f]$. Оно замкнуто относительно всех алгебраических операций в алгебре $\operatorname{End}(V)$. Такие подмножества принято называть *подалгебрами.* Важно отметить, что подалгебра $\mathbb{K}[f]$ коммутативна, то есть для любых двух полиномов P и Q,операторы вида (4.9) перестановочны:

$$P(f)\,Q(f) = Q(f)\,P(f). \tag{4.10}$$

Это проверяется непосредственным вычислением. Действительно, пусть оператор $P(f)$ и оператор $Q(f)$ имеют вид (4.9):

$$P(f) = \sum_{i=0}^{p} \alpha_i \cdot f^i, \qquad\qquad Q(f) = \sum_{j=0}^{q} \beta_j \cdot f^j.$$

Здесь мы считаем: $f^{\,0} = 1$. Это соотношение следует рассматривать как определение нулевой степени оператора f. Тогда

$$P(f)\,Q(f) = \sum_{i=0}^{p}\sum_{j=0}^{q} (\alpha_i \beta_j) \cdot f^{i+j} = Q(f)\,P(f).$$

Проделанные вычисления доказывают перестановочность операторов (4.10).

Теорема 4.5. *Пусть U — инвариантное подпространство оператора f. Тогда это подпространство инвариантно относительно любого оператора $P(f)$ из полиномиальной оболочки $\mathbb{K}[f]$.*

Док-во. Пусть $\mathbf{u}$ — некоторый произвольный вектор из инвариантного подпространства U. Рассмотрим следующие вектора: $\mathbf{u}_0 = \mathbf{u}$, $\mathbf{u}_1 = f(\mathbf{u})$, $\mathbf{u}_2 = f^2(\mathbf{u})$, $\ldots$, $\mathbf{u}_p = f^p(\mathbf{u})$. Каждый следующий вектор в этой последовательности векторов получается применением оператора f к предыдущему вектору: $\mathbf{u}_{i+1} = f(\mathbf{u}_i)$. Поэтому из $\mathbf{u}_0 \in U$ в силу инвариантности подпространства U вытекает $\mathbf{u}_1 \in U$. Из этого, в свою очередь, вытекает $\mathbf{u}_2 \in U$, и т. д. вплоть до $\mathbf{u}_p \in U$. Применив оператор $P(f)$ вида (4.9) к вектору $\mathbf{u} \in U$, получим вектор $P(f)\,\mathbf{u}$ следующего вида:

$$P(f)\,\mathbf{u} = \alpha_p \cdot \mathbf{u}_p + \ldots + \alpha_1 \cdot \mathbf{u}_1 + \alpha_0 \cdot \mathbf{u}_0.$$

Отсюда в силу $\mathbf{u}_i \in U$ находим, что $P(f)\,\mathbf{u} \in U$, что доказывает инвариантность U относительно оператора $P(f)$. $\square$

Любопытно отметить следующий факт: если λ — собственное число оператора f и $\mathbf{v}$ — соответствующий собственный вектор, то $P(f)\,\mathbf{v} = P(\lambda) \cdot \mathbf{v}$. Поэтому всякий собственный вектор $\mathbf{v}$ оператора f является собственным вектором для $P(f)$. Обратное, вообще говоря, неверно.

Пусть $\lambda_1, \ldots, \lambda_s$ — набор попарно различных собственных чисел оператора f. Рассмотрим операторы $h_i = f - \lambda_i \cdot 1$, которые принадлежат полиномиальной оболочке f. Перестановочность произвольных двух таких операторов вытекает из (4.10). Собственное подпространство V_{λ_i}, соответствующее собственному числу λ_i оператора f, определяется как ядро оператора h_i. Согласно определению 4.1 оно отлично от нуля, а из теорем 4.4 и 4.5 вытекает, что V_{λ_i} инвариантно относительно f и всех операторов h_j.

Теорема 4.6. *Пусть $\lambda_1, \ldots, \lambda_s$ набор попарно различных собственных чисел оператора $f: V \to V$. Тогда сумма соответствующих собственных пространств $V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_s}$ является прямой.*

Отметим, что набор попарно различных собственных чисел $\lambda_1, \ldots, \lambda_s$ в теореме может быть полным набором собственных чисел оператора f или же он может включать лишь часть из собственных чисел этого оператора. На результат теоремы это обстоятельство не влияет.

Док-во. Обозначим через W сумму собственных подпространств

$$W = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_s}. \tag{4.11}$$

Для доказательства того, что сумма (4.11) прямая, надо для произвольного $\mathbf{w} \in W$ установить однозначность разложения

$$\mathbf{w} = \mathbf{v}_1 + \ldots + \mathbf{v}_s, \text{ где } \mathbf{v}_i \in V_{\lambda_i}. \tag{4.12}$$

Построим оператор f_i, определив его по следующей формуле:

$$f_i = \prod_{r \neq i}^{s} h_r.$$

Оператор f_i принадлежит полиномиальной оболочке оператора f, причем

$$f_i(\mathbf{v}_j) = \left(\prod_{r \neq i}^{s} (\lambda_j - \lambda_r) \right) \cdot \mathbf{v}_j. \tag{4.13}$$

Это вытекает из $\mathbf{v}_j \in V_{\lambda_j}$, в силу чего $h_r(\mathbf{v}_j) = (\lambda_j - \lambda_r) \cdot \mathbf{v}_j$. Из формулы (4.13) видим, что $f_i(\mathbf{v}_j) = \mathbf{0}$ для всех $j \neq i$. Применив оператор f_i к обеим частям разложения (4.12), мы получим

$$f_i(\mathbf{w}) = \left(\prod_{r \neq i}^{s} (\lambda_i - \lambda_r) \right) \cdot \mathbf{v}_i.$$

Отсюда для вектора $\mathbf{v}_i$ в разложении (4.12) выводим

$$\mathbf{v}_i = \frac{f_i(\mathbf{w})}{\displaystyle\prod_{r \neq i}^{s} (\lambda_i - \lambda_r)}. \tag{4.14}$$

Формула (4.14) однозначно определяет все компоненты разложения (4.12) по заданному вектору $\mathbf{w} \in W$. Следовательно, сумма собственных подпространств (4.11) прямая. $\square$

Определение 4.3. Оператор $f : V \to V$ называется *диагонализуемым*, если существует базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в пространстве V, в котором матрица этого оператора диагональна.

Теорема 4.7. *Оператор $f : V \to V$ диагонализуем тогда и только тогда, когда сумма всех его собственных подпространств совпадает с V.*

Док-во. Пусть оператор f диагонализуем. Выберем базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$, в котором его матрица F диагональна, то есть в ней могут быть отличны от нуля только элементы вида F^i_i. Тогда соотношение (1.2), определяющее матрицу F, принимает вид: $f(\mathbf{e}_i) = F^i_i \cdot \mathbf{e}_i$. Значит, каждый из базисных векторов $\mathbf{e}_i$ выбранного базиса является собственным вектором для f, а $\lambda_i = F^i_i$ есть соответствующее собственное число. Разложение произвольного вектора $\mathbf{v}$ по такому базису есть разложение по собственным векторам оператора f. Сгруппировав вместе и просуммировав в таком разложении слагаемые с совпадающими собственными числами, мы получим разложение

$$\mathbf{w} = \mathbf{v}_1 + \ldots + \mathbf{v}_s, \text{ где } \mathbf{v}_i \in V_{\lambda_i},$$

из которого вытекает $V_{\lambda_1} + \ldots + V_{\lambda_s} = V$. В одну сторону утверждение теоремы 4.7 доказано.

Пусть, наоборот, $\lambda_1, \ldots, \lambda_s$ — полный набор собственных чисел оператора f и пусть известно, что $V_{\lambda_1} + \ldots + V_{\lambda_s} = V$. Из теоремы 4.6 мы знаем, что эта сумма прямая. Поэтому, выбрав базисы в каждом из собственных подпространств и объединив их вместе, мы получим базис в V (см. теорему 6.3 из первой главы). Полученный базис есть базис из собственных векторов, действие f на каждый базисный вектор сводится к умножению этого вектора на соответствующее собственное число. Поэтому матрица оператора f в таком базисе диагональна. На ее диагонали стоят собственные числа f. $\square$

Пусть оператор $f : V \to V$ диагонализуем и пусть выбран базис, где его матрица диагональна. Тогда матрица H_λ в формуле (4.3) тоже диагональна. Отсюда сразу же выводим соотношение

$$\det(f - \lambda \cdot 1) = \prod_{i=1}^{n} (F^i_i - \lambda).$$

Из этого соотношения видим, что характеристический полином диагонализуемого оператора раскладывается в произведение линейных множителей и все его корни принадлежат полю $\mathbb{K}$. Значит, все характеристические числа диагонализуемого оператора являются собственными. Это условие является необходимым условием диагонализуемости оператора f, но оно далеко не достаточное. Даже в случае алгебраически замкнутого поля комплексных чисел $\mathbb{K} = \mathbb{C}$ имеются недиагонализуемые операторы, действующие в пространстве над полем комплексных чисел $\mathbb{C}$.

§ 5. Нильпотентные операторы.

Определение 5.1. Линейный оператор $f : V \to V$ называется *нильпотентным*, если для любого вектора $\mathbf{v} \in V$ существует целое положительное число k, такое, что $f^k(\mathbf{v}) = \mathbf{0}$.

Согласно определению 5.1 для любого вектора $\mathbf{v}$ существует число k (зависящее от $\mathbf{v}$), для которого $f^k(\mathbf{v}) = \mathbf{0}$. Выбор такого числа не ограничен сверху, если m больше k, то из $f^k(\mathbf{v}) = \mathbf{0}$ вытекает $f^m(\mathbf{v}) = \mathbf{0}$. Поэтому существует некоторое минимальное число $k = k_{\min}$, для которого $f^k(\mathbf{v}) = \mathbf{0}$. Это минимальное число $k_{\min}$, которое, разумеется, зависит от $\mathbf{v}$, принято называть *высотой вектора* $\mathbf{v}$ относительно нильпотентного оператора f. Высоту нулевого вектора принято считать равной нулю, высота ненулевого вектора больше нуля. Обозначим высоту вектора $\mathbf{v}$ через $\nu(\mathbf{v})$ и положим

$$\nu(f) = \max_{\mathbf{v} \in V} \nu(\mathbf{v}). \tag{5.1}$$

Для каждого вектора $\mathbf{v} \in V$ его высота конечна, однако, максимум в (5.1) может оказаться бесконечным, ибо число векторов в пространстве V обычно бывает бесконечным.

Определение 5.2. В случае, когда максимум в формуле (5.1) конечен, нильпотентный оператор f называется оператором *конечной высоты*, а число $\nu(f)$ называется *высотой нильпотентного оператора* f.

Теорема 5.1. *В конечномерном пространстве V высота $\nu(f)$ всякого нильпотентного оператора $f\colon V \to V$ конечна.*

Док-во. Выберем базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в V и рассмотрим высоты всех базисных векторов $\nu(\mathbf{e}_1), \ldots, \nu(\mathbf{e}_n)$. Обозначим

$$m = \max\{\nu(\mathbf{e}_1), \ldots, \nu(\mathbf{e}_n)\}.$$

Выберем произвольный вектор $\mathbf{v} \in V$ и рассмотрим его разложение по базису: $\mathbf{v} = v^1 \cdot \mathbf{e}_1 + \ldots + v^n \cdot \mathbf{e}_n$. Применив оператор f^m к $\mathbf{v}$, выводим

$$f^m(\mathbf{v}) = \sum_{i=1}^{n} v^i \cdot f^m(\mathbf{e}_i) = \mathbf{0}. \tag{5.2}$$

Из формулы (5.2) видим, что высоты всех векторов пространства V ограничены сверху одним числом m. Это доказывает конечность высоты нильпотентного оператора: $\nu(f) = m < \infty$. $\square$

Теорема 5.2. *Если $f\colon V \to V$ — нильпотентный оператор в пространстве V и если U — инвариантное подпространство оператора f, то сужение f_U и фактороператор $f_{V/U}$ нильпотентны.*

Док-во. Всякий вектор $\mathbf{u}$ из подпространства $U \subset V$ является элементом V. Поэтому для него существует целое число $k > 0$, такое, что $f^k(\mathbf{u}) = \mathbf{0}$. Действие сужения f_U на векторы из U совпадает с действием исходного оператора f. Поэтому мы можем записать соотношение

$$(f_U)^k\, \mathbf{u} = f^k(\mathbf{u}) = \mathbf{0},$$

которое доказывает нильпотентность сужения f_U. Для доказательства нильпотентности фактороператора рассмотрим некоторый произвольный класс Q

из факторпространства V/U. Пусть $Q = \mathrm{Cl}_U(\mathbf{v})$, где $\mathbf{v}$ — некоторый фиксированный вектор из Q, и пусть $k = \nu(\mathbf{v})$ — высота вектора $\mathbf{v}$ относительно оператора f. Тогда, применив $(f_{V/U})^k$ к Q, мы получим

$$(f_{V/U})^k\, Q = \mathrm{Cl}_U(f^k(\mathbf{v})) = \mathbf{0}.$$

Отсюда видим, что фактороператор $f_{V/U}$ также нильпотентен. Теорема полностью доказана. $\square$

Теорема 5.3. *Нильпотентный оператор $f : V \to V$ не может иметь собственных чисел, отличных от нуля.*

Док-во. Пусть λ — собственное число оператора f и пусть $\mathbf{v} \neq 0$ — соответствующий собственный вектор. Тогда для него $f(\mathbf{v}) = \lambda \cdot \mathbf{v}$. С другой стороны из нильпотентности f вытекает существование числа $k > 0$, такого, что $f^k(\mathbf{v}) = \mathbf{0}$. Из этих двух условий получаем

$$f^k(\mathbf{v}) = \lambda^k \cdot \mathbf{v} = \mathbf{0}.$$

Но $\mathbf{v} \neq \mathbf{0}$, поэтому $\lambda^k = 0$. Это уравнение относительно λ имеет единственный корень $\lambda = 0$. $\square$

В конечномерном случае доказанная теорема может быть усилена. Сформулируем ее усиление так.

Теорема 5.4. *В конечномерном пространстве V размерности $\dim V = n$ всякий нильпотентный оператор f имеет ровно одно собственное число $\lambda = 0$ кратности n.*

Док-во. Доказательство теоремы проведем индукцией по размерности пространства $n = \dim V$. В случае $n = 1$ всякий вектор $\mathbf{v} \neq \mathbf{0}$ из V является собственным вектором оператора f. В силу предыдущей теоремы соответствующее собственное число не может отличаться от нуля. База индукции доказана.

Предположим, что теорема доказана для всех конечномерных пространств размерности меньше n. Рассмотрим пространство V размерности $n = \dim V$. Фиксируем некоторый вектор $\mathbf{v} \in V$ и обозначим через $k = \nu(\mathbf{v})$ его высоту относительно оператора f. Тогда $f^k(\mathbf{v}) = \mathbf{0}$, но $\mathbf{y} = f^{k-1}(\mathbf{v}) \neq \mathbf{0}$. Отсюда для ненулевого вектора $\mathbf{y}$ получаем

$$f(\mathbf{y}) = f(f^{k-1}(\mathbf{v})) = f^k(\mathbf{v}) = \mathbf{0} = 0 \cdot \mathbf{y}.$$

Значит, $\mathbf{y}$ есть собственный вектор оператора f, а $\lambda = 0$ есть собственное число этого оператора. Рассмотрим собственное подпространство $U = V_0$, отвечающее собственному числу 0. Обозначим через $m = \dim U \neq 0$ размерность подпространтства U. Сужение f_U равно нулю, поэтому для характеристического полинома оператора f_U имеем

$$\det(f_U - \lambda \cdot 1) = (-\lambda)^m.$$

Из этого равенства, применяя теорему 3.6, получаем характеристический полином исходного оператора f:

$$\det(f - \lambda \cdot 1) = (-\lambda)^m \det(f_{V/U} - \lambda \cdot 1). \tag{5.3}$$

Фактороператор $f_{V/U}$ действует в факторпространстве V/U, размерность которого равна $n - m$, что меньше n. В силу теоремы 5.2 фактороператор $f_{V/U}$ нильпотентен. Поэтому к нему можно применить предположение индукции. Для характеристического полинома, определенного по $f_{V/U}$, получаем

$$\det(f_{V/U} - \lambda \cdot 1) = (-\lambda)^{n-m}. \tag{5.4}$$

Сравнивая полученные выше два соотношения (5.3) и (5.4), находим характеристический полином нильпотентного оператора f:

$$\det(f - \lambda \cdot \mathbf{1}) = (-\lambda)^n.$$

Значит, единственное собственное число $\lambda = 0$ оператора f имеет кратность $n = \dim V$. Теорема доказана. $\square$

Пусть $f\colon V \to V$ — нильпотентный линейный оператор. Рассмотрим вектор $\mathbf{v} \in V$, обозначим через $k = \nu(\mathbf{v})$ высоту этого вектора. С вектором $\mathbf{v}$ связана цепочка из k векторов в пространстве V:

$$\mathbf{v}_1 = f^{k-1}(\mathbf{v}),\ v_2 = f^{k-2}(\mathbf{v}),\ \ldots,\ \mathbf{v}_k = f^0(\mathbf{v}) = \mathbf{v}. \tag{5.5}$$

Вектора цепочки (5.5) связаны соотношениями: $\mathbf{v}_i = f(v_{\mathbf{i}-1})$. Подействуем на все вектора цепочки (5.5) оператором f. При этом вектор $\mathbf{v}_1$ перейдет в ноль. Результаты применения f к остальным векторам образуют новую цепочку векторов пространства V, длина которой равна $k - 1$:

$$\mathbf{w}_1 = f^{k-1}(\mathbf{v}),\ \mathbf{w}_2 = f^{k-2}(\mathbf{v}),\ \ldots,\ \mathbf{w}_{k-1} = f(\mathbf{v}). \tag{5.6}$$

Сравнение цепочек (5.5) и (5.6) показывает, что вторая может быть получена из первой удалением вектора $\mathbf{v}_k$.

Вектор $\mathbf{v}_1$ называется *крайним вектором* или *собственным вектором* цепочки (5.5). Остальные векторы называются *присоединенными векторами.* Если крайние векторы двух цепочек не совпадают, то нет совпадения и в присоединенных векторах. Имеется даже более сильный результат, известный как теорема о «линейной независимости цепочек».

Теорема 5.5. *Из линейной независимости крайних векторов в некотором количестве цепочек вида* (5.5) *вытекает линейная независимость всех векторов в этих цепочках.*

Док-во. Рассмотрим s цепочек вида (5.5). Для нумерации векторов в этих цепочках используем два индекса $v_{i,j}$. Первый индекс — номер цепочки, а второй — номер вектора в цепочке. Обозначим через $k_1, \ldots, k_s$ длины рассматриваемых цепочек. Цепочки векторов можно без ограничения общности считать пронумерованными в порядке убывания их длин:

$$k_1 \geqslant k_2 \geqslant \ldots \geqslant k_s \geqslant 1. \tag{5.7}$$

Пусть $k = \max\{k_1, \ldots, k_s\}$. Доказательство теоремы проведем индукцией по числу k. При $k = 1$ длины всех цепочек равны 1, поэтому все векторы в цепочках крайние. Утверждение теоремы в этом случае справедливо тривиальным образом.

Пусть теорема доказана для цепочек, длины которых не превосходят $k - 1$. Выбрав набор из s цепочек, длина которых ограничена числом k, рассмотрим линейную комбинацию из всех векторов, входящих в эти цепочки:

$$\sum_{i=1}^{s}\sum_{j=1}^{k_s} \alpha_{i,j} \cdot \mathbf{v}_{i,j} = \mathbf{0}. \tag{5.8}$$

Из равенства нулю линейной комбинации (5.8) мы должны вывести тривиальность этой линейной комбинации. Применим оператор f к обеим частям равенства (5.8). Для этого используем следующие очевидные соотношения:

$$f(\mathbf{v}_{i,j}) = \begin{cases} \mathbf{0}, & \text{при } j = 1, \\ \mathbf{v}_{i,j-1}, & \text{при } j > 1. \end{cases}$$

Результат применения f к обеим частям равенства (5.8) при учете неравенств (5.7) можно записать в следующем виде:

$$\sum_{i=1}^{s}\sum_{j=1}^{k_s} \alpha_{i,j} \cdot f(\mathbf{v}_{i,j}) = \sum_{i=1}^{r}\sum_{j=2}^{k_r} \alpha_{i,j} \cdot \mathbf{v}_{i,j-1} = \mathbf{0}. \tag{5.9}$$

В типичной ситуации $r = s$. Однако, иногда вектора из некоторых цепочек могут совсем выпасть из суммирования. Это происходит, если часть цепочек имела длину 1. В таком случае $r < s$ и $k_{r+1} = \ldots = k_s = 1$. Все длины в (5.7) не могут быть равны 1, ибо $k > 1$.

Сделав сдвиг индекса $j + 1 \to j$ в последней сумме, преобразуем формулу (5.9) к следующему виду:

$$\sum_{i=1}^{r}\sum_{j=1}^{k_r-1} \alpha_{i,j+1} \cdot \mathbf{v}_{i,j} = \mathbf{0}. \tag{5.10}$$

Соотношение (5.10) представляют собой линейную комбинацию векторов из цепочек, длина которых сократилась на единицу по сравнению с исходной. Применив предположение индукции, мы получим линейную независимость векторов в формуле (5.10) и равенство нулю всех коэффициентов этой линейной комбинации.

Тривиальность линейной комбинации (5.10) означает зануление большей части слагаемых в сумме (5.8). В результате этого она преобразуется к виду

$$\sum_{i=1}^{s} \alpha_{i,1} \cdot \mathbf{v}_{i,1} = \mathbf{0}. \tag{5.11}$$

В линейной комбинации (5.11) остались лишь крайние векторы исходных цепочек. А они линейно независимы по условию теоремы. Следовательно, линейная комбинация (5.11) также тривиальна. Из тривиальности (5.10) и (5.11) вытекает тривиальность исходной линейной комбинации (5.8). Индукционный переход выполнен и вся теорема в целом доказана. □

Пусть $f\colon V \to V$ — нильпотентный оператор и пусть $\mathbf{v}$ — вектор высоты $k = \nu(\mathbf{v})$ в пространстве V. Рассмотрим цепочку (5.5), порожденную вектором $\mathbf{v}$. Обозначим через $U(v)$ линейную оболочку векторов цепочки (5.5):

$$U(\mathbf{v}) = \langle \mathbf{v}_1, \ldots, \mathbf{v}_k \rangle. \tag{5.12}$$

В силу доказанной выше теоремы 5.5 подпространство $U(\mathbf{v})$ конечномерно, причем $\dim U(\mathbf{v}) = k$ и вектора цепочки (5.5) составляют базис этого пространства. Из следующих очевидных соотношений

$$\begin{array}{l} f(\mathbf{v}_1) = \mathbf{0}, \\ f(\mathbf{v}_2) = \mathbf{v}_1, \\ \dots\dots\dots\dots\dots \\ f(\mathbf{v}_k) = \mathbf{v}_{k-1}, \end{array} \tag{5.13}$$

вытекающих непосредственно из определения цепочки (5.5), легко видеть, что подпространство (5.12) инвариантно относительно оператора f. Это обстоятельство позволяет рассмотреть сужение $f_{U(\mathbf{v})}$ оператора f на подпространство (5.12). Соотношения (5.13) определяют вид матрицы такого сужения в базисе, образованном векторами цепочки (5.5):

$$J_k(0) = \left\| \begin{matrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{matrix} \right\|. \tag{5.14}$$

Матрица вида (5.14) называется *жордановым блоком* или же *жордановой клеткой* нильпотентного оператора. На главной диагонали такой матрицы расположены нули, следующая параллельная диагональ выше заполнена единицами, а все оставшееся пространство матрицы состоит из нулей. Матрица (5.14) имеет размер $k \times k$, при $k = 1$ она вырождается в чисто нулевую матрицу с единственным элементом $J_1(0) = \|\, 0 \,\|$.

Пусть $f\colon V \to V$ — нильпотентный оператор. Продолжим изучение цепочек вида (5.5). Для этого рассмотрим следующие подпространства в V:

$$U_k = \operatorname{Ker} f \cap \operatorname{Im} f^{k-1}. \tag{5.15}$$

Если $\mathbf{u} \in U_k$, то $\mathbf{u} \in \operatorname{Im} f^{k-1}$ и $\mathbf{u} = f^{k-1}(\mathbf{v})$ для некоторого вектора $\mathbf{v}$. Это значит, что $\mathbf{u}$ содержится в цепочке вида (5.5). Из условия $\mathbf{u} \in \operatorname{Ker} f$ вытекает, что $f(\mathbf{u}) = f^k(\mathbf{v}) = \mathbf{0}$. Следовательно, высота вектора $\mathbf{v}$ равна k и вектор $\mathbf{u}$ есть крайний вектор в цепочке (5.5), порожденной вектором $\mathbf{v}$. Для подпространств (5.15) имеют место включения

$$V_0 = U_1 \supseteq U_2 \supseteq \ldots \supseteq U_k \supseteq \ldots, \tag{5.16}$$

где $V_0 = \operatorname{Ker} f$ есть собственное подпространство, отвечающее единственному собственному числу нильпотентного оператора $\lambda = 0$. Включения (5.16) вытекают из того, что всякую цепочку (5.5) длины k с крайним вектором $\mathbf{u} = f^{k-1}(\mathbf{v})$ можно превратить в цепочку длины $k-1$, отбросив вектор $\mathbf{v}_k = \mathbf{v}$ (см. переход от (5.5) к (5.6)). Тогда для вектора $\mathbf{v}' = f(\mathbf{v})$ имеем $\mathbf{u} = f^{k-2}(\mathbf{v}')$, что дает $U_k \subset U_{k-1}$ при $k > 1$.

В конечномерном пространстве V высота любого вектора $\mathbf{v} \in V$ ограничена сверху высотой оператора f:

$$\nu(\mathbf{v}) \leq \nu(f) = m < \infty$$

(см. теорему 5.1). Поэтому $U_{m+1} = \{\mathbf{0}\}$ и последовательность включений (5.16) заканчивается на m-ом шаге:

$$V_0 = U_1 \supseteq U_2 \supseteq \ldots \supseteq U_m \supseteq \{\mathbf{0}\}. \tag{5.17}$$

Последовательности вложенных друг в друга подпространств вида (5.16) или (5.17) иногда называют *флагами*, а сами подпространства *флаговыми подпространствами.*

Теорема 5.6. *Для всякого нильпотентного оператора f в конечномерном пространстве V существует базис в пространстве V, образованный цепочками векторов вида* (5.5). *Такой базис называется каноническим или жордановым нормальным базисом нильпотентного оператора.*

Док-во. Доказательство теоремы опирается на факт конечности флага (5.17). Выберем базис в самом маленьком из подпространств U_m. Дополним этот базис до базиса в U_{m-1}, затем до базиса в U_{m-2} и далее по цепочке (5.17). В результате этого мы получим базис $\mathbf{e}_1, \ldots, \mathbf{e}_s$ в пространстве $V_0 = \operatorname{Ker} f$. Каждый из векторов такого базиса является крайним вектором в некоторой цепочке вида (5.5). Для базисных векторов из U_m длина такой цепочки равна m, для дополнительных векторов из U_{m-1}, она равна $m - 1$ и далее по убывающей.

Объединим все вектора из всех таких цепочек и будем нумеровать их двумя индексами $\mathbf{e}_{i,j}$, где i — номер цепочки, а j — номер вектора в цепочке. При этом для исходных векторов $\mathbf{e}_1, \ldots, \mathbf{e}_s$ получим

$$\mathbf{e}_1 = \mathbf{e}_{1,1},\ \ldots,\ \mathbf{e}_s = \mathbf{e}_{s,1}.$$

Докажем, что полученный набор векторов является базисом в V. Линейная независимость такого набора векторов вытекает из теоремы 5.5. Остается доказать, что произвольный вектор $\mathbf{v} \in V$ может быть по ним разложен. Доказательство произведем индукцией по высоте вектора $\mathbf{v}$.

Если $k = \nu(\mathbf{v}) = 1$, то $\mathbf{v} \in \operatorname{Ker} f = V_0$, тогда он раскладывается по базису $\mathbf{e}_1, \ldots, \mathbf{e}_s$ в V_0. База индукции доказана.

Предположим, что все вектора высоты, меньшей чем k, могут быть разложены в линейную комбинацию векторов $\mathbf{e}_{i,j}$. Рассмотрим вектор $\mathbf{v}$ высоты k и положим $\mathbf{u} = f^{k-1}(\mathbf{v})$. Тогда $f(\mathbf{u}) = \mathbf{0}$ и $\mathbf{u}$ есть крайний вектор в цепочке длины k, заданной вектором $\mathbf{v}$. Значит, $\mathbf{u}$ принадлежит подпространству U_k

вида (5.15) и он может быть разложен по построенному выше базису:

$$\mathbf{u} = \sum_{i=1}^{r} \alpha_i \cdot \mathbf{e}_i. \tag{5.18}$$

В разложении (5).18 участвуют не все вектора $\mathbf{e}_1, \ldots, \mathbf{e}_s$, а лишь те из них, которые принадлежат U_k и, следовательно, являются крайними векторами в цепочках длины не меньше k. Отсюда

$$\mathbf{e}_1 = f^{k-1}(\mathbf{e}_{1,k}), \ \ldots, \ \mathbf{e}_r = f^{k-1}(\mathbf{e}_{r,k}).$$

Это обстоятельство позволяет переписать (5.18) в виде

$$f^{k-1}(\mathbf{v}) = \sum_{i=1}^{r} \alpha_i \cdot f^{k-1}(\mathbf{e}_{i,k}). \tag{5.19}$$

Пользуясь коэффициентами разложения (5.19), из исхоного вектора $\mathbf{v}$ высоты k построим новый вектор $\mathbf{v}'$:

$$\mathbf{v}' = \mathbf{v} - \sum_{i=1}^{r} \alpha_i \cdot \mathbf{e}_{i,k}. \tag{5.20}$$

Применяя к вектору $\mathbf{v}'$ оператор f^{k-1} и учитывая (5.19), получаем

$$f^{k-1}(\mathbf{v}') = f^{k-1}(\mathbf{v}) - \sum_{i=1}^{r} \alpha_i \cdot f^{k-1}(\mathbf{e}_{i,k}) = \mathbf{0}.$$

Значит, высота вектора $\mathbf{v}'$ меньше k и он может быть разложен по векторам $\mathbf{e}_{i,j}$. Но из разложимости $\mathbf{v}'$ в силу (5.20) вытекает разложимость исходного вектора $\mathbf{v}$. Индукционный переход и вся теорема в целом доказаны. $\square$

В базисе из цепочек, существование которого вытекает из доказанной теоремы 5.6, матрица нильпотентного оператора f имеет вид

$$F = \begin{Vmatrix} J_{k_1}(0) & & & \\ & J_{k_2}(0) & & \\ & & \ddots & \\ & & & J_{k_s}(0) \end{Vmatrix}. \tag{5.21}$$

Матрица (5.21) блочно-диагональна, диагональные блоки ее имеют вид жордановых клеток (5.14), а все остальное пространство заполнено нулями. Это обстоятельство нетрудно понять. Действительно, каждая цепочка с крайним вектором $\mathbf{e}_i$ порождает инвариантное подпространство $U(\mathbf{v})$ вида (5.12), где $\mathbf{v} = \mathbf{e}_{i,k_i}$. В силу теоремы 5.6 все пространство V есть прямая сумма таких инвариантных подпространств:

$$V = U(\mathbf{e}_{1,k_1}) \oplus \ldots \oplus U(\mathbf{e}_{s,k_s}).$$

Матрица (5.21) называется *жордановой нормальной формой* матрицы нильпотентного оператора, а теорема 5.6 известна как *теорема о приведении* матрицы нильпотентного оператора к жордановой нормальной форме. Если базис из цепочек построен в строгом соответствии с доказательством теоремы 5.6, то размеры жордановых клеток идут в порядке убывания:

$$k_1 \geqslant k_2 \geqslant \ldots \geqslant k_s.$$

Однако, перестановка векторов $\mathbf{e}_1, \ldots, \mathbf{e}_s$ может нарушить этот порядок, что на практике чаще всего и происходит.

Теорема 5.7. *Высота нильпотентного оператора f в конечномерном пространстве V не превосходит размерности этого пространства $n = \dim V$, то есть $f^n = 0$.*

Док-во. Выше при доказательстве теоремы 5.1 мы заметили, что высота $\nu(f)$ нильпотентного оператора f совпадает с наибольшей из высот базисных векторов. Пользуясь результатом теоремы 5.6, выберем базис из цепочек. Высота вектора из цепочки вида (5.5) не превосходит длины этой цепочки, поэтому высота базисных векторов в базисе из цепочек не может быть больше количества базисных векторов. Это дает $\nu(f) \leqslant n = \dim V$. Но высота любого вектора $\mathbf{v}$ не превосходит высоты оператора f, поэтому $f^n(\mathbf{v}) = \mathbf{0}$ для всех $\mathbf{v} \in V$. Отсюда $f^n = 0$. Теорема доказана. $\square$

§ 6. Корневые подпространства. Теорема о сумме корневых подпространств.

Определение 6.1. Корневым подпространством оператора $f\colon V \to V$, соответствующим его собственному числу λ, называется множество

$$V(\lambda) = \{\mathbf{v} \in V\colon\ \exists k\,((k \in \mathbb{N})\ \ \&\ \ ((f - \lambda \cdot \mathbf{1})^k\,\mathbf{v} = \mathbf{0}))\},$$

состоящее из векторов, которые зануляются некоторой целой положительной степенью оператора $h_\lambda = f - \lambda \cdot 1$.

Для каждого целого положительного числа k определим подпространство $V(k, \lambda) = \operatorname{Ker}(h_\lambda)^k$. При $k = 1$ подпространство $V(1, \lambda)$ совпадает с собственным подпространством V_λ. Заметим, что из $(h_\lambda)^k\,\mathbf{v} = \mathbf{0}$ вытекает $(h_\lambda)^{k+1}\,\mathbf{v} = \mathbf{0}$. Поэтому имеет место цепочка включений

$$V(1, \lambda) \subseteq V(2, \lambda) \subseteq \ldots \subseteq V(k, \lambda) \subseteq \ldots \tag{6.1}$$

Нетрудно видеть, что все подпространства, составляющие цепочку включений (6.1), содержатся в корневом пространстве $V(\lambda)$. Более того, $V(\lambda)$ есть объединение подпространств (6.1):

$$V(\lambda) = \bigcup_{k=1}^{\infty} V(k, \lambda) = \sum_{k=1}^{\infty} V(k, \lambda). \tag{6.2}$$

В данном случае сумма подпространств $V(k, \lambda)$ совпадает с их объединением. Действительно, пусть вектор $\mathbf{v}$ принадлежит сумме этих подпространств:

$$\mathbf{v} = \mathbf{v}_{k_1} + \ldots + \mathbf{v}_{k_s}, \ \text{ где }\ \mathbf{v}_{k_s} \in V(k_s, \lambda). \tag{6.3}$$

Положим $k = \max\{k_1, \ldots, k_s\}$, тогда из цепочки включений (6.1) видим, что $\mathbf{v}_{k_i} \in V(k, \lambda)$. Поэтому вектор $\mathbf{v}$ из (6.3) принадлежит $V(k, \lambda)$, а значит, и объединению всех подпространств вида $V(k, \lambda)$.

Доказательство совпадения суммы и объединения в (6.2) основано только на включениях (6.1). Поэтому мы фактически доказали более общую теорему.

Теорема 6.1. *Сумма произвольной цепочки вложенных друг в друга подпространств совпадает с их объединением.*

Теорема 6.1 показывает, что множество $V(\lambda)$ из определения 6.1 действительно является подпространством в пространстве V. Корневое подпространство $V(\lambda)$ отлично от нуля, ибо оно содержит в себе соответствующее собственное подпространство V_λ.

Теорема 6.2. *Корневое подпространство $V(\lambda)$ оператора f инвариантно относительно самого оператора f и относительно всех операторов $P(f)$ из его полиномиальной оболочки.*

Док-во. Пусть $\mathbf{v} \in V(\lambda)$. Тогда существует целое положительное число k, для которого $(h_\lambda)^k \, \mathbf{v} = \mathbf{0}$. Рассмотрим вектор $\mathbf{w} = f(\mathbf{v})$. Для него имеем

$$(h_\lambda)^k \, \mathbf{w} = (h_\lambda)^k \, f \, \mathbf{v} = f \, (h_\lambda)^k \, \mathbf{v} = \mathbf{0}.$$

Здесь мы использовали перестановочность операторов h_λ и f, которая вытекает из принадлежности h_λ полиномиальной оболочке оператора f. Из полученного соотношения видим, что $\mathbf{w} = f(\mathbf{v}) \in V(\lambda)$. Инвариантность подпространства $V(\lambda)$ относительно f доказана. Его инвариантность относительно $P(f)$ теперь вытекает из теоремы 4.5. $\square$

Теорема 6.3. *Пусть λ и μ — собственные числа линейного оператора $f\colon V \to V$. Тогда сужение оператора $h_\lambda = f - \lambda \cdot \mathbf{1}$ на корневое подпространство $V(\mu)$ является*

(1) *биективным при $\mu \neq \lambda$;*
(2) *нильпотентным при $\mu = \lambda$.*

Док-во. Докажем первое утверждение теоремы. Мы уже знаем что подпространство $V(\mu)$ инвариантно относительно h_λ. Для удобства записи обозначим через $h_{\lambda,\mu}$ сужение оператора h_λ на подпространство $V(\mu)$. Это оператор, действующий из $V(\mu)$ в $V(\mu)$. Найдем его ядро

$$\operatorname{Ker} h_{\lambda,\mu} = \{\mathbf{v} \in V(\mu) \colon \ \ h_\lambda(\mathbf{v}) = \mathbf{0}\} = \operatorname{Ker} h_\lambda \cap V(\mu).$$

Но ядро оператора h_λ есть собственное подпространство V_λ. Поэтому мы имеем $\operatorname{Ker} h_{\lambda,\mu} = V_\lambda \cap V(\mu)$.

Пусть $\mathbf{v}$ — некоторый произвольный вектор из $\operatorname{Ker} h_{\lambda,\mu}$. В силу сказанного выше он принадлежит V_λ, поэтому

$$f(\mathbf{v}) = \lambda \cdot \mathbf{v}. \tag{6.4}$$

Условие же $\mathbf{v} \in V(\mu)$ означает существование некоторого целого положительного числа k, для которого выполняется равенство

$$(h_\mu)^k\,\mathbf{v} = (f - \mu\cdot 1)^k\,\mathbf{v} = \mathbf{0}. \tag{6.5}$$

Из (6.4) видим, что $h_\mu(\mathbf{v}) = f(\mathbf{v}) - \mu\cdot\mathbf{v} = (\lambda - \mu)\cdot\mathbf{v}$. Соединив это с формулой (6.5), получаем следующее соотношение:

$$(h_\mu)^k\,\mathbf{v} = (\lambda - \mu)^k\cdot\mathbf{v} = \mathbf{0}.$$

Отсюда при $\lambda \neq \mu$ немедленно получаем $\mathbf{v} = \mathbf{0}$, что означает зануление $\operatorname{Ker} h_{\lambda,\mu} = \{\mathbf{0}\}$. Значит, оператор

$$h_{\lambda,\mu}\colon\, V(\mu) \to V(\mu)$$

инъективен. Его сюръективность, а значит, и биективность вытекает из инъективности в силу теоремы 1.3.

Для доказательства второго утверждения теоремы положим $\mu = \lambda$ и рассмотрим оператор $h_{\lambda,\lambda}$, полученный сужением h_λ на $V(\lambda)$. Действие $h_{\lambda,\lambda}$ на вектора из $V(\lambda)$ совпадает с действием исходного оператора h_λ на эти вектора. Поэтому из самого определения корневого подпространства $V(\lambda)$ видим, что для любого вектора $v \in V(\lambda)$ существует целое положительное число k, такое, что выполняется следующее равенство:

$$(h_{\lambda,\lambda})^k\,\mathbf{v} = (f - \lambda\cdot 1)^k\,\mathbf{v} = \mathbf{0}.$$

А эторавенство, как раз, и означает нильпотентность оператора $h_{\lambda,\lambda}$. Теорема полностью доказана. $\square$

Теорема 6.4. *Пусть $\lambda_1, \ldots, \lambda_s$ — некоторый набор попарно различных собственных чисел оператора $f\colon\, V \to V$. Тогда сумма корневых подпространств $V(\lambda_1) \oplus \ldots \oplus V(\lambda_s)$ прямая.*

Док-во. Доказательство этой теоремы сходно с доказательством теоремы 4.6. Обозначим через W сумму указанных в теореме подпространств:

$$W = V(\lambda_1) + \ldots + V(\lambda_s). \tag{6.6}$$

Для доказательства того, что сумма (6.6) прямая, надо для произвольного $\mathbf{w} \in W$ установить однозначность разложения

$$\mathbf{w} = \mathbf{v}_1 + \ldots + \mathbf{v}_s, \text{ где } \mathbf{v}_i \in V(\lambda_i). \tag{6.7}$$

Допустим, что существует некоторое другое разложение вектора $\mathbf{w}$:

$$\mathbf{w} = \tilde{\mathbf{v}}_1 + \ldots + \tilde{\mathbf{v}}_s, \text{ где } \tilde{\mathbf{v}}_i \in V(\lambda_i). \tag{6.8}$$

Вычтем второе разложение из первого разложения и для краткости обозначим $\mathbf{w}_i = (\mathbf{v}_i - \tilde{\mathbf{v}}_i) \in V(\lambda_i)$. В результате этого получим:

$$\mathbf{w}_1 + \ldots + \mathbf{w}_s = \mathbf{0}. \tag{6.9}$$

Положим $h_r = f - \lambda_r \cdot 1$. По определению корневого подпространства $V(\lambda_r)$ для каждого вектора $\mathbf{w}_r$ в разложении (6.9) существует целое положительное число k_r, такое, что $(h_r)^{k_r}\,\mathbf{w}_r = \mathbf{0}$. Определим операторы

$$f_i = \prod_{r \neq i}^{s} (h_r)^{k_r}. \tag{6.10}$$

Из перестановочности операторов $h_1, \ldots, h_s$, входящих в полиномиальную оболочку оператора f, и из соотношения $(h_r)^{k_r}\,\mathbf{w}_r = \mathbf{0}$ получаем:

$$f_i(\mathbf{w}_j) = 0 \quad \text{для всех} \quad j \neq i.$$

Применим оператор f_i вида (6.10) к обеим частям равенства (6.9). При этом из суммы в левой части этого равенства останется лишь одно i-ое слагаемое, что дает $f_i(\mathbf{w}_i) = \mathbf{0}$. Запишем полученное выражение в развернутом виде:

$$\left(\prod_{r \neq i}^{s} (h_r)^{k_r} \right) \mathbf{w}_i = \mathbf{0}. \tag{6.11}$$

Вектор $\mathbf{w}_i$ принадлежит корневому подпространству $V(\lambda_i)$, которое инвариантно относительно действия всех операторов вида h_r, представленных в соотношении (6.11). Поэтому операторы h_r в (6.11) можно заменить их сужениями $h_{r,i}$ на подпространство $V(\lambda_i)$:

$$\left(\prod_{r \neq i}^{s} (h_{r,i})^{k_r} \right) \mathbf{w}_i = \mathbf{0}. \tag{6.12}$$

Согласно теореме 6.3 сужения $h_{r,i}$ биективны при $r \neq i$. Но именно такие операторы $h_{r,i}$ фигурируют в формуле (6.12). Произведение (композиция) биективных операторов биективно, а действие биективного оператора на ненулевой вектор отлично от нуля. Это рассуждение показывает, что $\mathbf{w}_i = \mathbf{0}$. Значит, $\mathbf{v}_i = \tilde{\mathbf{v}}_i$ и разложения (6.7) и (6.8) совпадают. Однозначность разложения (6.7) и вся теорема в целом доказаны. □

Теорема 6.5. *Пусть f — линейный оператор в конечномерном пространстве V над полем $\mathbb{K}$ и пусть его характеристический полином раскладывается на линейные множители в поле $\mathbb{K}$. Тогда сумма всех корневых подпространств $V(\lambda_1) \oplus \ldots \oplus V(\lambda_s)$ этого оператора равна V.*

Док-во. Пусть $\lambda_1, \ldots, \lambda_s$ есть полный набор попарно различных собственных чисел оператора f. Из того, что характеристический полином этого оператора раскладывается на линейные множители в поле $\mathbb{K}$ заключаем, что

$$\det(f - \lambda \cdot \mathbf{1}) = \prod_{i=1}^{s} (\lambda_i - \lambda)^{n_i},$$

где $n_1, \ldots, n_s$ — кратности собственных чисел оператора f. Обозначим через W сумму всех корневых подпространств оператора f: $W = V(\lambda_1) \oplus \ldots \oplus V(\lambda_s)$.

Корневые подпространства $V(\lambda_1), \ldots, V(\lambda_s)$ отличны от нуля. Поэтому для подпространства W имеем $W \neq \{\mathbf{0}\}$.

Допустим, что $W \neq V$. Подпространство W инвариантно относительно f как сумма инвариантных подпространств $V(\lambda_i)$ (см. теорему 3.2). В силу теоремы 4.5 оно инвариантно и относительно оператора $h_\lambda = f - \lambda \cdot 1$. Применим к оператору h_λ теорему 3.5. Это дает

$$\det(f - \lambda \cdot 1) = \det(f_W - \lambda \cdot 1)\,\det(f_{V/W} - \lambda \cdot 1). \tag{6.13}$$

Здесь мы учли, что $1_W = 1$ и $1_{V/W} = 1$, а также применили теорему 3.4. Вывод: характеристический полином оператора f есть произведение характеристических полиномов сужения f_W и фактороператора $f_{V/W}$. Левая часть (6.13) раскладывается на линейные множители в поле $\mathbb{K}$, поэтому каждый из многочленов в правой части (6.13) также раскладывается в произведение линейных множителей. Пусть λ_q — одно из собственных чисел фактороператора $f_{V/W}$ и пусть $Q \in V/W$ — соответствующий этому собственному числу собственный вектор фактороператора $f_{V/W}$. В силу (6.13) число λ_q содержится в списке $\lambda_1, \ldots, \lambda_s$ собственных чисел оператора f. Из сделанного нами допущения $W \neq V$ заключаем, что факторпространство V/W нетривиально: $V/W \neq \{\mathbf{0}\}$, и класс Q отличен от нуля. Пусть $\mathbf{v} \in Q$ — некоторый представитель этого класса. Из $Q \neq \mathbf{0}$ вытекает, что $\mathbf{v} \notin W$. Класс Q является собственным вектором для фактороператора $f_{V/W}$, поэтому

$$(f_{V/W} - \lambda_q \cdot 1)\,Q = \mathrm{Cl}_W((f - \lambda_q \cdot 1)\,\mathbf{v}) = \mathbf{0}. \tag{6.14}$$

Для удобства записи обозначим $h_r = f - \lambda_r \cdot 1$ для всех $r = 1, \ldots, s$. Такое обозначение мы уже использовали ранее при доказательстве предыдущей теоремы. Соотношение (6.14) означает, что

$$(f - \lambda_q \cdot 1)\,\mathbf{v} = h_q(\mathbf{v}) = \mathbf{w} \in W. \tag{6.15}$$

Из разложения $W = V(\lambda_1) \oplus \ldots \oplus V(\lambda_s)$ для вектора $\mathbf{w}$ в формуле (6.15) получаем следующее разложение

$$h_q(\mathbf{v}) = \mathbf{w} = \mathbf{v}_1 + \ldots + \mathbf{v}_s, \text{ где } \mathbf{v}_i \in V(\lambda_i). \tag{6.16}$$

Рассмотрим сужение оператора h_q на корневое подпространство $V(\lambda_i)$, такое сужение мы обозначали через $h_{q,i}$ (см. доказательство теоремы 6.4). В силу теоремы 6.3 операторы $h_{q,i}\colon V(\lambda_i) \to V(\lambda_i)$ биективны при всех $i \neq q$. Поэтому для всех $\mathbf{v}_1, \ldots, \mathbf{v}_s$ в (6.16), кроме $\mathbf{v}_q$ можно подобрать такие $\tilde{\mathbf{v}}_i \in V(\lambda_i)$, что $\mathbf{v}_i = h_{q,i}(\tilde{\mathbf{v}}_i)$. Подставив это в (6.16), получим

$$\mathbf{w} = h_q(\mathbf{v}) = \mathbf{v}_q + \sum_{i \neq q}^{s} h_q(\tilde{\mathbf{v}}_i). \tag{6.17}$$

Основываясь на формуле (6.17), определим новый вектор $\tilde{\mathbf{v}}_q$:

$$\tilde{\mathbf{v}}_q = \mathbf{v} - \sum_{i \neq q}^{s} \tilde{\mathbf{v}}_i. \tag{6.18}$$

Для этого вектора из (6.17) имеем $h_q(\tilde{\mathbf{v}}_q) = \mathbf{v}_q \in V(\lambda_q)$. Из определения корневого подпространства $V(\lambda_q)$ вытекает существование целого положительного k, такого, что $(h_q)^k\,\mathbf{v}_q = \mathbf{0}$. Отсюда $(h_q)^{k+1}\,\tilde{\mathbf{v}}_q = \mathbf{0}$, что дает $\tilde{\mathbf{v}}_q \in V(\lambda_q)$. Теперь, возвращаясь к формуле (6.18), из нее получаем

$$\mathbf{v} = \sum_{i=1}^{s} \tilde{\mathbf{v}}_i, \text{ где } \mathbf{v}_i \in V(\lambda_i). \tag{6.19}$$

Из формулы (6.19) и разложения $W = V(\lambda_1) \oplus \ldots \oplus V(\lambda_s)$ вытекает $\mathbf{v} \in W$, что противоречит первоначальному условию $\mathbf{v} \notin W$, которое вытекает из допущения $W \neq V$. Полученное противоречие доказывает требуемое совпадение $W = V$. Теорема доказана. $\square$

§ 7. Жорданов нормальный базис линейного оператора. Теорема Гамильтона-Кэли.

Пусть $f\colon V \to V$ — линейный оператор в конечномерном пространстве V. Пусть V раскладывается в сумму корневых подпространств оператора f:

$$V = V(\lambda_1) \oplus \ldots \oplus V(\lambda_s). \tag{7.1}$$

Положим $h_i = f - \lambda_i \cdot 1$ и обозначим через $h_{i,j}$ сужение h_i на $V(\lambda_j)$. Согласно теореме 6.3 сужение $h_{i,i}$ является нильпотентным оператором в соответствующем корневом подпространстве $V(\lambda_i)$. Поэтому для него в $V(\lambda_i)$ можно подобрать жорданов нормальный базис (см. теорему 5.6). Оператор $h_{i,i}$ имеет в этом базисе матрицу вида (5.21), составленную из блоков вида (5.14).

Определение 7.1. *Жордановым нормальным базисом* оператора $f\colon V \to V$ называется базис, составленный из жордановых нормальных базисов нильпотентных операторов $h_{i,i}$ в корневых подпространствах $V(\lambda_i)$.

Из этого определения видим, что оператор f в конечномерном пространстве V имеет жорданов нормальный базис только если имеет место разложение (7.1). Теорема 6.5 дает достаточное условие существования жорданова нормального базиса.

Пусть оператор f в конечномерном пространстве V имеет жорданов нормальный базис. Подпространства $V(\lambda_i)$ в (7.1) инвариантны относительно f. Обозначим через f_i сужение f на $V(\lambda_i)$. Матрица оператора f в жордановом нормальном базисе блочно-диагональна

$$F = \begin{Vmatrix} F_1 & & & \\ & F_2 & & \\ & & \ddots & \\ & & & F_s \end{Vmatrix}, \tag{7.2}$$

а диагональные блоки F_i определяются операторами f_i. Операторы f_i и $h_{i,i}$

связаны соотношением $f_i = h_{i,i} + \lambda_i \cdot 1$, поэтому

$$F_i = \begin{Vmatrix} J_{k_1}(\lambda_i) & & & \\ & J_{k_2}(\lambda_i) & & \\ & & \ddots & \\ & & & J_{k_r}(\lambda_i) \end{Vmatrix}. \tag{7.3}$$

Число диагональных блоков в блочно-диагональной матрице (7.3) определяется числом цепочек в жордановом нормальном базисе нильпотентного оператора $h_{i,i}$, а сами они имеют вид:

$$J_k(\lambda) = \begin{Vmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{Vmatrix}. \tag{7.4}$$

Матрица вида (7.4) называется *жордановым блоком* или же *жордановой клеткой* с диагональю λ. Матрица (7.4) имеет размер $k \times k$, при $k = 1$ она вырождается в матрицу с единственным элементом $J_0(\lambda) = \|\,\lambda\,\|$.

Вид матрицы оператора f в жордановом нормальном базисе, определяемый соотношениями (7.2), (7.3) и (7.4), называется *жордановой нормальной формой* матрицы линейного оператора. Задача построения нормального базиса и нахождения матрицы F известна как задача *приведения матрицы линейного оператора к жордановой нормальной форме.*

Возможность приведения матрицы линейного оператора к жордановой нормальной форме имеет несколько важных следствий. Заметим, что матрица вида (7.4) является верхне-треугольной. Следовательно, верхне-треугольными являются и матрицы вида (7.3), а значит, и вся матрица оператора f в жордановом нормальном базисе. На диагонали верхне-треугольной матрицы (7.2) стоят собственные числа оператора f, причем, число λ_i встречается n_i-раз, где $n_i = \dim V(\lambda_i)$. Из курса алгебры известно, что детерминант верхне-треугольной матрицы равен произведению всех ее диагональных элементов. Для характеристического полинома оператора f, имеющего жорданов нормальный базис, это обстоятельство дает:

$$\det(f - \lambda \cdot \mathbf{1}) = \prod_{i=1}^{s} (\lambda_i - \lambda)^{n_i}. \tag{7.5}$$

Теорема 7.1. *Матрица линейного оператора в конечномерном пространстве допускает приведение к жордановой нормальной форме тогда и только тогда, когда его характеристический полином раскладывается в произведение линейных множителей в поле $\mathbb{K}$.*

Док-во. Необходимость утверждения теоремы вытекает из формулы (7.5), а достаточность этого утверждения обеспечивается теоремами 5.6 и 6.5. $\square$

В случае поля комплексных чисел любой многочлен раскладывается в произведение линейных множителей. Поэтому матрица любого линейного оператора в комплексном векторном пространстве допускает приведение к жордановой нормальной форме.

Теорема 7.2. *Кратность собственного числа λ оператора f в конечномерном пространстве V равна размерности корневого подпространства $V(\lambda)$.*

Для оператора f, характеристический полином которого допускает полное разложение на линейные множители, утверждение теоремы немедленно вытекает из формулы (7.5). Этот факт имеет место и в случае, когда характеристический полином оператора f имеет один или несколько корней в поле $\mathbb{K}$, но не допускает полного разложения на линейные множители. Такой более общий случай может быть сведен к рассмотренному выше на основе механизма расширения поля $\mathbb{K}$. Этот прием мы не рассматриваем. Следующая теорема известна как теорема Гамильтона-Кэли.

Теорема 7.3. *Пусть $P(\lambda)$ — характеристический полином линейного оператора f в конечномерном пространстве V. Тогда $P(f) = 0$.*

Док-во. Мы докажем теорему только для случая, когда характеристический полином $P(\lambda)$ допускает полное разложение на линейные множители

$$P(\lambda) = \prod_{i=1}^{s} (\lambda_i - \lambda)^{n_i}. \tag{7.6}$$

Положим $h_i = f - \lambda_i \cdot 1$ и обозначим через $h_{i,j}$ сужение h_i на корневое подпрпостранство $V(\lambda_j)$. Тогда из формулы (7.6) выводим

$$P(f) = \prod_{i=1}^{s} (h_i)^{n_i}.$$

Рассмотрим действие $P(f)$ на произвольный вектор $\mathbf{v} \in V$. Теоремы 6.5 позволяет разложить $\mathbf{v}$ в сумму $\mathbf{v} = \mathbf{v}_1 + \ldots + \mathbf{v}_s$, где $\mathbf{v}_i \in V(\lambda_i)$. Поэтому

$$P(f)\,\mathbf{v} = P(f)\,\mathbf{v}_1 + \ldots + P(f)\,\mathbf{v}_s. \tag{7.7}$$

Корневое подпространство $V(\lambda_j)$ инвариантно относительно операторов h_i. Отсюда для j-го слагаемого в разложении (7.7) имеем

$$P(f)\,\mathbf{v}_j = \prod_{i=1}^{s} (h_i)^{n_i}\,\mathbf{v}_j = \prod_{i=1}^{s} (h_{i,j})^{n_i}\,\mathbf{v}_j.$$

Пользуясь перестановочностью операторов h_i и их сужений $h_{i,j}$, находим

$$P(f)\,\mathbf{v}_j = \prod_{i \neq j}^{s} (h_{i,j})^{n_i}\,(h_{j,j})^{n_j}\,\mathbf{v}_j. \tag{7.8}$$

Но оператор $h_{j,j}$ нильпотентен в пространстве $V(\lambda_j)$, причем $n_j = \dim V(\lambda_j)$. Пименив теорему 5.7, находим $(h_{j,j})^{n_j}\,\mathbf{v}_j = \mathbf{0}$. Теперь из (7.7) и (7.8) получаем $P(f)\,\mathbf{v} = \mathbf{0}$ для произвольного вектора $\mathbf{v} \in V$. Это завершает доказательство теоремы для операторов, характеристические полиномы которых раскладывается на линейные множители (7.6). Общий случай сводится к этому при помощи механизма расширения поля $\mathbb{K}$, который мы не рассматриваем. $\square$

ГЛАВА III

СОПРЯЖЕННОЕ ПРОСТРАНСТВО.

§ 1. Линейные функционалы. Векторы и ковекторы. Сопряженное пространство.

Определение 1.1. Пусть V — линейное векторное пространство над полем $\mathbb{K}$. Числовая функция $y = f(\mathbf{v})$ с аргументами из V и значениями из поля $\mathbb{K}$ называется *линейным функционалом* на пространстве V, если:

(1) $f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2)$ для любых $\mathbf{v}_1, \mathbf{v}_2 \in V$;
(2) $f(\alpha \cdot \mathbf{v}) = \alpha\, f(\mathbf{v})$ для любого $\mathbf{v} \in V$ и любого числа α из поля $\mathbb{K}$.

Определение линейного функционала очень похоже на определение линейного отображения (см. определение 8.1 из первой главы). Сравнивая эти два определения, видим, что всякий линейный функционал f есть линейное отображение $f\colon V \to \mathbb{K}$ и, наоборот, всякое такое линейное отображение есть линейный функционал. При этом поле $\mathbb{K}$ рассматривается как одномерное линейное пространство над $\mathbb{K}$.

Как линейные отображения из пространства V в поле $\mathbb{K}$, линейные функционалы составляют пространство $\operatorname{Hom}(V, \mathbb{K})$, которое принято называть *дуальным пространством* или *сопряженным пространством* для пространства V. Сопряженное пространство $\operatorname{Hom}(V, \mathbb{K})$ обозначают через V^*. Пространство гомоморфизмов $\operatorname{Hom}(V, W)$ обычно определяется заданием двух пространств V и W. Однако, сопряженное пространство $V^* = \operatorname{Hom}(V, \mathbb{K})$ является исключением, оно определяется только заданием V, а наличие поля $\mathbb{K}$ вытекает из самого определения линейного векторного пространства V.

Вспомним, что $V^* = \operatorname{Hom}(V, \mathbb{K})$ есть линейное векторное пространство над $\mathbb{K}$. В случае конечномерности V размерность сопряженного пространства определяется при помощи теоремы 10.4 из первой главы: $\dim V^* = \dim V$. Структуру линейного векторного пространства в $V^* = \operatorname{Hom}(V, \mathbb{K})$ определяют две алгебраические операции: операция поточечного сложения и операция поточечного умножения на числа (см. определения 10.1 и 10.2 из первой главы). Определения этих операций для случая линейных функционалов уместно переформулировать отдельно.

Определение 1.2. Пусть f и g — два линейных функционала из V^*. Суммой функционалов f и g называется функционал h, значения которого определяются формулой $h(\mathbf{v}) = f(\mathbf{v}) + g(\mathbf{v})$ для всех $\mathbf{v} \in V$.

Определение 1.3. Произведением числа $\alpha \in \mathbb{K}$ на линейный функционал $f \in V^*$ называется функционал $h\colon V \to W$, значения которого вычисляются по формуле $h(\mathbf{v}) = \alpha\, f(\mathbf{v})$ для всех $\mathbf{v} \in V$.

Пусть V — конечномерное линейное векторное пространство над полем $\mathbb{K}$ и пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — базис в V. Тогда произвольный вектор $\mathbf{v} \in V$ может быть разложен по этому базису. Запишем это разложение так:

$$\mathbf{v} = v^1 \cdot \mathbf{e}_1 + \ldots + v^n \cdot \mathbf{e}_n. \tag{1.1}$$

Рассмотрим i-ую координату вектора $\mathbf{v}$. В силу однозначности разложения (1.1) при фиксированном выборе базиса в V координата v^i — это число, однозначно определяемое вектором $\mathbf{v}$. Рассмотрим отображения $h^i : V \to \mathbb{K}$, определив их формулой $h^i(\mathbf{v}) = v^i$. При сложении векторов их координаты в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$ складываются, а при умножении вектора на число его координаты умножаются на это число (см. соотношения (5.4) из первой главы). Отсюда вытекает линейность определенных выше отображений $h^i : V \to \mathbb{K}$. Значит, с каждым выбором базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в пространстве V связано некоторое семейство n функционалов из V^*. Функционалы $h^1, \ldots, h^n$ называются *координатными функционалами* базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Для них

$$h^i(\mathbf{e}_j) = \delta^i_j, \tag{1.2}$$

где δ^i_j — символ Кронекера. Соотношения (1.2) называются *соотношениями биортогональности.*

Соотношения биортогональности доказываются очень просто: если разложить вектор $\mathbf{e}_j$ по базису $\mathbf{e}_1, \ldots, \mathbf{e}_n$, то j-ая координата его будет равна 1, а остальные — нулю. Число $h^i(\mathbf{e}_j)$ есть i-ая координата вектора $\mathbf{e}_j$, поэтому $h^i(\mathbf{e}_j) = 1$ при $i = j$ и $h^i(\mathbf{e}_j) = \mathbf{0}$ в остальных случаях.

Теорема 1.1. *Координатные функционалы $h^1, \ldots, h^n$ линейно независимы и составляют базис в V^*.*

Док-во. Рассмотрим линейную комбинацию координатных функционалов базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$, равную нулю:

$$\alpha_1 \cdot h^1 + \ldots + \alpha_n \cdot h^n = 0. \tag{1.3}$$

Правая часть (1.3) есть нулевой функционал. Его значение на базисном векторе $\mathbf{e}_j$ равно нулю. Поэтому

$$\alpha_1\, h^1(\mathbf{e}_j) + \ldots + \alpha_n\, h^n(\mathbf{e}_j) = 0. \tag{1.4}$$

Воспользуемся соотношениями биортогональности (1.2). В силу этих соотношений из n слагаемых в правой части (1.4) остается лишь одно j-тое слагаемое, причем $h_j(\mathbf{e}_j) = 1$. Поэтому из (1.4) выводим $\alpha_j = 0$. Из произвольности j теперь вытекает тривиальность линейной комбинации (1.3) и линейная независимость координатных функционалов $h^1, \ldots, h^n$.

Для завершения доказательства теоремы теперь можно было бы использовать $\dim V^* = \dim V = n$ и сослаться на четвертый пункт в теореме 4.5 из первой главы. Однако, мы завершим доказательство теоремы, показав напрямую, что произвольный функционал $f \in V^*$ может быть разложен в линейную комбинацию координатных функционалов $h^1, \ldots, h^n$. Рассмоторим произвольный вектор v из пространства V. Тогда из разложения (1.1) для значения функционала f на векторе $\mathbf{v}$ получаем

$$f(\mathbf{v}) = v^1 f(\mathbf{e}_1) + \ldots + v^n f(\mathbf{e}_n) = f(\mathbf{e}_1)\, h^1(\mathbf{v}) + \ldots + f(\mathbf{e}_n)\, h^n(\mathbf{v}).$$

Здесь $f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)$ — набор чисел из $\mathbb{K}$. В силу произвольности вектора $\mathbf{v} \in V$ полученное равенство можно переписать как равенство функционалов:

$$f = f(\mathbf{e}_1) \cdot h^1 + \ldots + f(\mathbf{e}_n) \cdot h^n. \tag{1.5}$$

Формула (1.5) показывает разложимость произвольного функционала f в линейную комбинацию координатных функционалов. Коэффициенты такой линейной комбинации оказываются равными значениям функционала f на базисных векторах $\mathbf{e}_1, \ldots, \mathbf{e}_n$. $\square$

Определение 1.4. Базис $h^1, \ldots, h^n$ в V^*, построенный из координатных функционалов базиса $\mathbf{e}_1, \ldots \mathbf{e}_n$ в V, называется *дуальным* или *сопряженным базисом* для $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

Определение 1.5. Пусть f — линейный функционал в конечномерном пространстве V и пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — базис этого пространства. Числа $f_1, \ldots, f_n$, определяемые линейным функционалом f по формуле

$$f_i = f(\mathbf{e}_i), \tag{1.6}$$

называются *координатами* или *компонентами* линейного функционала f в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

Из формулы (1.5) видим, что числа (1.6) являются коэффициентами разложения функционала f в сопряженном базисе.

Алгебраические операции сложения и умножения на число в пространствах V и V^* связаны следующими соотношениями:

$$\begin{aligned} &f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2), && f(\alpha \cdot \mathbf{v}) = \alpha\, f(\mathbf{v});\\ &(f_1 + f_2)(\mathbf{v}) = f_1(\mathbf{v}) + f_2(\mathbf{v}), && (\alpha \cdot f)(\mathbf{v}) = \alpha\, f(\mathbf{v}). \end{aligned} \tag{1.7}$$

Векторы и функционалы входят в эти соотношения равноправным образом. Тот факт, что в записи $f(\mathbf{v})$ функционал играет роль функции, а вектор — роль аргумента, не имеет существенного значения. Поэтому для записи значения функционала f на векторе $\mathbf{v}$ иногда применяют обозначение:

$$f(\mathbf{v}) = \langle f \,|\, \mathbf{v} \rangle. \tag{1.8}$$

С записью (1.8) связана своя терминология. Функционалы из пространства V^* называют *ковекторами*, а само выражение $\langle f \,|\, \mathbf{v} \rangle$ называют *спариванием* или *сверткой* или же *скалярным произведением* вектора с ковектором.

Скалярное произведение (1.8) обладает свойством билинейности: оно линейно по первому аргументу и линейно по второму аргументу. Это вытекает из соотношений (1.7), которые теперь записываются так:

$$\begin{aligned} &\langle f_1 + f_2 \,|\, \mathbf{v} \rangle = \langle f_1 \,|\, \mathbf{v} \rangle + \langle f_2 \,|\, \mathbf{v} \rangle, && \langle \alpha \cdot f \,|\, \mathbf{v} \rangle = \alpha\, \langle f \,|\, \mathbf{v} \rangle;\\ &\langle f \,|\, \mathbf{v}_1 + \mathbf{v}_2 \rangle = \langle f \,|\, \mathbf{v}_1 \rangle + \langle f \,|\, \mathbf{v}_2 \rangle, && \langle f \,|\, \alpha \cdot \mathbf{v} \rangle = \alpha\, \langle f \,|\, \mathbf{v} \rangle. \end{aligned} \tag{1.9}$$

Понятие билинейности мы уже встречали (см. теорему 1.1 из второй главы).

Свойства (1.9) скалярного произведения (1.8) аналогичны свойствам скалярного произведения векторов в геометрическом пространстве — оно обычно изучается в курсе аналитической геометрии (см., например, [5]). Однако, в отличие от такого «геометрического» скалярного произведения, скалярное произведение (1.8) не является симметричным: аргументы в нем принадлежат разным пространствам и их нельзя переставлять. Ковекторы в скалярном произведении (1.8) всегда пишутся слева, а векторы — справа.

Следующее определение продиктовано стремлением углубить аналогию между (1.8) и обычным «геометрическим» скалярным произведением.

Определение 1.7. Вектор $\mathbf{v}$ и ковектор f называются *ортогональными* друг другу, если $\langle f\,|\,\mathbf{v}\rangle = 0$.

Теорема 1.2. *Пусть $U \subsetneq V$ — некоторое подпространство в конечномерном векторном пространстве V и пусть $\mathbf{v} \notin U$. Тогда найдется функционал f из V^*, такой, что $f(\mathbf{v}) \neq 0$ и $f(\mathbf{u}) = 0$ для всех $\mathbf{u} \in U$.*

Док-во. Пусть $\dim V = n$ и $\dim U = s$. Выберем базис $\mathbf{e}_1, \ldots, \mathbf{e}_s$ в подпространстве U. Добавим к векторам $\mathbf{e}_1, \ldots, \mathbf{e}_s$ вектор $\mathbf{v} = \mathbf{e}_{s+1}$. Полученная система векторов будет линейно независимой. Это вытекает из условия $\mathbf{v} \notin U$ и пункта (4) в теореме 3.1 из первой главы. Рассмотрим линейную оболочку полученной системы векторов $W = \langle \mathbf{e}_1, \ldots, \mathbf{e}_{s+1}\rangle$. Ясно, что W есть подпространство в V, оно содержит в себе исходное подпространство U и имеет на единицу большую размерность. Вектора $\mathbf{e}_1, \ldots, \mathbf{e}_{s+1}$ образуют базис в W. Если $W \neq V$, то дополним базис $\mathbf{e}_1, \ldots, \mathbf{e}_{s+1}$ до базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в пространстве V и рассмотрим координатные функционалы $h^1, \ldots, h^n$, определенные таким выбором базиса в V. Положим $f = h^{s+1}$. Тогда из соотношений биортогональности (1.2) выводим

$$f(\mathbf{v}) = h^{s+1}(\mathbf{e}_{s+1}) = 1 \quad \text{и} \quad f(\mathbf{e}_i) = 0 \quad \text{для} \quad i = 1, \ldots, s.$$

Будучи равным нулю на базисных векторах подпространства U, выбранный функционал f зануляется на всех векторах $\mathbf{u} \in U$. А его значение на векторе $\mathbf{v}$ равно единице. □

Рассмотрим случай $U = \{\mathbf{0}\}$ в доказанной теореме. Тогда для любого ненулевого вектора $\mathbf{v}$ имеем $\mathbf{v} \notin U$. Это позволяет сформулировать следующее следствие из теоремы 1.2.

Следствие. *Для всякого вектора $\mathbf{v} \neq 0$ в конечномерном пространстве V существует линейный функционал из V^*, такой, что $f(\mathbf{v}) \neq 0$.*

Пусть V — линейное векторное пространство над числовым полем $\mathbb{K}$ и пусть $W = V^*$ — сопряженное пространство для V. Из сказанного выше мы знаем, что W тоже есть линейное пространство над полем $\mathbb{K}$. Поэтому у него есть сопряженное пространство W^*. По отношению к V это пространство будет двойным сопряженным V^{**}. Мы можем рассмотреть также тройное сопряженное пространство, четверное и т. д. Возникает бесконечная последовательность сопряженных пространств. Однако, сейчас мы увидим, что в случае конечномерного пространства V в рассмотрении кратных сопряженных пространств нет особой необходимости.

Пусть $\mathbf{v} \in V$. Каждому элементу $f \in V^*$ поставим в соответствие число $f(\mathbf{v}) \in \mathbb{K}$. Это задает отображение $\varphi_v\colon V^* \to \mathbb{K}$, линейность которого вытекает из следующей серии вычислений:

$$\varphi_{\mathbf{v}}(f_1 + f_2) = (f_1 + f_2)(\mathbf{v}) = f_1(\mathbf{v}) + f_2(\mathbf{v}) = \varphi_{\mathbf{v}}(f_1) + \varphi_{\mathbf{v}}(f_2),$$

$$\varphi_{\mathbf{v}}(\alpha \cdot f) = (\alpha \cdot f)(\mathbf{v}) = \alpha\, f(\mathbf{v}) = \alpha\, \varphi_{\mathbf{v}}(f).$$

Значит, $\varphi_{\mathbf{v}}$ есть линейный функционал на пространстве V^* или, другими словами, элемент двойного сопряженного пространства V^{**}. Функционал $\varphi_{\mathbf{v}}$ определяется заданием вектора $\mathbf{v} \in V$. Поэтому сопоставление вектору $\mathbf{v}$ функционала $\varphi_{\mathbf{v}} \in V^{**}$ мы можем рассматривать как задание отображения

$$h\colon V \to V^{**}, \text{ где } h(\mathbf{v}) = \varphi_{\mathbf{v}} \text{ для всех } \mathbf{v} \in V. \tag{1.10}$$

Отображение (1.10) линейно. Для доказательства этого мы должны установить справедливость следующих тождеств:

$$h(\mathbf{v}_1 + \mathbf{v}_2) = h(\mathbf{v}_1) + h(\mathbf{v}_2), \qquad\qquad h(\alpha \cdot \mathbf{v}) = \alpha\, h(\mathbf{v}). \tag{1.11}$$

Результатом применения h к векторам пространства V являются функционалы из V^{**}. Поэтому проверка равенств (1.11) состоит в проверке совпадения результатов применения правой и левой частей этих равенств к произвольному ковектору f из пространства V^*. Тогда из следующих выкладок

$$\begin{gathered}h(\mathbf{v}_1 + \mathbf{v}_2)(f) = \varphi_{\mathbf{v}_1+\mathbf{v}_2}(f) = f(\mathbf{v}_1) + f(\mathbf{v}_2) = \varphi_{\mathbf{v}_1}(f) + \\ + \varphi_{\mathbf{v}_2}(f) = h(\mathbf{v}_1)(f) + h(\mathbf{v}_2)(f) = (h(\mathbf{v}_1) + h(\mathbf{v}_2))(f),\end{gathered}$$

$$\begin{gathered}h(\alpha \cdot \mathbf{v})(f) = \varphi_{\alpha\cdot\mathbf{v}}(f) = f(\alpha \cdot \mathbf{v}) = \alpha\, f(\mathbf{v}) = \\ = \alpha\varphi_{\mathbf{v}}(f) = \alpha\, h(\mathbf{v})(f) = (\alpha \cdot h(\mathbf{v}))(f).\end{gathered}$$

получаем соотношения (1.11), которые и означают линейность отображения h, заданного соотношением (1.10).

Теорема 1.3. *Для конечномерного пространства V отображение* (1.10) *биективно. Оно устанавливает изоморфизм пространств V и V^{**}, который называется каноническим изоморфизмом этих пространств.*

Док-во. Начнем с доказательства инъективности отображения (1.10). Для этого рассмотрим его ядро $\operatorname{Ker} h$. Пусть вектор $\mathbf{v}$ — некоторый произвольный вектор из ядра отображения h. Тогда $\varphi = h(\mathbf{v}) = 0$. Но $\varphi \in V^{**}$, это значит, что φ есть линейный функционал на пространстве V^*. Поэтому равенство $\varphi = 0$ означает, что $\varphi(f) = 0$ для любого ковектора f из V^*. Используя это обстоятельство, из формулы (1.10) выводим

$$h(\mathbf{v})(f) = \varphi_{\mathbf{v}}(f) = f(\mathbf{v}) = 0 \text{ для всех } f \in V^*. \tag{1.12}$$

Применим следствие из теоремы 1.2. Если бы вектор $\mathbf{v}$ был отличен от нуля, то для него существовал бы функционал f, такой, что $f(\mathbf{v}) \neq 0$. Это противоречило бы (1.12). Поэтому $\mathbf{v} = 0$, что доказывает тривиальность ядра $\operatorname{Ker} h = \{\mathbf{0}\}$ и устанавливает инъективность отображения h.

Для доказательства сюръективности отображения (1.10) используем теорему 9.4 из первой главы. Согласно этой теореме

$$\dim(\operatorname{Ker} h) + \dim(\operatorname{Im} h) = \dim V.$$

Но в силу уже доказанной инъективности $\dim(\operatorname{Ker} h) = 0$. Отсюда получаем $\dim(\operatorname{Im} h) = \dim V$. Но образ $\operatorname{Im} h$ есть подпространство в пространстве V^{**}, причем $\dim V^{**} = \dim V^{*} = \dim V$. Значит, $\dim(\operatorname{Im} h) = \dim V^{**}$ и $\operatorname{Im} h = V^{**}$ (см. третий пункт теоремы 4.5 из первой главы). Это завершает доказательство сюръективности отображения h и всей теоремы в целом. □

Канонический изоморфизм $h\colon V \to V^{**}$ из (1.10) обладает тем свойством, что для произвольного вектора $\mathbf{v} \in V$ и для произвольного ковектора $f \in V^{*}$ имеет место следующее соотношение

$$\langle h(\mathbf{v}) \,|\, f\rangle = \langle f \,|\, \mathbf{v}\rangle. \tag{1.13}$$

Равенство (1.13) выводится из определения отображения h. Действительно, $\langle h(\mathbf{v}) \,|\, f\rangle = h(\mathbf{v})(f) = \varphi_{\mathbf{v}}(f) = f(\mathbf{v}) = \langle f \,|\, \mathbf{v}\rangle$. Соотношение (1.13) выделяет канонический изоморфизм из числа всех возможных изоморфизмов, связывающих пространство V с пространством V^{**}.

§ 2. Преобразование координат ковектора при замене базиса.

Пусть V — некоторое конечномерное линейное векторное пространство и пусть V^{*} — соответствующее ему сопряженное пространство. Если рассматривать пространство V^{*} изолированно, вне всякой связи с пространством V, то выбор базиса, и замена базиса в V^{*} ничем не отличается выбора и замены базиса в любом другом линейном векторном пространстве. Однако, сопряженное пространство V^{*} практически никогда не рассматривается само по себе. Построение этого пространства следует рассматривать как расширение теории исходного пространства V.

Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — базис в конечномерном пространстве V. С каждым таким базисом связан сопряженный базис из координатных функционалов $h^1, \ldots, h^n$ в пространстве ковекторов V^{*}. Выбор другого базиса $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ в V определяет другой сопряженный базис $\tilde{h}^1, \ldots, \tilde{h}^n$ в V^{*}. Обозначим через S матрицу перехода из базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в базис $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$. Аналогичным образом обозначим через P матрицу перехода из базиса $h^1, \ldots, h^n$ в базис $\tilde{h}^1, \ldots, \tilde{h}^n$. Компоненты этих матриц входят в разложения базисных векторов «волнистых» базисов по «неволнистым» базисам:

$$\tilde{\mathbf{e}}_j = \sum_{i=1}^{n} S^i_j \cdot \mathbf{e}_i, \qquad\qquad \tilde{h}^r = \sum_{s=1}^{n} P^r_s \cdot h^s. \tag{2.1}$$

Отметим, что вторая формула в (2.1) нарушает стандарт записи подобных соотношений, введенный формулой (5.5) в первой главе. Причина в том, что вектора сопряженных базисов нумеруются верхними индексами, что также нестандартно. Это объясняется тем, что сопряженное пространство и сопряженный базис являются вторичными (производными) понятиями, по отношению к пространству V и базисам в нем. Аналогичное нарушение стандарта мы уже встречали при построении базиса E^i_j в пространстве $\operatorname{Hom}(V, W)$ при доказательстве теоремы 10.4 в первой главе.

Несмотря на нарушение стандарта в индексации базисных векторов, вторая формула (2.1) не нарушает общих принципов тензорной нотации: свободный индекс r имеют одинаковую позицию в обеих частях равенства, а индекс суммирования s встречается дважды — один раз наверху, второй раз — внизу.

Теорема 2.1. *Матрица перехода P из сопряженного базиса $h^1, \ldots, h^n$ в сопряженный базис $\tilde{h}^1, \ldots, \tilde{h}^n$ является обратной для матрицы перехода S из базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в базис $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$.*

Док-во. Для доказательства теорема воспользуемся соотношениями биортогональности (1.2). Подставляя (2.1) в эти соотношения, получим

$$\delta^r_j = h^r(\mathbf{e}_j) = \sum_{s=1}^{n}\sum_{i=1}^{n} P^r_s\, S^i_j\, h^s(\mathbf{e}_i) = \sum_{s=1}^{n}\sum_{i=1}^{n} P^r_s\, S^i_j\, \delta^s_i = \sum_{i=1}^{n} P^r_i\, S^i_j.$$

Полученное соотношение может быть записано в матричной форме. Оно имеет вид $P\,S = 1$, откуда $P = S^{-1}$. Теорема доказана. $\square$

Вспомним, что матрица обратного перехода T из базиса $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ в базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$ также является обратной для S. Поэтому при записи полной совокупности формул для связи базисов в V и V^* достаточно матриц S и T:

$$\begin{aligned}
\tilde{\mathbf{e}}_j &= \sum_{i=1}^{n} S^i_j \cdot \mathbf{e}_i, & \tilde{h}^r &= \sum_{s=1}^{n} T^r_s \cdot h^s,\\
\mathbf{e}_i &= \sum_{j=1}^{n} T^j_i \cdot \tilde{\mathbf{e}}_j, & h^s &= \sum_{r=1}^{n} S^s_r \cdot \tilde{h}^r.
\end{aligned} \tag{2.2}$$

Пусть f — ковектор из сопряженного пространства V^*. Рассмотрим его разложения по сопряженным базисам:

$$f = \sum_{s=1}^{n} f_s \cdot h^s, \qquad\qquad f = \sum_{r=1}^{n} \tilde{f}_r \cdot \tilde{h}^r. \tag{2.3}$$

Разложения (2.3) также отходят от стандарта, введенного формулой (5.1) из первой главы. Для координат ковекторов принят другой стандарт: они нумеруются нижними индексами и записываются в форме вектор-строк.

Соотношения (2.2) выражают связь между базисами в пространстве V и между соответствующими им сопряженными базисами в V^*. Они же, в конечном счете, определяют связь между координатами ковектора f в разложениях (2.3).

Теорема 2.2. *Связь между координатами ковектора f, разложенного по двум сопряженным базисам, соответствующим базисам $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ в пространстве V, дается формулами*

$$\tilde{f}_r = \sum_{s=1}^{n} S^s_r\, f_s, \qquad\qquad f_s = \sum_{j=1}^{n} T^r_s\, \tilde{f}_r, \tag{2.4}$$

где S — матрица прямого перехода из $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$, а T — матрица обратного перехода.

Док-во. Для доказательства первого соотношения (2.4) подставим в первое разложение (2.3) выражение для ковектора h^s, даваемое одним из соотношений (2.2). В результате этого получим

$$f = \sum_{s=1}^{n} f_s \cdot \left(\sum_{r=1}^{n} S^s_r \cdot \tilde{h}^r\right) = \sum_{r=1}^{n} \left(\sum_{s=1}^{n} S^s_r\, f_s\right) \cdot \tilde{h}^r.$$

Сравнив полученное разложение ковектора f со вторым разложением (2.3), получаем требуемое первое соотношение (2.4). Вторая формула в (2.4) выводится аналогично. $\square$

Отметим, что соотношения (2.4) могут быть выведены непосредственным образом из определения 1.5 и формулы (1.6) без использования понятия сопряженного базиса.

Теорема 2.3. *Скалярное произведение вектора $\mathbf{v}$ и ковектора f определяется их координатами по формуле*

$$\langle f \,|\, \mathbf{v} \rangle = \sum_{i=1}^{n} f_i\, v^i = f_1\, v^1 + \ldots + f_n\, v^n. \tag{2.5}$$

Док-во. Для доказательства (2.5) воспользуемся соотношением (1.6):

$$\langle f \,|\, \mathbf{v} \rangle = f(\mathbf{v}) = \sum_{i=1}^{n} f(\mathbf{e}_i)\, v^i = \sum_{i=1}^{n} f_i\, v^i.$$

В самой формуле (2.5) и в приведенных вычислениях мы считаем f разложенным по базису, который является сопряженным для базиса, в котором разложен вектор $\mathbf{v}$. $\square$

§ 3. Ортогональные дополнения в сопряженном пространстве.

Определение 3.1. Пусть S — некоторое подмножество в линейном векторном пространстве V. *Ортогональным дополнением* множества S в сопряженном пространстве V^* называется множество $S^\perp$, состоящее из ковекторов, которые ортогональны всем векторам из множества S.

Определение ортогонального дополнения $S^\perp$ можно изобразить формулой $S^\perp = \{f \in V^* \colon \forall \mathbf{v}\, ((\mathbf{v} \in S) \Rightarrow (\langle f \,|\, \mathbf{v} \rangle = 0))\}$.

Теорема 3.1. *Операция построения ортогональных дополнений для подмножеств $S \subset V$ в пространстве V^* обладает следующими свойствами:*

(1) *$S^\perp$ есть подпространство в V^*;*
(2) *из $S_1 \subset S_2$ вытекает $(S_2)^\perp \subset (S_1)^\perp$;*
(3) *$\langle S \rangle^\perp = S^\perp$, где $\langle S \rangle$ — линейная оболочка множества S;*
(4) $\left(\bigcup\limits_{i \in I} S_i\right)^{\perp} = \bigcap\limits_{i \in I} (S_i)^\perp.$

Док-во. Начнем с доказательства первого пункта. Для этого проверим два условия из определения подпространства. Пусть $f_1, f_2 \in S^{\perp}$, тогда $\langle f_1 \,|\, \mathbf{v}\rangle = 0$ и $\langle f_2 \,|\, \mathbf{v}\rangle = 0$ для всех $\mathbf{v} \in S$. Отсюда для всех $\mathbf{v} \in S$ имеем

$$\langle f_1 + f_2 \,|\, \mathbf{v}\rangle = \langle f_1 \,|\, \mathbf{v}\rangle + \langle f_2 \,|\, \mathbf{v}\rangle = 0,$$

что для суммы ковекторов $f_1 + f_2$ означает $f_1 + f_2 \in S^{\perp}$.

Пусть теперь $f \in S^{\perp}$. Тогда $\langle f \,|\, \mathbf{v}\rangle = 0$ для всех векторов $\mathbf{v} \in S$. Отсюда для ковектора $\alpha \cdot f$ выводим

$$\langle \alpha \cdot f \,|\, \mathbf{v}\rangle = \alpha \,\langle f \,|\, \mathbf{v}\rangle = 0.$$

Значит, $\alpha \cdot f \in S^{\perp}$. Первый пункт в теореме 3.1 доказан.

Для доказательства включения $(S_2)^{\perp} \subset (S_1)^{\perp}$ во втором пункте теоремы 3.1 рассмотрим произвольный ковектор f из $(S_2)^{\perp}$. Из условия $f \in (S_2)^{\perp}$ вытекает $\langle f \,|\, \mathbf{v}\rangle = 0$ для всякого $\mathbf{v} \in S_2$. Но $S_1 \subset S_2$, поэтому равенство $\langle f \,|\, \mathbf{v}\rangle = 0$ имеет место для всякого $\mathbf{v} \in S_1$. Отсюда $f \in (S_1)^{\perp}$. Значит, из $f \in (S_2)^{\perp}$ вытекает $f \in (S_1)^{\perp}$, что доказывает требуемое включение.

При доказательстве третьего пункта теоремы, заметим, что линейная оболочка множества S содержит в себе это множество: $S \subset \langle S\rangle$. Применив уже доказанный пункт (2), получаем включение $\langle S\rangle^{\perp} \subset S^{\perp}$.

Остается установить противоположное включение. Линейная оболочка $\langle S\rangle$ состоит из всевозможных конечных линейных комбинаций вида

$$\mathbf{v} = \alpha_1 \cdot \mathbf{v}_1 + \ldots + \alpha_r \cdot \mathbf{v}_r, \tag{3.1}$$

где $\mathbf{v}_i \in S$. Пусть $f \in S^{\perp}$, тогда $\langle f \,|\, \mathbf{v}\rangle = 0$ для всех $\mathbf{v} \in S$. В частности, это относится и к векторам $\mathbf{v}_i$ из разложения (3.1). Для них $\langle f \,|\, \mathbf{v}_i\rangle = 0$. Тогда

$$\langle f \,|\, \mathbf{v}\rangle = \alpha_1 \,\langle f \,|\, \mathbf{v}_1\rangle + \ldots + \alpha_r \,\langle f \,|\, \mathbf{v}_r\rangle = 0.$$

То есть $\langle f \,|\, \mathbf{v}\rangle = 0$ для всех $\mathbf{v} \in \langle S\rangle$, что доказывает противоположное включение $S^{\perp} \subset \langle S\rangle^{\perp}$ и, тем самым, завершает доказательство совпадения $\langle S\rangle^{\perp} = S^{\perp}$.

Перейдем к доказательству четвертого пункта в теореме. С этой целью введем следующие обозначения:

$$S = \bigcup_{i \in I} S_i, \qquad\qquad \tilde{S} = \bigcap_{i \in I} (S_i)^{\perp}.$$

Пусть $f \in S^{\perp}$. Тогда $\langle f \,|\, \mathbf{v}\rangle = 0$ для всех $\mathbf{v} \in S$. Но $S_i \subset S$ для всякого $i \in I$. Поэтому $\langle f \,|\, \mathbf{v}\rangle = 0$ для всех $\mathbf{v} \in S_i$ и для всех $i \in I$. Это означает, что f принадлежит каждому из ортогональных дополнений $(S_i)^{\perp}$, поэтому он принадлежит их пересечению. Это доказывает включение $S^{\perp} \subset \tilde{S}$.

Обратно, из принадлежности $f \in (S_i)^{\perp}$ для всех $i \in I$ вытекает равенство нулю $\langle f \,|\, \mathbf{v}\rangle = 0$ для всех $\mathbf{v} \in S_i$ и для всех $i \in I$. Значит, это равенство нулю имеет место для всех векторов $\mathbf{v}$ из объединения всех множеств S_i. Это доказывает противоположное включение $\tilde{S} \subset S^{\perp}$.

Полученные два включения доказывают необходимое совпадение двух множеств $S^{\perp} = \tilde{S}$. Теорема 3.1 доказана. □

Определение 3.2. Пусть S — некоторое подмножество в сопряженном пространстве V^*. *Ортогональным дополнением* множества S в пространстве V называется множество $S^\perp$, состоящее из векторов, которые ортогональны всем ковекторам из множества S.

Это определение $S^\perp$ также можно записать в виде формального определения $S^\perp = \{\mathbf{v} \in V : \forall f \ ((f \in S) \Rightarrow (\langle f \,|\, \mathbf{v} \rangle = 0))\}$. Для него имеет место теорема, аналогичная теореме 3.1.

Теорема 3.2. *Операция построения ортогональных дополнений для подмножеств сопряженного пространства V^* в пространстве V обладает следующими свойствами:*

(1) *$S^\perp$ есть подпространство в V;*
(2) *из $S_1 \subset S_2$ вытекает $(S_2)^\perp \subset (S_1)^\perp$;*
(3) *$\langle S \rangle^\perp = S^\perp$, где $\langle S \rangle$ — линейная оболочка множества S;*
(4) $\left(\bigcup\limits_{i \in I} S_i\right)^\perp = \bigcap\limits_{i \in I} (S_i)^\perp.$

Доказательство этой теоремы является почти дословным повторением теоремы 3.1, поэтому его мы не приводим.

Теорема 3.3. *Пусть задано конечномерное векторное пространство V и два подпространства $U \subset V$ и $W \subset V^*$. Тогда условие $W = U^\perp$ в смысле определения 3.1 эквивалентно условию $U = W^\perp$ в смысле определения 3.2.*

Док-во. Пусть выполнено условие $W = U^\perp$. Тогда для всякого $w \in W$ и всякого $\mathbf{u} \in U$ имеет место ортогональность $\langle w \,|\, \mathbf{u} \rangle = 0$. Но $W^\perp$ — это множество, состоящее из векторов $\mathbf{v}$, перпендикулярных всем ковекторам $w \in W$. Значит, $\mathbf{u} \in U$ влечет $\mathbf{u} \in W^\perp$ и имеет место включение $U \subset W^\perp$.

Допустим, что совпадения нет: $U \neq W^\perp$. Тогда существует вектор $\mathbf{v}_0 \in W^\perp$, не принадлежащий подпространству U. Применим теорему 1.2, которая утверждает, что существует функционал f, равный нулю на всех векторах $\mathbf{u} \in U$, но отличный от нуля на векторе $\mathbf{v}_0$. Это дает $f \in W$ и $\langle f \,|\, \mathbf{v}_0 \rangle \neq 0$, что противоречит условию $\mathbf{v}_0 \in W^\perp$. Полученное противоречие доказывает совпадение $U = W^\perp$. Таким образом, из $W = U^\perp$ вытекает $U = W^\perp$.

Пусть, наоборот, выполнено условие $U = W^\perp$. Тогда для всякого $w \in W$ и всякого $\mathbf{u} \in U$ имеет место ортогональность $\langle w \,|\, \mathbf{u} \rangle = 0$. Но $U^\perp$ — это множество, состоящее из ковекторов f, перпендикулярных всем векторам $\mathbf{u} \in U$. Значит, $w \in W$ влечет $w \in U^\perp$ и имеет место включение $W \subset U^\perp$.

Допустим, что совпадения нет: $W \neq U^\perp$. Тогда существует ковектор $f_0 \in U^\perp$, не принадлежащий подпространству W. Применим теорему 1.2. В данном случае она утверждает, что существует функционал φ из пространства V^{**}, равный нулю на всех ковекторах $w \in W$, но отличный от нуля на ковекторе f_0. Воспользуемся существованием канонического изоморфизма $h\colon V \to V^{**}$ и рассмотрим вектор $\mathbf{v} = h^{-1}(\varphi)$. С учетом соотношения (1.13) это дает $\mathbf{v} \in U$ и $\langle f_0 \,|\, \mathbf{v} \rangle \neq 0$, что противоречит условию $f_0 \in U^\perp$. Полученное противоречие доказывает совпадение $W = U^\perp$. Таким образом, из $U = W^\perp$ вытекает $W = U^\perp$. Теорема доказана. $\square$

Утверждение теоремы 3.3 можно переформулировать так: в случае конечномерности пространства V для любых двух подпространств $U \in V$ и $W \in V^*$

выполнены соотношения:

$$(U^\perp)^\perp = U, \qquad\qquad (W^\perp)^\perp = W. \tag{3.2}$$

Для произвольных множеств $S \in V$ и $R \in V^*$ в случае конечномерности пространства V получаем соотношения

$$(S^\perp)^\perp = \langle S\rangle, \qquad\qquad (R^\perp)^\perp = \langle R\rangle. \tag{3.3}$$

Соотношения (3.3) выводятся из (3.2) и пункта (3) в теоремах 3.1 и 3.2.

Теорема 3.4. *В случае конечномерного пространства V для любого подпространства $U \subset V$ или $U \subset V^*$ выполнено соотношение*

$$\dim U + \dim U^\perp = \dim V.$$

Док-во. В силу соотношений (3.2) случай $U \subset V^*$ сводится к случаю $U \subset V$ заменой U на $U^\perp$. Поэтому рассмотрим только случай $U \subset V$.

Пусть $\dim V = n$ и $\dim U = s$. Выберем базис $\mathbf{e}_1, \ldots, \mathbf{e}_s$ в подпространстве U и дополним его до базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в пространстве V. Базисом $\mathbf{e}_1, \ldots, \mathbf{e}_n$ определяется сопряженный базис $h^1, \ldots, h^n$ в пространстве V^*. Задавая вектора их координатами в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$, а ковектора — координатами в сопряженном базисе $h^1, \ldots, h^n$, их скалярное произведение можно вычислять по формуле (2.5) из теоремы 2.3.

По построению базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ подпространство U состоит из векторов, первые s координат которых произвольны, а последние $n - s$ равны нулю. Поэтому условие $f \in U^\perp$ означает, что для ковектора f равенство

$$\langle f \,|\, \mathbf{v}\rangle = \sum_{i=1}^{s} f_i\, v^i = 0$$

должно выполняться тождественно при произвольном выборе чисел $v^1, \ldots, v^s$. Это возможно в том и только в том случае, когда первые s координат ковектора f равны нулю. Значения остальных координат этого ковектора ничем не ограничены. Поэтому пространство $U^\perp$ есть линейная оболочка последних $n - s$ векторов сопряженного базиса:

$$U^\perp = \langle h^{s+1}, \ldots, h^n\rangle.$$

Для размерности пространства $U^\perp$ это дает $\dim U^\perp = n - s$, откуда вытекает требуемое равенство $\dim U + \dim U^\perp = \dim V$. Теорема доказана. $\square$

Доказанная теорема 3.4 известна как теорема *о размерности ортогональных дополнений.* Как немедленное следствие из этой теоремы получаем

$$\begin{aligned} &\{0\}^\perp = V, && V^\perp = \{0\}, \\ &\{\mathbf{0}\}^\perp = V^*, && (V^*)^\perp = \{\mathbf{0}\}. \end{aligned} \tag{3.4}$$

Соотношения (3.4) имеют прозрачную интерпретацию, причем первые три из них могут быть доказаны непосредственно без учета конечномерности пространства V. Доказательство последнего использует следствие из теоремы 1.2, а сама эта теорема предполагает конечномерность V.

Теорема 3.5. *В случае конечномерного пространства V для любого семейства подпространств в V или в V^* выполнены соотношения*

$$\left(\sum_{i\in I} U_i\right)^{\perp}=\bigcap_{i\in I}(U_i)^{\perp},\qquad\qquad \left(\bigcap_{i\in I} U_i\right)^{\perp}=\sum_{i\in I}(U_i)^{\perp}.\tag{3.5}$$

Док-во. Сумма подпространств есть линейная оболочка их объединения. Поэтому первое из соотношений (3.5) является немедленным следствием пунктов (3) и (4) в теоремах 3.1 и 3.2. Оно не использует конечномерности пространства V.

Второе соотношение (3.5) вытекает из первого после замены U_i на $(U_i)^{\perp}$:

$$\left(\sum_{i\in I}(U_i)^{\perp}\right)^{\perp}=\bigcap_{i\in I}((U_i)^{\perp})^{\perp}=\bigcap_{i\in I}U_i.$$

Остается лишь перейти к ортогональным дополнениям в обеих частях полученного равенства и применить результат теоремы 3.3. Теорема доказана. $\square$

§ 4. Сопряженное отображение.

Определение 4.1. Пусть $f : V \to W$ — линейное отображение. Отображение $\varphi : W^* \to V^*$ называется *сопряженным отображением* для f, если для любого $\mathbf{v} \in V$ и для любого $\mathbf{w} \in W^*$ выполнено соотношение $\langle\varphi(w)\,|\,\mathbf{v}\rangle = \langle w\,|\,f(\mathbf{v})\rangle$.

Вопрос о существовании сопряженного отображения решается самим же определением 4.1. Действительно, для того, чтобы задать отображение $\varphi\colon W^* \to V^*$ надо для каждого функционала $w \in W^*$ указать соответствующий ему функционал $h = \varphi(w) \in V^*$. Но задать функционал из V^* — это значит определить его действие на произвольный вектор $\mathbf{v} \in V$. В свете этого рассуждения определяющее соотношение для сопряженного отображения можно переписать следующим образом:

$$h(\mathbf{v}) = \langle h\,|\,\mathbf{v}\rangle = \langle\varphi(w)\,|\,\mathbf{v}\rangle = \langle w\,|\,f(\mathbf{v})\rangle.$$

Проверка того, что функционал $h(\mathbf{v})$, заданный этим соотношением линеен, не представляет большого труда:

$$\begin{gathered}h(\mathbf{v}_1+\mathbf{v}_2)=\langle w\,|\,f(\mathbf{v}_1+\mathbf{v}_2)\rangle=\langle w\,|\,f(\mathbf{v}_1)+f(\mathbf{v}_2)\rangle=\\ =\langle w\,|\,f(\mathbf{v}_1)\rangle+\langle w\,|\,f(\mathbf{v}_2)\rangle=h(\mathbf{v}_1)+h(\mathbf{v}_2),\\[1ex] h(\alpha\cdot\mathbf{v})=\langle w\,|\,f(\alpha\cdot\mathbf{v})\rangle=\langle w\,|\,\alpha\cdot f(\mathbf{v})\rangle=\alpha\,\langle w\,|\,f(\mathbf{v})\rangle=\alpha\,h(\mathbf{v}).\end{gathered}$$

Теорема 4.1. *Для линейного отображения $f\colon V \to W$ сопряженное отображение $\varphi\colon W^* \to V^*$ также линейно.*

Док-во. В силу самого определения 4.1 для сопряженного отображения $\varphi\colon W^* \to V^*$ имеют место следующие соотношения:

$$\begin{gathered}\varphi(w_1+w_2)(\mathbf{v})=\langle w_1+w_2\,|\,f(\mathbf{v})\rangle=\langle w_1\,|\,f(\mathbf{v})\rangle+\\ +\langle w_2\,|\,f(\mathbf{v})\rangle=\varphi(w_1)(\mathbf{v})+\varphi(w_2)(\mathbf{v})=(\varphi(w_1)+\varphi(w_2))(\mathbf{v}),\end{gathered}$$

$$\begin{aligned}\varphi(\alpha\cdot w)(\mathbf{v})=\langle\alpha\cdot w\,|\,f(\mathbf{v})\rangle=\alpha\,\langle w\,|\,f(\mathbf{v})\rangle=\\=\alpha\,\varphi(w)(\mathbf{v})=(\alpha\cdot\varphi(w))(\mathbf{v}).\end{aligned}$$

Из этих выкладок в силу произвольности вектора $\mathbf{v}\in V$ в них получаем $\varphi(w_1+w_2)=\varphi(w_1)+\varphi(w_2)$ и $\varphi(\alpha\cdot w)=\alpha\cdot\varphi(w)$, что доказывает линейность сопряженного отображения φ. $\square$

Сопряженное отображение $\varphi\colon\, W^*\to V^*$ для $f\colon\, V\to W$ принято обозначать $\varphi=f^*$. Операция перехода от f к f^* обладает следующими свойствами:

$$(f+g)^*=f^*+g^*,\qquad (\alpha\cdot f)^*=\alpha\cdot f^*,\qquad (f\circ g)^*=g^*\circ f^*.$$

Первые два свойства естественно назвать линейностью, а последнее свойство делает операцию $f\to f^*$ некоторым аналогом транспонирования матриц. Перечисленные выше три свойства операции $f\to f^*$ доказываются непосредственным вычислением на базе определения 4.1. Мы не приводим этих выкладок, поскольку и сами свойства нигде далее использоваться не будут.

Теорема 4.2. *В случае конечномерности пространств V и W ядра и образы отображений $f\colon\, V\to W$ и $f^*\colon\, W^*\to V^*$ связаны следующими соотношениями:*

$$\begin{aligned}&\operatorname{Ker} f^*=(\operatorname{Im} f)^\perp, &&\operatorname{Ker} f=(\operatorname{Im} f^*)^\perp,\\&\operatorname{Im} f=(\operatorname{Ker} f^*)^\perp, &&\operatorname{Im} f^*=(\operatorname{Ker} f)^\perp.\end{aligned}\tag{4.1}$$

Док-во. Ядро $\operatorname{Ker} f^*$ — это множество функционалов из W^*, которые под действием f^* отображаются в нулевой функционал из V^*. Поэтому $w\in\operatorname{Ker} f^*$ эквивалентно равенству $f^*(w)(\mathbf{v})=0$ для всех $\mathbf{v}\in V$. В результате несложных преобразований это условие трансформируется следующим образом:

$$f^*(w)(\mathbf{v})=\langle f^*(w)\,|\,\mathbf{v}\rangle=\langle w\,|\,f(\mathbf{v})\rangle=0.$$

Значит, ядро $\operatorname{Ker} f^*$ есть множество ковекторов ортогональных векторам вида $f(\mathbf{v})$. Но вектора вида $f(\mathbf{v})\in W$ составляют образ $\operatorname{Im} f$. Поэтому $\operatorname{Ker} f^*=(\operatorname{Im} f)^\perp$. Первое соотношение (4.1) доказано. При его доказательстве конечномерность пространства W не используется. Оно верно и для бесконечномерных пространств.

Для доказательства второго соотношения рассмотрим ортогональное дополнение $(\operatorname{Im} f^*)^\perp$. Оно состоит из векторов, ортогональных всем ковекторам вида $f^*(w)$. То есть выполняются следующие равенства:

$$0=\langle f^*(w)\,|\,\mathbf{v}\rangle=\langle w\,|\,f(\mathbf{v})\rangle.$$

Используя конечномерность пространства W, применим следствие из теоремы 1.2. Оно означает, что выполнение равенства $\langle w\,|\,f(\mathbf{v})\rangle=0$ для всех $w\in W^*$ эквивалентно условию $f(\mathbf{v})=0$. Поэтому $(\operatorname{Im} f^*)^\perp=\operatorname{Ker} f$. Второе соотношение (4.1) доказано. Третье и четвертое соотношения (4.1) выводятся из первого и второго при помощи теоремы 3.3. При этом используется конечномерность пространств W и V. $\square$

Пусть пространства V и W конечномерны. Выберем базис $\mathbf{e}_1,\,\ldots\,,\mathbf{e}_n$ в V и базис $\tilde{\mathbf{e}}_1,\,\ldots\,,\tilde{\mathbf{e}}_m$ в пространстве W. Этим выбором определяется выбор сопряженных базисов $h^1, \ldots, h^n$ и $\tilde{h}^1, \ldots, \tilde{h}^m$ в пространствах V^* и W^*. Рассмотрим отображение $f\colon V \to W$ и сопряженное ему отображение $f^*\colon W^* \to V^*$. Матрицы отображений f и f^* определяется из разложений:

$$f(e_j) = \sum_{k=1}^{m} F_j^k\, \tilde{e}_k, \qquad\qquad f^*(\tilde{h}^i) = \sum_{q=1}^{n} \Phi_q^i\, h^q. \tag{4.2}$$

Второе соотношение (4.1) несколько отличается по структуре от первого соотношения. Это объясняется отличием в индексации базисных векторов сопряженных базисов. Однако, в нем реализована та же идея, что и в первом соотношении: отображение f^* применяется к базисному вектору одного пространства, а результат раскладывается по базису во втором.

Теорема 4.3. *Матрицы отображений f и f^*, определяемые соотношениями (4.2), совпадают, то есть $F_j^i = \Phi_j^i$.*

Док-во. Из самого определения сопряженного отображения имеем

$$\langle \tilde{h}^i \,|\, f(\mathbf{e}_j) \rangle = \langle f^*(\tilde{h}^i) \,|\, \mathbf{e}_j \rangle. \tag{4.3}$$

Вычислим отдельно левую и правую части этого равенства, используя (4.2):

$$\begin{aligned}
\langle \tilde{h}^i \,|\, f(\mathbf{e}_j) \rangle &= \sum_{k=1}^{m} F_j^k\, \langle \tilde{h}^i \,|\, \tilde{\mathbf{e}}_k \rangle = \sum_{k=1}^{m} F_j^k\, \delta_k^i = F_j^i,\\
\langle f^*(\tilde{h}^i) \,|\, \mathbf{e}_j \rangle &= \sum_{q=1}^{n} \Phi_q^i\, \langle h^q \,|\, \mathbf{e}_j \rangle = \sum_{q=1}^{n} \Phi_q^i\, \delta_j^q = \Phi_j^i.
\end{aligned}$$

Подставив полученные выражения в формулу (4.3), получаем требуемое совпадение матриц $F_j^i = \Phi_j^i$. $\square$

Замечание. В некоторых теоремах этой главы ограничения, связанные с конечномерностью пространств можно снять. Но доказательство усиленного варианта таких теорем опирается на использование аксиомы выбора (см. [1]).

ГЛАВА IV

БИЛИНЕЙНЫЕ И КВАДРАТИЧНЫЕ ФОРМЫ.

§ 1. Симметрические билинейные формы и квадратичные формы. Формула восстановления.

Определение 1.1. Пусть V — линейное векторное пространство над полем $\mathbb{K}$. Числовая функция $y = f(v, w)$ с двумя аргументами из V и значениями из поля $\mathbb{K}$ называется *билинейной формой* на пространстве V, если:

(1) $f(\mathbf{v}_1 + \mathbf{v}_2, \mathbf{w}) = f(\mathbf{v}_1, \mathbf{w}) + f(\mathbf{v}_2, \mathbf{w})$;
(2) $f(\alpha \cdot \mathbf{v}, \mathbf{w}) = \alpha\, f(\mathbf{v}, \mathbf{w})$;
(3) $f(\mathbf{v}, \mathbf{w}_1 + \mathbf{w}_2) = f(\mathbf{v}, \mathbf{w}_1) + f(\mathbf{v}, \mathbf{w}_2)$;
(4) $f(\mathbf{v}, \alpha \cdot \mathbf{w}) = \alpha\, f(\mathbf{v}, \mathbf{w})$.

Билинейная форма $f(\mathbf{v}, \mathbf{w})$ линейна по первому аргументу при фиксированном втором аргументе, она также линейна по второму аргументу при фиксированном первом аргументе.

Определение 1.2. Билинейная форма $f(\mathbf{v}, \mathbf{w})$ называется *симметричной*, если $f(\mathbf{v}, \mathbf{w}) = f(\mathbf{w}, \mathbf{v})$.

Определение 1.3. Билинейная форма $f(\mathbf{v}, \mathbf{w})$ называется *кососимметричной*, если $f(\mathbf{v}, \mathbf{w}) = -f(\mathbf{w}, \mathbf{v})$.

Кососимметричные билинейные формы иногда называют также *антисимметричными*. Из любой билинейной формы $f(\mathbf{v}, \mathbf{w})$ можно получить симметричную билинейную форму

$$f_+(\mathbf{v}, \mathbf{w}) = \frac{f(\mathbf{v}, \mathbf{w}) + f(\mathbf{w}, \mathbf{v})}{2}. \tag{1.1}$$

Из нее также можно построить и кососимметричную форму

$$f_-(\mathbf{v}, \mathbf{w}) = \frac{f(\mathbf{v}, \mathbf{w}) - f(\mathbf{w}, \mathbf{v})}{2}. \tag{1.2}$$

Операцию (1.1) называют *симметрированием* формы f, а операцию (1.2) называют *альтернированием* формы. При этом любая билинейная форма раскладывается в сумму симметричной и антисимметричной форм:

$$f(\mathbf{v}, \mathbf{w}) = f_+(\mathbf{v}, \mathbf{w}) + f_-(\mathbf{v}, \mathbf{w}). \tag{1.3}$$

Теорема 1.1. *Разложение билинейной формы $f(\mathbf{v}, \mathbf{w})$ в сумму симметричной и антисимметричной форм единственно.*

Док-во. Рассмотрим некоторое разложение $f(\mathbf{v}, \mathbf{w})$ в сумму симметричной и антисимметричной билинейных форм:

$$f(\mathbf{v}, \mathbf{w}) = h_{+}(\mathbf{v}, \mathbf{w}) + h_{-}(\mathbf{v}, \mathbf{w}). \tag{1.4}$$

Из (1.4) в результате симметрирования и альтернирования билинейной формы f выводим следующие соотношения:

$$\begin{aligned} f(\mathbf{v}, \mathbf{w}) + f(\mathbf{w}, \mathbf{v}) &= (h_{+}(\mathbf{v}, \mathbf{w}) + h_{+}(\mathbf{w}, \mathbf{v})) + \\ &+ (h_{-}(\mathbf{v}, \mathbf{w}) + h_{-}(\mathbf{w}, \mathbf{v})) = 2\,h_{+}(\mathbf{v}, \mathbf{w}), \\ f(\mathbf{v}, \mathbf{w}) - f(\mathbf{w}, \mathbf{v}) &= (h_{+}(\mathbf{v}, \mathbf{w}) - h_{+}(\mathbf{w}, \mathbf{v})) + \\ &+ (h_{-}(\mathbf{v}, \mathbf{w}) - h_{-}(\mathbf{w}, \mathbf{v})) = 2\,h_{-}(\mathbf{v}, \mathbf{w}), \end{aligned}$$

Значит, $h_{+} = f_{+}$ и $h_{-} = f_{-}$, в силу чего разложение (1.4) совпадает с разложением (1.3). Теорема доказана. $\square$

Определение 1.4. Числовая функция $y = g(\mathbf{v})$ с одним векторным аргументом $\mathbf{v} \in V$ называется *квадратичной формой* на пространстве V, если $g(\mathbf{v}) = f(\mathbf{v}, \mathbf{v})$ для некоторой билинейной формы $f(\mathbf{v}, \mathbf{w})$.

Если $g(\mathbf{v}) = f(\mathbf{v}, \mathbf{v})$ то говорят, что квадратичная форма g *порождается* билинейной формой f. Для кососимметричной билинейной формы имеем: $f_{-}(\mathbf{v}, \mathbf{v}) = -f_{-}(\mathbf{v}, \mathbf{v})$. Отсюда $f_{-}(\mathbf{v}, \mathbf{v}) = 0$. Поэтому из (1.3) выводим

$$g(\mathbf{v}) = f(\mathbf{v}, \mathbf{v}) = f_{+}(\mathbf{v}, \mathbf{v}). \tag{1.5}$$

Одна и та же квадратичная форма может порождаться несколькими билинейными формами. Соотношение (1.5) показывает, что всякая квадратичная форма может быть порождена симметричной билинейной формой.

Теорема 1.2. *Для всякой квадратичной формы $g(\mathbf{v})$ существует единственная симметричная билинейная форма $f(\mathbf{v}, \mathbf{w})$, которая порождает $g(\mathbf{v})$.*

Док-во. Существование симметричной билинейной формы $f(\mathbf{v}, \mathbf{w})$, порождающей форму $g(\mathbf{v})$, вытекает из (1.5). Докажем единственность такой формы. Из $g(\mathbf{v}) = f(\mathbf{v}, \mathbf{v})$ и из симметричности билинейной формы f выводим

$$\begin{aligned} g(\mathbf{v} + \mathbf{w}) = f(\mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w}) &= f(\mathbf{v}, \mathbf{v}) + f(\mathbf{v}, \mathbf{w}) + \\ + f(\mathbf{w}, \mathbf{v}) + f(\mathbf{w}, \mathbf{w}) &= f(\mathbf{v}, \mathbf{v}) + 2\,f(\mathbf{v}, \mathbf{w}) + f(\mathbf{w}, \mathbf{w}). \end{aligned}$$

Но $f(\mathbf{v}, \mathbf{v})$ и $f(\mathbf{w}, \mathbf{w})$ в правой части этого соотношения можно заменить на $g(\mathbf{v})$ и $g(\mathbf{w})$. Отсюда получаем следующую формулу:

$$f(\mathbf{v}, \mathbf{w}) = \frac{g(\mathbf{v} + \mathbf{w}) - g(\mathbf{v}) - g(\mathbf{w})}{2}. \tag{1.6}$$

Формула (1.6) показывает, что значения симметричной билинейной формы $f(\mathbf{v}, \mathbf{w})$, однозначно определяются значениями квадратичной формы $g(\mathbf{v})$. Теорема доказана. $\square$

Формула (1.6) называется *формулой восстановления.* Квадратичную форму $g(\mathbf{v})$ и порождающую ее симметричную билинейную форму обозначают одной буквой: $g(\mathbf{v}) = g(\mathbf{v}, \mathbf{v})$. Более того, говоря о том, что задана квадратичная форма $g(\mathbf{v})$, сразу без специальных оговорок считают, что задана и соответствующая ей симметричная билинейная форма $g(\mathbf{v}, \mathbf{w})$.

Пусть $f(\mathbf{v}, \mathbf{w})$ — билинейная форма в конечномерном пространстве V и пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — базис в V. Числа f_{ij}, определяемые по формуле

$$f_{ij} = f(\mathbf{e}_i, \mathbf{e}_j), \tag{1.7}$$

называются *координатами* или *компонентами* формы f в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Числа (1.7) записывают в виде матрицы

$$F = \begin{Vmatrix} f_{11} & \ldots & f_{1n} \\ \vdots & \ldots & \vdots \\ f_{n1} & \ldots & f_{nn} \end{Vmatrix}, \tag{1.8}$$

которая называется матрицей билинейной формы f. Для элемента f_{ij} в матрице (1.8) первый индекс i определяет номер строки, а второй индекс j задает номер столбца. Симметричной билинейной форме $g(\mathbf{v}, \mathbf{w})$ соответствует симметричная матрица: $g_{ij} = g_{ji}$. Говоря о матрице квадратичной формы $g(\mathbf{v})$ в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$, мы будем понимать матрицу g_{ij} соответствующей ей симметричной билинейной формы $g(\mathbf{v}, \mathbf{w})$.

Пусть $v^1, \ldots, v^n$ и $w^1, \ldots, w^n$ — координаты векторов v и w в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Тогда значения билинейной формы $f(\mathbf{v}, \mathbf{w})$ и квадратичной формы $g(\mathbf{v})$ определяются по их матрицам из следующих соотношений:

$$f(v, w) = \sum_{i=1}^{n} \sum_{j=1}^{n} f_{ij}\, v^i\, w^j, \qquad\qquad g(v) = \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij}\, v^i\, v^j. \tag{1.9}$$

В случае диагональности матрицы g_{ij} формула для $g(\mathbf{v})$ содержит только квадраты координат вектора $\mathbf{v}$:

$$g(\mathbf{v}) = g_{11}\,(v^1)^2 + \ldots + g_{nn}\,(v^n)^2. \tag{1.10}$$

Этим оправдывается название «квадратичная форма». Приведение квадратичной формы к виду (1.10) за счет подходящего выбора выбора базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в пространстве V является одной из задач, которая решается в теории квадратичных форм.

Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ — два базиса в пространстве V. Обозначим через S матрицу перехода из первого базиса во второй и положим $T = S^{-1}$. Связь между компонентами билинейной формы $f(\mathbf{v}, \mathbf{w})$ в этих базисах легко выводится из (1.7). Для этого достаточно подставить соотношения (5.8) из первой главы в (1.7) и воспользоваться билинейностью формы:

$$f_{ij} = f(\mathbf{e}_i, \mathbf{e}_j) = \sum_{k=1}^{n} \sum_{q=1}^{n} T_i^k \, T_j^q \, f(\tilde{\mathbf{e}}_k, \tilde{\mathbf{e}}_q) = \sum_{k=1}^{n} \sum_{q=1}^{n} T_i^k \, T_j^q \, \tilde{f}_{kq}.$$

Аналогичным образом выводится формула, выражающая $\tilde{f}_{kq}$ через f_{ij}. Выпишем обе эти формулы:

$$f_{ij} = \sum_{k=1}^{n} \sum_{q=1}^{n} T_i^k \, T_j^q \, \tilde{f}_{kq}, \qquad \tilde{f}_{kq} = \sum_{i=1}^{n} \sum_{j=1}^{n} S_k^i \, S_q^j \, f_{ij}. \tag{1.11}$$

В матричной форме эти соотношения имеют следующий вид

$$F = T^{\mathrm{tr}} \, \tilde{F} \, T, \qquad \tilde{F} = S^{\mathrm{tr}} \, F \, S, \tag{1.12}$$

где через S^{tr} и T^{tr} обозначены матрицы, которые получаются из S и T при помощи транспонирования.

§ 2. Ортогональные дополнения относительно квадратичной формы.

Определение 2.1. Два вектора $\mathbf{v}$ и $\mathbf{w}$ в линейном векторном пространстве V называются *ортогональными друг другу относительно квадратичной формы* g, если $g(\mathbf{v}, \mathbf{w}) = 0$.

Определение 2.2. Пусть S — некоторое подмножество в линейном векторном пространстве V. *Ортогональным дополнением* множества S относительно квадратичной формы $g(\mathbf{v})$ называется множество $S_\perp \subset V$, состоящее из векторов, которые ортогональны всем векторам из множества S.

Определение ортогонального дополнения $S_\perp$ можно задать формулой $S_\perp = \{\mathbf{v} \in V : \forall \mathbf{w} \, ((\mathbf{w} \in S) \Rightarrow (g(\mathbf{v}, \mathbf{w}) = 0))\}$. Для ортогональных дополнений определяемых квадратичной формой $g(\mathbf{v})$ имеет место теорема, полностью аналогичная теоремам 3.1 и 3.2 из третьей главы.

Теорема 2.1. *Операция построения ортогональных дополнений относительно квадратичной формы $g(\mathbf{v})$ в линейном векторном пространстве V обладает следующими свойствами:*

(1) *$S_\perp$ есть подпространство в V;*
(2) *из $S_1 \subset S_2$ вытекает $(S_2)_\perp \subset (S_1)_\perp$;*
(3) *$\langle S \rangle_\perp = S_\perp$, где $\langle S \rangle$ — линейная оболочка множества S;*
(4) $\displaystyle\left(\bigcup_{i \in I} S_i\right)_\perp = \bigcap_{i \in I} (S_i)_\perp$.

Док-во. Начнем с доказательства первого пункта. Для этого проверим два условия из определения подпространства. Пусть $\mathbf{v}_1, \mathbf{v}_2 \in S_\perp$, тогда $g(\mathbf{v}_1, \mathbf{w}) = 0$ и $g(\mathbf{v}_2, \mathbf{w}) = 0$ для всех $\mathbf{w} \in S$. Отсюда для всех $\mathbf{w} \in S$ имеем

$$g(\mathbf{v}_1 + \mathbf{v}_2, \mathbf{w}) = g(\mathbf{v}_1, \mathbf{w}) + g(\mathbf{v}_2, \mathbf{w}) = 0,$$

что для суммы векторов $\mathbf{v}_1 + \mathbf{v}_2$ означает $\mathbf{v}_1 + \mathbf{v}_2 \in S_\perp$.

Пусть теперь $\mathbf{v} \in S_\perp$. Тогда $g(\mathbf{v}, \mathbf{w}) = 0$ для всех векторов $\mathbf{w} \in S$. Отсюда для вектора $\alpha \cdot \mathbf{v}$ выводим

$$g(\alpha \cdot \mathbf{v}, \mathbf{w}) = \alpha\, g(\mathbf{v}, \mathbf{w}) = 0.$$

Значит, $\alpha \cdot \mathbf{v} \in S_\perp$. Первый пункт в теореме 2.1 доказан.

Для доказательства включения $(S_2)_\perp \subset (S_1)_\perp$ во втором пункте теоремы рассмотрим произвольный вектор $\mathbf{v}$ из $(S_2)_\perp$. Из условия $\mathbf{v} \in (S_2)_\perp$ вытекает $g(\mathbf{v}, \mathbf{w}) = 0$ для всякого $\mathbf{w} \in S_2$. Но $S_1 \subset S_2$, поэтому равенство $g(\mathbf{v}, \mathbf{w}) = 0$ имеет место для всякого $\mathbf{w} \in S_1$. Отсюда $\mathbf{v} \in (S_1)_\perp$. Значит, из $\mathbf{v} \in (S_2)_\perp$ вытекает $\mathbf{v} \in (S_1)_\perp$, что доказывает требуемое включение.

При доказательстве третьего пункта теоремы, заметим, что линейная оболочка множества S содержит в себе это множество: $S \subset \langle S \rangle$. Применив уже доказанный пункт (2), получаем включение $\langle S \rangle_\perp \subset S_\perp$.

Остается установить противоположное включение. Вспомним, что линейная оболочка $\langle S \rangle$ состоит из всевозможных конечных линейных комбинаций

$$\mathbf{w} = \alpha_1 \cdot \mathbf{w}_1 + \ldots + \alpha_r \cdot \mathbf{w}_r, \tag{2.1}$$

где $\mathbf{w}_i \in S$. Пусть $\mathbf{v} \in S_\perp$, тогда $g(\mathbf{v}, \mathbf{w}) = 0$ для всех $\mathbf{w} \in S$. В частности, это относится и к векторам $\mathbf{w}_i$ из разложения (2.1). Для них $g(\mathbf{v}, \mathbf{w}_i) = 0$. Тогда

$$g(\mathbf{v}, \mathbf{w}) = \alpha_1\, g(\mathbf{v}, \mathbf{w}_1) + \ldots + \alpha_r\, g(\mathbf{v}, \mathbf{w}_r) = 0.$$

То есть $g(\mathbf{v}, \mathbf{w}) = 0$ для всех $\mathbf{w} \in \langle S \rangle$, что доказывает противоположное включение $S_\perp \subset \langle S \rangle_\perp$ и, тем самым, завершает доказательство совпадения множеств $\langle S \rangle_\perp = S_\perp$.

Перейдем к доказательству четвертого пункта в теореме 2.1. С этой целью введем следующие обозначения:

$$S = \bigcup_{i \in I} S_i, \qquad\qquad \tilde{S} = \bigcap_{i \in I} (S_i)_\perp.$$

Пусть $\mathbf{v} \in S_\perp$. Тогда $g(\mathbf{v}, \mathbf{w}) = 0$ для всех $\mathbf{w} \in S$. Но $S_i \subset S$ для всякого $i \in I$. Поэтому $g(\mathbf{v}, \mathbf{w}) = 0$ для всех $\mathbf{w} \in S_i$ и для всех $i \in I$. Это означает, что $\mathbf{v}$ принадлежит каждому из ортогональных дополнений $(S_i)_\perp$, а значит, он принадлежит их пересечению. Это доказывает включение $S_\perp \subset \tilde{S}$.

Обратно, из принадлежности $\mathbf{v} \in (S_i)_\perp$ для всех $i \in I$ вытекает равенство нулю $g(\mathbf{v}, \mathbf{w}) = 0$ для всех $\mathbf{w} \in S_i$ и для всех $i \in I$. Значит, это равенство нулю имеет место для всех векторов $\mathbf{w}$ из объединения всех множеств S_i. Это доказывает противоположное включение $\tilde{S} \subset S_\perp$.

Полученные два включения доказывают необходимое совпадение двух множеств $S_\perp = \tilde{S}$. Теорема 2.1 доказана. $\square$

Определение 2.3. *Ядром квадратичной формы* $g(\mathbf{v})$ в пространстве V называется множество $\operatorname{Ker} g = V_\perp$, которое состоит из векторов, ортогональных всем векторам пространства V относительно формы g.

Определение 2.4. Квадратичная форма, ядро которой нетривиально $\operatorname{Ker} g \neq \{\mathbf{0}\}$, называется *вырожденной.* В противоположном случае нулевого ядра $\operatorname{Ker} g = \{\mathbf{0}\}$ форма g называется *невырожденной.*

В силу теоремы 2.1 ядро формы $g(\mathbf{v})$ есть подпространство в пространстве V. Термин «ядро» для обозначения множества $V_\perp$ выбран не случайно. С каждой квадратичной формой связано некоторое отображение, ядром которого является множество $V_\perp$.

Определение 2.5. *Ассоциированным отображением* для квадратичной формы g называется отображение $a_g\colon V \to V^*$, которое каждому вектору $\mathbf{v} \in V$ ставит в соответствие функционал $f_\mathbf{v} \in V^*$, определяемый соотношением

$$f_\mathbf{v}(\mathbf{w}) = g(\mathbf{v}, \mathbf{w}). \tag{2.2}$$

Ассоциированное отображение $a_g\colon V \to V^*$ линейно, что вытекает из билинейности формы g. Его ядро $\operatorname{Ker} a_g$ совпадает с ядром формы g. Действительно, условие $\mathbf{v} \in \operatorname{Ker} a_g$ означает тождественное зануление функционала $f_\mathbf{v}$ в (2.2). Значит, вектор $\mathbf{v}$ ортогонален всем векторам $\mathbf{w} \in V$.

Ассоциированное отображение a_g устанавливает связь между ортогональным дополнением $S_\perp$ относительно формы g и ортогональным дополнением в сопряженном пространстве $S^\perp$, которое мы рассматривали в третьей главе.

Теорема 2.2. *Для всякого множества $S \subset V$ и для всякой квадратичной формы $g(\mathbf{v})$ в пространстве V, множество $S_\perp$ является полным прообразом множества $S^\perp$ относительно ассоциированного отображения, то есть имеет место равенство $S_\perp = a_g^{-1}(S^\perp)$.*

Док-во. Условие $\mathbf{v} \in S_\perp$ означает, что $g(\mathbf{v}, \mathbf{w}) = 0$ для всех векторов $\mathbf{w} \in S$. Но это условие можно записать так

$$g(\mathbf{v}, \mathbf{w}) = f_\mathbf{v}(\mathbf{w}) = a_g(\mathbf{v})(\mathbf{w}) = \langle a_g(\mathbf{v}) \,|\, \mathbf{w}\rangle = 0.$$

Значит, условие $v \in S_\perp$ эквивалентно $a_g(v) \in S^\perp$. Это доказывает необходимое равенство $S_\perp = a_g^{-1}(S^\perp)$. $\square$

Согласно определению 2.3 вектора из ядра $\operatorname{Ker} g$ ортогональны всем векторам пространства V относительно формы g. Поэтому $(\operatorname{Ker} g)_\perp = V$. Применим к множеству $S = \operatorname{Ker} g$ результат теоремы 2.2, что дает

$$a_g^{-1}((\operatorname{Ker} g)^\perp) = (\operatorname{Ker} g)_\perp = V.$$

Полученный результат становится более понятным, если переписать его в следующем эквивалентном виде

$$\operatorname{Im} a_g = a_g(V) \subseteq (\operatorname{Ker} g)^\perp. \tag{2.3}$$

Следствие 1. *Образ ассоциированного отображения a_g содержится в ортогональном дополнении к его ядру $(\operatorname{Ker} a_g)^\perp$, то есть $\operatorname{Im} a_g \subseteq (\operatorname{Ker} a_g)^\perp$.*

Это следствие из теоремы 2.2 вытекает из формулы (2.3), если учесть соотношение $\operatorname{Ker} g = \operatorname{Ker} a_g$. Для квадратичной формы g в конечномерном пространстве V оно может быть усилено.

Следствие 2. *Для квадратичной формы g в конечномерном линейном векторном пространстве V образ ассоциированного отображения a_g совпадает с ортогональным дополнением к его ядру:*

$$\operatorname{Im} a_g = (\operatorname{Ker} a_g)^{\perp}. \tag{2.4}$$

Док-во. Пользуясь теоремой 9.4 из первой главы вычислим размерность образа ассоциированного отображения:

$$\dim(\operatorname{Im} a_g) = \dim V - \dim(\operatorname{Ker} a_g).$$

Размерность ортогонального дополнения в сопряженном пространстве определяется теоремой 3.4 из третьей главы:

$$\dim(\operatorname{Ker} a_g)^{\perp} = \dim V - \dim(\operatorname{Ker} a_g).$$

Размерности этих пространств совпадают. Поэтому в силу следствия 1 и пункта (3) теоремы 4.5 из первой главы получаем соотношение (2.4). □

Теорема 2.3. *Пусть $U \subsetneq V$ — некоторое подпространство в векторном конечномерном пространстве V, содержащее в себе ядро квадратичной формы g, и пусть $\mathbf{v} \notin U$. Тогда найдется вектор $\mathbf{w} \in V$, такой, что $g(\mathbf{w}, \mathbf{v}) \neq 0$ и $g(\mathbf{w}, \mathbf{u}) = 0$ для всех векторов $\mathbf{u} \in U$.*

Док-во. Данная теорема является аналогом теоремы 1.2 из третьей главы и ее доказательство существенно опирается на эту теорему. Применив теорему 1.2 из главы III, получаем существование линейного функционала $f \in V^*$, такого, что $f(\mathbf{v}) \neq 0$ и $f(\mathbf{u}) = \langle f \,|\, \mathbf{u} \rangle = 0$ для всех $\mathbf{u} \in U$. В силу последнего условия, такой функционал принадлежит ортогональному дополнению $U^{\perp}$. Из включения $\operatorname{Ker} g \subset U$, пользуясь вторым пунктом в теореме 3.1 из третьей главы, получаем $U^{\perp} \subset (\operatorname{Ker} g)^{\perp}$. Значит, $f \in (\operatorname{Ker} g)^{\perp}$.

Теперь применим следствие 2 из теоремы 2.2. Из него получаем соотношение $(\operatorname{Ker} g)^{\perp} = \operatorname{Im} a_g$, следовательно, $f \in \operatorname{Im} a_g$. Поэтому существует вектор $\mathbf{w} \in V$, который ассоциированным отображением a_g отображается в функционал f, то есть мы имеем $f = a_g(\mathbf{w})$. Тогда

$$\begin{aligned} &g(\mathbf{w}, \mathbf{v}) = a_g(\mathbf{w})(\mathbf{v}) = f(\mathbf{v}) \neq 0, \\ &g(\mathbf{w}, \mathbf{u}) = a_g(\mathbf{w})(\mathbf{u}) = f(\mathbf{u}) = 0 \ \text{ для всех } \ \mathbf{u} \in U. \end{aligned}$$

Полученные соотношения завершают доказательство теоремы 2.3. □

Теорема 2.4. *Пусть V — конечномерное линейное векторное пространство и пусть U и W — два его подпространства, содержащие ядро квадратичной формы g. Тогда условия $W = U_{\perp}$ и $U = W_{\perp}$ эквивалентны.*

Док-во. Теорема 2.4 является аналогом теоремы 3.3 из третьей главы. Доказательства этих теорем также очень похожи.

Пусть выполнено условие $W = U_{\perp}$. Тогда для всякого вектора $\mathbf{w} \in W$ и всякого вектора $\mathbf{u} \in U$ имеет место соотношение ортогональности $g(\mathbf{w}, \mathbf{u}) = 0$. Множество $W_{\perp}$ образовано векторами, ортогональными всем векторам из W, поэтому имеет место включение $U \subset W_{\perp}$.

Допустим, что совпадения нет: $U \neq W_\perp$. Тогда существует вектор $\mathbf{v}_0$, такой, что $\mathbf{v}_0 \in W_\perp$ и $\mathbf{v}_0 \notin U$. Это дает возможность применить теорему 2.3, которая утверждает, что существует вектор $\mathbf{v}$, такой, что $g(\mathbf{v}, \mathbf{v}_0) \neq 0$ и $g(\mathbf{v}, \mathbf{u}) = 0$ для всех $\mathbf{u} \in U$. Последнее условие означает $\mathbf{v} \in U_\perp = W$. Тогда другое условие $g(\mathbf{v}, \mathbf{v}_0) \neq 0$ противоречит первоначальному выбору $\mathbf{v}_0 \in W_\perp$. Полученное противоречие показывает, что допущение $U \neq W_\perp$ неверно и имеет место совпадение $U = W_\perp$. Значит, из $W = U_\perp$ вытекает $U = W_\perp$. Поскольку пространства U и W в условии теоремы совершенно равноправны, то из $U = W_\perp$ также вытекает $W = U_\perp$. Эти два условия эквивалентны. $\square$

Утверждение теоремы 2.3 можно переформулировать так: для подпространства $U \subset V$ в конечномерном пространстве V условие $\operatorname{Ker} g \subset U$ означает, что двукратное ортогональное дополнение относительно формы g совпадает с ним самим $(U_\perp)_\perp = U$. Для произвольного подмножества $S \subset V$ в конечномерном пространстве V можно вывести соотношение

$$(S_\perp)_\perp = \langle S \rangle + \operatorname{Ker} g. \tag{2.5}$$

Докажем соотношение (2.5). Заметим, что вектора из ядра $\operatorname{Ker} g$ ортогональны всем векторам пространства V, поэтому, добавив вектора из $\operatorname{Ker} g$ к множеству S, мы не изменим ортогонального дополнения этого множества. Запишем это обстоятельство следующим образом:

$$S_\perp = (S \cup \operatorname{Ker} g)_\perp.$$

Теперь применим третий пункт из теоремы (2.1). Это дает

$$S_\perp = (S \cup \operatorname{Ker} g)_\perp = \langle S \cup \operatorname{Ker} g \rangle_\perp = (\langle S \rangle + \operatorname{Ker} g)_\perp.$$

Но подпространство $U = \langle S \rangle + \operatorname{Ker} g$ содержит в себе ядро формы g. Поэтому $(U_\perp)_\perp = U$, что доказывает соотношение (2.5):

$$(S_\perp)_\perp = ((\langle S \rangle + \operatorname{Ker} g)_\perp)_\perp = \langle S \rangle + \operatorname{Ker} g.$$

Теорема 2.5. *В случае конечномерного пространства V для любого подпространства $U \subset V$ имеет место равенство*

$$\dim U + \dim U_\perp = \dim V + \dim(\operatorname{Ker} g \cap U), \tag{2.6}$$

где $U_\perp$ — ортогональное дополнение подпространства U относительно формы g с ядром $\operatorname{Ker} g$.

Док-во. Вектора из ядра $\operatorname{Ker} g$ ортогональны всем векторам пространства V, поэтому их добавление к U не изменяет пространства $U_\perp$. Обозначим $W = U + \operatorname{Ker} g$. Тогда в силу сказанного $U_\perp = W_\perp$. Для размерности подпространства W в силу теоремы 6.4 из первой главы имеем:

$$\dim W = \dim U + \dim(\operatorname{Ker} g) - \dim(\operatorname{Ker} g \cap U). \tag{2.7}$$

Теперь применим результат теоремы 2.2 к $S = W$. Это дает $W_\perp = a_g^{-1}(W^\perp)$. Заметим, что $\operatorname{Ker} g \subset W$, этим подпространство W отличается от исходного подпространства U. Применим пункт (2) теоремы 3.1 к включению $\operatorname{Ker} g \subset W$ и учтем при этом следствие 2 из теоремы 2.2. Это дает

$$W^{\perp} \subset (\operatorname{Ker} g)^{\perp} = \operatorname{Im} a_g.$$

Полученное включение $W^{\perp} \subset \operatorname{Im} a_g$ означает, что прообраз каждого элемента $f \in W^{\perp}$ при отображении a_g непуст, а равенство $W_{\perp} = a_g^{-1}(W^{\perp})$ показывает, что такой прообраз содержится в $W_{\perp}$. Поэтому из $W_{\perp} = a_g^{-1}(W^{\perp})$ вытекает равенство $a_g(W_{\perp}) = W^{\perp}$.

Рассмотрим сужение ассоциированного отображения a_g на подпространство $W_{\perp}$. Обозначим такое сужение через a:

$$a\colon\, W_{\perp} \to V^*. \tag{2.8}$$

Ядро отображения (2.8) совпадает с ядром ассоциированного отображения a_g, поскольку $\operatorname{Ker} a_g = \operatorname{Ker} g \subset W_{\perp}$. Для его образа имеем:

$$\operatorname{Im} a = a_g(W_{\perp}) = W^{\perp}.$$

Это дает возможность использовать теорему о «сумме размерностей ядра и образа» (см. теорему 9.4 из первой главы):

$$\dim(\operatorname{Ker} g) + \dim W^{\perp} = \dim W_{\perp} \tag{2.9}$$

Для определения размерности $W^{\perp}$ применим соотношение

$$\dim W + \dim W^{\perp} = \dim V, \tag{2.10}$$

которое вытекает из теоремы 3.4 из третьей главы. Теперь сложим равенства (2.7) с (2.9) и вычтем из полученного результата (2.10). С учетом совпадения $W_{\perp} = U_{\perp}$ это дает в точности соотношение (2.6), которое нам как раз и требовалось доказать. $\square$

Аналогом соотношений (3.4) из третьей главы в данном случае являются равенства $\{\mathbf{0}\}_{\perp} = V$ и $V_{\perp} = \operatorname{Ker} g$.

Теорема 2.6. *В случае конечномерного пространства V и квадратичной формы g в нем для любого семейства подпространств в V, каждое из которых содержит в себе ядро $\operatorname{Ker} g$, выполнены соотношения*

$$\left(\sum_{i\in I} U_i\right)_{\perp} = \bigcap_{i\in I} (U_i)_{\perp}, \qquad \left(\bigcap_{i\in I} U_i\right)_{\perp} = \sum_{i\in I} (U_i)_{\perp}. \tag{2.11}$$

Док-во. Для доказательства первого из двух соотношений (2.11) условие $\operatorname{Ker} g \subset U_i$ несущественно. Оно выводится из пунктов (3) и (4) теоремы 2.1, если учесть, что сумма подпространств есть линейная оболочка объединения этих подпространств.

Второе соотношение (2.11) выводится из первого. Из условия $\operatorname{Ker} g \in U_i$ выводим $((U_i)_{\perp})_{\perp} = U_i$, что является следствием теоремы 2.4. Обозначим $(U_i)_{\perp} = V_i$ и используем первое соотношение (2.11) для семейства подпространств V_i. В результате этого получаем

$$\left(\sum_{i\in I}(U_i)_\perp\right)_\perp=\left(\sum_{i\in I}V_i\right)_\perp=\bigcap_{i\in I}(V_i)_\perp=\bigcap_{i\in I}U_i.$$

Теперь остается перейти к ортогональным дополнениям левой и правой частей в полученном соотношении и вновь применить теорему 2.4. Это дает требуемое второе соотношение в (2.11). Теорема доказана. $\square$

§ 3. Приведение квадратичной формы к каноническому виду.

Определение 3.1. Подпространство U в линейном векторном пространстве V называется *регулярным относительно квадратичной формы* g, если имеет место включение $U\cap U_\perp\subseteq\operatorname{Ker}g$.

Теорема 3.1. *Пусть U — подпространство конечномерного пространства V, регулярное относительно квадратичной формы g. Тогда $U+U_\perp=V$.*

Док-во. Положим $W=U+U_\perp$ и вычислим размерность подпространства W, применив теорему 6.4 из первой главы:

$$\dim W=\dim U+\dim U_\perp-\dim(U\cap U_\perp).$$

Вектора из ядра ортогональны любому вектору v из V, поэтому $\operatorname{Ker}g\subseteq U_\perp$. Кроме того, из регулярности U относительно формы g имеем $U\cap U_\perp\subseteq\operatorname{Ker}g$. Поэтому для пересечения подпространств $U\cap U_\perp$ имеем

$$U\cap U_\perp=(U\cap U_\perp)\cap\operatorname{Ker}g=U\cap(U_\perp\cap\operatorname{Ker}g)=U\cap\operatorname{Ker}g.$$

В силу выведенного равенства $U\cap U_\perp=U\cap\operatorname{Ker}g$ полученное выше соотношение для размерности W можно переписать так:

$$\dim W=\dim U+\dim U_\perp-\dim(U\cap\operatorname{Ker}g). \tag{3.1}$$

Сравним (3.1) с формулой (2.6) из теоремы 2.5. Такое сравнение приводит нас к равенству $\dim W=\dim V$. Теперь, применив пункт (3) в теореме 4.5 из первой главы, получаем $W=V$. $\square$

Теорема 3.2. *Пусть U — подпространство конечномерного пространства V, регулярное относительно квадратичной формы g. Тогда, если $U_\perp\neq\operatorname{Ker}g$, то в подпространстве $U_\perp$ найдется вектор $\mathbf{v}$, для которого $g(\mathbf{v})\neq 0$.*

Док-во. Допустим противоположное. Если такого вектора $\mathbf{v}$ нет, то числовая функция $g(\mathbf{v})$ равна нулю тождественно на пространстве $U_\perp$. В силу формулы восстановления (1.6) равна нулю и функция $g(\mathbf{v},\mathbf{w})$ для всех векторов $\mathbf{v},\mathbf{w}\in U_\perp$.

Применим теорему 3.1 и разложим некоторый произвольный вектор $\mathbf{x}\in V$ в сумму $\mathbf{x}=\mathbf{u}+\mathbf{w}$, где $\mathbf{u}\in U$ и $\mathbf{w}\in U_\perp$. Тогда для произвольного вектора $\mathbf{v}\in U_\perp$ выводим следующую цепочку равенств:

$$g(\mathbf{v},\mathbf{x})=g(\mathbf{v},\mathbf{u}+\mathbf{w})=g(\mathbf{v},\mathbf{u})+g(\mathbf{v},\mathbf{w})=0+0=0.$$

Первое слагаемое $g(\mathbf{v}, \mathbf{u})$ зануляется по причине ортогональности U и $U_{\perp}$, а второе — в силу сделанного выше допущения. Из $g(\mathbf{v}, \mathbf{x}) = 0$ для всех $x \in V$ получаем $\mathbf{v} \in \operatorname{Ker} g$. Ввиду произвольности $\mathbf{v} \in U_{\perp}$ это означает $U_{\perp} \subseteq \operatorname{Ker} g$. Противоположное включение $\operatorname{Ker} g \subseteq U_{\perp}$ имеет место всегда. Поэтому $U_{\perp} = \operatorname{Ker} g$, что противоречит условию теоремы. Значит, сделанное в начале допущение было ошибочным. Это доказывает существование вектора $\mathbf{v} \in U_{\perp}$, для которого $g(\mathbf{v}) \neq 0$. □

Теорема 3.3. *Для всякой квадратичной формы g в конечномерном пространстве V существует базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$, в котором матрица этой квадратичной формы диагональна.*

Док-во. Случай $g = 0$ исключим сразу. Здесь все очевидно: тождественно нулевая квадратичная форма имеет нулевую матрицу в любом базисе, а нулевая матрица $n \times n$ диагональна.

Пусть $g \not\equiv 0$. Доказательство теоремы проведем индукцией по размерности пространства $\dim V = n$. В случае $n = 1$ утверждение теоремы тривиально: всякая матрица размера 1×1 диагональна.

Предположим, что теорема справедлива для всех квадратичных форм в пространствах размерности меньшей n. После этого рассмотрим подпространство $U = \operatorname{Ker} g$. Оно регулярно относительно формы g и $U_{\perp} = V$, что позволяет применить теорему 3.2. Согласно этой теореме существует вектор $\mathbf{v}_0 \notin U$, для которого $g(\mathbf{v}_0) \neq 0$. Рассмотрим подпространство W, получаемое из $U = \operatorname{Ker} g$ добавлением вектора $\mathbf{v}_0$:

$$W = \operatorname{Ker} g + \langle \mathbf{v}_0 \rangle = U \oplus \langle \mathbf{v}_0 \rangle. \tag{3.2}$$

Построенное подпространство W определяет собой следующую альтернативу: $W = V$ или же $W \neq V$.

В случае $W = V$ выберем базис $\mathbf{e}_1, \ldots, \mathbf{e}_s$ в ядре квадратичной формы $\operatorname{Ker} g$ и дополним его одним вектором $\mathbf{e}_{s+1} = \mathbf{v}_0$. Получится базис в V. Матрица квадратичной формы в таком базисе почти целиком состоит из нулей: при $i = 1, \ldots, s$ имеем $g_{ij} = g_{ji} = g(\mathbf{e}_i, \mathbf{e}_j) = 0$, ибо $\mathbf{e}_i \in \operatorname{Ker} g$. Единственный ненулевой элемент — это g_{ii}, где $i = s + 1$. Он расположен на диагонали и равен $g(\mathbf{e}_{s+1}, \mathbf{e}_{s+1}) = g(\mathbf{v}_0) \neq 0$.

В случае $W \neq V$ рассмотрим пересечение $W \cap W_{\perp}$. Пусть $\mathbf{w} \in W \cap W_{\perp}$. Тогда из (3.2) получаем $\mathbf{w} = \alpha \cdot \mathbf{v}_0 + \mathbf{u}$, где $\mathbf{u} \in \operatorname{Ker} g$. Будучи вектором из W и из $W_{\perp}$, вектор $\mathbf{w}$ должен быть ортогонален самому себе относительно g:

$$\begin{aligned} g(\mathbf{w}, \mathbf{w}) &= g(\alpha \cdot \mathbf{v}_0 + \mathbf{u}, \alpha \cdot \mathbf{v}_0 + \mathbf{u}) = \\ &= \alpha^2 g(\mathbf{v}_0, \mathbf{v}_0) + 2\alpha\, g(\mathbf{v}_0, \mathbf{u}) + g(\mathbf{u}, \mathbf{u}) = 0. \end{aligned} \tag{3.3}$$

Но $\mathbf{u} \in \operatorname{Ker} g$, поэтому $g(\mathbf{v}_0, \mathbf{u}) = 0$ и $g(\mathbf{u}, \mathbf{u}) = 0$. Но, в то же самое время, $g(\mathbf{v}_0, \mathbf{v}_0) = g(\mathbf{v}_0) \neq 0$. Поэтому из (3.3) получаем $\alpha = 0$. Это значит, что $\mathbf{w} = \mathbf{u} \in \operatorname{Ker} g$. Тем самым, мы установили включение $W \cap W_{\perp} \subseteq \operatorname{Ker} g$, которое означает регулярность подпространства W относительно формы g.

Применим теорему 3.1. Она дает разложение $V = W + W_{\perp}$. Заметим, что $\mathbf{v}_0 \in W$, но $\mathbf{v}_0 \notin W \cap W_{\perp}$. Это вытекает из условия $g(\mathbf{v}_0, \mathbf{v}_0) \neq 0$. Отсюда получаем $\mathbf{v}_0 \notin W_{\perp}$ и $W_{\perp} \neq V$.

Мы получили $W_\perp \neq V$. Значит, размерность $W_\perp$ меньше n. Формула (2.6) определяет точное значение размерности этого подпространства

$$\dim W_\perp = \dim V + \dim(U \cap \operatorname{Ker} g) - \dim U = n - 1.$$

Рассмотрев сужение квадратичной формы g на подпространство $W_\perp$, мы можем применить предположение индукции. Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}$ — базис подпространства $W_\perp$, в котором матрица сужения формы g на $W_\perp$ диагональна:

$$g_{ij} = g_{ji} = g(\mathbf{e}_i, \mathbf{e}_j) = 0 \quad \text{при} \quad i < j \leq n - 1. \tag{3.4}$$

Дополним этот базис ровно одним вектором $\mathbf{e}_n = \mathbf{v}_0$. В силу $\mathbf{v}_0 \notin W_\perp$ полученная система векторов $\mathbf{e}_1, \ldots, \mathbf{e}_n$ линейно независима и образует базис в V. Рассмотрим матрицу квадратичной формы g в построенном базисе. Для элементов такой матрицы имеются соотношения

$$g_{in} = g_{ni} = g(\mathbf{e}_i, \mathbf{e}_n) = 0 \quad \text{при} \quad i < n. \tag{3.5}$$

Они вытекают из ортогональности векторов $\mathbf{e}_i$ и $\mathbf{e}_n$ в (3.5). Действительно, $\mathbf{e}_n \in W$ и $\mathbf{e}_i \in W_\perp$. Соотношения (3.4) и (3.5), как раз, и означают диагональность матрицы квадратичной формы g в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Теорема доказана. $\square$

Пусть g — квадратичная форма в конечномерном пространстве V и пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — базис, в котором матрица формы g диагональна. Тогда значение формы g на векторе v можно вычислять по формуле (1.10). Часть диагональных элементов $g_{11}, \ldots, g_{nn}$ в матрице квадратичной формы может быть равна нулю. Пусть число таких элементов равно s. Базисные вектора $\mathbf{e}_1, \ldots, \mathbf{e}_n$ можно пронумеровать так, что

$$g_{11} = \ldots = g_{ss} = 0. \tag{3.6}$$

Первые s векторов базиса, которые соответствуют элементам (3.6), принадлежат ядру формы $\operatorname{Ker} g$. Действительно, если $\mathbf{w} = \mathbf{e}_i$ для $i = 1, \ldots, s$, то $g(\mathbf{v}, \mathbf{w}) = 0$ для всех векторов $\mathbf{v} \in V$. Это может быть выведено, например, из формул (1.9).

Наоборот, пусть $\mathbf{w} \in \operatorname{Ker} g$. Тогда для произвольного вектора $\mathbf{v} \in V$ имеет место следующее соотношение:

$$g(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij}\, v^i\, w^j = \sum_{i=s+1}^{n} g_{ii}\, v^i\, w^i = 0.$$

Ввиду произвольности $\mathbf{v} \in V$ это соотношение должно быть выполнено тождественно при произвольных значениях координат $v^{s+1}, \ldots, v^n$. Но $g_{ii} \neq 0$ при $i \geq s + 1$, поэтому $w^{s+1} = \ldots = w^n = 0$. Отсюда

$$\mathbf{w} = w^1 \cdot \mathbf{e}_1 + \ldots + w^s \cdot \mathbf{e}_s.$$

Вывод: всякий вектор w из ядра $\operatorname{Ker} g$ может быть разложен в линейную комбинацию первых s базисных векторов. Следовательно, вектора $e_1, \ldots, e_s$ составляют базис в ядре $\operatorname{Ker} g$. Приведенные здесь рассуждения доказывают следующее утверждение, которое мы сформулируем в виде теоремы.

Теорема 3.4. *Число нулей на диагонали матрицы квадратичной формы g, приведенной к диагональному виду, является геометрическим инвариантом формы g. Оно не зависит от способа приведения этой матрицы к диагональному виду и совпадает с размерностью ее ядра:* $s = \dim(\operatorname{Ker} g)$.

Определение 3.2. Число $s = \dim(\operatorname{Ker} g)$ называется *нулевым индексом инерции* квадратичной формы g.

Пусть g — квадратичная форма в линейном векторном пространстве над полем комплексных чисел $\mathbb{C}$, матрица которой диагональна в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Пусть s — нулевой индекс инерции формы g. Без ограничения общности мы можем считать, что первые s базисных векторов $\mathbf{e}_1, \ldots, \mathbf{e}_s$ образуют базис в ядре $\operatorname{Ker} g$. Введем числа $\gamma_1, \ldots, \gamma_n$, определив их по формуле:

$$\gamma_i = \begin{cases} 1 & \text{при } \ i \leqslant s, \\ \sqrt{g_{ii}} & \text{при } \ i > s. \end{cases} \tag{3.7}$$

Напомним, что из всякого комплексного числа можно извлечь квадратный корень, который также будет комплексным числом. Комплексные числа (3.7) отличны от нуля. Используем их для построения нового базиса:

$$\tilde{\mathbf{e}}_i = (\gamma_i)^{-1} \cdot \mathbf{e}_i, \ i = 1, \ldots, n. \tag{3.8}$$

Матрица квадратичной формы g остается диагональной и в новом базисе:

$$\tilde{g}_{ij} = g(\tilde{\mathbf{e}}_i, \tilde{\mathbf{e}}_j) = (\gamma_i\,\gamma_j)^{-1}\,g_{ij} = 0 \ \text{ при } \ i \neq j.$$

Для диагональных элементов матрицы квадратичной формы g прямым вычислением выводится формула:

$$\tilde{g}_{ii} = g(\tilde{\mathbf{e}}_i, \tilde{\mathbf{e}}_i) = (\gamma_i)^{-2}\,g_{ii} = \begin{cases} 0 & \text{при } \ i \leqslant s, \\ 1 & \text{при } \ i > s. \end{cases}$$

Матрица квадратичной формы g в таком базисе имеет вид, который принято называть *каноническим* для матрицы квадратичной формы над полем $\mathbb{C}$:

$$\mathcal{G} = \left\| \begin{matrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{matrix} \right\| \begin{matrix} \left.\vphantom{\begin{matrix}0\\ \ddots\\ 0\end{matrix}}\right\} s \\ \left.\vphantom{\begin{matrix}1\\ \ddots\\ 1\end{matrix}}\right\} n-s \end{matrix} \tag{3.9}$$

Матрица $\mathcal{G}$ в (3.8) диагональна, ее диагональ состоит из нулей и единиц.

В случае векторного пространства над полем вещественных чисел $\mathbb{R}$ канонический вид матрицы квадратичной формы отличается от (3.9). Пусть вновь $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — базис, в котором матрица формы g диагональна. Диагональные элементы матрицы делятся теперь на три группы: нулевые, положительные и отрицательные. Пусть s — число нулевых элементов и пусть r — число положительных элементов. Тогда остальные $n - s - r$ элементов отрицательны. Базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$, можно считать пронумерованным так, что $g_{ii} = 0$ при

$i = 1, \ldots, s$ и $g_{ii} > 0$ при $i = s+1, \ldots, s+r$. В поле вещественных чисел $\mathbb{R}$ мы можем извлечь квадратный корень только из неотрицательного числа, поэтому числа $\gamma_1, \ldots, \gamma_n$ приходится определять теперь несколько иначе, чем это было сделано выше в формуле (3.7):

$$\gamma_i = \begin{cases} 1 & \text{при} \quad i \leqslant s, \\ \sqrt{|g_{ii}|} & \text{при} \quad i > s. \end{cases} \tag{3.10}$$

По числам (3.10) построим новый базис $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$, используя для этого соотношения (3.9). В новом базисе $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ матрица квадратичной формы остается по-прежнему диагональной:

$$\mathcal{G} = \left\| \begin{array}{ccccccccc} 0 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 0 & & & & & & \\ & & & 1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & 1 & & & \\ & & & & & & -1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & -1 \end{array} \right\| \begin{array}{l} \left.\vphantom{\begin{array}{c}0\\ \ddots\\ 0\end{array}}\right\} s \\ \left.\vphantom{\begin{array}{c}1\\ \ddots\\ 1\end{array}}\right\} r_p \\ \left.\vphantom{\begin{array}{c}-1\\ \ddots\\ -1\end{array}}\right\} r_n \end{array} \tag{3.11}$$

Определение 3.3. Формула (3.11) определяет *канонический вид* матрицы квадратичной формы g в пространстве над полем действительных чисел $\mathbb{R}$. Числа r_p и r_n, определяющие количество положительных и отрицательных элементов на диагонали, называются *положительным индексом инерции* и *отрицательным индексом инерции* формы g соответственно.

Теорема 3.5. *Отрицательный и положительный индексы инерции r_p и r_n квадратичной формы g в пространстве над полем $\mathbb{R}$ являются геометрическими инвариантами этой формы и не зависят от способа приведения ее матрицы к диагональному виду* (3.11).

Док-во. Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — базис в пространстве V, в котором форма g имеет матрицу канонического вида (3.11). Рассмотрим следующие подпространства, связанные с каноническим базисом:

$$U_+ = \langle \mathbf{e}_1, \ldots, \mathbf{e}_{s+r_p} \rangle, \qquad\qquad U_- = \langle \mathbf{e}_{s+r_p+1}, \ldots, \mathbf{e}_n \rangle. \tag{3.12}$$

Подпространства U_+ и U_- имеют тривиальное пересечение, причем выполняются соотношения $\dim U_+ = s + r_p$, $\dim U_- = r_n$ и $U_+ \oplus U_- = V$.

Рассмотрим вектор $\mathbf{v} \in U_+$. Значение квадратичной формы g на этом векторе определяется формулой (1.10) и видом матрицы (3.11):

$$g(\mathbf{v}) = \sum_{i=s+1}^{s+r_p} (v^i)^2.$$

Сумма квадратов в правой части этого равенства есть величина неотрицательная, то есть $g(\mathbf{v}) \geqslant 0$ для всех $\mathbf{v} \in U_+$.

Теперь рассмотрим вектор $\mathbf{v}$ из подпространства U_-. Для такого вектора формула (1.10) записывается в виде

$$g(\mathbf{v}) = \sum_{i=s+r_p+1}^{n} (-(v^i)^2),$$

причем, если $\mathbf{v} \neq 0$, то, по меньшей мере, одно из слагаемых в этой сумме отлично от нуля. Значит, $g(\mathbf{v}) < 0$ для всех ненулевых векторов из U_-.

Пусть $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ — некоторый другой базис, в котором матрица квадратичной формы g также имеет канонический вид. Обозначим через $\tilde{s}$, $\tilde{r}_p$ и $\tilde{r}_n$ индексы инерции для g в этом базисе. Для нулевых индексов инерции имеет место совпадение $s = \tilde{s}$, ибо $s = \dim(\operatorname{Ker} g)$ и $\tilde{s} = \dim(\operatorname{Ker} g)$.

Докажем совпадение положительных и отрицательных индексов инерции. Для этого рассмотрим подпространства $\tilde{U}_+$ и $\tilde{U}_-$, задав их соотношениями вида (3.12) в базисе $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$. Если допустить, что $r_p \neq \tilde{r}_p$, то либо $r_p > \tilde{r}_p$, либо $r_p < \tilde{r}_p$. Пусть, для определенности, $r_p > \tilde{r}_p$. Тогда

$$\dim U_+ = s + r_p, \qquad\qquad \dim \tilde{U}_- = \tilde{r}_n = n - s - \tilde{r}_p.$$

Для суммы размерностей этих двух подпространств получаем: $\dim U_+ + \dim \tilde{U}_- = n + (r_p - \tilde{r}_p)$. Из сделанного выше допущения $r_p > \tilde{r}_p$ вытекает следующее неравенство для размерностей:

$$\dim U_+ + \dim \tilde{U}_- > \dim V \tag{3.13}$$

Кроме того, из естественного включения $U_+ + \tilde{U}_- \subseteq V$ вытекает оценка $\dim(U_+ + \tilde{U}_-) \leqslant \dim V$. Из этой оценки и неравенства (3.13) после применения теоремы 6.4 из первой главы получаем: $\dim(U_+ \cap \tilde{U}_-) > 0$. Значит, пересечение $U_+ \cap \tilde{U}_-$ отлично от нуля и содержит в себе некоторый ненулевой вектор $\mathbf{v} \in U_+ \cap \tilde{U}_-$. Из условий $\mathbf{v} \in U_+$ и $\mathbf{v} \in U_-$ вытекают следующие два прямо противоположных неравенства:

$$g(\mathbf{v}) \geqslant 0, \qquad\qquad g(\mathbf{v}) < 0.$$

Полученное противоречие показывает ошибочность допущения $r_p \neq \tilde{r}_p$. Значит, индексы инерции r_p и $\tilde{r}_p$ совпадают. Из совпадения $r_p = \tilde{r}_p$ и $s = \tilde{s}$ вытекает совпадение $r_n = \tilde{r}_n$. Теорема доказана. $\square$

Определение 3.4. Совокупность индексов инерции называется *сигнатурой* квадратичной формы. В комплексном случае $\mathbb{K} = \mathbb{C}$ сигнатура определяется двумя числами $(s, n - s)$, в вещественном случае $\mathbb{K} = \mathbb{R}$ она определяется тремя числами (s, r_p, r_n).

В случае векторного пространства над полем рациональных чисел $\mathbb{K} = \mathbb{Q}$ мы также можем диагонализовать матрицу квадратичной формы и разделить диагональные элементы на нулевые, положительные и отрицательные. Это определяет числа s, r_p и r_n. Эти числа являются геометрическими инвариантами формы g, что позволяет говорить о ее сигнатуре.

Однако, свести ненулевые элементы в диагонализованной матрице квадратичной формы только к $+1$ и к -1 в случае поля $\mathbb{K} = \mathbb{Q}$ не удается, поэтому число геометрических инвариантов здесь больше чем три. Мы не будем заниматься определением полного набора геометрических инвариантов квадратичной формы в случае $\mathbb{K} = \mathbb{Q}$ и изложением теории таких инвариантов. Это увело бы нас в сторону теории чисел к вопросам делимости, взаимной простоты, разложения на простые множители и т. д.

§ 4. Положительно определенные квадратичной формы. Критерий Сильвестра.

В данном параграфе мы рассматриваем квадратичные формы в линейных векторных пространствах над полем вещественных чисел $\mathbb{R}$. Однако, почти все результаты этого параграфа остаются справедливыми и для форм в векторных пространствах над полем рациональных чисел $\mathbb{Q}$.

Определение 4.1. Квадратичная форма g в пространстве V над полем $\mathbb{R}$ называется положительно определенной, если $g(\mathbf{v}) > 0$ для всякого ненулевого вектора $\mathbf{v} \in V$.

Теорема 4.1. *Квадратичная форма g в конечномерном пространстве V положительно определена тогда и только тогда, когда в ее сигнатуре (s, r_p, r_n) числа s и r_n равны нулю.*

Док-во. Пусть квадратичная форма g положительно определена и пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — базис, в котором матрица этой формы имеет канонический вид (3.11). Если $s \neq 0$, то для базисного вектора $\mathbf{e}_1 \neq 0$ имеем $g(\mathbf{e}_1) = g_{11} = 0$, что противоречит положительной определенности g. Если же $r_n \neq 0$, то для базисного вектора $\mathbf{e}_n \neq 0$ получаем: $g(\mathbf{e}_n) = g_{nn} = -1$, что также противоречит положительной определенности формы g. Значит, $s = r_n = 0$.

Теперь в обратную сторону. Пусть в сигнатуре формы g числа s и r_n равны нулю. Тогда в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$, где матрица формы g имеет вид (3.11), ее значение $g(\mathbf{v})$ вычисляется как сумма квадратов координат вектора $\mathbf{v}$:

$$g(\mathbf{v}) = (v^1)^2 + \ldots + (v^n)^2.$$

Это вытекает из формулы (1.10). Для ненулевого вектора $\mathbf{v}$, по меньшей мере, одна из координат отлична от нуля, поэтому $g(\mathbf{v}) > 0$. Это доказывает положительную определенность формы g. Теорема доказана. $\square$

Условие $s = \operatorname{Ker} g$, полученное в теореме 3.4, и условие $s = 0$ означают, что положительно определенная форма g в конечномерном пространстве V невырождена: $\operatorname{Ker} g = \{\mathbf{0}\}$. Этот факт имеет место и для форм в бесконечномерных пространствах.

Теорема 4.2. *Всякая положительно определенная квадратичная форма g невырождена.*

Док-во. Если допустить, что $\operatorname{Ker} g \neq \{\mathbf{0}\}$, то найдется ненулевой вектор $\mathbf{v} \in \operatorname{Ker} g$. Вектор $\mathbf{v}$ из ядра формы ортогонален всем векторам пространства V. Значит, он ортогонален и самому себе: $g(\mathbf{v}) = g(\mathbf{v}, \mathbf{v}) = 0$. Это противоречит положительной определенности формы g. Полученное противоречие, как раз, и доказывает утверждение теоремы. $\square$

Теорема 4.3. *Всякое подпространство $U \subset V$ регулярно относительно положительно определенной формы g в пространстве V.*

Док-во. В силу тривиальности ядра положительно определенной формы g, условие регулярности подпространства U относительно g эквивалентно равенству $U \cap U_{\perp} = \{\mathbf{0}\}$ (см. определение 3.1). Докажем это равенство. Пусть $\mathbf{v}$ — некоторый произвольный вектор из пересечения $U \cap U_{\perp}$. Из $v \in U_{\perp}$ вытекает, что он ортогонален всем векторам из U. Значит, он ортогонален и самому себе, ибо $\mathbf{v} \in U$. Отсюда $g(\mathbf{v}) = g(\mathbf{v}, \mathbf{v}) = 0$. В силу положительной определенности формы g равенство $g(\mathbf{v}) = 0$ имеет место только для нулевого вектора $\mathbf{v} = 0$. Следовательно, $U \cap U_{\perp} = \{\mathbf{0}\}$. Теорема доказана. $\square$

Теорема 4.4. *Для всякого подпространства $U \subset V$ и всякой положительно определенной формы g в конечномерном пространстве V имеет место разложение $V = U \oplus U_{\perp}$.*

Док-во. Разложение $V = U + U_{\perp}$ вытекает из теоремы 3.1. Остается доказать, что сумма в этом разложении прямая. Для суммы размерностей подпространств U и $U_{\perp}$ из теоремы 2.5 и из тривиальности ядра положительно определенной квадратичной формы $\operatorname{Ker} g = \{\mathbf{0}\}$ получаем

$$\dim U + \dim U_{\perp} = \dim V.$$

В силу полученного равенства для завершения доказательства теоремы остается лишь применить результат теоремы 6.3 из первой главы. $\square$

Пусть g — квадратичная форма в конечномерном пространстве V над полем вещественных чисел $\mathbb{R}$. Выберем некоторый произвольный базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в V и построим матрицу квадратичной формы g:

$$\mathcal{G} = \left\| \begin{matrix} g_{11} & \cdots & g_{1n} \\ \vdots & \ldots & \vdots \\ g_{n1} & \cdots & g_{nn} \end{matrix} \right\| \tag{4.1}$$

Зачеркнем $n-k$ последних столбцов и $n-k$ последних строк в матрице (4.1). Детерминант полученной матрицы называется *k-ым главным диагональным минором* в матрице $\mathcal{G}$ квадратичной формы g:

$$M_k = \det \begin{vmatrix} g_{11} & \cdots & g_{1k} \\ \vdots & \ldots & \vdots \\ g_{k1} & \cdots & g_{kk} \end{vmatrix} \tag{4.2}$$

При этом n-ый главный диагональный минор M_n совпадает с определителем исходной матрицы $\mathcal{G}$.

Теорема 4.5. *Для положительно определенной квадратичной формы g в конечномерном пространстве V, детерминант ее матрицы в любом базисе положителен.*

Док-во. Рассмотрим сначала базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$, в котором матрица квадратичной формы g имеет канонический вид (3.11). Согласно теореме 4.1 для положительно определенной квадратичной формы g ее матрица в каноническом базисе единична. Отсюда $\det \mathcal{G} = 1 > 0$.

Пусть теперь $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ — некоторый произвольный базис в пространстве V и пусть S — матрица перехода из базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в базис $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$. Применив формулу (1.12), получаем

$$\det \tilde{\mathcal{G}} = \det S^{\mathrm{tr}}\,(\det \mathcal{G})\,\det S = (\det S)^2.$$

В вещественном линейном векторном пространстве V, в котором определена форма g, матрицы перехода из одного базиса в другой вещественны. Поэтому $\det S \neq 0$ — вещественное число, а его квадрат $(\det S)^2$ — число положительное. Теорема доказана. □

Пусть вновь $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — некоторый произвольный базис в V, а g_{ij} — матрица положительно определенной квадратичной формы g в этом базисе. Рассмотрим следующее подпространство:

$$U_k = \langle \mathbf{e}_1, \ldots, \mathbf{e}_k \rangle.$$

Обозначим через h_k сужение формы g на подпространство U_k. Матрица формы h_k в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_k$ совпадает с диагональным блоком в матрице исходной формы g. Именно этот блок определяет главный минор M_k в формуле (4.2). Сужение положительно определенной формы g на любое подпространство является положительно определенной формой. Это позволяет применить теорему 4.5 к форме h_k, что дает $M_k > 0$.

Вывод: положительность диагональных главных миноров (4.2) есть необходимое условие положительной определенности квадратичной формы g. Оказывается, это условие является и достаточным, что составляет содержание следующей теоремы, известной как *критерий Сильвестра*.

Теорема 4.6 (Сильвестр). *Квадратичная форма g в конечномерном пространстве V положительно определена тогда и только тогда, когда все главные диагональные миноры в ее матрице положительны.*

Док-во. Из положительной определенности формы g вытекает положительность диагональных главных миноров в ее матрице. Этот факт нами уже доказан выше. Докажем обратное утверждение. Пусть все диагональные главные миноры (4.2) в матрице квадратичной формы g положительны. Докажем положительную определенность самой квадратичной формы g. Доказательство произведем индукцией по размерности пространства $n = \dim V$.

База индукции в случае $\dim V = 1$ очевидна. Здесь матрица квадратичной формы g имеет ровно один элемент g_{11}, совпадающий с единственным главным минором: $g_{11} = M_1$. Значение $g(v)$ определяется единственной координатой вектора v по формуле: $g(v) = g_{11}\,(v^1)^2$. Поэтому из $M_1 > 0$ вытекает положительная определенность формы g.

Предположим, что доказываемое нами утверждение выполнено для всех квадратичных форм в пространствах размерности меньшей, чем $n = \dim V$. Пусть g_{ij} — матрица квадратичной формы g в некотором базисе $e_1, \ldots, e_n$. Положим $U = \langle \mathbf{e}_1, \ldots, \mathbf{e}_{n-1} \rangle$. Обозначим через h сужение квадратичной формы g на подпространство U размерности $n - 1$. Матричные элементы h_{ij} в матрице формы h, вычисленные в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}$, совпадают с соответствующими элементами в матрице исходной квадратичной формы: $h_{ij} = g_{ij}$. Поэтому миноры $M_1, \ldots , M_{n-1}$ можно вычислять по матрице h_{ij}. Из положительности этих миноров в силу предположения индукции вытекает положительная определенность квадратичной формы h.

Пусть $\tilde{\mathbf{e}}_1, \ldots , \tilde{e}_{n-1}$ — базис, в котором матрица формы h имеет канонический вид (3.11). Применив теорему 4.1 к форме h, заключаем, что матрица $\tilde{h}_{ij}$ в каноническом базисе $\tilde{e}_1, \ldots , \tilde{e}_{n-1}$ является единичной. Дополним базис $\tilde{\mathbf{e}}_1, \ldots , \tilde{\mathbf{e}}_{n-1}$ подпространства U вектором $\mathbf{e}_n \notin U$. В результате этого получится базис $\tilde{\mathbf{e}}_1, \ldots , \tilde{\mathbf{e}}_{n-1}, \mathbf{e}_n$, в котором матрица исходной квадратичной формы g имеет следующий вид:

$$\mathcal{G}_1 = \begin{Vmatrix} 1 & \ldots & 0 & \tilde{g}_{1n} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 1 & \tilde{g}_{n-1\,n} \\ \tilde{g}_{n1} & \ldots & \tilde{g}_{n\,n-1} & g_{nn} \end{Vmatrix}. \tag{4.3}$$

Переход из базиса $\mathbf{e}_1, \ldots , \mathbf{e}_n$ в базис $\tilde{\mathbf{e}}_1, \ldots , \tilde{\mathbf{e}}_{n-1}, \mathbf{e}_n$ дается блочно-диагональной матрицей перехода S вида

$$S_1 = \begin{Vmatrix} S^1_1 & \ldots & S^1_n & 0 \\ \vdots & \ddots & \vdots & \vdots \\ S^{n-1}_1 & \ldots & S^{n-1}_n & 0 \\ 0 & \ldots & 0 & 1 \end{Vmatrix}. \tag{4.4}$$

Формула (1.12) определяет связь матрицы (4.3) с матрицей $\mathcal{G}$ квадратичной формы g в исходном базисе: $\mathcal{G}_1 = S^{\mathrm{tr}}\, \mathcal{G}\, S$. Из этой формулы получаем:

$$\det \mathcal{G}_1 = \det \mathcal{G}\, (\det S)^2 = M_n\, (\det S)^2. \tag{4.5}$$

В силу формулы (4.5) из положительности главного минора $M_n = \det \mathcal{G}$ в исходной матрице (4.1) вытекает положительность детерминанта матрицы (4.3), то есть мы имеем $\det \mathcal{G}_1 > 0$.

Вычислим детерминант матрицы (4.3) непосредственным образом. Для этого умножим первый столбец этой матрицы на $\tilde{g}_{1n}$ и вычтем из последнего столбца. Второй столбец умножим на $\tilde{g}_{2n}$ и также вычтем из последнего. Подобные действия проделаем n-1 раз с каждым из первых n-1 столбцов матрицы (4.3). Из курса общей алгебры известно, что такие преобразования не меняют детерминанта матрицы. Они упрощают саму матрицу делая ее нижне-треугольной. Это позволяет вычислить $\det \mathcal{G}_1$ в явной форме:

$$\det \mathcal{G}_1 = \det \begin{vmatrix} 1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 1 & 0 \\ \tilde{g}_{n1} & \ldots & \tilde{g}_{n\,n-1} & \tilde{g}_{nn} \end{vmatrix} = \tilde{g}_{nn}. \tag{4.6}$$

Элемент $\tilde{g}_{nn}$ в преобразованной матрице определяется сделанными выше преобразованиями. Для $\tilde{g}_{nn}$ получается формула

$$\tilde{g}_{nn} = g_{nn} - \sum_{i=1}^{n-1} g_{ni}\, g_{in} = g_{nn} - \sum_{i=1}^{n-1} (g_{in})^2. \tag{4.7}$$

Матрица квадратичной формы g в базисе $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_{n-1}, \mathbf{e}_n$ близка к диагональной. Завершим процесс диагонализации матрицы квадратичной формы g заменив вектор $\mathbf{e}_n$ вектором $\tilde{\mathbf{e}}_n \notin U$, определив этот вектор формулой

$$\tilde{\mathbf{e}}_n = \mathbf{e}_n - \sum_{i=1}^{n-1} g_{in} \cdot \tilde{\mathbf{e}}_i.$$

Переход из базиса $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_{n-1}, \mathbf{e}_n$ в базис $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ изменяет лишь последний базисный вектор. Поэтому единичный диагональный блок в (4.3) остается неизменным. Для недиагональных элементов $g(\tilde{\mathbf{e}}_k, \tilde{\mathbf{e}}_n)$ в новом базисе имеем

$$g(\tilde{\mathbf{e}}_k, \tilde{\mathbf{e}}_n) = \tilde{g}_{kn} - \sum_{i=1}^{n-1} g_{in}\, g(\tilde{\mathbf{e}}_k, \tilde{\mathbf{e}}_i) = \tilde{g}_{kn} - \sum_{i=1}^{n-1} g_{in}\, \tilde{h}_{ki} = 0.$$

Равенство нулю $g(\tilde{\mathbf{e}}_k, \tilde{\mathbf{e}}_n) = 0$ в этой формуле вытекает из единичности матрицы квадратичной формы h в каноническом базисе $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_{n-1}$. Для диагонального элемента $g(\tilde{\mathbf{e}}_n, \tilde{\mathbf{e}}_n)$ имеем

$$g(\tilde{\mathbf{e}}_n, \tilde{\mathbf{e}}_n) = g_{nn} - \sum_{i=1}^{n-1}\sum_{k=1}^{n-1} g_{in}\, g_{kn}\, h_{ik} = g_{nn} - \sum_{i=1}^{n-1} (g_{in})^2.$$

Сравнивая полученное выражение с (4.7), мы обнаруживаем совпадение $g(\tilde{\mathbf{e}}_n, \tilde{\mathbf{e}}_n) = \tilde{g}_{nn}$. Матрица квадратичной формы g в базисе $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ диагональна. Она имеет следующий вид:

$$\mathcal{G}_2 = \begin{Vmatrix} 1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 1 & 0 \\ 0 & \ldots & 0 & \tilde{g}_{nn} \end{Vmatrix}. \tag{4.8}$$

Соединив формулы (4.5) и (4.6), для элемента $\tilde{g}_{nn}$ в матрице (4.8) выводим

$$\tilde{g}_{nn} = M_n\, (\det S)^2.$$

Теперь из положительности диагонального главного минора M_n в исходной матрице (4.1) вытекает положительность элемента $\tilde{g}_{nn}$ и, следовательно, положительная определенность квадратичной формы g. Индукционный переход и теорема в целом доказаны. $\square$

ГЛАВА V

ЕВКЛИДОВЫ ПРОСТРАНСТВА.

§ 1. Норма и скалярное произведение. Угол между векторами. Ортонормированные базисы.

Определение 1.1. *Евклидовым векторным пространством* называется линейное векторное пространство V над полем вещественных чисел $\mathbb{R}$, в котором выбрана и зафиксирована некоторая положительно определенная квадратичная форма g.

Пусть (V, g) — евклидово векторное пространство. В пространстве V можно задать и другие положительно определенные формы, однако, только форма g жестко связана с V и задает структуру евклидового пространства в V. Пространства (V, g_1) и (V, g_2) совпадают как линейные векторные пространства, но они различны как евклидовы пространства.

В евклидовом векторном пространстве V вводится специальная терминология и специальные обозначения. Значения квадратичной формы $g(\mathbf{v})$ на векторах неотрицательно. Корень квадратный из $g(\mathbf{v})$ называется *нормой* вектора $\mathbf{v}$ или *длиной* вектора $\mathbf{v}$. Норма вектора $\mathbf{v}$ обозначается так:

$$|\mathbf{v}| = \sqrt{g(\mathbf{v})}. \tag{1.1}$$

С квадратичной формой g связана симметричная билинейная форма $g(\mathbf{v}, \mathbf{w})$, определяемая по формуле восстановления (1.6) из четвертой главы. Значение билинейной формы g на векторах $\mathbf{v}$ и $\mathbf{w}$ называется *скалярным произведением* векторов $\mathbf{v}$ и $\mathbf{w}$. Скалярное произведение векторов мы будем обозначать так:

$$(\mathbf{v} \,|\, \mathbf{w}) = g(\mathbf{v}, \mathbf{w}). \tag{1.2}$$

В силу обозначений (1.1) и (1.2), работая с некоторым фиксированным евклидовым пространством (V, g), мы можем совсем не использовать символ g соответствующей положительно определенной формы.

Скалярное произведение (1.2) связывает два вектора $\mathbf{v}, \mathbf{w} \in V$. Этим оно существенно отличается от скалярного произведения (1.8) из третьей главы, которое связывает вектор и ковектор. Скалярное произведение (1.2) обладает следующими свойствами:

(1) $(\mathbf{v}_1 + \mathbf{v}_2 \,|\, \mathbf{w}) = (\mathbf{v}_1 \,|\, \mathbf{w}) + (\mathbf{v}_2 \,|\, \mathbf{w})$,
(2) $(\alpha \cdot \mathbf{v} \,|\, \mathbf{w}) = \alpha\, (\mathbf{v} \,|\, \mathbf{w})$,
(3) $(\mathbf{v} \,|\, \mathbf{w}_1 + \mathbf{w}_2) = (\mathbf{v} \,|\, \mathbf{w}_1) + (\mathbf{v} \,|\, \mathbf{w}_2)$,
(4) $(\mathbf{v} \,|\, \alpha \cdot \mathbf{w}) = \alpha\, (\mathbf{v} \,|\, \mathbf{w})$,
(5) $(\mathbf{v} \,|\, \mathbf{w}) = (\mathbf{w} \,|\, \mathbf{v})$,
(6) $|\mathbf{v}|^2 = (\mathbf{v} \,|\, \mathbf{v}) \geq 0$, причем из $|\mathbf{v}| = 0$ вытекает $\mathbf{v} = \mathbf{0}$.

Свойства (1)-(4) отражают билинейность формы g в (1.2). Они полностью аналогичны свойствам скалярного произведения вектора и ковектора (см. формулы (1.9) в третьей главе).

Свойства (5) и (6) таких аналогов не имеют. Но именно они делают скалярное произведение (1.2) прямым обобщением обычного скалярного произведения геометрических векторов.

Теорема 1.1. *Из свойств* (1)-(6) *вытекают два дополнительных свойства*

(7) $|(\mathbf{v},\mathbf{w})| \leq |\mathbf{v}|\,|\mathbf{w}|$,

(8) $|\mathbf{v}+\mathbf{w}| \leq |\mathbf{v}| + |\mathbf{w}|$.

Свойство (7) *известно как неравенство Коши-Буняковского или неравенство Коши-Шварца, а свойство* (8) *называется неравенством треугольника.*

Док-во. Для доказательства неравенства (7) выберем два произвольных ненулевых вектора $\mathbf{v},\mathbf{w}\in V$ и рассмотрим числовую функцию $f(\alpha)$ числового аргумента α, определив ее следующей формулой:

$$f(\alpha) = |\mathbf{v} + \alpha\cdot\mathbf{w}|^2. \tag{1.3}$$

Пользуясь свойствами (1)-(6) скалярного произведения, находим, что функция $f(\alpha)$ есть полином второй степени по переменной α:

$$\begin{aligned} f(\alpha) = |\mathbf{v}+\alpha\cdot\mathbf{w}|^2 = (\mathbf{v}+\alpha\cdot\mathbf{w}\,|\,\mathbf{v}+\alpha\cdot\mathbf{w}) = \\ = (\mathbf{v}\,|\,\mathbf{v}) + 2\,\alpha\,(\mathbf{v}\,|\,\mathbf{w}) + \alpha^2\,(\mathbf{w}\,|\,\mathbf{w}). \end{aligned}$$

Функция (1.3) ограничена снизу: $f(\alpha) \geqslant 0$. Это вытекает из свойства (6). Вычислим точку минимума функции $f(\alpha)$, приравняв нулю ее производную:

$$f'(\alpha) = 2\,(\mathbf{v}\,|\,\mathbf{w}) + 2\,\alpha\,(\mathbf{w}\,|\,\mathbf{w}) = 0.$$

Это дает $\alpha_{\min} = -(\mathbf{v}\,|\,\mathbf{w})/(\mathbf{w}\,|\,\mathbf{w})$. Теперь запишем неравенство $f(\alpha)\geqslant 0$ для значения функции $f(\alpha)$ в точке минимума:

$$f(\alpha_{\min}) = \frac{|\mathbf{v}|^2\,|\mathbf{w}|^2 - (\mathbf{v}\,|\,\mathbf{w})^2}{|\mathbf{w}|^2} \geqslant 0. \tag{1.4}$$

Знаменатель дроби в (1.4) положителен, поэтому из неравенства (1.4) легко выводится свойство (7).

Для доказательства свойства (8) рассмотрим квадрат нормы вектора $\mathbf{v}+\mathbf{w}$. Для него прямым вычислением находим

$$|\mathbf{v}+\mathbf{w}|^2 = (\mathbf{v}+\mathbf{w}\,|\,\mathbf{v}+\mathbf{w}) = |v|^2 + 2\,(\mathbf{v}\,|\,\mathbf{w}) + |\mathbf{w}|^2. \tag{1.5}$$

Применив уже доказанное свойство (7), для правой части полученного выражения получаем оценку:

$$|\mathbf{v}|^2 + 2\,(\mathbf{v}\,|\,\mathbf{w}) + |\mathbf{w}|^2 \leqslant |\mathbf{v}|^2 + 2\,|\mathbf{v}|\,|\mathbf{w}| + |\mathbf{w}|^2 = (|\mathbf{v}| + |\mathbf{w}|)^2.$$

Из соотношения (1.5) и из полученного неравенства выводим: $|\mathbf{v}+\mathbf{w}|^2 \leq (|\mathbf{v}|+|\mathbf{w}|)^2$. Теперь свойство (8) получается простым извлечением квадратного корня. Такой прием корректен, ибо функция $y=\sqrt{x}$ монотонно возрастает на полупрямой $[0,\ +\infty[$. Теорема доказана. $\square$

Аналогия между (1.2) и обычным геометрическим скалярным произведением, а также доказанное выше неравенство $|(\mathbf{v}, \mathbf{w})| \leqslant |\mathbf{v}|\,|\mathbf{w}|$ позволяют ввести понятие *угла между векторами* в евклидовом пространстве.

Определение 1.2. Вещественное число φ, которое заключено в интервале $0 \leqslant \varphi \leqslant \pi$ и определяется формулой

$$\cos(\varphi) = \frac{(\mathbf{v}\,|\,\mathbf{w})}{|\mathbf{v}|\,|\mathbf{w}|}, \tag{1.6}$$

называется углом между двумя ненулевыми векторами $\mathbf{v}$ и $\mathbf{w}$ в евклидовом пространстве V.

В силу свойства (7) из теоремы 1.1 дробь в правой части (1.6) ограничена по модулю числом 1. Поэтому формула (1.6) корректна. Она однозначно определяет угол φ из указанного интервала $0 \leqslant \varphi \leqslant \pi$.

Определение 1.3. Два вектора $\mathbf{v}$ и $\mathbf{w}$ в евклидовом пространстве называются *ортогональными*, если угол между ними прямой ($\varphi = \pi/2$).

Определение 1.3 применимо только к ненулевым векторам $\mathbf{v}$ и $\mathbf{w}$. Более общим является определение 2.1 из четвертой главы. Переформулируем его применительно к евклидовым пространствам.

Определение 1.4. Два вектора $\mathbf{v}$ и $\mathbf{w}$ в евклидовом пространстве называются ортогональными, если их скалярное произведение равно нулю, то есть имеет место равенство $(\mathbf{v}\,|\,\mathbf{w}) = 0$.

Для ненулевых векторов $\mathbf{v}$ и $\mathbf{w}$ эти два определения 1.3 и 1.4 абсолютно эквивалентны.

Пусть $\mathbf{v}_1, \ldots, \mathbf{v}_m$ некоторая система векторов в евклидовом пространстве (V, g). Матрица g_{ij}, составленная из скалярных произведений этих векторов

$$g_{ij} = (\mathbf{v}_i\,|\,\mathbf{v}_j), \tag{1.7}$$

называется *матрицей Грама* системы векторов $\mathbf{v}_1, \ldots, \mathbf{v}_m$.

Теорема 1.2. *Система векторов $\mathbf{v}_1, \ldots, \mathbf{v}_m$ линейно зависима тогда и только тогда, когда детерминант матрицы Грама для нее равен нулю.*

Док-во. Пусть вектора $\mathbf{v}_1, \ldots, \mathbf{v}_m$ линейно зависимы. Тогда существует нетривиальная линейная комбинация этих векторов, равная нулю:

$$\alpha_1 \cdot \mathbf{v}_1 + \ldots + \alpha_m \cdot \mathbf{v}_m = \mathbf{0}. \tag{1.8}$$

Используя это, составим следующее выражение из компонент матрицы Грама:

$$\sum_{j=1}^{m} g_{ij}\,\alpha_j = \sum_{j=1}^{m} (\mathbf{v}_i\,|\,\mathbf{v}_j)\,\alpha_j = (\mathbf{v}_i\,|\,\alpha_1 \cdot \mathbf{v}_1 + \ldots + \alpha_m \cdot \mathbf{v}_m) = 0.$$

В силу произвольности индекса i, эта формула означает линейную зависимость столбцов матрицы g_{ij}. Из линейной зависимости столбцов вытекает равенство нулю детерминанта матрицы g_{ij} (этот факт хорошо известен из курса общей алгебры).

Пусть, наоборот, детерминант матрицы Грама равен нулю. Тогда столбцы этой матрицы линейно зависимы, следовательно, существует нетривиальная линейная комбинация столбцов этой матрицы, равная нулю:

$$\sum_{j=1}^{m} g_{ij}\,\alpha_j = 0. \tag{1.9}$$

Обозначим $\mathbf{v} = \alpha_1\cdot\mathbf{v}_1 + \ldots + \alpha_m\cdot\mathbf{v}_m$ и рассмотрим следующую двойную сумму, которая равна нулю в силу (1.9):

$$\begin{aligned}0 = \sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\,g_{ij}\,\alpha_j = \sum_{i=1}^{m}\alpha_i\,(\mathbf{v}_i\,|\,\alpha_1\cdot\mathbf{v}_1 + \ldots + \alpha_m\cdot\mathbf{v}_m) = \\ = (\alpha_1\cdot\mathbf{v}_1 + \ldots + \alpha_m\cdot\mathbf{v}_m\,|\,\mathbf{v}) = (\mathbf{v}\,|\,\mathbf{v}) = |\mathbf{v}|^2.\end{aligned}$$

Из полученного равенства $|\mathbf{v}|^2 = 0$ выводим $\mathbf{v} = 0$, что является следствием положительной определенности формы g, задающей структуру евклидового пространства в V. Из зануления вектора v возникает нетривиальная линейная комбинация (1.8) векторов $\mathbf{v}_1, \ldots, \mathbf{v}_m$, равная нулю. Значит, вектора $\mathbf{v}_1, \ldots, \mathbf{v}_m$ линейно зависимы. $\square$

Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — базис в конечномерном евклидовом векторном пространстве (V, g). Рассмотрим матрицу Грама системы базисных векторов. Знание компонент матрицы Грама позволяет вычислять норму (1.1) и скалярное произведение (1.2) через координаты векторов

$$|\mathbf{v}|^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}\,v^i\,v^j, \qquad (\mathbf{v}\,|\,\mathbf{w}) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}\,v^i\,w^j. \tag{1.10}$$

Базис называется *ортонормированным*, если матрица Грама для системы базисных векторов $\mathbf{e}_1, \ldots, \mathbf{e}_n$ единична:

$$g_{ij} = \begin{cases} 1 & \text{при} \ \ i = j, \\ 0 & \text{при} \ \ i \neq j. \end{cases} \tag{1.11}$$

Если же условие (1.11) не выполнено, то базис называется *косоугольным.* В ортонормированном базисе вектора имеют единичную длину и ортогональны друг другу, что сильно упрощает формулы (1.10):

$$|\mathbf{v}|^2 = \sum_{i=1}^{n}(v^i)^2, \qquad (\mathbf{v}\,|\,\mathbf{w}) = \sum_{i=1}^{n} v^i\,w^i. \tag{1.12}$$

Ортонормированные базисы существуют. Из формул (1.2) и (1.7) видим, что матрица Грама для системы базисных векторов $\mathbf{e}_1, \ldots, \mathbf{e}_n$ есть матрица квадратичной формы g в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Теорема 3.3 из четвертой главы утверждает существование базиса, в котором матрица формы g имеет канонический вид (формула 3.11 из главы IV). Для положительно определенной формы канонический вид задается единичной матрицей (см. теорему 4.1 из четвертой главы).

Теорема 4.8 о дополнении базиса из первой главы имеет обобщение на случай ортонормированных базисов в евклидовых пространствах.

Теорема 1.3. *Пусть* $\mathbf{e}_1, \ldots, \mathbf{e}_s$ — *ортонормированный базис в подпространстве* U *конечномерного евклидового пространства* (V, g). *Тогда он может быть дополнен до ортонормированного базиса* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *в пространстве* V.

Док-во. Рассмотрим ортогональное дополнение $U_\perp$ подпространства U. Согласно теореме 4.4 из четвертой главы имеет место разложение V в прямую сумму $V = U \oplus U_\perp$. Подпространство $U_\perp$ наследует структуру евклидового пространства с V в результате сужения формы g на $U_\perp$. Выберем ортонормированный базис $\mathbf{e}_{s+1}, \ldots, \mathbf{e}_n$ в $U_\perp$ и объединим два базиса в U и $U_\perp$. В результате этого получится базис в V (см. теорему 6.3 из первой главы). Вектора в построенном базисе имеют единичную длину и ортогональны друг другу. Следовательно, это ортонормированный базис в V, дополняющий базис $\mathbf{e}_1, \ldots, \mathbf{e}_s$ в подпространстве U. $\square$

Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_s$ и $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_s$ — два ортонормированных базиса в евклидовом пространстве и пусть S — матрица перехода из первого базиса во второй. Матрицы Грама этих двух базисов единичны, поэтому, применив формулы (1.12) из четвертой главы, для S выводим:

$$S^{\mathrm{tr}}\, S = 1, \qquad\qquad S^{-1} = S^{\mathrm{tr}}. \tag{1.13}$$

Квадратная матрица S, которая удовлетворяет соотношениям (1.13), называется *ортогональной*.

Из соотношений (1.13) для детерминанта ортогональной матрицы получаем: $(\det S)^2 = 1$. Поэтому ортогональные матрицы делятся два типа: с положительным детерминантом $\det S = 1$ и с отрицательным детерминантом $\det S = -1$. Такое разделение связано с понятием *ориентации*. Все базисы в линейном векторном пространстве над полем вещественных чисел $\mathbb{R}$ (не обязательно в евклидовом) можно разделить на два множества, которые условно называют «левоорентированными» и «правоорентированными». Матрица перехода из левого базиса в левый и из правого базиса в правый имеет положительный детерминант — она не меняет ориентации. Детерминант матрицы перехода из левого базиса в правый или, наоборот, из правого в левый отрицателен, такой переход изменяет ориентацию на противоположную. Говорят, что на вещественном векторном пространстве задана ориентация, если указан какой-то механизм, позволяющий как-то выделять один из двух типов базисов. Выделенные базисы при этом нвзывают «правоорентированными».

§ 2. Квадратичные формы в евклидовом пространстве. Диагонализация пары форм.

Пусть (V, g) — некоторое евклидово векторное пространство. Для всякой квадратичной формы φ в V определим следующее отношение:

$$\mu(\mathbf{v}) = \frac{|\varphi(\mathbf{v})|}{|\mathbf{v}|^2}. \tag{2.1}$$

Число $\mu(\mathbf{v})$ в (2.1) вещественно и неотрицательно. Заметим, что $\mu(\alpha\cdot\mathbf{v}) = \mu(\mathbf{v})$ для любого ненулевого числа $\alpha \in \mathbb{R}$. Это позволяет вектор $\mathbf{v}$ в (2.1) считать вектором единичной длины.

Обозначим через $\|\varphi\|$ верхнюю грань чисел $\mu(\mathbf{v})$ для всевозможных векторов единичной длины (такие вектора заметают *единичную сферу* в евклидовом векторном пространстве V). А именно, обозначим

$$\|\varphi\| = \sup_{|\mathbf{v}|=1} \mu(\mathbf{v}). \tag{2.2}$$

Определение 2.1. Величина $\|\varphi\|$, определяемая формулами (2.1) и (2.2) называется *нормой* квадратичной формы φ в евклидовом векторном пространстве V. Если норма $\|\varphi\|$ конечна, то говорят, что форма φ *ограничена.*

Теорема 2.1. *Если квадратичная форма φ ограничена, то для значений соответствующей симметричной билинейной формы φ имеется оценка*

$$|\varphi(\mathbf{v}, \mathbf{w})| \leqslant \|\varphi\| \, |\mathbf{v}| \, |\mathbf{w}|.$$

Док-во. Для вычисления $\varphi(\mathbf{v}, \mathbf{w})$ воспользуемся следующим равенством, которое является еще одним вариантом формулы восстановления:

$$4\alpha \, \varphi(\mathbf{v}, \mathbf{w}) = \varphi(\mathbf{v} + \alpha \cdot \mathbf{w}) - \varphi(\mathbf{v} - \alpha \cdot \mathbf{w}). \tag{2.3}$$

Из (2.3) вытекает следующее неравенство для $4\alpha \, \varphi(\mathbf{v}, \mathbf{w})$:

$$4\alpha \, \varphi(\mathbf{v}, \mathbf{w}) \leqslant |\varphi(\mathbf{v} + \alpha \cdot \mathbf{w})| + |\varphi(\mathbf{v} - \alpha \cdot \mathbf{w})|. \tag{2.4}$$

Теперь применим неравенство $|\varphi(\mathbf{u})| \leqslant \|\varphi\| \, |\mathbf{u}|^2$, вытекающее из (2.1) и (2.2), для оценки правой части (2.4). Это дает

$$4\alpha \, \varphi(\mathbf{v}, \mathbf{w}) \leqslant \|\varphi\| \, (|\mathbf{v} + \alpha \cdot \mathbf{w}|^2 + |(\mathbf{v} - \alpha \cdot \mathbf{w})|^2). \tag{2.5}$$

Выразим квадраты модулей через скалярные произведения:

$$|\mathbf{v} + \alpha \cdot \mathbf{w}|^2 = |\mathbf{v}|^2 \pm 2\alpha \, (\mathbf{v} \,|\, \mathbf{w}) + \alpha^2 \, |\mathbf{w}|^2.$$

Это позволяет упростить неравенство (2.5), сводя его к виду:

$$4\alpha \, \varphi(\mathbf{v}, \mathbf{w}) \leqslant 2\|\varphi\| \, (|\mathbf{v}|^2 + \alpha^2 \, |\mathbf{w}|^2).$$

Запишем полученное неравенство, немного преобразовав его:

$$f(\alpha) = \alpha^2 \|\varphi\| \, |\mathbf{w}|^2 - 2\alpha \, \varphi(\mathbf{v}, \mathbf{w}) + \|\varphi\| \, |\mathbf{w}|^2 \geqslant 0.$$

Числовая функция $f(\alpha)$ числового аргумента α есть полином второй степени по α. Найдем точку минимума этой функции $\alpha = \alpha_{\min}$ из условия зануления ее производной $f'(\alpha) = 0$:

$$\alpha_{\min} = \frac{\varphi(\mathbf{v}, \mathbf{w})}{\|\varphi\| \, |\mathbf{w}|^2}.$$

Теперь запишем неравенство $f(\alpha_{\min}) \geqslant 0$ для значения этой функции в точке ее минимума. Это дает оценку

$$\varphi(\mathbf{v}, \mathbf{w})^2 \leq \|\varphi\| \, |\mathbf{v}|^2 \, |\mathbf{w}|^2,$$

откуда выводится требуемая оценка для $|\varphi(\mathbf{v}, \mathbf{w})|$. Отметим, что аналогичный прием был использован при доказательстве неравенства Коши-Буняковского-Шварца в теореме 1.1. □

Теорема 2.2. *Всякая квадратичная форма φ в конечномерном евклидовом векторном пространстве V ограничена.*

Док-во. Выберем ортонормированный базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в V и разложим единичные вектора $\mathbf{v}$ по этому базису. Для координат таких векторов в силу соотношения (1.12) получим тождество

$$(v^1)^2 + \ldots + (v^n)^2 = 1,$$

откуда $|v^i| \leqslant 1$. Вычислим $\mu(\mathbf{v})$ в формуле (2.1) через координаты вектора $\mathbf{v}$:

$$\mu(\mathbf{v}) = |\varphi(\mathbf{v})| = \left|\sum_{i=1}^{n} \sum_{j=1}^{n} \varphi_{ij} \, v^i \, v^j\right|.$$

Оценки $|v^i| \leqslant 1$ для координат единичного вектора позволяют получить оценку для $\mu(\mathbf{v})$ через компоненты матрицы квадратичной формы φ:

$$\mu(\mathbf{v}) \leqslant \sum_{i=1}^{n} \sum_{j=1}^{n} |\varphi_{ij}|, \tag{2.6}$$

Правая часть (2.6) не зависит от $\mathbf{v}$, что доказывает ограниченность сверху для $\mu(\mathbf{v})$ и конечность нормы квадратичной формы $\|\varphi\|$. Теорема доказана. □

Теорема 2.3. *Для всякой квадратичной формы φ в конечномерном евклидовом векторном пространстве V верхняя грань в формуле* (2.2) *достигается, то есть существует вектор $\mathbf{v} \neq 0$, такой, что $|\varphi(\mathbf{v})| = \|\varphi\| \, |\mathbf{v}|^2$.*

Док-во. Из курса математического анализа знаем, что верхняя грань числового множества есть предел некоторой последовательности чисел, взятых из этого множества (см. [6]). Это означает, что существует последовательность векторов единичной длины $\mathbf{v}(1)$, $\mathbf{v}(2)$, ... в пространстве V, такая, что

$$\|\varphi\| = \lim_{s\to\infty} |\varphi(\mathbf{v}(s))|. \tag{2.7}$$

Выберем некоторый ортонормированный базис $e_1, \ldots, e_n$ в V и разложим каждый из векторов $v(s)$ в последовательности по этому базису. Равенство

$$(v^1(s))^2 + \ldots + (v^n(s))^2 = 1 \tag{2.8}$$

вытекает из (1.12) и из единичности длины всех векторов в последовательности. В силу (2.8) координаты векторов $v(s)$ ограничены:

$$-1 \leqslant v^i(s) \leqslant 1.$$

Из курса математического анализа знаем, что во всякой ограниченной числовой последовательности можно выбрать сходящуюся подпоследовательность. Проделав такую процедуру выбора n раз, мы можем выделить в последовательности $\mathbf{v}(1)$, $\mathbf{v}(2)$, ... такую подпоследовательность векторов, что все соответствующие им последовательности координат будут сходиться:

$$v^i = \lim_{k\to\infty} v^i(s_k). \tag{2.9}$$

Обозначим через $\mathbf{v}$ вектор, координаты которого определяются значениями пределов (2.9). Переходя к пределу при $k \to \infty$ в формуле (2.8), заключаем, что предельный вектор $\mathbf{v}$ будет единичным по длине: $|\mathbf{v}| = 1$.

Вычислим $|\varphi(\mathbf{v})|$ по матрице квадратичной формы φ в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и по координатам вектора $\mathbf{v}$ в этом базисе

$$|\varphi(\mathbf{v})| = \left|\sum_{i=1}^{n}\sum_{j=1}^{n} \varphi_{ij}\, v^i\, v^j\right| = \lim_{k\to\infty}\left|\sum_{i=1}^{n}\sum_{j=1}^{n} \varphi_{ij}\, v^i(s_k)\, v^j(s_k)\right|.$$

Продолжим эту цепочку вычислений, учитывая (2.7). Для $|\varphi(\mathbf{v})|$ это дает

$$|\varphi(\mathbf{v})| = \lim_{k\to\infty} |\varphi(v(s_k))| = \lim_{s\to\infty} |\varphi(v(s))| = \|\varphi\|. \tag{2.10}$$

Умножив предельный вектор $\mathbf{v}$ на некоторое число α, мы можем снять ограничение $|\mathbf{v}| = 1$. Тогда (2.10) запишется как $|\varphi(\mathbf{v})| = \|\varphi\|\,|\mathbf{v}|^2$. Теорема полностью доказана. $\square$

Теорема 2.4. *Для всякой квадратичной формы φ в конечномерном евклидовом векторном пространстве (V, g) существует ортонормированный базис, в котором матрица этой формы диагональна.*

Док-во. Доказательство проведем индукцией по размерности пространства V. В случае $\dim V = 1$ утверждение теоремы очевидно: всякая матрица размером 1×1 диагональна.

Предположим, что утверждение теоремы верно для всех квадратичных форм в евклидовых пространствах размерности меньшей, чем n. Пусть $\dim V = n$ и пусть φ — квадратичная форма в евклидовом пространстве (V, g). Применим результаты теорем 2.2 и 2.3, что позволяет найти единичный по длине вектор $\mathbf{v} \in V$, для которого $|\varphi(\mathbf{v})| = \|\varphi\|$. Пусть для определенности $\varphi(\mathbf{v}) \geq 0$, что позволяет убрать знак модуля: $\varphi(\mathbf{v}) = \|\varphi\|$. Если $\varphi(\mathbf{v}) < 0$, мы просто заменим φ на форму $\tilde{\varphi} = -\varphi$, поскольку формы, отличающиеся только знаком, диагонализуются одновременно.

Обозначим через U линейную оболочку вектора $\mathbf{v}$ и рассмотрим ортогональное дополнение $U_\perp$. Подпространства $U = \langle v\rangle$ и $U_\perp$ имеют нулевое пересечение, их сумма является прямой и совпадает с V (см. теорему 4.4 из четвертой главы). Рассмотрим вектор единичной длины $w \in U_\perp$ и составим из $\mathbf{v}$ и $\mathbf{w}$ еще один вектор $\mathbf{u}$, задав его в форме линейной комбинации:

$$\mathbf{u} = \cos(\alpha) \cdot \mathbf{v} + \sin(\alpha) \cdot \mathbf{w},$$

где α — числовой параметр. Длина вектора $\mathbf{u}$ также единична, это вытекает из ортогональности единичных векторов $\mathbf{v}$ и $\mathbf{w}$ и из тождества

$$\cos^2(\alpha) + \sin^2(\alpha) = 1.$$

Вычислим значение квадратичной формы φ на векторе $\mathbf{u}$ и рассмотрим его как числовую функцию параметра α:

$$\begin{aligned} f(\alpha) = \varphi(\mathbf{u}) = cos^2(\alpha)\,\varphi(\mathbf{v}) + \\ + 2\sin(\alpha)\cos(\alpha)\,\varphi(\mathbf{v}, \mathbf{w}) + \sin^2(\alpha)\,\varphi(\mathbf{w}). \end{aligned}$$

По самому выбору вектора $\mathbf{v}$ имеется оценка $\varphi(\mathbf{u}) \leq \varphi(\mathbf{v})$, причем равенство в этой оценке достигается при $\alpha = 0$, когда $\mathbf{u} = \mathbf{v}$. Следовательно, точка $\alpha = 0$ есть точка максимума для функции $f(\alpha)$. Вычислим производную этой функции в точке $\alpha = 0$ и приравняем ее к нулю. Это дает

$$f'(0) = 2\,\varphi(\mathbf{v}, \mathbf{w}) = 0. \tag{2.11}$$

Значит, $\varphi(\mathbf{v}, \mathbf{w}) = 0$ для всех векторов $\mathbf{w} \in U_\perp$. Применим предположение индукции к подпространству $U_\perp$, размерность которого на единицу меньше, чем размерность пространства V. Это позволяет найти ортонормированный базис $\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}$ в подпространстве $U_\perp$, в котором матрица формы φ диагональна: $\varphi(\mathbf{e}_i, \mathbf{e}_j) = 0$ при $i \neq j$. Дополним построенный базис из $U_\perp$ вектором $\mathbf{e}_n = \mathbf{v}$. Дополнительный вектор $\mathbf{e}_n$ имеет единичную длину. Он ортогонален векторам $\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}$. Поэтому построенный базис является ортонормированным базисом в V. Диагональность матрицы квадратичной формы φ в нем вытекает из (2.11). Теорема доказана. $\square$

Доказанная теорема 2.4 известна как теорема о приведении пары форм φ и g к диагональному виду. Для этого одна из них должна быть положительно определенной. Тогда положительно определенная форма g задает структуру евклидового пространства в V, что позволяет применить теорему 2.4. Ортонормированность базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ означает диагональность матрицы g в базисе построенном согласно теореме 2.4.

§ 3. Самосопряженные операторы. Теорема о спектре и базисе из собственных векторов.

Определение 3.1. Линейный оператор $f\colon V \to V$ в евклидовом векторном пространстве V называется *симметрическим* или *самосопряженным* если для любых двух векторов $\mathbf{v}, \mathbf{w} \in V$ выполнено равенство: $(\mathbf{v}\,|\,f(\mathbf{w})) = (f(\mathbf{v})\,|\,\mathbf{w})$.

Определение 3.2. Линейный оператор $h\colon V \to V$ в евклидовом векторном пространстве V называется *сопряженным* оператору $f : V \to V$, если для любых двух векторов $\mathbf{v}, \mathbf{w} \in V$ выполнено равенство: $(\mathbf{v}\,|\,f(\mathbf{w})) = (h(\mathbf{v})\,|\,\mathbf{w})$. Сопряженный оператор обозначается так: $h = f^+$.

В § 4 третьей главы мы ввели понятие сопряженного отображения. Там мы показали, что всякое линейное отображение $f: V \to W$ имеет сопряженное отображение $f^*: W^* \to V^*$. Для линейного оператора $f: V \to V$ отображение f^* есть линейный оператор в сопряженном пространстве V^*. Он связан с исходным оператором f соотношением

$$\langle f^*(\mathbf{u}) \,|\, \mathbf{v} \rangle = \langle \mathbf{u} \,|\, f(\mathbf{v}) \rangle, \tag{3.1}$$

которое выполнено для всех $\mathbf{u} \in V^*$ и всех $\mathbf{v} \in V$.

Структура евклидового пространства в V задается положительно определенной квадратичной формой g. С каждой такой формой связано ассоциированное отображение $a_g: V \to V^*$ (см. § 2 четвертой главы), для которого

$$\langle a_g(\mathbf{v}) \,|\, \mathbf{w} \rangle = g(\mathbf{v}, \mathbf{w}) = (\mathbf{v} \,|\, \mathbf{w}). \tag{3.2}$$

В случае конечномерности V из положительной определенности формы g вытекает биективность a_g. Пользуясь этим, для каждого линейного оператора $f: V \to V$ построим композицию: $h = a_g^{-1} \circ \circ f^* \circ a_g$. Из (3.1) и (3.2) выводим

$$\begin{aligned} &(h(\mathbf{v}) \,|\, \mathbf{w}) = \langle a_g \circ h(\mathbf{v}) \,|\, \mathbf{w} \rangle = \\ &\quad = \langle f^* \circ a_g(\mathbf{v}) \,|\, \mathbf{w} \rangle = \langle a_g(\mathbf{v}) \,|\, f(\mathbf{w}) \rangle = (\mathbf{v} \,|\, f(\mathbf{w})). \end{aligned} \tag{3.3}$$

Сравнивая полученное выражение (3.3) с определением 3.2, мы можем сформулировать следующую теорему.

Теорема 3.1. *Для всякого оператора f в конечномерном евклидовом пространстве (V, g) существует однозначно определенный сопряженный оператор f^+, который задается формулой $f^+ = a_g^{-1} \circ f^* \circ a_g$.*

Док-во. Факт существования оператора f^+ мы уже показали выше. Докажем его единственность. Пусть h — некоторый другой оператор, удовлетворяющий условию определения 3.2. Тогда для разницы $r = h - f^+$ выводим

$$\begin{aligned} &(r(\mathbf{v}) \,|\, \mathbf{w}) = (h(\mathbf{v}) \,|\, \mathbf{w}) - (f^+(\mathbf{v}) \,|\, \mathbf{w}) = \\ &\qquad\qquad = (\mathbf{v} \,|\, f(\mathbf{w})) - (\mathbf{v} \,|\, f(\mathbf{w})) = 0. \end{aligned} \tag{3.4}$$

В силу произвольности вектора $\mathbf{w}$ в (3.4), имеем $h(\mathbf{v}) \in \operatorname{Ker} g$. Однако, $\operatorname{Ker} g = \{\mathbf{0}\}$ для положительно определенной формы g, следовательно, $h(\mathbf{v}) = 0$ для любого $\mathbf{v} \in V$. Отсюда выводим $h = 0$, что доказывает единственность сопряженного оператора f^+ для оператора f. $\square$

Следствие. *Операция перехода от f к сопряженному оператору f^+ в пространстве эндоморфизмов $\operatorname{End}(V)$ конечномерного евклидового векторного пространства (V, g) обладает следующими свойствами:*

$$\begin{aligned} &(f + h)^+ = f^+ + h^+, &\qquad &(\alpha \cdot f)^+ = \alpha \cdot f^+, \\ &(f \circ h)^+ = h^+ \circ f^+, &\qquad &(f^+)^+ = f. \end{aligned}$$

Факт существования и единственности сопряженного оператора f^+ для всякого оператора $f \in \mathrm{End}(V)$ позволяет вывести все эти соотношения непосредственно из определения 3.2.

Соотношение $f^+ = a_g^{-1} \circ f^* \circ a_g$ можно изобразить в виде следующей коммутативной диаграммы:

$$\begin{array}{ccc} V & \xrightarrow{\ f^+\ } & V \\ {\scriptstyle a_g}\big\downarrow & & \big\downarrow{\scriptstyle a_g} \\ V^* & \xrightarrow[\ f^*\ ]{} & V^* \end{array}$$

Из сравнения определений 3.1 и 3.2 заключаем, что самосопряженный оператор f — это оператор, который является сопряженным самому себе: $f^+ = f$.

Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — базис в конечномерном евклидовом пространстве (V, g) и пусть $h^1, \ldots, h^n$ — сопряженный базис из координатных функционалов в V^*. Для всякого вектора $\mathbf{v} \in V$ имеется разложение, которое вытекает из самого определения координатных функционалов (см. § 1 третьей главы):

$$\mathbf{v} = h^1(\mathbf{v}) \cdot \mathbf{e}_1 + \ldots + h^n(\mathbf{v}) \cdot \mathbf{e}_n.$$

Воспользуемся этой формулой для вычисления матрицы ассоциированного отображения a_g. Для определения компонент такой матрицы надо применить a_g по очереди ко всем базисным векторам $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и полученные ковекторы разложить по сопряженному базису в V^*. Для вывода таких разложений найдем значение функционала $a_g(\mathbf{e}_i)$ на произвольном векторе $\mathbf{v} \in V$:

$$\begin{aligned} a_g(\mathbf{e}_i)(\mathbf{v}) &= \langle a_g(\mathbf{e}_i) \,|\, \mathbf{v}\rangle = g(\mathbf{e}_i, \mathbf{v}) = \\ &= g(\mathbf{e}_i, h^1(\mathbf{v}) \cdot e_1 + \ldots + h^n(\mathbf{v}) \cdot \mathbf{e}_n) = \sum_{j=1}^{n} g_{ij}\, h^j(\mathbf{v}). \end{aligned}$$

Из этого соотношения в силу произвольности вектора $\mathbf{v} \in V$ заключаем, что матрица ассоциированного отображения a_g в базисах $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $h^1, \ldots, h^n$ совпадает с матрицей $g_{ij} = g(\mathbf{e}_i, \mathbf{e}_j)$ квадратичной формы g в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

Матрица g_{ij} невырождена (см. теорему 1.2 или критерий Сильвестра в § 4 четвертой главы). Обозначим через g^{ij} компоненты матрицы, обратной к матрице g_{ij}. Матрица g^{ij} является матрицей обратного отображения a_g^{-1}:

$$a_g(\mathbf{e}_i) = \sum_{j=1}^{n} g_{ij}\, h^j, \qquad\qquad a_g^{-1}(h^j) = \sum_{j=1}^{n} g^{ij}\, \mathbf{e}_i. \tag{3.5}$$

Матрица, обратная симметричной, симметрична (это известный факт из общей алгебры). Поэтому $g^{ij} = g^{ji}$.

Матрица сопряженного отображения f^* была вычислена нами ранее (см. формулу (4.2) и теорему 4.3 из третьей главы). Применительно к данному случаю результаты третьей главы означают, что матрица оператора $f^* : V^* \to V^*$ в базисе из координатных функционалов $h^1, \ldots, h^n$ совпадает с матрицей исходного оператора f в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Соединим это с (3.5) и используем формулу $f^+ = a_g^{-1} \circ f^* \circ a_g$ из теоремы 3.1. Тогда для матрицы F^+ сопряженного оператора f^+ получаем:

$$(F^+)^i_j = \sum_{k=1}^{n} \sum_{q=1}^{n} g^{iq}\, F^k_q\, g_{kj}. \tag{3.6}$$

В матричном виде формула (3.6) может быть записана так: $F^+ = G^{-1}\, F^{\,\mathrm{tr}}\, G$, где G — матрица Грама того базиса, в котором вычисляются матрицы операторов f и f^+. Наиболее просто формула (3.6) выглядит в ортонормированных базисах. Здесь операция сопряжения оператора сводится к транспонированию его матрицы. Матрица самосопряженного оператора в ортонормированном базисе симметрична. Поэтому самосопряженные операторы часто называют симметрическими.

Пусть $f\colon\, V \to V$ — самосопряженный оператор в евклидовом пространстве V. С каждым таким оператором связана квадратичная форма φ_f:

$$\varphi_f(\mathbf{v}) = (\mathbf{v}\,|\, f(\mathbf{v})). \tag{3.7}$$

Пусть, наоборот, в конечномерном евклидовом пространстве (V, g) задана квадратичная форма φ. Форма φ определяет ассоциированное отображение $a_\varphi:\ V \to V^*$ (см. определение 2.5 из главы IV), которое для любых двух векторов $\mathbf{v}, \mathbf{w} \in V$ удовлетворяет соотношению

$$\langle a_\varphi(\mathbf{v})\,|\,\mathbf{w}\rangle = \varphi(\mathbf{v}, \mathbf{w}). \tag{3.8}$$

Положительно определенная форма g, задающая структуру евклидового пространства в V, также имеет ассоциированное отображение a_g. Отображение a_g биективно в силу невырожденности g (см. теорему 4.2 из четвертой главы). Это позволяет рассмотреть следующую композицию отображений a_g^{-1} и a_φ:

$$f_\varphi = a_g^{-1} \circ a_\varphi. \tag{3.9}$$

Композиция (3.9) есть оператор в V. Его называют *ассоциированным оператором квадратичной формы* φ в евклидовом пространстве. В силу обратимости отображения a_g, (3.2) можно преобразовать к виду $\langle \mathbf{u}\,|\,\mathbf{w}\rangle = (a_g^{-1}(\mathbf{u})\,|\,\mathbf{w})$. Соединив это с формулой (3.8), выводим

$$(f_\varphi(\mathbf{v})\,|\,\mathbf{w}) = (a_g^{-1}(a_\varphi(\mathbf{v}))\,|\,\mathbf{w}) = \langle a_\varphi(\mathbf{v})\,|\,\mathbf{w}\rangle = \varphi(\mathbf{v}, \mathbf{w}). \tag{3.10}$$

Теперь, используя симметричность формы $\varphi(\mathbf{v}, \mathbf{w})$ в правой части (3.10), это выражение можно преобразовать к виду

$$(f_\varphi(\mathbf{v})\,|\,\mathbf{w}) = \varphi(\mathbf{v}, \mathbf{w}) = \varphi(\mathbf{w}, \mathbf{v}) = (f_\varphi(\mathbf{w})\,|\,\mathbf{v}) = (\mathbf{v}\,|\, f_\varphi(\mathbf{w})). \tag{3.11}$$

Соотношение (3.11), которое выполняется тождественно для всех $v, w \in V$, в точности, означает самосопряженность оператора f_φ (см. определение 2.1).

Формула (3.7) каждому самосопряженному оператору f сопоставляет квадратичную форму φ_f, а формула (3.9) квадратичной форме φ ставит в соответствие самосопряженный оператор f_φ. Эти соответствия взаимно однозначны и являются обратными друг для друга. Действительно, применим формулу (3.7) к оператору (3.9) и учтем (3.10):

$$\varphi_f(\mathbf{v}) = (\mathbf{v}\,|\,f_\varphi(\mathbf{v})) = \varphi(\mathbf{v},\mathbf{v}) = \varphi(\mathbf{v}).$$

Теперь, наоборот, построим оператор $h = f_\varphi$ по квадратичной форме $\varphi = \varphi_f$. Для такого оператора h и двух произвольных векторов $\mathbf{v}, \mathbf{w} \in V$ из формулы (3.10) получаем следующие соотношения:

$$(h(\mathbf{v})\,|\,\mathbf{w}) = \varphi_f(\mathbf{v},\mathbf{w}) = (\mathbf{v}\,|\,f(\mathbf{w})) = (f(\mathbf{v})\,|\,\mathbf{w}).$$

В силу произвольности вектора $\mathbf{w} \in V$ и невырожденности формы g, определяющей евклидово скалярное произведение в пространстве V, из этого выводим $h(\mathbf{v}) = f(\mathbf{v})$.

Из сказанного выше следует вывод: задание самосопряженного оператора в конечномерном евклидовом векторном пространстве эквивалентно заданию квадратичной формы в этом пространстве. Этот факт позволяет применить результат теоремы 2.4 для описания самосопряженных операторов в конечномерном случае.

Теорема 3.2. *Все собственные числа самосопряженного оператора f в конечномерном евклидовом пространстве V вещественны и существует ортонормированный базис, составленный из собственных векторов этого оператора.*

Док-во. Для самосопряженного оператора f в V рассмотрим симметричную билинейную форму $\varphi_f(\mathbf{v},\mathbf{w})$, соответствующую квадратичной форме (3.7). Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — ортонормированный базис, в котором матрица формы φ_f диагональна. Тогда из формулы (3.7) выводим

$$\varphi_f(\mathbf{e}_i,\mathbf{e}_j) = (\mathbf{e}_i\,|\,f(\mathbf{e}_j)) = \sum_{k=1}^{n} F^k_j\, g_{ik} = F^i_k. \tag{3.12}$$

Из (3.12) видим, что матрицы оператора f и формы φ_f в таком базисе совпадают, что и доказывает утверждение теоремы. □

Доказанная теорема 3.2 известна как теорема о спектре и базисе из собственных векторов самосопряженного оператора. Основной результат этой теоремы — диагонализуемость самосопряженных операторов в конечномерном евклидовом пространстве. Характеристический полином самосопряженного оператора раскладывается на линейные множители в поле $\mathbb{R}$, а собственные подпространства такого оператора совпадают с соответствующими корневыми подпространствами и в сумме дают пространство V:

$$V = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_s}. \tag{3.13}$$

Теорема 3.3. *У всякого самосопряженного оператора собственные векторы, соответствующие различным собственным числам ортогональны.*

Док-во. Пусть f — самосопряженный оператор в евклидовом пространстве и пусть $\lambda \neq \mu$ — два его собственных числа. Рассмотрим соответствующие им собственные векторы $\mathbf{a}$ и $\mathbf{b}$:

$$f(\mathbf{a}) = \lambda \cdot \mathbf{a}, \qquad\qquad f(\mathbf{b}) = \mu \cdot \mathbf{b}.$$

Тогда для этих двух собственных векторов $\mathbf{a}$ и $\mathbf{b}$ имеем:

$$\lambda\,(\mathbf{a}\,|\,\mathbf{b}) = (f(\mathbf{a})\,|\,\mathbf{b}) = (\mathbf{a}\,|f(\mathbf{b})) = \mu\,(\mathbf{a}\,|\,\mathbf{b}).$$

Отсюда $(\lambda - \mu)\,(\mathbf{a}\,|\,\mathbf{b}) = 0$, но $\lambda - \mu \neq 0$. Поэтому $(\mathbf{a}\,|\,\mathbf{b}) = 0$. Это завершает доказательство теоремы. □

Пусть ядро самосопряженного оператора f нетривиально $\operatorname{Ker} f \neq \{\mathbf{0}\}$. Тогда в разложении (3.13) имеем $\lambda_1 = 0$ и при этом

$$\operatorname{Ker} f = V_{\lambda_1}, \qquad\qquad \operatorname{Im} f = V_{\lambda_2} \oplus \ldots \oplus V_{\lambda_s}.$$

Значит, ядро и образ самосопряженного оператора f ортогональны друг другу и в сумме дают все пространство V:

$$V = \operatorname{Ker} f \oplus \operatorname{Im} f. \tag{3.14}$$

§ 4. Изометрии и ортогональные операторы.

Определение 4.1. Линейное отображение $f\colon V \to W$ из одного евклидового векторного пространства (V, g) в другое евклидово векторное пространство (W, h) называется *изометрией*, если

$$(f(\mathbf{x})\,|\,f(\mathbf{y})) = (\mathbf{x}\,|\,\mathbf{y}) \tag{4.1}$$

для всех $\mathbf{x}, \mathbf{y} \in V$, то есть, если оно сохраняет величину евклидового скалярного произведения векторов.

Из (4.1) вытекает $|f(\mathbf{x})| = |\mathbf{x}|$, поэтому $f(\mathbf{x}) = 0$ влечет $|\mathbf{x}| = 0$, откуда $\mathbf{x} = \mathbf{0}$. Эти рассуждения показывают, что ядро изометрии нулевое $\operatorname{Ker} f = \{\mathbf{0}\}$ и что всякая изометрия инъективна. В силу формулы восстановления для квадратичных форм (см. формулу (1.6) в четвертой главе) для проверки изометричности линейного отображения $f : V \to W$ достаточно проверить сохранение норм $|f(\mathbf{x})| = |\mathbf{x}|$ для всех $\mathbf{x} \in V$.

Теорема 4.1. *Композиция изометрий есть изометрия.*

Док-во. Пусть отображения $h\colon U \to V$ и $f\colon V \to W$ являются изометриями. Тогда $|h(\mathbf{u})| = |\mathbf{u}|$ для всех $\mathbf{u} \in U$ и $|f(\mathbf{v})| = |\mathbf{v}|$ для всех $\mathbf{v} \in V$. Отсюда в результате простой цепочки вычислений находим, что

$$|f \circ h(\mathbf{u})| = |f(h(\mathbf{u}))| = |h(\mathbf{u})| = |\mathbf{u}|$$

для всех $\mathbf{u} \in U$. В силу формулы восстановления для квадратичных форм задание нормы определяет скалярное произведение в евклидовом пространстве. Поэтому полученное равенство доказывает изометричность композиции $f \circ h$ двух изометрий f и h. □

Определение 4.2. Биективная изометрия $f\colon V \to W$ называется *изоморфизмом евклидовых пространств.*

Теорема 4.2. *Изоморфизмы евклидовых векторных пространств обладают следующими свойствами:*

(1) *тождественное отображение* id_V *есть изоморфизм;*
(2) *композиция изоморфизмов есть изоморфизм;*
(3) *отображение, обратное изоморфизму, есть изоморфизм.*

Доказательство этой теоремы является простым следствием теоремы 4.1 и теоремы 8.1 из первой главы.

Определение 4.3. Два евклидовых векторных пространства V и W называются *изоморфными*, если существует изоморфизм $f\colon V \to W$.

Рассмотрим арифметическое векторное пространство $\mathbb{R}^n$, которое состоит из n-компонентных вектор-столбцов. Операции сложения и умножения на число таких столбцов выполняются покомпонентно (см. формулы (2.1) в первой главе). Определим на $\mathbb{R}^n$ квадратичную форму $g(\mathbf{x})$, полагая

$$g(\mathbf{x}) = (x^1)^2 + \ldots + (x^n)^2 = \sum_{i=1}^{n} (x^i)^2. \tag{4.2}$$

Форма (4.2) задает *стандартное скалярное произведение* и определяет *стандартную структуру евклидового пространства* в $\mathbb{R}^n$.

Теорема 4.3. *Всякое n-мерное евклидово векторное пространство V изоморфно пространству $\mathbb{R}^n$ со стандартным скалярным произведением, заданным формой* (4.2).

Для доказательства этой теоремы достаточно выбрать ортонормированный базис в V и рассмотреть отображение ψ, сопоставляющее вектору $\mathbf{v} \in V$ вектор-столбец из его координат (см. формулы (5.4) из первой главы).

Определение 4.4. Оператор f в евклидовом векторном пространстве V называется *ортогональным оператором*, если он биективен и задает изометрию $f\colon V \to V$ этого пространства в себя.

В силу теоремы 4.2 ортогональные операторы образуют группу, которая называется *ортогональной группой* евклидового пространства V и обозначается $\mathrm{O}(V)$. Группа $\mathrm{O}(V)$, очевидно, является подргуппой в группе автоморфизмов $\mathrm{Aut}(V)$. В случае $V = \mathbb{R}^n$, соответствующая ортогональная подгруппа, определяемая стандартным скалярным произведением в $\mathbb{R}^n$, обозначается $\mathrm{O}(n, \mathbb{R})$.

Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — ортонормированный базис в евклидовом пространстве V и пусть f — ортогональный оператор. Тогда

$$(f(\mathbf{e}_i) \,|\, f(\mathbf{e}_j)) = (\mathbf{e}_i \,|\, \mathbf{e}_j),$$

что вытекает из (4.1). Для матрицы оператора f в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$ это условие приводит к следующему соотношению:

$$\sum_{k=1}^{n} F_i^k \, F_j^k = \begin{cases} 1 & \text{при } i = j, \\ 0 & \text{при } i \neq j. \end{cases} \tag{4.3}$$

В матричной записи формула (4.3) приводит к ограничениям на матрицу F:

$$F^{\,\mathrm{tr}}\,F = 1, \qquad\qquad F^{\,-1} = F^{\,\mathrm{tr}}. \tag{4.4}$$

Соотношения (4.4) полностью идентичны соотношениям (1.13). Матрицы, удовлетворяющие таким соотношениям, как мы уже знаем, называются ортогональными. Как следствие из такого замечания сформулируем теорему.

Теорема 4.4. *Ортогональный оператор $f\colon\, V \to V$ в ортонормированном базисе $\mathbf{e}_1,\, \ldots\, ,\, \mathbf{e}_n$ евклидового векторного пространства V задается ортогональной матрицей.*

Как мы уже отмечали в § 1, детерминант ортогональной матрицы может быть равен 1 либо -1. Ортогональные операторы с единичным детерминантом в V образуют группу, которая называется *специальной ортогональной группой* евклидового векторного пространства V. Она обозначается $\mathrm{SO}(V)$. В случае $V = \mathbb{R}^n$, для такой группы принято обозначение $\mathrm{SO}(n, \mathbb{R})$.

Наиболее просто устроены операторы $f \in \mathrm{SO}(V)$ в двумерном случае: $\dim V = 2$. Если $\mathbf{e}_1$, $\mathbf{e}_2$ — ортонормированный базис в V, то из (4.3) и $\det F = 1$ легко выводится следующий вид ортогональной матрицы F:

$$F = \begin{Vmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{Vmatrix} \tag{4.5}$$

Матрица F вида (4.5) называется матрицей *двумерного поворота*, где числовой параметр φ имеет смысл угла поворота.

Рассмотрим ортогональные операторы $f \in SO(V)$ в случае $\dim V = 3$. Пусть $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ — ортонормированный базис в V. Матрица вида

$$F = \begin{Vmatrix} \cos(\varphi) & -\sin(\varphi) & 0 \\ \sin(\varphi) & \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{Vmatrix} \tag{4.6}$$

является ортогональной матрицей с единичным определителем. Задаваемое ею ортогональное преобразование, называется поворотом пространства вокруг вектора $\mathbf{e}_3$ на угол φ.

Теорема 4.5. *В трехмерном евклидовом пространстве V всякий ортогональный оператора f с единичным детерминантом* $\det f = 1$ *имеет собственное число* $\lambda = 1$.

Док-во. Рассмотрим характеристический полином оператора f. Это полином третьей степени по переменной λ с вещественными коэффициентами: $P(\lambda) = -\lambda^3 + F_1\,\lambda^2 - F_2\,\lambda + F_3$, причем $F_3 = \det f = 1$. Значения полинома нечетной степени по λ при больших положительных λ и при больших отрицательных λ отличаются знаком:

$$\lim_{\lambda\to-\infty} P(\lambda) = +\infty, \qquad\qquad \lim_{\lambda\to+\infty} P(\lambda) = -\infty.$$

Поэтому уравнение нечетной степени $P(\lambda) = 0$ с вещественными коэффициентами имеет, по меньшей мере, один вещественный корень $\lambda = \lambda_1$, который является собственным числом оператора f.

Пусть $\mathbf{e}_1 \neq 0$ — собственный вектор оператора f, отвечающий собственному числу λ_1. Тогда, применив условие изометричности $|\mathbf{v}| = |f(\mathbf{v})|$ к вектору $\mathbf{v} = \mathbf{e}_1$, получаем следующую цепочку равенств:

$$|\mathbf{e}_1| = |f(\mathbf{e}_1)| = |\lambda_1 \cdot \mathbf{e}_1| = |\lambda_1|\,|\mathbf{e}_1|.$$

Отсюда находим $|\lambda_1| = 1$. Следовательно, $\lambda_1 = 1$ или $\lambda_1 = -1$.

В случае $\lambda_1 = 1$ утверждение теоремы выполнено. Поэтому рассмотрим второй случай $\lambda_1 = -1$. Выделим из характеристического полинома $P(\lambda)$ линейный множитель. Запишем это так:

$$P(\lambda) = -\lambda^3 + F_1\,\lambda^2 - F_2\,\lambda + 1 = -(\lambda + 1)(\lambda^2 - \Phi_1\,\lambda - 1).$$

Тогда $F_1 = \Phi_1 - 1$ и $F_2 = -1 - \Phi_1$. Для нахождения оставшихся корней характеристического полинома $P(\lambda)$ рассмотрим следующее квадратное уравнение:

$$\lambda^2 - \Phi_1\,\lambda - 1 = 0.$$

Такое уравнение всегда имеет два вещественных корня λ_2 и λ_3, ибо его дискриминант положителен: $D = (\Phi_1)^2 + 4 > 0$. Из теоремы Виета имеем $\lambda_2\,\lambda_3 = -1$. Числа λ_2 и λ_3 являются собственными числами оператора f. Повторив для них те же рассуждения, что и для λ_1, получаем $|\lambda_2| = |\lambda_3| = 1$. Значит, одно из этих чисел равно минус единице, а другое — плюс единице. Таким образом, число $\lambda = 1$ есть собственное число оператора f в любом случае. Теорема доказана. $\square$

Теорема 4.5. *В трехмерном евклидовом пространстве V для всякого ортогонального оператора f с единичным детерминантом* $\det f = 1$ *существует ортонормированный базис, в котором матрица оператора f имеет вид* (4.6).

Док-во. В предположениях теоремы 4.5 оператор f имеет собственное число $\lambda_1 = 1$. Пусть $\mathbf{e}_1 \neq 0$ — собственный вектор оператора f, отвечающий этому собственному числу. Обозначим через U линейную оболочку собственного вектора $\mathbf{e}_1$ и рассмотрим ортогональное дополнение $U_\perp$. Это двумерное подпространство в трехмерном пространтстве V. Оно инвариантно относительно действия оператора f. Действительно, из условия $\mathbf{x} \in U_\perp$ вытекает $(\mathbf{x}\,|\,\mathbf{e}_1) = 0$. Запишем условие изометричности в форме соотношения (4.1) для вектора $\mathbf{x}$ и вектора $\mathbf{y} = \mathbf{e}_1$. Это дает

$$0 = (\mathbf{x}\,|\,\mathbf{e}_1) = (f(\mathbf{x})\,|\,f(\mathbf{e}_1)) = \lambda_1\,(f(\mathbf{x})\,|\,\mathbf{e}_1).$$

Но $\lambda_1 = 1$, поэтому $(f(\mathbf{x})\,|\,\mathbf{e}_1) = 0$. Значит, $f(\mathbf{x}) \in U_\perp$, что доказывает инвариантность подпространства $U_\perp$.

Рассмотрим сужение оператора f на инвариантное подпространство $U_\perp$. Такое сужение есть ортогональный оператор в двумерном пространстве $U_\perp$, причем детерминант сужения равен единице. Поэтому в ортонормированном базисе $\mathbf{e}_2$, $\mathbf{e}_3$ в $U_\perp$ матрица сужения имеет вид (4.5). Вектор $\mathbf{e}_1$ перпендикулярен векторам $\mathbf{e}_2$ и $\mathbf{e}_3$. Он может быть выбран единичным по длине. Тогда система векторов $\mathbf{e}_1,\ \mathbf{e}_2,\ \mathbf{e}_3$ образует ортонормированный базис в трехмерном пространстве V, в котором матрица оператора f имеет вид (4.6). Теорема доказана. $\square$

Результат доказанной теоремы означает, что всякий ортогональный оператор f с единичным детерминантом в трехмерном евклидовом пространстве V есть оператор поворота. Собственный вектор $\mathbf{e}_1$, отвечающий собственному числу $\lambda_1 = 1$, определяет ось поворота, а параметр φ в матрице (4.6) определяет угол поворота.

ГЛАВА VI

АФФИННЫЕ ПРОСТРАНСТВА.

§ 1. Точки и параллельные переносы. Аффинные пространства.

Пусть M — некоторое множество. Преобразованием множества M называется биективное отображение $p\colon\ M \to M$.

Определение 1.1. Пусть V — линейное векторное пространство. Скажем, что задано действие векторов из V на множестве M, если каждому элементу $\mathbf{v} \in V$ соответствует преобразование $p_{\mathbf{v}}$ множества M, причем

(1) $p_{\mathbf{0}} = \mathrm{id}_M$;
(2) $p_{\mathbf{v}+\mathbf{w}} = p_{\mathbf{v}} \circ p_{\mathbf{w}}$.

Из свойств (1) и (2) действия пространства V на множестве M вытекают еще два свойства такого действия:

(3) $p_{-\mathbf{v}} = p_{\mathbf{v}}^{-1}$;
(4) $p_{\mathbf{v}} \circ p_{\mathbf{w}} = p_{\mathbf{w}} \circ p_{\mathbf{v}}$.

Определение 1.2. Действие векторного пространства V на множестве M называется *транзитивным*, если для всяких двух элементов $A, B \in M$ существует вектор $\mathbf{v} \in V$, такой, что $p_{\mathbf{v}}(A) = B$.

Определение 1.3. Действие векторного пространства V на множестве M называется *свободным*, если для всякого элемента $A \in M$ из $p_{\mathbf{v}}(A) = A$ вытекает $\mathbf{v} = \mathbf{0}$.

Определение 1.4. Множество M называется *аффинным пространством* над полем $\mathbb{K}$, если на нем задано транзитивное и свободное действие некоторого линейного векторного пространства V над полем $\mathbb{K}$.

В силу сформулированного определения всякое аффинное пространство M неразрывно связано с некоторым линейным векторным пространством V. Поэтому аффинное пространство иногда обозначают в виде пары (M, V).

Элементы аффинного пространства принято называть *точками.* Мы будем обозначать их прописными буквами латинского алфавита: A, B, C и т. д. Само аффинное пространство иногда называют *точечным пространством.* Преобразование $p_{\mathbf{v}}$, заданное вектором $\mathbf{v} \in V$, называется *параллельным переносом* в аффинном пространстве M.

Пусть U — подпространство в V. Выберем некоторую точку $A \in M$ и определим подмножество $L \subset M$ так:

$$L = \{B \in M\colon\ \exists\,\mathbf{u}\ ((\mathbf{u} \in U)\ \ \&\ \ (B = p_{\mathbf{u}}(A)))\}. \tag{1.1}$$

Подмножество L в M, определяемое согласно (1.1), называется *линейным подмногообразием* в аффинном пространстве M. Подпространство $U \subset V$ называется *определяющим подпространством* линейного подмногообразия L. *Размерностью линейного подмногообразия* (1.1) называют размерность его определяющего подпространства. Одномерные линейные подмногообразия называются *прямыми*, двумерные — *плоскостями*. Если же размерность U на единицу меньше размерности V, то есть если $\dim(V/U) = 1$, то соответствующие линейное подмногообразие L называется *гиперплоскостью*. Линейные подмногообразия промежуточных размерностей специального названия не имеют.

Пусть $U = \langle \mathbf{a} \rangle$ — одномерное подпространство в пространстве V. Любой вектор $\mathbf{u} \in U$ имеет вид $\mathbf{u} = t \cdot \mathbf{a}$, где $t \in \mathbb{K}$. После выбора точки $A \in M$ подпространство U определяет прямую в M, проходящую через эту точку. Произвольная точка $A(t)$ прямой определяется формулой:

$$A(t) = p_{t \cdot \mathbf{a}}(A). \tag{1.2}$$

Формула (1.2) называется *параметрическим уравнением прямой* в аффинном пространстве, вектор $\mathbf{a}$ — направляющий вектор прямой, а число $t \in \mathbb{K}$ — параметр. В случае $\mathbb{K} = \mathbb{R}$, мы рассмотрим множество точек прямой (1.2), отвечающее значениям параметра t из интервала $[0, 1]$. Такое множество называется *отрезком*. Точки $A = A(0)$ и $B = A(1)$ — концевые точки отрезка. На отрезке AB можно выделить направление, выбрав одну из концевых точек в качестве *начальной точки*, а другую — в качестве *конечной точки*. Отрезок AB с выделенным направлением называется *направленным отрезком*. Направленные отрезки AB и BA считаются различными[1].

Пусть A и B — две точки аффинного пространства M. В силу транзитивности действия V на M существует вектор $\mathbf{v} \in V$, который определяет параллельный перенос $p_{\mathbf{v}}$, переводящий точку A в точку B: $p_{\mathbf{v}}(A) = B$. Докажем, что такой параллельный перенос единственен. Если $p_{\mathbf{w}}$ — другой параллельный перенос, для которого $p_{\mathbf{w}}(A) = B$, то для параллельного переноса $p_{\mathbf{w}-\mathbf{v}}$ прямым вычислением находим:

$$p_{\mathbf{w}-\mathbf{v}}(A) = p_{-\mathbf{v}} \circ p_{\mathbf{w}}(A) = p_{\mathbf{v}}^{-1}(p_{\mathbf{w}}(A)) = p_{\mathbf{v}}^{-1}(B) = A.$$

Отсюда в силу свободности действия V на M (см. определение 1.3) имеем $\mathbf{w} - \mathbf{v} = \mathbf{0}$. Значит, $\mathbf{w} = \mathbf{v}$, что доказывает единственность вектора $\mathbf{v}$, определяемого условием $p_{\mathbf{v}}(A) = B$.

Доказанный факт имеет важное следствие: если задано аффинное пространство (M, V), то вектора из V можно изображать направленными отрезками в M. Каждая пара точек $A, B \in M$ однозначно определяет вектор $\mathbf{a}$ из условия $p_{\mathbf{a}}(A) = B$. Вектор $\mathbf{a}$ может быть использован в качестве направляющего вектора прямой (1.2), проходящей через точки A и B. Направленный отрезок такой прямой с началом в точке A и с концом в точке B называется *геометрической реализацией* вектора $\mathbf{a}$. Он обозначается $\overrightarrow{AB}$.

Вектор $\mathbf{a}$ определяется своей геометрической реализацией $\overrightarrow{AB}$ однозначно. Однако, у вектора $\mathbf{a}$ может быть несколько геометрических реализаций.

[1] В случае $\mathbb{K} \neq \mathbb{R}$ под направленным отрезком $\overrightarrow{AB}$ понимается упорядоченная пара точек A и B, то есть такой отрезок не имеет внутренности.

Действительно, выбрав точку $C \neq A$, можно определить точку $D = p_{\mathbf{a}}(C)$ и построить геометрическую реализацию $\overrightarrow{CD}$ для вектора $\mathbf{a}$. Точки A и C определяют параллельный перенос $p_{\mathbf{b}}$ из условия $p_{\mathbf{b}}(A) = C$. Пользуясь свойством (4) параллельного переноса, легко видеть, что параллельный перенос $p_{\mathbf{b}}$ отображает отрезок AB в отрезок CD. Вывод: различные геометрические реализации вектора $\mathbf{a}$ отличаются параллельным переносом. Заметим, что $p_{\mathbf{a}+\mathbf{b}}(A) = D$. Поэтому $\overrightarrow{AD}$ есть геометрическая реализация вектора $\mathbf{a} + \mathbf{b}$. Отсюда вытекают известные правила сложения векторов — правило треугольника $\overrightarrow{AC} + \overrightarrow{CD} = \overrightarrow{AD}$ и правило параллелограма $\overrightarrow{AB} + \overrightarrow{AC} = \overrightarrow{AD}$.

Пусть O — некоторая фиксированная точка аффинного пространства M. Назовем ее *началом отсчета*. Тогда любая точка $A \in M$ определяет направленный отрезок $\overrightarrow{OA}$, который можно отождествить с некоторым вектором $\mathbf{r} \in V$ из условия $p_{\mathbf{r}}(O) = A$. Вектор $\mathbf{r} = \mathbf{r}_A$ называют *радиус-вектором* точки A. Если пространство V конечномерно, то в нем можно выбрать базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$, по которому можно раскладывать радиус-векторы точек $A \in M$.

Определение 1.5. *Репером* или *системой координат* в аффинном пространстве M называется совокупность, состоящая из точки $O \in M$ и базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ в V. Координаты радиус-вектора $r_A = \overrightarrow{OA}$ в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$ называются координатами точки A в системе координат $O, \mathbf{e}_1, \ldots, \mathbf{e}_n$.

Системы координат в аффинных пространствах играют ту же роль, какую базисы играют в линейных векторных пространствах. Пусть $O, \mathbf{e}_1, \ldots, \mathbf{e}_n$ и $O', \tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ — две системы координат в аффинном пространстве M. Связь между базисами $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ определяется матрицами прямого и обратного переходов: S и T. Точки O и O' определяют направленные отрезки $\overrightarrow{OO'}$ и $\overrightarrow{O'O}$. Им соответствуют вектора $\boldsymbol{\rho}, \tilde{\boldsymbol{\rho}} \in V$:

$$\rho = \overrightarrow{OO'} \qquad\qquad \tilde{\rho} = \overrightarrow{O'O}$$

Разложим первый вектор $\mathbf{a}$ по базису $\mathbf{e}_1, \ldots, \mathbf{e}_n$, а второй вектор $\tilde{\mathbf{a}}$ по второму базису $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ в векторном пространстве V:

$$\begin{aligned} \boldsymbol{\rho} &= \rho^1 \cdot \mathbf{e}_1 + \ldots + \rho^n \cdot \mathbf{e}_n, \\ \tilde{\boldsymbol{\rho}} &= \tilde{\rho}^1 \cdot \tilde{\mathbf{e}}_1 + \ldots + \tilde{\rho}^n \cdot \tilde{\mathbf{e}}_n. \end{aligned} \tag{1.3}$$

Рассмотрим произвольную точку $X \in M$. Следующие очевидные формулы

$$\overrightarrow{OX} = \overrightarrow{OO'} + \overrightarrow{O'X}, \qquad\qquad \overrightarrow{O'X} = \overrightarrow{O'O} + \overrightarrow{OX}$$

позволяют найти связь между координатами точки X в двух различных системах координат аффинного пространства:

$$x^i = \rho^i + \sum_{j=1}^{n} S^i_j\, \tilde{x}^j, \qquad\qquad \tilde{x}^i = \tilde{\rho}^i + \sum_{j=1}^{n} T^i_j\, x^j. \tag{1.4}$$

Хотя векторы $\boldsymbol{\rho}$ и $\tilde{\boldsymbol{\rho}}$ отличаются только знаком: $\boldsymbol{\rho} = -\tilde{\boldsymbol{\rho}}$, их координаты в формулах (1.4) отличаются значительно сильнее:

$$\rho^i = -\sum_{j=1}^{n} S^i_j\, \tilde{\rho}^j, \qquad\qquad \tilde{\rho}^i = -\sum_{j=1}^{n} T^i_j\, \rho^j.$$

Это связано с тем, что вектора $\boldsymbol{\rho}$ и $\tilde{\boldsymbol{\rho}}$ раскладываются по разным базисам (см. разложения (1.3)).

Изложенные выше факты из теории аффинных пространств показывают, что аффинные пространства — это правильный путь для геометризации линейной алгебры. Вектор — это алгебраический объект: вектора можно складывать, умножать на числа, можно организовывать линейные комбинации из них. В аффинном пространстве на первый план выходит понятие точки. Из точек формируются прямые, плоскости, и их многомерные обобщения — линейные подмногообразия. Здесь имеется естественное понятие параллельного переноса и можно ввести понятие параллельности линейных подмногообразий. Геометрия двумерных аффинных пространств называется *планиметрией*, геометрия трехмерных аффинных пространств называется *стереометрией.* Аффинные пространства произвольной размерности изучает математическая дисциплина называемая многомерной геометрией.

§ 2. Евклидовы точечные пространства. Квадрики в евклидовом пространстве.

Определение 2.1. Аффинное пространство (M, V) над полем вещественных чисел $\mathbb{R}$ называется *евклидовым точечным пространством*, если пространство V, действующее на M параллельными переносами, оснащено структурой евклидового векторного пространства, то есть если на V фиксирована некоторая положительно определенная квадратичная форма g.

В аффинных пространствах, рассмотренных в предыдущем параграфе, отсутствует одна очень важная компонента. В них нет понятия длины и нет понятия угла. Евклидова структура, задаваемая формой g, призвана восполнить этот недостаток. Пусть A и B две точки евклидового точечного пространства M. Они определяют вектор $\mathbf{v} \in V$ из условия $p_{\mathbf{v}}(A) = B$ (этот вектор принято отождествлять с направленным отрезком $\overrightarrow{AB}$). Норма вектора $\mathbf{v}$, определяемая квадратичной формой g, называется *длиной* отрезка АВ или же *расстоянием* между точками A и B: $|AB| = |\mathbf{v}| = \sqrt{g(\mathbf{v})}$. При этом из $|-\mathbf{v}| = |\mathbf{v}|$ вытекает $|AB| = |BA|$.

Пусть $\overrightarrow{AB}$ и $\overrightarrow{CD}$ — два направленных отрезка в евклидовом точечном пространстве. Они являются геометрическими реализациями двух векторов $\mathbf{v}$ и $\mathbf{w}$ из V. *Углом* между $\overrightarrow{AB}$ и $\overrightarrow{CD}$ называется угол между $\mathbf{v}$ и $\mathbf{w}$, определяемый по формуле (1.6) из пятой главы.

Определение 2.2. Система координат $O, \mathbf{e}_1, \ldots, \mathbf{e}_n$ в конечномерном евклидовом точечном пространстве (M, V, g) называется *декартовой прямоугольной системой координат* в M, если базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$ является ортонормированным базисом в пространстве V.

Определение 2.3. *Квадрикой* в евклидовом точечном пространстве M называется множество точек Q, координаты которых $x^1, \ldots, x^n$ в некоторой декартовой прямоугольной системе координат $O, \mathbf{e}_1, \ldots, \mathbf{e}_n$ удовлетворяют полиномиальному уравнению второй степени:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\, x^i\, x^j + 2\sum_{i=1}^{n} b^i\, x^i + c = 0. \tag{2.1}$$

Определение квадрики не является инвариантным. Оно привязано к некоторой декартовой прямоугольной системе координат $O, \mathbf{e}_1, \ldots, \mathbf{e}_n$. Однако, переход в другую систему координат приводит к линейной замене переменных в уравнении (2.1) (см. формулы (1.4)). Такая замена переменных меняет коэффициенты полинома (2.1), но не меняет самого вида этого уравнения. Квадрика остается квадрикой в любой системе координат.

Пусть $O', \tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ — некоторая другая декартова прямоугольная система координат в M. Рассмотрим переход из $O, \mathbf{e}_1, \ldots, \mathbf{e}_n$ в $O', \tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$. Матрицы перехода S и T в (1.4) оказываются ортогональными матрицами (см. формулы (1.13) из пятой главы). Вычислим коэффициенты уравнения квадрики в новой системе координат. Прямая подстановка первой формулы (1.4) в уравнение (2.1) дает:

$$\tilde{a}_{qp} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\, S^i_q\, S^j_p, \tag{2.2}$$

$$\tilde{b}_q = \sum_{i=1}^{n} b_i\, S^i_q + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\, \rho^j\, S^i_q, \tag{2.3}$$

$$\tilde{c} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\, \rho^i\, \rho^j + \sum_{i=1}^{n} b_i\, \rho^i + c. \tag{2.4}$$

Задача приведения квадрики к каноническому виду формулируется как задача о нахождении декартовой прямоугольной системы координат, в которой уравнение (2.1) имеет максимально простой канонический вид.

Формула (2.2) совпадает с формулой преобразования координат квадратичной формы при замене базиса (см. формулу (1.11) из четвертой главы). Вывод: с каждой квадрикой в M связана квадратичная форма a в V. Форма a, определяемая матрицей a_{ij} в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$ называется *главной квадратичной формой* квадрики (2.1).

Рассмотрим ассоциированный оператор f_a, заданный главной квадратичной формой (см. формулу (3.9) из пятой главы). Оператор f_a является самосопряженным оператором, он определяет разложение

$$V = \operatorname{Ker} f_a \oplus \operatorname{Im} f_a \tag{2.5}$$

на два ортогональных друг другу подпространства (см. формулу (3.14) из пятой главы). Матрица оператора f_a определяется формулой

$$F^i_j = \sum_{k=1}^{n} g^{ik}\, a_{kj}, \tag{2.6}$$

где g^{ik} — матрица обратная для матрицы Грама g_{kj} базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Рассмотрим также вектор $\mathbf{b}$, определив его координаты по формуле

$$b^i = \sum_{k=1}^{n} g^{ik}\, b_k. \tag{2.7}$$

Вектор $\mathbf{b}$, определенный своими координатами (2.7), существенно привязан к системе координат $O,\, \mathbf{e}_1,\, \ldots\,,\, \mathbf{e}_n$. Это связано с тем, что формула (2.3) отличается от правила преобразования координат ковектора при замене базиса (см. формулы (2.4) в третьей главе). Перепишем (2.3) в форме

$$\tilde{b}_q = \sum_{i=1}^{n} S_q^i \left(b_i + \sum_{j=1}^{n} a_{ij}\, \rho^j \right). \tag{2.8}$$

После этого рассмотрим разложение вектора $\mathbf{b}$ в сумму двух векторов $\mathbf{b} = \mathbf{b}^{(1)} + \mathbf{b}^{(2)}$ в соответствии с разложением (2.5) для пространства V. Это определяет разложение $b_i = b_i^{(1)} + b_i^{(2)}$, где числа $b_i^{(1)}$ при замене координат преобразуются по стандартному правилу

$$\tilde{b}_q^{(1)} = \sum_{i=1}^{n} S_q^i\, b_i^{(1)}. \tag{2.9}$$

Вектор $\mathbf{b}^{(2)}$ в разложении $\mathbf{b} = \mathbf{b}^{(1)} + \mathbf{b}^{(2)}$ можно занулить за счет правильного выбора системы отсчета. Для этого выберем вектор сдвига $\boldsymbol{\rho} = \overrightarrow{OO'}$ из условия $\mathbf{b}^{(2)} = -f_a(\boldsymbol{\rho})$. Такой вектор $\boldsymbol{\rho}$ существует, ибо $\mathbf{b}^{(2)} \in \operatorname{Im} f_a$. Для него

$$b_i^{(2)} + \sum_{j=1}^{n} a_{ij}\, \rho^j = 0, \tag{2.10}$$

что вытекает из (2.6) и (2.7). Подставив (2.10) в (2.8), получим следующие соотношения: $\tilde{\mathbf{b}}^{(2)} = 0$ и $\tilde{\mathbf{b}} = \tilde{\mathbf{b}}^{(1)}$.

Соотношения (2.9) показывают, что числа $b_i^{(1)}$ невозможно занулить (если, конечно, они не равны нулю с самого начала). Эти числа определяют вектор $\mathbf{b}^{(1)} \in \operatorname{Ker} f_a$, не зависящий от выбора системы координат. В итоге мы доказали следующую теорему о квадриках.

Теорема 2.1. *Со всякой квадрикой в евклидовом точечном пространстве (M, V, g) связан некоторый самосопряженный оператор f и вектор $b \in \operatorname{Ker} f$, такие, что в некоторой ортонормированной системе координат радиус-векторы $\mathbf{r}$ точек квадрики удовлетворяют уравнению*

$$(f(\mathbf{r}) \,|\, \mathbf{r}) + 2\,(\mathbf{b} \,|\, \mathbf{r}) + c = 0. \tag{2.11}$$

Самосопряженный оператор f задает главную квадратичную часть в уравнении квадрики (2.11). Он определяет деление квадрик на два основных типа:

(1) *невырожденные квадрики*, соответствующие случаю, когда ядро оператора f тривиально: $\operatorname{Ker} f = \{\mathbf{0}\}$,

(2) *вырожденные квадрики*, соответствующие случаю, когда ядро оператора f отлично от нуля: $\operatorname{Ker} f \neq \{\mathbf{0}\}$.

Для невырожденных квадрик вектор $\mathbf{b}$ в (2.11) равен нулю. Поэтому невырожденные квадрики делятся на три типа:

(1) *эллиптический тип*, когда $c \neq 0$ и квадратичная форма $(f(\mathbf{x}) \,|\, \mathbf{x})$ положительно определена либо она может быть сделана положительно определенной в результате замены f на $-f$,

(2) *гиперболический тип*, когда $c \neq 0$ и форма $(f(\mathbf{x}) \,|\, \mathbf{x})$ не является знакоопределенной,
(3) *конический тип*, когда $c = 0$.

Вырожденные квадрики принято подразделять на два типа:

(1) *параболический тип*, когда $\dim \operatorname{Ker} f = 1$ и $\mathbf{b} \neq \mathbf{0}$,
(2) *цилиндрический тип*, когда $\dim \operatorname{Ker} f > 1$ или $\mathbf{b} = \mathbf{0}$.

Уравнение невырожденной квадрики эллиптического типа может быть приведено к следующему каноническому виду

$$\frac{(x^1)^2}{(a_1)^2} + \ldots + \frac{(x^n)^2}{(a_n)^2} = \pm 1.$$

Каноническое уравнение невырожденной квадрики гиперболического типа отличается от предыдущего случая только знаками:

$$\frac{(x^1)^2}{(a_1)^2} \pm \ldots \pm \frac{(x^n)^2}{(a_n)^2} = \pm 1.$$

В коническом случае уравнение невырожденной квадрики может быть приведено к следующему каноническому виду:

$$\frac{(x^1)^2}{(a_1)^2} \pm \ldots \pm \frac{(x^n)^2}{(a_n)^2} = 0.$$

Уравнение вырожденной квадрики параболического типа приводится к следующему каноническому виду:

$$\frac{(x^1)^2}{(a_1)^2} \pm \ldots \pm \frac{(x^{n-1})^2}{(a_{n-1})^2} = 2\, x^n.$$

Если $\dim V > 1$, то в каноническое уравнении квадрики цилиндрического типа не входит одна из переменных. Это позволяет уменьшить размерность на единицу. Редуцированная квадрика может принадлежать любому из перечисленных типов. Если она вновь оказалась квадрикой цилиндрического типа, то редукцию размерности можно повторить. Процесс редуцирования может завершиться в промежуточной размерности на каком-то типе, отличном от цилиндрического. Если этого не случится, то процесс редукции доходит до размерности $\dim M = 1$.

В одномерном евклидовом точечном пространстве квадрик цилиндрического типа нет. Поэтому квадрики цилиндрического типа — это квадрики которые принадлежат одному из нецилиндрических типов после редукции размерности.

СПИСОК ЛИТЕРАТУРЫ.

1. Курош А. Г. *Курс общей алгебры*, издательство «Наука», Москва.
2. Шарипов Р. А. *Курс дифференциальной геометрии*[1], издание БашГосУниверситета, Уфа, 1996.
3. Шарипов Р. А. *Классическая электродинамика и теория относительности*, издание БашГосУниверситета, Уфа, 1997; см. также physics/0311011 в электронном архиве http://arXiv.org.
4. Кострикин А. И. *Введение в алгебру*, издательство «Наука», Москва, 1977.
5. Беклемишев Д. В. *Курс аналитической геометрии и линейной алгебры*, издательство «Наука», Москва, 1985.
6. Кудрявцев Л. Д. *Курс математического анализа, Т. 1 и Т. 2*, издательство «Высшая школа», Москва, 1985.
7. Шарипов Р. А. *Быстрое введение в тензорный анализ*, бесплатное интернет-издание math.HO/0403252 в электронном архиве http://arXiv.org, 2004.

[1] Ссылки [2] и [2] добавлены в 1998 году, ссылка [7] добалена в 2004 году.

Шарипов Руслан Абдулович
http://www.geocities.com/r-sharipov

КУРС ЛИНЕЙНОЙ АЛГЕБРЫ
И МНОГОМЕРНОЙ ГЕОМЕТРИИ

Учебное пособие

3 К 3 ЛР № 020259 от 30.10.1991

Подписано в печать 23.05.96. Формат 60×84/16. Бумага типографская № 2. Компьютерный набор. Отпечатано на ризографе. Усл. печ. л. 13,0. Уч.-изд. л. 10,8. Тираж 100. Заказ 249

Редакционно-издательский отдел Башкирского университета
Ротапринт Башкирского университета, 450074, Уфа, ул. Фрунзе, 32.