Д. И. МОЛДАВАНСКИЙ

ЛИНЕЙНАЯ АЛГЕБРА И ГЕОМЕТРИЯ

Учебное пособие

Иваново, 2003

Данное пособие написано по материалам лекций по одноименному курсу, который в течение ряда читался автором для студентов математического факультета Ивановского государственного университета.

СОДЕРЖАНИЕ

ГЛАВА I
ЛИНЕЙНЫЕ ПРОСТРАНСТВА И ЛИНЕЙНЫЕ ОТОБРАЖЕНИЯ

§ 1. ЛИНЕЙНЫЕ ПРОСТРАНСТВА И ИХ ПОДПРОСТРАНСТВА

Раздел математики, называемый линейной алгеброй, в конечном счете занимается изучением всего двух понятий — линейное пространство и линейное отображение. В этом параграфе мы знакомимся с первым из них.

Определение 1. *Пусть K — некоторое поле. Непустое множество L называется линейным (или векторным) пространством над полем K, если на этом множестве заданы внутренняя операция сложения $+$ и внешняя операция умножения (слева) на элементы из поля K так, что выполнены следующие свойства (или аксиомы):*

(1) $(\forall a, b \in L)(a + b = b + a)$;
(2) $(\forall a, b, c \in L)((a + b) + c = a + (b + c))$;
(3) $(\forall a, b \in L)(\exists x \in L)(a + x = b)$;
(4) $(\forall \alpha \in K)(\forall a, b \in L)(\alpha(a + b) = \alpha a + \alpha b)$;
(5) $(\forall \alpha, \beta \in K)(\forall a \in L)((\alpha + \beta)a = \alpha a + \beta a)$;
(6) $(\forall \alpha, \beta \in K)(\forall a \in L)((\alpha\beta)a = \alpha(\beta a)$;
(7) $(\forall a \in L)(1a = a)$.

Элементы множества L будут называться *векторами* и обозначаться малыми латинскими буквами; элементы поля K будем называть *скалярами* и обозначать малыми греческими буквами. Мы будем часто говорить просто о пространстве, опуская для краткости слово "линейное". Будет опускаться и упоминание о поле K, если из контекста ясно, над каким полем рассматривается данное линейное пространство.

В следующем предложении собраны простейшие свойства линейных пространств, вытекающие непосредственно из определения.

Предложение 1. *Пусть L — линейное пространство над полем K. Тогда*

(1) *Группоид $(L, +)$ является абелевой группой. В частности, в множестве L содержится единственный элемент, называемый нулевым вектором и обозначаемый символом 0, который (однозначно) определяется тем свойством, что для любого вектора $a \in L$ выполнено равенство $a + 0 = a$; кроме того, для любого вектора $a \in L$ существует единственный вектор, обозначаемый символом $-a$, называемый противоположным к вектору a и (однозначно) определямый тем свойством, что $a + (-a) = 0$.*
(2) *Для любых $\alpha \in K$ и $a \in L$ имеют место равенства $\alpha 0 = 0$ и $0a = 0$. Обратно, если $\alpha a = 0$, то или $\alpha = 0$, или $a = 0$.*
(3) *Для любых $\alpha \in K$ и $a \in L$ имеют место равенства $\alpha(-a) = -\alpha a$ и $(-\alpha)a = -\alpha a$.*
(4) *Полагая, как обычно, $a - b = a + (-b)$ $(a, b \in L)$, имеем для произвольных векторов $a, b \in L$ и произвольных скаляров $\alpha, \beta \in K$ справедливость равенств $\alpha(a - b) = \alpha a - \alpha b$ и $(\alpha - \beta)a = \alpha a - \beta a$.*

Доказательство. Все утверждения пункта (1) почти очевидны. В самом деле, вторая и третья аксиомы в определении линейного пространства означают в точности то, что группоид $(L,+)$ является группой, а первая аксиома говорит о том, что эта группа абелева. Остальные утверждения пункта (1) просто выражают известные свойства группы.

Для произвольного $\alpha \in K$ имеем

$$\alpha 0 = \alpha(0+0) = \alpha 0 + \alpha 0.$$

Прибавив к обеим частям равенства $\alpha 0 + \alpha 0 = \alpha 0$ элемент, противоположный к вектору $\alpha 0$, получаем $\alpha 0 = 0$. Равенство $0a = 0$ доказывается аналогично. Обратно, если $\alpha a = 0$ и $\alpha \neq 0$, то умножив обе части равенства $\alpha a = 0$ на скаляр α^{-1}, получаем $\alpha^{-1}(\alpha a) = \alpha^{-1}0$. Так как в силу аксиом (6) и (7) $\alpha^{-1}(\alpha a) = (\alpha^{-1}\alpha)a = 1a = a$ и по доказанному $\alpha^{-1}0 = 0$, отсюда следует, что $a = 0$, и все утверждения пункта (2) доказаны. Поскольку, далее,

$$\alpha a + \alpha(-a) = \alpha(a + (-a)) = \alpha 0 = 0 \quad \text{и} \quad \alpha a + (-\alpha)a = (\alpha + (-\alpha))a = 0a = 0,$$

получаем $\alpha(-a) = -\alpha a$ и $(-\alpha)a = -\alpha a$. Утверждения пункта (3) доказаны, а утверждения пункта (4) следуют теперь непосредственно из них и из определения разности векторов. $\square$

Приведем несколько примеров линейных пространств. Эти примеры в дальнейшем помогут нам лучше усвоить свойства линейных пространств и связанных с ними понятий. Кроме того, они достаточно убедительно свидетельствуют о том, что структура линейного пространства возникает при рассмотрении самых разных математических объектов, тем самым подтверждая универсальность развиваемой здесь теории.

1. Пусть V^2 — множество всех направленных отрезков евклидовой плоскости, начало которых совпадает с фиксированной точкой O этой плоскости. Обычные операции сложения отрезков (по правилу параллелограмма) и умножения отрезка на действительное число удовлетворяют, как легко видеть, всем аксиомам из определения 1. Таким образом, V^2 является линейным пространством над полем $\mathbb{R}$ действительных чисел. Аналогично определяется пространство V^3 всех направленных отрезков с общим началом трехмерного евклидова пространства.

2. Пусть K — произвольное поле и $n \geq 1$ — фиксированное натуральное число. Пусть K^n обозначает n-ую декартову степень множества K, т.е. K^n состоит из всевозможных последовательностей длины n элементов из K:

$$K^n = \left\{(\alpha_1, \alpha_2, \ldots, \alpha_n) \mid \alpha_i \in K (i = 1, 2, \ldots, n)\right\}.$$

Определим на этом множестве операции сложения и внешнего умножения на элементы из K покомпонентно: если $a, b \in K^n$, $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, $b = (\beta_1, \beta_2, \ldots, \beta_n)$, и $\lambda \in K$, полагаем

$$a + b = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n) \quad \text{и} \quad \lambda a = (\lambda\alpha_1, \lambda\alpha_2, \ldots, \lambda\alpha_n).$$

Непосредственная проверка показывает, что эти операции удовлетворяют всем аксиомам из определения 1, т.е. K^n является линейным пространством над полем K. Это пространство мы будем называть n-мерным координатным (или арифметическим) линейным пространством над полем K.

3. Пусть поле K является подполем поля L. Тогда на множестве L определены операции сложения и умножения. Если первую из них считать внутренней операцией на L, а произведение элементов из K на элементы из L рассматривать как внешнюю операцию, определенную на множестве L, то выполнимость аксиом из определения 1 становится очевидной, так как они в этом случае выражают определенные свойства поля. Таким образом, произвольное поле L можно считать линейным пространством над его подполем K. Конкретно, поле $\mathbb{R}$ действительных чисел является линейным пространством над полем $\mathbb{Q}$ рациональных чисел. Линейным пространством над полем $\mathbb{Q}$ является и поле $\mathbb{C}$ комплексных чисел; поле $\mathbb{C}$ является также линейным пространством над полем $\mathbb{R}$.

4. Предыдущий пример допускает следующее очевидное обобщение: если L — кольцо с единицей и K — подполе кольца L, то L является линейным пространством над полем K. В частности, множество $L = K[x]$ всех многочленов от переменной x над полем K является линейным пространством над K.

5. Пусть снова K — произвольное поле и $m, n \geq 1$ — фиксированные натуральные числа. Множество $M_{mn}(K)$ всех $m \times n$-матриц с элементами из поля K относительно обычных операций сложения матриц и умножения элементов из K на матрицу является линейным пространством над полем K. (Договоримся при $m = n$ писать $M_n(K)$ вместо $M_{nn}(K)$.)

6. Пусть M — непустое множество, K — поле и $F(M, K)$ обозначает множество всевозможных функций, определенных на M и принимающих значение в поле K. Определим на множестве $F(M, K)$ внутреннюю операцию сложения и внешнюю операцию умножения на элементы из K, полагая для $f, g \in F(M, K)$ и $\lambda \in K$

$$(f + g)(x) = f(x) + g(x) \quad \text{и} \quad (\lambda f)(x) = \lambda f(x) \qquad (x \in M).$$

Справедливость здесь всех аксиом линейного пространства из определения 1 проверяется непосредственно.

Введем теперь понятие, которое будет постоянно использоваться. Пусть L — линейное пространство над полем K, $n \geq 1$ — натуральное число и $a_1, a_2, \ldots, a_n$ и $\alpha_1, \alpha_2, \ldots, \alpha_n$ — две системы одинаковой длины n элементов из множества L и поля K соответственно. Выражение вида

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n \tag{1}$$

называется линейной комбинацией системы векторов $a_1, a_2, \ldots, a_n$ с коэффициентами $\alpha_1, \alpha_2, \ldots, \alpha_n$. Разумеется, выражение (1) совпадает с некоторым вектором $b \in L$, и этот вектор мы будем называть значением линейной комбинации (1). В большинстве случаев нет необходимости различать линейную комбинацию и ее значение. Тем не менее, следует иметь в виду, что значения разных линейных комбинаций могут совпадать. Например, для векторов $a_1 = (1,1,1)$, $a_2 = (1,3,-1)$, $a_3 = (-1,-2,0)$ пространства $\mathbb{Q}^3$ значением каждой из формально различных линейных комбинаций $0a_1 + 0a_2 + 0a_3$ и $1a_1 + 1a_2 + 2a_3$ является нулевой вектор (0,0,0).

Линейную комбинацию (1) можно записывать с использованием знака суммирования в виде $\sum_{i=1}^{n} \alpha_i a_i$. При этом из аксиом линейного пространства вытекают обычные правила обращения со знаком суммирования; в частности, постоянный (т.е. независящий от переменной суммирования) множитель можно вносить под знак суммирования и выносить из под него, т. е. для любого скаляра α имеет место равенство (значений)

$$\alpha \sum_{i=1}^{n} \alpha_i a_i = \sum_{i=1}^{n} \alpha \alpha_i a_i.$$

Мы ввели определение линейной комбинации для конечной непустой системы векторов. Договоримся, кроме того, значением линейной комбинации пустой системы векторов считать нулевой вектор; использование этого соглашения позволит упростить ряд формулировок. Заметим еще, что мы могли бы ввести формально линейную комбинацию и бесконечной системы векторов, но не делаем этого, поскольку в общем случае не существует никакого разумного способа определить значение такой линейной комбинации.

Одной из центральных задач при изучении линейных пространств (как, впрочем, и при изучении других алгебраических структур) является описание их "правильных" частей, т. е. таких частей линейных пространств, которые сами являются линейными пространствами. Уточнением интуитивного представления о правильной части линейного пространства является понятие линейного подпространства:

Определение 2. *Пусть L — линейное пространство над полем K. Непустое подмножество M множества L называется линейным подпространством пространства L, если выполнены следующие условия:*

(1) $(\forall a, b \in L)(a \in M \wedge b \in M \Rightarrow a + b \in M)$;
(2) $(\forall a \in L)(\forall \lambda \in K)(a \in M \Rightarrow \lambda a \in M)$.

Говоря более содержательно, первое из условий в определении подпространства требует, чтобы в подмножество M вместе с любыми двумя векторами входила и их сумма; мы будем называть это условие замкнутостью множества M относительно сложения. Второе условие требует, чтобы вместе с каждым вектором подмножество M содержало и произведение его на любой скаляр; это условие называют замкнутостью множества M относительно умножения на скаляры. Таким образом, по определению, подмножество M линейного пространства L является подпространством пространства L, если оно непусто и замкнуто относительно сложения и умножения на скаляры. Отсюда сразу же следует, что само линейное пространство L мы должны считать подпространством этого пространства. Очевидно, кроме того, что в произвольном линейном пространстве подмножество $\{0\}$, состоящее лишь из одного нулевого вектора, является подпространством; это подпространство мы будем называть нулевым.

Таким образом, в каждом линейном пространстве L имеется по меньшей мере два подпространства: само пространство L и нулевое подпространство (которые, впрочем, совпадают, если пространство L само является нулевым, т. е. содержит лишь один вектор). Приведем другие примеры подпространств в тех линейных пространствах, которые перечислены после предложения 1. Читателю рекомендуется самостоятельно убедиться в том, что указанные подмножества действительно удовлетворяют требованиям определения 2.

1. В линейном пространстве V^2 (или V^3) множество всех отрезков, принадлежащих некоторой фиксированной прямой, проходящей через общее начало O, является подпространством. Аналогично, подпространством пространства V^3 является множество всех отрезков плоскости, проходящей через общее начало O.

2. Пусть K — поле. Рассмотрим систему

$$\begin{cases} \alpha_{11}\,x_1 + \alpha_{12}\,x_2 + \cdots + \alpha_{1n}\,x_n = 0 \\ \alpha_{21}\,x_1 + \alpha_{22}\,x_2 + \cdots + \alpha_{2n}\,x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \cdots + \alpha_{mn}x_n = 0 \end{cases}$$

линейных однородных уравнений от n неизвестных, коэффициенты α_{ij} которой являются элементами поля K. Произвольное решение $(\beta_1, \beta_2, \ldots, \beta_n)$ этой системы (где $\beta_i \in K$) является элементом n-мерного координатного линейного пространства K^n. Множество всех решений данной системы уравнений является подпространством пространства K^n.

3. Множество $\mathbb{R}$ действительных чисел является подпространством линейного пространства $\mathbb{C}$ над полем $\mathbb{Q}$.

Множество $M = \left\{\alpha + \beta\sqrt{2} \;\middle|\; \alpha, \beta \in \mathbb{Q}\right\}$ является подпространством линейного пространства $\mathbb{R}$ над полем $\mathbb{Q}$.

4. В пространстве $K[x]$ многочленов от переменной x над полем K для любого $n \geq 0$ подпространством будет множество

$$K_n[x] = \left\{f(x) \in K[x] \;\middle|\; f(x) = 0 \vee \deg f(x) \leq n\right\},$$

состоящее из нулевого многочлена и всех многочленов степени, не превосходящей числа n.

5. Напомним, что квадратная матрица A называется симметричной, если $A^{\mathrm{t}} = A$, и кососимметричной, если $A^{\mathrm{t}} = -A$ (здесь и ниже A^{t} обозначает матрицу, транспонированную к матрице A). Множество всех симметричных матриц и множество всех кососимметричных матриц являются примерами подпространств линейного пространства $M_n(K)$ квадратных матриц порядка n над полем K. Заметим, что эти подпространства совпадают тогда и только тогда, когда характеристика поля K равна 2.

6. Рассмотрим пространство $F(M, K)$ функций, определенных на множестве M и принимающих значения в поле K. Множество всех функций, почти всюду равных

нулю (т.е. принимающих нулевое значение во всех точках из M, кроме быть может конечного их числа) является подпространством этого пространства.

Пусть $\alpha < \beta$ — произвольные действительные числа. Множество $C_{[\alpha,\,\beta]}$ всех функций, непрерывных на отрезке $[\alpha,\,\beta]$, является подпространством пространства $F(\mathbb{R},\mathbb{R})$ всех действительных функций от одной переменной.

Переходя к изучению свойств подпространств линейных пространств, заметим, прежде всего, что два условия в определении подпространства можно заменить одним:

Предложение 2. *Пусть L — линейное пространство над полем K. Непустое подмножество M множества L является подпространством пространства L тогда и только тогда, когда M содержит значение любой линейной комбинации произвольной системы двух векторов из множества M, т. е. тогда и только тогда, когда множество M удовлетворяет условию*

$$(\forall a, b \in L)(\forall \alpha, \beta \in K)(a \in M \wedge b \in M \Rightarrow \alpha a + \beta b \in M). \tag{2}$$

Доказательство. Пусть сначала подмножество M является подпространством пространства L, пусть a и b — произвольные элементы из M, α и β — произвольные скаляры из K. Тогда в силу замкнутости подпространства относительно умножения на скаляры векторы αa и βb содержатся в M, а в силу замкнутости подпространства относительно сложения и вектор $\alpha a + \beta b$ входит в M.

Обратно, предположим, что непустое подмножество M пространства L удовлетворяет условию (2). Тогда для любых $a, b \in M$ и $\alpha \in K$ имеем

$$a + b = 1a + 1b \in M \quad \text{и} \quad \alpha a = \alpha a + 0a \in M.$$

Таким образом, подмножество M оказывается замкнутым относительно сложения и умножения на скаляры, т.е. является подпространством пространства L. $\square$

Отметим, далее, следующие простые свойства подпространств:

Предложение 3. *Пусть M — подпространство линейного пространства L. Тогда*

(1) *нулевой вектор пространства L содержится в M, и вместе с каждым вектором a подпространство M содержит и противоположный ему вектор $-a$;*
(2) *подпространство M содержит значение любой линейной комбинации любой системы векторов из M.*

Доказательство. Утверждения пункта (1) почти очевидны: поскольку подмножество M непусто, существует вектор b, входящий в M, а потому и вектор $0 = 0b$ принадлежит множеству M. Если a — произвольный вектор из M, то и вектор $-a = (-1)a$ должен содержаться в M.

В пункте (2) утверждается, что для любых векторов $a_1, a_2, \ldots, a_n \in M$ и произвольных скаляров $\alpha_1, \alpha_2, \ldots, \alpha_n$ вектор

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n$$

является элементом множества M. Заметим, что при $n = 1$ это утверждение совпадает с одним из условий из определения подпространства, а при $n = 2$ оно доказано в предложении 2. Справедливость его в общем случае доказывается с помощью очевидных индуктивных рассуждений. Действительно, если при $n > 1$ предположить, что вектор $\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_{n-1} a_{n-1}$ принадлежит множеству M, то поскольку $\alpha_n a_n \in M$, утверждение о принадлежности этому множеству вектора

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n = (\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_{n-1} a_{n-1}) + \alpha_n a_n$$

следует из замкнутости подпространства M относительно сложения. $\square$

Определение 3. *Пусть L — линейное пространство над полем K и M — произвольное подмножество множества L. Линейной оболочкой множества M называется совокупность $l(M)$ значений всевозможных линейных комбинаций всевозможных систем векторов из множества M.*

Таким образом, если множество M непусто, то множество векторов, названное в этом определении его линейной оболочкой, может быть записано в следующем виде:

$$l(M) = \left\{ \sum_{i=1}^{n} \alpha_i a_i \;\middle|\; a_i \in M, \alpha_i \in K \ (i = 1, 2, \ldots, n), n = 1, 2, \ldots \right\}.$$

Если же множество M пусто, то ввиду соглашения, принятого выше, его линейная оболочка состоит из единственного элемента — нулевого вектора.

Теорема 1. *Линейная оболочка $l(M)$ произвольного подмножества M линейного пространства L является подпространством пространства L. Более того, $l(M)$ является наименьшим из подпространств пространства L, содержащих множество M.*

Доказательство. Говоря более подробно, в теореме утверждается, что
1) $l(M)$ является подпространством пространства L;
2) множество M содержится в множестве $l(M)$;
3) произвольное подпространство пространства L, содержащее множество M, содержит и его линейную оболочку $l(M)$.

Если множество M пусто (и потому $l(M) = \{0\}$) все эти утверждения очевидны. Но и в случае, когда M непусто, каждое из них доказывается без труда. Произвольный элемент $a \in M$ является значением линейной комбинации $1a$ системы, состоящей из одного вектора, взятого из M, и потому входит в множество $l(M)$. Таким образом, $M \subseteq l(M)$, и справедливость утверждения 2) установлена. В частности, мы видим, что и множество $l(M)$ непусто. Пусть a и b — произвольные элементы множества $l(M)$. Это означает, что

$$a = \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_m a_m \quad \text{и} \quad b = \beta_1 b_1 + \beta_2 b_2 + \cdots + \beta_n b_n$$

для подходящих скаляров $\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n$ из поля K и векторов $a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_n$ из множества M. Тогда и вектор

$$a + b = \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_m a_m + \beta_1 b_1 + \beta_2 b_2 + \cdots + \beta_n b_n,$$

будучи линейной комбинацией системы элементов из M, должен входить в $l(M)$. Аналогично, для любого $\lambda \in K$ вектор

$$\lambda a = (\lambda\alpha_1)a_1 + (\lambda\alpha_2)a_2 + \cdots + (\lambda\alpha_m)a_m$$

содержится в подмножестве $l(M)$. Таким образом, это (непустое) подмножество пространства L замкнуто относительно сложения и умножения на скаляры и потому является подпространством пространства L. Наконец, пусть N — произвольное подпространство пространства L такое, что $M \subseteq N$. Пусть a — произвольный вектор из $l(M)$ и пусть, как и выше, $a = \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_m a_m$, где $a_1, a_2, \ldots, a_m \in M$. Так как векторы $a_1, a_2, \ldots, a_m$ принадлежат подпространству N, в силу предложения 3 и вектор a лежит в подпространстве N. Следовательно, $l(M) \subseteq N$, и теорема доказана. $\square$

Отметим два следствия этой теоремы.

Следствие 1. *Пусть $M = \{a_1, a_2, \ldots, a_n\}$ — конечное подмножество линейного пространства L. Тогда*

$$l(M) = \left\{\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n \;\middle|\; \alpha_i \in K (i = 1, 2, \ldots, n)\right\}. \tag{3}$$

Следует пояснить, что здесь утверждается, что если множество M конечно (и организовано в систему), то для перечисления векторов, входящих в линейную оболочку множества M, достаточно взять значения всевозможных линейных комбинаций лишь одной этой системы, тогда как исходное определение линейной оболочки, требует перечислять линейные комбинации всевозможных систем, составленных из элементов множества M, т. е. всех систем, составленных из одного вектора множества M, из двух векторов, из трех векторов и т. д.

Для доказательства следствия обозначим через N множество, стоящее в правой части равенства (3). В силу определения линейной оболочки очевидно включение $N \subseteq l(M)$. С другой стороны, для любых векторов $a, b \in N$, $a = \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n$, $b = \beta_1 a_1 + \beta_2 a_2 + \cdots + \beta_n a_n$, и скаляра $\lambda \in K$ векторы

$$a + b = (\alpha_1 + \beta_1)a_1 + (\alpha_2 + \beta_2)a_2 + \cdots + (\alpha_n + \beta_n)a_n$$

и

$$\lambda a = (\lambda\alpha_1)a_1 + (\lambda\alpha_2)a_2 + \cdots + (\lambda\alpha_n)a_n$$

входят в N, так что N является подпространством пространства L. Кроме того, очевидное равенство $a_1 = 1a_1 + 0a_2 + \cdots + 0a_n$ говорит о том, что вектор a_1 входит в подпространство N, и аналогично для остальных векторов множества M, так что $M \subseteq N$. Из теоремы 1 теперь следует, что $l(M) \subseteq N$, и потому $l(M) = N$. $\square$

Следствие 2. *Подмножество M линейного пространства L является подпространством этого пространства тогда и только тогда, когда $M = l(M)$.*

Достаточность сформулированного условия очевидна. Докажем необходимость. Если M является подпространством пространства L, то ввиду очевидного включения $M \subseteq M$ из теоремы 1 следует, что $l(M) \subseteq M$. Так как противоположное включение утверждается той же теоремой 1, имеем $M = l(M)$. $\square$

Договоримся подпространство $l(M)$ называть также подпространством, порожденным множеством (или системой) M. При этом множество (или система) M будет называться множеством (или системой) порождающих подпространства $l(M)$. Иначе говоря, если A — подпространство линейного пространства L, то множество (или система) M векторов пространства L называется множеством (или системой) порождающих подпространства A, если $A = l(M)$. Если $A = L$, мы говорим о множестве или системе порождающих пространства L. Приведем несколько примеров.

1. Пусть a — ненулевой вектор линейного пространства V^2. Тогда подпространство $l(a)$ (здесь и в других аналогичных случаях мы используем сокращение $l(a)$ обозначения $l(\{a\})$, требуемого определением) состоит из всевозможных отрезков вида αa, где $\alpha \in \mathbb{R}$, т.е. совпадает с множеством отрезков прямой, определяемой отрезком a. Два направленных отрезка, не лежащих на одной прямой, составляют систему порождающих всего пространства V^2.

2. Рассмотрим следующую систему векторов n-мерного координатного пространства K^n над полем K:

$$\begin{aligned}
e_1 &= (1, 0, 0, \dots, 0, 0)\\
e_2 &= (0, 1, 0, \dots, 0, 0)\\
e_3 &= (0, 0, 1, \dots, 0, 0)\\
&\dots\dots\dots\dots\dots\dots\dots\\
e_{n-1} &= (0, 0, 0, \dots, 1, 0)\\
e_n &= (0, 0, 0, \dots, 0, 1).
\end{aligned}$$

Так как для любого вектора $a \in K^n$, $a = (\alpha_1, \alpha_2, \dots, \alpha_n)$, имеет место очевидное равенство

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n,$$

то система $e_1, e_2, \dots, e_n$ является системой порождающих пространства K^n.

3. Тот факт, что любое комплексное число записывается в виде $\alpha + \beta i = \alpha 1 + \beta i$, где α и β — действительные числа, означает, что числа 1 и i составляют систему порождающих линейного пространства $\mathbb{C}$ над полем $\mathbb{R}$.

4. В пространстве $K[x]$ линейная оболочка системы $1, x, x^2, \dots, x^n$ (где $n \geq 0$) совпадает с подпространством $K_n[x]$. Очевидным примером системы порождающих всего пространства $K[x]$ служит бесконечная система элементов $1, x, x^2, \dots$.

Линейное пространство L над полем K будем называть *конечно порожденным*, если это пространство обладает конечной системой порождающих. В только что рассмотренных примерах все пространства оказались конечно порожденными, кроме, быть может, пространства $K[x]$. Мы не можем утверждать, что пространство $K[x]$ не является конечно порожденным, лишь на том основании, что обнаружили в нем бесконечную систему порождающих: нельзя, a priori, исключить возможность того, что потрудившись немного больше, мы сумеем найти в этом пространстве и какую-нибудь конечную систему порождающих. Тем не менее, нетрудно показать, что это пространство действительно не является конечно порожденным; впрочем, позже это утверждение мы сможем получить, как простое следствие некоторых общих свойств линейных пространств.

Определив подпространство, порожденное данным множеством M векторов линейного пространства L, как линейную оболочку $l(M)$ этого множества, мы фактически указали, из каких элементов это подпространство состоит. Поэтому это определение можно назвать внутренним (по отношению к этому подпространству). Теорема 1 подсказывает возможность внешнего определения подпространства, порожденного множеством M, как наименьшего из всех подпространств пространства L, содержащих M. Но в отличие от внутреннего, это определение требует доказательства существования определяемого объекта: тот факт, что среди всех подпространств пространства L, содержащих множество M, существует наименьшее, т. е. принадлежащее всем остальным таким подпространствам, не вполне очевиден. Тем не менее, следующее простое утверждение делает его таковым:

Предложение 4. *Пересечение произвольного семейства подпространств линейного пространства является подпространством этого пространства.* $\square$

Теперь ясно, что если M — подмножество линейного пространства L, то для построения наименьшего из подпространств пространства L, содержащих множество M, следует взять пересечение всех подпространств пространства L, содержащих это множество.

Таким образом, мы располагаем двумя равносильными определениями подпространства пространства L, порожденного множеством M: внутренним определением, как линейной оболочки множества M, и внешним определением, как наименьшего из подпространств, содержащих это множество.

В отличие от пересечения, объединение двух подпространств линейного пространства почти всегда не является подпространством:

Предложение 5. *Пусть A и B — подпространства линейного пространства L. Множество $A \cup B$ является подпространством пространства L тогда и только тогда, когда или $A \subseteq B$, или $B \subseteq A$.* $\square$

Заметим, что объединение $A \cup B$ двух подмножеств A и B некоторого множества X можно охарактеризовать, как наименьшее из всех подмножеств множества X, содержащих эти подмножества A и B. Поэтому, по аналогии, объединением подпространств A и B линейного пространства L можно было бы назвать наименьшее из всех подпространств пространства L, содержащих и A, и B. Такое подпространство сейчас и будет определено, но с помощью более обозримой конструкции суммы подмножеств линейного пространства:

Определение 4. *Пусть $A_1, A_2, \dots, A_n$ — подмножества линейного пространства L. Суммой этих подмножеств называется множество $A_1 + A_2 + \dots + A_n$, состоящее из всевозможных векторов вида $a_1 + a_2 + \dots + a_n$, где $a_i \in A_i$ для всех $i = 1, 2, \dots, n$.*

Для обозначения суммы подмножеств $A_1, A_2, \dots, A_n$ мы будем использовать также и знак суммирования:

$$\sum_{i=1}^{n} A_i = \left\{ a_1 + a_2 + \dots + a_n \mid a_i \in A_i \ (i = 1, 2, \dots, n) \right\}.$$

Будем рассматривать также и сумму подмножеств линейного пространства, состоящую из одного слагаемого: как обычно, это позволит избежать дополнительных оговорок при формулировке ряда утверждений.

Поясним введенное понятие на простом примере суммы двух конечных подмножеств пространства $\mathbb{Q}^2$:

$$A = \{(1,1), (1,2), (-1,0)\} \quad \text{и} \quad B = \{(2,0), (2,-1)\}.$$

Так как $(1,1) + (2,0) = (3,1)$, $(1,1) + (2,-1) = (3,0)$, $(1,2) + (2,0) = (3,2)$, $(1,2) + (2,-1) = (3,1)$, $(-1,0) + (2,0) = (1,0)$ $(-1,0) + (2,-1) = (1,-1)$, то

$$A + B = \{(3,1), (3,0), (3,2), (1,0), (1,-1)\}.$$

Покажем теперь, что сумма произвольных подпространств линейного пространства является подпространством и что это подпространство действительно может служить аналогом теоретико-множественного объединения.

Теорема 2. *Сумма $\sum_{i=1}^{n} A_i$ подпространств $A_1, A_2, \dots, A_n$ линейного пространства L является подпространством этого пространства. Более того, $\sum_{i=1}^{n} A_i$ является наименьшим из всех подпространств пространства L, содержащих каждое из подпространств $A_1, A_2, \dots, A_n$.*

Доказательство. Как и в доказательстве теоремы 1, начнем с более детальной формулировки утверждений, которые нам предстоит доказывать. Теорема 2 утверждает, что

1) $\sum_{i=1}^{n} A_i$ является подпространством пространства L;
2) каждое из подпространств $A_1, A_2, \dots, A_n$ содержится в подпространстве $\sum_{i=1}^{n} A_i$;
3) произвольное подпространство пространства L, содержащее все подпространства $A_1, A_2, \dots, A_n$, содержит и их сумму $\sum_{i=1}^{n} A_i$.

Доказательства этих утверждений почти очевидны. Пусть a и b — произвольные векторы из множества $\sum_{i=1}^{n} A_i$. Тогда каждый из них должен допускать разложение вида

$$a = \sum_{i=1}^{n} a_i \quad \text{и} \quad b = \sum_{i=1}^{n} b_i,$$

где для любого $i = 1, 2, \ldots, n$ векторы a_i и b_i принадлежат подпространству A_i. Взяв еще произвольные скаляры α и β, получаем

$$\alpha a + \beta b = \alpha \sum_{i=1}^{n} a_i + \beta \sum_{i=1}^{n} b_i = \sum_{i=1}^{n} \alpha a_i + \sum_{i=1}^{n} \beta b_i = \sum_{i=1}^{n} (\alpha a_i + \beta b_i).$$

Так как для каждого $i = 1, 2, \ldots, n$ вектор $\alpha a_i + \beta b_i$ принадлежит подпространству A_i, то в соответствии с определением суммы подмножеств вектор $\alpha a + \beta b$ должен считаться элементом множества $\sum_{i=1}^{n} A_i$. Следовательно, это множество замкнуто относительно взятия линейных комбинаций произвольных систем из двух векторов и потому в силу предложения 2 является подпространством.

Фиксируем теперь произвольный номер k, $1 \le k \le n$, и возьмем вектор a из подпространства A_k. Полагая для $i = 1, 2, \ldots, n$

$$a_i = \begin{cases} a, & \text{если } i = k \\ 0, & \text{если } i \ne k, \end{cases}$$

очевидно имеем

$$a = a_1 + a_2 + \cdots + a_n \quad \text{и} \quad a_i \in A_i (i = 1, 2, \ldots, n),$$

что говорит о том, что вектор a лежит в подпространстве $\sum_{i=1}^{n} A_i$. Таким образом, каждое из подпространств $A_1, A_2, \ldots, A_n$ содержится в подпространстве $\sum_{i=1}^{n} A_i$.

Наконец, пусть N — произвольное подпространство пространства L, содержащее все подпространства A_1, A_2, ... , A_n, и пусть a — произвольный вектор из подпространства $\sum_{i=1}^{n} A_i$. Тогда $a = a_1 + a_2 + \cdots + a_n$, где каждое слагаемое a_i содержится в соответствующем подпространстве A_i, а потому — и в подпространстве N. Следовательно, $a \in N$ и включение $\sum_{i=1}^{n} A_i \subseteq N$ доказано. $\square$

Из теоремы 2 следует, в частности, что подпространство $\sum_{i=1}^{n} A_i$ порождается теоретико-множественным объединением $\bigcup_{i=1}^{n} A_i$ подпространств A_1, A_2, ... , A_n. В действительности, можно указать более "экономную" систему порождающих пространства $\sum_{i=1}^{n} A_i$:

Следствие. *Пусть подпространства $A_1, A_2, \ldots, A_n$ линейного пространства L порождаются множествами векторов $M_1, M_2, \ldots, M_n$ соответственно. Тогда подпространство $\sum_{i=1}^{n} A_i$ порождается множеством $M = \bigcup_{i=1}^{n} M_i$.*

В самом деле, так как $A_k = l(M_k)$, очевидное включение $M_k \subseteq l(M)$ дает ввиду теоремы 1 включение $A_k \subseteq l(M)$ для любого $k = 1, 2, \ldots, n$, откуда по теореме 2 следует включение $\sum_{i=1}^{n} A_i \subseteq l(M)$. Противоположное включение $l(M) \subseteq \sum_{i=1}^{n} A_i$ следует из теорем 1 и 2 и очевидного же включения $M \subseteq \sum_{i=1}^{n} A_i$. $\square$

Иногда сумма подпространств линейного пространства может обладать дополнительным важным свойством. Рассмотрим в связи с этим два примера разложения пространства $L = \mathbb{Q}^3$ в сумму двух подпространств.

Полагаем

$$M = \{(\alpha_1, \alpha_2, \alpha_3) \in L \mid \alpha_1 + \alpha_2 = 0\} \text{ и } N = \{(\alpha_1, \alpha_2, \alpha_3) \in L \mid \alpha_1 + \alpha_3 = 0\}.$$

Эти подмножества являются, очевидно, подпространствами пространства L. Так как произвольный вектор $a = (\alpha, \beta, \gamma)$ пространства L может быть записан в виде $a = b+c$, где $b = (\alpha, -\alpha, \alpha+\beta+\gamma) \in M$ и $c = (0, \alpha+\beta, -(\alpha+\beta)) \in N$, это пространство является суммой подпространств M и N. Представить вектор a в виде суммы двух слагаемых, взятых из подпространств M и N можно и по другому; например, $a = b' + c'$, где $b' = (\alpha - \beta, \beta - \alpha, \alpha + \gamma)$ и $c' = (\beta, \alpha, -\alpha)$.

С другой стороны, полагая

$$U = \{(\alpha_1, \alpha_2, \alpha_3) \in L \mid \alpha_3 = 0\} \quad \text{и} \quad V = \{(\alpha_1, \alpha_2, \alpha_3) \in L \mid \alpha_1 = \alpha_2 = 0\},$$

мы видим, что пространство L является суммой и подпространств U и V. Действительно, для произвольного вектора $a = (\alpha, \beta, \gamma)$ пространства L имеем $a = u + v$, где $u = (\alpha, \beta, 0) \in U$ и $v = (0, 0, \gamma) \in V$. Но в отличие от первого примера, указанное разложение вектора a является, очевидно, единственным возможным представлением этого вектора в виде суммы двух слагаемых, принадлежащих подпространствам U и V. Это и есть то свойство, которое отличает второе разложение пространства L от первого, и чтобы выразить это различие, мы будем говорить, что пространство L является прямой суммой подпространств U и V. Более точно, мы принимаем

Определение 5. *Линейное пространство L называется прямой суммой своих подпространств $L_1, L_2, \dots, L_n$, если каждый вектор $a \in L$ может быть единственным образом представлен в виде $a = a_1 + a_2 + \dots + a_n$, где $a_i \in L_i$ для всех $i = 1, 2, \dots, n$.*

Говоря подробнее, в соответствии с определением линейное пространство L является прямой суммой своих подпространств $L_1, L_2, \dots, L_n$, если, во-первых, каждый вектор из L представим в виде суммы векторов, взятых по одному из этих подпространств, т. е. $L = \sum_{i=1}^{n} L_i$, а во-вторых, указанное разложение вектора a является единственным, т. е. если

$$a = a_1 + a_2 + \dots + a_n \quad \text{и} \quad a = a_1' + a_2' + \dots + a_n',$$

где $a_i, a_i' \in L_i$ $(i = 1, 2, \dots, n)$, то $a_i = a_i'$ для всех $i = 1, 2, \dots, n$. В этом случае мы будем писать $L = L_1 \oplus L_2 \oplus \dots \oplus L_n$ или, короче, $\bigoplus_{i=1}^{n} L_i$. Таким образом, в только что рассмотренном примере имеем $L = M + N$, $L = U \oplus V$, но утверждение $L = M \oplus N$ неверно. Отметим еще, что при $n = 1$ $\sum_{i=1}^{n} L_i = \bigoplus_{i=1}^{n} L_i$.

Следующее утверждение показывает, что требование единственности в определении прямой суммы подпространств можно выразить двумя другими способами.

Теорема 3. *Пусть линейное пространство L является суммой своих подпространств $L_1, L_2, \dots, L_n$. Пространство L является прямой суммой этих подпространств тогда и только тогда, когда выполнено любое из следующих двух равносильных между собой условий:*

а) *для любых векторов $a_1 \in L_1,\ a_2 \in L_2, \dots,\ a_n \in L_n$ равенство $a_1 + a_2 + \dots + a_n = 0$ возможно лишь тогда, когда $a_1 = a_2 = \dots = a_n = 0$;*

б) *для каждого номера $k = 1, 2, \ldots, n$ пересечение подпространства L_k с суммой всех остальных подпространств семейства $L_1, L_2, \ldots, L_n$ является нулевым подпространством:*

$$L_k \cap \sum_{i \neq k} L_i = \{0\}.$$

Доказательство. Покажем сначала, что если $L = \bigoplus_{i=1}^{n} L_i$, то выполнено условие а).

Предположим, что для некоторых векторов $a_1 \in L_1,\ a_2 \in L_2, \ldots,\ a_n \in L_n$ имеет место равенство $a_1 + a_2 + \cdots + a_n = 0$. Тогда нулевой вектор пространства L можно представить в виде суммы n слагаемых двумя способами

$$\begin{aligned} 0 &= a_1 + a_2 + \cdots + a_n \quad \text{и} \\ 0 &= 0\ \ + 0\ \ + \cdots + 0, \end{aligned}$$

причем в каждом из этих разложений i-ое слагаемое входит в подпространство L_i. Требование единственности в определении прямой суммы дает теперь систему равенств $a_1 = 0, a_2 = 0, \ldots, a_n = 0$, т.е. условие а) выполняется.

Покажем теперь, что из условия а) следует, что пространство L является прямой суммой подпространств $L_1, L_2, \ldots, L_n$. Так как требуемое определением прямой суммы равенство $L = \sum_{i=1}^{n} L_i$ содержится в условии теоремы, нам остается доказать лишь единственность разложения вектора $a \in L$ в сумму слагаемых, взятых по одному из подпространств $L_1, L_2, \ldots, L_n$. Пусть для некоторого вектора $a \in L$ имеют место разложения

$$a = a_1 + a_2 + \cdots + a_n \quad \text{и} \quad a = a_1' + a_2' + \cdots + a_n',$$

где $a_i, a_i' \in L_i$ $(i = 1, 2, \ldots, n)$. Вычитая почленно из первого равенства второе, после очевидной группировки имеем

$$(a_1 - a_1') + (a_2 - a_2') + \cdots + (a_n - a_n') = 0.$$

Так как $a_i - a_i' \in L_i$ $(i = 1, 2, \ldots, n)$, из условия а) следует, что для каждого $i = 1, 2, \ldots, n$ выполнено равенство $a_i - a_i' = 0$, т.е. $a_i = a_i'$. Таким образом, единственность указанного разложения произвольного вектора пространства L установлена, и, следовательно, $L = \bigoplus_{i=1}^{n} L_i$.

Покажем наконец, что условия а) и б) равносильны. Предположим сначала, что имеет место а). Если вектор a пространства L входит в пересечение $L_k \cap \sum_{i \neq k} L_i$, то $a \in \sum_{i \neq k} L_i$ и потому $a = a_1 + a_2 + \cdots + a_{k-1} + a_{k+1} + \cdots + a_n$ для подходящих векторов $a_1 \in L_1,\ a_2 \in L_2, \ldots,\ a_{k-1} \in L_{k-1}, a_{k+1} \in L_{k+1}, \ldots,\ a_n \in L_n$. Тогда выполнено равенство

$$a_1 + a_2 + \cdots + a_{k-1} + (-a) + a_{k+1} + \cdots + a_n = 0,$$

i-ое слагаемое в левой части которого принадлежит подпространству L_i. Поэтому из условия а) следует, что каждое из этих слагаемых равно нулю, и в частности, $a = 0$. Таким образом, $L_k \cap \sum_{i \neq k} L_i = \{0\}$.

Обратно, пусть имеет место условие б) и пусть для некоторых векторов $a_1 \in L_1,\ a_2 \in L_2, \ldots,\ a_n \in L_n$ имеет место равенство $a_1 + a_2 + \cdots + a_n = 0$. Выбрав произвольный номер $k,\ 1 \le k \le n$, перепишем это равенство в виде

$$a_k = (-a_1) + (-a_2) + \cdots + (-a_{k-1}) + (-a_{k+1}) + \cdots + (-a_n).$$

Так как вектор, стоящий в правой части этого равенства, является, очевидно, элементом подпространства $\sum_{i \ne k} L_i$, мы видим, что вектор a_k входит в пересечение $L_k \cap \sum_{i \ne k} L_i$, по условию равное нулевому подпространству. Следовательно, $a_k = 0$, и так как номер k был выбран произвольным, условие а) доказано. $\square$

Из теоремы 3 следует ряд полезных утверждений о разложении линейного пространства в прямую сумму подпространств. Заметим, прежде всего, что для распознавания того, является ли сумма двух подпространств прямой, теорема 3 дает совсем простой критерий:

Следствие 1. *Сумма $A + B$ двух подпространств A и B линейного пространства L является прямой тогда и только тогда, когда $A \cap B = \{0\}$.* $\square$

Отметим, далее

Следствие 2. *Пусть $L = \bigoplus_{i=1}^n L_i$ и пусть $M = \sum_{i=1}^s L_i$, где $1 \le s \le n$. Тогда $M = \bigoplus_{i=1}^s L_i$.*

Ввиду теоремы 3 для доказательства этого утверждения достаточно убедиться в том, что разложение пространства M в сумму подпространств $L_1, L_2, \ldots, L_s$ удовлетворяет условию а) из формулировки этой теоремы. В самом деле, пусть для векторов $a_1 \in L_1,\ a_2 \in L_2, \ldots,\ a_s \in L_s$ имеет место равенство $a_1 + a_2 + \cdots + a_s = 0$. Полагая $a_{s+1} = a_{s+2} = \cdots = a_n = 0$, имеем, очевидное равенство

$$a_1 + a_2 + \cdots + a_s + a_{s+1} + \cdots + a_n = 0,$$

причем для всех $i = 1, 2, \ldots, n$ вектор a_i принадлежит подпространству L_i. Поскольку разложение пространства L в сумму подпространств $L_1, L_2, \ldots, L_n$ условию а) теоремы 3 удовлетворяет, каждое слагаемое в левой части этого равенства должно быть равным нулю. В частности, $a_1 = a_2 = \cdots = a_s = 0$. $\square$

Говоря содержательно, в следствии 2 утверждается, что сумма любой части слагаемых из разложения пространства L в прямую сумму также является прямой. Покажем теперь, что группировка прямых слагаемых или разложение их в прямую сумму приводят к разложению пространства в прямую сумму соответствующих подпространств.

Следствие 3. *Пусть $L = \bigoplus_{i=1}^n L_i$ и пусть $M = \sum_{i=1}^s L_i,\ N = \sum_{i=s+1}^n L_i$ где $1 \le s < n$. Тогда $L = M \oplus N$.*

Доказательство. Покажем сначала, что $L = M + N$. Действительно, произвольный вектор $a \in L$ может быть записан в виде $a = a_1 + a_2 + \cdots + a_n$, где для всех $i = 1, 2, \ldots, n$

$a_i \in L_i$. Полагая здесь $b = a_1 + a_2 + \cdots + a_s$ и $c = a_{s+1} + a_{s+2} + \cdots + a_n$, получаем $a = b + c$, $b \in M$, $c \in N$, что и требовалось.

Ввиду следствия 1 остается понять, что $M \cap N = \{0\}$. В самом деле, если a — произвольный вектор из пересечения $M \cap N$, то $a \in M$ и потому $a = a_1 + a_2 + \cdots + a_s$ для подходящих векторов $a_i \in L_i$ $(i = 1, 2, \ldots, s)$, а также $a \in N$ и потому $a = a_{s+1} + a_{s+2} + \cdots + a_n$ для подходящих векторов $a_i \in L_i$ $(i = s+1, s+2, \ldots, n)$. Отсюда получаем равенство

$$a_1 + a_2 + \cdots + a_s + (-a_{s+1}) + \cdots + (-a_n) = 0,$$

выполнимость которого в силу условия а) из теоремы 3 возможна лишь в том случае, когда все слагаемые в его левой части равны нулю. В частности, $a_1 = a_2 = \cdots = a_s = 0$, и потому $a = 0$. Таким образом, $M \cap N = \{0\}$, и по следствию 1 $L = M \oplus N$. □

Следствие 4. *Пусть $L = \bigoplus_{i=1}^{n} L_i$ и пусть для каждого $i = 1, 2, \ldots, n$ подпространство L_i, в свою очередь, является прямой суммой подпространств: $L_i = \bigoplus_{j=1}^{n_i} M_{ij}$. Тогда $L = \bigoplus_{ij} M_{ij}$.*

Доказательство. Покажем сначала, что $L = \sum_{ij} M_{ij}$. Так как для каждого $i = 1, 2, \ldots, n$ $L_i = \sum_{j=1}^{n_i} M_{ij}$, из теоремы 2 следует, что все подпространства $L_1, L_2, \ldots, L_n$ содержатся в подпространстве $\sum_{ij} M_{ij}$. Поскольку $L = \sum_{i=1}^{n} L_i$, по той же теореме имеем, что $L \subseteq \sum_{ij} M_{ij}$, и так как справедливость противоположного включения очевидна, имеем $L = \sum_{ij} M_{ij}$.

Предположим теперь, что сумма векторов $a_{ij} \in M_{ij}$ $(i = 1, 2, \ldots, n,\ j = 1, 2, \ldots, n_i)$ равна нулю, и запишем это равенство в виде $\sum_{i=1}^{n} \sum_{j=1}^{n_i} a_{ij} = 0$. Так как для каждого $i = 1, 2, \ldots, n$ вектор $\sum_{j=1}^{n_i} a_{ij}$ принадлежит подпространству L_i, из справедливости условия а) для исходного разложения пространства L следует, что для всех $i = 1, 2, \ldots, n$ выполнены равенства $\sum_{j=1}^{n_i} a_{ij} = 0$. Применяя теперь условие а) к соответствующим разложениям пространств $L_1, L_2, \ldots, L_n$, получаем равенство нулю каждого из векторов a_{ij}. Из теоремы 3 теперь следует, что $L = \bigoplus_{ij} M_{ij}$. □

§ 2. ЛИНЕЙНЫЕ ОТОБРАЖЕНИЯ ЛИНЕЙНЫХ ПРОСТРАНСТВ

В этом параграфе вводится второе из двух основных понятий линейной алгебры — понятие линейного отображения.

Определение 1. *Пусть L и M — линейные пространства над (одним и тем же) полем K. Отображение $\varphi : L \to M$ множества векторов пространства L в множество векторов пространства M называется линейным отображением пространства L в пространство M, если для любых векторов $a, b \in L$ и скаляра $\alpha \in K$ выполнены равенства*

$$(a+b)\varphi = a\varphi + b\varphi \quad \textit{и} \quad (\alpha a)\varphi = \alpha(a\varphi).$$

Биективное линейное отображение называется изоморфизмом. Пространство L называется изоморфным пространству M, если существует изоморфизм пространства L в пространство M.

Договоримся также линейное отображение пространства L в себя называть *линейным оператором* пространства L.

Следует подчеркнуть, что линейное отображение определяется лишь для линейных пространств над одним и тем же полем. Интуитивно ясно, что с помощью линейного отображения пространства L в пространство M можно (во всяком случае, некоторые) свойства пространства L переносить в пространство M, а изоморфные пространства должны иметь одинаковые свойства. Иначе говоря, изоморфизм линейных пространств будет служить уточнением интуитивного представления о том, какие линейные пространства следует считать одинаковыми. Дальнейшее покажет, что наши интуитивные ожидания во многом оправданы. Уже следующее простое предложение говорит о том, что отношение изоморфизма для линейных пространств обладает обычными свойствами отношения равенства.

Предложение 1. *Тождественное отображение линейного пространства в себя является линейным отображением (и потому изоморфизмом). Произведение двух линейных отображений является линейным отображением. Отображение, обратное к изоморфизму, является изоморфизмом.*

Доказательство. Линейность тождественного отображения ι_L множества векторов линейного пространства L очевидна: для любых векторов a и b и скаляра α имеем

$$(a+b)\iota_L = a + b = a\iota_L + b\iota_L \quad \text{и} \quad (\alpha a)\iota_L = \alpha a = \alpha(a\iota_L).$$

Пусть теперь L, M и N — линейные пространства над полем K и пусть $\varphi : L \to M$ — линейное отображение пространства L в пространство M и $\psi : M \to N$ — линейное отображение пространства M в пространство N. Утверждается, что отображение $\varphi\psi$ пространства L в пространство N также является линейным. Действительно, для любых векторов $a, b \in L$ и скаляра $\alpha \in K$ в силу определения умножения отображений и линейности отображений φ и ψ имеем

$$(a+b)(\varphi\psi) = ((a+b)\varphi)\psi = (a\varphi + b\varphi)\psi = (a\varphi)\psi + (b\varphi)\psi = a(\varphi\psi) + b(\varphi\psi)$$

и

$$(\alpha a)(\varphi\psi) = ((\alpha a)\varphi)\psi = (\alpha(a\varphi))\psi = \alpha((a\varphi)\psi) = \alpha(a(\varphi\psi)).$$

Покажем, наконец, что если $\varphi : L \to M$ — изоморфизм пространства L в пространство M, то обратное отображение φ^{-1} является изоморфизмом пространства M в пространство L. Для векторов a и b пространства M полагаем $a\varphi^{-1} = x$ и $b\varphi^{-1} = y$. Тогда $x, y \in L$, $x\varphi = a$ и $y\varphi = b$. Поскольку ввиду линейности отображения φ справедливо равенство $(x+y)\varphi = x\varphi + y\varphi = a+b$, то по определению обратного отображения имеем $(a+b)\varphi^{-1} = x+y = a\varphi^{-1}+b\varphi^{-1}$. Аналогично для произвольного скаляра $\alpha \in K$ имеем $(\alpha x)\varphi = \alpha(x\varphi) = \alpha a$, и потому $(\alpha a)\varphi^{-1} = \alpha x = \alpha(a\varphi^{-1})$. Таким образом, (биективное) отображение φ^{-1} линейно и потому является изоморфизмом пространства M в пространство L. $\square$

Рассмотрим несколько примеров линейных отображений линейных пространств. Мы только что видели, что тождественное отображение множества векторов линейного пространства L является линейным оператором этого пространства. Еще одним примером линейного оператора пространства L является нулевой оператор θ_L, отображающий произвольный вектор $a \in L$ в нулевой вектор (линейность этого отображения очевидна как, впрочем, и определяемого аналогичным образом нулевого отображения пространства L в любое другое пространство над тем же полем). Обобщая эти примеры, введем для произвольного скаляра λ из поля K отображение χ_λ пространства L над полем K в себя, действующее по правилу

$$a\chi_\lambda = \lambda a \qquad (a \in L).$$

Линейность отображения χ_λ проверяется непосредственно, и этот оператор называется *гомотетией с коэффициентом* λ линейного пространства L. Очевидно, что при $\lambda = 0$ оператор χ_λ совпадает с нулевым, а при $\lambda = 1$ с тождественным оператором.

Пусть $L = K^n$ — n-мерное координатное пространство над полем K. Для фиксированного номера i, $1 \le i \le n$, определим отображение π_i пространства L в поле K, рассматриваемое как линейное пространство над самим собой, по следующему правилу: для любого вектора $a = (\alpha_1, \alpha_2, \dots, \alpha_n)$ пространства L полагаем $a\pi_i = \alpha_i$. Это линейное (как легко проверить) отображение называется проектированием пространства L на i-ую компоненту.

Пусть линейное пространство L над полем K разложено в прямую сумму некоторых своих подпространств M и N, $L = M \oplus N$. Тогда произвольный вектор $a \in L$ однозначно записывается в виде $a = b+c$, где $b \in M$ и $c \in N$, и потому можно определить отображение φ пространства L в себя, переводящее вектор a в вектор b. Иначе говоря, для вычисления образа относительно отображения φ вектора пространства L следует разложить этот вектор в сумму двух слагаемых, принадлежащих подпространствам M и N, и взять первое из них. Покажем, что это отображение линейно. Пусть a и b — произвольные векторы пространства L и пусть образами этих векторов относительно отображения φ являются векторы c и d: $a\varphi = c$, $b\varphi = d$. По определению нашего отображения это означает, что векторы c и d лежат в подпространстве

M и для подходящих векторов u и v из подпространства N выполнены равенства $a = c + u$ и $b = d + v$. Так как тогда $a + b = (c + d) + (u + v)$ и $c + d \in M$, $u + v \in N$, имеем $(a + b)\varphi = c + d = a\varphi + b\varphi$. Точно так же поскольку для произвольного $\alpha \in K$ $\alpha a = \alpha c + \alpha u$ и $\alpha c \in M$, $\alpha u \in N$, имеем $(\alpha a)\varphi = \alpha c = \alpha(a\varphi)$. Отображение φ называется оператором проектирования пространства L на подпространство M параллельно подпространству N.

Укажем, наконец, на отображение дифференцирования в линейном пространстве $K[x]$ многочленов от переменной x над полем K, а именно — отображение, которое ставит в соответствие многочлену $f(x) \in K[x]$ его производную $f'(x)$. Известные свойства производной говорят о том, что это отображение является линейным оператором пространства $K[x]$.

Переходя к перечислению простейших свойств линейных отображений, покажем, прежде всего, что два требования в определении 1 можно заменить одним.

Предложение 2. *Пусть L и M — линейные пространства над полем K. Отображение $\varphi : L \to M$ пространства L в пространство M является линейным тогда и только тогда, когда для любых векторов $a, b \in L$ и любых скаляров $\alpha, \beta \in K$ выполнено равенство*

$$(\alpha a + \beta b)\varphi = \alpha(a\varphi) + \beta(b\varphi). \tag{1}$$

В самом деле, если отображение φ линейно, то применяя последовательно равенства из определения 1, получаем

$$(\alpha a + \beta b)\varphi = (\alpha a)\varphi + (\beta b)\varphi = \alpha(a\varphi) + \beta(b\varphi).$$

Обратно, два равенства из формулировки определения 1 получаются из (1) при $\alpha = 1$, $\beta = 1$ и $\alpha = 1$, $\beta = 0$ соответственно. $\square$

Отметим еще почти очевидное

Предложение 3. *Пусть $\varphi : L \to M$ — линейное отображение линейного пространства L в линейное пространство M. Тогда*

1) $0\varphi = 0$ и для любого вектора $a \in L$ $(-a)\varphi = -a\varphi$;

2) для любых векторов $a, b \in L$ $(a - b)\varphi = a\varphi - b\varphi$;

3) образ относительно отображения φ линейной комбинации произвольной системы векторов пространства L равен линейной комбинации с теми же коэффициентами образов векторов этой системы, т. е. для любых векторов $a_1, a_2, \ldots, a_n$ пространства L и любых скаляров $\alpha_1, \alpha_2, \ldots, \alpha_n$

$$(\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n)\varphi = \alpha_1(a_1\varphi) + \alpha_2(a_2\varphi) + \cdots + \alpha_n(a_n\varphi).$$

В самом деле, $0\varphi = (0a)\varphi = 0(a\varphi) = 0\varphi$ и $(-a)\varphi = ((-1)a)\varphi = (-1)(a\varphi) = -a\varphi$. Равенство из утверждения 3) (которое мы будем записывать и в виде $\left(\sum_{i=1}^{n} \alpha_i a_i\right)\varphi = \sum_{i=1}^{n} \alpha_i(a_i\varphi)$) доказывается очевидной индукцией по n. $\square$

Следующее утверждение подтверждает справедливость интуитивного представления о том, что подпространства — это "правильные" подмножества, а линейные отображения — это "правильные" отображения линейных пространств:

Теорема 1. *Образ и прообраз подпространства относительно линейного отображения является подпространством.*

Доказательство. Пусть $\varphi : L \to M$ — линейное отображение линейного пространства L в линейное пространство M и пусть A — подпространство пространства L. Покажем, что образ относительно φ подпространства A, т.е. множество $A\varphi = \{a\varphi \mid a \in A\}$ всех тех векторов пространства M, которые являются образами относительно φ элементов подпространства A, является подпространством пространства M. Для этого воспользуемся критерием предложения 1.2. Если векторы x и y пространства M принадлежат множеству $A\varphi$, то для подходящих векторов $a, b \in A$ $x = a\varphi$ и $y = b\varphi$. Поэтому для произвольных скаляров α и β имеем

$$\alpha x + \beta y = \alpha(a\varphi) + \beta(b\varphi) = (\alpha a + \beta b)\varphi,$$

и так как $\alpha a + \beta b \in A$ (поскольку A подпространство), то $\alpha x + \beta y \in A\varphi$. Таким образом, в силу предложения 1.2 множество $A\varphi$ является подпространством пространства M.

Пусть теперь B — подпространство пространства M. Покажем, что прообраз относительно φ множества B, т.е. множество $B\varphi^{-1} = \{a \in L \mid a\varphi \in B\}$ всех тех элементов пространства L, образы которых относительно φ лежат в множестве B, является подпространством пространства L. Возьмем произвольные векторы a и b из множества $B\varphi^{-1}$. Это означает, что векторы $a\varphi$ и $b\varphi$ принадлежат подпространству B, и потому для произвольных скаляров α и β имеем

$$(\alpha a + \beta b)\varphi = \alpha(a\varphi) + \beta(b\varphi) \in B.$$

Следовательно, вектор $\alpha a + \beta b$ принадлежит множеству $B\varphi^{-1}$, и ввиду предложения 1.2 множество $B\varphi^{-1}$ является подпространством пространства L. $\square$

Выделим следующий частный случай теоремы 1:

Следствие 1. *Пусть $\varphi : L \to M$ — линейное отображение линейного пространства L в линейное пространство M. Тогда образ $Im\,\varphi$ отображения φ (совпадающий, напомним, с множеством $L\varphi$) является подпространством пространства M.* $\square$

Отметим еще полезное

Следствие 2. *Пусть $\varphi : L \to M$ — линейное отображение линейного пространства L в линейное пространство M. Тогда образ линейной оболочки произвольного множества S векторов пространства L совпадает с линейной оболочкой образа множества S, т.е. $l(S)\varphi = l(S\varphi)$. В частности, образ множества порождающих пространства L является множеством порождающих подпространства $Im\,\varphi$ пространства M.*

Доказательство. Из включения $S \subseteq l(S)$ следует включение $S\varphi \subseteq l(S)\varphi$, и так как $l(S)\varphi$ является подпространством пространства M, из теоремы 1.1 следует, что $l(S\varphi) \subseteq l(S)\varphi$.

С другой стороны, включение $S\varphi \subseteq l(S\varphi)$ означает, что $S \subseteq (l(S\varphi))\varphi^{-1}$, и так как $(l(S\varphi))\varphi^{-1}$ является подпространством пространства L, снова по теореме 1.1 имеем включение $l(S) \subseteq (l(S\varphi))\varphi^{-1}$. Отсюда $l(S)\varphi \subseteq l(S\varphi)$, и равенство $l(S)\varphi = l(S\varphi)$ доказано.

Если, в частности, $L = l(S)$, то $\operatorname{Im}\varphi = L\varphi = l(S)\varphi = l(S\varphi)$. $\square$

В следствии 1 к теореме 1 рассматривается образ относительно отображения $\varphi : L \to M$ наибольшего подпространства пространства L. Другой предельный случай, прообраз наименьшего подпространства пространства M, играет весьма заметную роль в теории линейных пространств и заслуживает специального названия:

Определение 2. *Ядром линейного отображения $\varphi : L \to M$ линейного пространства L в линейное пространство M называется множество $Ker\,\varphi$ всех векторов пространства L, образ которых относительно φ равен нулю:*

$$Ker\,\varphi = \left\{a \in L \;\middle|\; a\varphi = 0\right\}.$$

Иначе говоря, ядро отображения φ является прообразом нулевого подпространства пространства M, и потому из теоремы 1 получаем очевидное

Следствие 3. *Ядро линейного отображения $\varphi : L \to M$ линейного пространства L в линейное пространство M является подпространством пространства L.* $\square$

О роли ядра линейного отображения говорит следующее простое утверждение:

Теорема 2. *Пусть $\varphi : L \to M$ — линейное отображение пространства L в пространство M. Для любых векторов $a, b \in L$ равенство $a\varphi = b\varphi$ имеет место тогда и только тогда, когда их разность $a - b$ принадлежит ядру отображения φ. В частности, отображение φ инъективно в точности тогда, когда его ядро совпадает с нулевым подпространством.*

Доказательство. Первое утверждение теоремы почти очевидно: для произвольных векторов $a, b \in L$ имеем

$$a\varphi = b\varphi \Leftrightarrow a\varphi - b\varphi = 0 \Leftrightarrow (a-b)\varphi = 0 \Leftrightarrow a - b \in \operatorname{Ker}\varphi.$$

Предположим теперь, что отображение φ инъективно и что вектор $a \in L$ лежит в ядре этого отображения. Тогда $a\varphi = 0$, и переписывая это равенство в виде $a\varphi = 0\varphi$, ввиду инъективности φ имеем $a = 0$. Таким образом, $\operatorname{Ker}\varphi$ содержит лишь нулевой вектор.

Обратно, пусть $\operatorname{Ker}\varphi = \{0\}$ и пусть для двух векторов $a, b \in L$ выполнено равенство $a\varphi = b\varphi$. Тогда $a - b \in \operatorname{Ker}\varphi$ и потому $a - b = 0$, т.е. $a = b$, что и доказывает инъективность отображения φ. $\square$

В предложении 1 было доказано, что произведение двух линейных отображений (если оно определено) является линейным отображением. Введем для линейных отображений новые операции: сложение и умножение на скаляр.

Определение 3. *Пусть L и M — линейные пространства над полем K, $\varphi : L \to M$ и $\psi : L \to M$ — два отображения множества L в множество M. Суммой отображений φ и ψ называется отображение $\varphi + \psi : L \to M$, определяемое по правилу*

$$a(\varphi + \psi) = a\varphi + a\psi \qquad (a \in L).$$

Произведением скаляра $\lambda \in K$ на отображение φ называется отображение $\lambda\varphi : L \to M$, определяемое по правилу

$$a(\lambda\varphi) = \lambda(a\varphi) \qquad (a \in L).$$

Вычислим, например, сумму двух гомотетий χ_λ и χ_μ пространства L. Для произвольного вектора $a \in L$ имеем, последовательно используя определение суммы отображений и свойства линейного пространства:

$$a(\chi_\lambda + \chi_\mu) = a\chi_\lambda + a\chi_\mu = \lambda a + \mu a = (\lambda + \mu)a = a\chi_{\lambda+\mu}.$$

Таким образом отображения $\chi_\lambda + \chi_\mu$ и $\chi_{\lambda+\mu}$ одинаково действуют на произвольный вектор пространства L и потому они равны. Следовательно, суммой двух гомотетий с коэффициентами λ и μ является гомотетия с коэффициентом $\lambda + \mu$, т. е. $\chi_\lambda + \chi_\mu = \chi_{\lambda+\mu}$. Аналогично доказывается, что для скаляра $\alpha \in K$ $\alpha\chi_\lambda = \chi_{\alpha\lambda}$. Отметим, заодно, и справедливость равенства $\chi_\lambda\chi_\mu = \chi_{\lambda\mu}$.

Теорема 3. *Если отображения $\varphi : L \to M$ и $\psi : L \to M$ линейного пространства L в линейное пространство M линейны, то линейными являются и отображения $\varphi + \psi$ и $\lambda\varphi$.*

Доказательство состоит в непосредственной проверке того, что рассматриваемые отображения удовлетворяют требованию предложения 2. Действительно, для любых векторов $a, b \in L$ и произвольных скаляров α и β имеем

$$(\alpha a + \beta b)(\varphi + \psi) = (\alpha a + \beta b)\varphi + (\alpha a + \beta b)\psi =$$
$$(\alpha(a\varphi) + \beta(b\varphi)) + (\alpha(a\psi) + \beta(b\psi)) = (\alpha(a\varphi) + \alpha(a\psi)) + (\beta(b\varphi) + \beta(b\psi)) =$$
$$\alpha(a\varphi + a\psi) + \beta(b\varphi + b\psi) = \alpha(a(\varphi + \psi)) + \beta(b(\varphi + \psi))$$

и

$$(\alpha a + \beta b)(\lambda\varphi) = (\lambda(\alpha a + \beta b)\varphi) = \lambda(\alpha(a\varphi) + \beta(b\varphi)) = \lambda(\alpha(a\varphi)) + \lambda(\beta(b\varphi)) =$$
$$\alpha(\lambda(a\varphi)) + \beta(\lambda(b\varphi)) = \alpha(a(\lambda\varphi)) + \beta(b(\lambda\varphi)). \quad \square$$

Если L и M линейные пространства над полем K, символом $\mathcal{L}(L, M)$ будем обозначать множество всех линейных отображений пространства L в пространство M. Множество $\mathcal{L}(L, L)$ всех линейных операторов пространства L договоримся обозначать символом $\mathcal{L}(L)$.

Теорема 4. *Пусть L и M — линейные пространства над полем K. Множество $\mathcal{L}(L, M)$ всех линейных отображений пространства L в пространство M с операциями сложения и умножения на скаляры, введенными в определении 3, является линейным пространством над полем K.*

Для доказательства этого утверждения достаточно убедиться в том, что сложение линейных отображений и умножение линейного отображения на скаляр удовлетворяют всем требованиям из определения линейного пространства.

Проверка коммутативности операции сложения не вызывает затруднений. Действительно, для произвольных отображений φ и ψ из множества $\mathcal{L}(L, M)$ и любого вектора $a \in L$ имеем

$$a(\varphi + \psi) = a\varphi + a\psi = a\psi + a\varphi = a(\psi + \varphi),$$

что и доказывает равенство $\varphi + \psi = \psi + \varphi$. Ассоциативность сложения проверяется аналогично. Полагая, далее, $\xi = (-1)\varphi + \psi$, получаем $\varphi + \xi = \psi$, так как

$$\begin{aligned} a(\varphi + \xi) = a\varphi + a\xi = a\varphi + a((-1)\varphi + \psi) = a\varphi + a((-1)\varphi) + a\psi \\ = a\varphi + (-1)(a\varphi) + a\psi = a\psi. \end{aligned}$$

Таким образом, и третья аксиома из определения линейного пространства выполняется. Для проверки выполнимости следующих трех аксиом возьмем произвольные скаляры α и β и с помощью столь же очевидных преобразований докажем справедливость равенств $\alpha(\varphi + \psi) = \alpha\varphi + \alpha\psi$, $(\alpha + \beta)\varphi = \alpha\varphi + \beta\varphi$ и $(\alpha\beta)\varphi = \alpha(\beta\varphi)$:

$$\begin{aligned} a(\alpha(\varphi + \psi)) = \alpha(a(\varphi + \psi)) = \alpha(a\varphi + a\psi) = \alpha(a\varphi) + \alpha(a\psi) = \\ a(\alpha\varphi) + a(\alpha\psi) = a(\alpha\varphi + \alpha\psi); \end{aligned}$$

$$a((\alpha + \beta)\varphi) = (\alpha + \beta)(a\varphi) = \alpha(a\varphi) + \beta(a\varphi) = a(\alpha\varphi) + a(\beta\varphi) = a(\alpha\varphi + \beta\varphi);$$
$$a((\alpha\beta)\varphi) = (\alpha\beta)(a\varphi) = \alpha(\beta(a\varphi)) = \alpha(a(\beta\varphi)) = a(\alpha(\beta\varphi)).$$

Наконец, последняя аксиома из определения линейного пространства просто очевидна: $a(1\varphi) = 1(a\varphi) = a\varphi$. □

В том частном случае, когда $L = M$, кроме операций сложения и умножения на скаляры, введенных в определении 3, на множестве $\mathcal{L}(L)$ имеется еще операция умножения: для любых двух отображений φ и ψ из линейного пространства $\mathcal{L}(L)$ всегда определено их произведение, состоящее в последовательном выполнении этих отображений и являющееся в силу предложения 1 линейным отображением, т. е. элементом пространства $\mathcal{L}(L)$. При этом умножение связано со сложением обычными дистрибутивными законами: для любых операторов $\varphi, \psi, \xi \in \mathcal{L}(L)$ имеют место равенства $\varphi(\psi + \xi) = \varphi\psi + \varphi\xi$ и $(\varphi + \psi)\xi = \varphi\xi + \psi\xi$. Действительно, для любого вектора $a \in L$ имеем

$$a(\varphi(\psi + \xi)) = (a\varphi)(\psi + \xi) = (a\varphi)\psi + (a\varphi)\xi = a(\varphi\psi) + a(\varphi\xi) = a(\varphi\psi + \varphi\xi),$$

$$a((\varphi+\psi)\xi)=(a(\varphi+\psi))\xi=(a\varphi+a\psi)\xi=(a\varphi)\xi+(a\psi)\xi=$$
$$a(\varphi\xi)+a(\psi\xi)=a(\varphi\xi+\psi\xi).$$

Это означает, что операции сложения и умножения элементов множества $\mathcal{L}(L)$ удовлетворяют всем аксиомам кольца. Заметим еще, что для любого скаляра α имеют место следующие легко проверяемые равенства $\alpha(\varphi\psi)=(\alpha\varphi)\psi=\varphi(\alpha\psi)$.

Таким образом, множество $\mathcal{L}(L)$ всех линейных операторов пространства L является носителем двух связанных между собой алгебраических структур — линейного пространства и кольца. Такое соединение заслуживает отдельного наименования:

Определение 4. *Пусть K — поле. Непустое множество A называется линейной алгеброй над полем K, если на этом множестве заданы две внутренние операции сложения $+$ и умножения $\cdot$ и внешняя операция умножения (слева) на элементы из поля K так, что выполнены следующие свойства:*

(1) *относительно операций сложения и умножения на элементы из поля K A является линейным пространством над K;*
(2) *относительно операций сложения и умножения A является кольцом;*
(3) *для любых элементов $a, b\in A$ и скаляра $\alpha\in K$ выполнены равенства $\alpha(ab)=(\alpha a)b=a(\alpha b)$.*

Линейная алгебра A называется коммутативной, если кольцо A коммутативно. Линейная алгебра A называется алгеброй с единицей, если единицу имеет кольцо A.

Таким образом, множество $\mathcal{L}(L)$ всех линейных операторов линейного пространства L над полем K является линейной алгеброй над тем же полем и даже линейной алгеброй с единицей (которой служит, очевидно, тождественный оператор пространства L). Линейными алгебрами (с обычными операциями) являются также пространство $M_n(K)$ квадратных матриц порядка n над полем K и пространство многочленов $K[x]$. Поле $\mathbb{C}$ комплексных чисел, рассматриваемое как линейное пространство над полем $\mathbb{R}$, является линейной алгеброй над $\mathbb{R}$. Все приведенные примеры являются алгебрами с единицей, а коммутативными среди них являются лишь две последние.

Существенным для нас свойством линейной алгебры над полем K является то, что в ней можно определить значение произвольного многочлена, коэффициенты которого принадлежат этому полю.

Определение 5. *Пусть A — линейная алгебра с единицей 1 над полем K и $f(x)=\alpha_0x^n+\alpha_1x^{n-1}+\dots+\alpha_{n-1}x+\alpha_n$ — многочлен от переменной x с коэффициентами из поля K. Значением многочлена $f(x)$ от элемента c алгебры A называется элемент $f(c)$ этой алгебры вида*

$$f(c)=\alpha_0c^n+\alpha_1c^{n-1}+\dots+\alpha_{n-1}c+\alpha_n1$$

Например, из вычислений, выполненных выше (перед формулировкой теоремы 3), легко следует, что $f(\chi_\lambda)=\chi_{f(\lambda)}$.

Предложение 4. *Пусть A — линейная алгебра с единицей 1 над полем K. Значение от элемента c алгебры A суммы и произведения двух многочленов равно соответственно сумме и произведению значений этих многочленов.*

Говоря более подробно, в предложении 4 утверждается, что если

$$f(x) = \alpha_0 x^m + \alpha_1 x^{m-1} + \cdots + \alpha_{m-1} x + \alpha_m \text{ и } g(x) = \beta_0 x^n + \beta_1 x^{n-1} + \cdots + \beta_{n-1} x + \beta_n$$

— произвольные многочлены над полем K и $u(x) = f(x) + g(x)$, $v(x) = f(x)g(x)$, то для любого элемента $c \in A$ имеют место равенства $u(c) = f(c) + g(c)$ и $v(c) = f(c)g(c)$. В самом деле, для вычисления в соответствии с определением, например, значения $v(c)$ многочлена $v(x)$, следует записать этот многочлен в виде

$$v(x) = \gamma_0 x^{m+n} + \gamma_1 x^{m+n-1} + \cdots + \gamma_{m+n-1} x + \gamma_{m+n}.$$

Но это выражение для многочлена $v(x)$ можно получить из его выражения вида

$$(\alpha_0 x^m + \alpha_1 x^{m-1} + \cdots + \alpha_{m-1} x + \alpha_m) \cdot (\beta_0 x^n + \beta_1 x^{n-1} + \cdots + \beta_{n-1} x + \beta_n),$$

раскрывая скобки и приводя подобные члены. Возможность выполнять эти преобразования основана на таких свойствах операций сложения и умножения многочленов, как ассоциативность, коммутативность и дистрибутивность умножения относительно сложения. Эти свойства выполняются в любом кольце и, в частности, — в алгебре A. Так как, к тому же, для любых $i = 0, 1, \ldots, m,\ j = 0, 1, \ldots, n$

$$(\alpha_i c^{m-i})(\beta_j c^{n-j}) = \alpha_i (c^{m-i}(\beta_j c^{n-j})) = \alpha_i(\beta_j (c^{m-i} c^{n-j})) = (\alpha_i \beta_j) c^{(m+n)-(i+j)},$$

то мы можем в алгебре A выражение

$$f(c)g(c) = (\alpha_0 \mathrm{c}^m + \alpha_1 \mathrm{c}^{m-1} + \cdots + \alpha_{m-1} \mathrm{c} + \alpha_m 1) \cdot (\beta_0 \mathrm{c}^n + \beta_1 \mathrm{c}^{n-1} + \cdots + \beta_{n-1} \mathrm{c} + \beta_n 1),$$

преобразовать в выражение

$$v(c) = \gamma_0 c^{m+n} + \gamma_1 c^{m+n-1} + \cdots + \gamma_{m+n-1} \mathrm{c} + \gamma_{m+n} 1.$$

Аналогично доказывается равенство $u(c) = f(c) + g(c)$. $\square$

Пример матричной алгебры $M_n(K)$ показывает, что линейные алгебры, вообще говоря, не являются коммутативными. Следующее утверждение, позволяющее, тем не менее, утверждать перестановочность некоторых элементов алгебры, будет неоднократно использоваться в дальнейшем.

Следствие. *Значения любых двух многочленов от одного и того же элемента линейной алгебры являются перестановочными элементами этой алгебры.*

Действительно, поскольку кольцо многочленов является коммутативным, то из равенства $v(x) = f(x)g(x)$ следует равенство $v(x) = g(x)f(x)$. Из предложения 4 теперь имеем $f(c)g(c) = v(c) = g(c)f(c)$. $\square$

§ 3. ЛИНЕЙНАЯ ЗАВИСИМОСТЬ ВЕКТОРОВ. БАЗИС И РАЗМЕРНОСТЬ ЛИНЕЙНОГО ПРОСТРАНСТВА

В этом параграфе вводится понятие линейной зависимости систем векторов линейного пространства, являющееся основным инструментом при изучении свойств линейных пространств.

Определение 1. *Пусть L — линейное пространство над полем K и пусть $a_1, a_2, \ldots, a_m$ — конечная система векторов пространства L. Будем говорить, что вектор $b \in L$ линейно выражается через систему $a_1, a_2, \ldots, a_m$, если он является значением некоторой линейной комбинации этой системы.*

Будем говорить также, что система $b_1, b_2, \ldots, b_n$ векторов пространства L линейно выражается через систему $a_1, a_2, \ldots, a_m$, если каждый ее вектор b_i линейно выражается через систему $a_1, a_2, \ldots, a_m$. Две системы векторов будем называть эквивалентными, если каждая из них линейно выражается через другую.

Таким образом, если в системе $a_1, a_2, \ldots, a_m$ число векторов $m \geq 1$, то вектор b линейно выражается через эту систему тогда и только тогда, когда $b = \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_m a_m$ для подходящих скаляров $\alpha_1, \alpha_2, \ldots, \alpha_m \in K$. Если же $m = 0$, т.е. система пустая, то из соглашений, принятых в § 1, следует, что через эту систему линейно выражается нулевой вектор пространства L и только он.

Разумеется, эти понятия можно выразить на языке линейной оболочки системы векторов:

Предложение 1. *Пусть L — линейное пространство над полем K. Тогда*

(1) *вектор $b \in L$ линейно выражается через систему $a_1, a_2, \ldots, a_m$ тогда и только тогда, когда*
$$b \in l(a_1, a_2, \ldots, a_m);$$

(2) *система $b_1, b_2, \ldots, b_n$ векторов пространства L линейно выражается через систему $a_1, a_2, \ldots, a_m$ тогда и только тогда, когда*
$$l(b_1, b_2, \ldots, b_n) \subseteq l(a_1, a_2, \ldots, a_m);$$

(3) *системы $a_1, a_2, \ldots, a_m$ и $b_1, b_2, \ldots, b_n$ эквивалентны тогда и только тогда, когда*
$$l(a_1, a_2, \ldots, a_m) = l(b_1, b_2, \ldots, b_n). \quad \square$$

Следующие простейшие свойства введенного отношения между векторами и системами векторов линейного пространства являются очевидными следствиями предложения 1, хотя могут быть доказаны и непосредственно:

Предложение 2. *Пусть L — линейное пространство над полем K. Тогда*

(1) *каждый вектор системы $a_1, a_2, \ldots, a_m$ линейно выражается через эту систему;*

(2) *если вектор b пространства L линейно выражается через некоторую подсистему системы $a_1, a_2, \ldots, a_m$, то он выражается и через всю систему $a_1, a_2, \ldots, a_m$;*

(3) *если вектор b пространства L линейно выражается через систему $a_1, a_2, \ldots, a_m$, а система $a_1, a_2, \ldots, a_m$ линейно выражается через систему $b_1, b_2, \ldots, b_n$, то вектор b линейно выражается через систему векторов $b_1, b_2, \ldots, b_n$;*

(4) *если система векторов $a_1, a_2, \ldots, a_m$ линейно выражается через систему $b_1, b_2, \ldots, b_n$, а система $b_1, b_2, \ldots, b_n$ линейно выражается через систему $c_1, c_2, \ldots, c_k$, то система $a_1, a_2, \ldots, a_m$ линейно выражается через систему $c_1, c_2, \ldots, c_k$;*

(5) *каждая система векторов пространства L эквивалентна самой себе; если система $a_1, a_2, \ldots, a_m$ эквивалентна системе $b_1, b_2, \ldots, b_n$, то и система $b_1, b_2, \ldots, b_n$ эквивалентна системе $a_1, a_2, \ldots, a_m$; две системы, эквивалентные третьей, эквивалентны между собой.* □

Отметим еще

Предложение 3. *Пусть вектор b пространства L линейно выражается через систему $a_1, a_2, \ldots, a_m$ и не выражается через ее подсистему $a_1, a_2, \ldots, a_{m-1}$. Тогда вектор a_m линейно выражается через систему $a_1, a_2, \ldots, a_{m-1}, b$.* □

Определение 2. *Конечная система векторов линейного пространства называется линейно зависимой, если хотя бы один вектор этой системы линейно выражается через подсистему, составленную из остальных ее векторов. В противном случае система называется линейно независимой.*

Заметим, что из этого определения и соглашений, принятых ранее, следует, в частности, что система, состоящая из одного вектора a, линейно зависима тогда и только тогда, когда $a = 0$: зависимость этой системы означает, что ее единственный вектор a должен линейно выражаться через подсистему, составленную из остальных векторов этой системы, т. е. — через пустую систему. Ясно также, что в соответствии с этим определением пустая система векторов должна считаться линейно независимой, поскольку в ней вообще нет ни одного вектора, а потому нельзя найти вектор, удовлетворяющий требованию определения линейно зависимой системы. Для непустой системы векторов можно дать другое определение понятий линейной зависимости и независимости, равносильное, разумеется, приведенному в определении 2, но более удобное в ряде случаев.

Предложение 4. *Система $a_1, a_2, \ldots, a_m$ (где $m \geq 1$) векторов пространства L является линейно зависимой тогда и только тогда, когда существуют скаляры $\alpha_1, \alpha_2, \ldots, \alpha_m$, хотя бы один из которых отличен от нуля, такие, что выполнено равенство*

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_m a_m = 0.$$

Система $a_1, a_2, \ldots, a_m$ векторов пространства L является линейно независимой тогда и только тогда, когда для любых скаляров $\alpha_1, \alpha_2, \ldots, \alpha_m$ равенство

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_m a_m = 0.$$

возможно лишь в том случае, когда каждый из этих скаляров равен нулю.

Доказательство. Очевидно, что нам достаточно доказать лишь первое из утверждений этого предложения. Предположим сначала, что система $a_1, a_2, \ldots, a_m$ является линейно зависимой. Это означает по определению, что хотя бы один из векторов этой системы линейно выражается через подсистему, составленную из остальных ее векторов; будем для простоты (и без потери общности) считать, что вектор a_1 линейно выражается через систему $a_2, a_3, \ldots, a_m$. Тогда для подходящих скаляров $\alpha_2, \alpha_3, \ldots, \alpha_m$ должно выполняться равенство

$$a_1 = \alpha_2 a_2 + \alpha_3 a_3 + \cdots + \alpha_m a_m.$$

Прибавив к обеим частям его вектор $-a_1$, получаем

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_m a_m = 0,$$

где $\alpha_1 = -1$ и потому отличен от нуля.

Обратно, пусть для некоторых скаляров $\alpha_1, \alpha_2, \ldots, \alpha_m$ имеет место равенство

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_m a_m = 0,$$

причем хотя бы один из этих скаляров — будем, опять-таки для простоты, считать, что это α_1 — отличен от нуля. Прибавив к обеим частям этого равенства вектор $-\alpha_1 a_1$ и умножив их на скаляр $-\alpha_1^{-1}$, получаем выражение

$$a_1 = \left(-\frac{\alpha_2}{\alpha_1}\right) a_2 + \left(-\frac{\alpha_3}{\alpha_1}\right) a_3 + \cdots + \left(-\frac{\alpha_m}{\alpha_1}\right) a_m$$

вектора a_1 через систему $a_2, a_3, \ldots, a_m$, что и доказывает линейную зависимость системы $a_1, a_2, \ldots, a_m$. $\square$

Доставлямый предложением 4 критерий линейной зависимости и линейной независимости непустой конечной системы векторов линейного пространства можно выразить более наглядно. Нулевой вектор пространства можно представить как значение линейной комбинации произвольной системы $a_1, a_2, \ldots, a_m$ векторов этого пространства очевидным способом, взяв в качестве коэффициентов нулевые скаляры:

$$0 = 0a_1 + 0a_2 + \cdots + 0a_m.$$

Будем говорить о таком представлении нулевого вектора, как о тривиальном, и называть линейную комбинацию вида $0a_1 + 0a_2 + \cdots + 0a_m$ тривиальной. Предложение 4 утверждает, что существование нетривиального выражения нулевого вектора через данную систему равносильно ее линейной зависимости. Говоря подробнее, система векторов является линейно зависимой тогда и только тогда, когда некоторая нетривиальная комбинация этой системы равна нулю, и система векторов является линейно независимой тогда и только тогда, когда только тривиальная комбинация этой системы может быть равна нулю.

Из предложения 2 (или из предложения 4) легко следует

Предложение 5. *Всякая подсистема линейно независимой системы векторов линейно независима.* $\square$

Приведем еще полезное

Предложение 6. *Если система $a_1, a_2, \ldots, a_m$ векторов пространства L линейно независима, то для любого вектора $b \in L$ система $a_1, a_2, \ldots, a_m, b$ является линейно зависимой тогда и только тогда, когда вектор b линейно выражается через систему $a_1, a_2, \ldots, a_m$.*

Действительно, достаточность условия очевидна. Обратно, если система a_1, a_2, $\ldots$,a_m, b линейно зависима, то хотя бы один ее вектор должен линейно выражаться через подсистему, составленную из остальных. Утверждается, что одним из таких векторов обязательно является вектор b. В самом деле, пусть один из векторов $a_1, a_2, \ldots, a_m$ линейно выражается через подсистему остальных векторов рассматриваемой системы; для простоты обозначений (и без потери общности) будем считать, что вектор a_1 линейно выражается через систему $a_2, a_3, \ldots, a_m, b$. Поскольку система $a_1, a_2, \ldots, a_m$ линейно независима, вектор a_1 не может выражаться через систему $a_2, a_3, \ldots, a_m$. Поэтому ввиду предложения 3 вектор b линейно выражается через систему $a_1, a_2, \ldots, a_m$. $\square$

Более глубокие свойства линейной зависимости и независимости систем векторов содержатся в следующем утверждении, именуемом обычно теоремой Штейница о замене:

Теорема 1. *Пусть система*

$$a_1, a_2, \ldots, a_m \tag{1}$$

векторов линейного пространства L линейно независима и линейно выражается через систему

$$b_1, b_2, \ldots, b_n. \tag{2}$$

Тогда $m \le n$ и в системе (2) можно найти подсистему из m векторов, заменив которую системой (1), получим систему векторов, эквивалентную системе (2).

Доказательство. Проведем индукцию по числу m векторов системы (1). Все утверждения теоремы очевидны, если $m = 0$, т. е. система (1) пустая. Рассмотрим еще случай $m = 1$ (в данном рассуждении это необязательно, но поможет лучше понять индуктивный шаг). В этом случае линейно независимая система (1) состоит из одного вектора a_1, и этот вектор линейно выражается через систему (2). Поэтому система

$$a_1, b_1, b_2, \ldots, b_n$$

является линейно зависимой, и следовательно найдется номер r, $1 \le r \le n$, такой, что система

$$a_1, b_1, b_2, \ldots, b_{r-1}$$

линейно независима, а система

$$a_1, b_1, b_2, \dots, b_{r-1}, b_r$$

линейно зависима. По предложению 6 вектор b_r линейно выражается через систему $a_1, b_1, b_2, \dots, b_{r-1}$, и потому система

$$b_1, b_2, \dots, b_{r-1}, a_1, b_{r+1}, \dots, b_n, \tag{3}$$

полученная из системы (2) заменой вектора b_r на вектор a_1, эквивалентна, очевидно, системе (2).

Предположим теперь, что $m > 1$ и что для любой пары систем векторов пространства L, удовлетворяющей условиям теоремы, выполняется ее заключение, если число векторов в первой из этих систем меньше, чем m. Это индуктивное предположение применимо, очевидно, к подсистеме $a_1, a_2, \dots, a_{m-1}$ системы (1). Из него следует, что $m - 1 \leq n$ и что система (2) (после подходящей перенумерации ее векторов) эквивалентна системе

$$a_1, a_2, \dots, a_{m-1}, b_m, b_{m+1} \dots, b_n. \tag{4}$$

Так как вектор a_m линейно выражается через систему (2), он должен выражаться и через систему (4). Но этот вектор не может выражаться через ее подсистему $a_1, a_2, \dots, a_{m-1}$, и потому в систему (4) кроме векторов этой подсистемы должен входить еще хотя бы один вектор. Это означает, что $m \leq n$. Кроме того, существует номер r, $m \leq r \leq n$ такой, что вектор a_m линейно выражается через систему

$$a_1, a_2, \dots, a_{m-1}, b_m, b_{m+1} \dots, b_r,$$

но не выражается через систему

$$a_1, a_2, \dots, a_{m-1}, b_m, b_{m+1} \dots, b_{r-1}.$$

Из предложения 3 следует, что вектор b_r линейно выражается через систему

$$a_1, a_2, \dots, a_{m-1}, b_m, b_{m+1} \dots, b_{r-1}, a_m,$$

и потому система (4) эквивалентна системе

$$a_1, a_2, \dots, a_{m-1}, b_m, b_{m+1} \dots, b_{r-1}, a_m, b_{r+1}, \dots, b_n. \tag{5}$$

Так как система (2) эквивалентна системе (5), индуктивный шаг закончен. $\square$

Отметим, прежде всего, очевидное и важное

Следствие 1. *Если две конечные линейно независимые системы векторов линейного пространства эквивалентны, то они содержат одинаковое число векторов.* □

Справедливо также и частичное обращение этого следствия. Предположим, что системы (1) и (2) в теореме Штейница содержат одинаковое число векторов, т. е. $m = n$. Тогда результатом замены в системе (2) ее подсистемы, состоящей из m векторов, на систему (1) является, очевидно, система (1). Таким образом, в этом случае системы (1) и (2) должны быть эквивалентны. Кроме того, система (2) должна быть линейно независимой. В самом деле, в противном случае эта система была бы эквивалентна некоторой своей подсистеме, состоящей из $m-1$ векторов. Ввиду первого утверждения теоремы Штейница это невозможно, так как через эту подсистему выражалась бы тогда система (1). Таким образом, мы доказали

Следствие 2. *Если линейно независимая система $a_1, a_2, \ldots, a_m$ векторов линейного пространства L линейно выражается через систему $b_1, b_2, \ldots, b_m$ с тем же числом векторов, то эти системы эквивалентны и вторая также линейно независима.* □

Распространим теперь понятия линейной выразимости и линейной зависимости на произвольные не обязательно конечные системы векторов.

Определение 3. *Пусть $(a_i)_{i\in I}$ — произвольная система векторов линейного пространства L. Будем говорить, что вектор $b \in L$ линейно выражается через эту систему, если он линейно выражается через некоторую конечную подсистему системы $(a_i)_{i\in I}$.*

Система $(a_i)_{i\in I}$ называется линейно зависимой, если хотя бы один ее вектор линейно выражается через подсистему, состоящую из остальных векторов этой системы. В противном случае система $(a_i)_{i\in I}$ называется линейно независимой.

Легко видеть, что свойства понятия линейной выразимости, перечисленные в предложении 2, остаются верными и в этом более общем случае. То же справедливо и для предложения 5. Заметим, кроме того, что определение линейно зависимой и линейно независимой системы векторов в общем случае дословно совпадает с определением 2, сформулированным для конечных систем. Оно использует понятие выразимости вектора через данную систему, и для его формулировки нам достаточно было договориться о значении этого понятия для произвольных систем векторов. Оказывается, что линейная зависимость или независимость произвольной системы векторов определяется поведением конечных подсистем этой системы.

Предложение 7. *Система $(a_i)_{i\in I}$ векторов линейного пространства L является линейно зависимой тогда и только тогда, когда линейно зависимой является некоторая конечная подсистема этой системы.*

Доказательство. Предположим сначала, что система $(a_i)_{i\in I}$ является линейно зависимой. Тогда существует индекс $i_0 \in I$ такой, что вектор a_{i_0} линейно выражается через систему $(a_i)_{i\in J}$, где $J = I \setminus \{i_0\}$. По определению это означает, что вектор

a_{i_0} выражается через некоторую конечную систему $a_{i_1}, a_{i_2}, \ldots, a_{i_n}$, где $i_1, i_2, \ldots, i_n$ — попарно различные элементы множества J. Очевидно теперь, что $a_{i_0}, a_{i_1}, a_{i_2}, \ldots, a_{i_n}$ — линейно зависимая конечная подсистема системы $(a_i)_{i \in I}$.

Обратно, пусть система $(a_i)_{i \in I}$ содержит некоторую конечную подсистему $a_{j_1}, a_{j_2}, \ldots, a_{j_m}$, которая линейно зависима и потому хотя бы один вектор которой, скажем, a_{j_1} линейно выражается через подсистему $a_{j_2}, a_{j_3}, \ldots, a_{j_m}$ этой системы, составленную из остальных ее векторов. Полагая $J = I \setminus \{j_1\}$, мы видим, что так как $j_2, j_3, \ldots, j_m \in J$, вектор a_{j_1} линейно выражается через конечную подсистему системы $(a_i)_{i \in J}$. Таким образом, в соответствии с определением 3 вектор a_{j_1} линейно выражается через подсистему $(a_i)_{i \in J}$ системы $(a_i)_{i \in I}$, и потому эта система линейно зависима. $\square$

Определение 4. *Пусть $(a_i)_{i \in I}$ — произвольная система векторов линейного пространства L и $J \subseteq I$. Подсистема $(a_i)_{i \in J}$ системы $(a_i)_{i \in I}$ называется максимальной линейно независимой подсистемой системы $(a_i)_{i \in I}$, если она линейно независима и любая подсистема системы $(a_i)_{i \in I}$, содержащая систему $(a_i)_{i \in J}$ и не совпадающая с ней, линейно зависима (т. е. для любого множества индексов K такого, что $J \subset K \subseteq I$, система $(a_i)_{i \in K}$ линейно зависима).*

Теорема 2. *а) Каждая система векторов линейного пространства L обладает максимальной линейно независимой подсистемой. Более того, произвольная линейно независимая подсистема данной системы векторов содержится в некоторой ее максимальной линейно независимой подсистеме.*

б) Произвольная линейно независимая подсистема данной системы векторов тогда и только тогда является ее максимальной линейно независимой подсистемой, когда она эквивалентна этой системе.

в) Если хотя бы одна максимальная линейно независимая подсистема данной системы векторов является конечной, то конечной будет и любая другая максимальная линейно независимая ее подсистема и число векторов во всех максимальных линейно независимых подсистемах этой системы одно и то же.

Доказательство. Поскольку каждая система векторов пространства L обладает хотя бы одной линейно независимой подсистемой (по крайней мере, таковой является пустая подсистема), утверждение а) для конечных систем векторов очевидно: присоединяя по очереди к данной линейно независимой подсистеме не входящие в нее векторы всей системы, мы либо обнаружим линейно независимую подсистему, содержащую на один вектор больше, чем исходная подсистема, либо убедимся в том, что эта подсистема является максимальной линейно независимой. Поэтому через конечное число таких шагов, не превышающее, во всяком случае, количества векторов системы, будет построена ее максимальная линейно независимая подсистема. Для бесконечной системы векторов доказательство утверждения а) проводится с использованием леммы Цорна следующим образом.

Семейство всех линейно независимых подсистем данной системы $(a_i)_{i \in I}$ векторов пространства L частично упорядочено отношением "быть подсистемой". Пусть $(a_i)_{i \in J_\lambda}$ $(\lambda \in \Lambda)$ — произвольная цепь в этом семействе; это означает, что для каждого $\lambda \in \Lambda$ система $(a_i)_{i\in J_\lambda}$ линейно независима и для любых $\lambda_1, \lambda_2 \in \Lambda$ одно из множеств J_{λ_1} и J_{λ_2} содержится в другом. Пусть $J = \bigcup_{\lambda\in\Lambda} J_\lambda$. Покажем, что система $(a_i)_{i\in J}$ линейно независима и, следовательно, является верхней гранью рассматриваемой цепи. В самом деле, если бы эта система оказалась линейно зависимой, то в силу предложения 6 она содержала бы некоторую конечную линейно зависимую подсистему $a_{j_1}, a_{j_2}, \ldots, a_{j_m}$. Здесь каждый индекс j_i содержится в некотором множестве J_{λ_i} $(i = 1, 2, \ldots, m)$. Так как наше семейство является цепью, в этой конечной системе множеств существует наибольшее, скажем, J_{λ_m}, и потому система $a_{j_1}, a_{j_2}, \ldots, a_{j_m}$ является подсистемой системы $(a_i)_{i\in J_{\lambda_m}}$, что противоречит ее линейной независимости. Таким образом, семейство линейно независимых подсистем системы $(a_i)_{i\in I}$ индуктивно упорядочено, и утверждение а) вытекает из леммы Цорна.

Докажем теперь утверждение б). Пусть $(a_i)_{i\in J}$ — линейно независимая подсистема системы $(a_i)_{i\in I}$. Предположим сначала, что эти системы эквивалентны. Пусть множество индексов K таково, что $J \subset K \subseteq I$, и пусть $k \in K \setminus J$. Так как вектор a_k линейно выражается через систему $(a_i)_{i\in J}$, он выражается и через подсистему системы $(a_i)_{i\in K}$, состоящую из всех остальных ее векторов, и потому система $(a_i)_{i\in K}$ линейно зависима. Таким образом, $(a_i)_{i\in J}$ — максимальная линейно независимая подсистема системы $(a_i)_{i\in I}$.

Для доказательства обратного достаточно показать, что если $(a_i)_{i\in J}$ — максимальная линейно независимая подсистема системы $(a_i)_{i\in I}$, то произвольный вектор a_k системы $(a_i)_{i\in I}$ линейно выражается через систему $(a_i)_{i\in J}$. Это очевидно, если $k \in J$. Если же $k \notin J$, то система $(a_i)_{i\in K}$, где $K = J \cup \{k\}$, является линейно зависимой и в силу предложения 7 должна содержать конечную линейно зависимую подсистему. Так как система $(a_i)_{i\in J}$ линейно независима, одним из векторов этой подсистемы является вектор a_k, т. е. она имеет вид $a_k, a_{j_1}, a_{j_2}, \ldots, a_{j_m}$, где $j_1, j_2, \ldots, j_m \in J$. Следовательно, система $a_{j_1}, a_{j_2}, \ldots, a_{j_m}$, являясь подсистемой линейно независмой системы $(a_i)_{i\in J}$, линейно независима, и по предложению 6 вектор a_k должен через нее линейно выражаться, а потому этот вектор линейно выражается и через систему $(a_i)_{i\in J}$.

Утверждение в) легко следует из утверждения б) и теоремы Штейница. В самом деле, если данная система векторов пространства L содержит конечную максимальную линейно независимую подсистему, состоящую, скажем, из n векторов, то поскольку ввиду утверждения б) через эту подсистему линейно выражается любая другая подсистема данной системы, то в силу теоремы Штейница любая конечная линейно независимая подсистема нашей системы не может содержать больше, чем n векторов. Отсюда, в свою очередь, следует, что в данной системе векторов нельзя найти бесконечную линейно независимую подсистему: такая подсистема содержала бы конечные подсистемы со сколь угодно большим числом векторов, причем все они были бы также линейно независимыми. Мы видим, таким образом, что если одна из максимальных линейно независимых подсистем данной системы векторов является конечной, то конечной будет и любая другая максимальная линейно независимая ее подсистема. Остается заметить, что любые две максимальные линейно независимые подсистемы данной системы эквивалентны, так как в силу утверждения б) каждая из них эквивалентна всей системе. Поэтому если они конечны, число векторов в них одно и то же в силу следствия 1 из теоремы Штейница. □

Теорема 2 утверждает, в частности, что любая система векторов линейного пространства обладает максимальными линейно независимыми подсистемами, причем либо все такие ее подсистемы бесконечны, либо все они конечны и содержат одно и тоже число векторов. Это позволяет ввести понятие ранга системы векторов:

Определение 5. *Будем говорить, что система $(a_i)_{i \in I}$ векторов линейного пространства L имеет конечный ранг, если максимальные линейно независимые подсистемы этой системы конечны. При этом число векторов в максимальной линейно независимой подсистеме системы $(a_i)_{i \in I}$ называется рангом этой системы.*

При рассмотрении системы всех векторов линейного пространства для понятий, введенных выше, используется специальная терминология:

Определение 6. *Базисом линейного пространства L называется линейно независимая система порождающих этого пространства.*

Линейное пространство L называется конечномерным, если оно обладает конечным базисом. При этом число векторов базиса пространства L называется размерностью этого пространства и обозначается $\dim L$.

Таким образом, базисом линейного пространства L мы называем такую линейно независимую систему его векторов, через которую линейно выражаются все векторы пространства L, т. е. — максимальную линейно независимую подсистему системы всех векторов этого пространства. Поэтому из теоремы 2 следует, что в каждом линейном пространстве существует базис. Конечномерность пространства равносильна конечности ранга системы всех векторов этого пространства, и в этом случае ранг этой системы (т. е. число векторов в произвольном базисе пространства) мы называем размерностью пространства L. Договоримся еще называть пространство L n-мерным, если его размерность равна n.

В соответствии с определением размерность конечномерного линейного пространства является неотрицательным целым числом. При этом линейное пространство является 0-мерным тогда и только тогда, когда его базисом служит пустая система векторов. Поэтому из наших соглашений следует, что $\dim L = 0$ в точности тогда, когда $L = \{0\}$.

Найдем теперь базисы некоторых линейных пространств, перечисленных в параграфе 1.

1. В параграфе 1 отмечено, что система из двух направленных отрезков пространства V^2, не лежащих на одной прямой, порождает это пространство. Так как такая система является, как легко понять, линейно независимой, она составляет базис пространства V^2, так что это пространство конечномерно и $\dim V^2 = 2$. Аналогично, $\dim V^3 = 3$.

2. Покажем, что система порождающих n-мерного координатного пространства

K^n над полем K

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0, 0) \\ e_2 &= (0, 1, 0, \ldots, 0, 0) \\ e_3 &= (0, 0, 1, \ldots, 0, 0) \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ e_{n-1} &= (0, 0, 0, \ldots, 1, 0) \\ e_n &= (0, 0, 0, \ldots, 0, 1). \end{aligned}$$

является линейно независимой. Если, в самом деле, для некоторых скаляров $\alpha_1, \alpha_2, \ldots, \alpha_n \in K^n$ имеем равенство $\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n = 0$, то поскольку $\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, из него следует, что $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Таким образом, система $e_1, e_2, \ldots, e_n$ является базисом пространства K^n, и потому $\dim K^n = n$. Отметим, что прилагательное "n-мерное", до сих пор лишь формально входившее в определение пространства K^n, приобрело теперь содержательный смысл. Указанный базис пространства K^n будем называть стандартным (или естественным).

3. Так как равенство $\alpha 1 + \beta i = 0$, где α и β — действительные числа, возможно лишь в случае, когда $\alpha = 0$ и $\beta = 0$, то система порождающих 1, i линейного пространства $\mathbb{C}$ над полем $\mathbb{R}$ линейно независима и потому является базисом этого пространства. Следовательно, $\dim \mathbb{C} = 2$.

4.Для любого натурального числа n система $1, x, x^2, \ldots, x^n$ векторов пространства $K[x]$ многочленов от переменной x над полем K линейно независима. В самом деле, равенство $\alpha_0 1 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_n x^n = 0$ означает, что многочлен $\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_n x^n$ является нулевым, т. е. все его коэффициенты равны нулю. Так как любая конечная подсистема бесконечной системы $1, x, x^2, \ldots$ содержится в некоторой подсистеме $1, x, x^2, \ldots, x^n$ и потому линейно независима, то и вся система $1, x, x^2, \ldots$, порождающая пространство $K[x]$, линейно независима. Следовательно, система $1, x, x^2, \ldots$ является базисом пространства $K[x]$, и это пространство не является конечномерным. Одновременно мы видим, что $\dim K_n[x] = n$.

5. Найдем, наконец, базис пространства $M_{mn}(K)$ всех $m \times n$-матриц над полем K. Пусть для $i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n$ символ e_{ij} обозначает $m \times n$-матрицу, у которой на пересечении i-ой строки и j-го столбца стоит единица, а остальные вхождения равны нулю. Для любой матрицы $a = (\alpha_{ij})$ пространства $M_{mn}(K)$ имеем, очевидно,

$$a = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} e_{ij}.$$

Из этого равенства следует, что система матриц e_{ij} $(i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n)$ порождает пространство $M_{mn}(K)$ и что лишь тривиальная линейная комбинация этой системы может быть равна нулю. Таким образом, она является базисом пространства $M_{mn}(K)$ и $\dim M_{mn}(K) = mn$.

В следующем утверждении собраны некоторые свойства конечномерных линейных пространств.

Теорема 3. *а) Пусть S — система порождающих линейного пространства L и S' — максимальная линейно независимая подсистема системы S. Тогда S' является базисом пространства L. В частности, произвольное линейное пространство с конечной системой порождающих является конечномерным, а его размерность совпадает с рангом данной системы порождающих.*

б) Произвольная линейно независимая система векторов линейного пространства может быть дополнена до базиса этого пространства.

в) Если $b_1, b_2, \ldots, b_m$ — линейно независимая система векторов n-мерного пространства L, то $m \leq n$ и $m = n$ в том и только в том случае, когда система $b_1, b_2, \ldots, b_m$ является базисом пространства L.

г) Если $a_1, a_2, \ldots, a_n$ — система порождающих n-мерного пространства L, то эта система является базисом пространства L.

Доказательство. Утверждение а) вполне очевидно: по теореме 2 система S' эквивалентна системе порождающих S пространства L и потому сама является (линейно независимой) системой порождающих этого пространства.

Утверждение б) содержится непосредственно в утверждении а) теоремы 2, примененном к системе всех векторов пространства L. Тем не менее, для случая конечномерного пространства приведем здесь еще одно доказательство, указывающее заодно и конкретный алгоритм нахождения базиса пространства, включающего заданную линейно независимую систему векторов этого пространства.

Пусть $b_1, b_2, \ldots, b_m$ — данная линейно независимая система векторов n-мерного пространства L и пусть $a_1, a_2, \ldots, a_n$ — некоторый базис этого пространства. Пусть

$$S = (b_1, b_2, \ldots, b_m, a_1, a_2, \ldots, a_n)$$

— соединение этих систем и пусть

$$S' = (b_1, b_2, \ldots, b_m, a_{i_1}, \ldots, a_{i_{n-m}})$$

— максимальная линейно независимая подсистема системы S, содержащая систему $b_1, b_2, \ldots, b_m$ (Для записи системы S' в таком виде придется, возможно, перенумеровать элементы системы $a_1, a_2, \ldots, a_n$.) Так как S является, очевидно, системой порождающих пространства L, то в силу утверждения а) система S' — искомый базис пространства L.

Утверждение в) вытекает непосредственно из утверждения б): так как систему $b_1, b_2, \ldots, b_m$ можно дополнить до некоторого базиса пространства L и этот базис состоит из n векторов, то очевидно, что $m \leq n$. Если при этом $m = n$, то дополнительных векторов в этом базисе оказаться не может, а это означает, что уже исходная система $b_1, b_2, \ldots, b_m$ является базисом нашего пространства. Обратно, если эта система является базисом, то $m = n$, так как число векторов произвольного базиса n-мерного пространства равно n.

Аналогичные соображения показывают, что утверждение г) следует непосредственно из утверждения а): число векторов в максимальной линейно независимой подсистеме системы $a_1, a_2, \ldots, a_n$ должно совпадать с размерностью пространства L, т. е. — с числом n. Следовательно, эта система является линейно независимой, а потому — базисом пространства L. $\square$

Выделим важный частный случай утверждения в) доказанной теоремы.

Следствие 1. *В n-мерном линейном пространстве любая система векторов, число элементов которой равно $n+1$, линейно зависима.* $\square$

Если мы хотим убедиться в том, что данная система векторов линейного пространства является его базисом, то нам следует доказать, что эта система обладает двумя свойствами, требуемыми определением 6: во-первых, она должна быть линейно независимой, а во-вторых, — порождать все пространство. Утверждения в) и г) теоремы 3 показывают, что если нам известна размерность пространства и число векторов данной системы совпадает с этой размерностью, то достаточно проверить любое одно из этих свойств. Иначе говоря, имеет место

Следствие 2. *В n-мерном линейном пространстве L для любой системы векторов $a_1, a_2, \ldots, a_n$, содержащей n векторов, следующие утверждения равносильны:*
а) система $a_1, a_2, \ldots, a_n$ линейно независима;
б) система $a_1, a_2, \ldots, a_n$ порождает пространство L. $\square$

Отметим еще

Следствие 3. *Пусть L — конечномерное линейное пространство и M — подпространство пространства L. Тогда пространство M также является конечномерным и* $\dim M \le \dim L$. *При этом,* $\dim M = \dim L$ *тогда и только тогда, когда* $M = L$.

В самом деле, базис пространства M (существующий в силу теоремы 2) является линейно независимой системой элементов пространства L. Поэтому все сформулированные утверждения являются следствиями утверждения в) теоремы 3. $\square$

Рассмотрим, далее, как ведет себя размерность при суммировании подпространств. Отметим, прежде всего, почти очевидное

Предложение 8. *Пусть линейное пространство L является суммой подпространств $L_1, L_2, \ldots, L_n$. Если каждое из пространств $L_1, L_2, \ldots, L_n$ конечномерно, то конечномерным является и пространство L, причем*

$$\dim L \le \sum_{i=1}^{n} \dim L_i.$$

Действительно, соединение S базисов $S_1, S_2, \ldots, S_n$ подпространств $L_1, L_2, \ldots, L_n$ соответственно является в силу следствия к теореме 1.2 системой порождающих пространства $L = \sum_{i=1}^{n} L_i$, и нам остается сослаться на утверждение а) теоремы 3. $\square$

Для размерности суммы двух подпространств можно получить более точное выражение, часто называемое равенством Грассмана:

Теорема 4. *Пусть подпространства L_1 и L_2 линейного пространства L конечномерны. Тогда их сумма $L_1 + L_2$ и пересечение $L_1 \cap L_2$ также являются конечномерными пространствами и*

$$\dim\,(L_1 + L_2) = \dim L_1 + \dim L_2 - \dim\,(L_1 \cap L_2)\,. \tag{6}$$

Доказательство. Конечномерность подпространства $L_1 + L_2$ содержится в предложении 8, а конечномерность подпространства $L_1 \cap L_2$, лежащего, например, в конечномерном подпространстве L_1, вытекает из следствия 3 теоремы 3.

Обозначим через k размерность подпространства $L_1 \cap L_2$ и выберем некоторый базис $c_1, c_2, \ldots, c_k$ этого подпространства. Система $c_1, c_2, \ldots, c_k$ является линейно независимой системой векторов пространства L_1 и потому может быть дополнена до некоторого базиса

$$a_1, a_2, \ldots, a_m, c_1, c_2, \ldots, c_k \tag{7}$$

этого пространства. Точно так же дополним эту систему до базиса

$$b_1, b_2, \ldots, b_n, c_1, c_2, \ldots, c_k \tag{8}$$

пространства L_2. Из этих обозначений следует, в частности, что $\dim L_1 = m + k$ и $\dim L_2 = n + k$. Поэтому для доказательства равенства (6) достаточно показать, что система

$$a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_n, c_1, c_2, \ldots, c_k \tag{9}$$

является базисом подпространства $L_1 + L_2$.

Покажем сначала, что векторы системы (9) порождают это подпространство, т. е. что $L_1 + L_2$ совпадает с линейной оболочкой

$$l(a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_n, c_1, c_2, \ldots, c_k)$$

этой системы. Так как система (7) является подсистемой системы (9), линейная оболочка первой содержится в линейной оболочке второй. Таким образом, подпространство L_1 содержится в линейной оболочке системы (9). То же справедливо и для подпространства L_2, и из теоремы 1.2 следует, что

$$L_1 + L_2 \subseteq l(a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_n, c_1, c_2, \ldots, c_k).$$

С другой стороны, каждый вектор системы (9) входит или в подпространство L_1, или в подпространство L_2, и потому система (9) содержится в подпространстве $L_1 + L_2$. Отсюда и из теоремы 1.1 следует противоположное включение, и совпадение подпространства $L_1 + L_2$ с линейной оболочкой системы (9) доказано.

Остается установить линейную независимость системы (9). Для этого предположим, что для некоторых скаляров $\alpha_1, \alpha_2 \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n, \gamma_1, \gamma_2, \ldots, \gamma_k$ выполнено равенство

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_m a_m + \beta_1 b_1 + \beta_2 b_2 + \cdots + \beta_n b_n + \gamma_1 c_1 + \gamma_2 c_2 + \cdots + \gamma_k c_k = 0. \tag{10}$$

Введя обозначение $a = \alpha_1 a_1 + \alpha_2 a_2 + \dots + \alpha_m a_m$, мы видим, что $a \in L_1$, но поскольку

$$a = (-\beta_1)b_1 + (-\beta_2)b_2 + \dots + (-\beta_n)b_n + (-\gamma_1)c_1 + (-\gamma_2)c_2 + \dots + (-\gamma_k)c_k,$$

имеем также $a \in L_2$. Следовательно, вектор a содержится в подпространстве $L_1 \cap L_2$ и потому должен выражаться через базис $c_1, c_2, \dots, c_k$ этого подпространства. Значит для подходящих скаляров $\delta_1, \delta_2, \dots, \delta_k$ должно выполняться равенство

$$a = \delta_1 c_1 + \delta_2 c_2 + \dots + \delta_k c_k,$$

или, что равносильно,

$$\alpha_1 a_1 + \alpha_2 a_2 + \dots + \alpha_m a_m + (-\delta_1)c_1 + (-\delta_2)c_2 + \dots + (-\delta_k)c_k = 0.$$

Так как система (7) линейно независима, отсюда следует, в частности, что $\alpha_1 = \alpha_2 = \dots = \alpha_m = 0$, и равенство (10) принимает вид

$$\beta_1 b_1 + \beta_2 b_2 + \dots + \beta_n b_n + \gamma_1 c_1 + \gamma_2 c_2 + \dots + \gamma_k c_k = 0.$$

Из линейной независимости системы (8) теперь получаем

$$\beta_1 = \beta_2 = \dots = \beta_n = \gamma_1 = \gamma_2 = \dots = \gamma_k = 0,$$

и линейная независимость системы (9) доказана. Итак, эта система является базисом подпространства $L_1 + L_2$, и теорема доказана. $\square$

Доказанная теорема дает еще один критерий того, когда сумма двух (конечномерных) пространств является прямой:

Следствие. *Сумма $L_1 + L_2$ конечномерных подпространств L_1 и L_2 линейного пространства L является прямой тогда и только тогда, когда $\dim\,(L_1 + L_2) = \dim L_1 + \dim L_2$.*

В самом деле, в силу следствия 1 к теореме 1.3 равенство $L_1 + L_2 = L_1 \oplus L_2$ равносильно условию $L_1 \cap L_2 = \{0\}$, что, в свою очередь, равносильно тому, что $\dim\,(L_1 \cap L_2) = 0$. А это ввиду равенства Грассмана равносильно равенству из формулировки следствия. $\square$

Аналогичный критерий имеет место и для произвольного числа слагаемых. На самом деле, с помощью понятий базиса и размерности требование единственности в определении прямой суммы подпространств можно в дополнение к теореме 1.3 выразить еще двумя способами.

Теорема 5. *Пусть конечномерное линейное пространство L является суммой подпространств $L_1, L_2, \dots, L_n$. Тогда следующие условия равносильны:*

(1) $L = \bigoplus_{i=1}^n L_i$;
(2) *соединение произвольных базисов подпространств $L_1, L_2, \dots, L_n$ является базисом пространства L;*
(3) $\dim L = \sum_{i=1}^n \dim L_i$.

Доказательство. Покажем сначала, что из условия (1) следует условие (2). Пусть $S_i = (a_{i1}, a_{i2}, \ldots, a_{im_i})$ — базис подпространства L_i $(i = 1, 2, \ldots, n)$. Так как $L = \sum_{i=1}^{n} L_i$, в силу следствия к теореме 1.2 соединение

$$S = (a_{11}, a_{12}, \ldots, a_{1m_1}, a_{21}, a_{22}, \ldots, a_{2m_2}, \ldots, a_{n1}, a_{n2}, \ldots, a_{nm_n})$$

систем $S_1, S_2, \ldots, S_n$ является системой порождающих пространства L. Покажем, что система S линейно независима. Пусть для некоторых скаляров α_{ij} $(j = 1, 2, \ldots, m_i,\ i = 1, 2, \ldots, n)$ имеет место равенство

$$\alpha_{11}a_{11} + \alpha_{12}a_{12} + \cdots + \alpha_{1m_1}a_{1m_1} + \alpha_{21}a_{21} + \cdots + \alpha_{nm_n}a_{nm_n}. \tag{11}$$

Полагая для каждого $i = 1, 2, \ldots, n$ $b_i = \sum_{j=1}^{m_i} \alpha_{ij}a_{ij}$, мы перепишем равенство (11) в виде $b_1 + b_2 + \cdots + b_n$, и так как $b_i \in L_i$ и $L = \bigoplus_{i=1}^{n} L_i$, из теоремы 1.3 следует, что $b_1 = b_2 = \cdots = b_n = 0$, т. е. для каждого $i = 1, 2, \ldots, n$

$$\alpha_{i1}a_{i1} + \alpha_{i2}a_{i2} + \cdots + \alpha_{im_i}a_{im_i} = 0.$$

Поскольку система S_i линейно независима, отсюда следует, что $\alpha_{ij} = 0$ $(j = 1, 2, \ldots, m_i)$. Таким образом, все коэффициенты в равенстве (11) равны 0, и линейная независимость системы S доказана; тем самым доказано, что эта система является базисом пространства L, и закончено доказательство выполнимости условия (2).

То, что из условия (2) следует условие (3), очевидно. Покажем теперь, что из условия (3) следует условие (1), а именно, индукцией по числу n подпространств $L_1, L_2, \ldots, L_n$ покажем, что если $\dim L = \sum_{i=1}^{n} \dim L_i$, то $L = \bigoplus_{i=1}^{n} L_i$. При $n = 1$ это тривиально, а при $n = 2$ содержится в следствии к теореме 4.

Пусть $n \geq 2$ и пусть для любого конечномерного пространства, являющегося суммой меньшего, чем n, числа подпространств, доказываемое утверждение справедливо. Покажем, что тогда оно справедливо и для пространства $L = \sum_{i=1}^{n} L_i$. Для этого обозначим через M подпространство, порожденное первыми $n - 1$ слагаемыми этой суммы, т. е. $M = \sum_{i=1}^{n-1} L_i$; ясно, что тогда $L = M + L_n$. Так как по предложению 8 $\dim M \leq \sum_{i=1}^{n-1} \dim L_i$, то используя равенство Грассмана, получаем

$$\sum_{i=1}^{n} \dim L_i = \dim L = \dim M + \dim L_n - \dim (M \cap L_n) \leq$$
$$\sum_{i=1}^{n-1} \dim L_i + \dim L_n - \dim (M \cap L_n) = \sum_{i=1}^{n} \dim L_i - \dim (M \cap L_n).$$

Отсюда следует, что $\dim (M \cap L_n) \leq 0$, и так как размерность любого линейного пространства является неотрицательным числом, имеем $\dim (M \cap L_n) = 0$, т.е. $M \cap L_n = \{0\}$. Это означает, что $L = M \oplus L_n$, и по следствию к теореме 4 $\dim L = \dim M + \dim L_n$. Поэтому $\dim M = \sum_{i=1}^{n-1} \dim L_i$, и по индуктивному предположению $M = \bigoplus_{i=1}^{n-1} L_i$. Следствие 4 к теореме 1.3 теперь позволяет утверждать, что $L = \bigoplus_{i=1}^{n} L_i$, и индуктивный шаг закончен. $\square$

Из доказанной теоремы следует, что произвольному разбиению базиса конечномерного линейного пространства на непересекающиеся части соответствует разложение этого пространства в прямую сумму подпространств, порождаемых этими частями. В частности, отметим

Следствие. *Пусть $a_1, a_2, \dots, a_n$ — базис линейного пространства L. Тогда*

(1) $L = \bigoplus_{i=1}^n L_i$, *где* $L_i = l(a_i)$ *— подпространство, порожденное вектором* a_i $(i = 1, 2, \dots, n)$*;*

(2) *если для некоторого* s, $1 \le s < n$ $M = l(a_1, \dots, a_s)$ *и* $N = l(a_{s+1}, \dots a_n)$, *то* $L = M \oplus N$. $\square$

Перейдем теперь к изучению поведения понятий линейной зависимости и размерности при линейных отображениях линейных пространств. Приведем, прежде всего, почти очевидное

Предложение 9. *Пусть* $\varphi : L \to M$ *— линейное отображение линейного пространства* L *в линейное пространство* M*. Если система* $a_1, a_2, \dots, a_n$ *векторов пространства* L *линейно зависима, то линейно зависимой является и система их образов* $a_1\varphi, a_2, \varphi \dots, a_n\varphi$*. Если система* $a_1, a_2, \dots, a_n$ *линейно независима, а отображение* φ *инъективно, то и система* $a_1\varphi, a_2, \varphi \dots, a_n\varphi$ *является линейно независимой.*

В самом деле, напомним, что в силу предложения 2.3 образ относительно отображения φ линейной комбинации произвольной системы векторов пространства L равен линейной комбинации образов векторов этой системы, т. е. для любых скаляров $\alpha_1, \alpha_2, \dots, \alpha_n$ имеет место равенство

$$(\alpha_1 a_1 + \alpha_2 a_2 + \dots + \alpha_n a_n)\varphi = \alpha_1(a_1\varphi) + \alpha_2(a_2\varphi) + \dots + \alpha_n(a_n\varphi).$$

Поскольку, к тому же, $0\varphi = 0$, то из равенства

$$\alpha_1 a_1 + \alpha_2 a_2 + \dots + \alpha_n a_n = 0$$

следует, равенство

$$\alpha_1(a_1\varphi) + \alpha_2(a_2\varphi) + \dots + \alpha_n(a_n\varphi) = 0,$$

а если φ инъективно, то справедливо и обратное. Таким образом, возможность нетривиального выражения нулевого вектора пространства L через систему $a_1, a_2, \dots, a_n$ означает существование такого же выражения (с теми же коэффициентами) нулевого вектора пространства M через систему $a_1\varphi, a_2, \varphi \dots, a_n\varphi$. Если φ иннъективно, то верно и обратное: из линейной зависимости второй системы следует линейная зависимость первой. $\square$

В связи с только что доказанным утверждением следует заметить, что если отображение φ не является инъективным, то в пространстве L всегда можно указать линейно независимую систему векторов, образы которых составляют систему линейно зависимую: простейшим таким примером служит система, состоящая из одного ненулевого вектора, взятого из ядра отображения φ.

Предложение 10. *Пусть $\varphi : L \to M$ — линейное отображение линейного пространства L в линейное пространство M и пусть пространство L n-мерно. Тогда подпространства $\operatorname{Im}\varphi$ и $\operatorname{Ker}\varphi$ являются конечномерными и размерность каждого из них не превосходит числа n.*

Действительно, поскольку $\operatorname{Ker}\varphi$ является подпространством пространства L, в этом случае достаточно вспомнить следствие 3 к теореме 3. Далее, ввиду следствия 2 к теореме 2.1 подпространство $\operatorname{Im}\varphi$ пространства M порождается образами векторов некоторого базиса пространства L, и требуемое утверждение о размерности этого подпространства получается из п. а) теоремы 3. □

Утверждение, сформулированное в предложении 10, позволяет ввести следующие числовые характеристики линейного отображения конечномерного линейного пространства в некоторое другое пространство:

Определение 7. *Пусть $\varphi : L \to M$ — линейное отображение линейного пространства L в линейное пространство M, причем пространство L конечномерно. Рангом отображения φ называется размерность образа этого отображения, т. е. число $\dim(\operatorname{Im}\varphi)$. Дефектом отображения φ называется размерность ядра этого отображения, т. е. число $\dim(\operatorname{Ker}\varphi)$.*

Следующее утверждение уточняет оценки из предложения 10.

Теорема 6. *Сумма ранга и дефекта произвольного линейного отображения конечномерного линейного пространства L в некоторое пространство M равна размерности пространства L.*

Доказательство. Пусть $\varphi : L \to M$ — линейное отображение конечномерного линейного пространства L в линейное пространство M. Обозначим через r ранг отображения φ, $r = \dim(\operatorname{Im}\varphi)$, и выберем базис $b_1, b_2, \ldots, b_r$ подпространства $\operatorname{Im}\varphi$ пространства M. У каждого из векторов $b_1, b_2, \ldots, b_r$ (как и у любого вектора из $\operatorname{Im}\varphi$) имеется прообраз относительно отображения φ, и поэтому для каждого $i = 1, 2, \ldots, r$ можно выбрать вектор $a_i \in L$ такой, что $a_i\varphi = b_i$. Так как система $b_1, b_2, \ldots, b_r$ линейно независима, то ввиду предложения 9 и система $a_1, a_2, \ldots, a_r$ линейно независима. Поэтому если $A = l(a_1, a_2, \ldots, a_r)$ — линейная оболочка системы $a_1, a_2, \ldots, a_r$, то эта система является базисом подпространства A и потому $\dim A = r$.

Покажем, что пространство L является прямой суммой подпространств A и $\operatorname{Ker}\varphi$. Заметим сразу же, что тем самым ввиду теоремы 5 будет доказано равенство

$$\dim L = \dim A + \dim(\operatorname{Ker}\varphi),$$

и так как $\dim A = \dim(\operatorname{Im}\varphi)$, будет доказано и утверждение теоремы.

Пусть u — произвольный вектор пространства L. Тогда его образ $u\varphi$ является элементом подпространства $\operatorname{Im}\varphi$ и потому линейно выражается через его базис $b_1, b_2, \ldots, b_r$. Следовательно, для подходящих скаляров $\beta_1, \beta_2, \ldots, \beta_r$ имеет место равенство

$$u\varphi = \beta_1 b_1 + \beta_2 b_2 + \cdots + \beta_r b_r.$$

Рассмотрим вектор a, равный линейной комбинации с теми же коэффициентами векторов системы $a_1, a_2, \ldots, a_r$:

$$a = \beta_1 a_1 + \beta_2 a_2 + \cdots + \beta_r a_r.$$

Ясно, что $a \in A$. Кроме того, имеем, очевидно,

$$a\varphi = \beta_1(a_1\varphi) + \beta_2(a_2\varphi) + \cdots + \beta_r(a_r\varphi)\beta_1 b_1 + \beta_2 b_2 + \cdots + \beta_r b_r = u\varphi,$$

и потому теорема 2.2 (или непосредственная проверка) говорит о том, что вектор $c = u - a$ принадлежит ядру $\operatorname{Ker}\varphi$ отображения φ. Таким образом, вектор u раскладывается в сумму двух слагаемых $u = a + c$, где $a \in A$ и $c \in \operatorname{Ker}\varphi$. Этим доказано, что $L = A + \operatorname{Ker}\varphi$.

Покажем теперь, что пересечение подпространств A и $\operatorname{Ker}\varphi$ совпадает с нулевым подпространством. Пусть u — произвольный вектор из этого пересечения. Тогда $u \in A$, и потому

$$u = \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_r a_r$$

для подходящих скаляров $\alpha_1, \alpha_2, \ldots, \alpha_r$. Но вектор u входит также и в подпространство $\operatorname{Ker}\varphi$, и потому образ его относительно φ равен нулю. Так как

$$u\varphi = \alpha_1(a_1\varphi) + \alpha_2(a_2\varphi) + \cdots + \alpha_r(a_r\varphi) = \alpha_1 b_1 + \alpha_2 b_2 + \cdots + \alpha_r b_r,$$

получаем равенство

$$\alpha_1 b_1 + \alpha_2 b_2 + \cdots + \alpha_r b_r = 0.$$

Из линейной независимости системы $b_1, b_2, \ldots, b_r$ теперь следует, что $\alpha_1 = \alpha_2 = \cdots = \alpha_r = 0$, откуда $u = 0$. Таким образом, $A \cap \operatorname{Ker}\varphi = 0$, и ввиду следствия 1 к теореме 1.3 $L = A \oplus \operatorname{Ker}\varphi$. $\square$

Из этой теоремы следует, что линейные отображения конечномерных линейных пространств ведут себя так же, как отображения конечных множеств:

Следствие 1. *Пусть $\varphi : L \to M$ — линейное отображение конечномерного линейного пространства L в конечномерное линейное пространство M, причем $\dim L = \dim M$. Отображение φ является инъективным тогда и только тогда, когда оно сюръективно.*

В самом деле, инъективность отображения φ ввиду теоремы 2.2 равносильна тому, что его ядро является нулевым подпространством, т. е. тому, что $\dim(\operatorname{Ker}\varphi) = 0$. А это, в свою очередь, ввиду теоремы 6 означает, что $\dim(\operatorname{Im}\varphi) = \dim L = \dim M$, т. е. (см. следствие 3 к теореме 3) $L = M$. $\square$

Доказанное следствие допускает равносильную формулировку, известную как конечномерная альтернатива Фредгольма:

Следствие 2. *Пусть $\varphi : L \to M$ — линейное отображение конечномерного линейного пространства L в конечномерное линейное пространство M, причем $\dim L = \dim M$. Тогда выполняется одно и только одно из следующих двух утверждений:*

(1) *уравнение $x\varphi = b$ имеет решение для любого вектора $b \in M$;*
(2) *уравнение $x\varphi = 0$ имеет хотя бы одно ненулевое решение.* $\square$

В заключение этого параграфа укажем на универсальный способ построения линейных отображений конечномерных линейных пространств.

Пусть L и M — линейные пространства над полем K,

$$a_1, a_2, \ldots, a_n \quad \text{и} \quad b_1, b_2, \ldots, b_n$$

— две системы векторов пространств L и M соответственно. Возникает естественный вопрос, существует ли линейное отображение $\varphi : L \to M$ такое, что для любого $i = 1, 2, \ldots, n$ имеет место равенство $a_i\varphi = b_i$. Из предложения 9 немедленно следует, что если система $a_1, a_2, \ldots, a_n$ линейно зависима, а система $b_1, b_2, \ldots, b_n$ независима, то такое линейное отображение существовать не может. Следующее утверждение показывает, в частности, что в том случае, когда система $a_1, a_2, \ldots, a_n$ линейно независима, такое отображение существует для любой системы $b_1, b_2, \ldots, b_n$.

Предложение 11. *Пусть L и M — линейные пространства над полем K, $a_1, a_2, \ldots, a_m$ — система векторов пространства L, $b_1, b_2, \ldots, b_m$ — система векторов пространства M. Если система $a_1, a_2, \ldots, a_m$ порождает пространство L, то существует не более одного линейного отображения $\varphi : L \to M$, такого, что для любого $i = 1, 2, \ldots, m$ выполнено равенство $a_i\varphi = b_i$. Если, к тому же, система $a_1, a_2, \ldots, a_m$ линейно независима (т.е. является базисом пространства L), то при любых $b_1, b_2, \ldots, b_m$ линейное отображение с указанным свойством существует.*

Доказательство. Предположим сначала, что φ и ψ — два линейных отображения пространства L в пространство M таких, что для любого $i = 1, 2, \ldots, m$

$$a_i\varphi = b_i \quad \text{и} \quad a_i\psi = b_i.$$

Если система $a_1, a_2, \ldots, a_m$ порождает пространство L, то произвольный вектор $a \in L$ можно записать в виде линейной комбинации $a = \sum_{i=1}^{m} \beta_i a_i$. Тогда

$$a\varphi = \sum_{i=1}^{m} \beta_i (a_i\varphi) = \sum_{i=1}^{m} \beta_i (a_i\psi) = a\psi.$$

Следовательно, отображения φ и ψ одинаково действуют на каждый вектор пространства L и потому $\varphi = \psi$.

Пусть теперь система $a_1, a_2, \ldots, a_m$ является базисом пространства L. Определим отображение φ множества L в множество M следующим образом. Пусть $a = \sum_{i=1}^{m} \lambda_i a_i$ выражение вектора $a \in L$ в виде линейной комбинации системы $a_1, a_2, \ldots, a_m$;

нетрудно видеть, что ввиду линейной независимости этой системы такое выражение вектора a является единственным. Поэтому полагая $a\varphi = \sum_{i=1}^{m} \lambda_i b_i$, мы получаем вполне определенное отображение L в M.

Вычислим, например образ относительно этого отображения вектора a_1. Поскольку $a_1 = 1a_1 + 0a_2 + \cdots + 0a_m$, то в соответствии с нашим определением имеем, очевидно, $a_1\varphi = 1b_1 + 0b_2 + \cdots + 0b_m = b_1$. Ясно, что и для любого $i = 1, 2, \ldots, m$ отображение φ вектор a_i переводит в вектор b_i, и нам остается лишь проверить его линейность. Взяв произвольные векторы $a = \sum_{i=1}^{m} \alpha_i a_i$ и $b = \sum_{i=1}^{m} \beta_i a_i$ пространства L и произвольные скаляры $\alpha, \beta \in L$, получаем $\alpha a + \beta b = \sum_{i=1}^{m} (\alpha\alpha_i + \beta\beta_i) a_i$ и потому

$$(\alpha a + \beta b)\varphi = \sum_{i=1}^{m} (\alpha\alpha_i + \beta\beta_i) b_i = \alpha \sum_{i=1}^{m} \alpha_i b_i + \beta \sum_{i=1}^{m} \beta_i b_i = \alpha(a\varphi) + \beta(b\varphi).$$

Таким образом, отображение φ действительно является линейным, и все утверждения теперь доказаны. $\square$

§ 4. КООРДИНАТЫ ВЕКТОРА И МАТРИЦА ЛИНЕЙНОГО ОТОБРАЖЕНИЯ. КЛАССИФИКАЦИЯ ЛИНЕЙНЫХ ПРОСТРАНСТВ

В этом параграфе рассматриваются только конечномерные линейные пространства. Мы увидим, что зафиксировав в таком пространстве некоторый базис, можно ввести связанную с этим базисом систему координат аналогично тому, как это делается в элементарной геометрии. Затем будет показано, что каждому линейному отображению одного линейного пространства в другое можно, зафиксировав некоторые базисы этих пространств, сопоставить вполне определенную матрицу. Это позволит сводить решение ряда задач о векторах и линейных отображениях к соответствующим задачам о пространстве K^n и использовать вычислительные методы, разработанные для этого пространства.

Начнем с необходимой нам здесь формулировки уже упоминавшегося очевидного свойства базиса конечномерного линейного пространства.

Предложение 1. *Пусть система $S = (e_1, e_2, \dots, e_n)$ является базисом линейного пространства L над полем K. Тогда каждый вектор $a \in L$ может быть выражен в виде линейной комбинации системы S единственным образом.* $\square$

Заметим, что справедливо и обратное утверждение. Более того, можно утверждать даже, что если хотя бы один вектор линейного пространства однозначно выражается через некоторую систему его порождающих, то эта система является базисом пространства.

Определение 1. *Пусть $S = (e_1, e_2, \dots, e_n)$ — базис линейного пространства L над полем K. Пусть*

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n$$

— (единственное) выражение вектора $a \in L$ в виде линейной комбинации системы S. Тогда элемент

$$(a)_S = (\alpha_1, \alpha_2, \dots, \alpha_n)$$

пространства K^n называется координатной строкой вектора a в базисе S, а компоненты вектора $(a)_S$ называются координатами вектора a в этом базисе.

Таким образом, каждому вектору a пространства L ставится в соответствие вполне определенный элемент $(a)_S$ n-мерного координатного пространства K^n (вот и оправдание прилагательного "координатное" в названии этого пространства). Это соответствие вполне определяется тем, что равенства $a = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n$ и $(a)_S = (\alpha_1, \alpha_2, \dots, \alpha_n)$ равносильны. Некоторые свойства введенного соответствия содержатся в следующем утверждении.

Предложение 2. *Пусть $S = (e_1, e_2, \dots, e_n)$ — базис пространства L над полем K. Отображение $\varphi : L \to K^n$, сопоставляющее вектору $a \in L$ его координатную* *строку $(a)_S$ в базисе S, биективно. Для любых векторов $a, b \in L$ и любого скаляра $\lambda \in K$ имеют место равенства*

$$(a+b)_S = (a)_S + (a)_S \quad \text{и} \quad (\lambda a)_S = \lambda(a)_S.$$

Короче говоря, отображение φ является изоморфизмом пространства L на пространство K^n.

Доказательство. Предположим, что для некоторых векторов $a, b \in L$ имеет место равенство $a\varphi = b\varphi$. По определению отображения φ это означает, что если

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n \quad \text{и} \quad b = \beta_1 e_1 + \beta_2 e_2 + \cdots + \beta_n e_n$$

— выражения этих векторов в виде линейной комбинации элементов базиса S, то

$$a\varphi = (a)_S = (\alpha_1, \alpha_2, \ldots, \alpha_n), \quad b\varphi = (b)_S = (\beta_1, \beta_2, \ldots, \beta_n)$$

и потому в пространстве K^n выполнено равенство

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\beta_1, \beta_2, \ldots, \beta_n).$$

Поскольку два вектора из K^n равны тогда и только тогда, когда равны их соответствующие компоненты, отсюда следует, что для любого $i = 1, 2, \ldots, n$ $\alpha_i = \beta_i$, а значит $a = b$. Мы доказали, таким образом, что отображение φ инъективно.

Пусть теперь $(\gamma_1, \gamma_2, \ldots, \gamma_n)$ — произвольный элемент пространства K^n. Полагая

$$c = \gamma_1 e_1 + \gamma_2 e_2 + \cdots + \gamma_n e_n,$$

имеем, очевидно, $c\varphi = (c)_S = (\gamma_1, \gamma_2, \ldots, \gamma_n)$. Тем самым доказана сюръективность, а значит и биективность отображения φ.

Сохраняя введенные выше обозначения для координатных строк векторов a и b, далее имеем

$$\begin{aligned} a+b &= (\alpha_1+\beta_1)e_1 + (\alpha_2+\beta_2)e_2 + \cdots + (\alpha_n+\beta_n)e_n \quad \text{и} \\ \lambda a &= (\lambda\alpha_1)e_1 + (\lambda\alpha_2)e_2 + \cdots + (\lambda\alpha_n)e_n, \end{aligned}$$

откуда, вспоминая определения операций в пространстве K^n, получаем

$$\begin{aligned} (a+b)_S &= (\alpha_1+\beta_1, \alpha_2+\beta_2, \ldots, \alpha_n+\beta_n) = \\ &= (\alpha_1, \alpha_2, \ldots, \alpha_n) + (\beta_1, \beta_2, \ldots, \beta_n) = (a)_S + (b)_S \qquad \text{и} \\ (\lambda a)_S &= (\lambda\alpha_1, \lambda\alpha_2, \ldots, \lambda\alpha_n) = \lambda(\alpha_1, \alpha_2, \ldots \alpha_n) = \lambda(a)_S. \quad \square \end{aligned}$$

Поскольку при изоморфизме линейных пространств линейно зависимая или линейно независимая система векторов переходит в линейно зависимую или линейно независимую систему соответственно, получаем очевидное

Следствие. *Система $a_1, a_2, \dots, a_m$ векторов n-мерного линейного пространства L над полем K линейно независима тогда и только тогда, когда линейно независимой в пространстве K^n является система $(a_1)_S, (a_2)_S, \dots, (a_m)_S$ их координатных строк (в базисе S пространства L).* $\square$

Таким образом, введение координат позволяет, в частности, для распознавания линейной зависимости данной системы векторов произвольного n-мерного линейного пространства воспользоваться соответствующими вычислительными алгоритмами для пространства K^n.

В пространстве K^n, как и в любом другом конечномерном линейном пространстве, можно ввести координаты, фиксируя какой–либо базис этого пространства. При этом, среди всех базисов выделяется стандартный базис пространства K^n. Напомним, что стандартным (или естественным) базисом пространства K^n мы назвали систему S векторов вида

$$\begin{aligned} e_1 &= (1, 0, 0, \dots, 0, 0) \\ e_2 &= (0, 1, 0, \dots, 0, 0) \\ e_3 &= (0, 0, 1, \dots, 0, 0) \\ &\dots\dots\dots\dots\dots\dots\dots \\ e_{n-1} &= (0, 0, 0, \dots, 1, 0) \\ e_n &= (0, 0, 0, \dots, 0, 1). \end{aligned}$$

Если $a = (\alpha_1, \alpha_2, \dots, \alpha_n)$ — произвольный вектор пространства K^n, то очевидное равенство

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n$$

говорит о том, что

$$(a)_S = (\alpha_1, \alpha_2, \dots, \alpha_n) = a.$$

Таким образом, координатная строка произвольного вектора a пространства K^n, вычисленная в стандартном базисе этого пространства, совпадает с самим вектором a.

Вычисление координатной строки вектора зависит от выбора базиса пространства. Посмотрим, как изменится координатная строка вектора при изменении базиса.

Пусть $S = (e_1, e_2, \dots, e_n)$ и $S' = (e'_1, e'_2, \dots, e'_n)$ — два базиса линейного пространства L над полем K. Найдем выражения векторов второго базиса в виде линейной комбинации первого:

$$\begin{aligned} e'_1 &= \tau_{11} e_1 + \tau_{12} e_2 + \dots + \tau_{1n} e_n \\ e'_2 &= \tau_{21} e_1 + \tau_{22} e_2 + \dots + \tau_{2n} e_n \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ e'_n &= \tau_{n1} e_1 + \tau_{n2} e_2 + \dots + \tau_{nn} e_n \end{aligned}$$

Матрицу

$$T = \begin{pmatrix} \tau_{11} & \tau_{12} & \dots & \tau_{1n} \\ \tau_{21} & \tau_{22} & \dots & \tau_{2n} \\ \dots & \dots & \dots & \dots \\ \tau_{n1} & \tau_{n2} & \dots & \tau_{nn} \end{pmatrix},$$

составленную из коэффициентов этих разложений, будем называть матрицей перехода от базиса S к базису S'. Точнее говоря, мы принимаем следующее определение:

Определение 2. *Пусть S и S' — два базиса n-мерного пространства L. Матрицей перехода от системы координат в базисе S к системе координат в базисе S' называется $n \times n$-матрица, строками которой служат координатные строки соответствующих векторов базиса S', вычисленные в базисе S.*

Сохраняя все обозначения, предшествующие определению 2, введем еще обозначения для координатных строк в базисах S и S' вектора $a \in L$:

$$(a)_S = (\alpha_1, \alpha_2, \dots, \alpha_n), \quad (a)_{S'} = (\beta_1, \beta_2, \dots, \beta_n).$$

Это означает, что

$$a = \sum_{i=1}^{n} \alpha_i e_i \quad \text{и} \quad a = \sum_{i=1}^{n} \beta_i e_i'.$$

Поскольку, к тому же, для каждого $i = 1, 2, \dots, n$ $e_i' = \sum_{j=1}^{n} \tau_{ij} e_j$, имеем

$$a = \sum_{i=1}^{n} \beta_i e_i' = \sum_{i=1}^{n} \beta_i \left(\sum_{j=1}^{n} \tau_{ij} e_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_i \tau_{ij} e_j = \sum_{j=1}^{n} \sum_{i=1}^{n} \beta_i \tau_{ij} e_j =$$

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{n} \beta_i \tau_{ij} \right) e_j.$$

В силу однозначности выражения вектора a в виде линейной комбинации системы S, отсюда получаем равенства

$$\alpha_j = \sum_{i=1}^{n} \beta_i \tau_{ij} \quad (j = 1, 2, \dots, n).$$

Эта система равенств означает, что матрица–строка $(\alpha_1\, \alpha_2\, \dots\, \alpha_n)$ равна произведению матрицы–строки $(\beta_1\, \beta_2\, \dots\, \beta_n)$ на матрицу T:

$$(\alpha_1\, \alpha_2\, \dots\, \alpha_n) = (\beta_1\, \beta_2\, \dots\, \beta_n) \cdot T.$$

Таким образом, нами доказано

Предложение 3. *Пусть S и S' — два базиса пространства L и T — матрица перехода от системы координат в базисе S к системе координат в базисе S'. Тогда для любого вектора $a \in L$ выполнено равенство $(a)_S = (a)_{S'} \cdot T$.* $\square$

Этим утверждением и выражается связь между координатными строками вектора в "старом" и "новом" базисах: если эти строки интерпретировать как матрицы, то координатная строка вектора $a \in L$ в старом базисе равна произведению координатной строки этого вектора в новом базисе на матрицу перехода.

Следует отметить, что в ряде учебных пособий по линейной алгебре матрицей перехода от базиса S к базису S' называют матрицу, строки которой являются координатными строками векторов базиса старого S, вычисленными в новом базисе S'. При таком определении в предложении 3 следует поменять местами строки $(a)_S$ и $(a)_{S'}$. Впрочем, мы сейчас поймем, что при этом наша матрица перехода просто заменится на обратную к ней.

Сохраняя попрежнему все предыдущие обозначения, рассмотрим еще матрицу R перехода (в смысле определения 2) от базиса S' к базису S. Тогда по предложению 3 для любого вектора $a \in L$ наряду с равенством $(a)_S = (a)_{S'} \cdot T$ должно выполняться равенство $(a)_{S'} = (a)_S \cdot R$. Отсюда

$$(a)_S = (a)_{S'} \cdot T = ((a)_S \cdot R) \cdot T = (a)_S \cdot (RT)\,.$$

Таким образом, для любого вектора $a \in L$ имеет место равенство $(a)_S \cdot (RT) = (a)_S = (a)_S \cdot E$ (где E — единичная матрица). Так как при этом $(a)_S$ может быть любым элементом пространства K^n, отсюда, как известно, следует равенство $RT = E$, а это означает, что матрица T обратима и $R = T^{-1}$. Мы приходим к следующему утверждению:

Предложение 4. *Пусть S и S' — два базиса пространства L и T — матрица перехода от базиса S к базису S'. Тогда матрица T обратима и обратная к ней матрица T^{-1} служит матрицей перехода от базиса S' к базису S. Для любой обратимой матрицы U (подходящего порядка) и любого базиса S пространства L существует базис S'' этого пространства такой, что U является матрицей перехода от базиса S к базису S''.*

Доказательства требует лишь последнее утверждение. Пусть $U = (\upsilon_{ij})$ и $S = (e_1, e_2, \ldots, e_n)$. Для каждого $i = 1, 2, \ldots, n$ полагаем

$$e''_i = \sum_{j=1}^{n} \upsilon_{ij} e_j.$$

Легко видеть, что полученная система векторов $S'' = (e''_1, e''_2, \ldots, e''_n)$ линейно независима. Действительно, ввиду следствия к предложению 2 для этого достаточно убедиться в линейной независимости системы $(e''_1)_S, (e''_2)_S, \ldots, (e''_n)_S$ координатных строк векторов системы S''. Но система $(e''_1)_S, (e''_2)_S, \ldots, (e''_n)_S$ совпадает с системой строк матрицы U, и ее линейная независимость следует из обратимости этой матрицы.

Так как число векторов системы S'' совпадает с размерностью пространства L, эта система является базисом пространства L. Очевидно, что матрица U является матрицей перехода от базиса Sк базису S''. $\square$

Перейдем теперь к построению матрицы линейного отображения. Пусть L и M — линейные пространства над полем K и φ — линейное отображение пространства L в пространство M. Фиксируем некоторый базис $S = (a_1, a_2, \ldots, a_m)$ пространства L и некоторый базис $S' = (b_1, b_2, \ldots, b_n)$ пространства M. Выразим образ относительно отображения φ каждого вектора a_i из базиса S в виде линейной комбинации системы S':

$$\begin{array}{l} a_1\varphi = \alpha_{11}b_1 \, + \, \alpha_{12}b_2 \, + \cdots + \, \alpha_{1n}b_n \\ a_2\varphi = \alpha_{21}b_1 \, + \, \alpha_{22}b_2 \, + \cdots + \, \alpha_{2n}b_n \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_m\varphi = \alpha_{m1}b_1 + \alpha_{m2}b_2 + \cdots + \alpha_{mn}b_n. \end{array}$$

Это означает, что координатные строки в базисе S' образов векторов из системы S имеют вид

$$(a_i\varphi)_{S'} = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in}) \quad (i = 1, 2, \ldots, m).$$

Тогда матрица

$$A_\varphi = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} \end{pmatrix},$$

составленная из компонент этих координатных строк, будет называться матрицей отображения φ в базисах S и S'. Точнее говоря, мы принимаем следующее

Определение 3. *Матрицей линейного отображения φ m-мерного линейного пространства L в n-мерное линейное пространство M, вычисленной в базисах S и S' пространств L и M соответственно, называется $m \times n$-матрица, строками которой являются координатные строки образов относительно φ соответствующих векторов базиса S, вычисленные в базисе S'.*

Матрицей линейного оператора φ m-мерного линейного пространства L, вычисленной в базисе S этого пространства, называется $m \times m$-матрица, строками которой являются координатные строки образов относительно φ соответствующих векторов базиса S, вычисленные в этом же базисе пространства L.

Рассмотрим несколько примеров.

1. Пусть $L = \mathbb{R}^3$, $M = \mathbb{R}^2$ и отображение $\varphi : L \to M$ переводит вектор $a = (\alpha_1, \alpha_2, \alpha_3)$ из L в вектор $a\varphi = (\alpha_1 + \alpha_2, \alpha_2 + \alpha_3)$. Линейность этого отображения проверяется непосредственно. Вычислим его матрицу в стандартных базисах $S = (e_1, e_2, e_3)$ и $S' = (f_1, f_2)$ данных пространств (где, напомним,

$$e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1) \quad \text{и} \quad f_1 = (1, 0), f_2 = (0, 1)).$$

Непосредственные вычисления дают

$$\begin{aligned} e_1\varphi &= (1,0) = 1\,f_1 + 0\,f_2,\\ e_2\varphi &= (1,1) = 1\,f_1 + 1\,f_2,\\ e_3\varphi &= (0,1) = 0\,f_1 + 1\,f_2. \end{aligned}$$

Поэтому матрицей отображения в этих базисах будет матрица

$$A_\varphi = \begin{pmatrix} 1 & 0\\ 1 & 1\\ 0 & 1 \end{pmatrix}.$$

2. В пространстве V^2 выберем базис, состоящий из двух взаимно перпендикулярных отрезков e_1 и e_2 одинаковой длины, причем при повороте плоскости против часовой стрелки на угол, равный $\dfrac{\pi}{2}$, отрезок e_1 переходит в отрезок e_2, а отрезок e_2 — в отрезок $-e_1$. Поэтому матрица оператора φ, совпадающего с этим поворотом, имеет вид $A_\varphi = \begin{pmatrix} 0 & 1\\ -1 & 0 \end{pmatrix}$. Пусть еще ψ обозначает оператор, переводящий произвольный вектор $a \in V^2$ в его ортогональную проекцию на прямую, определяемую отрезком e_1. Тогда $e_1\psi = e_1$, $e_2\psi = 0$ и потому $A_\psi = \begin{pmatrix} 1 & 0\\ 0 & 0 \end{pmatrix}$.

3. Пусть φ — оператор дифференцирования пространства $L = \mathbb{R}_3[x]$. Тогда в базисе $1, x, x^2, x^3$ этого пространства матрица оператора φ имеет вид

$$A_\varphi = \begin{pmatrix} 0 & 0 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & 3 & 0 \end{pmatrix}.$$

Предложение 5. *Пусть L и M — линейные пространства с базисами S и S' соответственно, $\varphi : L \to M$ — линейное отображение и A_φ — матрица отображения φ в базисах S и S'. Тогда для любого вектора $a \in L$ имеет место равенство*

$$(a\varphi)_{S'} = (a)_S A_\varphi.$$

Доказательство. Введем обозначения для данных в формулировке предложения объектов. Пусть $S = (a_1, a_2, \ldots, a_m)$, $S' = (b_1, b_2, \ldots, b_n)$ и $A_\varphi = (\alpha_{ij})$. Тогда по определению матрицы отображения φ имеем

$$a_i\varphi = \sum_{j=1}^{n} \alpha_{ij} b_j \quad (i = 1, 2, \ldots, m).$$

Пусть a — произвольный вектор пространства L и $a = \sum_{i=1}^{m} \beta_i a_i$ — его выражение в виде линейной комбинации системы S, так что $(a)_S = (\beta_1\, \beta_2\, \ldots\, \beta_m)$. Тогда

$$a\varphi = \sum_{i=1}^{m} \beta_i\,(a_i\varphi) = \sum_{i=1}^{m} \beta_i \left(\sum_{j=1}^{n} \alpha_{ij} b_j \right) = \sum_{i=1}^{m} \sum_{j=1}^{n} (\beta_i \alpha_{ij}) b_j =$$

$$\sum_{j=1}^{n} \sum_{i=1}^{m} (\beta_i \alpha_{ij}) b_j = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} \beta_i \alpha_{ij} \right) b_j.$$

Отсюда следует, что

$$(a\varphi)_{S'} = (\sum_{i=1}^{m} \beta_i \alpha_{i1}, \sum_{i=1}^{m} \beta_i \alpha_{i2}, \ldots, \sum_{i=1}^{m} \beta_i \alpha_{in}) = (\beta_1\, \beta_2\, \ldots\, \beta_m) \cdot A_\varphi = (a)_S A_\varphi,$$

и предложение доказано. $\square$

Предложение 5 указывает на полезность введения матрицы линейного отображения: вычисление образа данного вектора или, что практически то же самое, — его координатной строки сводится к умножению соответствующих матриц. Пусть, например, φ — оператор из примера 3 выше. Пусть $f(x) = x^3 - 4x + 2$ — данный элемент пространства $L = \mathbb{R}_3[x]$. Тогда его координатная строка в рассматриваемом базисе S пространства L имеет вид $(f(x))_S = (2, -4, 0, 1)$. Так как

$$(2\ -4\,0\,1) \cdot \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix} = (-4\,0\,3\,0),$$

координатная строка его образа относительно равна $(-4, 0, 3, 0)$, и потому образом относительно отображения φ многочлена $f(x)$, т.е. его производной, является многочлен $3x^2 - 4$.

Покажем теперь, что действиям с линейными отображениями соответствуют действия с матрицами, подобно тому, как операциям на множестве векторов соответствовали операции с координатными строками векторов. Следующее утверждение является аналогом предложения 2.

Предложение 6. *Пусть L и M — линейные пространства над полем K, имеющие размерность m и n соответственно, S — базис пространства L и S' — базис пространства M. Отображение $\theta : \mathcal{L}(L, M) \to M_{mn}(K)$, сопоставляющее линейному отображению $\varphi : L \to M$ его матрицу A_φ в базисах S и S', является биективным. Кроме того, для любых $\varphi, \psi \in \mathcal{L}(L, M)$ и $\alpha \in K$ выполняются равенства*

$$A_{\varphi+\psi} = A_\varphi + A_\psi \quad \text{и} \quad A_{\alpha\varphi} = \alpha A_\varphi.$$

Иначе говоря, отображение θ является изоморфизмом линейного пространства $\mathcal{L}(L, M)$ на линейное пространство $M_{mn}(K)$.

Доказательство. Пусть $S = (a_1, a_2, \ldots, a_m)$ и $S' = (b_1, b_2, \ldots, b_n)$. Если θ-образы отображений φ и ψ из пространства $\mathcal{L}(L, M)$ совпадают, т.е. равны матрицы A_φ и A_ψ, то равными являются соответствующие строки этих матриц. Следовательно, $(a_i\varphi)_{S'} = (a_i\psi)_{S'}$ для всех $i = 1, 2, \ldots, m$. Из предложения 2 теперь следует, что для каждого i имеет место равенство $a_i\varphi = a_i\psi$, а это в силу предложения 3.11 означает, что $\varphi = \psi$. Мы доказали, что отображение θ инъективно.

Пусть, далее, $B = (\beta_{ij})$ — произвольная матрица из пространства $M_{mn}(K)$. Определим векторы $c_1, c_2, \ldots, c_m$ пространства M, полагая

$$c_i = \sum_{j=1}^{n} \beta_{ij} b_j \quad (i = 1, 2, \ldots, m).$$

Ввиду предложения 3.11 существует линейное отображение $\varphi \in \mathcal{L}(L, M)$ такое, что для любого $i = 1, 2, \ldots, m$ имеет место равенство $a_i\varphi = c_i$. Так как координатная строка $(c_i\varphi)_{S'} = (\beta_{i1}, \beta_{i2}, \ldots, \beta_{in})$ образа относительно φ вектора a_i совпадает с соответствующей строкой матрицы B, эта матрица и является матрицей отображения φ в базисах S и S', т.е. $\varphi\theta = B$. Таким образом, доказана и сюръективность отображения θ.

Из предложения 5 следует, что для любых отображений $\varphi, \psi \in \mathcal{L}(L, M)$ и любого вектора $a \in L$ выполнено равенство

$$(a(\varphi + \psi))_{S'} = (a)_S A_{\varphi+\psi}.$$

Так как, с другой стороны,

$$(a(\varphi + \psi))_{S'} = (a\varphi + a\psi)_{S'} = (a\varphi)_{S'} + (a\psi)_{S'} = \\ (a)_S A_\varphi + (a)_S A_\psi = (a)_S (A_\varphi + A_\psi),$$

мы видим, что для любого вектора $a \in L$, а потому и для любого элемента $(a)_S$ пространства K^m выполняется равенство

$$(a)_S (A_{\varphi+\psi}) = (a)_S (A_\varphi + A_\psi.$$

Поэтому матрицы $A_{\varphi+\psi}$ и $A_\varphi + A_\psi$ совпадают. Равенство $A_{\alpha\varphi} = \alpha A_\varphi$ доказывается аналогично. $\square$

В дополнение к предложению 6 заметим, что если произведение двух линейных отображений определено, то и матрица произведения в подходящих базисах будет произведением матриц сомножителей. Более точно, имеет место

Предложение 7. *Пусть $\varphi : L \to M$ и $\psi : M \to N$ — линейные отображения. Пусть S, S' и S'' — фиксированные базисы линейных пространств L, M и N соответственно, A — матрица отображения φ в базисах S и S' , B — матрица отображения ψ в базисах S' и S'' и C — матрица отображения $\varphi\psi$ в базисах S и S''. Тогда $C = AB$.*

Действительно, для произвольного вектора $a \in L$ в силу предложения 5 имеем

$$(a\varphi)_{S'} = (a)_S A, \quad ((a\varphi)\psi)_{S''} = (a\varphi)_{S'} B \quad \text{и} \quad (a(\varphi\psi))_{S''} = (a)_S C,$$

откуда ввиду равенства $a(\varphi\psi) = a(\varphi)\psi$ получаем

$$(a)_S C = (a\varphi)_{S'} B = ((a)_S A)B = (a)_S(AB).$$

Отсюда, как и прежде, следует, что $C = AB$. $\square$

Напомним, что множество $\mathcal{L}(L)$ всех операторов пространства L является линейной алгеброй так же, как и множество $M_n(K)$ всех квадратных матриц порядка n над полем K. Эти алгебры оказываются изоморфными в смысле следующего естественного определения:

Определение 4. *Пусть A и B — линейные алгебры над полем K. Отображение $\varphi : A \to B$ называется изоморфизмом алгебры A на алгебру B, если φ является изоморфизмом линейного пространства A на линейное пространство B (т.е. φ линейно и биективно) и для любых элементов $a, b \in A$ выполнено равенство $(ab)\varphi = a\varphi\, b\varphi$. Алгебры A и B называются изоморфными, если существует изоморфизм A на B.*

Из предложений 6 и 7 получаем очевидное

Следствие. *Для любого n-мерного линейного пространства L над полем K отображение θ, определенное в предложении 6, является изоморфизмом алгебры $\mathcal{L}(L)$ на алгебру $M_n(K)$.*

В предложениях 6 и 7 установлена связь между операциями на множестве линейных отображений с операциями на множестве матриц. Еще одна связь между линейными отображениями и их матрицами содержится в следующем утверждении:

Предложение 8. *Ранг линейного отображения равен рангу его матрицы.*

Доказательство. Пусть φ — линейное отображение пространства L в пространство M и A_φ — матрица этого отображения в базисах S и S' пространств L и M соответственно. Пусть $S = (a_1, a_2, \ldots, a_m)$. Ранг матрицы A_φ равен рангу системы ее строк, т.е. рангу системы векторов

$$(a_1\varphi)_{S'},\ (a_2\varphi)_{S'}, \ldots,\ (a_m\varphi)_{S'}$$

пространства K^n (где $n = \dim M$). Так как отображение пространства M на пространство K^n, сопоставляющее произвольному вектору его координатную строку, является изоморфизмом и потому сохраняет ранг произвольной системы векторов, ранг матрицы A_φ равен рангу системы векторов

$$a_1\varphi,\ a_2\varphi, \ldots,\ a_m\varphi$$

пространства M. А так как эта последняя является системой порождающих подпространства $\operatorname{Im}\varphi$ пространства M, то ее ранг совпадает с размерностью этого подпространства, т.е. — с рангом отображения φ. $\square$

Выше во всех формулировках, где речь шла о матрице линейного отображения, фиксировались базисы линейных пространств, участвующие в вычислении матрицы. Тем самым, в частности, подчеркивалась очевидная зависимость матрицы линейного отображения от выбора базисов. Настало время выяснить характер этой зависимости.

Предложение 9. *Пусть φ — линейное отображение пространства L_1 в пространство L_2, S_1 и S'_1 — два базиса пространства L_1 и T_1 — матрица перехода от базиса S_1 к базису S'_1, S_2 и S'_2 — два базиса пространства L_2 и T_2 — матрица перехода от базиса S_2 к базису S'_2. Пусть, наконец, A_φ — матрица отображения φ в базисах S_1 и S_2 и A'_φ — матрица отображения φ в базисах S'_1 и S'_2. Тогда $A'_\varphi = T_1 A_\varphi T_2^{-1}$.*

Доказательство. По предложению 5 для произвольного вектора $a \in L_1$ имеем

$$(a\varphi)_{S_2} = (a)_{S_1} A_\varphi \quad \text{и} \quad (a\varphi)_{S'_2} = (a)_{S'_1} A'_\varphi.$$

Кроме того, из предложения 4.3 следует, что

$$(a)_{S_1} = (a)_{S'_1} T_1 \quad \text{и} \quad (a\varphi)_{S_2} = (a\varphi)_{S'_2} T_2.$$

Отсюда

$$(a)_{S'_1} A'_\varphi T_2 = (a\varphi)_{S'_2} T_2 = (a\varphi)_{S_2} = (a)_{S_1} A_\varphi = (a)_{S'_1} T_1 A_\varphi.$$

Таким образом, для любой матрицы – строки $(a)_{S'_1}$ имеет место равенство

$$(a)_{S'_1} \left(A'_\varphi T_2\right) = (a)_{S'_1} \left(T_1 A_\varphi\right).$$

Поэтому $A'_\varphi T_2 = T_1 A_\varphi$, откуда после умножения обеих частей справа на матрицу T_2^{-1} и получаем $A'_\varphi = T_1 A_\varphi T_2^{-1}$. $\square$

В частности, для операторов пространства L имеем

Следствие. *Пусть φ — линейный оператор пространства L, S и S' — два базиса этого пространства и T — матрица перехода от базиса S к базису S'. Пусть A_φ — матрица оператора φ в базисе S и A'_φ — матрица оператора φ в базисе S'. Тогда $A'_\varphi = T A_\varphi T^{-1}$.* $\square$

Последнее утверждение делает естественным введение следующего понятия:

Определение 5. *Матрица $A \in M_n(K)$ называется сопряженной с матрицей $B \in M_n(K)$, если существует обратимая матрица $U \in M_n(K)$ такая, что $A = UBU^{-1}$.*

Пусть матрица A сопряжена с матрицей B, т.е. $A = UBU^{-1}$. Тогда $B = U^{-1}AU = U^{-1}A(U^{-1})^{-1}$ и, значит, матрица B сопряжена с матрицей A. Пусть еще матрица B сопряжена с матрицей C, $B = VCV^{-1}$. Тогда $A = UBU^{-1} = A = U(VCV^{-1})U^{-1} = (UV)C(UV)^{-1}$, и матрица сопряжена с матрицей . Наконец, так как каждая матрица, очевидно, сопряжена с самой собой, нами доказано следующее

Предложение 10. *Отношение сопряженности матриц является отношением эквивалентности.* $\square$

Таким образом, если базис пространства L не предполагать фиксированным, то каждому оператору этого пространства соответствует класс сопряженных матриц. Естественно возникает задача нахождения такого базиса пространства, матрица данного оператора в котором имеет простейший и в некотором смысле канонический вид. Для операторов линейных пространств эта задача будет подробно обсуждаться в следующей главе. Для отображений же одного пространства в другое она легко решается по той причине, что базисы в этих пространствах можно выбирать произвольным образом независимо один от другог.

Предложение 11. *Пусть φ — линейное отображение пространства L в пространство M. Существуют базисы S и S' пространств L и M соответственно, матрица отображения φ в которых имеет вид*

$$\begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix},$$

где число единиц на главной диагонали совпадает с рангом отображения φ.

Доказательство. Как и в доказательстве теоремы 3.6 выберем некоторый базис $b_1, b_2, \dots, b_r$ подпространства $\operatorname{Im}\varphi$ пространства M и векторы $a_1, a_2, \dots, a_r$ пространства L такие, что для всех $i = 1, 2 \dots r \quad a_i\varphi = b_i$. Если векторы $a_{r+1}, a_{r+2}, \dots, a_m$ составляют базис подпространства $\operatorname{Ker}\varphi$ пространства L, то из доказательства теоремы 3.6 и из теоремы 3.5 вытекает, что система $S = (a_1, a_2, \dots, a_m)$ является базисом пространства L. Наконец, дополним систему $b_1, b_2, \dots, b_r$ до базиса S' пространства M. Очевидно, что матрица отображения φ в построенных базисах S и S' имеет требуемый вид. $\square$

В заключение этого параграфа покажем, что введение координат в конечномерном пространстве позволяет легко решить задачу классификации линейных пространств, т. е. нахождения условий, при которых два линейных пространства являются изоморфными.

Предложение 12. *Если два линейных пространства изоморфны и одно из них конечномерно, то и другое является конечномерным. Конечномерные линейные пространства над полем K изоморфны тогда и только тогда, когда их размерности совпадают.*

Доказательство. Пусть L и M — изоморфные линейные пространства над полем K и $\varphi : L \to M$ — изоморфизм пространства L на пространство M. Если пространство L конечномерно, то по предложению 3.8 подпространство $\operatorname{Im}\varphi$ пространства M конечномерно и $\dim(\operatorname{Im}\varphi) \leq \dim L$. Так как отображение φ сюръективно, то $\operatorname{Im}\varphi = M$ и потому пространство M конечномерно и $\dim M \leq \dim L$. Рассматривая вместо отображения φ отображение $\varphi^{-1} : M \to L$, приходим к неравенству $\dim L \leq \dim M$, так что $\dim L = \dim M$. Обратно, если L и M — n-мерные пространства, то каждое из них в силу предложения 2 изоморфно пространству K^n, и потому они изоморфны между собой. $\square$

§ 5. ЛИНЕЙНЫЕ ФУНКЦИОНАЛЫ

В этом параграфе мы рассмотрим важный частный случай общего понятия линейного отображения.

Определение 1. *Пусть L — линейное пространство над полем K. Линейное отображение пространства L в поле K, рассматриваемое как линейное пространство над самим собой, называется линейным функционалом на пространстве L. Линейное пространство $\mathcal{L}(L, K)$ всех функционалов на L будем называть двойственным (или сопряженным) к пространству L и обозначать символом L^*.*

Для линейных функционалов мы будем использовать обычную функциональную запись, в отличие от записи остальных линейных отображений как правых операторов. Таким образом, элементы множества L^* можно определить как функции $f(x)$ от одной переменной, определенные на множестве L, принимающие значения в поле K и такие, что для любых векторов $a, b \in L$ и скаляра $\lambda \in K$ выполнены равенства

$$f(a+b) = f(a) + f(b) \quad \text{и} \quad f(\lambda a) = \lambda f(a)$$

(определение линейности в функциональной записи). Заметим также, что в соответствии с общим определением суммы линейных отображений и умножения скаляра на линейное отображение для любых функционалов $f, g \in L^*$ и скаляра $\lambda \in K$ функционалы $f+g$ и λf на пространстве L определяются равенствами

$$(f+g)(a) = f(a) + g(a) \quad \text{и} \quad (\lambda f)(a) = \lambda f(a) \quad (a \in L).$$

Приведем несколько примеров.

1. Проектирование π_i пространства K^n на i-ую компоненту является, очевидно, линейным функционалом на этом пространстве. Другой пример функционала на том же пространстве получим, сопоставляя каждому вектору из K^n сумму всех его компонент.

2. Пусть $L = K[x]$ и λ — фиксированный элемент поля K. Сопоставляя каждому многочлену $u(x) \in L$ его значение при $x = \lambda$, получаем линейный функционал на L.

3. След $\operatorname{tr} a$ матрицы $a = (\alpha_{ij})$ из пространства $M_n(K)$ есть, по определению, сумма ее диагональных элементов: $\operatorname{tr} a = \sum_{i=1}^{n} \alpha_{ii}$. Очевидно, что $\operatorname{tr} a$ есть линейный функционал на пространстве $M_n(K)$.

Из предложения 3.11 следует универсальный способ задания функционала на кнечномерном линейном пространстве:

Предложение 1. *Пусть $a_1, a_2, \ldots, a_n$ — базис пространства L над полем K. Для любой системы $\lambda_1, \lambda_2, \ldots, \lambda_n$ скаляров из поля K существует единственный функционал $f(x)$ на пространстве L такой, что для каждого $i = 1, 2, \ldots, n$ выполнено равенство $f(a_i) = \lambda_i$. Более того, произвольный функционал на пространстве L может быть задан указанным способом.* $\square$

Говоря более содержательно, предложение 1 утверждает, что произвольная функция, определенная на элементах некоторого базиса пространства L и принимающая значения в поле K, может быть однозначно доопределена (или, как говорят, продолжена по линейности) до некоторого функционала на этом пространстве. Это приводит к следующему утверждению:

Предложение 2. *Пусть L — конечномерное линейное пространство, M — подпространство пространства L и a — вектор пространства L, не входящий в подпространство M. Тогда существует функционал $f \in L^*$ такой, что $f(a) = 1$ и для любого $c \in M$ $f(c) = 0$.*

В частности, для любого ненулевого вектора $a \in L$ существует функционал $f \in L^*$ такой, что $f(a) \neq 0$. Иначе говоря, если вектор $a \in L$ таков, что для любого $f \in L^*$ выполнено равенство $f(a) = 0$, то $a = 0$.

Доказательство. Пусть система векторов $a_1, a_2, \dots, a_m$ является базисом подпространства M. Так как вектор a не принадлежит подпространству M, система векторов $a, a_1, a_2, \dots, a_m$ линейно независима (докажите это). Поэтому ее можно дополнить до базиса $a, a_1, a_2, \dots, a_m, a_{m+1}, a_{m+2}, \dots, a_{n-1}$ всего пространства L ($n = \dim L$). В силу предложения 1 существует функционал $f \in L^*$ такой, что $f(a) = 1$ и для любого $i = 1, 2, \dots, n-1$ $f(a_i) = 0$. Если $c \in M$, то $c = \sum_{i=1}^m \alpha_i a_i$ для подходящих $\alpha_1, \alpha_2, \dots \alpha_m \in K$ и потому

$$f(c) = f(\sum_{i=1}^m \alpha_i a_i) = \sum_{i=1}^m \alpha_i f(a_i) = 0.$$

Таким образом, функционал f искомый. Последнее утверждение предложения 1 сводится к уже доказанному, если в качестве M взять нулевое подпространство. □

Предложение 3. *Пусть $S = (e_1, e_2, \dots, e_n)$ — базис линейного пространства L. Для каждого $i = 1, 2, \dots, n$ определим функционал $e_i^* \in L^*$, полагая*

$$e_i^*(e_j) = \begin{cases} 1, & \text{если } j = i \\ 0, & \text{если } j \neq i. \end{cases}$$

Тогда система $S^* = (e_1^*, e_2^*, \dots, e_n^*)$ элементов пространства L^* является базисом этого пространства, причем координатная строка в базисе S^* любого функционала $f \in L^*$ имеет вид

$$(f)_{S^*} = (f(e_1), f(e_2), \dots, f(e_n)).$$

Доказательство. Прежде всего заметим, что в силу предложения 1 функционалы $e_1^*, e_2^*, \dots, e_n^*$, определенные в предложении 3, действительно существуют и определены однозначно: функционал e_i^* мы определяем, задавая его значения от векторов $e_1, e_2, \dots, e_n$ базиса S пространства L. Покажем, что система S^* является базисом пространства L^*.

Пусть для некоторых скаляров $\alpha_1, \alpha_2, \dots \alpha_n$ имеет место равенство

$$\alpha_1 e_1^* + \alpha_2 e_2^* + \dots + \alpha_n e_n^* = 0$$

(в правой части которого символ 0 обозначает нулевой функционал, т. е. — нулевое линейное отображение, значение которого от любого вектора пространства L равно 0). Вычисляя значение правой и левой частей этого равенства от произвольного вектора e_j базиса S, получаем в соответствии с определением суммы функционалов и произведения скаляра на функционал:

$$0 = (\alpha_1 e_1^* + \alpha_2 e_2^* + \cdots + \alpha_n e_n^*)(e_j) = \alpha_1 e_1^*(e_j) + \alpha_2 e_2^*(e_j) + \cdots + \alpha_n e_n^*(e_j)$$
$$= \alpha_j e_j^*(e_j) = \alpha_j.$$

Следовательно, $\alpha_j = 0$ для всех $j = 1, 2, \ldots, n$, и потому система S^* линейно независима.

Пусть, теперь, f — произвольный элемент пространства L^* и пусть $f(e_i) = \beta_i$ $(i = 1, 2, \ldots, n)$. Полагая

$$g = \beta_1 e_1^* + \beta_2 e_2^* + \cdots + \beta_n e_n^*,$$

для произвольного $j = 1, 2, \ldots, n$ имеем, очевидно,

$$g(e_j) = \beta_1 e_1^*(e_j) + \beta_2 e_2^*(e_j) + \cdots + \beta_n e_n^*(e_j) = \beta_j e_j^*(e_j) = \beta_j = f(e_j),$$

и потому в силу предложения 1 $f = g$. Следовательно, функционал f линейно выражается через систему S^*, так что эта система порождает пространство L^* и утверждение о том, что S^* — базис пространства L^*, доказано. Более того, доказанное только что равенство

$$f = \beta_1 e_1^* + \beta_2 e_2^* + \cdots + \beta_n e_n^*,$$

означает, что

$$(f)_{S^*} = (\beta_1, \beta_2, \ldots, \beta_n) = (f(e_1), f(e_2), \ldots, f(e_n)),$$

и все утверждения предложения доказаны. □

Договоримся о терминологии:

Определение 2. *Базис S^* пространства L^*, построенный в предложении 2, называется двойственным (или сопряженным) к базису S пространства L.*

Предложение 3 говорит, частности, о том, что пространство L^*, двойственное к конечномерному пространству L, имеет ту же размерность, что и L. Поэтому из предложения 4.12 вытекает

Следствие. *Произвольное конечномерное пространство L изоморфно двойственному к нему пространству L^*.* □

Для линейных пространств, не являющихся конечномерными, последнее утверждение оказывается неверным. Если же пространство L конечномерно, то из него следует, что пространство L^*, в свою очередь, изоморфно двойственному к нему пространству, которое естественно обозначать L^{**} и называть дважды двойственным к пространству L. Разумеется, в этом случае пространства L и L^{**} также оказываются изоморфными. Существование изоморфизма между пространствами L и L^*, L^* и L^{**} и т.д. обеспечивается просто совпадением их размерностей, а не природой их элементов. В отличие от этого, сейчас будет построен изоморфизм конечномерного пространства L на дважды двойственное к нему пространство L^{**}, основанный на более глубокой связи между ними, чем простое совпадение размерностей. На это указывает хотя бы то обстоятельство, что в построении этого изоморфизма не будут участвовать базисы пространств L и L^{**}. Подчеркивая это, мы будем называть его каноническим изоморфизмом.

Теорема. *Пусть L — конечномерное линейное пространство над полем K. Для каждого вектора $a \in L$ определим отображение g_a пространства L^* в поле K, полагая для любого $f \in L^*$ $g_a(f) = f(a)$. Это отображение является линейным функционалом на пространстве L^*, т.е. $g_a \in L^{**}$. Отображение $\varphi : L \to L^{**}$, сопоставляющее вектору $a \in L$ функционал g_a, является изоморфизмом пространства L на пространство L^{**}.*

Доказательство. Для произвольных функционалов f_1, f_2 и скаляров $\alpha, \beta \in K$ имеем

$$g_a(\alpha f_1 + \beta f_2) = (\alpha f_1 + \beta f_2)(a) = \alpha f_1(a) + \beta f_2(a) = \alpha g_a(f_1) + \beta g_a(f_2),$$

и отображение g_a действительно является линейным функционалом на L^*. Если еще a и b — произвольные векторы из L и $f \in L^*$,

$$g_{\alpha a + \beta b}(f) = f(\alpha a + \beta b) = \alpha f(a) + \beta f(b) = \alpha g_a(f) + \beta g_b(f) = (\alpha g_a + \beta g_b)(f),$$

и потому

$$(\alpha a + \beta b)\varphi = g_{\alpha a + \beta b} = \alpha g_a + \beta g_b = \alpha(a\varphi) + \beta(b\varphi),$$

так что и отображение φ является линейным. Для доказательства его инъективности предположим, что вектор a содержится в его ядре. Это означает, что соответствующий вектору a функционал g_a является нулевым, т.е. для любого $f \in L^*$ $g_a(f) = 0$. По определению отображения g_a это означает, что для любого $f \in L^*$ $f(a) = 0$, и из предложения 1 теперь следует, что $a = 0$. Таким образом, ядро отображения φ является нулевым подпространством, и инъективность отображения φ доказана. Так как $\dim L = \dim L^{**}$, из следствия 1 к теореме 3.6 вытекает теперь сюръективность отображения φ. $\square$

Существование канонического изоморфизма пространства L на пространство L^{**} позволяет в определенном смысле отождествлять эти пространства и считать пространство L двойственным к пространству L^*. На этом основан так называемый принцип двойственности: если в некотором верном утверждении о пространствах L и L^* поменять их местами (формально заменив векторы на функционалы, а функционалы на векторы), то полученное утверждение, двойственное исходному, автоматически оказывается также справедливым. Продемонстрируем обоснованность этого принципа на примере доказательства утверждения, двойственного предложению 2.

Предложение 2^*. *Пусть* L *— конечномерное линейное пространство,* M *— подпространство пространства* L^* *и* f *— элемент пространства* L^**, не входящий в подпространство* M*. Тогда существует вектор* $a \in L$ *такой, что* $f(a) = 1$ *и для любого* $h \in M$ $h(a) = 0$*.*

Действительно, применяя предложение 2 к пространству L^*, приходим к существованию такого функционала g на этом пространстве, что $g(f) = 1$ и для любого $h \in M$ $g(h) = 0$. Пусть при каноническом изоморфизме пространства L на пространство L^{**} вектор $a \in L$ является прообразом этого элемента $g \in L^{**}$, т. е. $g = g_a$. Вектор является искомым, так как $f(a) = g_a(f) = g(f) = 1$ и $h(a) = g_a(h) = g(h) = 0$ для любого $h \in M$. $\square$

ГЛАВА II

ЛИНЕЙНЫЕ ОПЕРАТОРЫ ЛИНЕЙНЫХ ПРОСТРАНСТВ

§ 1. ИНВАРИАНТНЫЕ ПОДПРОСТРАНСТВА

Начнем с определения одного из основных понятий этой главы.

Определение 1. *Пусть φ — линейный оператор линейного пространства L над полем K. Подпространство M пространства L называется инвариантным относительно φ (или, короче, φ-инвариантным), если $M\varphi \subseteq M$ (т.е. для любого вектора $a \in L$ из того, что $a \in M$, следует, что и $a\varphi \in M$).*

Непосредственно из определения следует, очевидно, что в любом линейном пространстве L нулевое подпространство так же, как и все пространство L, являются инвариантными относительно каждого оператора $\varphi \in \mathcal{L}(L)$. Рассмотрим еще несколько примеров инвариантных подпространств.

1. Поскольку гомотетия χ_λ линейного пространства L переводит произвольный вектор $a \in L$ в вектор λa, а любое подпространство замкнуто относительно умножений на скаляры, то произвольное подпространство M пространства L является χ_λ-инвариантным.

2. Если φ — оператор дифференцирования пространства $K[x]$, то для любого $n \geq 0$ подпространство $K_n[x]$ φ-инвариантно.

3. Пусть линейное пространство L является прямой суммой своих подпространств M и N, $L = M \oplus N$, и φ — оператор проектирования на M параллельно N. Тогда подпространства M и N φ-инвариантны.

4. Пусть в пространстве V^2 φ — оператор поворота на угол $\pi/2$ против часовой стрелки. Тогда φ-инвариантными подпространствами будут лишь нулевое пространство и все пространство V^2.

Перечислим, далее, ряд простых свойств инвариантных подпространств.

Предложение 1. *Пусть φ — оператор линейного пространства L и M — φ-инвариантное подпространство этого пространства. Для любого перестановочного с φ оператора ψ пространства L образ и прообраз относительно ψ подпространства M являются φ-инвариантными подпространствами.*

В частности, для любого многочлена $f(x) \in K[x]$ образ и прообраз подпространства M относительно оператора $f(\varphi)$ являются φ-инвариантными подпространствами пространства L.

Доказательство. Пусть $\psi \in \mathcal{L}(L)$ — такой оператор пространства L, что $\varphi\,\psi = \psi\,\varphi$, и пусть вектор $a \in L$ принадлежит образу $M\psi$ подпространства M относительно оператора ψ. Это означает, что $a = b\psi$ для подходящего вектора $b \in M$. Тогда ввиду φ-инвариантности подпространства M имеем $b\varphi \in M$, и так как

$$a\varphi = (b\psi)\varphi = b(\psi\varphi) = b(\varphi\psi) = (b\varphi)\psi,$$

мы видим, что $a\varphi \in M\psi$. Значит подпространство $M\psi$ является φ-инвариантным.

Предположим теперь, что вектор a лежит в прообразе $M\psi^{-1}$ подпространства M относительно оператора ψ. Тогда $a\psi \in M$, и так как

$$(a\varphi)\psi = a(\varphi\psi)) = a(\psi\varphi) = (a\psi)\varphi \in M,$$

вектор $a\varphi$ лежит в подпространстве $M\psi^{-1}$. Таким образом, φ-инвариантность и этого подпространства доказана. Для доказательства последнего утверждения предложения 1 достаточно вспомнить (см. следствие к предложению I.2.4), что значения любых двух многочленов от одного и того же оператора пространства L перестановочны. $\square$

Поскольку образ произвольного оператора пространства L совпадает по определению с образом этого пространства, а ядро — с прообразом нулевого подпространства и подпространства L и $\{0\}$ φ-инвариантны для любого оператора $\varphi \in \mathcal{L}(L)$, получаем очевидное

Следствие. *Для любого перестановочного с φ оператора ψ пространства L подпространства $Im\,\psi$ и $Ker\,\psi$ пространства L являются инвариантными относительно φ. В частности, для любого оператора $\varphi \in \mathcal{L}(L)$ и любого многочлена $f(x) \in K[x]$ подпространства $Im\,f(\varphi)$ и $Ker\,f(\varphi)$ пространства L φ-инвариантны.*

Справедливость следующего утверждения проверяется непосредственно.

Предложение 2. *Пусть L — линейное пространство над полем K и M — подпространство пространства L. Если M является инвариантным относительно каждого из двух операторов φ и ψ пространства L, то M инвариантно и относительно их суммы $\varphi + \psi$ и произведения $\varphi\psi$. Кроме того, для любого скаляра $\lambda \in K$ подпространство M инвариантно относительно оператора $\lambda\varphi$. Поэтому всякое φ-инвариантное подпространство является $f(\varphi)$-инвариантным для любого многочлена $f(x) \in K[x]$.*

Если подпространство M линейного пространства L инвариантно относительно оператора φ этого пространства, то ограничение $\overline{\varphi} = \varphi|_M$ отображения φ на подмножество M является отображением этого подмножества в себя. Из предложения 2 следует, что в этом случае для любого $\lambda \in K$ и $\overline{\lambda\varphi}$ является отображением множества M в себя, а если M инвариантно еще и относительно оператора ψ, то таковыми же являются отображения $\overline{\varphi + \psi}$ и $\overline{\varphi\psi}$.

Предложение 3. *Пусть L — линейное пространство над полем K, φ — оператор этого пространства и M — φ-инвариантное подпространство пространства L. Пусть $\overline{\varphi} = \varphi|_M$ — ограничение отображения φ на подмножество M. Тогда $\overline{\varphi}$ является линейным оператором пространства M. Кроме того, для любого скаляра $\lambda \in K$ $\overline{\lambda\varphi} = \lambda\overline{\varphi}$.*

Если подпространство M инвариантно и относительно другого оператора ψ пространства L, то $\overline{\varphi + \psi} = \overline{\varphi} + \overline{\psi}$ и $\overline{\varphi\psi} = \overline{\varphi}\overline{\psi}$. Поэтому для любого многочлена $f(x) \in K[x]$ $\overline{f(\varphi)} = f(\overline{\varphi})$.

Доказательство. Поскольку по определению ограничения отображения операторы $\overline{\varphi}$ и φ одинаково действуют на произвольный вектор из подпространства M, для любых векторов $a, b \in M$ и скаляров $\alpha, \beta \in K$ с учетом того, что $\alpha a + \beta b \in M$, имеем

$$(\alpha a + \beta b)\overline{\varphi} = (\alpha a + \beta b)\varphi = \alpha(a\varphi) + \beta(b\varphi) = \alpha(a\overline{\varphi}) + \beta(b\overline{\varphi}),$$

чем линейность отображения $\overline{\varphi}$ доказана. Далее, очевидно имеем для произвольного вектора $a \in M$:

$$\begin{gathered}
a\overline{\lambda\varphi} = a(\lambda\varphi) = \lambda(a\varphi) = \lambda(a\overline{\varphi}) = a(\lambda\overline{\varphi}),\\
a\overline{\varphi + \psi} = a(\varphi + \psi) = a\varphi + a\psi = a\overline{\varphi} + a\overline{\psi} = a(\overline{\varphi} + \overline{\psi}),\\
a\overline{\varphi\psi} = a(\varphi\psi) = (a\varphi)\psi = (a\overline{\varphi})\overline{\psi} = a(\overline{\varphi}\overline{\psi}).
\end{gathered}$$

Последнее утверждение предложения 3 является очевидным следствием доказанных соотношений. $\square$

Покажем теперь, что существование в пространстве L ненулевого подпространства, отличного от пространства L и инвариантного относительно оператора φ пространства L позволяет подходящим выбором базиса упростить вид матрицы оператора φ.

Пусть φ — оператор n-мерного пространства L и M — m-мерное φ-инвариантное подпространство пространства L, где $1 \le m < n$. Выберем базис $S = (a_1, a_2, \ldots, a_n)$ пространства L так, чтобы подсистема $S_1 = (a_1, a_2, \ldots, a_m)$ первых m векторов системы S составляла базис подпространства M (это можно сделать, выбрав произвольный базис в M и дополнив его до базиса всего пространства L). Для вычисления матрицы оператора φ в этом базисе следует выразить каждый из векторов $a_1\varphi, a_2\varphi, \ldots, a_n\varphi$ в виде линейной комбинации системы S. При этом следует учесть, что поскольку векторы $a_1, a_2, \ldots, a_m$ принадлежат подпространству M и это подпространство инвариантно относительно φ, векторы $a_1\varphi, a_2\varphi, \ldots, a_m\varphi$ являются линейными комбинациями системы S_1. Поэтому в линейные выражения этих векторов через систему S векторы $a_{m+1}, a_{m+2}, \ldots, a_n$ входят с нулевыми коэффициентами, и следовательно искомые выражения имеют вид

$$\begin{array}{llllllll}
a_1\,\varphi & = \alpha_{11}\,a_1 & + \ \alpha_{12}\,a_2 & + \ \cdots + \alpha_{1m}\,a_m & & & \\
a_2\,\varphi & = \alpha_{21}\,a_1 & + \ \alpha_{22}\,a_2 & + \ \cdots + \alpha_{2m}\,a_m & & & \\
\multicolumn{4}{l}{\dotfill} & & & \\
a_m\,\varphi & = \alpha_{m1}\,a_1 & + \ \alpha_{m2}\,a_2 & + \ \cdots + \alpha_{mm}\,a_m & & & \\
a_{m+1}\varphi & = \alpha_{m+1,1}a_1 & + \alpha_{m+1,2}a_2 & + \cdots + \alpha_{m+1,m}a_m & + \alpha_{m+1,m+1}a_{m+1} & + \cdots + \alpha_{m+1,n}a_n \\
\multicolumn{6}{l}{\dotfill} \\
a_n\,\varphi & = \alpha_{n1}\,a_1 & + \alpha_{n2}\,a_2 & + \cdots + \alpha_{nm}\,a_m & + \ \alpha_{n,m+1}\,a_{m+1} & + \cdots + \alpha_{nn}\,a_n
\end{array}$$

Следовательно, матрица A оператора φ в базисе S имеет вид

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2m} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} & 0 & \dots & 0 \\ a_{m+1\,1} & a_{m+1\,2} & \dots & a_{m+1\,m} & a_{m+1\,m+1} & \dots & a_{m+1\,n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} & a_{nm+1} & \dots & a_{nn} \end{pmatrix}.$$

Иначе говоря, матрица A является матрицей блочного вида $A = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$, где в левом верхнем углу расположена квадратная матрица порядка m

$$B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix},$$

в левом нижнем углу расположена $(n-m)\times m$-матрица

$$C = \begin{pmatrix} a_{m+1,1} & a_{m+1,2} & \dots & a_{n-m,m} \\ a_{m+2,1} & a_{m+2,2} & \dots & a_{m+2,m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix},$$

в правом нижнем углу — $(n-m)\times(n-m)$-матрица

$$D = \begin{pmatrix} a_{m+1,m+1} & a_{m+1,m+2} & \dots & a_{m+1,n} \\ a_{m+2,m+1} & a_{m+2,m+2} & \dots & a_{m+2,n} \\ \dots & \dots & \dots & \dots \\ a_{n,m+1} & a_{n,m+2} & \dots & a_{n,n} \end{pmatrix}$$

и в правом верхнем углу — $m\times(n-m)$-матрица, все элементы которой равны нулю. Матрицу A такого вида называют полураспавшейся. Очевидно, кроме того, что матрица B совпадает с матрицей в базисе S_1 ограничения $\overline{\varphi} = \varphi|_M$ оператора φ на подпространство M. Таким образом, имеет место

Предложение 4. *Пусть φ — оператор n-мерного пространства L и M — m-мерное φ-инвариантное подпространство, где $1 \leq m < n$. Пусть S — такой базис пространства L, что первые m его векторов составляют базис S_1 пространства M. Тогда матрица оператора φ в базисе S имеет полураспавшийся вид $A = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$, где B — матрица оператора $\overline{\varphi} = \varphi|_M$ в базисе S_1.* $\square$

Предположим, далее, что (в предыдущих обозначениях) подпространство N, порожденное остальными векторами $a_{m+1}, a_{m+2}, \dots, a_n$ базиса S, тоже является инвариантным относительно оператора φ. Ясно, что тогда и матрица C является нулевой, и потому матрица A оператора φ в базисе S принимает распавшийся или блочно-диагональный вид $A = \begin{pmatrix} B & 0 \\ 0 & D \end{pmatrix}$. Матрица D в этом случае является, разумеется, матрицей ограничения оператора φ на подпространство N. Таким образом, справедливо

Предложение 5. *Пусть φ — оператор конечномерного пространства L и пространство L является прямой суммой φ-инвариантных подпространств L_1 и L_2. Пусть базис S пространства L получен соединением произвольных базисов S_1 и S_2 подпространств L_1 и L_2 соответственно. Тогда матрица A оператора φ в базисе S имеет блочно-диагональный вид $A = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}$, где B_i — матрица ограничения $\varphi_i = \varphi|_{L_i}$ оператора φ на подпространство L_i ($i = 1, 2$).* $\square$

Последнее утверждение делает естественным введение следующего понятия:

Определение 2. *Пусть φ — оператор конечномерного пространства L и пусть пространство L является прямой суммой φ-инвариантных подпространств L_1, L_2, $\ldots$, L_n. Пусть $\varphi_i = \varphi|_{L_i}$ — ограничение оператора φ на подпространство L_i ($i = 1, 2, \ldots, n$). Тогда оператор φ будем называть прямой суммой операторов φ_1, φ_2, $\ldots$, φ_n и записывать $\varphi = \varphi_1 \oplus \varphi_2 \oplus \cdots \oplus \varphi_n$.*

Например, пусть пространство L является прямой суммой подпространств M и N и φ — оператор проектирования на M параллельно N. Тогда для любого вектора $a \in M$ имеем $a\varphi = a$, и потому подпространство M φ-инвариантно и ограничение оператора φ на M совпадает с тождественным оператором ι_M пространства M. Так как для вектора $a \in N$ имеем $a\varphi = 0$, подпространство N также φ-инвариантно и ограничение φ на N совпадает с нулевым оператором θ_N пространства N. Таким образом, $\varphi = \iota_M \oplus \theta_N$.

Из предложения 5 очевидной индукцией получаем

Предложение 6. *Пусть оператор φ конечномерного пространства L является прямой суммой операторов φ_1, φ_2, $\ldots$, φ_n. Тогда в подходящем базисе S пространства L матрица A оператора φ имеет блочно-диагональный вид*

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_n \end{pmatrix},$$

где A_i — матрица оператора φ_i.

Более подробно, пусть L_1, L_2, $\ldots$, L_n — φ-инвариантные подпространства пространства L и $L = \bigoplus_{i=1}^{n} L_i$. Пусть S_i — базис пространства L_i и A_i — матрица оператора $\varphi_i = \varphi|_{L_i}$ в этом базисе ($i = 1, 2, \ldots, n$). Если базис S пространства L является соединением базисов S_1, S_2, $\ldots$, S_n , то матрица A оператора φ в базисе S имеет указанный вид. $\square$

Очевидно, что имеет место и обратное: если матрица A оператора φ линейного пространства L в некотором базисе S этого пространства имеет блочно-диагональный вид из формулировки предложения 6, то система векторов S разбивается на такие части $S_1, S_2, \dots, S_n$ (S_i состоит из векторов, соответствующих тем строкам матрицы A, в которых располагается блок A_i), что если $L_i = l(S_i)$, то $\varphi = \varphi_1 \oplus \varphi_2 \oplus \dots \oplus \varphi_n$, где $\varphi_i = \varphi|_{L_i}$.

Возможность разложения оператора в прямую сумму позволяет сводить решение различных вопросов об операторе к решению их для его прямых слагаемых. Здесь может быть полезным следующее утверждение, непосредственно вытекающее из наших определений и предложения 3:

Предложение 7. *Пусть оператор φ линейного пространства L над полем K является прямой суммой операторов $\varphi_1, \varphi_2, \dots, \varphi_n$. Тогда для любого многочлена $f(x) \in K[x]$ оператор $f(\varphi)$ равен прямой сумме операторов $f(\varphi_1), f(\varphi_2), \dots, f(\varphi_n)$.* $\square$

Следует отметить, что прямая сумма операторов не является частным случаем сложения операторов, определенного в параграфе I.2, в отличие от ситуации с подпространствами линейных пространств, где прямая сумма подпространств является, в частности, суммой этих подпространств. Действительно, операция сложения линейных отображений была определена для отображений с одной и той же областью определения и, к тому же, для любых таких отображений, а области определения прямых слагаемых оператора не имеют общих ненулевых элементов. Кроме того, мы определили пока лишь разложимость данного оператора в прямую сумму, что не дает возможности говорить о прямой сумме наперед заданных операторов. Этот пробел можно, впрочем, легко устранить:

Предложение 8. *Пусть линейное пространство L является прямой суммой своих подпространств $L_1, L_2, \dots, L_n$ и для каждого $i = 1, 2, \dots, n$ задан оператор φ_i пространства L_i. Тогда существует единственный оператор φ пространства L, являющийся прямой суммой операторов $\varphi_1, \varphi_2, \dots, \varphi_n$.*

Действительно, если a — произвольный вектор пространства L и $a = a_1 + a_2 + \dots + a_n$, где $a_i \in L_i$ $(i = 1, 2, \dots, n)$, полагаем

$$a\varphi = a_1\varphi_1 + a_2\varphi_2 + \dots + a_n\varphi_n.$$

Легко проверить, что так определенное отображение φ является линейным, подпространства $L_1, L_2, \dots, L_n$ инвариантны относительно φ и его ограничение на подпространство L_i совпадает с оператором φ_i. Таким образом, $\varphi = \varphi_1 \oplus \varphi_2 \oplus \dots \oplus \varphi_n$. Единственность такого отображения очевидна. $\square$

§ 2. СОБСТВЕННЫЕ ВЕКТОРЫ И СОБСТВЕННЫЕ ЗНАЧЕНИЯ ЛИНЕЙНОГО ОПЕРАТОРА. ХАРАКТЕРИСТИЧЕСКИЙ МНОГОЧЛЕН

В предыдущем параграфе мы увидели, что если оператор пространства L раскладывается в прямую сумму, то подходящим выбором базиса этого пространства матрицу оператора можно сделать блочно-диагональной. Среди таких матриц наиболее простой вид имеет диагональная матрица, т.е. матрица вида

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Тем не менее, как будет видно из дальнейшего, далеко не всякий оператор может иметь такую матрицу, и потому уместно принять следующее

Определение 1. *Оператор φ конечномерного линейного пространства L называется диагонализируемым, если его матрица в некотором базисе пространства L является диагональной.*

Следующее понятие естественно возникает в связи со свойством диагонализируемости оператора, но его применение при изучении операторов линейных пространств окажется гораздо более широким.

Определение 2. *Пусть L — линейное пространство над полем K и φ — оператор пространства L. Ненулевой вектор $a \in L$ называется собственным вектором оператора φ, принадлежащим собственному значению $\lambda \in K$, если $a\varphi = \lambda a$.*

Если говорить более подробно, этим определением введено два понятия: собственный вектор линейного оператора и собственное значение линейного оператора. Мы называем вектор a пространства L собственным вектором оператора φ, если $a \neq 0$ и образ вектора a относительно φ совпадает с λa для некоторого скаляра $\lambda \in K$. Равенством $a\varphi = \lambda a$ этот скаляр λ определяется однозначно (если $a\varphi = \lambda_1 a$ и $a\varphi = \lambda_2 a$, то $\lambda_1 a = \lambda_2 a$, $(\lambda_1 - \lambda_2)a = 0$, и так как $a \neq 0$, $\lambda_1 - \lambda_2 = 0$), и мы называем его собственным значением оператора φ и говорим, что вектор a принадлежит этому собственному значению. Таким образом, с другой стороны, элемент $\lambda \in K$ мы называем собственным значением оператора φ, если ему принадлежит хотя бы один собственный вектор, т.е. если в пространстве L существует вектор $a \neq 0$ такой, что $a\varphi = \lambda a$.

Следующее утверждение становится очевидным, стоит лишь вспомнить определение матрицы линейного оператора.

Предложение 1. *Матрица оператора φ линейного пространства L, вычисленная в базисе S этого пространства, диагональна тогда и только тогда, когда каждый вектор из S является собственным вектором оператора φ.* $\square$

Рассмотрим несколько примеров.

1. Если $\varphi = \chi_\lambda$ — гомотетия пространства L с коэффициентом λ, то каждый ненулевой вектор пространства L является собственным вектором оператора φ, принадлежащим собственному значению λ.

2. Оператор поворота на угол $\pi/2$ против часовой стрелки пространства V^2 не имеет, очевидно, собственных векторов.

3. Определим оператор φ пространства $L = \mathbb{R}^2$ следующим образом: если $a = (\alpha, \beta)$ — произвольный вектор из L, полагаем $a\varphi = (\alpha, \alpha + 2\beta)$. Равенство $a\varphi = \lambda a$ будет выполняться для некоторого $\lambda \in \mathbb{R}$ тогда и только тогда, когда $\alpha = \lambda\alpha$ и $\alpha + 2\beta = \lambda\beta$. Пусть $a \neq 0$. Если $\alpha \neq 0$, то первое из этих равенств дает $\lambda = 1$, и потому второе принимает вид $\alpha + \beta = 0$. Если же $\alpha = 0$, то $\beta \neq 0$ и потому из второго равенства получаем $\lambda = 2$. Следовательно, собственными значениями оператора φ являются числа 1 и 2, причем все собственные векторы, принадлежащие первому из них имеют вид $(-\beta, \beta)$, где $\beta \neq 0$, а все собственные векторы, принадлежащие второму, имеют вид $(0, \beta)$ где $\beta \neq 0$. Полагая в обоих случаях $\beta = 1$, получаем векторы $a_1 = (-1, 1)$ и $a_2 = (0, 1)$ соответственно. Так как система a_1, a_2 линейно независима, она составляет базис пространства L, и матрица оператора φ в этом базисе имеет вид $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$.

Характеризация и вычисление собственных значений оператора конечномерного пространства основаны на следующем понятии.

Определение 3. *Пусть A — квадратная матрица порядка n над полем K. Характеристическим многочленом матрицы A называется определитель матрицы $xE - A$ (где E — единичная матрица порядка n) .*

Таким образом, если $A = (\alpha_{ij})$, то характеристическим многочленом матрицы A нами назван определитель матрицы

$$f(x) = \begin{vmatrix} x - \alpha_{11} & -\alpha_{12} & \dots & -\alpha_{1n} \\ -\alpha_{21} & x - \alpha_{22} & \dots & -\alpha_{2n} \\ \dots & \dots & \dots & \dots \\ -\alpha_{n1} & -\alpha_{n2} & \dots & x - \alpha_{nn} \end{vmatrix},$$

составленной из многочленов, входящих в кольцо $K[x]$, и поэтому сам являющийся многочленом с коэффициентами из поля K.

Рассмотрим этот многочлен подробнее. Одним из слагаемых развернутой записи данного определителя является произведение его диагональных элементов

$$(x - \alpha_{11})(x - \alpha_{22})(x - \alpha_{nn}) = x^n - x^{n-1} \sum_{i=1}^{n} \alpha_{ii} + \cdots .$$

Легко понять, что все остальные слагаемые, входящие в запись определителя, являются многочленами от x, степень которых уже не превосходит числа $n - 2$. Действительно, любое слагаемое развернутой записи определителя, отличное от произведения диагональных элементов, содержит элемент из i-ой строки и j-ого столбца, где $i \neq j$. Но тогда в это слагаемое не могут войти по меньшей мере два элемента диагонали

$x - \alpha_{ii}$ и $x - \alpha_{jj}$, так как первый находится в той строке, а второй — в том столбце, представители которых уже вошли в рассматриваемое слагаемое.

Итак, если многочлен $f(x)$ является характеристическим многочленом матрицы A порядка n, то степень многочлена $f(x)$ равна n, а старший коэффициент равен 1 (т.е. многочлен $f(x)$ унитарен). Коэффициентом при x^{n-1} является след матрицы A, взятый с противоположным знаком, а свободный член равен $(-1)^n \det A$ (так как это — значение $f(x)$ при $x = 0$). Естественный вопрос, всякий ли унитарный многочлен является характеристическим многочленом некоторой матрицы, сейчас будет решен утвердительно.

Определение 4. *Пусть $f(x) = x^n + \alpha_1 x^{n-1} + \dots + \alpha_{n-1} x + \alpha_n$ — произвольный унитарный многочлен степени n над полем K. Матрицей Фробениуса этого многочлена называется $n \times n$-матрица вида*

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \hdotsfor{6} \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -\alpha_n & -\alpha_{n-1} & -\alpha_{n-2} & \dots & -\alpha_2 & -\alpha_1 \end{pmatrix}$$

Предложение 2. *Характеристический многочлен матрицы Фробениуса произвольного унитарного многочлена $f(x)$ совпадает с этим многочленом.*

Доказательство состоит в непосредственном вычислении определителя соответствующей матрицы. Сохраняя обозначения определения 4, имеем

$$\det(xE - A) = \begin{vmatrix} x & -1 & 0 & \dots & 0 & 0 \\ 0 & x & -1 & \dots & 0 & 0 \\ 0 & 0 & x & \dots & 0 & 0 \\ \hdotsfor{6} \\ 0 & 0 & 0 & \dots & x & -1 \\ \alpha_n & \alpha_{n-1} & \alpha_{n-2} & \dots & \alpha_2 & x + \alpha_1 \end{vmatrix}.$$

Выполним следующие элементарные преобразования столбцов этого определителя: К первому столбцу прибавляем последовательно

второй столбец, умноженный на x,

третий столбец, умноженный на x^2 и т.д. и, наконец,

последний столбец, умноженный на x^{n-1}.

В результате имеем определитель

$$\det(xE - A) = \begin{vmatrix} 0 & -1 & 0 & \dots & 0 & 0 \\ 0 & x & -1 & \dots & 0 & 0 \\ 0 & 0 & x & \dots & 0 & 0 \\ \hdotsfor{6} \\ 0 & 0 & 0 & \dots & x & -1 \\ f(x) & \alpha_{n-1} & \alpha_{n-2} & \dots & \alpha_2 & x + \alpha_1 \end{vmatrix},$$

раскладывая который по элементам первого столбца, получаем

$$\det(xE-A) = (-1)^{n+1} f(x) \begin{vmatrix} -1 & 0 & \dots & 0 & 0 \\ x & -1 & \dots & 0 & 0 \\ 0 & x & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x & -1 \end{vmatrix} = f(x). \quad \square$$

Прежде, чем вернуться к линейным операторам, отметим еще

Предложение 3. *Характеристические многочлены сопряженных матриц равны.*

В самом деле, если матрицы A и B сопряжены, т.е. для некоторой обратимой матрицы U выполнено равенство $A = UBU^{-1}$, то

$$\begin{aligned} &\det(xE-A) = \det(xE - UBU^{-1}) = \det(U(xE)U^{-1} - UBU^{-1}) = \\ &\det(U(xE-B)U^{-1}) = \det U \det(xE-B) \det(U^{-1}) = \\ &\det U \det(U^{-1}) \det(xE-B) = \det(xE-B). \end{aligned}$$

(Здесь мы последовательно использовали перестановочность матрицы U со скалярной матрицей xE, дистрибутивный закон в кольце матриц, перестановочность определителей матриц, как элементов поля K, и тот факт, что определители взаимно обратных матриц взаимно обратны.) $\square$

Поскольку матрицы линейного оператора пространства L, вычисленные в различных базисах этого пространства, являются сопряженными, предложение 3 обеспечивает корректность следующего определения.

Определение 5. *Пусть φ — оператор конечномерного линейного пространства L. Характеристическим многочленом оператора φ называется характеристический многочлен матрицы этого оператора в некотором базисе пространства L.*

Рассмотрим примеры.

1. Так как матрица гомотетии χ_λ n-мерного линейного пространства (в произвольном базисе) имеет вид

$$\lambda E = \begin{pmatrix} \lambda & 0 & 0 & \dots & 0 \\ 0 & \lambda & 0 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix},$$

то характеристический многочлен этого оператора равен

$$\det(xE - \lambda E) = \begin{vmatrix} x-\lambda & 0 & 0 & \dots & 0 \\ 0 & x-\lambda & 0 & \dots & 0 \\ 0 & 0 & x-\lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & x-\lambda \end{vmatrix} = (x-\lambda)^n.$$

2. В параграфе II.5 было показано, что матрица поворота пространства V^2 на угол $\dfrac{\pi}{2}$ против часовой стрелки в подходящем базисе имеет вид $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Поэтому характеристический многочлен этого оператора равен $\begin{vmatrix} x & -1 \\ 1 & x \end{vmatrix} = x^2 + 1$.

3. В том же параграфе мы видели, что матрица оператора дифференцирования пространства $\mathbb{R}_3[x]$ равна

$$A_\varphi = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix}.$$

Поэтому характеристический многочлен этого оператора есть x^4.

4. В начале этого параграфа был рассмотрен оператор пространства $\mathbb{R}^2$, отображающий вектор (α, β) в вектор $(\alpha, \alpha + 2\beta)$. Было показано, что в одном из базисов пространства $\mathbb{R}^2$ матрица этого оператора имеет вид $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, и потому его характеристическим многочленом является многочлен $(x-1)(x-2)$. Непосредственные вычисления показывают, что в стандартном базисе пространства $\mathbb{R}^2$ этот оператор имеет матрицу $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$, характеристический многочлен которой также (разумеется) равен многочлену $(x-1)(x-2)$.

Теорема 1. *Пусть φ — оператор конечномерного пространства L над полем K. Множество собственных значений оператора φ совпадает с множеством принадлежащих полю K корней его характеристического многочлена.*

Доказательство. Напомним, что по определению скаляр $\lambda \in K$ является собственным значением оператора φ тогда и только тогда, когда для некоторого ненулевого вектора $a \in L$ выполняется равенство $a\varphi = \lambda a$. Так как оно равносильно равенству $a(\chi_\lambda - \varphi) = 0$, означающему, что вектор a принадлежит ядру оператора $\chi_\lambda - \varphi$, то λ является собственным значением оператора φ в точности тогда, когда $\operatorname{Ker}(\chi_\lambda - \varphi)$ — ненулевое подпространство пространства L или, что то же самое, $\dim \operatorname{Ker}(\chi_\lambda - \varphi) > 0$. Так как по теореме I.3.6

$$\dim \operatorname{Im}(\chi_\lambda - \varphi) + \dim \operatorname{Ker}(\chi_\lambda - \varphi) = n,$$

где $n = \dim L$, мы видим, что скаляр λ является собственным значением оператора φ в точности тогда, когда ранг оператора $\chi_\lambda - \varphi$ меньше числа n. Так как в силу предложения I.5.5 ранг линейного отображения совпадает с рангом его матрицы, то скаляр λ является собственным значением оператора φ в точности тогда, когда ранг матрицы $A_{\chi_\lambda - \varphi}$ оператора $\chi_\lambda - \varphi$ меньше числа n, т.е. меньше порядка этой матрицы. Поскольку ранг квадратной матрицы меньше ее порядка тогда и только тогда, когда определитель этой матрицы равен нулю, и

$$A_{\chi_\lambda - \varphi} = A_{\chi_\lambda} - A_\varphi = \lambda E - A_\varphi,$$

мы получаем, наконец, что скаляр λ является собственным значением оператора φ тогда и только тогда, когда $\det(\lambda E - A_\varphi) = 0$, т.е. тогда и только тогда, когда λ является корнем характеристического многочлена $\det(xE - A_\varphi)$ оператора φ. $\square$

Доказанная теорема достаточно наглядно иллюстрируется примерами, приведенными выше. Так, отсутствие собственных векторов у оператора поворота пространства V^2 объясняется теперь тем, что характеристический многочлен x^2+1 этого оператора не имеет действительных корней. Эта теорема дает нам обоснование следующего алгоритма для нахождения собственных векторов линейного оператора φ пространства L:

1) вычислить матрицу A_φ линейного оператора φ в некотором базисе S пространства L (впрочем, в большинстве случаев оператор задается именно своей матрицей);

2) найти характеристический многочлен этой матрицы и отыскать все его корни, принадлежащие основному полю (тем самым окажутся найденными все собственные значения оператора φ);

3) для каждого собственного значения λ найти все ненулевые решения матричного уравнения $(a)_S(\lambda E - A_\varphi) = 0$ и тем самым — координатные строки в базисе S всех собственных векторов оператора φ, принадлежащих собственному значению λ.

Рассмотрим конкретный пример. Пусть требуется найти собственные векторы и собственные значения оператора φ пространства $\mathbb{R}^3$, заданного в стандартном базисе этого пространства матрицей

$$A = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 3 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Вычислим характеристический многочлен матрицы A:

$$f(x) = \begin{vmatrix} x-1 & -1 & 1 \\ 1 & x-3 & 1 \\ 1 & -1 & x-1 \end{vmatrix} = \begin{vmatrix} x-1 & x-2 & 1 \\ 1 & x-2 & 1 \\ 1 & 0 & x-1 \end{vmatrix} = (x-2)\begin{vmatrix} x-1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & x-1 \end{vmatrix} =$$

$$= (x-2)\begin{vmatrix} x-2 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & x-1 \end{vmatrix} = (x-2)^2(x-1).$$

(Здесь были выполнены следующие преобразования: первый столбец определителя прибавили ко второму, общий множитель $(x-2)$ элементов второго столбца вынесли за знак определителя, вторую строку вычли из первой и разложили определитель по элементам первой строки.) Итак, характеристический многочлен оператора φ имеет 1-кратный корень 1 и 2-кратный корень 2, т.е. у оператора φ ровно два собственных значения 1 и 2. Найдем теперь собственные векторы оператора φ. Так как координатная строка в стандартном базисе пространства $\mathbb{R}^3$ вектора $a = (\alpha_1, \alpha_2, \alpha_3)$ совпадает с этим вектором, вектор a удовлетворяет условию $a\varphi = a$ тогда и только тогда, когда выполняется матричное равенство

$$(\alpha_1\,\alpha_2\,\alpha_3)\begin{pmatrix} 1 & 1 & -1 \\ -1 & 3 & -1 \\ -1 & 1 & 1 \end{pmatrix} = (\alpha_1\,\alpha_2\,\alpha_3),$$

равносильное следующей системе линейных однородных уравнений:

$$\begin{cases} \alpha_1 - \alpha_2 - \alpha_3 = \alpha_1 \\ \alpha_1 + 3\alpha_2 + \alpha_3 = \alpha_2 \\ -\alpha_1 - 3\alpha_2 + \alpha_3 = \alpha_3 \end{cases} \quad \text{или} \quad \begin{cases} - \alpha_2 - \alpha_3 = 0 \\ \alpha_1 + 2\alpha_2 + \alpha_3 = 0 \\ -\alpha_1 - 3\alpha_2 + = 0 \end{cases}.$$

Приведя ее к ступенчатому виду, получим систему

$$\begin{cases} \alpha_1 + \alpha_2 \phantom{{}+\alpha_3} = 0 \\ \phantom{\alpha_1 + {}} \alpha_2 + \alpha_3 = 0 \end{cases},$$

множество решений которой состоит из векторов вида $(\alpha, -\alpha, \alpha)$, где α — произвольное действительное число. Таким образом, все собственные векторы оператора φ, принадлежащие собственному значению 1, имеют вид $(\alpha, -\alpha, \alpha)$, где α — произвольное действительное число, отличное от нуля. В частности, при $\alpha = 1$ получаем вектор $a = (1, -1, 1)$, а остальные собственные векторы оператора φ, принадлежащие собственному значению 1, имеют вид αa, где α — произвольное действительное число, отличное от нуля.

Аналогично, вектор $a = (\alpha_1, \alpha_2, \alpha_3)$ удовлетворяет условию $a\varphi = 2a$ тогда и только тогда, когда выполняется матричное равенство

$$(\alpha_1\,\alpha_2\,\alpha_3)\begin{pmatrix} 1 & 1 & -1 \\ -1 & 3 & -1 \\ -1 & 1 & 1 \end{pmatrix} = 2(\alpha_1\,\alpha_2\,\alpha_3),$$

равносильное, как показывают непосредственные вычисления, уравнению $\alpha_1 + \alpha_2 + \alpha_3 = 0$. Фундаментальная система решений системы, состоящей из одного этого уравнения, включает два вектора $b = (-1, 1, 0)$ и $c = (-1, 0, 1)$, и все собственные векторы оператора φ, принадлежащие собственному значению 2, имеют вид $\beta b + \gamma c$, где β и γ — произвольные действительные числа, не равные одновременно нулю.

Задача решена. Заметим еще, что система a, b, c линейно независима и потому составляет базис пространства $\mathbb{R}^2$. Матрица оператора φ в этом базисе имеет вид

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

так что наши вычисления позволили, в частности, обнаружить диагонализируемость оператора φ.

Укажем теперь одно достаточное условие диагонализируемости оператора конечномерного линейного пространства.

Теорема 2. *Пусть φ — оператор n-мерного линейного пространства L над полем K. Система собственных векторов оператора φ, принадлежащих попарно различным собственным значениям, линейно независима.*

В частности, если характеристический многочлен оператора φ имеет n различных корней (принадлежащих полю K), то этот оператор диагонализируем.

Доказательство. Пусть
$$a_1, a_2, \dots, a_m \tag{1}$$
— система собственных векторов оператора φ, принадлежащих попарно различным собственным значениям $\lambda_1, \lambda_2, \dots, \lambda_m$ соответственно. Это означает, что для каждого $i = 1, 2, \dots, m$ имеем $a_i \neq 0$ и $a_i\varphi = \lambda_i a_i$ и для любых $i \neq j$ $\lambda_i \neq \lambda_j$. Линейную независимость системы (1) будем доказывать индукцией по m. При $m = 1$ наше утверждение очевидно (система, состоящая из одного ненулевого вектора линейно независима), и мы предположим, что $m > 1$ и что любая система из $m - 1$ векторов, удовлетворяющая условиям теоремы, линейно независима. Из этого предположения следует, в частности, что подсистема $a_1, a_2, \dots, a_{m-1}$ системы (1) линейно независима.

Пусть для некоторых скаляров $\alpha_1, \alpha_2, \dots, \alpha_m \in K$ выполнено равенство
$$\alpha_1 a_1 + \alpha_2 a_2 + \dots + \alpha_m a_m = 0. \tag{2}$$
Вычисляя образы относительно оператора φ обеих частей этого равенства, получаем
$$\alpha_1(a_1\varphi) + \alpha_2(a_2\varphi) + \dots + \alpha_m(a_m\varphi) = 0,$$
или, с учетом равенств $a_i\varphi = \lambda_i a_i$,
$$(\alpha_1\lambda_1)a_1 + (\alpha_2\lambda_2)a_2 + \dots + (\alpha_m\lambda_m)a_m = 0. \tag{3}$$
Умножим обе части равенства (2) на скаляр $-\lambda_m$ и прибавим полученные выражения к соответствующим частям равенства (3):
$$(\alpha_1(\lambda_1 - \lambda_m))a_1 + (\alpha_2(\lambda_2 - \lambda_m))a_2 + \dots + (\alpha_{m-1}(\lambda_{m-1} - \lambda_m))a_{m-1} = 0.$$

Так как система $a_1, a_2, \dots, a_{m-1}$ линейно независима, отсюда следует, что для каждого $i = 1, 2, \dots, m-1$ имеет место равенство $\alpha_i(\lambda_i - \lambda_m) = 0$, и поскольку для любого из указанных значений i $\lambda_i \neq \lambda_m$, получаем, что для всех $i = 1, 2, \dots, m-1$ $\alpha_i = 0$. Тогда равенство (2) принимает вид $\alpha_m a_m = 0$, и так как $a_m \neq 0$, имеем $\alpha_m = 0$. Следовательно, система (1) является линейно независимой, и первое утверждение теоремы доказано.

Предположим теперь, что характеристический многочлен оператора φ имеет n различных корней $\lambda_1, \lambda_2, \dots, \lambda_n$, принадлежащих полю K. По теореме 1 для каждого $i = 1, 2, \dots, n$ существует собственный вектор a_i оператора φ, принадлежащий собственному значению λ_i. Тогда по доказанному система $a_1, a_2, \dots, a_n$ линейно независима и поскольку число ее векторов равно размерности пространства L, она является базисом этого пространства. В силу предложения 1 матрица оператора φ в этом базисе диагональна. $\square$

Рассмотренный перед этой теоремой пример показывает, что приведенное в ней достаточное условие диагонализируемости оператора не является необходимым.

В заключение этого параграфа приведем два простых свойства характеристических многочленов линейных операторов.

Предложение 4. *Пусть φ — оператор конечномерного линейного пространства L и M — φ-инвариантное подпространство этого пространства. Характеристический многочлен ограничения $\overline{\varphi} = \varphi|_M$ оператора φ на подпространство M является делителем характеристического многочлена оператора φ.*

Если оператор φ является прямой суммой операторов $\varphi_1, \varphi_2, \ldots, \varphi_n$, то характеристический многочлен оператора φ равен произведению характеристических многочленов операторов $\varphi_1, \varphi_2, \ldots, \varphi_n$.

В самом деле, если M — φ-инвариантное подпространство пространства L, то матрица оператора φ в подходящем базисе L имеет вид $A = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$, где B — матрица оператора $\overline{\varphi} = \varphi|_M$ (предложение 1.4). Тогда матрица $xE - A$ совпадает с матрицей

$$\begin{pmatrix} xE_1 - B & 0 \\ -C & xE_2 - D \end{pmatrix},$$

где E_1 и E_2 — единичные матрицы подходящих размеров. Вычисляя определитель этой матрицы с помощью теоремы Лапласа, имеем

$$\det(xE - A) = \det(xE_1 - B) \cdot \det(xE_2 - D)$$

Аналогично, если $\varphi = \varphi_1 \oplus \varphi_2 \oplus \cdots \oplus \varphi_n$, то в подходящем базисе пространства L матрица A оператора φ имеет блочно-диагональный вид

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_n \end{pmatrix},$$

где A_i — матрица оператора φ_i. Поэтому

$$xE - A = \begin{pmatrix} xE_1 - A_1 & & & \\ & xE_2 - A_2 & & \\ & & \ddots & \\ & & & xE_n - A_n \end{pmatrix},$$

и, следовательно, характеристический многочлен $\det(xE - A)$ матрицы A равен произведению характеристических многочленов $\det(xE_1 - A_1)$, $\det(xE_2 - A_2)$, … , $\det(xE_n - A_n)$ матриц A_1, A_2, … , A_n соответственно. $\square$

Предложение 5. *Пусть φ — оператор конечномерного пространства L и пусть пространство L является суммой φ-инвариантных подпространств $L_1, L_2, \dots, L_n$. Пусть $\varphi_i = \varphi|_{L_i}$ — ограничение оператора φ на подпространство L_i ($i = 1, 2, \dots, n$). Тогда характеристический многочлен оператора φ является делителем произведения характеристических многочленов операторов $\varphi_1, \varphi_2, \dots, \varphi_n$.*

Доказательство. При $n = 1$ утверждение тривиально. Рассмотрим случай, когда $n = 2$, т.е. $L = L_1 + L_2$. Выберем базис S_0 пересечения $L_0 = L_1 \cap L_2$ подпространств L_1 и L_2 и дополним его до базиса S_0S_1 подпространства L_1 и до базиса S_0S_2 подпространства L_2. Тогда (см. доказательство теоремы I.3.4) система $S_0S_1S_2$ является базисом пространства L. Так как пересечение φ-инвариантных подпространств является φ-инвариантным подпространством, подпространство L_0 инвариантно относительно оператора φ, а потому и относительно оператора φ_1. Следовательно, матрица оператора φ_1 в базисе S_0S_1 пространства L_1 имеет вид $\begin{pmatrix} B_0 & 0 \\ C_1 & D_1 \end{pmatrix}$ (где B_0 — матрица оператора $\varphi_0 = \varphi|_{L_0}$), и потому характеристический многочлен оператора φ_1 равен произведению характеристических многочленов $f_0(x)$ и $f_1(x)$ матриц B_0 и D_1 соответственно.

Аналогично, матрица оператора φ_2 в базисе S_0S_2 пространства L_2 имеет вид $\begin{pmatrix} B_0 & 0 \\ C_2 & D_2 \end{pmatrix}$, и потому характеристический многочлен оператора φ_2 равен произведению характеристических многочленов $f_0(x)$ и $f_2(x)$ матриц B_0 и D_2 соответственно.

Легко, далее, видеть, что матрица оператора φ в базисе $S_0S_1S_2$ пространства L имеет вид
$$\begin{pmatrix} B_0 & 0 & 0 \\ C_1 & D_1 & 0 \\ C_2 & 0 & D_2 \end{pmatrix},$$
и потому характеристический многочлен оператора φ равен произведению многочленов $f_0(x)f_1(x)f_2(x)$. Справедливость доказываемого утверждения при $n = 2$ теперь очевидна.

Предположим теперь, что $n > 2$. Так как сумма φ-инвариантных подпространств является φ-инвариантным подпространством, подпространство $L' = \sum_{i=1}^{n-1} L_i$ φ-инвариантно. Если φ' — ограничение оператора φ на подпространство L', то очевидно, что для любого $i = 1, 2, \dots, n-1$ оператор φ_i является ограничением на подпространство L_i оператора φ'. Поэтому по индукции можно считать, что характеристический многочлен оператора φ' является делителем произведения характеристических многочленов операторов $\varphi_1, \varphi_2, \dots, \varphi_{n-1}$. Так как $L = L' + L_n$, в силу рассмотренного случая двух слагаемых характеристический многочлен оператора φ является делителем произведения характеристических многочленов операторов φ' и φ_n. Следовательно, характеристический многочлен оператора φ является делителем произведения характеристических многочленов операторов $\varphi_1, \varphi_2, \dots, \varphi_n$. $\square$

§ 3. АННУЛЯТОРЫ. ТЕОРЕМА ГАМИЛЬТОНА - КЭЛИ

Определение 1. *Пусть φ — оператор линейного пространства L над полем K. Многочлен $f(x) \in K[x]$ называется аннулятором вектора $a \in L$, если образ вектора a относительно оператора $f(\varphi)$ равен нулю. Многочлен $f(x)$ называется аннулятором оператора φ, если оператор $f(\varphi)$ является нулевым.*

Если многочлен $f(x)$ является аннулятором оператора φ (или вектора a), мы будем говорить также, что оператор φ (или вектор a) аннулируется многочленом $f(x)$. При этом, говоря об аннуляторе вектора, мы всегда должны иметь в виду некоторый оператор пространства L; чтобы избежать неопределенности, мы будем говорить иногда, что вектор a аннулируется многочленом $f(x)$ с помощью оператора φ.

Ясно, что многочлен $f(x)$ аннулирует оператор φ пространства L тогда и только тогда, когда он является аннулятором (с помощью φ) каждого вектора $a \in L$. Понятно также, что нулевым многочленом аннулируется произвольный оператор, а потому и произвольный вектор пространства. Нулевой вектор аннулируется любым многочленом с помощью любого оператора. С другой стороны очевидно, что нулевой оператор аннулируется в точности теми многочленами, свободный член у которых равен нулю. Рассмотрим, далее, несколько конкретных примеров.

1. Пусть χ_λ — гомотетия линейного пространства L и $f(x) = x - \lambda$. Тогда поскольку $\chi_\lambda = \lambda\iota_L$, имеем $f(\chi_\lambda) = \chi_\lambda - \lambda\iota_L = \theta_L$, так что оператор χ_λ аннулируется многочленом $f(x) = x - \lambda$. Поэтому и каждый вектор пространства L аннулируется многочленом $f(x) = x - \lambda$ с помощью оператора χ_λ. Последнее утверждение обобщается следующим образом: произвольный вектор $a \in L$ аннулируется с помощью оператора φ пространства L линейным унитарным многочленом $x - \lambda$ тогда и только тогда, когда либо $a = 0$, либо a является собственным вектором оператора φ, принадлежащим собственному значению λ.

2. Если φ — оператор поворота пространства V^2 на угол $\pi/2$ против часовой стрелки, то оператор φ^2 поворачивает каждый вектор на угол π, а потому для любого вектора $a \in V^2$ имеет место равенство $a\varphi^2 = -a$ или, что равносильно, $a(\varphi^2 + \iota) = 0$. Поэтому все векторы пространства и, следовательно, оператор φ аннулируются многочленом $x^2 + 1$.

3. Пусть φ — оператор дифференцирования пространства $K[x]$. Легко видеть, что для любых неотрицательных целых чисел n и k

$$x^n\varphi^k = \begin{cases} 0, & \text{если } k > n \\ n(n-1)\cdots(n-k+1)x^{n-k}, & \text{если } 1 \le k \le n \\ x^n, & \text{если } k = 0. \end{cases}$$

Поэтому для любого многочлена $f(x) = \alpha_0 x^n + \alpha_1 x^{n-1} + \cdots + \alpha_{n-1}x + \alpha_n$ над полем K имеем $x^n f(\varphi) = \beta_0 + \beta_1 x + \cdots + \beta_{n-1}x^{n-1} + \beta_n x^n$, где

$$\beta_k = \begin{cases} n(n-1)\cdots(k+1)\alpha_k, & \text{если } 0 \le k < n \\ \alpha_n, & \text{если } k = n. \end{cases}$$

В частности, если $f(x)$ — ненулевой многочлен, то и оператор $f(\varphi)$ отличен от нулевого; иначе говоря, единственным аннулятором оператора φ является нулевой многочлен. С другой стороны, каждый вектор пространства $K[x]$ обладает ненулевыми аннуляторами: произвольный многочлен степени n, очевидно, аннулируется с помощью φ многочленом x^{n+1}.

4. Пусть $L = M \oplus N$ и φ — оператор проектирования на подпространство M параллельно подпространству N. Тогда, как мы видели, $\varphi = \iota_{_M} \oplus \theta_{_N}$. Поскольку для любого многочлена $f(x)$ имеет место равенство $f(\varphi) = f(\iota_{_M}) \oplus f(\theta_{_N})$, оператор φ аннулируется многочленом $x^2 - x$.

При рассмотрении примера 3 мы обнаружили, что оператор дифференцирования линейного пространства $K[x]$ аннулируется только нулевым многочленом. Сейчас мы увидим, что существование такого оператора оказалось возможным потому, что пространство $K[x]$ не является конечномерным.

Предложение 1. *Если линейное пространство L над полем K конечномерно, то для любого оператора пространства L существует ненулевой аннулятор.*

Доказательство. Пусть $\dim L = n$. Тогда отображение, сопоставляющее произвольному оператору φ пространства L матрицу этого оператора (в фиксированном базисе пространства L), является изоморфизмом пространства $\mathcal{L}(L)$ всех линейных операторов пространства L на пространство $M_n(K)$ всех квадратных матриц порядка n над полем K, и потому $\dim \mathcal{L}(L) = \dim M_n(K)$. Так как $\dim M_n(K) = n^2$, то и $\dim \mathcal{L}(L) = n^2$. Следовательно, для любого оператора φ пространства L система элементов

$$\iota = \varphi^0, \varphi^1, \ldots, \varphi^m$$

пространства $\mathcal{L}(L)$ (где $m = n^2$) должна быть линейно зависимой, а потому должны найтись скаляры $\alpha_0, \alpha_1, \ldots, \alpha_m \in K$, не равные нулю одновременно, такие, что

$$\alpha_0\varphi^0 + \alpha_1\varphi + \cdots + \alpha_m\varphi^m = 0.$$

Это и означает, очевидно, что многочлен $f(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_m x^m$ отличен от нуля и аннулирует оператор φ. $\square$

Отметим очевидное

Следствие. *Для любого вектора a конечномерного линейного пространства и любого оператора φ этого пространства найдется ненулевой многочлен, аннулирующий с помощью φ вектор a.*

Если для оператора (или вектора) линейного пространства L существует ненулевой аннулятор, то среди всех многочленов, аннулирующих этот оператор (или вектор) найдется многочлен наименьшей степени. Поскольку, к тому же, произведение скаляра на аннулирующий многочлен снова является аннулирующим многочленом, то для любого оператора и любого вектора в этом случае существует минимальный аннулятор в смысле следующего определения:

Определение 2. *Минимальным аннулятором оператора φ (вектора a) линейного пространства L называется унитарный многочлен, аннулирующий этот оператор (вектор) и имеющий наименьшую степень среди всех ненулевых аннуляторов оператора φ (вектора a).*

Из предложения 1 следует, что для любого вектора и любого оператора конечномерного линейного пространства существует минимальный аннулятор. Минимальный аннулятор оператора называют также *минимальным многочленом* этого оператора.

Заметим еще, что по аналогии с определением 1 можно ввести понятие аннулирующего многочлена квадратной матрицы. Так как сопоставление оператору n-мерного линейного пространства L над полем K матрицы этого оператора является изоморфизмом линейной алгебры $\mathcal{L}(L)$ на алгебру $M_n(K)$, всякий аннулятор оператора будет аннулятором и его матрицы и наоборот. Поэтому следствие к предложению 1 гарантирует существование ненулевого аннулятора произвольной квадратной матрицы. В частности, для любой квадратной матрицы существует минимальный многочлен, т. е. ненулевой аннулятор наименьшей степени.

Легко видеть, что всякий многочлен, делящийся на аннулятор оператора или вектора, сам будет аннулировать этот оператор или вектор. Действительно, если $f(x) = g(x)h(x)$ и $g(\varphi) = 0$, то $f(\varphi) = g(\varphi)h(\varphi) = 0$; соответственно, из $ag(\varphi) = 0$ получаем $af(\varphi) = a(g(\varphi)h(\varphi)) = (ag(\varphi))h(\varphi)) = 0$. Если аннулятор является минимальным, справедливо и обратное:

Предложение 2. *Минимальный аннулятор оператора или вектора является делителем каждого аннулятора этого оператора или вектора соответственно.*

Пусть, в самом деле, $p(x)$ — минимальный аннулятор оператора φ пространства L или вектора $a \in L$. Разделим произвольный многочлен $f(x)$ на многочлен $p(x)$ с остатком:

$$f(x) = p(x)q(x) + r(x), \quad \text{где либо } r(x) = 0, \text{ либо } \deg r(x) < \deg p(x).$$

Тогда $r(x) = f(x) - p(x)q(x)$, и потому если $p(\varphi) = f(\varphi) = 0$, то $r(\varphi) = 0$, а если $ap(\varphi) = af(\varphi) = 0$, то

$$ar(\varphi) = a(f(\varphi) - p(\varphi)q(\varphi)) = af(\varphi) - a(p(\varphi)q(\varphi)) = 0.$$

В любом случае отсюда следует, что $r(x) = 0$, так как предположив противное, мы получим ненулевой аннулятор $r(x)$ оператора φ или вектора a, степень которого меньше степени соответствующего минимального аннулятора. $\square$

Определение 3. *Пусть φ — оператор линейного пространства L. Циклическим подпространством, порожденным вектором $a \in L$, называется наименьшее φ-инвариантное подпространство пространства L, содержащее вектор a.*

Таким образом, говоря более подробно, циклическим подпространством пространства L, порожденным вектором a, называется подпространство M этого пространства, (однозначно) определяемое следующими свойствами: M инвариантно относительно φ, содержит вектор a и содержится в любом φ-инвариантном подпространстве, содержащем вектор a. Используя то, что пересечение произвольного семейства φ-инвариантных подпространств пространства L является φ-инвариантным подпространством, можно равносильным образом определить циклическое подпространство пространства L, порожденное вектором a, как пересечение всех φ-инвариантных подпространств, содержащих этот вектор. Так как хотя бы одно φ-инвариантное подпространство пространства L, содержащее вектор a, всегда найдется (скажем, само пространство L), это замечание делает факт существования циклического подпространства, порожденного произвольным вектором a, очевидным.

Теорема 1. *Пусть φ — оператор конечномерного линейного пространства L, a — произвольный вектор этого пространства. Размерность циклического подпространства M, порожденного вектором a, равна степени минимального аннулятора вектора a, а характеристический многочлен ограничения $\overline{\varphi} = \varphi|_M$ оператора φ на подпространство M совпадает с минимальным аннулятором вектора a.*

Доказательство. Пусть $p(x) = x^m + \alpha_1 x^{m-1} + \cdots + \alpha_{m-1}x + \alpha_m$ — минимальный аннулятор вектора a. Так как $a \in M$ и подпространство M φ-инвариантно, все векторы системы

$$e_1 = a, e_2 = a\varphi, e_3 = a\varphi^2, \ldots, e_m = a\varphi^{m-1} \tag{1}$$

принадлежат подпространству M, и мы докажем, что система (1) является базисом этого подпространства.

Для доказательства линейной независимости системы (1) предположим, что для некоторых скаляров $\beta_0, \beta_1, \ldots, \beta_{m-1}$ имеет место равенство

$$\beta_0 a + \beta_1(a\varphi) + \cdots + \beta_{m-1}(a\varphi^{m-1}) = 0.$$

Используя определения сложения операторов и умножения скаляра на оператор, мы можем переписать его в виде

$$a(\beta_0\iota + \beta_1\varphi + \cdots + \beta_{m-1}\varphi^{m-1}) = 0,$$

и это, очевидно означает, что вектор a аннулируется многочленом $g(x) = \beta_0 + \beta_1 x + \cdots + \beta_{m-1}x^{m-1}$. Если бы многочлен $g(x)$ был бы ненулевым, то его степень не превосходила бы числа $m-1$, т.е. оказалась бы меньше, чем степень минимального аннулятора $p(x)$ вектора a. Поэтому все коэффициенты $\beta_0, \beta_1, \ldots, \beta_{m-1}$ многочлена $g(x)$ должны быть равны нулю, и линейная независимость системы (1) доказана. Покажем теперь, что эта система порождает подпространство M.

Очевидно, что линейная оболочка N системы (1) содержится в подпространстве M. Для доказательства противоположного включения достаточно показать, что подпространство N инвариантно относительно φ. Для этого, в свою очередь, достаточно установить, что φ-образы всех векторов системы (1) принадлежат подпространству N. Так как при любом k, $1 \le k \le m-1$, имеем, очевидно,

$$e_k\varphi = e_{k+1}, \tag{2}$$

то следует рассмотреть лишь вектор $e_m\varphi = a\varphi^m$.

Так как многочлен $p(x)$ является аннулятором вектора a, имеем

$$a(\varphi^m + \alpha_1\varphi^{m-1} + \dots + \alpha_{m-1}\varphi + \alpha_m\iota) = 0,$$

или, что равносильно,

$$a\varphi^m + \alpha_1(a\varphi^{m-1}) + \dots + \alpha_{m-1}(a\varphi) + \alpha_m(a\iota) = 0.$$

Переписывая последнее равенство в виде

$$a\varphi^m = -\alpha_1(a\varphi^{m-1}) - \dots - \alpha_{m-1}(a\varphi) - \alpha_m(a\iota)$$

и вспоминая введенные обозначения, получаем

$$e_m\varphi = (-\alpha_m)e_1 + (-\alpha_{m-1})e_2 + \dots + (-\alpha_2)e_{m-1} + (-\alpha_1)e_m. \tag{3}$$

Таким образом, все векторы $e_1\varphi, e_2\varphi, \dots, e_m\varphi$ лежат в подпространстве N, и это, как отмечено выше, означает φ-инвариантность подпространства N, откуда следует справедливость включения $N\varphi \subseteq N$, а потому и равенства $M = N$.

Итак, мы показали, что система (1) является базисом подпространства M. Отсюда $\dim M = m = \deg p(x)$. Кроме того, из равенств (2) и (3) следует, что матрица оператора $\overline{\varphi}$ в базисе $e_1, e_2, \dots, e_m$ пространства M имеет вид

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -\alpha_m & -\alpha_{m-1} & -\alpha_{m-2} & \dots & -\alpha_2 & -\alpha_1 \end{pmatrix},$$

т.е. совпадает с матрицей Фробениуса многочлена $p(x)$. Второе утверждение теоремы теперь вытекает из предложения 2.2. $\square$

Простым следствием теоремы 1 является важное утверждение, известное как теорема Гамильтона – Кэли:

Теорема 2. *Произвольный оператор конечномерного линейного пространства аннулируется своим характеристическим многочленом.*

Доказательство. Пусть $f(x)$ — характеристический многочлен оператора φ конечномерного линейного пространства L. Возьмем произвольный вектор $a \in L$ и обозначим через M циклическое подпространство пространства L, порожденное вектором a, через $p(x)$ — минимальный аннулятор вектора a и через $\overline{\varphi}$ — ограничение оператора φ на подпространство M. Тогда по теореме 1 характеристический многочлен оператора $\overline{\varphi}$ совпадает с многочленом $p(x)$, а в силу предложения 2.4 он должен быть делителем характеристического многочлена $f(x)$ оператора φ. Следовательно, многочлен $f(x)$ делится на аннулятор вектора a и потому также аннулирует этот вектор. Итак, многочлен $f(x)$ аннулирует с помощью оператора φ каждый вектор пространства L, а это и означает, что оператор φ аннулируется многочленом $f(x)$. $\square$

Пусть A — квадратная матрица порядка n над полем K. Возьмем произвольное n-мерное линейное пространство L над этим полем, фиксируем некоторый его базис и найдем такой оператор φ пространства L, матрица которого в этом базисе совпадает с A. Так как характеристический многочлен матрицы A будет тогда характеристическим многочленом оператора φ и так как линейный оператор и его матрица аннулируются одними и теми же многочленами, мы приходим к следующему утверждению — классической формулировке теоремы Гамильтона – Кэли:

Следствие. *Пусть A — квадратная матрица над некоторым полем и $f(x)$ — характеристический многочлен этой матрицы. Тогда матрица $f(A)$ является нулевой.* $\square$

§ 4. РАЗЛОЖИМОСТЬ ЛИНЕЙНОГО ПРОСТРАНСТВА В ПРЯМУЮ СУММУ ПРИМАРНЫХ ПОДПРОСТРАНСТВ. КОРНЕВЫЕ ПОДПРОСТРАНСТВА

В параграфе 2 было показано, что характеристический многочлен прямой суммы операторов равен произведению характеристических многочленов слагаемых. Возникает естественный вопрос о справедливости обратного утверждения: верно ли, в частности, что если характеристический многочлен $f(x)$ оператора φ имеет вид $f(x) = g(x)h(x)$, то оператор φ является прямой суммой двух операторов с характеристическими многочленами $g(x)$ и $h(x)$ соответственно? Рассмотрим, в связи с этим, следующий пример.

Пусть L — 2-мерное линейное пространство и пусть оператор φ пространства L задан в некотором базисе этого пространства матрицей $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Тогда характеристический многочлен оператора φ есть $(x-1)^2$. Покажем, что пространство L не может быть разложено в прямую сумму двух ненулевых φ-инвариантных подпространств. Пусть, напротив, L_1 и L_2 — такие ненулевые φ-инвариантные подпространства пространства L, что $L = L_1 \oplus L_2$. Тогда подпространства L_1 и L_2 1-мерны и если φ_1 и φ_2 — ограничения оператора φ на подпространства L_1 и L_2 соответственно, то характеристические многочлены этих операторов должны быть линейными, а в силу предложения 2.5 их произведение равно $(x-1)^2$. Таким образом, характеристический многочлен каждого из операторов φ_1 и φ_2 равен $x-1$. Из теоремы Гамильтона - Кэли следует теперь, что операторы $\varphi_1 - \iota$ и $\varphi_2 - \iota$ являются нулевыми, т. е. каждый из операторов φ_1 и φ_2 является тождественным. Так как отсюда очевидным образом следует тождественность оператора φ, мы пришли к противоречию.

Таким образом, вопрос, сформулированный выше, решается отрицательно; тем не менее, в этом параграфе мы увидим, что ответ на него положителен, если многочлены $g(x)$ и $h(x)$ взаимно просты. Начнем с ряда вспомогательных результатов.

Предложение 1. *Если вектор a линейного пространства L над полем K аннулируется с помощью оператора φ этого пространства двумя взаимно простыми многочленами, то $a = 0$.*

В самом деле, пусть $f(x)$ и $g(x)$ — такие многочлены над полем K, что $af(\varphi) = ag(\varphi) = 0$ и наибольший общий делитель их равен 1. Тогда найдутся многочлены $u(x), v(x) \in K[x]$ такие, что $f(x)u(x) + g(x)v(x) = 1$. Вычисляя значения обеих частей этого равенства при $x = \varphi$, получаем $f(\varphi)u(\varphi) + g(\varphi)v(\varphi) = \iota$, откуда

$$a = a\iota = a\left(f(\varphi)u(\varphi) + g(\varphi)v(\varphi)\right) = \left(af(\varphi)\right)u(\varphi) + \left(ag(\varphi)\right)v(\varphi) = 0. \quad \square$$

Предложение 2. *Пусть вектор a линейного пространства L над полем K аннулируется с помощью оператора φ этого пространства многочленом вида*

$$f(x) = g_1(x)g_2(x)\cdots g_n(x),$$

где многочлены $g_1(x), g_2(x), \cdots, g_n(x)$ попарно взаимно просты. Тогда найдутся векторы $b_1, b_2, \ldots, b_n$ пространства L такие, что $a = b_1 + b_2 + \cdots + b_n$, причем для каждого $i = 1, 2, \ldots, n$ вектор b_i аннулируется многочленом $g_i(x)$.

Доказательство. Будем доказывать это утверждение индукцией по числу n сомножителей в указанном разложении многочлена $f(x)$. При $n = 1$ утверждение просто тривиально. Рассмотрим еще случай $n = 2$.

Пусть вектор a аннулируется произведением $f(x) = g_1(x)g_2(x)$ двух взаимно простых многочленов $g_1(x)$ и $g_2(x)$. Выбирая многочлены $u(x)$ и $v(x)$ из $K[x]$, удовлетворяющими соотношению $g_2(x)u(x) + g_1(x)v(x) = 1$, как и выше, получим равенство

$$a = a\iota = a\left(g_2(\varphi)u(\varphi)\right) + a\left(g_1(\varphi)v(\varphi)\right).$$

Полагаем $b_1 = a\left(g_2(\varphi)u(\varphi)\right)$ и $b_2 = a\left(g_1(\varphi)u(\varphi)\right)$. Тогда $a = b_1 + b_2$ и

$$b_1 g_1(\varphi) = \left(a\left(g_2(\varphi)u(\varphi)\right)g_1(\varphi)\right) = a\left(\left(g_2(\varphi)u(\varphi)\right)g_1(\varphi)\right) = \\ = a\left(\left(g_2(\varphi)g_1(\varphi)\right)u(\varphi)\right) = a\left(f(\varphi)u(\varphi)\right) = 0.$$

Аналогично, $b_2 g_2(\varphi) = 0$, так что векторы b_1 и b_2 искомые.

Предположим теперь, что $n > 2$ и для любого вектора пространства L, аннулируемого многочленом, который равен произведению меньшего, чем n, числа попарно взаимно простых сомножителей, доказываемое утверждение справедливо. Покажем, что тогда оно справедливо и для многочлена $f(x)$.

Пусть $f'(x) = g_1(x)g_2(x)\cdots g_{n-1}(x)$. Тогда $f(x) = f'(x)g_n(x)$ и многочлены $f'(x)$ и $g_n(x)$ являются взаимно простыми. Поэтому в силу рассмотренного случая двух сомножителей найдутся векторы a' и b_n такие, что $a = a' + b_n$, вектор a' аннулируется многочленом $f'(x)$ и вектор b_n аннулируется многочленом g_n. По индуктивному предположению для вектора a' существуют векторы $b_1, b_2, \ldots, b_{n-1}$ такие, что $a' = b_1 + b_2 + \cdots + b_{n-1}$ и вектор b_i аннулируется многочленом $g_i(x)$ $(i = 1, 2, \ldots n-1)$. Тогда $a = b_1 + b_2 + \cdots + b_n$ — искомое разложение вектора a. □

Определение 1. *Линейное пространство L над полем K с действующим на L оператором φ называется примарным (относительно φ), если характеристический многочлен оператора φ является степенью некоторого неприводимого над K многочлена.*

Мы увидим, что произвольное линейное пространство с оператором φ раскладывается в прямую сумму примарных подпространств. Но сначала укажем простой критерий примарности линейного пространства.

Предложение 3. *Пусть φ — оператор конечномерного пространства L над полем K и пусть минимальный аннулятор каждого вектора пространства L является некоторой степенью одного и того же неприводимого над K многочлена $p(x)$. Тогда и характеристический многочлен оператора φ совпадает с некоторой степенью многочлена $p(x)$.*

В самом деле, если векторы $a_1, a_2, \ldots, a_n$ составляют систему порождающих пространства L и для каждого $i = 1, 2, \ldots n$ M_i обозначает циклическое подпространство пространства L, порожденное вектором a_i, то $L = \sum_{i=1}^{n} M_i$, и потому из предложения 2.5 следует, что характеристический многочлен оператора φ является делителем произведения характеристических многочленов ограничений оператора φ на подпространства $M_1, M_2, \ldots, M_n$, т.е. — с учетом теоремы 3.1 — делителем некоторой степени многочлена $p(x)$. $\square$

Теорема 1. *Пусть характеристический многочлен $f(x)$ оператора φ конечномерного линейного пространства L над полем K имеет вид*

$$f(x) = p_1(x)^{k_1} p_2(x)^{k_2} \cdots p_n(x)^{k_n},$$

где $p_1(x), p_2(x), \ldots, p_n(x)$ — попарно различные (унитарные) неприводимые над полем K многочлены, $k_1, k_2, \ldots, k_n$ — положительные целые числа. Тогда оператор φ равен прямой сумме операторов $\varphi_1, \varphi_2, \ldots, \varphi_n$ с характеристическими многочленами $p_1(x)^{k_1}, p_2(x)^{k_2}, \ldots p_n(x)^{k_n}$ соответственно.

Доказательство. Для каждого $i = 1, 2, \ldots, n$ полагаем $L_i = \operatorname{Ker} p_i(\varphi)^{k_i}$; таким образом, вектор $a \in L$ принадлежит подпространству L_i тогда и только тогда, когда он аннулируется при помощи φ многочленом $p_i(x)^{k_i}$. Так как по теореме Гамильтона – Кэли произвольный вектор $a \in L$ аннулируется многочленом $f(x)$, из предложения 2 следует, что в пространстве L существуют такие векторы $b_1, b_2, \ldots, b_n$, что $a = b_1 + b_2 + \cdots + b_n$ и вектор b_i аннулируется многочленом $p_i(x)^{k_i}$, т.е. $b_i \in L_i$ $(i = 1, 2, \ldots, n)$. Это означает, что пространство L является суммой подпространств $L_1, L_2, \ldots, L_n$.

Для доказательства того, что эта сумма прямая, для каждого $i = 1, 2, \ldots, n$ введем многочлен

$$g_i(x) = p_1(x)^{k_1} \cdots p_{i-1}(x)^{k_{i-1}} p_{i+1}(x)^{k_{i+1}} \cdots p_n(x)^{k_n}.$$

Очевидно, что тогда для каждого i $(p_i(x)^{k_i}, g_i(x)) = 1$, а если $j \neq i$, то $p_j(x)^{k_j} | g_i(x)$. В частности, если $j \neq i$, то для любого вектора $b \in L_j$ имеем $bg_i(\varphi) = 0$. Пусть теперь вектор a принадлежит пересечению подпространств L_i и $\sum_{j \neq i} L_j$. Тогда вектор $a \in \sum_{j \neq i} L_j$ и потому равен сумме некоторых векторов $b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n$, принадлежащих соответственно подпространствам $L_1, \ldots, L_{i-1}, L_{i+1}, \ldots, L_n$. Так как каждое из этих слагаемых аннулируется многочленом $g_i(x)$, вектор a тоже аннулируется многочленом $g_i(x)$. Но вектор $a \in L_i$ и потому аннулируется также и многочленом $p_i(x)^{k_i}$. Из предложения 1 теперь следует, что $a = 0$. Следовательно,

$$L_i \cap \sum_{j \neq i} L_j = \{0\},$$

и потому $L = L_1 \oplus L_2 \oplus \cdots \oplus L_n$.

Ввиду следствия к предложению 1.1 каждое из подпространств $L_1, L_2, \ldots, L_n$ является φ-инвариантным, так что оператор φ равен прямой сумме операторов $\varphi_1, \varphi_2, \ldots, \varphi_n$ — своих ограничений на соответствующие подпространства $L_1, L_2, \ldots, L_n$.

Поскольку, далее, каждый вектор из подпространства L_i аннулируется многочленом $p_i(x)^{k_i}$, из предложения 3 следует, что характеристический многочлен оператора φ_i имеет вид $p_i(x)^{l_i}$ для некоторого целого $l_i \geq 0$. Тогда по предложению 2.4 характеристический многочлен оператора φ должен иметь вид

$$f(x) = p_1(x)^{l_1} p_2(x)^{l_2} \cdots p_n(x)^{l_n}.$$

Ввиду единственности канонической записи многочлена отсюда следует, что для каждого $i = 1, 2, \ldots, n$ должно выполняться равенство $l_i = k_i$, и теорема доказана. □

Приведем два следствия из доказанной теоремы. Первое из них относится к вопросу, сформулированному в начале этого параграфа.

Следствие 1. *Пусть характеристический многочлен $f(x)$ оператора φ является произведением взаимно простых многочленов $g(x)$ и $h(x)$. Тогда оператор φ раскладывается в прямую сумму операторов ψ и ξ с характеристическими многочленами $g(x)$ и $h(x)$ соответственно.*

В самом деле, сохраняя все обозначения из формулировки и доказательства теоремы 1, можно предполагать без потери общности, что для некоторого числа s, $1 \leq s \leq n$,

$$g(x) = p_1(x)^{k_1} p_2(x)^{k_2} \cdots p_s(x)^{k_s} \quad \text{и} \quad h(x) = p_{s+1}(x)^{k_{s+1}} p_{s+2}(x)^{k_{s+2}} \cdots p_n(x)^{k_n}.$$

Тогда полагая $M = \sum_{i=1}^{s} L_i$ и $N = \sum_{i=s+1}^{n} L_i$, по свойствам прямой суммы подпространств будем иметь

$$M = \bigoplus_{i=s+1}^{n} L_i, \; N = \bigoplus_{i=s+1}^{n} L_i \text{ и } L = M \oplus N.$$

Полагая, далее, $\psi = \varphi|_M$ и $\xi = \varphi|_N$, получаем, очевидно, $\varphi = \psi \oplus \xi$, $\psi = \varphi_1 \oplus \varphi_2 \oplus \cdots \oplus \varphi_s$ и $\xi = \varphi_{s+1} \oplus \varphi_{s+2} \oplus \cdots \oplus \varphi_n$. Отсюда и из предложения 2.4 следует, что характеристическими многочленами операторов ψ и ξ являются многочлены $g(x)$ и $h(x)$ соответственно. □

Следствие 2. *Пусть φ — оператор линейного пространства L над полем K. Всякий неприводимый над K делитель характеристического многочлена оператора φ является делителем и минимального многочлена этого оператора.*

Более важное применение теоремы 1 связано со следующим понятием:

Определение 1. *Пусть φ — оператор линейного пространства L над полем K. Вектор a называется корневым вектором оператора φ, принадлежащим значению $\lambda \in K$, если a аннулируется при помощи φ многочленом вида $(x-\lambda)^r$, где $r \geq 0$.*

Множество $L(\lambda)$ всех корневых векторов оператора φ, принадлежащих значению λ, непусто, так как содержит, очевидно, нулевой вектор пространства L. Если вектор $a \in L$ входит в это множество, т. е. $a(\varphi - \lambda\iota)^r = 0$ для некоторого $r \geq 0$, то для любого $\alpha \in K$

$$(\alpha a)(\varphi - \lambda\iota)^r = \alpha(a(\varphi - \lambda\iota)^r) = 0,$$

так что $\alpha a \in L(\lambda)$. Кроме того, если векторы a_1 и a_2 принадлежат множеству $L(\lambda)$ и потому аннулируются при помощи оператора φ многочленами $(x-\lambda)^{r_1}$ и $(x-\lambda)^{r_2}$ соответственно, то каждый из этих векторов, а значит и их сумма аннулируется многочленом $(x-\lambda)^r$, где $r \geq \max\{r_1, r_2\}$. Поэтому $a_1 + a_2 \in L(\lambda)$, $L(\lambda)$ является подпространством пространства L и называется *корневым подпространством пространства L, принадлежащим значению λ*. Нетрудно получить более точную информацию об этом подпространстве.

Теорема 2. *Пусть φ — оператор конечномерного линейного пространства L над полем K. Для произвольного скаляра $\lambda \in K$ корневое подпространство $L(\lambda)$ является ненулевым тогда и только тогда, когда λ — собственное значение оператора φ. При этом размерность подпространства $L(\lambda)$ совпадает с кратностью корня λ характеристического многочлена оператора φ.*

Доказательство. Пусть вектор a принадлежит подпространству $L(\lambda)$ и $a \neq 0$. Тогда минимальный аннулятор вектора a имеет вид $(x-\lambda)^s$, где $s > 0$. Из теоремы 3.1 и предложения 2.4 теперь следует, что характеристический многочлен $f(x)$ оператора φ делится на многочлен $(x-\lambda)^s$, и потому λ является корнем многочлена $f(x)$. Следовательно, λ — собственное значение оператора φ.

Обратно, если λ — корень многочлена $f(x)$ и если

$$f(x) = p_1(x)^{k_1} p_2(x)^{k_2} \cdots p_n(x)^{k_n}$$

— каноническое разложение этого многочлена, как в формулировке теореме 1, то один из его неприводимых множителей $p_1(x), p_2(x), \ldots, p_n(x)$ должен совпасть с многочленом $x-\lambda$. Будем (без потери общности) считать, что этим множителем является $p_1(x)$ (и потому λ является k_1-кратным корнем многочлена $f(x)$). Сохраняя для обозначений $L_1, L_2, \ldots, L_n$ те же значения, что и в доказательстве теоремы 1, покажем, что $L(\lambda) = L_1$.

В самом деле, поскольку каждый вектор подпространства L_1 аннулируется многочленом $(x-\lambda)^{k_1}$, включение $L_1 \subseteq L(\lambda)$ очевидно. Пусть $a \in L(\lambda)$ и пусть $a = b_1 + b_2 + \cdots + b_n$, где $b_i \in L_i$ для $i = 1, 2, \ldots n$. Так как вектор a аннулируется некоторым многочленом вида $(x-\lambda)^r$, имеем

$$0 = a(\varphi - \lambda\iota)^r = b_1(\varphi - \lambda\iota)^r + b_2(\varphi - \lambda\iota)^r + \cdots + b_n(\varphi - \lambda\iota)^r.$$

Здесь каждое слагаемое $b_i(\varphi-\lambda\iota)^r$ входит в подпространство L_i (ввиду его φ-инвариантности), и так как $L=\bigoplus_{i=1}^n L_i$, получаем

$$b_i(\varphi-\lambda\iota)^r = 0 \quad (i=1,2,\dots,n).$$

Таким образом, каждый из векторов b_i аннулируется многочленом $(x-\lambda)^r$. Так как, кроме того, эти векторы аннулируются многочленами $p_1(x)^{k_1}, p_2(x)^{k_2}, \dots, p_n(x)^{k_n}$ соответственно и при $i \geq 2$ многочлены $p_i(x)^{k_i}$ и $(x-\lambda)^r$ взаимно просты, из предложения 1 следует, что $b_2=\dots=b_n=0$. Следовательно, вектор $a=b_1$ принадлежит подпространству L_1, и включение $L(\lambda)\subseteq L_1$ доказано. Тем самым доказано и равенство $L(\lambda)=L_1$. Ввиду теоремы 1 это означает, что характеристический многочлен ограничения оператора φ на подпространство $L(\lambda)$ совпадает с $(x-\lambda)^{k_1}$, и так как степень характеристического многочлена оператора равна размерности пространства, на котором действует этот оператор, имеем $\dim L(\lambda)=k_1$. $\square$

Сформулируем, наконец, следующий частный случай теоремы 1:

Теорема 3. *Пусть φ — оператор конечномерного линейного пространства L над полем K и пусть характеристический многочлен $f(x)$ оператора φ раскладывается в $K[x]$ на линейные множители. Тогда пространство L раскладывается в прямую сумму корневых подпространств. Более подробно, пусть*

$$f(x)=(x-\lambda_1)^{k_1}(x-\lambda_2)^{k_2}\cdots(x-\lambda_n)^{k_n},$$

где $k_1,k_2,\dots,k_n>0$ и $\lambda_1,\lambda_2,\dots,\lambda_n$ — попарно различные элементы из K. Тогда

$$L=\bigoplus_{i=1}^n L(\lambda_i)$$

и характеристический многочлен ограничения φ_i оператора φ на подпространство $L(\lambda_i)$ равен $(x-\lambda_i)^{k_i}$ $(i=1,2,\dots,n)$. $\square$

§5. НОРМАЛЬНАЯ ФОРМА ЖОРДАНА МАТРИЦЫ ЛИНЕЙНОГО ОПЕРАТОРА

Из предыдущей главы мы знаем, что матрица линейного оператора конечномерного линейного пространства зависит от выбора базиса этого пространства и что две матрицы могут являться матрицами одного и того же оператора тогда и только тогда, когда они сопряжены. Вопросу выбора базиса пространства, матрица оператора в котором имела бы наиболее простой вид, были посвящены предыдущие параграфы этой главы. Результатом этих рассмотрений явилась теорема 3.3, которая утверждает, что разложению характеристического многочлена оператора в произведение степеней различных неприводимых многочленов соответствует разложение оператора в прямую сумму и, следовательно, возможность выбора такого базиса, матрица оператора в котором имеет блочно-диагональный вид. Тем не менее, эта теорема ничего не говорит о строении диагональных блоков такой матрицы и потому не позволяет различать операторы с одинаковыми характеристическими многочленами.

Говоря более подробно, здесь имеется в виду следующее. Если матрицы двух операторов линейного пространства в подходящих базисах оказались равными, то понятно, что эти операторы имеют одни и те же свойства, т.е. их следует считать в определенном смысле одинаковыми (если их матрицы совпали в одном и том же базисе, то эти операторы просто равны). Хотелось бы уметь распознавать, являются ли одинаковыми два оператора, заданные матрицами в некоторых базисах нашего пространства, т.е. являются ли их матрицы сопряженными. Отрицательный ответ получается сразу, если характеристические многочлены операторов (или их матриц) различны. Но если они совпадают, решения поставленной задачи у нас пока нет. Мы смогли бы ее решить (по крайней мере, теоретически) выбором канонического представителя в каждом классе сопряженных между собой матриц. Это означает выделение таких матриц определенного вида, что любая матрица сопряжена с одной из них и по двум таким матрицам всегда можно узнать, сопряжены ли они между собой. В этом параграфе мы увидим, что матрицы нормальной жордановой формы позволяют решить эту задачу для некоторых операторов линейных пространств (в частности, для любого оператора линейного пространства над алгебраически замкнутым полем).

В следующем определении рассматриваются линейные пространства и матрицы над некоторым фиксированным полем K.

Определение 1. *Жордановой клеткой порядка r с собственным значением λ называется $r \times r$-матрица вида*

$$J_r(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix},$$

если $r \geq 2$, и 1×1-матрица (λ), единственным элементом которой является скаляр λ, если $r = 1$.

Жордановой матрицей (или матрицей, имеющей нормальную жорданову форму) называется блочно-диагональная матрица

$$A = \begin{pmatrix} J_{r_1}(\lambda_1) & & & \\ & J_{r_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{r_n}(\lambda_n) \end{pmatrix},$$

диагональные блоки которой являются жордановыми клетками.

Если φ — некоторый оператор конечномерного пространства L, то базис S этого пространства называется жордановым для оператора φ, если матрица этого оператора в базисе S является жордановой.

Отметим сразу же, что если матрица A оператора φ в базисе S пространства L имеет жорданову форму, указанную в определении, то система S является соединением систем $S_1, S_2, \ldots, S_n$, где S_i составлена из векторов системы S, соответствующих тем строкам матрицы A, в которых расположена клетка $J_{r_i}(\lambda_i)$. Очевидно, что перестановка подсистем $S_1, S_2, \ldots, S_n$ системы S вызовет соответствующую перестановку диагональных блоков матрицы A. Например, если базис S' пространства L является соединением систем $S_2, S_1, S_3, \ldots, S_n$, то матрица оператора φ в этом базисе имеет вид

$$A = \begin{pmatrix} J_{r_2}(\lambda_2) & & & & \\ & J_{r_1}(\lambda_1) & & & \\ & & J_{r_3}(\lambda_3) & & \\ & & & \ddots & \\ & & & & J_{r_n}(\lambda_n) \end{pmatrix}.$$

Таким образом, меняя порядок следования диагональных клеток в нормальной жордановой матрице оператора, мы снова получаем жорданову матрицу того же оператора; ниже будет показано, что справедливо и обратное, т. е. если для оператора линейного пространства существует жорданов базис, то любые две жордановы матрицы этого оператора различаются лишь порядком следования жордановых клеток.

Так как характеристический многочлен жордановой клетки $J_r(\lambda)$ равен $(x-\lambda)^r$, то характеристический многочлен жордановой матрицы

$$A = \begin{pmatrix} J_{r_1}(\lambda_1) & & & \\ & J_{r_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{r_n}(\lambda_n) \end{pmatrix}$$

имеет вид $(x-\lambda_1)^{r_1}(x-\lambda_2)^{r_2}\cdots(x-\lambda_n)^{r_n}$. Поэтому если в пространстве L существует жорданов базис для оператора φ, характеристический многочлен этого оператора должен раскладываться в кольце многочленов $K[x]$ на линейные множители. Весь этот параграф посвящен доказательству того, что это необходимое условие существования жорданова базиса является и достаточным. Точнее говоря, будет доказана

Теорема. *Пусть φ — оператор конечномерного линейного пространства L над полем K и пусть характеристический многочлен оператора φ раскладывается над K на линейные множители. Тогда в пространстве L существует жорданов базис для оператора φ. При этом жорданова матрица оператора φ определена однозначно с точностью до перестановки диагональных блоков.*

Доказательство этой теоремы займет весь остаток этого параграфа. Сначала будет показано, что вопрос о существовании и единственности жордановой матрицы оператора сводится к соответствующему вопросу для операторов специального вида — нильпотентных операторов.

Пусть характеристический многочлен оператора φ имеет вид

$$f(x) = (x - \lambda_1)^{k_1}(x - \lambda_2)^{k_2} \cdots (x - \lambda_n)^{k_n},$$

где $k_1, k_2, \ldots, k_n$ — положительные целые числа и $\lambda_1, \lambda_2, \ldots \lambda_n$ — попарно различные элементы поля K. Тогда в силу теоремs 4.3 пространство L является прямой суммой корневых подпространств $L(\lambda_1), L(\lambda_2), \ldots L(\lambda_n)$, причем характеристический многочлен ограничения φ_i оператора φ на подпространство L_i есть $(x-\lambda_i)^{k_i}$ $(i = 1, 2, \ldots, n)$.

Предложение 1. *Базис S пространства L является жордановым для оператора φ тогда и только тогда, когда он с точностью до перестановки его векторов совпадает с соединением некоторых жордановых базисов для операторов $\varphi_1, \varphi_2, \ldots, \varphi_n$ корневых подпространств $L(\lambda_1), L(\lambda_2), \ldots L(\lambda_n)$ соответственно.*

Доказательство. Так как блочно-диагональная матрица, диагональные блоки которой имеют жорданову нормальную форму, является, очевидно, жордановой, то соединение жордановых базисов подпространств $L(\lambda_1), L(\lambda_2), \ldots L(\lambda_n)$ будет жордановым базисом пространства L.

Обратно, предположим, что матрица A оператора φ в базисе S пространства L имеет жорданову нормальную форму. Так как собственное значение каждой жордановой клетки матрицы A является собственным значением оператора φ, то оно должно совпадать с одним из скаляров $\lambda_1, \lambda_2, \ldots \lambda_n$. Меняя местами жордановы клетки (с помощью подходящих перестановок векторов базиса S), приведем матрицу A к виду

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_n \end{pmatrix},$$

где для любого $i = 1, 2, \ldots, n$ A_i — жорданова матрица порядка l_i, все клетки которой относятся к собственному значению λ_i, и потому ее характеристический многочлен имеет вид $(x-\lambda_i)^{l_i}$. Считая (без потери общности), что уже исходная матрица A имеет такой вид, имеем представление базиса S в виде соединения таких систем векторов $S_1, S_2, \ldots, S_n$, что ограничение ψ_i оператора φ на подпространство $L_i = l(S_i)$ имеет в базисе S_i матрицу A_i, и потому его характеристический многочлен равен $(x - \lambda_i)^{l_i}$

($i = 1, 2, \ldots, n$). Из теоремы Гамильтона – Кэли теперь следует, что для любого $i = 1, 2, \ldots, n$ $L_i \subseteq L(\lambda_i)$, откуда, в частности, получаем $l_i = \dim L_i \leq \dim L(\lambda_i) = k_i$. Так как

$$L = \bigoplus_{i=1}^{n} L_i \quad \text{и} \quad L = \bigoplus_{i=1}^{n} L(\lambda_i),$$

то $\sum_{i=1}^{n} l_i = \sum_{i=1}^{n} k_i$, откуда для любого $i = 1, 2, \ldots, n$ получаем $l_i = k_i$ и потому $L_i = L(\lambda_i)$. Таким образом, базис является соединением жордановых базисов $S_1, S_2, \ldots, S_n$ корневых подпространств $L(\lambda_1), L(\lambda_2), \ldots L(\lambda_n)$. $\square$

Предложение 1 показывает, что существование жорданова базиса и единственность нормальной формы матрицы линейного оператора достаточно доказать для корневых подпространств данного пространства. Иначе говоря, мы можем предполагать, что пространство L совпадает с одним из своих корневых подпространств, т.е. что характеристический многочлен оператора φ имеет вид $(x-\lambda)^k$. Это и есть первый шаг сведения, о которм шла речь сразу после формулировки теоремы. Второй шаг сведения основан на следующем утверждении:

Предложение 2. *Если S — жорданов базис для оператора φ линейного пространства L, то для любого скаляра $\lambda \in K$ S является жордановым базисом и для оператора $\varphi + \lambda\iota$.*

Действительно, если матрица оператора φ в базисе S имеет вид

$$A = \begin{pmatrix} J_{r_1}(\lambda_1) & & & \\ & J_{r_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{r_n}(\lambda_n) \end{pmatrix}$$

то матрица оператора $\varphi + \lambda\iota$, равная $A + \lambda E$, совпадает с матрицей

$$A = \begin{pmatrix} J_{r_1}(\lambda_1 + \lambda) & & & \\ & J_{r_2}(\lambda_2 + \lambda) & & \\ & & \ddots & \\ & & & J_{r_n}(\lambda_n + \lambda) \end{pmatrix}. \quad \square$$

Таким образом, проблема существования жорданова базиса и единственности нормальной формы матрицы линейного оператора φ, характеристический многочлен которого имеет вид $(x - \lambda)^k$, сводится к соответствующей проблеме для оператора $\psi = \varphi - \lambda\iota$. Так как по теореме Гамильтона – Кэли оператор φ аннулируется многочленом $(x - \lambda)^k$, имеем

$$\psi^k = (\varphi - \lambda\iota)^k = 0,$$

так что оператор нильпотентен в смысле следующего определения.

Определение 2. *Оператор φ линейного пространства L называется нильпотентным, если для некоторого целого числа $m \geq 1$ имеет место равенство $\varphi^m = 0$. Наименьшее такое число m называется показателем нильпотентности оператора φ.*

(Например, оператор дифференцирования пространства многочленов $K_3[x]$ является нильпотентым ; показатель нильпотентности этого оператора равен 4.)

Итак, из предложений 1 и 2 следует, что теорему достаточно доказать для произвольного нильпотентного оператора. Начнем с очевидного замечания о таких операторах.

Предложение 3. *Единственным собственным значением нильпотентного оператора является нулевой скаляр.*

В самом деле, если $\varphi^m = 0$ и a — такой ненулевой вектор пространства L, что $a\varphi = \lambda a$, то поскольку тогда для любого числа $k \geq 0$ $a\varphi^k = \lambda^k a$, получаем $\lambda^m a = 0$, и так как $a \neq 0$, $\lambda = 0$.

С другой стороны, нуль является, очевидно, собственным значением нулевого оператора. Если же показатель нильпотентности оператора φ равен m и $m > 1$, то существует вектор $b \in L$ такой, что $b\varphi^{m-1} \neq 0$. Тогда вектор $a = b\varphi^{m-1}$ является собственным вектором оператора φ, принадлежащим собственному значению 0. $\square$

Таким образом, если S — жорданов базис для нильпотентного оператора φ, то все клетки жордановой матрицы этого оператора в базисе S относятся к собственному значению 0. Посмотрим, как действует оператор φ на векторы базиса S. Предположим сначала, что матрицей оператора в базисе $S = (e_1, e_2, \dots e_r)$ является жорданова клетка

$$J_r(0) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}.$$

Это означает, очевидно, что φ действует на векторы базиса S следующим образом:

$$e_1\varphi = e_2, e_2\varphi = e_3, \dots, e_{r-1}\varphi = e_r, e_r\varphi = 0.$$

Такое действие оператора на векторы базиса можно наглядно изобразить графически в виде следующей диаграммы:

$$\begin{array}{c} e_1 \\ \downarrow \\ e_2 \\ \downarrow \\ \vdots \\ e_{r-1} \\ \downarrow \\ e_r \end{array}$$

Ясно, что в общем случае, когда жорданова матрица нильпотентного оператора φ состоит из нескольких клеток Жордана, соответствующий базис пространства можно расположить в диаграмму, состоящую уже из нескольких таких столбиков — по одному для каждой клетки. Говоря точнее, имеет место почти очевидное

Предложение 4. *Пусть φ — нильпотентный оператор конечномерного пространства L. Базис S пространства L является жордановым для оператора φ тогда и только тогда, когда векторы из S располагаются в диаграмму относительно действия φ. При этом, число столбцов диаграммы равно числу клеток Жордана, составляющих матрицу оператора φ в базисе S, а число векторов в столбце — порядку соответствующей клетки.* $\square$

Например, если матрица оператора 10-мерного пространства в базисе $e_1, \ldots, e_{10}$ имеет вид

$$\begin{pmatrix} 0&1&0&0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&1&0&0&0&0\\ 0&0&0&0&0&0&1&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0 \end{pmatrix},$$

т.е. состоит из четырех жордановых клеток порядков 4, 3, 2 и 1, то векторы этого базиса располагаются в следующую диаграмму относительно действия φ:

$$\begin{array}{cccc} e_1 & & & \\ \downarrow & & & \\ e_2 & e_5 & & \\ \downarrow & \downarrow & & \\ e_3 & e_6 & e_8 & \\ \downarrow & \downarrow & \downarrow & \\ e_4 & e_7 & e_9 & e_{10} \end{array}$$

Таким образом, нам достаточно показать, что для любого нильпотентного оператора пространства L существует базис, который можно расположить в диаграмму, а вид диаграммы опеределен однозачно с точностью до перестановки столбцов.

Пусть, например, снова φ — оператор дифференцирования пространства $K_3[x]$. Так как $x^3\varphi = 3x^2, 3x^2\varphi = 6x, 6x\varphi = 6, 6\varphi = 0$, то векторы $e_1 = x^3, e_2 = 3x^2, e_3 = 6x, e_4 = 6$ составляющие линейно независимую систему, т.е. базис этого пространства, располагаются в диаграмму из одного столбца. Поэтому матрица оператора φ в этом базисе является жордановой клеткой $J_4(0)$.

Предположим, что некоторая система векторов пространства расположена в диаграмму относительно действия оператора φ. Будем считать, что находящиеся на одном уровне векторы диаграммы составляют ее строку. Строки диаграммы будем нумеровать снизу вверх, полагая номер нижней строки (состоящей из векторов, образом которых относительно φ является нулевой вектор) равным 1. Высотой вектора назовем номер строки, в которой он расположен. Высотой столбца будем считать высоту самого верхнего вектора (или вершины) этого столбца. Так, в приведенном выше примере расположенного в диаграмму базиса 10-мерного пространства высота векторов e_4, e_7, e_9, e_{10} равна 1, высота векторов e_3, e_6, e_8 равна 2, высота векторов e_2, e_5 равна 3 и e_1 является единственным вектором высоты 4.

Переходя к изучению свойств системы векторов, расположенной в диаграмму, отметим, прежде всего, очевидное

Предложение 5. *Пусть высота вектора a диаграммы равна m. Тогда при $m > 1$ вектор $a\varphi$ имеет высоту $m-1$, а при $m = 1$ вектор $a\varphi$ нулевой. Поэтому для любого целого числа $k \geq 1$ вектор $a\varphi^k$ имеет высоту $m - k$, если $k < m$, и $a\varphi^k = 0$, если $k \geq m$.* $\square$

Предложение 6. *Пусть φ — нильпотентный оператор линейного пространства L и пусть S — расположенный в диаграмму (т.е. жорданов для оператора φ) базис пространства L. Для любого целого числа $k \geq 1$ ядро оператора φ^k совпадает с подпространством пространства L, порожденным теми векторами системы S, высота которых не превосходит k.*

Доказательство. Занумеруем векторы $e_1, e_2, \ldots, e_n$ базиса S таким образом, что для некоторого числа r, $1 \leq r \leq n$, высота векторов $e_1, e_2, \ldots, e_r$ не превосходит числа k, а высота векторов $e_{r+1}, \ldots, e_n$ больше, чем k. Тогда для любого $i = 1, 2, \ldots r$ в силу предложения 5 имеем $e_i\varphi^k = 0$, т.е. $e_i \in \operatorname{Ker}\varphi^k$. Поэтому $l(e_1, e_2, \ldots, e_r) \subseteq \operatorname{Ker}\varphi^k$.

Обратно, пусть a — произвольный вектор из ядра оператора φ^k и пусть

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n$$

— выражение этого вектора через векторы базиса S. Действуя на обе части этого равенства оператором φ^k получаем, очевидно,

$$\alpha_{r+1}(e_{r+1}\varphi^k) + \alpha_{r+2}(e_{r+2}\varphi^k) + \cdots + \alpha_n(e_n\varphi^k) = 0.$$

Но векторы системы $e_{r+1}\varphi^k, e_{r+2}\varphi^k, \ldots, e_n\varphi^k$ попарно различны: для любых i и j, $r + 1 \leq i < j \leq n$, если векторы e_i и e_j стоят в различных столбцах диаграммы, то и векторы $e_i\varphi^k$ и $e_j\varphi^k$ из тех же, т.е. из разных ее столбцов, а если векторы e_i и e_j из одного столбца, то их высоты различны, а потому различны высоты и векторов $e_j\varphi^k$ и $e_j\varphi^k$. Таким образом, система $e_{r+1}\varphi^k, e_{r+2}\varphi^k, \ldots, e_n\varphi^k$, будучи подсистемой базиса S пространства L, линейно независима. Отсюда $\alpha_{r+1} = \alpha_{r+2} = \cdots = \alpha_n = 0$ и потому

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_r e_r \in l(e_1, e_2, \ldots, e_r). \quad \square$$

Из предложения 6 следует, что если в пространстве существует жорданов базис для нильпотентного оператора, то вид диаграммы, которую он образует, (а потому и вид жордановой матрицы оператора) однозначно определяется этим оператором. Действительно, во–первых, отметим

Следствие 1. *Если S — расположенный в диаграмму жорданов базис для нильпотентного оператора φ пространства L, то максимальная высота столбцов диаграммы совпадает с показателем нильпотентности оператора φ.*

В самом деле, если максимальная высота столбцов диаграммы равна m, то ядро оператора φ^m порождается всеми элементами системы S и потому совпадает с L. Следовательно, $\varphi^m = 0$. С другой стороны, произвольный вектор a высоты m не выражается через подсистему системы S, состоящую из всех тех векторов из S, высота которых не превосходит числа $m-1$, и потому не входит в порождаемое этой подсистемой ядро оператора φ^{m-1}. Поэтому $a\varphi^{m-1} \neq 0$, т.е. $\varphi^{m-1} \neq 0$. $\square$

Следствие 2. *Если для нильпотентного оператора φ линейного пространства L существует жорданов базис, то вид диаграммы (т.е. число и высоты ее столбцов), образуемой этим базисом, однозначно определяется оператором φ. Поэтому число клеток Жордана и их порядки в жордановой матрице оператора φ определены однозначно.*

Доказательство. Если m — показатель нильпотентности оператора φ, то диаграмма, образованная векторами жорданова базиса для этого оператора, состоит из m строк. Пусть для $i = 1, 2, \dots m$ через s_i обозначено число векторов в i-ой строке диаграммы. Заметим сначала, что числа $s_1, s_2, \dots, s_m$ однозначно определяют вид диаграммы. Действительно, число всех столбцов диаграммы равно s_1. Кроме того, если $i < m$, то поскольку под каждым вектором высоты $i+1$ расположен вектор высоты i, имеем $s_{i+1} \leq s_i$ и разность $s_i - s_{i+1}$ совпадает с числом столбцов диаграммы, имеющих высоту i.

Из предложения 6 следует, далее, что числа $s_1, s_2, \dots, s_m$ должны удовлетворять следующей системе равенств

$$\begin{cases} s_1 & = \dim \operatorname{Ker} \varphi \\ s_1 + s_2 & = \dim \operatorname{Ker} \varphi^2 \\ s_1 + s_2 + s_3 & = \dim \operatorname{Ker} \varphi^3 \\ \dots\dots\dots\dots\dots\dots & \\ s_1 + s_2 + s_3 + \dots + s_m & = \dim \operatorname{Ker} \varphi^m. \end{cases}$$

Если ее рассматривать как систему линейных уравнений от неизвестных $s_1, s_2, \dots, s_m$, то поскольку ее матрица имеет вид

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix},$$

эта система уравнений является крамеровской и, следовательно, имеет единственное решение. Так как ее расширенная матрица однозначно определяется оператором φ, наше утверждение доказано. $\square$

Последнее следствие показывает, как найти жорданову матрицу произвольного нильпотентного оператора (если существование такой матрицы, т.е. жорданова базиса для нильпотентного оператора пока принять на веру). Поясним соответствующую процедуру на примере.

Пусть оператор φ линейного пространства L в некотором базисе этого пространства задан матрицей

$$A = \begin{pmatrix} 0 & 1 & -2 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Характеристический многочлен матрицы A совпадает с многочленом x^5, и потому (в силу теоремы Гамильтона – Кэли) оператор φ нильпотентен. Поэтому в соответствии с предложением 6 и доказательством следствия 2 из этого предложения для нахождения вида жордановой матрицы оператора φ достаточно найти показатель m его нильпотентности и размерности ядер операторов $\varphi, \varphi^2, \dots \varphi^m$. В действительности, в ряде случаев (как в этом примере) можно обойтись меньшим объемом вычислений.

Ранг оператора φ совпадает с рангом матрицы A и потому, очевидно, равен 3. Следовательно, дефект этого оператора, т. е. размерность его ядра, равен 2 (поскольку $\dim L = 5$). Значит, диаграмма, построенная из жорданова базиса для этого оператора, состоит из двух столбцов. Для вычисления размерности ядра оператора φ^2 найдем его матрицу, равную квадрату матрицы A:

$$A^2 = \begin{pmatrix} 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ранг этой матрицы равен 2 и потому $\dim \operatorname{Ker} \varphi^2 = 3$. Следовательно, в первых двух строках диаграммы находится 3 вектора, т.е. во второй строке — всего один вектор. Это означает, что один из столбцов диаграммы имеет высоту 1. Поэтому другой должен быть высоты 4, и строение диаграммы полностью описано. Тем самым найден и вид жордановой матрицы нашего оператора: она состоит из двух жордановых клеток порядков 4 и 1, т.е. равна матрице

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Для завершения доказательства теоремы нам понадобится еще одно свойство системы векторов, расположенной в диаграмму относительно действия нильпотентного оператора.

Предложение 7. *Пусть φ — нильпотентный оператор линейного пространства L и пусть некоторая система S векторов пространства L расположена в диаграмму относительно действия оператора φ. Если система векторов, стоящих в первой строке диаграммы, линейно независима, то и вся система S является линейно независимой.*

Доказательство. Предположим, напротив, что система $S = (e_1, e_2, \dots, e_n)$ является линейно зависимой и потому

$$\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n = 0$$

для некоторых скаляров $\alpha_1, \alpha_2, \dots, \alpha_n$, среди которых хотя бы один отличен от нуля. Пусть k — максимальная высота в диаграмме тех векторов системы S, коэффициент при которых отличен от нуля, и пусть $i_1, i_2, \dots, i_r$ — номера всех векторов этой системы, имеющих высоту k. Подействовав на обе части этого равенства оператором φ^{k-1}, получаем

$$\alpha_1(e_1\varphi^{k-1}) + \alpha_2(e_2\varphi^{k-1}) + \cdots + \alpha_n(e_n\varphi^{k-1}) = 0.$$

Заметим, далее, что если номер i не входит в множество номеров $\{i_1, i_2, \dots, i_r\}$, то $\alpha_i(e_i\varphi^{k-1}) = 0$. Действительно, в этом случае либо высота вектора e_i больше, чем k, и потому $\alpha_i = 0$, либо высота этого вектора не превосходит числа $k-1$ и потому $e_i\varphi^{k-1} = 0$. Поэтому последнее равенство имеет вид

$$\alpha_{i_1}(e_{i_1}\varphi^{k-1}) + \alpha_{i_2}(e_{i_2}\varphi^{k-1}) + \cdots + \alpha_{i_r}(e_{i_r}\varphi^{k-1}) = 0,$$

и так как все векторы $e_{i_1}\varphi^{k-1}, e_{i_2}\varphi^{k-1}, \dots, e_{i_r}\varphi^{k-1}$ расположены в первой строке диаграммы и среди коэффициентов $\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_r}$ хотя бы один отличен от нуля, приходим к противоречию с условием. $\square$

Нам будет удобно воспользоваться следующим обобщением понятия линейной независимости системы векторов:

Определение 3. *Пусть M — подпространство линейного пространства L. Система $a_1, a_2, \dots, a_r$ векторов пространства L называется линейно независимой над M, если произвольная линейная комбинация $\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_r a_r$ этой системы входит в подпространство M лишь в том случае, когда $\alpha_1 = \alpha_2 = \cdots = \alpha_r = 0$. Система $a_1, a_2, \dots, a_r$ называется максимальной линейно независимой над M, если она линейно независима над M и не содержится ни в одной, отличной от нее линейно независимой над M системе.*

Очевидно, что произвольная система векторов является линейно независимой тогда и только тогда, когда она линейно независима над нулевым подпространством. Следующее утверждение также достаточно очевидно.

Предложение 8. *Пусть M — подпространство линейного пространства L. Система $a_1, a_2, \ldots, a_r$ векторов пространства L является линейно независимой над M (максимальной линейно независимой над M) тогда и только тогда, когда для любого базиса $e_1, e_2, \ldots, e_m$ подпространства M система*

$$a_1, a_2, \ldots, a_r, e_1, e_2, \ldots, e_m$$

линейно независима (соответственно, является базисом пространства L). $\square$

Отметим, наконец,

Предложение 9. *Пусть φ — нильпотентный оператор с показателем нильпотентности m линейного пространства L. Для произвольного целого числа k, $1 \leq k \leq m$, полагаем $F_k = Ker\varphi^k$; кроме того, пусть F_0 обозначает нулевое подпространство. Если при $k \geq 2$ система $a_1, a_2, \ldots, a_r$ векторов подпространства F_k линейно независима над F_{k-1}, то система их φ-образов $a_1\varphi, a_2\varphi, \ldots, a_r\varphi$ (очевидно, принадлежащих подпространству) F_{k-1} линейно независима над F_{k-2}.*

В самом деле, если

$$\alpha_1(a_1\varphi) + \alpha_2(a_2\varphi) + \cdots + \alpha_r(a_r\varphi) \in F_{k-2},$$

то по определению

$$(\alpha_1(a_1\varphi) + \alpha_2(a_2\varphi) + \cdots + \alpha_r(a_r\varphi))\varphi^{k-2} = 0.$$

Так как

$$\begin{aligned}(\alpha_1(a_1\varphi) + \alpha_2(a_2\varphi) + \cdots + \alpha_r(a_r\varphi))\varphi^{k-2} &= ((\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_r a_r)\varphi)\varphi^{k-2} \\ &= (\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_r a_r)\varphi^{k-1},\end{aligned}$$

мы видим, что $\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_r a_r \in F_{k-1}$, и потому $\alpha_1 = \alpha_2 = \cdots = \alpha_r = 0$. $\square$

Мы можем, наконец, приступить непосредственно к доказательству существования жорданова базиса для нильпотентного оператора. Сейчас будет описана по шагам процедура построения некоторой системы векторов нашего пространства. При этом из самого построения будет видно, что полученная система векторов располагаема в диаграмму. По окончании процедуры станет ясно, что построенная система является базисом нашего пространства.

Итак, пусть φ — нильпотентный оператор показателя нильпотентности m линейного пространства L. Пусть, как и в предложении 9, $F_k = \operatorname{Ker}\varphi^k$ $(1 \leq k \leq m)$; мы будем предполагать вычисленными базисы подпространств F_k. Заметим, что если вектор $a \in L$ принадлежит подпространству F_k, но не принадлежит подпространству F_{k-1}, то $a\varphi^k = 0$, а векторы $a, a\varphi, \ldots, a\varphi^{k-1}$ отличны от нуля и потому составляют диаграмму – столбец высоты k.

1-ый шаг. В пространстве $L = F_m$ выбираем максимальную линейно независимую над F_{m-1} систему векторов

$$a_{11}, a_{12}, \ldots, a_{1k_1}. \qquad (1)$$

(В соответствии с предложением 8 в систему (1) можно внести векторы, дополняющие базис подпространства F_{m-1} до базиса пространства $L = F_m$.) Каждый из векторов системы (1) служит вершиной столбца высоты m, и мы вносим все векторы этих столбцов в будущий базис. Таким образом, часть диаграммы, построенная на первом шаге, имеет вид

$$\begin{array}{cccc}
a_{11} & a_{12} & \ldots & a_{1k_1} \\
\downarrow & \downarrow & \ldots & \downarrow \\
a_{11}\varphi & a_{12}\varphi & \ldots & a_{1k_1}\varphi \\
\downarrow & \downarrow & \ldots & \downarrow \\
a_{11}\varphi^2 & a_{12}\varphi^2 & \ldots & a_{1k_1}\varphi^2 \\
\downarrow & \downarrow & \ldots & \downarrow \\
\ldots & \ldots & \ldots & \ldots \\
\downarrow & \downarrow & \ldots & \downarrow \\
a_{11}\varphi^{m-2} & a_{12}\varphi^{m-2} & \ldots & a_{1k_1}\varphi^{m-2} \\
\downarrow & \downarrow & \ldots & \downarrow \\
a_{11}\varphi^{m-1} & a_{12}\varphi^{m-1} & \ldots & a_{1k_1}\varphi^{m-1}
\end{array}$$

Так как векторы системы (1) вместе с базисом подпространства F_{m-1} составляют базис пространства L, очевидно, что векторы, внесенные в диаграмму на первом шаге, вместе с подпространством F_{m-1} порождают пространство L.

2-ой шаг. В силу предложения 9 система φ-образов векторов системы (1) линейно независима над F_{m-2}. Дополняем ее до максимальной линейно независимой над F_{m-2} системы

$$a_{11}\varphi, a_{12}\varphi, \ldots, a_{1k_1}\varphi, a_{21}, a_{22}, \ldots, a_{2k_2} \qquad (2)$$

векторов пространства F_{m-1}. Векторы $a_{21}, a_{22}, \ldots, a_{2k_2}$ объявляем вершинами столбцов высоты $m-1$ и вносим в диаграмму все векторы этих столбцов. Диаграмма к концу второго шага принимает вид

$$\begin{array}{cccccccc}
a_{11} & a_{12} & \ldots & a_{1k_1} & & & & \\
\downarrow & \downarrow & \ldots & \downarrow & & & & \\
a_{11}\varphi & a_{12}\varphi & \ldots & a_{1k_1}\varphi & a_{21} & a_{22} & \ldots & a_{2k_2} \\
\downarrow & \downarrow & \ldots & \downarrow & \downarrow & \downarrow & \ldots & \downarrow \\
a_{11}\varphi^2 & a_{12}\varphi^2 & \ldots & a_{1k_1}\varphi^2 & a_{21}\varphi & a_{22}\varphi & \ldots & a_{2k_2}\varphi \\
\downarrow & \downarrow & \ldots & \downarrow & \downarrow & \downarrow & \ldots & \downarrow \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\downarrow & \downarrow & \ldots & \downarrow & \downarrow & \downarrow & \ldots & \downarrow \\
a_{11}\varphi^{m-2} & a_{12}\varphi^{m-2} & \ldots & a_{1k_1}\varphi^{m-2} & a_{21}\varphi^{m-3} & a_{22}\varphi^{m-3} & \ldots & a_{2k_2}\varphi^{m-3} \\
\downarrow & \downarrow & \ldots & \downarrow & \downarrow & \downarrow & \ldots & \downarrow \\
a_{11}\varphi^{m-1} & a_{12}\varphi^{m-1} & \ldots & a_{1k_1}\varphi^{m-1} & a_{21}\varphi^{m-2} & a_{22}\varphi^{m-2} & \ldots & a_{2k_2}\varphi^{m-2}
\end{array}$$

Векторы, внесенные в диаграмму в результате выполнения первого и второго шагов, вместе с векторами подпространства F_{m-2} порождают пространство L.

Может случиться, что на втором шаге новых столбцов в диаграмме не появится. Это произойдет, если φ-образы системы (1) вместе с базисом подпространства F_{m-2} уже составляют базис подпространства F_{m-1}. А означает это, что будущая диаграмма не содержит столбцов высоты $m-1$.

На третьем шаге систему φ-образов системы (2) дополняем до максимальной линейно независимой над F_{m-3} системы

$$a_{11}\varphi^2, a_{12}\varphi^2, \ldots, a_{1k_1}\varphi^2, a_{21}\varphi, a_{22}\varphi, \ldots, a_{2k_2}\varphi, a_{31}, a_{32}, \ldots, a_{3k_3} \qquad (3)$$

векторов пространства F_{m-2} и вносим в диаграмму векторы столбцов высоты $m-2$ с вершинами $a_{31}, a_{32}, \ldots, a_{3k_3}$.

Продолжая аналогично, на $(m-1)$-ом шаге мы внесем в диаграмму векторы столбцов высоты 2 (или убедимся в том, что таких столбцов диаграмма не содержит). Векторы, внесенные в диаграмму к концу этого шага вместе с подпространством F_1 порождают все пространство L, а векторы, занимающие первую строку этой диаграммы, составляют линейно независимую систему векторов подпространства F_1. Дополнив ее до базиса этого подпространства и включив дополняющие векторы $a_{m1}, a_{m2}, \ldots, a_{mk_m}$ в диаграмму в качестве столбцов высоты 1, мы завершим построение.

Итак, мы построили некоторую систему векторов пространства L, расположенную относительно действия оператора φ в диаграмму следующего вида:

$$\begin{array}{ccccccccccc}
a_{11} & \ldots & a_{1k_1} & & & & & & & & \\
\downarrow & \ldots & \downarrow & & & & & & & & \\
a_{11}\varphi & \ldots & a_{1k_1}\varphi & a_{21} & \ldots & a_{2k_2} & & & & & \\
\downarrow & \ldots & \downarrow & \downarrow & \ldots & \downarrow & & & & & \\
a_{11}\varphi^2 & \ldots & a_{1k_1}\varphi^2 & a_{21}\varphi & \ldots & a_{2k_2}\varphi & & & & & \\
\downarrow & \ldots & \downarrow & \downarrow & \ldots & \downarrow & & & & & \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & & & & & \\
\downarrow & \ldots & \downarrow & \downarrow & \ldots & \downarrow & & & & & \\
a_{11}\varphi^{m-2} & \ldots & a_{1k_1}\varphi^{m-2} & a_{21}\varphi^{m-3} & \ldots & a_{2k_2}\varphi^{m-3} & & & & & \\
\downarrow & \ldots & \downarrow & \downarrow & \ldots & \downarrow & & & & & \\
a_{11}\varphi^{m-1} & \ldots & a_{1k_1}\varphi^{m-1} & a_{21}\varphi^{m-2} & \ldots & a_{2k_2}\varphi^{m-2} & \ldots & a_{m1} & \ldots & a_{mk_m}. &
\end{array}$$

Остается понять, что построенная система является базисом этого пространства. В самом деле, как только что было отмечено, каждый вектор пространства L линейно выражается через векторы, внесенные в диаграмму к концу $(m-1)$-го шага, и векторы подпространства F_1. А после m-го шага диаграмма приобрела подсистему, являющуюся базисом этого подпространства, и мы теперь можем выразить произвольный вектор из L через векторы этой диаграммы. Следовательно, векторы диаграммы составляют систему порождающих пространства L. Линейная независимость этой системы вытекает из предложения 7, так как система векторов, расположенных в первой строке, линейно независима по построению.

Проиллюстрируем описанную процедуру построения жорданова базиса на примере того оператора, для которого выше была найдена жорданова матрица. Для этого необходимо уточнить ситуацию: Будем считать, что φ — оператор линейного пространства $\mathbb{R}^5$, заданный матрицей

$$A = \begin{pmatrix} 0 & 1 & -2 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

в стандартном базисе этого пространства. Прежде, чем приступать непосредственно к процедуре построения жорданова базиса, необходимо найти базисы ядер степеней оператора φ. (Так как нам уже известна жорданова матрица этого конкретного оператора, мы могли бы вычислять базисы не всех ядер; тем не менее, будем поступать так, как это необходимо в общей ситуации.) Непосредственные вычисления показывают, что

базис подпространства $F_1 = \operatorname{Ker}\varphi$ состоит из векторов $(0,0,0,1,0)$ и $(0,0,0,0,1)$;

базис подпространства $F_2 = \operatorname{Ker}\varphi^2$ состоит из векторов $(0,0,1,0,0)$, $(0,0,0,1,0)$ и $(0,0,0,0,1)$;

базис подпространства $F_3 = \operatorname{Ker}\varphi^3$ состоит из векторов $(0,1,0,0,0)$, $(0,0,1,0,0)$, $(0,0,0,1,0)$ и $(0,0,0,0,1)$.

Подпространство $F_4 = \operatorname{Ker}\varphi^4$ совпадает со всем пространством $\mathbb{R}^5$.

Первый шаг построения начинается с выбора векторов, дополняющих базис подпространства F_3 до базиса всего пространства; в данном случае в качестве такого вектора может быть, очевидно, выбран вектор $a = (1,0,0,0,0)$. Он и будет вершиной столбца высоты 4. Остальными векторами этого столбца являются векторы $a\varphi = (0,1,-2,1,1)$, $a\varphi^2 = (0,0,1,-1,-2)$, $a\varphi^3 = (0,0,0,1,1)$.

На втором шаге следует присоединить вектор $a\varphi$ к базису подпространства F_2 и найти векторы, дополняющие эту (линейно независимую) систему до базиса подпространства F_3. В данном примере, присоединив вектор $a\varphi$ к базису подпространства F_2, мы уже получаем базис подпространства F_3 (поскольку число векторов в полученной системе совпадает с его размерностью). Поэтому никаких дополнительных векторов здесь не появляется, и это означает, что в диаграмме нет столбцов высоты 3. Точно так же обнаруживается, что в диаграмме нет столбцов высоты 2, поскольку вектор $a\varphi^2$ вместе с базисом подпространства F_1 образует базис подпространства F_2.

На последнем шаге следует единственный вектор $a\varphi^3 = (0,0,0,1,1)$, оказавшийся в первой строке, дополнить до базиса ядра оператора φ. В данном случае можно опять обойтись без вычислений: к этому вектору можно присоединить любой из составляющих вычисленный выше базис подпространства F_1.

Итак, следующая система векторов составляет жорданов базис для оператора φ пространства $\mathbb{R}^5$:

$$a_1 = (1,0,0,0,0), a_2 = (0,1,-2,1,1), a_3 = (0,0,1,-1,-2), a_4 = (0,0,0,1,1),$$
$$a_5 = (0,0,0,1,0).$$

ГЛАВА III

ЛИНЕЙНЫЕ ПРОСТРАНСТВА СО СКАЛЯРНЫМ ПРОИЗВЕДЕНИЕМ

§ 1. БИЛИНЕЙНЫЕ ФУНКЦИИ И ИХ МАТРИЦЫ

Определение 1. *Пусть L — линейное пространство над полем K. Функция $f(x, y)$ от двух переменных, определенная на множестве векторов пространства L и принимающая значения в поле K, называется билинейной, если она линейна по каждому аргументу, т.е. для любых векторов $a, b, c \in L$ и скаляра $\lambda \in K$ выполнены следующие равенства:*

(1) $f(a + b, c) = f(a, c) + f(b, c)$;
(2) $f(\lambda a, c) = \lambda f(a, c)$;
(3) $f(a, b + c) = f(a, b) + f(a, c)$;
(4) $f(a, \lambda c) = \lambda f(a, c)$.

Два первых равенства означают линейность по первому аргументу, а два последних — линейность по второму. Фиксируя один из аргументов в билинейной функции, получаем, очевидно, линейное отображение пространства L в поле K, т.е. — линейный функционал на пространстве L.

Рассмотрим несколько примеров.

1. На пространстве V^2 направленных отрезков плоскости обычное скалярное произведение $f(a, b) = |a| \cdot |b| \cdot \cos(\angle(a, b))$ является билинейной функцией.

2. Пусть $L = K^n$ — n-мерное координатное пространство над полем K. Полагая для векторов $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ и $b = (\beta_1, \beta_2, \ldots, \beta_n)$ этого пространства

$$f(a, b) = \sum_{i=1}^{n} \alpha_i \beta_i,$$

получаем билинейную функцию на L.

3. Пусть $C_{[\alpha,\beta]}$ — пространство действительных функций от одной переменной, непрерывных на отрезке $[\alpha, \beta]$. Функция

$$f(\varphi, \psi) = \int_{\alpha}^{\beta} \varphi(x)\psi(x)\, dx \qquad (\varphi(x), \psi(x) \in C_{[\alpha,\beta]})$$

является билинейной.

Отметим еще, что функция $f(x, y)$, определенная на произвольном линейном пространстве L так, что $f(a, b) = 0$ для любых $a, b \in L$, очевидно, является билинейной; будем называть ее нулевой билинейной функцией.

Перечислим теперь простейшие свойства билинейных функций.

Предложение 1. *Пусть $f(x,y)$ — билинейная функция на линейном пространстве L над полем K. Тогда*

а) для любых векторов $a,b\in L$

$$f(a,o)=f(0,b)=0 \quad \text{и} \quad f(-a,b)=fa,-b)=-f(a,b);$$

б) для любых векторов $a_1,a_2,\ldots,a_m,b_1,b_2,\ldots,b_n\in L$ и любых скаляров $\alpha_1,\alpha_2,\ldots,\alpha_m,\beta_1,\beta_2,\ldots,\beta_n\in K$

$$f(\sum_{i=1}^{m}\alpha_i a_i,\ \sum_{j=1}^{n}\beta_j b_j)=\sum_{i=1}^{m}\sum_{j=1}^{n}\alpha_i\beta_j f(a_i,b_j).$$

Оба утверждения непосредственно получаются из определения билинейной функции. Приведем, тем не менее, подробные вычисления для обоснования второго из них. Последовательно используя линейность по каждому аргументу и тот факт, что при линейном отображении образ линейной комбинации равен линейной комбинации образов, получаем

$$f(\sum_{i=1}^{m}\alpha_i a_i,\ \sum_{j=1}^{n}\beta_j b_j)=\sum_{i=1}^{m}\alpha_i f(a_i,\ \sum_{j=1}^{n}\beta_j b_j)=\sum_{i=1}^{m}\alpha_i\sum_{j=1}^{n}\beta_j f(a_i,b_j)=$$
$$\sum_{i=1}^{m}\sum_{j=1}^{n}\alpha_i\beta_j f(a_i,b_j) \quad \square$$

Определение 2. *Пусть $f(x,y)$ — билинейная функция на конечномерном линейном пространстве L и $S=(e_1,e_2,\ldots,e_n)$ — базис этого пространства. Матрицей функции $f(x,y)$ в базисе S называется квадратная матрица $A=(\alpha_{ij})$ порядка n, где $\alpha_{ij}=f(e_i,e_j)$ $(i,j=1,2,\ldots,n)$.*

Таким образом, в более подробной записи матрица функции $f(x,y)$ имеет вид

$$\begin{pmatrix} f(e_1,e_1) & f(e_1,e_2) & \ldots & f(e_1,e_n)\\ f(e_2,e_1) & f(e_2,e_2) & \ldots & f(e_2,e_n)\\ \ldots & \ldots & \ldots & \ldots\\ f(e_n,e_1) & f(e_n,e_2) & \ldots & f(e_n,e_n)\end{pmatrix}.$$

Следующее утверждение показывает, в частности, что билинейная функция своей матрицей однозначно определяется.

Предложение 2. *Пусть $f(x,y)$ — билинейная функция на конечномерном линейном пространстве L над полем K и A — матрица функции $f(x,y)$ в базисе S этого пространства. Тогда для любых векторов $a,b\in L$*

$$f(a,b)=(a)_S\cdot A\cdot((b)_S)^t. \tag{1}$$

Наоборот, если A — произвольная квадратная матрица над полем K, порядок которой совпадает с размерностью пространства L, то функция $f(x,y)$, определяемая формулой (1), билинейна, а ее матрица в базисе S совпадает с матрицей A.

(Следует отметить определенную условность в записи равенства (1). В левой части его расположен элемент поля K. В правой же части — матрица первого порядка, равная произведению матрицы – строки $(a)_S$, матрицы A и матрицы – столбца $((b)_S)^t$ (транспонированной строки $(b)_S$). Таким образом, условность состоит в отождествлении матрицы первого порядка с ее элементом.)

Доказательство. Введем более детальные обозначения. Пусть $S = (e_1, e_2, \dots, e_n)$ и $A = (\alpha_{ij})$; таким образом $\alpha_{ij} = f(e_i, e_j)$. Пусть, далее, $(a)_S = (\alpha_1, \alpha_2, \dots, \alpha_n)$ и $(b)_S = (\beta_1, \beta_2, \dots, \beta_n)$ — координатные строки векторов a и b. Это означает, что

$$a = \sum_{i=1}^{n} \alpha_i e_i \quad \text{и} \quad b = \sum_{i=1}^{n} \beta_i e_i,$$

и потому

$$f(a,b) = f(\sum_{i=1}^{n} \alpha_i e_i,\ \sum_{j=1}^{n} \beta_j e_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j f(e_i, e_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j \alpha_{ij}.$$

С другой стороны,

$$(a)_S A = (\sum_{i=1}^{n} \alpha_i \alpha_{i1}, \sum_{i=1}^{n} \alpha_i \alpha_{i2}, \cdots \sum_{i=1}^{n} \alpha_i \alpha_{in}),$$

и потому

$$(a)_S A((b)_S)^t = (\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j \alpha_{ij}).$$

Таким образом, первое утверждение доказано.

Обратно, если функция $f(x,y)$ определена для произвольной квадратной матрицы $A = (\alpha_{ij})$ порядка n равенством (1), то билинейность этой функции является следствием свойств матричных операций. Так, из дистрибутивного закона для матриц следует, что для любых векторов $a, b, c \in L$

$$f(a+b,c) = (a+b)_S A(c)_S^t = ((a)_S + (b)_S) A(c)_S^t = (a)_S A(c)_S^t + (b)_S A(c)_S^t = f(a,c) + f(b,c).$$

Кроме того, для $\lambda \in K$ имеем

$$f(\lambda a, c) = (\lambda a)_S A(c)_S^t = (\lambda (a)_S) A(c)_S^t = \lambda((a)_S A(c)_S^t) = \lambda f(a,c).$$

Аналогично доказывается линейность по второму аргументу. Элемент, стоящий в i-ой строке и j-ом столбце матрицы этой функции, вычисленной в базисе S, равен

$$f(e_i, e_j) = (e_i)_S A (e_j)_S^t,$$

откуда с учетом вида координатных строк векторов e_i и e_j получаем $f(e_i, e_j) = \alpha_{ij}$, так что искомая матрица совпадает с A. $\square$

Выясним теперь, как изменяется матрица билинейной функции при изменении базиса пространства.

Пусть S и S' — два базиса пространства L, A и A' — матрицы билинейной функции $f(x, y)$ в базисах S и S' соответственно. Тогда в силу предложения 2 для произвольных векторов $a, b \in L$ имеем

$$f(a, b) = (a)_S A (b)_S^t.$$

Если T — матрица перехода от базиса S к базису S', то $(a)_S = (a)_{S'} T$ и $(b)_S = (b)_{S'} T$, откуда получаем

$$f(a, b) = (a)_S A (b)_S^t = ((a)_{S'} T) A ((b)_{S'} T)^t = ((a)_{S'} T) A (T^t (b)_{S'}^t) = (a)_{S'} (T A T^t)(b)_{S'}^t.$$

Из второй части предложения 2 теперь следует, что матрица A' функции $f(x, y)$ в базисе S' должна совпадать с $T \cdot A \cdot T^t$. Таким образом, мы доказали

Предложение 3. *Пусть $f(x, y)$ — билинейная функция на линейном пространстве L, S и S' — два базиса пространства L, и A и A' — матрицы функции $f(x, y)$ в базисах S и S' соответственно. Тогда*

$$A' = T A T^t,$$

где T — матрица перехода от базиса S к базису S'. $\square$

Рассматривая линейное пространство вместе с фиксированной определенной на нем билинейной функцией, мы получаем новый класс математических систем, свойства которых определяются не только структурой линейного пространства, но и билинейной функцией. Поэтому следует договориться о том, какие линейные пространства с билинейными функциями мы будем считать одинаковыми. Напомним, что для линейных пространств эту задачу решает понятие изоморфизма. Следующее определение дает адекватное уточнение интуитивного представления того, когда такие системы следует считать одинаковыми.

Определение 4. *Пусть L_1 и L_2 — линейные пространства над (одним и тем же) полем K с билинейными функциями $f_1(x, y)$ и $f_2(x, y)$ соответственно. Отображение $\varphi : L_1 \to L_2$ множества L_1 в множество L_2 называется изометрией, если φ — изоморфизм линейного пространства L_1 на линейное пространство L_2 (т.е. является линейным и биективным) и для любых векторов $a, b \in L_1$ выполнено равенство*

$$f_1(a, b) = f_2(a\varphi, b\varphi).$$

Пространство L_1 *с билинейной функцией* $f_1(x, y)$ *называется изометричным пространству* L_2 *с билинейной функцией* $f_2(x, y)$ *, если существует изометрия пространства* L_1 *на пространство* L_2.

Очевидно, что тождественное отображение множества векторов линейного пространства является изометрией, и потому отношение изометричности линейных пространств рефлексивно. Легко проверить также, что произведение двух изометрий и отображение, обратное к изометрии, являются изометриями. Поэтому отношение изометричности линейных пространств транзитивно и симметрично.

Задача классификации линейных пространств с билинейными функциями состоит в нахождении условий, необходимых и достаточных для их изометричности. В третьем параграфе эта задача будет решена для линейных пространств над полем действительных чисел с симметричными билинейными функциями.

§ 2. СИММЕТРИЧНЫЕ БИЛИНЕЙНЫЕ ФУНКЦИИ. ОРТОГОНАЛЬНОСТЬ

Определение 1. *Билинейная функция $f(x, y)$ на линейном пространстве L называется симметричной, если для любых векторов $a, b \in L$ выполняется равенство $f(a, b) = f(b, a)$.*

Из этого определения и определения матрицы билинейной функции получаем почти очевидное

Предложение 1. *Билинейная функция $f(x, y)$ на линейном пространстве L является симметричной тогда и только тогда, когда матрица этой функции (в произвольном базисе пространства L) симметрична.*

В самом деле, необходимость условия просто очевидна. Предположим, с другой стороны, что матрица A функции $f(x, y)$, вычисленная в некотором базисе S пространства L симметрична. Это означает, что $A = A^t$, и потому, используя предложение 2 (и симметричность матрицы порядка 1), получаем для произвольных $a, b \in L$

$$f(a, b) = (a)_S A(b)_S^t = ((a)_S A(b)_S^t)^t = ((b)_S^t)^t A^t (a)_S^t = (b)_S A(a)_S^t = f(b, a). \quad \square$$

Определение 2. *Пусть L — линейное пространство с симметричной билинейной функцией $f(x, y)$. Вектор $a \in L$ будем называть ортогональным вектору $b \in L$ (и записывать это в виде $a \perp b$), если $f(a, b) = 0$. Подмножество $M \subseteq L$ назовем ортогональным подмножеству $N \subseteq L$ (и будем писать $M \perp N$), если каждый вектор $a \in M$ ортогонален каждому вектору $b \in N$. Наконец, ортогональным дополнением подмножества $M \subseteq L$ будем называть совокупность $M^\perp$ всех векторов пространства L, ортогональных каждому вектору из M.*

Заметим, что симметричность функции $f(x, y)$ обеспечивает симметричность отношения ортогональности векторов пространства L. Достаточно очевидные свойства этого отношения содержатся в следующем утверждении:

Предложение 2. *Ортогональное дополнение произвольного подмножества $M \subseteq L$ является подпространством пространства L. Кроме того, $M^\perp = l(M)^\perp$.*

Доказательство. Начнем с того, что произвольный вектор a, ортогональный каждому вектору системы $b_1, b_2, \ldots, b_n$, будет ортогонален и любой линейной комбинации этой системы: если для любого $i = 1, 2, \ldots, n$ $f(a, b_i)$, то для произвольных скаляров $\alpha_1, \alpha_2, \ldots, \alpha_n$ имеем

$$f(a, \sum_{i=1}^{n} \alpha_i b_i) = \sum_{i=1}^{n} \alpha_i f(a, b_i) = 0.$$

Поэтому множество $M^\perp$ замкнуто относительно взятия линейных комбинаций и, очевидно, непусто ($0 \in M^\perp$), так что является подпространством пространства L.

Из сделанного замечания следует также, что если $a \in M^\perp$, т.е. вектор a ортогонален каждому вектору из множества M, то a является ортогональным и каждому вектору из подпространства $l(M)$, т.е. $a \in l(M)^\perp$. Тем самым доказано включение $M^\perp \subseteq l(M)^\perp$. Так как $M \subseteq l(M)$, противоположное включение очевидно. $\square$

Определение 3. *Система векторов $a_1, a_2, \ldots, a_n$ линейного пространства L с симметричной билинейной функцией $f(x, y)$ называется ортогональной, если $a_i \perp a_j$ для любых $i \neq j$. Базис пространства L называется ортогональным, если он является ортогональной системой.*

Легко видеть, что существование ортогонального базиса в линейном пространстве с симметричной билинейной функцией позволяет сделать предельно простым вид матрицы этой функции; тем самым упрощается и вычисление значений функции:

Предложение 3. *Базис $S = (e_1, e_2, \ldots, e_n)$ линейного пространства L с симметричной билинейной функцией $f(x, y)$ является ортогональным тогда и только тогда, когда матрица функции $f(x, y)$ в базисе S диагональна.*

Если матрица функции $f(x, y)$ в базисе S имеет вид

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

и $(a)_S = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ и $(b)_S = (\beta_1, \beta_2, \ldots, \beta_n)$ — координатные строки в базисе S векторов a и b пространства L, то

$$f(a, b) = \sum_{i=1}^{n} \lambda_i \alpha_i \beta_i.$$

В самом деле, первое утверждение является непосредственным следствием определений матрицы билинейной функции и ортогонального базиса. Поскольку

$$a = \sum_{i=1}^{n} \alpha_i e_i \quad \text{и} \quad b = \sum_{i=1}^{n} \beta_i e_i$$

и

$$f(e_i, e_j) = \begin{cases} \lambda_i, & \text{если } i = j \\ 0, & \text{если } i \neq j, \end{cases}$$

по формуле из предложения 1 имеем

$$f(\sum_{i=1}^{n} \alpha_i e_i, \sum_{j=1}^{n} \beta_j e_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j f(e_i, e_j) = \sum_{i=1}^{n} \alpha_i \beta_i \lambda_i. \quad \square$$

Обсуждение проблемы существования ортогонального базиса начнем с рассмотрения примера.

Пусть $L = \mathbb{Z}_2^2$ — 2-мерное координатное пространство над полем $\mathbb{Z}_2$ классов вычетов целых чисел по модулю 2. Так как поле $\mathbb{Z}_2$ состоит из двух элементов 0 и 1, в пространстве L имеется в точности 4 вектора:

$$0 = (0,0), e_1 = (1,0), e_2 = (0,1), e_3 = (1,1).$$

Рассмотрим на пространстве L билинейную функцию $f(x,y)$, заданную в его стандартном базисе $S = (e_1, e_2)$ матрицей $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Это означает, что

$$f(e_1,e_1) = 0, f(e_2,e_2) = 0, f(e_1,e_2) = f(e_2,e_1) = 1.$$

Отсюда

$$\begin{aligned} f(e_1,e_3) &= f(e_1, e_1+e_2) = f(e_1,e_1) + f(e_1,e_2) = 1, \\ f(e_2,e_3) &= f(e_2, e_1+e_2) = f(e_2,e_1) + f(e_2,e_2) = 1, \\ f(e_3,e_3) &= f(e_1+e_2, e_3) = f(e_1,e_3) + f(e_2,e_3) = 1+1 = 0. \end{aligned}$$

Из этих вычислений и симметричности функции $f(x,y)$ следует, что два ненулевых вектора пространства L составляют ортогональную систему тогда и только тогда, когда они равны. Таким образом, в пространстве L ортогонального базиса не существует.

Сейчас мы увидим, что настоящая (и единственная) причина того, что на пространстве $L = \mathbb{Z}_2^2$ можно указать симметричную билинейную функцию, не допускающую ортогонального базиса, состоит в том, что характеристика основного поля этого пространства равна 2.

Теорема. *Пусть L — конечномерное линейное пространство над полем K с симметричной билинейной функцией $f(x,y)$. Если характеристика поля K не равна 2, то в пространстве L существует ортогональный базис.*

Доказательство. Проведем индукцию по числу $n = \dim L$. При $n = 1$ справедливость нашего утверждения не требует доказательства, так как в произвольная система векторов, состоящая из одного вектора, в соответствии с определением должна считаться ортогональной.

Пусть $n > 1$ и пусть для всех пространств меньшей размерности утверждение теоремы справедливо. Докажем, что тогда утверждение теоремы выполнено и для пространства L. Если функция $f(x,y)$ является нулевой (т.е. $f(a,b) = 0$ для любых $a, b \in L$), то доказывать снова нечего, так как в этом случае любая система векторов ортогональна. Будем считать функцию $f(x,y)$ ненулевой и докажем, что в этом случае найдется вектор $a \in L$ такой, что $f(a,a) \neq 0$.

Предположим, напротив, что для любого вектора $a \in L$ выполняется равенство $f(a,b) = 0$. Тогда для произвольных векторов $a, b \in L$ имеем

$$f(a+b, a+b) = 0,$$

и так как
$$f(a+b,a+b)=f(a,a)+f(a,b)+f(b,a)+f(b,b),$$
$f(a,a)=f(b,b)=0$ и $f(a,b)=f(b,a)$, получаем равенство $2f(a,b)=0$. Так как характеристика поля отлична от 2, отсюда следует, что для любых $a,b\in L$ $f(a,b)=0$, а это противоречит нашему предположению о том, что функция $f(x,y)$ ненулевая.

Итак, выберем в качестве e_1 такой вектор пространства L, что $f(e_1,e_1)\neq 0$, и пусть $M=l(e_1)$ — подпространство пространства L, порожденное вектором e_1. Отметим, что поскольку $e_1\neq 0$, $\dim M=1$. Покажем, что пространство L является прямой суммой подпространства M и его ортогонального дополнения $M^{\perp}$.

Для произвольного вектора $a\in L$ полагаем
$$\alpha=\frac{f(a,e_1)}{f(e_1,e_1)}\quad \text{и}\quad b=a-\alpha e_1.$$

Тогда
$$f(b,e_1)=f(a-\alpha e_1,e_1)=f(a,e_1)-\alpha f(e_1,e_1)=0,$$
т.е. $b\perp e_1$, и потому из предложения 2 следует, что $b\in M^{\perp}$.

Таким образом, произвольный вектор $a\in L$ допускает разложение вида $a=\alpha e_1+b$, где $\alpha e_1\in M$ и $b\in M^{\perp}$. Это означает, что $L=M+M^{\perp}$. Предположим теперь, что $a\in M\cap M^{\perp}$. Тогда поскольку вектор a входит в M, он имеет вид $a=\beta e_1$ для некоторого $\beta\in K$, а поскольку $a\in M^{\perp}$, должно выполняться равенство $f(a,e_1)=0$. С учетом вида вектора a это равенство равносильно равенству $\alpha f(e_1,e_1)=0$. Так как $f(e_1,e_1)\neq 0$, отсюда следует $\alpha=0$, а потому и $a=0$. Итак, пересечение подпространств M и $M^{\perp}$ оказалось нулевым, и потому $L=M\oplus M^{\perp}$.

Из доказанного следует, в частности, что $\dim M^{\perp}=n-1$. Кроме того, $M^{\perp}$ можно считать пространством с билинейной функцией $f(x,y)$ (точнее говоря, — с ограничением функции $f(x,y)$ на подпространство $M^{\perp}$). Поэтому из индуктивного предположения следует, что в пространстве $M^{\perp}$ найдется ортогональный базис $e_2,\dots e_n$. Очевидно тогда, что система $e_1,e_2,\dots e_n$ является базисом пространства L и притом ортогональным, так как векторы $e_2,\dots e_n$ принадлежат ортогональному дополнению вектора e_1. Индуктивный шаг закончен. $\square$

§ 3. ДЕЙСТВИТЕЛЬНЫЕ ЛИНЕЙНЫЕ ПРОСТРАНСТВА С СИММЕТРИЧНОЙ БИЛИНЕЙНОЙ ФУНКЦИЕЙ

Действительным линейным пространством будем называть линейное пространство над полем $\mathbb{R}$ действительных чисел. Поле $\mathbb{R}$ является упорядоченным, и это позволяет говорить о таких свойствах билинейных функций на действительных пространствах, которые не имеют смысла в общем случае и которые приближают общее понятие билинейной функции к понятию скалярного произведения, знакомому нам по элементарной геометрии.

Определение 1. *Пусть L — действительное линейное пространство с симметричной билинейной функцией $f(x, y)$. Вектор $a \in L$ называется*
положительным, если $f(a, a) > 0$,
отрицательным, если $f(a, a) < 0$, и
изотропным, если $f(a, a) = 0$.

Определение 2. *Симметричная билинейная функция $f(x, y)$ на действительном линейном пространстве L называется положительно (отрицательно) определенной, если для любого ненулевого вектора $a \in L$ $f(a, a) > 0$ (соответственно, $f(a, a) < 0$).*

Функция $f(x, y)$ называется положительно (отрицательно) полуопределенной, если для любого вектора $a \in L$ $f(a, a) \geq 0$ (соответственно, $f(a, a) \leq 0$).

Так как характеристика поля $\mathbb{R}$ действительных чисел равна 0, из теоремы, доказанной в предыдущем параграфе следует, что в произвольном конечномерном действительном пространстве с симметричной билинейной функцией существует ортогональный базис. По виду векторов такого базиса определенность и полуопределенность функции легко распознаются:

Предложение 1. *Пусть $S = (e_1, e_2, \ldots, e_n)$ — ортогональный базис действительного линейного пространства L с симметричной билинейной функцией $f(x, y)$. Тогда*

а) функция $f(x, y)$ положительно (отрицательно) определена тогда и только тогда, когда каждый вектор базиса S положителен (соответственно, отрицателен);

б) функция $f(x, y)$ положительно (отрицательно) полуопределена тогда и только тогда, когда каждый вектор из базиса S либо положителен (соответственно, отрицателен), либо изотропен.

В самом деле, необходимость условий в каждом случае очевидна. Достаточность их видна из того, что если $(a)_S = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ — координатная строка в базисе S вектора a пространства L, то в соответствии с предложением 2.3

$$f(a, a) = \sum_{i=1}^{n} \lambda_i \alpha_i^2,$$

где $\lambda_i = f(e_i, e_i)$ $(i = 1, 2, \ldots, n)$. $\square$

В конечномерном действительном пространстве можно выбрать ортогональный базис, удовлетворяющий дополнительному требованию:

Предложение (и определение) 2. *В произвольном конечномерном действительном пространстве L с симметричной билинейной функцией $f(x, y)$ существует ортогональный базис $S = (e_1, e_2, \dots, e_n)$, для любого вектора e_i которого значение $f(e_i, e_i)$ равно либо 1, либо -1, либо 0. Такой базис называется ортонормированным.*

Действительно, если $a_1, a_2, \dots, a_n$ — ортогональный базис пространства L, то полагая для каждого $i = 1, 2, \dots, n$

$$e_i = \begin{cases} \frac{1}{\sqrt{|f(a_i,a_i)|}} a_i & \text{если } f(a_i, a_i) \neq 0 \\ a_i & \text{если } f(a_i, a_i) = 0, \end{cases}$$

получим систему векторов, снова составляющую, как легко видеть, ортогональный базис нашего пространства. Кроме того, для произвольного $i = 1, 2, \dots, n$ из $f(a_i, a_i) = 0$ следует, очевидно, что $f(e_i, e_i) = 0$, а если $f(a_i, a_i) \neq 0$, то

$$f(e_i, e_i) = f(\frac{1}{\sqrt{|f(a_i, a_i)|}} a_i, \frac{1}{\sqrt{|f(a_i, a_i)|}} a_i) = \left(\frac{1}{\sqrt{|f(a_i, a_i)|}}\right)^2 f(a_i, a_i) =$$
$$\frac{1}{|f(a_i, a_i)|} f(a_i, a_i) = \pm 1.$$

Таким образом, базис $e_1, e_2, \dots, e_n$ является ортонормированным. $\square$

Несмотря на простоту критерия положительной определенности симметричной билинейной функции, доставляемого предложением 1, пользоваться им не очень удобно, так как билинейная функция, как правило, задается матрицей в случайном базисе пространства, и чтобы воспользоваться предложением 1, нам пришлось бы переходить к какому–либо ортогональному базису этого пространства. Более прямой способ распознавания положительной определенности билинейной функции на действительном конечномерном пространстве доставляет следующий классический результат — *критерий Сильвестра положительной определенности симметричной билинейной функции*:

Теорема 1. *Пусть $f(x, y)$ — симметричная билинейная функция на n-мерном действительном пространстве L и пусть*

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix}$$

— матрица этой функции в некотором базисе S пространства L. Функция $f(x, y)$ является положительно определенной тогда и только тогда, когда все главные миноры

$$\Delta_k = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1k} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2k} \\ \dots & \dots & \dots & \dots \\ \alpha_{k1} & \alpha_{k2} & \dots & \alpha_{kk} \end{vmatrix} \quad (k = 1, 2, \dots n)$$

матрицы A положительны.

Доказательство. Будем проводить индукцию по числу n. Пусть $n = 1$. Тогда базис S пространства L состоит из единственного вектора e, а матрица функции $f(x, y)$ в этом базисе имеет вид $A = (\alpha)$, где $\alpha = f(e, e)$. Поэтому если функция $f(x, y)$ положительно определена, то $\alpha > 0$ и потому единственный главный минор $\Delta_1 = \det A = \alpha$ матрицы A положителен.

Обратно, пусть $\Delta_1 > 0$, т.е. $\alpha > 0$. Произвольный вектор $a \in L$ имеет вид $a = \lambda e$ для подходящего числа λ, и потому

$$f(a, a) = f(\lambda e, \lambda e) = \lambda^2 f(e, e) = \lambda^2 \alpha.$$

Следовательно, если $a \neq 0$, т.е. $\lambda \neq 0$, то $f(a, a) > 0$. Значит функция $f(x, y)$ является положительно определенной.

Предположим теперь, что $n > 1$ и что для всех пространств меньшей размерности утверждение теоремы справедливо. Для доказательства справедливости теоремы для пространства L введем еще несколько обозначений.

Пусть $S = (e_1, e_2, \ldots, e_n)$ и пусть $M = l(e_1, e_2, \ldots, e_{n-1})$ — подпространство пространства L, порожденное векторами $e_1, e_2, \ldots, e_{n-1}$. Так как тогда система

$$S_1 = (e_1, e_2, \ldots, e_{n-1})$$

является базисом пространства M, имеем $\dim M = n - 1$. Кроме того, если $f_1(x, y)$ обозначает ограничение функции $f(x, y)$ на подпространство M, то очевидно, что матрица A_1 функции $f_1(x, y)$ в базисе S_1 имеет вид

$$A_1 = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n-1} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n-1} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{n-11} & \alpha_{n-12} & \ldots & \alpha_{n-1n-1} \end{pmatrix},$$

т.е. получается из матрицы A вычеркиванием последней строки и последнего столбца. Поэтому последовательность главных миноров матрицы A_1 получается отбрасыванием последнего члена в последовательности главных миноров матрицы A.

Приступая непосредственно к доказательству справедливости утверждения теоремы для пространства L, предположим сначала, что функция $f(x, y)$ является положительно определенной. Ясно, что тогда положительно определенной будет и функция $f_1(x, y)$, и по индуктивному предположению все главные миноры матрицы A_1, т.е. главные миноры $\Delta_1, \Delta_2, \ldots, \Delta_{n-1}$ матрицы A должны быть положительными. Таким образом, нам остается лишь установить положительность минора $\Delta_n = \det A$.

Пусть $S' = (e'_1, e'_2, \ldots, e'_n)$ — некоторый ортогональный базис пространства L и

$$A' = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

— матрица функции $f(x,y)$ в базисе S'. Так как функция $f(x,y)$ положительно определена и $\lambda_i = f(e'_i, e'_i)$, то для всех $i = 1, 2, \ldots, n$ $\lambda_i > 0$. Если T — матрица перехода от базиса S' к базису S, то $A = TA'T^t$ и потому

$$\det A = \det(TA'T^t) = \det T \det A' \det T^t = \det T \det T^t \det A' = (\det T)^2 \lambda_1 \lambda_2 \cdots \lambda_n > 0.$$

Обратно, предположим, что все главные миноры $\Delta_1, \Delta_2, \ldots, \Delta_n$ матрицы A положительны. Так как тогда будут положительными все главные миноры матрицы A_1, из индуктивного предположения следует, что функция $f_1(x,y)$ положительно определена. Выберем теперь ортогональный базис $b_1, b_2, \ldots, b_{n-1}$ пространства M. Так как тогда все числа $f(b_i, b_i) = f_1(b_i, b_i)$ положительны, можно ввести в рассмотрение вектор

$$e = e_n - \sum_{i=1}^{n-1} \beta_i b_i,$$

где

$$\beta_i = \frac{f(e_n, b_i)}{f(b_i, b_i)} \quad (i = 1, 2, \ldots n-1).$$

Очевидно, что вектор e не принадлежит подпространству M, так как в противном случае в M находился бы и вектор $e_n = e + \sum_{i=1}^{n-1} \beta_i b_i$. Поэтому система $S'' = (e_1, e_2, \ldots, e_{n-1}, e)$ линейно независима и, следовательно, является базисом пространства L. Далее, для произвольного $k = 1, 2, \ldots, n-1$, учитывая ортогональность системы $b_1, b_2, \ldots, b_{n-1}$, имеем

$$f(e, b_k) = f(e_n - \sum_{i=1}^{n-1} \beta_i b_i, b_k) = f(e_n, b_k) - \sum_{i=1}^{n-1} \beta_i f(b_i, b_k) = f(e_n, b_k) - \beta_k f(b_k, b_k) = 0.$$

Значит вектор e ортогонален каждому вектору из базиса $b_1, b_2, \ldots, b_{n-1}$ подпространства M, а потому e ортогонален любому вектору из этого подпространства. В частности, для любого $k = 1, 2, \ldots, n-1$ $f(e, e_k) = 0$. Поэтому матрица функции $f(x,y)$ в базисе S'' имеет вид

$$A'' = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n-1} & 0 \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n-1} & 0 \\ \multicolumn{5}{c}{\dotfill} \\ \alpha_{n-11} & \alpha_{n-12} & \ldots & \alpha_{n-1n-1} & 0 \\ 0 & 0 & \ldots & 0 & \lambda \end{pmatrix},$$

где $\lambda = f(e,e)$. Если теперь T_1 — матрица перехода от базиса S к базису S'', то $A'' = T_1 A T_1^t$. Отсюда

$$\det A'' = \det(T_1 A T_1^t) = \det T_1 \det A \det T_1^t = (\det T_1)^2 \det A.$$

Так как $\det A'' = \lambda\Delta_{n-1}$, имеем

$$\lambda\Delta_{n-1} = (\det T_1)^2\Delta_n,$$

откуда следует, что $\lambda > 0$. Записывая теперь произвольный вектор $a \in L$ в виде $a = b + \alpha e$, где $b \in M$, получаем

$$f(a,a) = f(b+\alpha e, b+\alpha e) = f(b,b) + f(\alpha e, \alpha e) = f(b,b) + \alpha^2 f(e,e) = f(b,b) + \alpha^2\lambda.$$

Поэтому если $a \neq 0$, то или $b \neq 0$, или $\alpha \neq 0$, и потому $f(a,a) > 0$. Следовательно, функция $f(x,y)$ является положительно определенной, и индукция завершена. $\square$

Следствие. *В обозначениях из формулировки теоремы 1 функция $f(x,y)$ является отрицательно определенной тогда и только тогда, когда все числа $\Delta_1, \Delta_2, \ldots, \Delta_n$ отличны от нуля и их знаки чередуются, причем $\Delta_1 < 0$.*

Доказательство. Функция $f(x,y)$ отрицательно определена тогда и только тогда, когда функция $f'(x,y) = -f(x,y)$ (являющаяся, очевидно, билинейной и симметричной) положительно определена. В силу теоремы 1 это равносильно положительности всех главных миноров $\Delta'_1, \Delta'_2, \ldots, \Delta'_n$ матрицы A' функции $f(x,y)$. Так как $A' = -A$, $\Delta'_k = (-1)^k\Delta_k$ $(k = 1, 2, \ldots, n)$, откуда и следует наше утверждение. $\square$

Докажем еще один классический результат о симметричных билинейных функциях на действительных пространствах — закон инерции для билинейных функций.

Теорема 2. *Пусть $f(x,y)$ — симметричная билинейная функция на конечномерном линейном пространстве L и пусть S и S' — два ортогональных базиса пространства L. Число положительных, отрицательных и изотропных векторов базиса S совпадает соответственно с числом положительных, отрицательных и изотропных векторов базиса S'.*

Доказательство. Начнем с введения ряда обозначений. Пусть

r_1 — число положительных векторов базиса S,

r_{-1} — число отрицательных векторов базиса S,

r_0 — число изотропных векторов базиса S.

Пусть r'_1, r'_{-1}, r'_0 — аналогичные числа для базиса S'. Очевидно, что тогда

$$r_1 + r_{-1} + r_0 = \dim L = r'_1 + r'_{-1} + r'_0.$$

Так как базис S пространства L ортогонален, матрица A функции $f(x,y)$ в этом базисе является диагональной, причем число ненулевых элементов ее диагонали равно $r_1 + r_{-1}$. Следовательно, ранг матрицы A равен $r_1 + r_{-1}$. Аналогично, ранг матрицы A' функции $f(x,y)$ в базисе S' равен $r'_1 + r'_{-1}$. Так как $A' = TAT^t$, где T — матрица перехода от базиса S к базису S' и потому обратима, ранги матриц A и A' совпадают. Поэтому $r_1 + r_{-1} = r'_1 + r'_{-1}$ и, следовательно, $r_0 = r'_0$. Остается доказать, что $r_1 = r'_1$ (поскольку тогда и $r_{-1} = r'_{-1}$), и рассуждая от противного, мы можем, не теряя общности, предположить, что $r_1 > r'_1$.

Пусть M — подпространство пространства L, порожденное всеми положительными векторами базиса S; очевидно, что $\dim M = r_1$, а в силу предложения 1 ограничение функции $f(x, y)$ на подпространство M является положительно определенной функцией, т.е. для любого ненулевого вектора $a \in M$ $f(a, a) > 0$. Пусть также N — подпространство пространства L, порожденное всеми отрицательными и изотропными векторами базиса S'. Тогда $\dim N = r'_{-1} + r'_0$ и по предложению 1 ограничение функции $f(x, y)$ на подпространство N является отрицательно полуопределенной функцией, т.е. для любого вектора $a \in N$ $f(a, a) \le 0$. В частности, $M \cap N = \{0\}$. Отсюда следует, что

$$\dim (M + N) = \dim M + \dim N - \dim (M \cap N) = r_1 + r'_{-1} + r'_0 > r'_1 + r'_{-1} + r'_0 = \dim L.$$

Это противоречит тому, что размерность подпространства $M + N$ конечномерного пространства L не может превосходить размерности этого пространства. Таким образом, $r_1 = r'_1$, что и требовалось. $\square$

Доказанный закон инерции для симметричных билинейных функций на действительных линейных пространствах позволяет ввести следующее понятие:

Определение 3. *Пусть $f(x, y)$ — симметричная билинейная функция на действительном конечномерном линейном пространстве L. Сигнатурой функции $f(x, y)$ называется тройка чисел (r_1, r_{-1}, r_0), где числа r_1, r_{-1} и r_0 равны соответственно количеству положительных, отрицательных и изотропных векторов в ортогональном базисе пространства L.*

Понятие сигнатуры билинейной функции позволяет классифицировать конечномерные действительные линейные пространства с симметричной билинейной функцией.

Теорема 3. *Пусть L_1 и L_2 — конечномерные действительные линейные пространства с симметричными билинейными функциями $f_1(x, y)$ и $f_2(x, y)$ соответственно. Пространства L_1 и L_2 изометричны тогда и только тогда, когда сигнатуры функций $f_1(x, y)$ и $f_2(x, y)$ совпадают.*

Доказательство. Предположим сначала, что пространства L_1 и L_2 являются изометричными. Пусть $\varphi : L_1 \to L_2$ — изометрия и пусть $S = (a_1, a_2, \ldots, a_n)$ какой–нибудь ортогональный базис пространства L_1. Так как отображение φ является, в частности, изоморфизмом линейных пространств, система $S\varphi = (a_1\varphi, a_2\varphi, \ldots, a_n\varphi)$ является базисом пространства L_2. Поскольку, далее, по определению изометрии для любых $i, j = 1, 2, \ldots, n$ выполняются равенства

$$f_1(a_i, a_j) = f_2(a_i\varphi, a_j\varphi),$$

то система $S\varphi$ является ортогональной, и каждый вектор $a_i\varphi$ является положительным, отрицательным или изотропным тогда и только тогда, когда соответствующим свойством обладает вектор a_i. Следовательно, сигнатуры функций $f_1(x, y)$ и $f_2(x, y)$ совпадают.

Предположим теперь, что сигнатуры функций $f_1(x, y)$ и $f_2(x, y)$ (а значит — и размерности пространств L_1 и L_2) совпадают. Поэтому если $S_1 = (a_1, a_2, \dots, a_n)$ и $S_2 = (b_1, b_2, \dots, b_n)$ — ортонормированные базисы пространств L_1 и L_2 соответственно, то выполнив, если нужно, подходящую перестановку векторов одного из них, мы можем считать, что обе функции имеют в соответствующем базисе одну и ту же матрицу вида

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

(где каждое λ_i равно 1, -1 или 0).

Поскольку система $S_1 = (a_1, a_2, \dots, a_n)$ является базисом пространства L_1, из предложения I.5.1 следует, что существует линейное отображение $\varphi : L_1 \to L_2$ такое, что $a_i\varphi = b_i$ для любого $i = 1, 2, \dots, n$, а поскольку система $S_2 = (b_1, b_2, \dots, b_n)$ является базисом пространства L_2, то легко видеть, что φ — изоморфизм. Если теперь $a = \sum_{i=1}^{n} \alpha_i a_i$ и $b = \sum_{i=1}^{n} \beta_i a_i$ — произвольные векторы пространства L_1, то в силу предложения 2.3 имеем

$$f_1(a, b) = \sum_{i=1}^{n} \alpha_i \beta_i \lambda_i.$$

С другой стороны, поскольку

$$a\varphi = \sum_{i=1}^{n} \alpha_i b_i \quad \text{и} \quad b\varphi = \sum_{i=1}^{n} \beta_i b_i,$$

то снова из преложения 2.3 получаем

$$f_2(a\varphi, b\varphi) = \sum_{i=1}^{n} \alpha_i \beta_i \lambda_i.$$

Таким образом, $f_1(a, b) = f_2(a\varphi, b\varphi)$ и отображение φ является изометрией пространства L_1 на пространство L_2. $\square$

§ 4. ЕВКЛИДОВЫ ПРОСТРАНСТВА

Определение 1. *Действительное линейное пространство L с симметричной билинейной функцией $f(x,y)$ называется евклидовым, если функция $f(x,y)$ положительно определена.*

В этом случае функцию $f(x,y)$ будем называть скалярным произведением и для векторов $a, b \in L$ значение $f(a,b)$ этой функции записывать в виде (a,b).

Все пространства с билинейными функциями, перечисленные в параграфе 1, являются, как легко проверить, примерами евклидовых пространств, при условии, разумеется, что n-мерное координатное пространство строится над полем $\mathbb{R}$ действительных чисел.

Скалярное произведение векторов $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ и $b = (\beta_1, \beta_2, \ldots, \beta_n)$ пространства $\mathbb{R}^n$, вычисляемое по правилу

$$f(a,b) = \sum_{i=1}^{n} \alpha_i \beta_i,$$

будем называть стандартным скалярным произведением, а термин n-мерное координатное евклидово пространство будет обозначать пространство $\mathbb{R}^n$, рассматриваемое вместе со стандартным скалярным произведением.

Многие понятия и утверждения, введенные и доказанные в предыдущих параграфах этой главы для произвольных линейных пространств с симметричной билинейной функцией, в случае евклидовых пространств приобретают специфический вид. Так, из предложения 3.2 с учетом положительной определенности скалярного произведения получаем

Предложение 1. *В произвольном конечномерном евклидовом пространстве существует ортонормированный базис, т.е. такой базис $e_1, e_2, \ldots, e_n$, что*

$$(e_i, e_j) = \begin{cases} 1, & \text{если } i = j \\ 0, & \text{если } i \neq j. \end{cases} \quad \square$$

Сигнатура скалярного произведения на n-мерном евклидовом пространстве имеет вид $(n, 0, 0)$, и потому из теоремы 3.3 получаем

Предложение 2. *Конечномерные евклидовы пространства изометричны тогда и только тогда, когда равны их размерности.* $\square$

Отсюда следует, в частности, что произвольное n-мерное евклидово пространство L изометрично n-мерному координатному евклидову пространству $\mathbb{R}^n$. Это, впрочем, можно усмотреть и непосредственно: отображение, сопоставляющее вектору $a \in L$ его координатную строку $(a)_S$ в некотором ортонормированном базисе S пространства L, является изометрией этого пространства на пространство $\mathbb{R}^n$ со стандартным скалярным произведением. В частности, одномерное евклидово пространство можно отождествить с действительной прямой, а дмумерное и трехмерное — соответственно с евклидовой плоскостью и евклидовым пространством элементарной геометрии.

Метрические свойства евклидовых пространств основываются на следующем утверждении:

Теорема 1. *Для любых векторов a и b евклидова пространства L выполнено неравенство*

$$(a,b)^2 \le (a,a)(b,b). \tag{1}$$

При этом, равенство здесь имеет место в том и только в том случае, когда векторы a и b составляют линейно зависимую систему.

Неравенство (1) является обобщением целого ряда классических неравенств, доказанных известными математиками в разное время и в различных конкретных ситуациях. Поэтому его часто называют неравенством Коши – Буняковского – Шварца. Два частных случая этого неравенства будут приведены ниже.

Доказательство. Заметим, прежде всего, что утверждение теоремы очевидно, если хотя бы один из векторов a или b равен 0. Действительно, в этом случае обе части неравенства (1) обращаются в нуль и система a, b линейно зависима. Будем предполагать поэтому, что оба вектора a и b отличны от нуля.

Так как скалярное произведение является положительно определенной билинейной функцией, для любого действительного числа λ должно выполняться неравенство

$$(\lambda a + b, \lambda a + b) \ge 0, \tag{2}$$

причем знак равенства имеет место тогда и только тогда, когда $\lambda a + b = 0$. Используя билинейность и симметричность скалярного произведения, перепишем неравенство (2) в виде

$$\lambda^2(a,a) + 2\lambda(a,b) + (b,b) \ge 0. \tag{2'}$$

Выполнимость неравенства (2') при любом $\lambda \in \mathbb{R}$ означает, что квадратный трехчлен $x^2(a,a) + 2x(a,b) + (b,b)$ может принимать лишь неотрицательные значения. Так как старший коэффициент (a,a) этого трехчлена положителен, это имеет место тогда и только тогда, когда его дискриминант неположителен, т.е. (после сокращения на 4)

$$(a,b)^2 - (a,a)(b,b) \le 0.$$

Таким образом, неравенство (1) выполняется. Кроме того, мы видим, что знак равенства в (1) имеет место тогда и только тогда, когда наш квадратный трехчлен имеет действительный корень λ_0. Это означает, что при $\lambda = \lambda_0$ неравенство (2) обращается в равенство, что, как отмечалось, равносильно тому, что $\lambda_0 a + b = 0$. Наконец, поскольку векторы a и b отличны от нуля, существование λ такого, что $\lambda a + b = 0$, равносильно линейной зависимости системы a, b. $\square$

В частном случае n-мерного координатного евклидова пространства $\mathbb{R}^n$ доказанное неравенство превращается в классическое неравенство Коши:

Следствие 1. *Для любого $n \geq 1$ и любых действительных чисел $\alpha_1, \alpha_2, \ldots, \alpha_n$ и $\beta_1, \beta_2, \ldots, \beta_n$ выполняется неравенство*

$$\left(\sum_{i=1}^{n} \alpha_i \beta_i\right)^2 \leq \left(\sum_{i=1}^{n} \alpha_i^2\right) \left(\sum_{i=1}^{n} \beta_i^2\right).$$

Равенство здесь имеет место тогда и только тогда, когда векторы $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ и $(\beta_1, \beta_2, \ldots, \beta_n)$ пропорциональны. $\square$

Применяя теорему 1 к евклидову пространству функций, непрерывных на отрезке $[\alpha, \beta]$, получаем неравенство Буняковского – Шварца:

Следствие 2. *Для любых действительных функций $f(x)$ и $g(x)$, непрерывных на отрезке $[\alpha, \beta]$, выполняется неравенство*

$$\left(\int_{\alpha}^{\beta} f(x) g(x)\, dx\right)^2 \leq \int_{\alpha}^{\beta} f(x)^2\, dx \int_{\alpha}^{\beta} g(x)^2\, dx$$

Равенство здесь имеет место тогда и только тогда, когда функции $f(x)$ и $g(x)$ пропорциональны. $\square$

Теорема 1 позволяет ввести в произвольном евклидовом пространстве метрические соотношения:

Определение 2. *Длиной вектора a евклидова пространства L называется действительное число $|a| = \sqrt{(a, a)}$.*

Положительная определенность скалярного произведения обеспечивает существование длины для любого вектора a пространства L; при этом $|a| \geq 0$ и $|a| = 0$ тогда и только тогда, когда $a = 0$. Неравенство (1) теперь может быть записано в виде

$$|(a, b)| \leq |a| \cdot |b|. \tag{1'}$$

(Здесь и далее некоторое неудобство может быть вызвано тем обстоятельством, что для длины вектора и абсолютной величины числа используется одно и тоже обозначение. Тем не менее, никаких недоразумений от этого не может случиться, так как всегда понятно, обозначение какого объекта помещено между вертикальными черточками.)

Длина вектора евклидова пространства L обладает обычными свойствами длины. Действительно, для любых $\lambda \in \mathbb{R}$ и $a \in L$ имеем

$$|\lambda a| = \sqrt{(\lambda a, \lambda a)} = \sqrt{\lambda^2 (a, a)} = |\lambda|\, |a|.$$

Кроме того, выполнено неравенство треугольника:

Теорема 2. *Для любых векторов a и b евклидова пространства L выполнено неравенство*

$$|a+b| \leq |a| + |b|. \tag{3}$$

При этом, равенство здесь имеет место в том и только в том случае, когда либо один из векторов a или b равен нулю, либо $a = \lambda b$ для некоторого действительного числа $\lambda > 0$.

Доказательство. Для произвольных векторов $a, b \in L$ из определения длины вектора с учетом билинейности и симметричности скалярного произведения получаем

$$|a+b|^2 = (a+b, a+b) = (a,a) + 2(a,b) + (b,b).$$

Используя очевидное неравенство $(a,b) \leq |(a,b)|$ и неравенство (1'), отсюда имеем

$$|a+b|^2 \leq (a,a) + 2|(a,b)| + (b,b) \leq |a|^2 + 2|a|\,|b| + |b|^2 = (|a| + |b|)^2.$$

Так как в полученном неравенстве $|a+b|^2 \leq (|a|+|b|)^2$ основаниями квадратов являются неотрицательные числа, оно равносильно неравенству (3).

Из приведенных рассуждений следует также, что знак равенства в (3) имеет место тогда и только тогда, когда $(a,b) = |(a,b)|$ и неравенство (1) становится равенством. Первое из этих условий означает, что $(a,b) \geq 0$, а второе равносильно линейной зависимости системы a, b. Если предполагать, что каждый из векторов этой системы отличен от нуля, то ее линейная зависимость равносильна существованию действительного числа $\lambda \neq 0$ такого, что $a = \lambda b$. Так как тогда $(a,b) = \lambda(b,b)$, то $\lambda > 0$. $\square$

Определение 3. *Пусть a и b — ненулевые векторы евклидова пространства L. Величиной угла между векторами a и b будем называть действительное число φ такое, что*

$$\cos\varphi = \frac{(a,b)}{|a|\,|b|} \quad \text{и} \quad 0 \leq \varphi \leq \pi.$$

Заметим, что для ненулевых векторов a и b неравенство (1') может быть переписано в виде

$$-1 \leq \frac{(a,b)}{|a|\,|b|} \leq 1,$$

так что число φ всегда существует (и определено однозначно).

Непосредственно из определения следует, что ненулевые векторы a и b ортогональны тогда и только тогда, когда угол между ними равен $\pi/2$. Угол между векторами a и $-a$ равен π.

Следующее утверждение естественно назвать теоремой Пифагора.

Предложение 3. *Если $a_1, a_2, \ldots, a_m$ — ортогональная система векторов евклидова пространства L, то*

$$\left|\sum_{i=1}^{n} a_i\right|^2 = \sum_{i=1}^{n} |a_i|^2.$$

В самом деле,

$$|\sum_{i=1}^{n} a_i|^2 = (\sum_{i=1}^{n} a_i, \sum_{j=1}^{n} a_j) = \sum_{i=1}^{n}\sum_{j=1}^{n}(a_i, a_j) = \sum_{i=1}^{n}(a_i, a_i) = \sum_{i=1}^{n} |a_i|^2. \quad \square$$

В этой общей ситуации можно получать и другие известные утверждения элементарной геометрии. Например, справедливость теоремы косинусов устанавливается следующим образом:

$$|a-b|^2 = (a-b, a-b) = (a,a) - 2(a,b) + (b,b) = |a|^2 + |b|^2 - 2|a|\,|b|\cos\varphi.$$

(здесь векторы a и b отличны от нуля и φ — величина угла между ними).

Возвращаясь к свойствам евклидовых пространств, выражаемым на языке линейной алгебры, отметим, прежде всего,

Предложение 4. *Произвольная ортогональная система ненулевых векторов евклидова пространства линейно независима.*

В самом деле, если $a_1, a_2, \ldots, a_m$ — ортогональная система ненулевых векторов и если для некоторых действительных чисел $\alpha_1, \alpha_2, \ldots, \alpha_m$ выполнено равенство

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_m a_m = 0,$$

то для любого числа k, $1 \le k \le m$, имеем

$$0 = (\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_m a_m, a_k) = (\alpha_1 a_1, a_k) + (\alpha_2 a_2, a_k) + \cdots + (\alpha_m a_m, a_k) = \\ \alpha_1(a_1, a_k) + \alpha_2(a_2, a_k) + \cdots + \alpha_m(a_m, a_k) = \alpha_k(a_k, a_k),$$

и так как $(a_k, a_k) > 0$, то $\alpha_k = 0$. $\square$

Как уже отмечалось выше, из результатов параграфа 2 следует, что в произвольном конечномерном евклидовом пространстве (а потому — и в любом его подпространстве) существует ортогональный базис. Один из способов построения такого базиса, называемый алгоритмом ортогонализации Грама – Шмидта, доставляется следующим утверждением:

Теорема 3. *Пусть $a_1, a_2, \ldots, a_m$ — линейно независимая система векторов евклидова пространства L. Полагаем $b_1 = a_1$ и если для некоторого k, $1 < k \le m$, векторы $b_1, b_2, \ldots, b_{k-1}$ уже определены, полагаем*

$$b_k = a_k - \sum_{i=1}^{k-1} \frac{(a_k, b_i)}{(b_i, b_i)} b_i.$$

Тогда система векторов $b_1, b_2, \ldots, b_m$ ортогональна и ее линейная оболочка совпадает с линейной оболочкой системы $a_1, a_2, \ldots, a_m$.

Доказательство. Индукцией по k покажем справедливость следующих двух утверждений:

1) система $b_1, b_2, \ldots, b_k$ ортогональна;

2) $l(b_1, b_2, \ldots, b_k) = l(a_1, a_2, \ldots, a_k)$.

При $k = 1$ оба утверждения очевидны. Предположим, что для некоторого числа k, где $1 < k \leq m$, система векторов $b_1, b_2, \ldots, b_{k-1}$ ортогональна и порождает то же подпространство, что и система $a_1, a_2, \ldots, a_{k-1}$. Тогда для произвольного числа l, $1 \leq l \leq m$ имеем

$$(b_k, b_l) = (a_k - \sum_{i=1}^{k-1} \frac{(a_k, b_i)}{(b_i, b_i)} b_i, b_l) = (a_k, b_l) - \sum_{i=1}^{k-1} \frac{(a_k, b_i)}{(b_i, b_i)} (b_i, b_l)$$
$$= (a_k, b_l) - \frac{(a_k, b_l)}{(b_l, b_l)} (b_l, b_l) = 0,$$

так что и система $b_1, b_2, \ldots, b_k$ ортогональна. Далее, из индуктивного предположения следует, что векторы $b_1, b_2, \ldots, b_{k-1}$ лежат в линейной оболочке системы $a_1, a_2, \ldots, a_k$. Из определения вектора b_k видно, что тогда и этот вектор принадлежит линейной оболочке той же системы $a_1, a_2, \ldots, a_k$. Отсюда следует, что

$$l(b_1, b_2, \ldots, b_k) \subseteq l(a_1, a_2, \ldots, a_k).$$

Противоположное включение доказывается аналогично. $\square$

Следствие. *Произвольная ортогональная система ненулевых векторов конечномерного евклидова пространства может быть дополнена до ортогонального базиса этого пространства.*

Доказательство. Пусть $a_1, a_2, \ldots, a_m$ — ортогональная система ненулевых векторов конечномерного евклидова пространства L. Ввиду предложения 5 эта система является линейно независимой и потому может быть дополнена до некоторого базиса

$$a_1, a_2, \ldots, a_m, a_{m+1}, \ldots, a_n$$

пространства L. Система $b_1, b_2, \ldots, b_m, b_{m+1}, \ldots, b_n$, полученная из этого базиса с помощью алгоритма ортогонализации Грама - Шмидта, является ортогональным базисом нашего пространства. Остается заметить, что для любого $k = 1, 2, \ldots, m$ имеет место равенство $b_k = a_k$.

В самом деле, при $k = 1$ это очевидно. Если для некоторого k, $1 < k \leq m$, равенства $b_1 = a_1, b_2 = a_2, \ldots, b_{k-1} = a_{k-1}$ справедливы, то с учетом ортогональности системы $a_1, a_2, \ldots, a_m$ имеем

$$b_k = a_k - \sum_{i=1}^{k-1} \frac{(a_k, b_i)}{(b_i, b_i)} b_i = a_k - \sum_{i=1}^{k-1} \frac{(a_k, a_i)}{(a_i, a_i)} a_i = a_k. \quad \square$$

Теорема 4. *Конечномерное евклидово пространство является прямой суммой произвольного своего подпространства и его ортогонального дополнения.*

Доказательство. Пусть M — подпространство евклидова пространства L. Так как ортогональное дополнение к нулевому подпространству совпадает с L, а ортогональное дополнение к L совпадает с нулевым подпространством, и в этих случаях утверждение теоремы тривиально, мы будем предполагать в дальнейшем, что $M \neq \{0\}$ и $M \neq L$.

Обозначим через L' сумму подпространств M и $M^\perp$. Произвольный вектор a, принадлежащий пересечению $M \cap M^\perp$ подпространств M и $M^\perp$, входит, в частности, в подпространство $M^\perp$ и потому должен быть ортогональным с каждым вектором из подпространства M. Так как, к тому же, $a \in M$, отсюда следует, что $(a, a) = 0$, и поскольку скалярное произведение является положительно определенной билинейной функцией, то $a = 0$. Таким образом $M \cap M^\perp = \{0\}$, и это означает, что $L' = M \oplus M^\perp$.

Покажем теперь, что $L' = L$. Для этого фиксируем некоторые ортогональные базисы $a_1, a_2, \ldots, a_r$ и $b_1, b_2, \ldots, b_s$ подпространств M и $M^\perp$ соответственно. Тогда система

$$a_1, a_2, \ldots, a_r, b_1, b_2, \ldots, b_s$$

является базисом подпространства L' и притом, очевидно, ортогональным. Поэтому ввиду следствия к предложению 4 ее можно дополнить до ортогонального базиса

$$a_1, a_2, \ldots, a_r, b_1, b_2, \ldots, b_s, c_1, \ldots, c_t$$

пространства L. Предположение о том, что подпространство L' не совпадает с пространством L, означает теперь, что $t \geq 1$. Но тогда вектор c_1, будучи ортогональным, в частности, каждому вектору системы $a_1, a_2, \ldots, a_r$, порождающей подпространство M, должен быть ортогональным и к любому вектору этого подпространства. Таким образом, вектор c_1 входит в подпространство $M^\perp$, что невозможно, так как система $b_1, b_2, \ldots, b_s, c_1$ является линейно независимой. $\square$

Теорема 4 позволяет получить дальнейшие результаты о метрических свойствах евклидовых пространств. Пусть M — произвольное подпространство конечномерного евклидова пространства L. Тогда $L = M \oplus M^\perp$ и потому произвольный вектор $a \in L$ однозначно представим в виде $a = b + c$, где $b \in M$, $c \in M^\perp$. Вектор b называется ортогональной проекцией вектора a на подпространство M, а вектор c — перпендикуляром, опущенным из вектора a на это подпространство.

Договоримся также расстоянием между двумя непустыми подмножествами M и N евклидова пространства L называть действительное число

$$d(M, N) = \inf \left\{ |x - y| \;\middle|\; x \in M,\ y \in N \right\}.$$

Имеет место следующий аналог хорошо известного утверждения из элементарной геометрии:

Предложение 5. *Пусть M — подпространство конечномерного евклидова пространства L. Расстояние от вектора $a \in L$ до подпространства M равно длине перпендикуляра, опущенного из a на подпространство M.*

В самом деле, если $a = b + c$, где $b \in M$ и $c \in M^{\perp}$, то для любого вектора $x \in M$ вектор $b - x$ ортогонален вектору c, и потому с использованием теоремы Пифагора получаем

$$|a - x|^2 = |(b - x) + c|^2 = |b - x|^2 + |c|^2 \geq |c|^2,$$

причем при $x = b$ имеет место равенство. $\square$

§ 5. ЛИНЕЙНЫЕ ОПЕРАТОРЫ ЕВКЛИДОВЫХ ПРОСТРАНСТВ

В этом параграфе мы увидим, что наличие в евклидовом пространстве дополнительной структуры — скалярного произведения — позволяет получить и дополнительную информацию о строении линейных операторов такого пространства. Тем не менее, мы начнем с одного факта, имеющего место для любых действительных пространств.

Теорема 1. *Для любого оператора φ конечномерного действительного пространства L существует такое φ-инвариантное подпространство пространства L, размерность которого равна 1 или 2.*

Доказательство. Пусть L — n-мерное линейное пространство над полем $\mathbb{R}$ действительных чисел и пусть φ — произвольный линейный оператор пространства L. Фиксируем некоторый базис $S = (a_1, a_2, \ldots, a_n)$ пространства L и обозначим через A матрицу в этом базисе оператора φ. Пусть также $f(x)$ — характеристический многочлен матрицы A (и оператора φ) и $\lambda = \rho + \sigma i$, где $\rho, \sigma \in \mathbb{R}$, — какой-либо комплексный корень этого многочлена.

Если число λ является действительным, то существование искомого подпространства пространства L почти очевидно. В самом деле, в этом случае λ является собственным значением оператора φ, и потому в пространстве L существует собственный вектор a оператора φ, принадлежащий этому собственному значению. Тогда подпространство $M = l(a)$, порожденное вектором a, является 1-мерным φ-инвариантным подпространством пространства L.

В общем случае рассмотрим n-мерное координатное пространство $\mathbb{C}^n$ над полем $\mathbb{C}$ комплексных чисел. Обозначим через ψ оператор пространства $\mathbb{C}^n$, матрица которого в естественном базисе этого пространства совпадает с A. Так как тогда многочлен $f(x)$ является характеристическим многочленом оператора ψ, число λ будет собственным значением этого оператора, и потому существует ненулевой вектор $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ пространства $\mathbb{C}^n$, такой, что $a\psi = \lambda a$.

Для каждого номера $k = 1, 2, \ldots, n$ запишем k-ую компоненту α_k вектора a в алгебраической форме: $\alpha_k = \beta_k + \gamma_k i$, где $\beta_k, \gamma_k \in \mathbb{R}$. Имеем тогда $a = b + ci$, где $b = (\beta_1, \beta_2, \ldots, \beta_n)$, $c = (\gamma_1, \gamma_2, \ldots, \gamma_n)$. Поэтому

$$\lambda a = (\rho + \sigma i)(b + ci) = (\rho b - \sigma c) + (\rho c + \sigma b)i.$$

С другой стороны, поскольку матрица оператора ψ вычислена в естественном базисе пространства $\mathbb{C}^n$ (и потому $a\psi = aA$), равенство $a\psi = \lambda a$ дает

$$\lambda a = aA = (b + ci)A = bA + ciA.$$

Сравнивая действительные и мнимые части у двух полученных выражений вектора λa, имеем

$$bA = \rho b - \sigma c \quad \text{и} \quad cA = \rho c + \sigma b.$$

Возвращаясь теперь к линейному пространству L, полагаем

$$x = \sum_{k=1}^{n} \beta_k a_k \quad \text{и} \quad y = \sum_{k=1}^{n} \gamma_k a_k.$$

Таким образом, x и y — такие векторы пространства L, что $(x)_S = b$ и $(y)_S = c$. Поэтому

$$(x\varphi)_S = (x)_S A = \rho b - \sigma c = \rho(x)_S - \sigma(y)_S = (\rho x - \sigma y)_S,$$

откуда имеем $x\varphi = \rho x - \sigma y$. Аналогично доказывается, что $y\varphi = \rho y + \sigma x$. Эти равенства говорят о том, что если $M = l(x, y)$ — подпространство пространства L, порожденное векторами x и y, то $x\varphi, y\varphi \in M$, и потому подпространство M φ-инвариантно. Так как, к тому же, $\dim M \leq 2$, подпространство M искомое. $\square$

Введем теперь первое из основных понятий этого параграфа.

Определение 1. *Пусть φ — оператор евклидова пространства L. Оператор ψ пространства L называется сопряженным к оператору φ, если для любых векторов $a, b \in L$ имеет место равенство $(a\varphi, b) = (a, b\psi)$.*

Теорема 2. *Для любого оператора φ конечномерного евклидова пространства L существует и притом единственный оператор, сопряженный к φ. Матрица оператора, сопряженного к φ, вычисленная в произвольном ортонормированном базисе пространства L, получается транспонированием матрицы оператора φ, вычисленной в том же базисе.*

Доказательство. Покажем сначала, что не существует двух различных операторов пространства L, каждый из которых сопряжен к данному оператору φ. Пусть операторы ψ_1 и ψ_2 таковы, что для любых векторов $a, b \in L$ имеют место равенства $(a\varphi, b) = (a, b\psi_1)$ и $(a\varphi, b) = (a, b\psi_2)$. Тогда $(a, b\psi_1) - (a, b\psi_2) = 0$, и так как

$$(a, b\psi_1) - (a, b\psi_2) = (a, b\psi_1 - b\psi_2),$$

мы видим, что для любых векторов a и b пространства L выполняется равенство $(a, b\psi_1 - b\psi_2) = 0$. В частности, $(b\psi_1 - b\psi_2, b\psi_1 - b\psi_2) = 0$, и поскольку скалярное произведение является положительно определенной билинейной функцией, отсюда следует, что для любого вектора $b \in L$, $b\psi_1 - b\psi_2 = 0$, т. е. $b\psi_1 = b\psi_2$. Следовательно,$\psi_1 = \psi_2$.

Для доказательства остальных утверждений теоремы фиксируем произвольный ортонормированный базис $e_1, e_2, \ldots, e_n$ пространства L. Пусть $A = (\alpha_{ij})$ — матрица оператора φ в этом базисе и пусть ψ — оператор, матрица которого в том же базисе совпадает с матрицей A^t, полученной транспонированием матрицы A. Пусть еще $a = \sum_{i=1}^{n} \alpha_i e_i$ и $b = \sum_{i=1}^{n} \beta_i e_i$ — произвольные векторы пространства L. Тогда

$$a\varphi = \sum_{i=1}^{n} \alpha_i (e_i\varphi) = \sum_{i=1}^{n} \alpha_i \left(\sum_{j=1}^{n} \alpha_{ij} e_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_{ij} e_j$$
$$= \sum_{j=1}^{n} \sum_{i=1}^{n} \alpha_i \alpha_{ij} e_j = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \alpha_i \alpha_{ij} \right) e_j,$$

и потому ввиду ортонормированности базиса $e_1, e_2, \ldots, e_n$

$$(a\varphi, b) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \alpha_i \alpha_{ij} \right) \beta_j = \sum_{j=1}^{n} \sum_{i=1}^{n} \alpha_i \alpha_{ij} \beta_j.$$

Аналогично,

$$b\psi = \sum_{j=1}^{n} \beta_j (e_j \psi) = \sum_{j=1}^{n} \beta_j \left(\sum_{i=1}^{n} \alpha_{ij} e_j \right) = \sum_{j=1}^{n} \sum_{i=1}^{n} \beta_j \alpha_{ij} e_i$$
$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_j \alpha_{ij} e_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \beta_j \alpha_{ij} \right) e_i,$$

и потому

$$(a, b\psi) = \sum_{i=1}^{n} \alpha_i \left(\sum_{j=1}^{n} \beta_j \alpha_{ij} \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j \alpha_{ij}.$$

Итак, мы видим, что для любых векторов $a, b \in L$ имеет место равенство $(a\varphi, b) = (a, b\psi)$, и потому построенный нами оператор ψ является сопряженным к оператору φ. Кроме того, построение оператора ψ делает последнее утверждение теоремы очевидным. $\square$

Оператор евклидова пространства L, сопряженный к оператору φ, будем обозначать символом φ^*. Таким образом, оператор φ^* однозначно определяется оператором φ и тем условием, что для любых векторов $a, b \in L$ выполнено равенство $(a\varphi, b) = (a, b\varphi^*)$. Так как ввиду симметричности скалярного произведения имеем

$$(b\varphi^*, a) = (a, b\varphi^*) = (a\varphi, b) = (b, a\varphi),$$

то оператор $\varphi^{**} = (\varphi^*)^*$, сопряженный к оператору φ^*, совпадает с оператором φ. Кроме того, для любых операторов φ и ψ пространства L имеем

$$(a(\varphi\psi), b) = ((a\varphi)\psi), b) = (a\varphi, b\psi^*) = (a, (b\psi^*)\varphi^*) = (a, b(\psi^*\varphi^*)),$$

откуда следует, что оператор $(\varphi\psi)^*$, сопряженный к произведению операторов φ и ψ, совпадает с произведением $\psi^*\varphi^*$ операторов, сопряженных к сомножителям, но перемножаемых в другом порядке. Таким образом, нами доказано

Следствие. *Для любых операторов φ и ψ евклидова пространства L выполнены равенства $\varphi^{**} = \varphi$ и $(\varphi\psi)^* = \psi^*\varphi^*$.*

Определение 2. *Оператор φ евклидова пространства L называется самосопряженным или симметричным, если $\varphi^* = \varphi$.*

Следующее утверждение вытекает непосредственно из теоремы 2.

Предложение 1. *Оператор конечномерного евклидова пространства L симметричен тогда и только тогда, когда симметричной является матрица этого оператора, вычисленная в произвольном ортонормированном базисе пространства L.* $\square$

Отметим еще

Предложение 2. *Пусть φ — оператор евклидова пространства L и M — φ-инвариантное подпространство этого пространства. Тогда ортогональное дополнение $M^\perp$ подпространства M является φ^*-инвариантным. В частности, если подпространство M инвариантно относительно симметричного оператора φ, то и ортогональное дополнение $M^\perp$ подпространства M является инвариантным относительно этого оператора.*

В самом деле, для любого вектора $a \in M^\perp$ и произвольного $x \in M$ имеем $x\varphi \in M$ и потому

$$(a\varphi^*, x) = (a, x\varphi) = 0,$$

так что $a\varphi^* \in M^\perp$. $\square$

Теперь мы можем полностью прояснить геометрическую природу симметричных операторов евклидовых пространств:

Теорема 3. *Пусть φ — симметричный оператор конечномерного евклидова пространства L. Существует ортонормированный базис пространства L, матрица оператора φ в котором является диагональной.*

Говоря более наглядно, в теореме утверждается, что для произвольного симметричного оператора евклидова пространства можно найти такой ортонормированный базис, на векторах которого оператор действует как растяжение.

Доказательство. Проведем индукцию по размерности n пространства L. При $n = 1$ утверждение теоремы тривиально. Пусть $n > 1$ и пусть утверждение теоремы выполняется для любого симметричного оператора любого $n-1$-мерного евклидова пространства. Для проведения индуктивного шага покажем сначала, что в пространстве L существует 1-мерное φ-инвариантное подпространство.

Теорема 1 гарантирует нам существование φ-инвариантного подпространства M, размерность которого равна 1 или 2. Пусть $\dim M = 2$. Из определения 2 непосредственно видно, что ограничение φ_1 оператора φ на подпространство M является симметричным оператором пространства M. Поэтому ввиду предложения 1 матрица оператора φ_1 в ортонормированном базисе этого пространства имеет вид $\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$, а потому его характеристический многочлен есть

$$f(x) = \begin{vmatrix} x-\alpha & -\beta \\ -\beta & x-\gamma \end{vmatrix} = x^2 - (\alpha+\gamma)x + (\alpha\gamma - \beta^2).$$

Так как дискриминант $(\alpha+\gamma)^2 - 4(\alpha\gamma-\beta^2) = (\alpha-\gamma)^2 + \beta^2$ этого квадратного трехчлена неотрицателен, характеристический многочлен оператора φ_1 имеет действительный корень, а потому в пространстве M существует собственный вектор a этого оператора. Поскольку вектор a является, очевидно, собственным вектором и оператора φ, подпространство $l(a)$, порожденное вектором a, является искомым φ-инвариантным одномерным подпространством.

Итак, в любом случае существует одномерное подпространство A пространства L, инвариантное относительно оператора φ. Поэтому (см. предложение 2 и теорему 4.4) пространство L является прямой суммой φ-инвариантных подпространств A и $A^\perp$ размерности 1 и $n-1$ соответственно. Так как ограничения оператора φ на эти подпространства являются симметричными операторами, в каждом из этих подпространств существуют ортонормированные базисы, матрицы в которых соответствующих ограничений оператора φ диагональны. Соединение этих базисов дает ортонормированный базис пространства L, матрица оператора φ в котором также диагональна. Этим завершен индуктивный переход, и теорема доказана. $\square$

Из этой теоремы можно получить важный и интересный результат о корнях характеристического многочлена симметричной действительной матрицы:

Следствие. *Все корни характеристического многочлена произвольной симметричной действительной матрицы являются действительными числами.*

Доказательство. Пусть A — произвольная симметричная действительная матрица порядка n. Возьмем n-мерное евклидово пространство L и фиксируем в нем некоторый ортонормированный базис S. Пусть φ — оператор пространства L, матрица которого в базисе S совпадает с матрицей A. Этот оператор в силу предложения 1 является симметричным, и по теореме 3 в пространстве L существует ортонормированный базис S', матрица A' оператора φ в котором является диагональной. Очевидно, что характеристический многочлен матрицы A' раскладывается над полем $\mathbb{R}$ действительных чисел на линейные множители. Остается вспомнить, что матрицы A и A' связаны соотношением $A' = TAT^{-1}$, где T — матрица перехода от базиса S к базису S', и потому характеристические многочлены этих матриц совпадают. $\square$

Определение 3. *Оператор φ евклидова пространства L называется ортогональным, если для любых векторов $a, b \in L$ выполнено равенство $(a\varphi, b\varphi) = (a, b)$.*

Понятие ортогонального оператора евклидова пространства обобщает понятие движения из элементарной геометрии, т.е. такого преобразования плоскости или пространства, которое сохраняет расстояния между точками. В самом деле, имеет место

Предложение 3. *Оператор φ евклидова пространства L является ортогональным тогда и только тогда, когда для любого вектора $a \in L$ выполнено равенство $|a\varphi| = |a|$.*

В самом деле, если оператор φ ортогонален, то для любого вектора $a \in L$

$$|a\varphi| = \sqrt{(a\varphi, a\varphi)} = \sqrt{(a, a)} = |a|.$$

Обратное утверждение является очевидным следствием легко проверяемого равенства

$$(a, b) = \frac{1}{2}((a + b, a + b) - (a, a) - (b, b)). \quad \square$$

Приведем еще два простых свойства ортогональных операторов.

Предложение 4. *Если число $\lambda \in \mathbb{R}$ является собственным значением ортогонального оператора φ евклидова пространства L, то $\lambda = \pm 1$. В частности, всякий ортогональный оператор инъективен.*

Пусть, в самом деле, для некоторого ненулевого вектора $a \in L$ имеет место равенство $a\varphi = \lambda a$. Тогда

$$(a, a) = (a\varphi, a\varphi) = (\lambda a, \lambda a) = \lambda^2 (a, a),$$

и так как $(a, a) > 0$, получаем $\lambda^2 = 1$, т.е. $\lambda = \pm 1$. Для доказательства второго утверждения достаточно заметить, что ненулевые векторы, входящие в ядро произвольного оператора линейного пространства, являются собственными векторами этого оператора, принадлежащими собственному значению, равному 0. $\square$

Предложение 5. *Пусть φ — ортогональный оператор конечномерного евклидова пространства L и M — φ-инвариантное подпространство этого пространства. Тогда и ортогональное дополнение $M^{\perp}$ подпространства M является φ-инвариантным.*

Если вектор $a \in L$ принадлежит подпространству $M^{\perp}$, то для любого вектора $b \in M$ имеет место равенство $(a, b) = 0$, а потому и $(a\varphi, b\varphi) = 0$. Так как ограничение оператора φ на подпространство M является, очевидно, ортогональным оператором пространства M, то ввиду предложения 4 и конечномерности этого пространства отображение φ действует на подпространстве M сюръективно. Это означает, что вектор $b\varphi$ вместе с вектором b пробегает все подпространство M, и потому вектор $a\varphi$ ортогонален ко всем векторам подпространства M. Следовательно, $a\varphi \in M^{\perp}$. $\square$

Определение 4. *Квадратная действительная матрица A называется ортогональной, если $AA^t = E$ (где, напомним, A^t — матрица, транспонированная к матрице A).*

Очевидно, что в силу этого определения матрица A является ортогональной тогда и только тогда, когда она обратима и обратная к ней матрица A^{-1} совпадает с матрицей A^t. Поэтому в определении 4 вместо равенства $AA^t = E$ можно было бы потребовать выполнимость равенства $A^tA = E$. Далее, транспонируя обе части равенства $A^{-1} = A^t$, получаем $(A^{-1})^t = (A^t)^t = A = (A^{-1})^{-1}$, так что матрица, обратная к ортогональной, также ортогональна. Кроме того, если $A^{-1} = A^t$ и $B^{-1} = B^t$, то $(AB)^{-1} = B^{-1}A^{-1} = B^tA^t = (AB)^t$, так что произведение двух ортогональных матриц тоже является ортогональной матрицей. Все сказанное означает, что множество $\mathcal{O}(n)$ всех ортогональных матриц порядка n является подгруппой группы $GL_n(\mathbb{R})$ всех обратимых матриц порядка n.

Еще одну характеризацию ортогональных матриц доставляет

Предложение 6. *Для квадратной действительной матрицы A следующие утверждения равносильны:*

(1) $A \in \mathcal{O}(n)$;

(2) *система строк матрицы A является ортонормированным базисом n-мерного координатного евклидова пространства $\mathbb{R}^n$;*
(3) *система столбцов матрицы A является ортонормированным базисом n-мерного координатного евклидова пространства $\mathbb{R}^n$.*

Легко видеть, в самом деле, что если $a_1, a_2, \ldots, a_n$ — система строк матрицы A, то равенство $AA^t = E$ равносильно системе равенств $(a_i, a_j) = \delta_{ij} (i, j = 1, 2, \ldots, n)$, которая и означает, что система $a_1, a_2, \ldots, a_n$ ортонормирована. Остается напомнить, что ортонормированная система векторов евклидова пространства линейно независима.

Аналогично, равенство $A^t A = E$ равносильно системе равенств $(b_i, b_j) = \delta_{ij} (i, j = 1, 2, \ldots, n)$, где $b_1, b_2, \ldots, b_n$ — система столбцов матрицы A. □

Теорема 4. *Следующие утверждения для оператора φ конечномерного евклидова пространства L равносильны:*

(1) *оператор φ ортогонален;*
(2) $\varphi\varphi^* = \iota$;
(3) *матрица оператора φ в произвольном ортонормированном базисе пространства L ортогональна;*
(4) *образ произвольного ортонормированного базиса пространства L относительно оператора φ является ортонормированным базисом пространства L.*
(5) *существует ортонормированный базис пространства L, образ которого относительно оператора φ является ортонормированным базисом этого пространства.*

Доказательство. Если оператор φ ортогонален, то для любых векторов a и b пространства L с учетом определения сопряженного оператора имеем

$$(a, b) = (a\varphi, b\varphi) = (a, (b\varphi)\varphi^*) = (a, b(\varphi\varphi^*)),$$

и это означает, что для любых векторов $a, b \in L$ имеет место равенство $(a\iota, b) = (a, b(\varphi\varphi^*))$, что, в свою очередь, говорит о том, что оператор $\varphi\varphi^*$ сопряжен тождественному оператору ι пространства L. Так как тождественный оператор самосопряжен, отсюда следует, что $\varphi\varphi^* = \iota$, и импликация (1) $\Rightarrow$ (2) доказана.

Импликация (2) $\Rightarrow$ (3) почти очевидна: если A — матрица оператора φ в некотором ортонормированном базисе пространства L, то по теореме 2 матрицей оператора φ^* в том же базисе является транспонированная к ней матрица A^t, и так как матрица произведения $\varphi\varphi^*$ этих операторов равна произведению матриц AA^t, из условия $\varphi\varphi^* = \iota$ следует условие $AA^t = E$, что и означает ортогональность матрицы A.

Для доказательства импликации (3) $\Rightarrow$ (4) выберем произвольный ортонормированный базис $a_1, a_2, \ldots, a_n$ пространства L и обозначим через $A = (\alpha_{ij})$ матрицу оператора φ в этом базисе. Полагаем также для $i = 1, 2, \ldots, n$ $b_i = a_i\varphi$ и покажем, что если матрица A ортогональна, то система $b_1, b_2, \ldots, b_n$ является ортонормированной (и вследствие этого, в силу предложения 4.4 автоматически оказывается базисом пространства L). В самом деле, по определению матрицы линейного оператора

$b_i = \sum_{j=1}^n \alpha_{ij} a_j$ $(i = 1, 2, \ldots, n$ и потому для произвольных i и j

$$(b_i, b_j) = \left(\sum_{k=1}^n \alpha_{ik} a_k, \sum_{l=1}^n \alpha_{jl} a_l \right) = \sum_{k=1}^n \sum_{l=1}^n \alpha_{ik} \alpha_{jl} (a_k, a_l).$$

Так как система векторов $a_1, a_2, \ldots, a_n$ является ортонормированной, отсюда следует, что

$$(b_i, b_j) == \sum_{k=1}^n \alpha_{ik} \alpha_j,$$

и из предложения 6 теперь имеем, что $(b_i, b_j) = 0$, если $i \neq j$, и $(b_i, b_j) = 1$, если $i = j$.

Импликация (4) $\Rightarrow$ (5) просто очевидна, и мы переходим к доказательству заключительной импликации (5) $\Rightarrow$ (1). Пусть $S = (e_1, e_2, \ldots, e_n)$ и $S' = (e'_1, e'_2, \ldots, e'_n)$ — два ортонормированных базиса пространства таких, что для любого $i = 1, 2, \ldots, n$ $e_i\varphi = e'_i$. Пусть a и b — произвольные векторы пространства L и $a = \sum_{i=1}^n \alpha_i e_i$ и $b = \sum_{i=1}^n \beta_i e_i$ — их выражения через векторы базиса S. Тогда $a\varphi = \sum_{i=1}^n \alpha_i e'_i$ и $b\varphi = \sum_{i=1}^n \beta_i e'_i$ — выражения этих векторов через векторы базиса S'. Поскольку базисы S и S' ортонормированы, из предложения 2.3 следует, что $(a, b) = \sum_{i=1}^n \alpha_i \beta_i$ и $(a\varphi, b\varphi) = \sum_{i=1}^n \alpha_i \beta_i$. Таким образом, для произвольных векторов $a, b \in L$ имеет место равенство $(a, b) = (a\varphi, b\varphi)$, т. е. оператор φ является ортогональным. $\square$

Перейдем теперь к выяснению геометрической структуры ортогональных операторов евклидовых пространств. Начнем с ортогонального оператора φ, действующего на одномерном пространстве (т. е. на действительной прямой). Так как произвольный оператор одномерного линейного пространства является гомотетией, $\varphi = \chi_\lambda$ для некоторого $\lambda \in \mathbb{R}$, а так как при этом λ является собственным значением оператора φ, то в силу предложения 4 $\lambda = \pm 1$. Таким образом, оператор φ совпадает либо с тождественным отображением, либо с симметрией относительно начала координат.

Пусть теперь φ — ортогональный оператор 2-мерного евклидова пространства L. Фиксируем некоторый ортонормированный базис e_1, e_2 этого пространства, и пусть $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ — матрица оператора φ в этом базисе. Если $\beta = \gamma$, то матрица A симметрична и потому симметричным является оператор φ. Тогда по теореме 3 в пространстве L существует ортонормированный базис e'_1, e'_2, матрица оператора φ в котором имеет вид $A' = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, где $\lambda_i = \pm 1$. Поэтому оператор φ является либо тождественным (при $\lambda_1 = \lambda_2 = 1$), либо симметрией относительно начала координат (при $\lambda_1 = \lambda_2 = -1$), либо симметрией относительно прямой, определяемой вектором e'_1 (при $\lambda_1 = 1, \lambda_2 = -1$), либо симметрией относительно прямой, определяемой вектором e'_2 (при $\lambda_1 = -1, \lambda_2 = 1$).

Предположим теперь, что $\beta \neq \gamma$. По теореме 4 матрица A является ортогональной, и в силу предложения 6 должны выполняться равенства

$$\begin{aligned} \alpha\beta + \gamma\delta &= 0 \\ \alpha\gamma + \beta\delta &= 0 \\ \alpha^2 + \beta^2 &= 1 \\ \gamma^2 + \delta^2 &= 1. \end{aligned} \tag{1}$$

ПРИЛОЖЕНИЕ

Темы и примерные задачи практических занятий

1. Определение и простейшие свойства линейных пространств и их подпространств.

1.1 На множестве $\mathbb{R}_+$ всех положительных действительных чисел определим внутреннюю операцию $\oplus$ и внешнюю операцию $*$ умножения на действительные числа по правилам

$$a \oplus b = ab \quad \text{и} \quad \alpha * a = a^\alpha \quad (a, b \in \mathbb{R}_+, \ \alpha \in \mathbb{R}.$$

Показать, что относительно этих операций множество $\mathbb{R}_+$ является линейным пространством над полем $\mathbb{R}$.

1.2. На множестве $\mathbb{Q}^2$ всех упорядоченных пар рациональных чисел определим внутреннюю и внешнюю операции, полагая для $a = (\alpha, \beta)$, $b = (\gamma, \delta)$ и $\lambda \in \mathbb{Q}$

$$a + b = (\alpha + \gamma, \beta + \delta), \quad \lambda a = (\lambda\alpha, \beta).$$

Проверить, какие из аксиом линейного пространства здесь выполнены, а какие нет.

1.3. Пусть L — линейное пространство над полем K. Доказать, что для любых $a, b \in L$ и $\alpha, \beta \in K$

а) $\alpha a = 0$ тогда и только тогда, когда $\alpha = 0$ или $a = 0$;

б) $\alpha a = \beta a$ тогда и только тогда, когда $\alpha = \beta$ или $a = 0$;

в) $\alpha a = \alpha b$ тогда и только тогда, когда $\alpha = 0$ или $a = b$.

1.4*. Доказать, что не существует линейного пространства, аддитивная группа которого совпадает с аддитивной группой $\mathbb{Z}$ целых чисел.

1.5. Доказать, что в определении линейного пространства коммутативность сложения векторов является следствием остальных аксиом.

1.6. Пусть поле $\mathbb{R}$ действительных чисел рассматривается как линейное пространство над полем $\mathbb{Q}$ рациональных чисел. Для каждого из следующих подмножеств множества $\mathbb{R}$ выяснить, является ли оно подпространством пространства $\mathbb{R}$:

а) множество $\mathbb{R}_+ = \{a \in \mathbb{R} \mid a > 0\}$ всех положительных действительных чисел;

б) множество $\mathbb{Z}$ всех целых чисел;

в) множество чисел вида $a + b\sqrt{2}$, где a и b — произвольные рациональные числа.

1.7. Пусть $K[x]$ — линейное пространство многочленов от переменной x над полем K. Для фиксированных скаляров $\alpha, \beta \in K$ рассмотрим множество $M_{\alpha\beta}$ всех тех многочленов из $K[x]$, значение которых при $x = \alpha$ равно β:

$$M_{\alpha\beta} = \{f(x) \in K[x] \mid f(\alpha) = \beta\}.$$

При каких α и β множество $M_{\alpha\beta}$ является подпространством пространства $K[x]$?

1.8. Доказать, что пересечение двух подпространств линейного пространства L является подпространством этого пространства.

1.9. Пусть A и B — подпространства линейного пространства L. Доказать, что их теоретико-множественное объединение $A \cup B$ является подпространством пространства L тогда и только тогда, когда одно из них содержится в другом, т.е. или $A \subseteq B$ или $B \subseteq A$.

1.10. Пусть $L = \mathbb{Z}_2^4$ — четырехмерное координатное линейное пространство над полем $\mathbb{Z}_2$ классов вычетов целых чисел по модулю 2. Выписать все векторы, входящие в линейную оболочку каждой из следующих систем векторов этого пространства:

а) $(1,0,1,0)$;

б) $(1,0,1,1),\ (1,1,1,1)$;

в) $(1,1,0,1),\ (1,0,1,0),\ (0,1,1,1)$.

1.11. Найти условия, необходимые и достаточные для того, чтобы вектор $a = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ линейного пространства $\mathbb{Q}^4$ принадлежал линейной оболочке каждой из следующих систем векторов этого пространства:

а) $(1,0,0,1),\ (0,0,1,0)$;

б) $(1,0,0,2),\ (0,0,1,0)$;

с) $(1,-1,0,0),\ (0,0,1,-1)$.

1.12. Описать многочлены, принадлежащие линейной оболочке следующих систем векторов пространства $\mathbb{Q}[x]$:

а) $1,\ x,\ x^2$;

б) $1-x^2,\ x-x^2,\ 1+x+x^2$;

в) $1-x^2,\ x-x^2,\ 2-x-x^2$.

1.13. Пусть M — подпространство пространства $\mathbb{R}^5$, состоящее из всевозможных решений данной системы линейных однородных уравнений. Найти систему порождающих пространства M.

$$\text{а)}\ \begin{cases} 2x_1 + 2x_2 - 2x_3 + 5x_4 - x_5 = 0 \\ 2x_1 + 2x_2 - x_3 + 5x_4 = 0 \\ 3x_1 + 3x_2 - 2x_3 + 8x_4 = 0 \\ x_1 + x_2 - x_3 + 2x_4 - x_5 = 0 \\ x_1 + x_2 + 3x_4 + x_5 = 0 \end{cases}$$

$$\text{б)}\ \begin{cases} 2x_1 + 2x_2 + 6x_3 - 2x_4 + 5x_5 = 0 \\ 2x_1 + 3x_2 + 7x_3 + 6x_5 = 0 \\ x_1 + x_2 + 3x_3 - x_4 + 2x_5 = 0 \\ 3x_1 + 3x_2 + 9x_3 - 3x_4 + 8x_5 = 0 \\ x_1 + 2x_2 + 4x_3 + x_4 + 3x_5 = 0 \end{cases}$$

1.14. Найти систему порождающих линейного пространства из задачи 1.

1.15. Доказать, что следующие условия для ненулевого линейного пространства L равносильны:

(1) L обладает системой порождающих, состоящей из одного элемента;

(2) всякое подпространство пространства L либо является нулевым, либо совпадает с пространством L.

2. Суммы и прямые суммы подпространств.

2.1. Для данных подпространств L_1 и L_2 линейного пространства L показать, что $L = L_1 + L_2$. Выяснить, верно ли, что $L = L_1 \oplus L_2$.

$$\begin{array}{lll}
\text{а)} & L = \mathbb{R}^2, & L_1 = \{(\alpha,\beta) \mid \alpha+\beta = 0\}, \\
 & & L_2 = \{(\alpha,\beta) \mid \alpha = 0\}. \\
\text{б)} & L = \mathbb{R}^3, & L_1 = \{(\alpha,\beta,\gamma) \mid \alpha+\beta+\gamma = 0\}, \\
 & & L_2 = \{(\alpha,\beta,\gamma) \mid \alpha = 0\}. \\
\text{в)} & L = \mathbb{R}^n, & L_1 = \{(\alpha_1,\alpha_2,\ldots,\alpha_n) \mid \alpha_1+\alpha_2+\cdots+\alpha_n = 0\}, \\
 & & L_2 = \{(\alpha_1,\alpha_2,\ldots,\alpha_n) \mid \alpha_1 = \alpha_2 = \cdots = \alpha_n\}.
\end{array}$$

2.2. Пусть A и B — подпространства линейного пространства L. Доказать, что $A + B = B$ тогда и только тогда, когда $A \subseteq B$.

2.3. Пусть a и b — ненулевые векторы линейного пространства L, $M = l(a)$ и $N = l(b)$. Показать, что если $M \neq N$, то $M + N = M \oplus N$.

2.4. Пусть линейное пространство L является суммой своих подпространств A и B. Доказать, что $L = A \oplus B$ тогда и только тогда, когда хотя бы один вектор $c \in L$ однозначно представим в виде $c = a + b$, где $a \in A$ и $b \in B$.

2.5. Многочлен $f(x) \in \mathbb{R}[x]$ назовем четным (нечетным), если $f(-x) = f(x)$ (соответственно, $f(-x) = -f(x)$). Доказать, что множества M всех четных и N всех нечетных многочленов являются подпространствами линейного пространства $\mathbb{R}[x]$ и что $\mathbb{R}[x] = M \oplus N$. Верно ли последнее утверждение для линейного пространства $\mathbb{Z}_2[x]$?

2.6. Пусть $p(x)$ — многочлен степени $m > 0$ и M — множество всех многочленов из $K[x]$, кратных многочлену $p(x)$. Показать, что M — подпространство пространства $K[x]$. Показать, что $K[x] = M \oplus K_{m-1}[x]$.

2.7. Пусть $M_n(\mathbb{R})$ — линейное пространство $n \times n$-матриц над полем $\mathbb{R}$ действительных чисел, A и B — соответственно множества всех симметричных и кососимметричных матриц из $M_n(\mathbb{R})$. Показать, что A и B — подпространства пространства $M_n(\mathbb{R})$ и что $M_n(\mathbb{R}) = A \oplus B$. Верно ли это утверждение для пространства $M_n(\mathbb{Z}_2)$?

3. Определение и простейшие свойства линейных отображений.

3.1. Выяснить, является ли линейным каждое из отображений линейного пространства $\mathbb{R}^3$ в себя, определяемых следующим действием на произвольный вектор $a = (\alpha, \beta, \gamma)$ этого пространства:

а) $a\varphi = (\gamma, \alpha, \beta)$;
б) $a\varphi = (\alpha, \beta, \gamma^2)$;
в) $a\varphi = (\alpha, \beta, 0)$;
г) $a\varphi = (\alpha - 1, \beta, \gamma)$;
д) $a\varphi = (\alpha + \beta, \beta + \gamma, \gamma + \alpha)$.

Для отображений, являющихся линейными, найти системы порождающих образа и ядра.

3.2. Пусть отображение φ переводит произвольный многочлен $f(x) \in K[x]$ в многочлен $f_1(x)$. Выяснить, является ли φ линейным отображением, если

а) $f_1(x) = f(-x)$;
б) $f_1(x) = f(x+1)$;
в) $f_1(x) = f(x) + 1$;
г) $f(x) = f'(x)$.

3.3. Пусть M — подпространство линейного пространства L и пусть отображение φ множества L в себя определено по правилу

$$a\varphi = \begin{cases} a, & \text{если } a \in M \\ 0, & \text{если } a \notin M. \end{cases}$$

Выяснить, когда отображение φ является линейным.

3.4. Пусть $L = K^m$ и $M = K^n$ — соответственно m- и n-мерное координатные пространства над полем K и пусть A — фиксированная $m \times n$-матрица над этим полем. Показать, что отображение φ пространства L в пространство M, определенное по правилу $a\varphi = aA$ ($a \in L$), является линейным. В следующих примерах найти системы порождающих образа и ядра отображения φ:

а) $L = \mathbb{Q}^3, M = \mathbb{Q}^4, \qquad A = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 1 & 1 & 2 \\ 1 & 1 & 0 & 1 \end{pmatrix}$

б) $L = M = \mathbb{Q}^3, \qquad A = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 1 & 3 & 3 \end{pmatrix}$

в) $L = M = \mathbb{Q}^4, \qquad A = \begin{pmatrix} 1 & 2 & 1 & -1 \\ -1 & 1 & 2 & 2 \\ 1 & 3 & -1 & 0 \\ 0 & 1 & -2 & 1 \end{pmatrix}$

3.5. Пусть $\varphi : L \to M$ — линейное отображение пространства L в пространство M , A и B — подпространства пространства L. Доказать, что $(A+B)\varphi = A\varphi + B\varphi$.

3.6. Пусть $\varphi : L \to M$ — линейное отображение пространства L в пространство M, A и B — подпространства пространства L. Доказать, что $(A \cap B)\varphi \subseteq A\varphi \cap B\varphi$. Может ли включение быть строгим ?

Доказать, что если φ инъективно, то $(A \cap B)\varphi = A\varphi \cap B\varphi$.

3.7. Доказать, что для любых операторов φ и ψ линейного пространства L

$$\mathrm{Ker}\varphi \subseteq \mathrm{Ker}\varphi\psi.$$

Доказать, что если операторы φ и ψ перестановочны, то

$$\mathrm{Ker}\varphi + \mathrm{Ker}\psi \subseteq \mathrm{Ker}\varphi\psi.$$

3.8. Пусть $\varphi : L \to M$ — линейное отображение пространства L в пространство M. Доказать, что для любых подпространств A и B пространства L равенство $A\varphi = B\varphi$ имеет место тогда и только тогда, когда

$$A + \mathrm{Ker}\varphi = B + \mathrm{Ker}\varphi.$$

3.9. Пусть линейное пространство L является прямой суммой подпространств A и B , $L = A \oplus B$. Пусть φ — оператор проектирования на A параллельно B (т. е. для любого вектора $c \in L$ образ a вектора c при отображении φ определяется из условий $c = a + b$, $a \in A$, $b \in B$). Найти образ и ядро оператора φ. Доказать, что $\varphi^2 = \varphi$ (оператор, удовлетворяющий такому равенству, называется идемпотентным). Доказать, что произвольный идемпотентный оператор линейного пространства является проектированием, соответствующим подходящему разложению пространства в прямую сумму двух подпространств.

3.10. Пусть линейное пространство L является прямой суммой подпространств A и B , $L = A \oplus B$. Пусть φ — оператор проектирования на A параллельно B, ψ — оператор проектирования на B параллельно A. Доказать, что $\varphi + \psi$ является тождественным оператором пространства L.

3.11. Пусть линейное пространство L над полем K является прямой суммой подпространств A и B , $L = A \oplus B$. Определим отображение $\varphi : L \to L$ по следующему правилу: для произвольного вектора $c \in L$ находим разложение вида $c = a + b$, где $a \in A, b \in B$, и полагаем $c\varphi = a - b$.

а) Показать, что отображение φ является линейным оператором пространства L (оно называется оператором отражения относительно подпространства A параллельно подпространству B).

б) Показать, что оператор φ удовлетворяет равенству $\varphi^2 = \iota$ (оператор линейного пространства, удовлетворяющий этому равенству называется инволютивным).

в) Доказать, что если в поле K $1+1 \neq 0$, то произвольный инволютивный оператор линейного пространства L является отражением, соответствующим подходящему разложению пространства в прямую сумму двух подпространств.

г) Пусть $L = \mathbb{Z}_2^2$ — 2-мерное координатное пространство над полем $\mathbb{Z}_2$ и пусть отображение $\varphi : L \to L$ определено по правилу $(\alpha, \beta)\varphi = (\alpha, \alpha + \beta)$. Показать, что φ — инволютивный оператор пространства L, не являющийся отражением.

3.12. Пусть φ и ψ — ненулевые операторы пространства L. Доказать, что если $\mathrm{Ker}\varphi \neq \mathrm{Ker}\psi$, то система $\varphi,\ \psi$ линейно независима (в пространстве $\mathcal{L}(L, L)$).

3.13. Пусть L и M — линейные пространства над полем K и линейные отображения

$$\varphi_1,\ \varphi_2,\ \ldots,\ \varphi_n \in \mathcal{L}(L, M)$$

таковы, что для некоторого вектора $a \in L$ система

$$b_1 = a\varphi_1,\ b_2 = a\varphi_2,\ \ldots,\ b_n = a\varphi_n$$

векторов пространства L линейно независима. Доказать, что тогда линейно независимой будет и система

$$\varphi_1,\ \varphi_2,\ \ldots,\ \varphi_n$$

векторов пространства $\mathcal{L}(L, M)$.

4. Линейная зависимость систем векторов линейного пространства. Базис.

I. Стандартные вычислительные задачи в пространстве K^n.

I.4.1. Нахождение базиса и размерности линейной оболочки данной системы векторов.

Найти базис и размерность подпространства действительного координатного пространства соответствующей размерности, порожденного данной системой векторов.

а) $(1, 2, 2, -1),\ (2, 3, 2, 5),\ (-1, 4, 3, -1),\ (2, 9, 3, 5)$;

б) $(-3, 1, 5, 3, 2),\ (2, 3, 0, 1, 0),\ (1, 2, 3, 2, 1),\ (3, -5, -1, -3, -1),\ (3, 0, 1, 0, 0)$;

в) $(-1, 1, 0, 1, 2),\ (3, 3, 2, 1, 0),\ (1, 2, 1, 1, 1),\ (2, 1, 1, 0, -1)$.

4.2. Включение линейно независимой системы векторов подпространства в его базис. Нахождение прямого дополнения для данного подпространства.

4.3. Вычисление размерности суммы и пересечения двух подпространств.

Найти размерность суммы и пересечения подпространств Aи B действительного координатного пространства соответствующей размерности, порожденных системами векторов $a_1, a_2, \ldots, a_m$ и $b_1, b_2, \ldots, b_m$ соответственно.

а) $a_1 = (1, 2, 0, 1),\ a_2 = (2, 1, 3, -1),\ a_3 = (1, 1, 1, 0)$,
$b_1 = (1, 0, 1, 0),\ b_2 = (1, 0, 2, -1)$;

б) $a_1 = (1,1,1,1),\ a_2 = (1,-1,1,-1),\ a_3 = (1,3,1,3)$,
$b_1 = (1,2,0,2),\ b_2 = (1,2,1,2),\ b_3 = (3,1,3,1)$.

4.4. Нахождение системы линейных однородных уравнений, множество решений которой совпадает с данным подпространством.

4.5. Нахождение базиса суммы и пересечения двух подпространств.

4.6. Нахождение базисов образа и ядра и вычисление ранга и дефекта линейного отображения.

II. Задачи на исследование и доказательство

4.7. Найти все значения λ, при которых из линейной независимости системы a, b векторов линейного пространства L следует линейная независимость системы $\lambda a + b,\ a + \lambda b$.

4.8. Доказать, что следующие утверждения о системе $S = (a_1, a_2, \dots, a_n)$ векторов линейного пространства L равносильны:
а) система S линейно независима;
б) хотя бы один вектор пространства L однозначно выражается через систему S;
в) произвольный вектор из линейной оболочки системы S однозначно выражается через эту систему.

4.9. Пусть M — подпространство линейного пространства L и $a_1,\ a_2,\ \dots,\ a_m$ — базис подпространства M. Доказать, что произвольный вектор $b \in L$ входит в подпространство M тогда и только тогда, когда система $a_1,\ a_2,\ \dots,\ a_m,\ b$ линейно зависима.

4.10. Пусть a, b, c — линейно независимая система векторов линейного пространства L. Выяснить, являются ли независимыми следующие системы векторов этого пространства:
а) $a,\ a+b,\ a+b+c$;
б) $a-b,\ b-c,\ c-a$;
в) $a+b,\ b+c,\ c+a$;
г) $\alpha a - \beta b,\ \gamma b - \alpha c,\ \beta c - \gamma a$;

4.11. Пусть системы

$$a_1,\ a_2,\ \dots,\ a_m \quad \text{и} \quad b_1,\ b_2,\ \dots,\ b_m$$

векторов линейного пространства L эквивалентны. Доказать, что если одна из них линейно независима, то и другая должна быть линейно независимой.

4.12. Пусть системы векторов

$$a_1,\ a_2,\ \dots,\ a_m \quad \text{и} \quad b_1,\ b_2,\ \dots,\ b_n$$

линейного пространства L имеют одинаковые ранги. Доказать, что если одна из них линейно выражается через другую, то эти системы эквивалентны.

4.13. Пусть A и B — подпространства линейного пространства L, $a_1, a_2, \ldots, a_m$ — базис пространства A, $b_1, b_2, \ldots, b_n$ — базис пространства B.

1) Показать, что найдется натуральное число s, $0 \leq s \leq n$ такое, что (при подходящей перенумерации системы $b_1, b_2, \ldots, b_m$) система

$$a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_s$$

является базисом подпространства $A + B$.

2) Выразим через этот базис остальные векторы $b_{s+1}, b_{s+2}, \ldots, b_n$ базиса подпространства B,

$$b_i = \alpha_{i1}a_1 + \alpha_{i2}a_2 + \cdots + \alpha_{im}a_m + \beta_{i1}b_1 + \cdots + \beta_{is}b_s \ (i = s+1, \ldots, n),$$

и для каждого $i = s+1, \ldots, n$ полагаем

$$c_i = \alpha_{i1}a_1 + \alpha_{i2}a_2 + \cdots + \alpha_{im}a_m.$$

Показать, что система $c_{s+1}, c_{s+2}, \ldots, c_n$ является базисом пересечения $A \cap B$ подпространств A и B.

4.14. Показать, что произвольное подпространство M пространства n-мерного координатного пространства K^n совпадает с множеством всех решений некоторой системы линейных однородных уравнений. Для этого

Взять некоторый базис $a_1, a_2, \ldots, a_r$ подпространства M,

$$\begin{aligned} a_1 &= (\alpha_{11}, \alpha_{12}, \ldots, \alpha_{1n}) \\ a_2 &= (\alpha_{21}, \alpha_{22}, \ldots, \alpha_{2n}) \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ a_r &= (\alpha_{r1}, \alpha_{r2}, \ldots, \alpha_{rn}), \end{aligned}$$

и рассмотреть следующую систему линейных однородных уравнений:

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n = 0 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{2n}x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{r1}x_1 + \alpha_{r2}x_2 + \cdots + \alpha_{rn}x_n = 0. \end{cases}$$

Показать, что ее фундаментальная система решений состоит из $n-r$ векторов. Пусть

$$\begin{aligned} b_1 \ &= (\beta_{11},\ \beta_{12},\ \ldots,\ \beta_{1n}) \\ b_2 \ &= (\beta_{21},\ \beta_{22},\ \ldots,\ \beta_{2n}) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ b_{n-r} &= (\beta_{n-r,1}, \beta_{n-r,2}, \ldots, \beta_{n-r,n}) \end{aligned}$$

— фундаментальная система решений этой системы. Показать, что тогда подпространство M соспадает с множеством решений системы уравнений

$$\begin{cases} \beta_{11} \quad x_1 \ + \beta_{12} \quad x_2 \ + \cdots + \beta_{1n} \quad x_n = 0 \\ \beta_{21} \quad x_1 \ + \beta_{22} \quad x_2 \ + \cdots + \beta_{2n} \quad x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \beta_{n-r,1} x_1 + \beta n - r, 2x_2 + \cdots + \beta_{n-r,n} x_n = 0. \end{cases}$$

4.15. Найти размерность линейной оболочки системы векторов

$$a_1 = (1,1,1,\lambda), a_2 = (2,2,2,\lambda), a_3 = (3,3,3,\lambda), a_4 = (4,4,4,4)$$

пространства R^4. Существует ли базис этой линейной оболочки, не зависящий от параметра λ ?

4.16. Найти в зависимости от значения параметра λ размерность подпространства линейного пространства R^4, натянутого на векторы

$$a_1 = (1,2,1,0), a_2 = (1,3,1,0), a_3 = (2,4,1,1), a_4 = (5,3,\lambda,2).$$

При каких значениях параметра λ система векторов a_1, a_2, a_3, a_4 образует базис в R^4?

4.17. Подпространство L пространства R^4 порождается векторами

$$a_1 = (1,1,-1,1), a_2 = (2,1,0,1), a_3 = (-1,0,1,1), a_4 = (1,1,1,\lambda).$$

Найти λ, если известно, что вектор $b = (1,1,1,1)$ не входит в подпространство L. Какова в этом случае размерность L ?

4.18. Пусть L — конечномерное линейное пространство. Доказать, что для любого подпространства A пространства L найдется такое подпространство B пространства L, что $L = A \oplus B$.

5. Координаты, матрица перехода.

Проверка того, что каждая из двух данных систем векторов из K^n составляет базис, нахождение матрицы перехода от одного к другому и проверка вычислением координат конкретного вектора.

6. Матрица линейного отображения.

Вычисление матрицы линейных отображений, заданных правилами вычисления образа. Вычисление образов векторов относительно отображения, заданного матрицей.

6.1. Пусть оператор φ линейного пространства $\mathbb{R}^3$ переводит произвольный вектор (α, β, γ) этого пространства в вектор $(\alpha+\beta-2\gamma, 2\alpha-\beta+\gamma, \alpha-2\beta+3\gamma)$.

1) Найти матрицу оператора φ в стандартном базисе пространства $\mathbb{R}^3$ и вычислить двумя способами образ относительно φ вектора $a = (2, 3, -1)$, воспользовавшись исходным определением отображения φ и основным свойством матрицы линейного отображения.

2) Найти ранг и дефект, базисы образа и ядра оператора φ.

6.2. Пусть отображение $\varphi \in \mathcal{L}(L, M)$ задано указанием образов элементов базиса $S = (a_1, a_2, \dots, a_r)$ пространства L, выраженных в виде линейных комбинаций базиса $S' = (b_1, a_b, \dots, b_s)$ пространства M. Требуется выписать матрицу этого отображения.

7. Инвариантные подпространства. Собственные векторы и собственные значения линейного оператора.

7.1. Доказать, что сумма и пересечение φ-инвариантных подпространств являются φ-инвариантными подпространствами.

7.2. Пусть φ — оператор линейного пространства L. Доказать, что все подпространства пространства L инвариантны относительно φ тогда и только тогда, когда φ является гомотетией.

7.3. Пусть A — подпространство пространства L, порожденное ненулевым вектором $a \in L$. Доказать, что A инвариантно относительно оператора φ пространства L тогда и только тогда, когда вектор a является собственным вектором оператора φ.

7.4. Пусть L — линейное пространство над полем K и $\varphi \in \mathcal{L}(L)$. Если для некоторого $\lambda \in K$ подпространство A пространства L содержит $\mathrm{Im}(\varphi - \lambda\iota)$, то A инвариантно относительно φ.

7.5. Доказать, что если в n-мерном линейном пространстве L существует хотя бы один собственный вектор оператора $\varphi \in \mathcal{L}(L)$, то в L есть $(n-1)$-мерное φ-инвариантное подпространство.

7.6. Пусть φ — оператор конечномерного линейного пространства L над полем K и пусть характеристический многочлен оператора φ неприводим над K. Доказать, что произвольное φ-инвариантное подпространство пространства L совпадает либо с L, либо с нулевым подпространством.

7.7. Пусть λ — k-кратный корень характеристического многочлена оператора φ конечномерного линейного пространства L. Доказать, что если $a_1, a_2, \dots, a_m$ — линейно независимая система собственных векторов оператора φ, принадлежащих собственному значению λ, то $m \le k$.

7.8. Пусть φ — оператор линейного пространства L над полем K, a и b — собственные векторы оператора φ, принадлежащие различным собственным значениям. Доказать, что если α и β — ненулевые скаляры из K, то вектор $\alpha a + \beta b$ не является собственным вектором оператора φ.

7.9. Пусть оператор φ линейного пространства L обратим. Доказать, что всякий собственный вектор оператора φ является собственным вектором и оператора φ^{-1}.

7.10. Пусть φ — оператор пространства L над полем K и пусть a — собственный вектор оператора φ, принадлежащий собственному значению λ. Доказать, что для любого многочлена $f(x) \in K[x]$ вектор a является собственным вектором оператора φ, принадлежащим собственному значению $f(\lambda)$.

7.11. Пусть φ — оператор линейного пространства L и пусть скаляр λ^2 является собственным значением оператора φ^2. Доказать, что хотя бы один из скаляров λ и $-\lambda$ является собственным значением этого оператора.

7.12. Пусть оператор φ n-мерного линейного пространства имеет n попарно различных собственных значений $\lambda_1, \lambda_2, \dots, \lambda_n$. Доказать, что произвольный оператор ψ, перестановочный с φ, диагонализируем. Для этого показать, что произвольный собственный вектор оператора φ является собственным вектором и оператора ψ.

7.13. Пусть φ — оператор линейного пространства L и пусть $f(x)$ — минимальный аннулятор вектора $a \in L$, $g(x)$ — минимальный аннулятор вектора $b \in L$. Доказать, что если многочлены $f(x)$ и $g(x)$ взаимно просты, то минимальный аннулятор вектора $a + b$ равен произведению $f(x)g(x)$.

7.14. Доказать, что произвольный аннулятор оператора φ делится на его минимальный аннулятор.

7.15. Доказать, что оператор φ обратим тогда и только тогда, когда свободный член его минимального аннулятора отличен от нуля.

7.16. Пусть φ — оператор линейного пространства L над полем K, и $\lambda \in K$ — собственное значение оператора φ. Геометрической кратностью λ назовем размерность подпространства $M(\lambda) = \{a \in L \mid a\varphi = \lambda a\}$. Доказать, что оператор φ диагонализируем тогда и только тогда, когда когда его характеристический многочлен раскладывается над полем K на линейные множители и кратность каждого его корня λ совпадает с его геометрической кратностью.

7.17. Пусть φ — оператор линейного пространства L над полем K. Для произвольного $\lambda \in K$ полагаем $M(\lambda) = \{a \in L \mid a\varphi = \lambda a\}$. Доказать, что

1) $M(\lambda)$ является подпространством пространства L;

2) для любого оператора ψ, перестановочного с φ, $M(\lambda)$ инвариантно относительно ψ;

3) если $\lambda_1 \neq \lambda_2$, то $M(\lambda_1 \cap M(\lambda_2) = \{0\}$.

8. Жорданова нормальная форма матрицы линейного оператора.

Корневые подпространства, нахождение базисов корневых подпространств. Вычисление жордановой матрицы линейного оператора и жорданова базиса пространства.

9. Линейные и билинейные функции

9.1. Пусть функция $f(x,y)$ на линейном пространстве $\mathbb{R}^3$ определена по правилу $f(a,b) = \alpha_1\beta_1 + 2\alpha_2\beta_2 - \alpha_3\beta_3$, где $a = (\alpha_1,\alpha_2,\alpha_3), b = (\beta_1,\beta_2,\beta_3)$ — произвольные элементы из $\mathbb{R}^3$. Доказать, что функция $f(x,y)$ билинейна. Найти ее матрицу в естественном базисе пространства $\mathbb{R}^3$.

9.2. Пусть $f(x,y)$ — билинейная функция на линейном пространстве L, φ — линейный оператор этого пространства. Доказать, что функция $g(x,y)$, определенная на пространстве L по правилу $g(a,b) = f(a\varphi, b)$ $(a, b \in L)$, билинейна.

9.3. Пусть $u(x)$ и $v(x)$ — линейные функционалы на линейном пространстве L. Доказать, что функция $f(x,y)$, определенная на L по правилу $f(a,b) = u(a)\cdot v(b)$ $(a, b \in L)$, билинейна.

9.4. Пусть $f(x,y)$ — симметричная билинейная функция на линейном пространстве L. Доказать, что для любых подпространств M и N пространства L

1) $(M+N)^\perp = M^\perp \cap N^\perp$;
2) $(M\cap N)^\perp \supseteq M^\perp + N^\perp$;
3) $M \subseteq (M^\perp)^\perp$.

9.5. Пусть $f(x,y)$ — билинейная функция, заданная в базисе e_1, e_2 пространства L матрицей $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Пусть M и N — подпространства пространства L, порожденные векторами e_1 и e_2 соответственно. Найти подпространства $M^\perp$, $N^\perp$ и показать, что $(M\cap N)^\perp \neq M^\perp + N^\perp$ и $M \neq (M^\perp)^\perp$.

9.6. Билинейная функция $f(x,y)$ задана в естественном базисе пространства $\mathbb{R}^4$ матрицей

$$\begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & 0 & 1 & -1 \\ 0 & 1 & -1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix}.$$

Найти ортогональный базис пространства и вычислить сигнатуру функции $f(x,y)$.

9.7. Пусть $f(x,y)$ — симметричная билинейная функция на линейном пространстве L. Доказать, что если ортогональная система $a_1, a_2, \dots a_n$ векторов пространства L линейно зависима, то хотя бы один из векторов этой системы изотропен.

9.8. Пусть L — конечномерное линейное пространство над полем K. Для произвольных подмножеств $A \subseteq L$ и $F \subseteq L^*$ полагаем

$$N^0(A) = \{f \in L^* \mid (\forall a \in A)(f(a) = 0)\} \text{ и } N^*(F) = \{a \in L \mid (\forall f \in F)(f(a) = 0)\}.$$

1) Доказать, что для любых $A \subseteq L$ и $F \subseteq L^*$ $N^0(A)$ и $N^*(F)$ — подпространства пространств L^* и L соответственно.

2) Доказать, что если A и F — подпространства пространств L и L^* соответственно, то $N^*(N^0(A)) = A$ и $N^0(N^*(F)) = F$.

9.9. Предполагая обозначения предыдущей задачи, доказать, что для произвольных подпространств A и B конечномерного пространства L и произвольных подпространств F и G пространства L^*

1) $N^0(A+B) = N^0(A) \cap N^0(B)$ и $N^*(F+G) = N^*(F) \cap N^*(G)$.

2) $\dim A + \dim N^0(A) = \dim L$ и $\dim F + \dim N^*(F) = \dim L$.

9.10. Пусть L и M — конечномерные линейные пространства над полем K и $\varphi : L \to M$ — линейное отображение. Для любого функционала $f \in M^*$ определим отображение $f^\varphi : L \to K$, полагая для произвольного вектора $a \in L$ $f^\varphi(a) = f(a\varphi)$. Доказать, что

1) $f^\varphi \in L^*$;

2) отображение $\varphi^* : M^* \to L^*$, сопоставляющее функционалу $f \in M^*$ функционал $f^\varphi \in L^*$, линейно;

3) отображение $\sigma : \mathcal{L}(L, M) \to \mathcal{L}(M^*, L^*)$, сопоставляющее линейному отображению $\varphi \in \mathcal{L}(L, M)$ отображение $\varphi^* \in \mathcal{L}(M^*, L^*)$, линейно.

Пусть A — матрица отображения $\varphi \in \mathcal{L}(L, M)$ в базисах S_1 и S_2 пространств L и M соответственно. Найти матрицу отображения φ^* в сопряженных базисах S_2^* и S_1^*.

9.11. Пусть $f(x, y)$ — билинейная функция на линейном пространстве L. Для произвольного вектора $a \in L$ определим отображение f_a пространства L в поле K, полагая для $x \in L$ $f_a(x) = f(a, x)$.

1) Доказать, что $f_a \in L^*$.

2) Показать, что отображение $\varphi : L \to L^*$, сопоставляющее вектору $a \in L$ функционал f_a, является линейным.

3) Если A — матрица функции $f(x, y)$ в базисе S пространства L, то A является и матрицей отображения φ в базисах S и S^* пространств L и L^*.

9.12. В обозначениях задач 11 и 8 показать, что для произвольного подпространства M пространства L выполнено равенство $M^\perp = N^*(M\varphi)$.

Ядро отображения φ называется ядром функции $f(x, y)$. Функция $f(x, y)$ называется невырожденной, если ее ядро является нулевым подпространством.

9.13. Пусть $f(x, y)$ — симметричная билинейная функция на линейном пространстве L. Показать, что для любых подпространств A и B пространства L

1) $(A^\perp)^\perp = A$;

2) $(A \cap B)^\perp = A^\perp + B^\perp$.