Лекции 1-21
по Линейной алгебре и геометрии
Лектор: О.В. Куликова (1 курс, 2 поток, 2020 год)

1. Векторные пространства

1.1. Основные определения

Для полноты изложения начнем с напоминания определений и утверждений из 1-го семестра (см. [1] гл. 1-3).

Определение 1.1. *Множество V называется* **векторным** *(или* **линейным) пространством** *над полем F, если*
а) каждым двум элементам $x, y \in V$ поставлен в соответствие элемент $z \in V$, называемый суммой элементов x и y; сумма элементов x и y обозначается через $x + y$,
б) каждому элементу $x \in V$ и каждому элементу $\lambda \in F$ поставлен в соответствие элемент $\lambda x \in V$, называемый произведением элемента λ на элемент x,
в) эти операции (линейные операции) удовлетворяют следующим требованиям (аксиомам):
1) $x + y = y + x$ для любого $x, y \in V$ (коммутативность);
2) $x + (y + z) = (x + y) + z$ для любых $x, y, z \in V$ (ассоциативность);
3) в V существует такой элемент 0 (нуль), что $x + 0 = x$ для любого $x \in V$;
4) для любого элемента $x \in V$ существует такой элемент $-x \in V$ (противоположный элемент), что $x + (-x) = 0$;
5) $\lambda(x + y) = \lambda x + \lambda y$ для любых $\lambda \in F$, $x, y \in V$;
6) $(\lambda + \mu)x = \lambda x + \mu x$ для любых $\lambda, \mu \in F$, $x \in V$;
7) $(\lambda\mu)x = \lambda(\mu x)$ для любых $\lambda, \mu \in F$, $x \in V$;
8) $1x = x$ для любого $x \in V$.

Отметим, что аксиомы 1)-4) векторного пространства V говорят о том, что относительно сложения V является абелевой группой.

Элементы векторного пространства называются *векторами.* То обстоятельство, что это слово часто употребляется в более узком смысле, не должно нас смущать. Геометрические представления, связанные с этим словом, помогут нам уяснить, а иногда и предвидеть, ряд результатов. Элементы поля F, в отличие от векторов, мы будем иногда, допуская вольность речи, называть числами, даже если F не является числовым полем.

ПРИМЕР 1. Векторы в смысле элементарной геометрии будем отныне называть *геометрическими векторами.* Операции над ними удовлетворяют всем аксиомам векторного пространства, что, собственно, и послужило основой для данного выше определения. Пространство геометрических векторов евклидовой плоскости (соотв., трехмерного евклидова пространства) будем обозначать через V^2 (соотв., V^3). Подчеркнем, что это векторное пространство над полем $\mathbb{R}$.

ПРИМЕР 2. Множество $M_{m\times n}(F)$ прямоугольных матриц размера $m \times n$ с коэффициентами из поля F является векторным пространством над F относительно стандартных операций сложения матриц и умножения матриц на число.

ПРИМЕР 3. Множество F^n упорядоченных последовательностей $(x_1, \dots, x_n)$ длины n с элементами x_i из поля F является векторным пространством над F относительно операций:

$$(a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n),$$

$$\lambda(a_1, a_2, \dots, a_n) = (\lambda a_1, \lambda a_2, \dots, \lambda a_n).$$

Это векторное пространство, по сути, есть векторное пространство матриц-строк $M_{1\times n}(F)$. Отличие лишь формальное, так как первое определено как множество упорядоченных последовательностей чисел, а второе как множество матриц. Но элементы матрицы также записываются в определенном порядке. Ниже F^n часто будем называть пространством строк длины n.

ПРИМЕР 4. Множество F^M всех функций на множестве M со значениями в поле F является векторным пространством над F относительно обычных операций над функциями:

$$(f + g)(x) = f(x) + g(x),$$

$$(\lambda f)(x) = \lambda f(x).$$

ПРИМЕР 5. Множество $C[a, b]$ непрерывных функций на отрезке $[a, b]$ образует векторное пространство над $\mathbb{R}$ относительно обычных операций над функциями.

Но множество непрерывных функций на отрезке $[a, b]$ таких, что $|f(x)| \leqslant 1$ не будет образовывать векторное пространство: из того, что $|f_1(x)| \leqslant 1$ и $|f_2(x)| \leqslant 1$, не следует $|f_1(x) + f_2(x)| \leqslant 1$.

ПРИМЕР 6. Множество $F_n[x]$ всех многочленов степени, не превышающей натурального числа n, с обычными операциями сложения многочленов и умножения их на числа образует векторное пространство над F.

Заметим, что множество многочленов степени, равной n, не образует векторного пространства, так как сумма двух многочленов степени n может оказаться многочленом более низкой степени: например, $(t^n + t) + (-t^n + t) = 2t$.

ПРИМЕР 7. Множество $F[x_1, \dots, x_n]$ всех многочленов от n переменных с коэффициентами из поля F является векторным пространством над F относительно обычных операций над многочленами.

ПРИМЕР 8. Пусть F – подполе поля L. Тогда L можно рассматривать как векторное пространство над F, определив умножение элементов из L на элементы из F просто как умножение в L. В частности, поле $\mathbb{C}$ есть в этом смысле векторное пространство над $\mathbb{R}$.

Укажем некоторые следствия из аксиом векторного пространства. Их доказательство проводилось в 1-м семестре. Символом 0 обозначается как нуль поля F, так и нулевой вектор, т.е. нуль аддитивной группы V (смотри третью аксиому векторного пространства), это не приводит к путанице.

Следствия из аксиом векторного пространства:

1) Вектор нуль единствен.

2) Противоположный вектор единствен.

3) Для любых $a, b \in V$ уравнение $x + a = b$ имеет единственное решение, равное $b + (-a)$. Это решение называется *разностью* векторов b и a и обозначается $b - a$.

4) Сумма произвольного числа (а не только трех) векторов не зависит от расстановки скобок. (Пользуясь этим свойством, скобки обычно вообще опускают.)

5) $\lambda 0 = 0$ для любого $\lambda \in F$ (здесь 0 – нулевой вектор).

6) $\lambda(-a) = -\lambda a$ для любых $\lambda \in F$, $a \in V$.

7) $\lambda(a - b) = \lambda a - \lambda b$ для любых $\lambda \in F$, $a, b \in V$.

8) $0a = 0$ для любого $a \in V$ (здесь 0 слева – число, справа – вектор).

9) $(-1)a = -a$ для любого $a \in V$.

10) $(\lambda - \mu)a = \lambda a - \mu a$ для любых $\lambda, \mu \in F$, $a \in V$.

Определение 1.2. *Непустое подмножество U векторного пространства V над полем F называется* **подпространством**, *если оно замкнуто относительно сложения и относительно умножения на произвольные числа из F, т.е.*

1) $\forall x, y \in U \Rightarrow x + y \in U$,

2) $\forall \lambda \in F\ \forall x \in U \Rightarrow \lambda x \in U$.

Замечание. *В определении подпространства вместо условия непустоты множества U и наличия условия 1) можно потребовать, чтобы U являлось подгруппой аддитивной группы V. Действительно, в определении подгруппы требуется, чтобы*

$$x, y \in U \Rightarrow x + y \in U, \qquad x \in U \Rightarrow -x \in U \qquad \text{и} \qquad 0 \in U.$$

Первое из этих условий совпадает с условием 1) из определения подпространства. При наличии условия 2) из определения подпространства условие $x \in U \Rightarrow -x \in U$ выполняется автоматически, так как $(-1)x = -x$. При наличии того же условия 2) условие $0 \in U$ равносильно тому, что

U не пусто, так как если существует хотя бы один вектор x в U, то $0 = 0x \in U$.

Подпространство векторного пространства само является векторным пространством относительно тех же операций. Это дает множество новых примеров векторных пространств.

ПРИМЕР 9. В пространстве V^3 геометрических векторов множество векторов, параллельных заданной плоскости или прямой, является подпространством.

ПРИМЕР 10. В пространстве $\mathbb{R}^{[a,b]}$ всех функций на заданном отрезке $[a, b]$ числовой прямой множество непрерывных функций является подпространством.

ПРИМЕР 11. В пространстве F^n строк длины n множество строк с нулями на четных местах образует подпространство, а множество с единицами на четных местах не образует подпространства.

ПРИМЕР 12. В пространстве квадратных матриц $M_n(F)$ порядка n множество матриц с нулями под главной диагональю образует подпространство. Также в пространстве $M_n(F)$ множество симметрических (соотв., кососимметрических) матриц, т.е. матриц A таких, что $A = A^T$ (соотв., $A = -A^T$), образует подпространство.

ПРИМЕР 13. В пространстве $F[x_1, \dots, x_n]$ всех многочленов от n переменных с коэффициентами из поля F множество симметрических многочленов, т.е. таких, которые не изменяются ни при каких перестановках переменных, образует подпространство.

В каждом векторном пространстве V есть два "тривиальных" подпространства: само пространство V и нулевое подпространство (состоящее из одного нулевого вектора).

Основным понятием теории векторных пространств является понятие линейной зависимости векторов, также подробно изученное в 1-м семестре.

Определение 1.3. *Конечная система векторов $\{v_1, \dots, v_m\} \subset V$ называется* **линейно зависимой**, *если найдутся не все равные нулю числа $\lambda_1, \dots, \lambda_m$ такие, что $\lambda_1 v_1 + \ldots + \lambda_m v_m = 0$. Бесконечное подмножество $S \subset V$ называется* **линейно зависимым**, *если некоторое непустое конечное подмножество $\{v_1, \dots, v_m\} \subset S$ является линейно зависимым. Подмножество, не являющееся линейно зависимым, называется* **линейно независимым**.

Определение 1.4. *Всякое выражение вида $\lambda_1 v_1 + \ldots + \lambda_m v_m$, где $\lambda_1, \dots, \lambda_m \in F$, называется* **линейной комбинацией** *векторов $\{v_1, \dots, v_m\} \subset V$. Если* *не все* $\lambda_1, \ldots, \lambda_m$ *равны нулю, то линейная комбинация* $\lambda_1 v_1 + \ldots + \lambda_m v_m$ *называется* **нетривиальной**, *в противном случае линейная комбинация называется* **тривиальной**. *Говорят, что вектор* $w \in V$ **линейно выражается** *через векторы* $\{v_1, \ldots, v_m\}$, *если он равен некоторой их линейной комбинации. Говорят, что вектор* $w \in V$ **линейно выражается** *через некоторое подмножество* $S \subset V$, *если* w *линейно выражается через некоторое конечное подмножество* $\{v_1, \ldots, v_m\} \subset S$.

Очевидно, что если система векторов содержит линейно зависимую подсистему, то она сама линейно зависима.

Заметим, что понятие системы векторов отличается от понятия множества векторов тем, что, во-первых, векторы системы предполагаются занумерованными и, во-вторых, среди них могут быть равные. Однако, свойство системы векторов быть линейно зависимой или независимой не зависит от нумерации векторов в ней.

В 1-м семестре были доказаны следующие леммы о линейно зависимых и линейно независимых системах.

Лемма 1.1. *(Критерий линейной зависимости)*

Векторы $v_1, \ldots, v_m$ $(m > 1)$ *линейно зависимы тогда и только тогда, когда хотя бы один из них линейно выражается через остальные.*

Лемма 1.2. *Пусть векторы* $v_1, \ldots, v_m$ *линейно независимы. Вектор* w *линейно выражается через* $v_1, \ldots, v_m$ *тогда и только тогда, когда векторы* $v_1, \ldots, v_m, w$ *линейно зависимы.*

Лемма 1.3. *Пусть вектор* w *линейно выражается через векторы* $v_1, \ldots, v_m$. *Это выражение единственно тогда и только тогда, когда векторы* $v_1, \ldots, v_m$ *линейно независимы.*

Пусть $S \subset V$ – какое-то подмножество. Совокупность всевозможных (конечных) линейных комбинаций векторов из S называется **линейной оболочкой** этого подмножества и будет обозначаться через $\langle S \rangle$. Это наименьшее подпространство пространства V, содержащее S, т.е. любое подпространство, содержащее S, содержит и всю линейную оболочку $\langle S \rangle$. Говорят, что пространство V **порождается** множеством S, если $\langle S \rangle = V$. Векторное пространство называется **конечномерным**, если оно порождается конечным числом векторов, и **бесконечномерным** в противном случае.

Бесконечномерные пространства составляют предмет специального изучения. В данном курсе мы будем заниматься в основном конечномерными пространствами.

Лемма 1.4. *(Основная лемма о линейной зависимости)*

Если векторное пространство V *порождается* n *векторами, то всякие* $m > n$ *векторов пространства* V *линейно зависимы.*

Определение 1.5. *Система векторов* $\{e_1, \dots, e_n\} \subset V$ *называется* **базисом** *векторного пространства* V, *если каждый вектор* $x \in V$ *единственным образом выражается через* $e_1, \dots, e_n$. *Коэффициенты этого выражения называются* **координатами** *вектора* x *в базисе* $\{e_1, \dots, e_n\}$.

Ввиду леммы 1.3 это определение базиса можно переформулировать следующим образом: **Базисом** векторного пространства V называется всякая линейно независимая система векторов, порождающая пространство V.

В 1-м семестре была доказана следующая теорема.

Теорема 1.1. *Всякое конечномерное векторное пространство* V *обладает базисом. Более точно, из всякого конечного порождающего множества* $S \subset V$ *можно выбрать базис пространства* V.

Также из основной леммы о линейной зависимости была получена:

Теорема 1.2. *Все базисы конечномерного векторного пространства* V *имеют одно и тоже число векторов.*

Это число называется **размерностью** пространства V и обозначается $\dim V$.

ПРИМЕР 14. Пространство V^2 (соотв., V^3) имеет размерность 2 (соотв., 3).

ПРИМЕР 15. В качестве базиса пространства F^n можно выбрать единичные строки $(1, 0, 0, \dots, 0, 0), (0, 1, 0, \dots, 0, 0), \dots, (0, 0, 0, \dots, 0, 1)$ и $\dim F^n = n$.

ПРИМЕР 16. В пространстве матриц $M_{m\times n}(F)$ размера $m \times n$ в качестве базиса можно выбрать все матрицы, у которых на одном каком-либо месте стоит единица, а на остальных местах нули. Таким образом, $\dim M_{m\times n}(F) = mn$.

ПРИМЕР 17. Поле комплексных чисел как векторное пространство над $\mathbb{R}$ имеет базис $\{1, i\}$ и, соответственно, размерность 2. Алгебра кватернионов как векторное пространство над $\mathbb{R}$ имеет базис $\{1, i, j, k\}$ и, соответственно, размерность 4.

ПРИМЕР 18. Поле $\mathbb{R}$ как векторное пространство над $\mathbb{Q}$ бесконечномерно. В самом деле, если бы оно было конечномерным, то вещественное число определялось бы конечным набором рациональных чисел – своих координат в некотором базисе этого пространства. Но тогда множество всех вещественных чисел было бы счетным, что неверно.

ПРИМЕР 19. В пространстве $F_n[x]$ многочленов степени $\leqslant n$ простейшим базисом является совокупность векторов $1, x, x^2, \dots, x^n$ и $\dim F_n[x] = n + 1$.

ЗАДАЧА. Доказать,что пространство всех непрерывных функций на любом промежутке числовой прямой бесконечномерно.

Из основной леммы о линейной зависимости следует, что в любом (конечном или бесконечном) множестве S векторов конечномерного векторного пространства V имеется максимальное линейно независимое подмножество, т.е.

такое линейно независимое подмножество, которое становится линейно зависимым при добавлении к нему любого вектора из оставшихся векторов множества S. Более того, любое линейно независимое подмножество множества S можно дополнить до максимального линейно независимого подмножества.

Лемма 1.5. *Всякое максимальное линейно независимое подмножество множества S конечномерного векторного пространства V является базисом линейной оболочки $\langle S \rangle$ этого множества.*

В частности, применяя соображения, изложенные выше, к $S = V$, получаем, что базисом конечномерного векторного пространства V является всякая максимальная линейно независимая система векторов из V. Также получаем следующую теорему.

Теорема 1.3. *Всякую линейно независимую систему векторов конечномерного векторного пространства V можно дополнить до базиса.*

В частности, любой ненулевой вектор можно включить в базис, а любые n линейно независимых векторов векторного пространства размерности n уже составляют базис.

Напомним также следующую теорему из 1-го семестра, устанавливающую монотонность размерности.

Теорема 1.4. *Всякое подпространство U конечномерного векторного пространства V также конечномерно, причем $\dim U \leqslant \dim V$. Более того, если $U \neq V$, то $\dim U < \dim V$.*

1.2. Координаты векторов. Преобразование координат при переходе к другому базису.

Пусть V – векторное пространство над полем F, $\dim V = n$. И пусть $\mathcal{E} = \{e_1, \ldots, e_n\}$ – какой-либо фиксированный базис. Тогда для любого вектора $v \in V$ существуют числа $\lambda_1, \ldots, \lambda_n \in F$ такие, что $v = \lambda_1 e_1 + \ldots + \lambda_n e_n$. Набор чисел $(\lambda_1, \ldots, \lambda_n)$ есть набор координат вектора v, причем при заданном базисе координаты вектора v определены однозначно.

Если в базисе $\mathcal{E}$ вектор x имеет координаты $(x_1, \ldots, x_n)$, а вектор y – координаты $(y_1, \ldots, y_n)$, т.е.

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n,$$

$$y = y_1 e_1 + y_2 e_2 + \ldots + y_n e_n,$$

то

$$x + y = (x_1 + y_1) e_1 + (x_2 + y_2) e_2 + \ldots + (x_n + y_n) e_n,$$

т.е. вектор $x + y$ имеет координаты $(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$. Аналогично получаем, что вектор λx имеет координаты $(\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$.

Таким образом, при сложении векторов x и y их координаты складываются. При умножении вектора x на число λ его координаты умножаются на это число.

Поэтому столбец (соотв., строка) координат линейной комбинации векторов есть линейная комбинация их столбцов (соотв., строк) координат с теми же коэффициентами. Отсюда следует, что *векторы линейно зависимы тогда и только тогда, когда линейно зависимы их столбцы (соотв., строки) координат.*

ПРИМЕР 1. Рассмотрим пространство F^n строк длины n. Возьмем базис

$$e_1 = (1, 1, 1, \ldots, 1),$$

$$e_2 = (0, 1, 1, \ldots, 1),$$

$$\ldots\ldots\ldots\ldots\ldots\ldots$$

$$e_n = (0, 0, 0, \ldots, 1).$$

Найдем координаты $(\chi_1, \chi_2, \ldots, \chi_n)$ вектора $x = (x_1, x_2, \ldots, x_n)$ в этом базисе. По определению

$$x = \chi_1 e_1 + \chi_2 e_2 + \ldots + \chi_n e_n,$$

т.е. $(x_1, x_2, \ldots, x_n) = \chi_1(1, 1, 1, \ldots, 1) + \chi_2(0, 1, 1, \ldots, 1) + \ldots + \chi_n(0, 0, \ldots, 1) =$ $= (\chi_1, \chi_1 + \chi_2, \ldots, \chi_1 + \chi_2 + \ldots + \chi_n)$. Таким образом, числа $\chi_1, \chi_2, \ldots, \chi_n$ находятся из следующей системы линейных уравнений:

$$\begin{cases} \chi_1 = x_1, \\ \chi_1 + \chi_2 = x_2, \\ \quad . \quad . \quad . \quad . \quad . \\ \chi_1 + \chi_2 + \ldots + \chi_n = x_n, \end{cases}$$

откуда $\chi_1 = x_1, \chi_2 = x_2 - x_1, \ldots, \chi_n = x_n - x_{n-1}$.

Рассмотрим теперь в F^n базис из единичных строк:

$$e'_1 = (1, 0, 0, \ldots, 0),$$

$$e'_2 = (0, 1, 0, \ldots, 0),$$

$$\ldots\ldots\ldots\ldots\ldots\ldots$$

$$e'_n = (0, 0, 0, \ldots, 1).$$

В этом базисе связь между координатами вектора $x = (x_1, x_2, \ldots, x_n)$ и числами $x_1, x_2, \ldots, x_n$, определяющими этот вектор, наиболее проста:
$x = (x_1, x_2, \ldots, x_n) = x_1(1, 0, 0, \ldots, 0) + x_2(0, 1, 0, \ldots, 0) + \ldots + x_n(0, 0, 0, \ldots, 1) =$

$= x_1 e'_1 + x_2 e'_2 + \ldots + x_n e'_n$, т.е. в F^n числа $x_1, x_2, \ldots, x_n$ можно трактовать как координаты вектора $x = (x_1, x_2, \ldots, x_n)$ в базисе $e'_1 = (1, 0, 0, \ldots, 0)$, $e'_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $e'_n = (0, 0, 0, \ldots, 1)$.

ПРИМЕР 2. Рассмотрим пространство $\mathbb{R}_n[x]$ многочленов степени $\leqslant n$. Координатами многочлена $f(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_n x^n$ в простейшем базисе $e_1 = 1, e_2 = x, e_3 = x^2, \ldots, e^{n+1} = x^n$ являются, как легко видеть, его коэффициенты $(\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_n)$.

Выберем теперь другой базис:

$$e'_1 = 1, e'_2 = (x-a), e'_2 = (x-a)^2, \ldots, e'_{n+1} = (x-a)^n.$$

Каждый многочлен $f(x)$ может быть представлен по формуле Тейлора в виде:

$$f(x) = f(a) + f'(a)(x-a) + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Таким образом, в этом базисе $f(x)$ имеет координаты $(f(a), f'(a), \ldots, \frac{f^{(n)}(a)}{n!})$.

Как видно из этих примеров, один и тот же вектор в разных базисах может иметь разные координаты. Нашей ближайшей целью является установление связи между координатами одного и того же вектора относительно двух различных базисов.

Пусть даны два базиса $\mathcal{E} = \{e_1, \ldots, e_n\}$ и $\mathcal{E}' = \{e'_1, \ldots, e'_n\}$ векторного пространства V. Выразим каждый вектор $e'_j \in \mathcal{E}'$ через базис $\mathcal{E}$:

$$e'_j = t_{1,j}e_1 + t_{2,j}e_2 + \ldots + t_{n,j}e_n, \ j = 1, 2, \ldots, n.$$

Из коэффициентов $t_{i,j} \in F$ составим квадратную матрицу $T_{\mathcal{E}\to\mathcal{E}'} = (t_{i,j})$. При этом j-й столбец матрицы $T_{\mathcal{E}\to\mathcal{E}'}$ есть столбец координат вектора e'_j в базисе $\mathcal{E}$. Полученная матрица $T_{\mathcal{E}\to\mathcal{E}'}$ называется **матрицей перехода** от базиса $\mathcal{E}$ к базису $\mathcal{E}'$. Отметим, что $\det T_{\mathcal{E}\to\mathcal{E}'} \neq 0$, так как столбцами матрицы $T_{\mathcal{E}\to\mathcal{E}'}$ являются координаты линейно независимых векторов $\{e'_1, \ldots, e'_n\}$.

Рассмотрим произвольный вектор $x \in V$. Пусть $(x_1, \ldots, x_n)$ – его координаты в базисе $\mathcal{E}$, $(x'_1, \ldots, x'_n)$ – его координаты в базисе $\mathcal{E}'$, т.е.

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n \qquad (*)$$

и

$$x = x'_1e'_1 + x'_2e'_2 + \ldots + x'_ne'_n.$$

Подставляя в правую часть последнего равенства вместо каждого e'_j его выражение через базис $\mathcal{E}$, получаем

$$x = x'_1(t_{1,1}e_1 + t_{2,1}e_2 + \ldots + t_{n,1}e_n) + x'_2(t_{1,2}e_1 + t_{2,2}e_2 + \ldots + t_{n,2}e_n) + \ldots +$$
$$+ x'_n(t_{1,n}e_1 + t_{2,n}e_2 + \ldots + t_{n,n}e_n).$$

Далее раскрываем скобки в полученной правой части и собираем вместе коэффициенты при каждом векторе e_i:

$$x = (x'_1t_{1,1} + x'_2t_{1,2} + \ldots + x'_nt_{1,n})e_1 + (x'_1t_{2,1} + x'_2t_{2,2} + \ldots + x'_nt_{2,n})e_2 + \ldots +$$
$$+(x'_1t_{n,1} + x'_2t_{n,2} + \ldots + x'_nt_{n,n})e_n. \quad (**)$$

Каждое из равенств $(*)$ и $(**)$ является разложением вектора x по базису $\mathcal{E}$. В силу единственности разложения вектора по базису получаем, что

$$\begin{cases} x_1 = x'_1 t_{1,1} + x'_2 t_{1,2} + \ldots + x'_n t_{1,n}, \\ x_2 = x'_1 t_{2,1} + x'_2 t_{2,2} + \ldots + x'_n t_{2,n}, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ x_n = x'_1 t_{n,1} + x'_2 t_{n,2} + \ldots + x'_n t_{n,n}, \end{cases}$$

или в матричном виде

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} t_{1,1} & t_{1,2} & \ldots & t_{1,n} \\ t_{2,1} & t_{2,2} & \ldots & t_{2,n} \\ \cdot & \cdot & \cdot & \cdot \\ t_{n,1} & t_{n,2} & \ldots & t_{n,n} \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}.$$

Положим

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad X' = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}.$$

Тогда полученную формулу преобразования координат при переходе от базиса $\mathcal{E}$ к базису $\mathcal{E}'$ можно записать так:

$$X = T_{\mathcal{E}\to\mathcal{E}'} X'. \tag{1}$$

Свойства матриц перехода:

1) Пусть $\mathcal{E}, \mathcal{E}'$ – два базиса векторного пространства V. Тогда

$$T_{\mathcal{E}'\to\mathcal{E}} = T^{-1}_{\mathcal{E}\to\mathcal{E}'}.$$

2) Пусть $\mathcal{E}, \mathcal{E}', \mathcal{E}''$ – три базиса векторного пространства V. Тогда

$$T_{\mathcal{E}\to\mathcal{E}''} = T_{\mathcal{E}\to\mathcal{E}'} T_{\mathcal{E}'\to\mathcal{E}''}.$$

3) Пусть $T = (t_{i,j})$ – произвольная квадратная невырожденная матрица, $\mathcal{E}$ – произвольный базис векторного пространства V. Тогда система векторов $\{e'_1, \ldots, e'_n\}$, полученных по формуле:

$$e'_j = t_{1,j} e_1 + t_{2,j} e_2 + \ldots + t_{n,j} e_n, \ j = 1, 2, \ldots, n,$$

будет также являться базисом в V. Другими словами, любая квадратная невырожденная матрица T является матрицей перехода из данного базиса $\mathcal{E}$ векторного пространства V в некоторый другой базис $\mathcal{E}'$.

ЗАДАЧА. Докажите эти свойства.

1.3. Изоморфизм векторных пространств.

Для сравнения двух векторных пространств V и W над одним и тем же полем F вводится понятие изоморфизма.

Определение 1.6. *Векторные пространства V и W над полем F называются* **изоморфными**, *если существует биективное отображение*

$$\varphi : V \to W,$$

сохраняющее операции сложения и умножения на число:

1) $\varphi(x+y) = \varphi(x) + \varphi(y)$ для любых $x, y \in V$,

2) $\varphi(\lambda x) = \lambda \varphi(x)$ для любых $\lambda \in F, x \in V$.

В этом случае пишут $V \cong W$. Само отображение φ называется при этом **изоморфизмом** *пространств V и W.*

Отметим некоторые **свойства изоморфизма**:

(i) $\varphi(0) = 0$, *причем в нуль пространства W переходит только нуль пространства V.*

Доказательство. Свойство (i) следует из биективности и из 2-й аксиомы изоморфизма (см. определение 1.6), если положить $\lambda = 0$. □

(ii) *Существует обратное отображение $\varphi^{-1} : W \to V$, причем оно также является изоморфизмом.*

Доказательство. Существование $\varphi^{-1} : W \to V$ следует из биективности отображения $\varphi : V \to W$. Проверим аксиомы изоморфизма (см. определение 1.6). Для произвольных $x', y' \in W$ найдутся $x, y \in V$ такие, что

$$\varphi(x) = x', \quad \varphi(y) = y',$$

а значит,

$$\varphi(x+y) = x' + y' \quad \text{и} \quad \varphi(\lambda x) = \lambda x'$$

для произвольного $\lambda \in F$. По определению обратного отображения имеем

$$\varphi^{-1}(x') = x, \quad \varphi^{-1}(y') = y, \quad \varphi^{-1}(x'+y') = x+y$$

и

$$\varphi^{-1}(\lambda x') = \lambda x.$$

Следовательно,

$$\varphi^{-1}(x'+y') = \varphi^{-1}(x') + \varphi^{-1}(y') \quad \text{и} \quad \varphi^{-1}(\lambda x') = \lambda \varphi^{-1}(x').$$

□

(iii) *Пусть U, V, W – векторные пространства над одним и тем же полем, причем $U \cong V$ и $V \cong W$. Тогда $U \cong W$.*

Проверьте свойство (iii) самостоятельно!

Как мы увидим ниже, описание векторных пространств с точностью до изоморфизма весьма просто. В частности, верно следующее утверждение

Утверждение 1.1. *Всякое векторное пространство V над полем F размерности n изоморфно пространству F^n строк длины n.*

Доказательство. Пусть $\{e_1, \dots, e_n\}$ – базис пространства V. Рассмотрим отображение

$$\varphi : V \to F^n,$$

ставящее в соответствие каждому вектору $x = x_1e_1 + \ldots + x_ne_n$ строку его координат $(x_1, \dots, x_n)$ в базисе $\{e_1, \dots, e_n\}$. В силу единственности разложения вектора по базису это отображение корректно определено. Очевидно, что это биективное отображение.

Проверим свойства 1) и 2) определения изоморфизма. Если

$$x = x_1e_1 + \ldots + x_ne_n,$$

$$y = y_1e_1 + \ldots + y_ne_n,$$

то

$$x + y = (x_1 + y_1)e_1 + \ldots + (x_n + y_n)e_n,$$
$$\lambda x = (\lambda x_1)e_1 + \ldots + (\lambda x_n)e_n.$$

Значит,

$$\varphi(x + y) = (x_1 + y_1, \dots, x_n + y_n) = (x_1, \dots, x_n) + (y_1, \dots, y_n) = \varphi(x) + \varphi(y),$$
$$\varphi(\lambda x) = (\lambda x_1, \dots, \lambda x_n) = \lambda(x_1, \dots, x_n) = \lambda\varphi(x).$$

Отсюда следует, что φ – изоморфизм. ■

Следующая теорема дает исчерпывающее описание всех конечномерных векторных пространств.

Теорема 1.5. *Конечномерные векторные пространства над одним и тем же полем изоморфны тогда и только тогда, когда они имеет одинаковую размерность.*

Доказательство. Пусть $\varphi : V \to W$ – изоморфизм векторных пространств V и W и $\{e_1, \dots, e_n\}$ – базис пространства V. Проверим, что $\{\varphi(e_1), \dots, \varphi(e_n)\}$ – базис пространства W, откуда будет следовать, что $\dim\, V = \dim\, W$.

1) Система векторов $\{\varphi(e_1), \dots, \varphi(e_n)\}$ линейно независима.

Действительно, допустим, что это не так. Тогда найдутся не все равные нулю числа $\lambda_1, \dots, \lambda_n \in F$ такие, что

$$\lambda_1\varphi(e_1) + \ldots + \lambda_n\varphi(e_n) = 0.$$

Так как изоморфизм сохраняет операции, получаем

$$\varphi(\lambda_1 e_1 + \ldots + \lambda_n e_n) = 0.$$

А так как при изоморфизме в нуль переходит только нуль, отсюда следует, что

$$\lambda_1 e_1 + \ldots + \lambda_n e_n = 0,$$

что противоречит линейной независимости базиса $\{e_1, \ldots, e_n\}$.

2) Любой вектор $x' \in W$ линейно выражается через $\{\varphi(e_1), \ldots, \varphi(e_n)\}$.

В самом деле, рассмотрим $x = \varphi^{-1}(x') \in V$. Этот вектор раскладывается по базису $\{e_1, \ldots, e_n\}$ пространства V:

$$x = x_1 e_1 + \ldots + x_n e_n.$$

Применяя к этому равенству изоморфизм φ и пользуясь тем, что изоморфизм сохраняет операции, получаем

$$x' = \varphi(x) = \varphi(x_1 e_1 + \ldots + x_n e_n) = x_1 \varphi(e_1) + \ldots + x_n \varphi(e_n).$$

Обратно, согласно Утверждению 1.1, всякое n-мерное векторное пространство над полем F изоморфно F^n; следовательно, все такие пространства изоморфны между собой. ■

Таким образом, с точностью до изоморфизма для любого n есть только одно пространство размерности n, например, F^n. И в любом рассуждении мы вправе заменить произвольное n-мерное пространство над полем F пространством строк F^n.

2. Сумма и пересечение подпространств.

Пусть V – векторное пространство над полем F, U и W – его подпространства. Легко проверить, что теоретико-множественное пересечение $U \cap W$, т.е. совокупность векторов, принадлежащих обоим этом подпространствам, также есть подпространство в V. Это наибольшее подпространство, содержащееся как в U, так и в W.

Теоретико-множественное объединение $U \cup W$, т.е. совокупность векторов, принадлежащих либо подпространству U, либо подпространству W,напротив, не обязано быть подпространством, как показывает следующий пример.

ПРИМЕР 1. Рассмотрим плоскость V^2 и стандартный прямоугольный базис $\{i, j\}$ плоскости. Пусть $U = \langle i \rangle$, т.е. это ось OX, $W = \langle j \rangle$, т.е. это ось OY. Тогда $U \cup W = OX \cup OY = \{\lambda i \mid \lambda \in \mathbb{R}\} \cup \{\mu j \mid \mu \in \mathbb{R}\}$ содержит векторы i и j, но не содержит их суммы $i+j$, что противоречит определению подпространства.

Если какое-то подпространство содержит U и W, то оно обязано содержать и всевозможные суммы $u + w$, где $u \in U, w \in W$.

Определение 2.1. **Суммой** $U + W$ *подпространств* U *и* W *называется совокупность векторов вида* $u + w$, *где* $u \in U, w \in W$.

Легко проверить, что сумма $U+W$ подпространств действительно является подпространством в V. Это наименьшее подпространство, содержащее как U, так и W.

Отметим, что в общем случае запись вектора $x \in U+W$ в виде $u+w$ может быть неоднозначной.

ПРИМЕР 2. Рассмотрим пространство V^3 и стандартный прямоугольный базис $\{i, j, k\}$ пространства. Пусть $U = \langle i, j \rangle$, т.е. это плоскость OXY, $W = \langle j, k \rangle$, т.е. это плоскость OYZ. Сумма $U + W = \langle i, j, k \rangle$ совпадет со всем пространством V^3. При этом вектор $i + j + k \in U + W$ можно представить двумя способами в виде $u+w$, где $u \in U, w \in W$: как сумму вектора $i+j \in U$ и вектора $k \in W$ и как сумму вектора $i \in U$ и вектора $j + k \in W$. Отметим, что в данном примере пересечением $U \cap W = \langle j \rangle$ является ось OY.

Определение 2.2. *Сумма подпространств* U *и* W *называется* **прямой** *и обозначается* $U \oplus W$, *если каждый вектор* $v \in U \oplus W$ *единственным образом представим в виде* $u+w$, *где* $u \in U, w \in W$. *При этом* u *называется проекцией вектора* v *на* U, w *– проекцией вектора* v *на* W.

ПРИМЕР 3. Плоскость V^2 является прямой суммой осей OX и OY.

Утверждение 2.1. *Сумма подпространств* U *и* W *является прямой тогда и только тогда, когда* $U \cap W = \{0\}$.

Доказательство. Пусть сумма подпространств U и W является прямой. Допустим, что в пересечении $U \cap W$ содержится ненулевой вектор v. Тогда $v \in U + W$ представим двумя способами в виде $u + w$, где $u \in U, w \in W$:

$$v = 0 + v, 0 \in U, v \in W,$$

$$v = v + 0, v \in U, 0 \in W,$$

что противоречит определению прямой суммы подпространств U и W. Это противоречие доказывает, что $U \cap W = \{0\}$.

Обратно, пусть $U \cap W = \{0\}$. Допустим, что сумма подпространств U и W не является прямой. Значит, найдется вектор $v \in U + W$, который имеет два различных представления: $v = u_1 + w_1$ и $v = u_2 + w_2$, где $u_1, u_2 \in U, w_1, w_2 \in W$. Отсюда следует, что $u_1 + w_1 = u_2 + w_2$. Перенося u_2 влево, а w_1 вправо, получаем вектор $u_1 - u_2 = w_2 - w_1$, который принадлежит $U \cap W$, так как $u_1 - u_2 \in U, w_2 - w_1 \in W$. В силу того, что в пересечении подпространств U и W содержится только нулевой вектор, получаем $u_1 - u_2 = 0, w_2 - w_2 = 0$. Откуда $u_1 = u_2, w_1 = w_2$, что противоречит тому, что $v = u_1 + w_1$ и $v = u_2 + w_2$ были двумя различными представлениями вектора v. ■

ПРИМЕР 4. Рассмотрим пространство $V = M_n(F)$ квадратных матриц порядка n. Множество симметрических (соотв, кососимметрических) матриц образует подпространство U (соотв., W) в пространстве $M_n(F)$. При условии, что характеристика поля $\operatorname{char} F$ отлична от 2, всякая матрица A может быть представлена в виде суммы симметрической и кососимметрической матриц:

$$A = \frac{A + A^T}{2} + \frac{A - A^T}{2}.$$

С другой стороны, при том же условии очевидно, что матрица, являющаяся одновременно симметрической и кососимметрической, равна нулю. Это означает, что пространство матриц раскладывается в прямую сумму подпространства симметрических и подпространства кососимметрических матриц, т.е. $M_n(F) = U \oplus W$.

ПРИМЕР 5. Рассмотрим пространство V всех функций на вещественной прямой. Множество четных (соотв., нечетных) функций образует подпространство U (соотв., W) в пространстве V всех функций. Всякая функция f может быть представлена в виде суммы четной и нечетной функций:

$$f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2}.$$

С другой стороны, функция, являющаяся одновременно четной и нечетной, тождественно равна нулю. Это означает, что пространство всех функций на вещественной прямой является прямой суммой подпространства четных и нечетных функций. (В этом примере векторное пространство и оба подпространства бесконечномерны.)

Определение 2.3. **Суммой** $V_1 + \cdots + V_m$ *подпространств* $V_1, \ldots, V_m \subset V$ *называется совокупность векторов вида* $v_1 + \cdots + v_m$, *где* $v_i \in V_i$.

Это наименьшее подпространство, содержащее все подпространства $V_1, \ldots, V_m$.

Определение 2.4. *Сумма* $V_1 + \cdots + V_m$ *подпространств* $V_1, \ldots, V_m \subset V$ *называется* **прямой** *и обозначается* $V_1 \oplus \cdots \oplus V_m$, *если каждый вектор* $v \in V_1 \oplus \cdots \oplus V_m$ *однозначно представляется в виде* $v_1 + \cdots + v_m$, *где* $v_i \in V_i$; *вектор* v_i *называется* **проекцией** *вектора* v *на подпространство* V_1.

Напрашивающееся обобщение Утверждения 2.1 для любого числа подпространств неверно. Как показывает следующий пример, из того, что все попарные пересечения $V_i \cap V_j$ подпространств равны нулевому подпространству $\{0\}$, не следует, что сумма подпространств $V_1 + V_2 + \ldots + V_m$ является прямой.

ПРИМЕР 6. Рассмотрим пространство V^3 и три его подпространства $V_1 = \langle i \rangle$, $V_2 = \langle i - j \rangle$, $V_3 = \langle i + j \rangle$. Очевидно, $V_1 \cap V_2 = V_1 \cap V_3 = V_2 \cap V_3 = \{0\}$. Однако сумма $V_1 + V_2 + V_3$ не является прямой, так как вектор $i \in V_1 + V_2 + V_3$ имеет два различных представления: $i = i + 0 + 0$ и $i = 0 + \frac{1}{2}(i - j) + \frac{1}{2}(i + j)$.

Обобщение Утверждения 2.1 на случай m слагаемых дает следующий результат.

Утверждение 2.2. *Сумма подпространств* $V_1 + V_2 + \ldots + V_m$ *векторного пространства* V *является прямой тогда и только тогда, когда*

$$V_i \cap (V_{i+1} + \ldots + V_m) = \{0\}$$

для любого $i = 1, 2, \ldots, m - 1$.

Доказательство. Пусть сумма $V_1 + V_2 + \ldots + V_m$ – прямая. Предположим, что для некоторого i существует ненулевой вектор

$$v \in V_i \cap (V_{i+1} + \ldots + V_m) \subset V_1 + V_2 + \ldots + V_m.$$

Тогда имеются два различных представления этого вектора:

$$v = v + 0, \quad v \in V_i, \quad 0 \in V_{i+1} + \ldots + V_m$$

и

$$v = 0 + v, \quad 0 \in V_i, \quad v \in V_{i+1} + \ldots + V_m,$$

что противоречит определению прямой суммы.

Обратно, пусть $V_i \cap (V_{i+1} + \ldots + V_m) = \{0\}$ для любого $i = 1, 2, \ldots, m - 1$, но сумма $V_1 + V_2 + \ldots + V_m$ не является прямой. Тогда для некоторого вектора v из суммы $V_1 + V_2 + \ldots + V_m$ существуют два различных разложения

$$v = u_1 + \ldots + u_m, \quad u_i \in V_i$$

и

$$v = w_1 + \ldots + w_m, \quad w_i \in V_i.$$

Откуда $u_1 + \ldots + u_m = w_1 + \ldots + w_m$. Перенося члены в левую часть и группируя, получим $v_1 + \ldots + v_m = 0$, где $v_i = u_i - w_i \in V_i$, причем не все слагаемые равны нулю. Пусть v_i – первое ненулевое слагаемое. Тогда

$$V_i \ni v_i = -v_{i+1} - \ldots - v_m \in V_{i+1} + \ldots + V_m,$$

что противоречит тому, что $V_i \cap (V_{i+1} + \ldots + V_m) = \{0\}$. ■

Замечание. *Часто используется другое, формально более сильное, однако эквивалентное, условие: пересечение каждого подпространства V_i с суммой всех остальных равно нулю.*

ПРИМЕР 7. Пусть $\{e_1, \ldots, e_n\}$ – базис векторного пространства V. Тогда $V = \langle e_1 \rangle \oplus \ldots \oplus \langle e_n \rangle$.

Подчеркнем, что проекция вектора на подпространство V_i зависит не только от этого подпространства, но и от остальных слагаемых разложения $V_1 + V_2 + \ldots + V_m$.

ЗАДАЧА. Приведите пример, подтверждающий это высказывание.

Нашей дальнейшей целью будет доказательство следующей важной теоремы.

Теорема 2.1. *Для любых двух подпространств U и W конечномерного векторного пространства V над полем F справедлива формула*

$$\dim\,(U + W) = \dim U + \dim W - \dim\,(U \cap W). \tag{2}$$

Формула (2) называется **формулой Грассмана.**

Доказательство. Мы воспользуемся Теоремой 1.3 о том, что всякую линейно независимую систему векторов конечномерного векторного пространства можно дополнить до базиса этого пространства. Выберем базис $\mathcal{E} = \{e_1, \ldots, e_k\}$ в подпространстве $U \cap W$ и дополним его векторами $\mathcal{F} = \{f_1, \ldots, f_l\}$ до базиса подпространства U и векторами $\mathcal{G} = \{g_1, \ldots, g_m\}$ до базиса подпространства W. Для доказательства формулы (2) достаточно показать, что $\mathcal{E} \cup \mathcal{F} \cup \mathcal{G}$ – базис суммы $U + W$ подпространств U и V.

Любой вектор v из суммы $U + W$ есть $u + w$, где $u \in U, w \in W$. Вектор u выражается через базис $\mathcal{E} \cup \mathcal{F}$ подпространства U, вектор w – через базис $\mathcal{E} \cup \mathcal{G}$ подпространства W. Поэтому $u + w$ выражается через систему векторов $\mathcal{E} \cup \mathcal{F} \cup \mathcal{G}$, которая в свою очередь принадлежит сумме $U + W$. Значит, $U + W$ равно линейной оболочке $\langle \mathcal{E}, \mathcal{F}, \mathcal{G} \rangle$ системы векторов $\mathcal{E} \cup \mathcal{F} \cup \mathcal{G}$.

Остается показать, что система $\mathcal{E} \cup \mathcal{F} \cup \mathcal{G}$ линейно независима. Предположим противное. Пусть имеется нетривиальная линейная комбинация векторов из системы $\mathcal{E} \cup \mathcal{F} \cup \mathcal{G}$, равная нулю. Группируя в ней слагаемые, получим

$$(\alpha_1 e_1 + \ldots + \alpha_k e_k) + (\beta_1 f_1 + \ldots + \beta_l f_l) + (\gamma_1 g_1 + \ldots + \gamma_m g_m) = 0,$$

т.е.

$$a + b + c = 0,$$

где

$$a = \alpha_1 e_1 + \ldots + \alpha_k e_k \in \langle \mathcal{E} \rangle, \;\; \alpha_1, \ldots, \alpha_k \in F,$$

$$b = \beta_1 f_1 + \ldots + \beta_l f_l \in \langle \mathcal{F} \rangle, \;\; \beta_1, \ldots, \beta_l \in F,$$

$$c = \gamma_1 g_1 + \ldots + \gamma_m g_m \in \langle \mathcal{G} \rangle, \;\; \gamma_1, \ldots, \gamma_m \in F.$$

С одной стороны, имеем

$$b = -a - c \in \langle \mathcal{E}, \mathcal{G} \rangle = W,$$

с другой стороны,

$$b \in \langle \mathcal{F} \rangle \subset U,$$

таким образом, $b \in U \cap W = \langle \mathcal{E} \rangle$. Однако система $\mathcal{E} \cup \mathcal{F}$ линейно независима, так что $\langle \mathcal{E} \rangle \cap \langle \mathcal{F} \rangle = \{0\}$. Отсюда

$$b = 0,$$

а значит, $\beta_1 = 0, \ldots, \beta_l = 0$. Остается равенство

$$a + c = 0.$$

Отсюда в силу линейной независимости множества $\mathcal{E} \cup \mathcal{G}$ вытекает, что

$$a = 0 \;\text{ и }\; c = 0,$$

а значит, $\alpha_1 = 0, \ldots, \alpha_m = 0, \gamma_1 = 0, \ldots, \gamma_m = 0$. Линейная независимость системы $\mathcal{E} \cup \mathcal{F} \cup \mathcal{G}$ доказана. Тем самым доказана и формула Грассмана. ∎

ЗАДАЧА. Приведите пример векторного пространства V и его трех подпространств V_1, V_2, V_3, для которых

$$\dim\,(V_1 + V_2 + V_3) \neq \dim\,V_1 + \dim\,V_2 + \dim\,V_3 - \dim\,(V_1 \cap V_2) -$$

$$-\dim\,(V_1 \cap V_3) - \dim\,(V_2 \cap V_3) + \dim\,(V_1 \cap V_2 \cap V_3).$$

Следствие 2.1. *Сумма двух подпространств U и W конечномерного векторного пространства V является прямой тогда и только тогда, когда*

$$\dim\,(U + W) = \dim\,U + \dim\,W.$$

Из следствия 2.1 индукцией по количеству слагаемых в прямой сумме $V_1 \oplus \ldots \oplus V_m$ легко вытекает также следующая формула:

$$\dim\,(V_1 \oplus \ldots \oplus V_m) = \dim\,V_1 + \ldots + \dim\,V_m.$$

3. Сопряженное пространство.

3.1. Линейные функции.

Пусть V – векторное пространство над полем F. Одновременно с V часто рассматривают другое, тесно связанное с ним пространство, так называемое сопряженное пространство. Для того чтобы сформулировать определение сопряженного пространства, нам понадобится определение линейной функции.

Определение 3.1. Линейной функцией *(или* **линейной формой***) называется отображение $\varphi : V \to F$, обладающее свойствами:*
1) $\varphi(x + y) = \varphi(x) + \varphi(y)$ для любых $x, y \in V$;
2) $\varphi(\lambda x) = \lambda\varphi(x)$ для любых $\lambda \in F$, $x \in V$.

ПРИМЕР 1. В пространстве геометрических векторов V^3 отображение

$$\varphi(x) = (x, a),$$

где $a \in V^3$ – некоторый фиксированный вектор, является линейной функцией, так как скалярное произведение линейно по первому аргументу.

ПРИМЕР 2. В пространстве $C[a, b]$, векторами которого являются непрерывные функции $f(x)$, заданные на отрезке $[a, b]$, отображение

$$\varphi(f) = \int_a^b f(t)dt$$

является линейной функцией. Действительно, условие 1) означает, что интеграл суммы равен сумме интегралов, а условие 2) означает, что постоянный множитель можно выносить за знак интеграла.

ПРИМЕР 3. В пространстве $C[a, b]$ отображение

$$\varphi(f) = f(x_0),$$

где x_0 – произвольное фиксированное число из отрезка $[a, b]$, является линейной функцией. Проверьте это!

ПРИМЕР 4. Из свойств производных следует, что отображение

$$\varphi(f) = f'(x_0) \ \ (x_0 \in \mathbb{R})$$

является линейной функцией на векторном пространстве $C'(\mathbb{R})$ дифференцируемых функций на вещественной прямой.

ПРИМЕР 5. **Следом** квадратной матрицы называется сумма ее диагональных элементов. След матрицы $A \in M_n(F)$ обозначается через $tr\ A$. Отображение

$$\varphi(A) = \mathrm{tr}\ A$$

является линейной функцией на пространстве $M_n(F)$ квадратных матриц. Проверьте это!

Отметим, что определитель также является функцией на $M_n(F)$, но эта функция не является линейной, так как $\det\,(A+B) \neq \det\,A + \det\,B$.

ПРИМЕР 6. В пространстве F^n строк длины n отображение (i-я проекция)

$$\varphi_i(a_1, a_2, \ldots, a_n) = a_i$$

является линейной функцией.

Множество всех функций на пространстве V со значениями в поле F образует векторное пространство над F относительно операций обычного поточечного сложения функций и умножения функции на число:

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

$$(\lambda\varphi)(v) = \lambda\varphi(v),$$

где φ, ψ – произвольные функции на V, v – произвольный вектор из V, а λ – произвольное число из F. Множество всех линейных функций образует подпространство в этом векторном пространстве, так как сумма линейных функций и произведение линейной функции на число есть снова линейная функция.

Определение 3.2. *Пространство всех линейных функций на V называется* **сопряженным пространством** *по отношению к V и обозначается через V^*.*

Рассмотрим n-мерное пространство V и выберем в нем произвольный базис $\mathcal{E} = \{e_1, \ldots, e_n\}$. Значение линейной функции $\varphi \in V^*$ на векторе $x \in V$ может быть выражено через координаты этого вектора $x_1, \ldots, x_n$:

$$\varphi(x) = \varphi(x_1e_1 + \ldots + x_ne_n) = x_1\varphi(e_1) + \ldots + x_n\varphi(e_n) = a_1x_1 + \ldots + a_nx_n, \quad (3)$$

где $a_i = \varphi(e_i)$. Числа $a_1 = \varphi(e_1), \ldots, a_n = \varphi(e_n)$ не зависят от вектора x, а определяются только функцией φ и базисом $\mathcal{E}$. Линейная функция φ однозначно определяется своими значениями $a_1, \ldots, a_n$ на базисных векторах. Эти значения называются **коэффициентами** функции φ в базисе $\mathcal{E}$. Коэффициенты могут быть произвольными: для любых n чисел $a_1, \ldots, a_n \in F$ функция φ, определяемая формулой $\varphi(x) = a_1x_1 + \ldots + a_nx_n$, является линейной. Нетрудно проверить также, что если $a_1, \ldots, a_n$ – коэффициенты линейной функции φ в базисе $\mathcal{E}$, а $b_1, \ldots, b_n$ – коэффициенты линейной функции ψ в этом же базисе, то $a_1 + b_1, \ldots, a_n + b_n$ – коэффициенты суммы $\varphi + \psi$, а $\lambda a_1, \ldots, \lambda a_n$ – коэффициенты $\lambda\varphi$, где $\lambda \in F$. Отсюда следует, что V^* изоморфно пространству F^n строк длины n. Из этого получаем два утверждения:

Утверждение 3.1. *Для любого конечномерного векторного пространства V имеем* $\dim V = \dim V^*$.

Утверждение 3.2. *Для любого конечномерного векторного пространства V имеем $V \cong V^*$.*

Отметим, что если V бесконечномерно, то V и V^* не являются изоморфными.

Пусть $\mathcal{E} = \{e_1, \dots, e_n\}$ – произвольный базис пространства V. Определим систему функций $\varepsilon^1, \dots, \varepsilon^n \in V^*$, полагая

$$\varepsilon^i(e_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \ (i, j = 1, 2, \dots, n),$$

или, используя символ Кронекера,

$$\varepsilon^i(e_j) = \delta^i_j,$$

т.е. строка коэффициентов функции ε^i есть i-я строка единичной матрицы. Отсюда легко следует, что функции $\varepsilon^1, \dots, \varepsilon^n$ линейно независимы. Так как пространство V^* n-мерно, эти функции составляют в нем базис.

Определение 3.3. *Базис $\{\varepsilon^1, \dots, \varepsilon^n\}$ пространства V^*, определяемый формулой*

$$\varepsilon^i(e_j) = \delta^i_j,$$

называется **сопряженным** *по отношению к базису $\{e_1, \dots, e_n\}$ пространства V.*

Имея в V и V^* сопряженные базисы, легко вычислить координаты любого вектора. Пусть $\{e_1, \dots, e_n\}$ – базис в V, $\{\varepsilon^1, \dots, \varepsilon^n\}$ – сопряженный базис относительно $\{e_1, \dots, e_n\}$. Для произвольного вектора $x = x_1e_1 + \ldots + x_ne_n \in V$ верно, что

$$\varepsilon^i(x) = \varepsilon^i(x_1e_1 + \ldots + x_ne_n) = x_1\varepsilon^i(e_1) + \ldots + x_n\varepsilon^i(e_n) = x_i.$$

Следовательно, координаты x_i вектора x в базисе $\{e_1, \dots, e_n\}$ вычисляется по формуле $x_i = \varepsilon^i(x)$ и

$$x = \varepsilon^1(x)e_1 + \ldots + \varepsilon^n(x)e_n.$$

Аналогично для произвольной линейной функции $\varphi = a_1\varepsilon^1 + \ldots + a_n\varepsilon^n \in V^*$ верно, что

$$\varphi(e_i) = a_1\varepsilon^1(e_i) + \ldots + a_n\varepsilon^n(e_i) = a_i.$$

Следовательно координаты a_i вектора $\varphi \in V^*$ в базисе $\{\varepsilon^1, \dots, \varepsilon^n\}$ вычисляются по формуле $a_i = \varphi(e_i)$ и

$$\varphi = \varphi(e_1)\varepsilon^1 + \ldots + \varphi(e_n)\varepsilon^n.$$

Отметим, что коэффициенты линейной функции φ в базисе $\mathcal{E}$ совпадают с ее координатами в базисе, сопряженном к $\mathcal{E}$.

3.2. Преобразование координат в V^*.

Утверждение 3.3. *Пусть даны два базиса* $\mathcal{E} = \{e_1, \ldots, e_n\}$ *и* $\widetilde{\mathcal{E}} = \{\tilde{e_1}, \ldots, \tilde{e_n}\}$ *векторного пространства* V. *Рассмотрим базисы* $\mathcal{F} = \{\varepsilon^1, \ldots, \varepsilon^n\}$ *и* $\widetilde{\mathcal{F}} = \{\tilde{\varepsilon}^1, \ldots, \tilde{\varepsilon}^n\}$ *пространства* V^*, *сопряженные к* $\mathcal{E}$ *и* $\widetilde{\mathcal{E}}$ *соответственно. Если* $T_{\mathcal{E}\to\widetilde{\mathcal{E}}} = T = (t_{i,j})$ *– матрица перехода от базиса* $\mathcal{E}$ *к базису* $\widetilde{\mathcal{E}}$, *то матрица перехода* $T_{\widetilde{\mathcal{F}}\to\mathcal{F}}$ *от базиса* $\widetilde{\mathcal{F}}$ *к базису* $\mathcal{F}$ *равняется матрице* T^t.

Доказательство. Обозначим через $S = (s_{i,j})$ – матрицу перехода $T_{\mathcal{F}\to\widetilde{\mathcal{F}}}$ от базиса $\mathcal{F}$ к базису $\widetilde{\mathcal{F}}$, а через $s^t_{i,j}$ – элементы матрицы S^t, т.е. $s^t_{i,j} = s_{j,i}$. Тогда для каждых $k, l = \overline{1,n}$

$$\tilde{\varepsilon}^k = \sum_{i=1}^{n} s_{i,k}\varepsilon^i,$$

$$\tilde{e_l} = \sum_{j=1}^{n} t_{j,l}e_j.$$

Использую равенства $\tilde{\varepsilon}^k(\tilde{e_l}) = \delta^k_l$ и $\varepsilon^i(e_j) = \delta^i_j$, получаем

$$\delta^k_l = \sum_{i=1}^{n} s_{i,k}\varepsilon^i(\sum_{j=1}^{n} t_{j,l}e_j) = \sum_{i,j=1}^{n} s_{i,k}t_{j,l}\varepsilon^i(e_j) = \sum_{i=1}^{n} s_{i,k}t_{i,l} = \sum_{i=1}^{n} s^t_{k,i}t_{i,l}.$$

Следовательно, $S^tT = E$. А значит, $S^{-1} = T^t$. Откуда получаем требуемое, так как $T_{\widetilde{\mathcal{F}}\to\mathcal{F}} = S^{-1}$. ■

3.3. Взаимозаменяемость V и V^*.

В предыдущем изложении V и V^* играли различную роль. Мы покажем, что они совершенно равноправны, т.е. что все теоремы останутся справедливыми, если мы поменяем V и V^* ролями.

Пространство V^* – такое же векторное пространство, как и любое другое, и, следовательно, имеет сопряженное пространство $V^{**} = (V^*)^*$, элементы которого – линейные функции на V^*. Покажем, что второе сопряженное пространство $V^{**} = (V^*)^*$ оказывается естественно изоморфным пространству V.

Для любого вектора $x \in V$ рассмотрим функцию f_x на V^*, определенную по формуле

$$f_x(\varphi) = \varphi(x), \varphi \in V^*.$$

Из определения операций в V^* следует, что f_x – линейная функция.

Утверждение 3.4. *Отображение* $\Phi : V \to V^{**}$, *переводящее каждый вектор* $x \in V$ *в линейную функцию* $f_x \in V^{**}$, *является изоморфизмом конечномерного пространства* V *и пространства* V^{**}.

Доказательство. 1) Докажем, что $\Phi(x+y) = \Phi(x)+\Phi(y)$ для любых векторов $x, y \in V$. Так как $\Phi(x+y) = f_{x+y}$, $\Phi(x)+\Phi(y) = f_x + f_y$, требуется доказать, что линейная функция f_{x+y} равна линейной функции $f_x + f_y$. Для любого вектора $\varphi \in V^*$ имеем

$$f_{x+y}(\varphi) = \varphi(x+y) = \varphi(x)+\varphi(y) = f_x(\varphi)+f_y(\varphi) = (f_x+f_y)(\varphi).$$

Откуда получаем равенство $f_{x+y} = f_x + f_y$, так как две функции равны тогда и только тогда, когда равны значения этих функций для каждого аргумента.

Равенство $\Phi(\lambda x) = \lambda\Phi(x)$ для любых $x \in V, \lambda \in F$ доказывается аналогично.

2) Остается проверить, что отображение Φ биективно. Пусть $\mathcal{E} = \{e_1, \ldots, e_n\}$ – некоторый базис пространства V и $\{\varepsilon^1, \ldots, \varepsilon^n\}$ – сопряженный базис пространства V^*. Тогда

$$f_{e_i}(\varepsilon^j) = \varepsilon^j(e_i) = \delta_i^j,$$

так что $\{f_{e_1}, \ldots, f_{e_n}\}$ – базис пространства V^{**}, сопряженный базису $\{\varepsilon^1, \ldots, \varepsilon^n\}$.

Так как

$$\Phi(x) = \Phi(x_1e_1+\ldots+x_ne_n) = x_1\Phi(e_1)+\ldots+x_n\Phi(e_n) = x_1f_{e_1}+\ldots+x_nf_{e_n},$$

то отображение Φ переводит вектор x с координатами $x_1, \ldots, x_n$ в базисе $\{e_1, \ldots, e_n\}$ пространства V в вектор $f_x = \Phi(x)$ с такими же координатами в базисе $\{f_{e_1}, \ldots, f_{e_n}\}$ пространства V^{**}. Следовательно, Φ биективно. ■

Таким образом, можно отождествить пространства V и V^{**} посредством указанного изоморфизма, т.е. рассматривать каждый вектор $x \in V$ одновременно и как линейную функцию на V^* и писать $x(\varphi)$ вместо $f_x(\varphi)$. При таком соглашении пространства V и V^* будут играть совершенно симметричную роль.

Следствие 3.1. *Всякий базис пространства V^* сопряжен некоторому базису пространства V.*

Часто вместо $\varphi(x)$ и $x(\varphi)$ пишут $<\varphi|x>$, желая подчеркнуть равноправие векторов и функций. Фиксация φ приводит к линейной функции на векторах, фиксация x приводит к линейной функции на линейных функциях.

Замечание 3.1. *Если пространство V бесконечномерно, то пространство V^* и, тем более, V^{**} имеет бóльшую размерность.*

3.4. Подпространства и линейные функции.

Из 1-го семестра знаем, что множество решений произвольной однородной системы линейных уравнений с n неизвестными образует подпространство в пространстве строк длины n. Оказывается, всякое подпространство является множеством решений некоторой однородной системы линейных уравнений.

Определение 3.4. *Пусть на произвольном векторном пространстве V над полем F определена линейная функция φ. Назовем* **ядром** *линейной функции φ множество*

$$\mathrm{Ker}\ \varphi = \{v \in V | \varphi(v) = 0\}.$$

Утверждение 3.5. *Ядро любой линейной функции φ является подпространством в векторном пространстве V над полем F.*

Доказательство. Для начала, отметим, что $\mathrm{Ker}\ \varphi$ не пусто, так как

$$0 \in \mathrm{Ker}\ \varphi.$$

Далее, если $u, v \in \mathrm{Ker}\ \varphi$, то $\varphi(u) = 0$ и $\varphi(v) = 0$. Поэтому

$$\varphi(u + v) = \varphi(u) + \varphi(v) = 0 + 0 = 0,$$

что означает $u + v \in \mathrm{Ker}\ \varphi$.

Если $\lambda \in F$, $u \in \mathrm{Ker}\ \varphi$, то $\varphi(\lambda u) = \lambda\varphi(u) = \lambda 0 = 0$. Значит, $\lambda u \in \mathrm{Ker}\ \varphi$.

Это доказывает, что $\mathrm{Ker}\ \varphi$ является подпространством в V. ■

Теорема 3.1. *Пусть V – конечномерное векторное пространство над полем F. Тогда*

1) всякое подпространство U пространства V есть пересечение ядер некоторого конечного множества линейных функций;

2) при фиксированном базисе каждое подпространство U пространства V есть множество векторов, координаты которых являются решениями подходящей однородной системы линейных уравнений.

Доказательство.

1) Выберем в подпространстве U базис $\{e_1, \ldots, e_k\}$. Дополним этот базис векторами $\{e_{k+1}, \ldots, e_n\}$ до базиса всего пространства V (Теорема 1.3). Пусть $\{\varepsilon^1, \ldots, \varepsilon^n\}$ – базис, сопряженный к базису $\mathcal{E} = \{e_1, \ldots, e_k, e_{k+1}, \ldots, e_n\}$. Докажем, что $U = \mathrm{Ker}\ \varepsilon^{k+1} \cap \ldots \cap \mathrm{Ker}\ \varepsilon^n$.

Действительно, возьмем произвольный вектор $x \in V$ и выразим его через базис $\mathcal{E}$ в виде:

$$x = x_1 e_1 + \ldots + x_k e_k + x_{k+1} e_{k+1} + \ldots + x_n e_n.$$

Учитывая, что $x_i = \varepsilon^i(x)$, получаем

$$x \in U \Leftrightarrow \left\{ \begin{array}{rcl} x_{k+1} & = & 0; \\ \cdots & & \\ x_n & = & 0; \end{array} \right. \Leftrightarrow \left\{ \begin{array}{rcl} \varepsilon^{k+1}(x) & = & 0; \\ \cdots & & \\ \varepsilon^n(x) & = & 0; \end{array} \right. \Leftrightarrow x \in \mathrm{Ker}\ \varepsilon^{k+1} \cap \ldots \cap \mathrm{Ker}\ \varepsilon^n,$$

что и требовалось.

2) Пусть $\widetilde{\mathcal{E}} = \{\widetilde{e}_1, \ldots, \widetilde{e}_n\}$ – некоторый фиксированный базис, T – матрица перехода от базиса $\mathcal{E} = \{e_1, \ldots, e_k, e_{k+1}, \ldots, e_n\}$ к базису $\widetilde{\mathcal{E}}$. Если $x = \tilde{x}_1\tilde{e}_1 + \ldots + \tilde{x}_n\tilde{e}_n$, то по формуле (1) преобразования координат вектора при переходе к другому базису имеем

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = T \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix}.$$

Обозначим через $\breve{T}$ матрицу, состоящую из последних $n-k$ строк матрицы T. Так как

$$\begin{pmatrix} x_{k+1} \\ \vdots \\ x_n \end{pmatrix} = \breve{T} \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix}$$

и

$$x \in U \Leftrightarrow \left\{ \begin{array}{rcl} x_{k+1} & = & 0; \\ \cdots & & \\ x_n & = & 0; \end{array} \right.$$

получаем, что

$$x \in U \Leftrightarrow \breve{T} \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix} = 0,$$

т.е. координаты $\begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix}$ векторов $x \in U$ в базисе $\widetilde{\mathcal{E}}$ являются множеством решений однородной системы линейных уравнений с матрицей $\breve{T}$. ■

4. Линейные отображения и линейные операторы.

Мы переходим к рассмотрению основного предмета нынешнего семестра – к рассмотрению линейных операторов. Теория линейных операторов – это ядро линейной алгебры и главный источник ее многочисленных приложений.

4.1. Линейные отображения.

Пусть V и W – векторные пространства над полем F.

Определение 4.1. *Отображение* $\varphi : V \to W$ *называется* **линейным**, *если* φ *сохраняет операции сложения и умножения на число:*
1) $\varphi(x+y) = \varphi(x) + \varphi(y)\ \ \forall x, y \in V$;
2) $\varphi(\lambda x) = \lambda\varphi(x)\ \ \forall \lambda \in F, \forall x \in V$.

Это определение отличается от определения изоморфизма векторных пространств тем, что в нем не требуется биективности.

Простейшие свойства линейных отображений:
1) $\varphi(0) = 0$;
2) $\varphi(-x) = -\varphi(x)\ \ \forall x \in V$;
3) $\varphi(x-y) = \varphi(x) - \varphi(y)\ \ x, y \in V$.
Докажите эти свойства!

ПРИМЕР 1. Линейные функции являются линейными отображениями из векторного пространства V в поле F, рассматриваемое как векторное пространство над самим собой.

ПРИМЕР 2. Поворот есть линейное отображение (и даже изоморфизм) пространства V^2 геометрических векторов в себя.

ПРИМЕР 3. Ортогональное проектирование на плоскость определяет линейное отображение (но не изоморфизм) пространства V^3 в пространство геометрических векторов этой плоскости.

ПРИМЕР 4. Дифференцирование является линейным отображением пространства непрерывно дифференцируемых функций на заданном промежутке числовой прямой в пространство непрерывных функций на этом промежутке.

Линейное отображение $\varphi : V \to W$ однозначно определяется образами базисных векторов пространства V. В самом деле, пусть $\mathcal{E} = \{e_1, \dots, e_n\}$ – базис пространства V. Тогда для любого вектора $x = \sum_{i=1}^n x_i e_i$ пространства V имеем

$$\varphi(x) = \sum_{i=1}^n x_i \varphi(e_i).$$

С другой стороны, если $w_1, \ldots, w_n \in W$ – произвольные векторы, то отображение, определяемое по формуле

$$\varphi(x) = \sum_{i=1}^{n} x_i w_i,$$

как легко видеть, является линейным и $\varphi(e_i) = w_i$.

Пусть $\mathcal{F} = \{f_1, \ldots, f_m\}$ – базис пространство W. Для каждого $j = \overline{1, n}$ образ $\varphi(e_j)$ вектора $e_j \in \mathcal{E}$ принадлежит пространству W, а значит, его можно разложить по базису $\mathcal{F}$ этого пространства:

$$\varphi(e_j) = \sum_{i=1}^{m} a_{i,j} f_i.$$

Матрица A порядка $m \times n$ с элементами $a_{i,j}$ называется **матрицей линейного отображения относительно базисов $\mathcal{E}$ и $\mathcal{F}$**. Другими словами, столбцы матрицы линейного отображения (в их естественном порядке) – это столбцы координат векторов $\varphi(e_1), \ldots, \varphi(e_n)$ в базисе $\mathcal{F}$.

Для произвольного вектора $x = \sum_{j=1}^{n} x_j e_j$ из пространства V разложим его образ $y = \varphi(x)$ по базису $\mathcal{F}$ пространства W:

$$y = \sum_{i=1}^{m} y_i f_i.$$

С другой стороны, имеем

$$y = \varphi(x) = \varphi(\sum_{j=1}^{n} x_j e_j) = \sum_{j=1}^{n} x_j \varphi(e_j) = \sum_{j=1}^{n} x_j (\sum_{i=1}^{m} a_{i,j} f_i) =$$

$$= \sum_{i=1}^{m} (\sum_{j=1}^{n} a_{i,j} x_j) f_i.$$

В силу единственности разложения вектора по базису отсюда получаем, что $y_i = \sum_{j=1}^{n} a_{i,j} x_j$ для каждого $i = \overline{1, m}$ или в матричной форме

$$Y = AX, \tag{4}$$

где

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ – столбец координат вектора } x \text{ в базисе } \mathcal{E},$$

$$\text{а } Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \text{ – столбец координат вектора } y = \varphi(x) \text{ в базисе } \mathcal{F}.$$

ПРИМЕР 5. Найдем матрицу ортогонального проектирования из примера 3. В плоскости проектирования выберем любой базис $\{e_1, e_2\}$ и дополним его ортогональным вектором e_3 до базиса пространства. Так как при проектировании векторы e_1 и e_2 переходят сами в себя, а вектор e_1 – в нуль, то искомая матрица (относительно выбранных базисов) имеет вид

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Наконец, обсудим вопрос об изменении матрицы линейного отображения $\varphi : V \to W$ при изменении базиса $\mathcal{E} = \{e_1, \ldots, e_n\}$ на базис $\widetilde{\mathcal{E}} = \{\tilde{e}_1, \ldots, \tilde{e}_n\}$ в пространстве V и базиса $\mathcal{F} = \{f_1, \ldots, f_m\}$ на базис $\widetilde{\mathcal{F}} = \{\tilde{f}_1, \ldots, \tilde{f}_m\}$ в пространстве W. Итак, пусть $\widetilde{A}$ – матрица линейного отображения φ относительно базисов $\widetilde{\mathcal{E}}$ и $\widetilde{\mathcal{F}}$. Тогда по формуле (4) линейное отображение φ в базисах $\widetilde{\mathcal{E}}$ и $\widetilde{\mathcal{F}}$ запишется так:

$$\widetilde{Y} = \widetilde{A}\widetilde{X}, \tag{5}$$

где $\widetilde{X}$ – столбец координат вектора x в базисе $\widetilde{\mathcal{E}}$, а $\widetilde{Y}$ – столбец координат вектора $y = \varphi(x)$ в базисе $\widetilde{\mathcal{F}}$.

Если $T = T_{\mathcal{E}\to\widetilde{\mathcal{E}}}$ – матрица перехода от $\mathcal{E}$ к $\widetilde{\mathcal{E}}$, а $S = S_{\mathcal{F}\to\widetilde{\mathcal{F}}}$ – матрица перехода от $\mathcal{F}$ к $\widetilde{\mathcal{F}}$, то по формуле (1)

$$X = T\widetilde{X},$$

$$Y = S\widetilde{Y}.$$

Подставим в формулу (4) вместо X и Y их выражение через $\widetilde{X}$ и $\widetilde{Y}$:

$$S\widetilde{Y} = AT\widetilde{X}.$$

Отсюда получаем

$$\widetilde{Y} = S^{-1}AT\widetilde{X}.$$

Сравнивая полученное выражение с (5), находим, что

$$\widetilde{A}\widetilde{X} = S^{-1}AT\widetilde{X}.$$

Так как это равенство верно для любого $\widetilde{X}$, в частности, для столбца координат каждого базисного вектора $\tilde{e}_i$, то получаем равенство матриц

$$\widetilde{A} = S^{-1}AT. \tag{6}$$

Эта формула называется **формулой преобразования матрицы линейного отображения φ при переходе к другим базисам**.

Лемма 4.1. *При умножении матрицы на невырожденную матрицу ее ранг не изменяется.*

Доказательство леммы. Пусть A – произвольная матрица размера $m \times n$, C – произвольная матрица размера $n \times k$. Тогда $\mathrm{rk}\ AC \leqslant \mathrm{rk}\ A$. Действительно, из 1-го семестра известно, что ранг матрицы равен рангу ее столбцов, т.е. размерности линейной оболочки ее столбцов. Столбцы матрицы AC являются линейными комбинациями столбцов матрицы A, значит, линейная оболочка столбцов матрицы AC является подпространством в линейной оболочке столбцов матрицы A. Поэтому из теоремы 1.4 имеем $\mathrm{rk}\ AC \leqslant \mathrm{rk}\ A$.

Применяя эту формулу к матрицам AC и C^{-1} в случае квадратной невырожденной матрицы C, имеем $\mathrm{rk}\ (AC)C^{-1} \leqslant \mathrm{rk}\ AC$. Следовательно,

$$\mathrm{rk}\ A = \mathrm{rk}\ (AC)C^{-1} \leqslant \mathrm{rk}\ AC \leqslant \mathrm{rk}\ A,$$

что доказывает равенство $\mathrm{rk}\ A = \mathrm{rk}\ AC$. ■

Из леммы 4.1 и формулы 6 получаем, что ранг матрицы линейного отображения не зависит от выбора базисов.

В отличие от изоморфизма линейное отображение не обязано быть ни сюръективным, ни инъективным. Нарушение этих свойств приводит к возможности связать с каждым линейным отображением два подпространства – его образ и ядро.

Определение 4.2. Образом *линейного отображения* $\varphi : V \to W$ *называется подмножество*

$$\mathrm{Im}\ \varphi = \{y \in W \mid \exists x \in V,\ y = \varphi(x)\} \subset W,$$

а **ядром** *– подмножество*

$$\mathrm{Ker}\ \varphi = \{x \in V \mid \varphi(x) = 0\} \subset V.$$

Утверждение 4.1. *Образ* $\mathrm{Im}\ \varphi$ *является подпространством в пространстве* W*, а ядро* $\mathrm{Ker}\ \varphi$ *является подпространством в* V*.*

Доказательство. Так как $\varphi(0) = 0$, то $\mathrm{Im}\ \varphi$ и $\mathrm{Ker}\ \varphi$ – непустые подмножество. Остается только проверить, что эти подмножества замкнуты относительно операций сложения векторов и умножения вектора на число.

Сначала, проверим замкнутость $\mathrm{Im}\ \varphi$ относительно этих операций. Пусть y, w – произвольные векторы из $\mathrm{Im}\ \varphi$, т. е. найдутся такие $x, z \in V$, что $\varphi(x) = y$, $\varphi(z) = w$. Так как $x + z \in V$ и

$$y + w = \varphi(x) + \varphi(z) = \varphi(x + z),$$

то $y + w \in \mathrm{Im}\ \varphi$. Так как для любого $\lambda \in F$ верно, что $\lambda x \in V$ и

$$\lambda y = \lambda\varphi(x) = \varphi(\lambda x),$$

то $\lambda y \in \mathrm{Im}\ \varphi$.

Теперь рассмотрим ядро $\mathrm{Ker}\ \varphi$. Если x, z – произвольные векторы из $\mathrm{Ker}\ \varphi$, то $\varphi(x) = 0, \varphi(z) = 0$. Следовательно,

$$\varphi(x + z) = \varphi(x) + \varphi(z) = 0 + 0 = 0$$

и для любого $\lambda \in F$

$$\varphi(\lambda x) = \lambda\varphi(x) = \lambda 0 = 0,$$

что означает, что $x + z, \lambda x \in \mathrm{Ker}\ \varphi$. ■

ПРИМЕР 6. Ядром отображения проектирования из примера 3 является совокупность векторов, ортогональных плоскости проектирования, а образом – вся плоскость проектирования.

ПРИМЕР 7. Ядром отображения дифференцирования из примера 4 является совокупность постоянных функций, а образом – пространство всех непрерывных функций. Последнее следует из существования первообразной у каждой непрерывной функции, доказанного в курсе "Математический анализ".

Утверждение 4.2. *Линейное отображение $\varphi : V \to W$ инъективно тогда и только тогда, когда* $\mathrm{Ker}\ \varphi = \{0\}$.

Доказательство. Так как $\varphi(0) = 0$, то из инъективности φ следует, что $\varphi(x) \neq 0$ для любого ненулевого вектора x, т.е. $\mathrm{Ker}\ \varphi = \{0\}$.

Обратно, пусть $\mathrm{Ker}\ \varphi = \{0\}$. Допустим, что φ не инъективно. Тогда существуют различные векторы x и y из V такие, что $\varphi(x) = \varphi(y)$. Отсюда в силу линейности отображения φ имеем $\varphi(x - y) = \varphi(x) - \varphi(y) = 0$. Значит, $x - y \in \mathrm{Ker}\ \varphi = \{0\}$, т.е. $x - y = 0$. Это противоречит тому, что x и y – различные векторы. ■

Утверждение 4.3. *Пусть $\varphi : V \to W$ – линейное отображение конечномерного пространства V и $\mathcal{E} = \{e_1, \ldots, e_n\}$ – базис пространства V. Тогда*

$$\mathrm{Im}\ \varphi = \langle \varphi(e_1), \ldots, \varphi(e_n) \rangle. \tag{7}$$

Кроме того, если W – также конечномерное векторное пространство, $\mathcal{F} = \{f_1, \ldots, f_m\}$ – его базис, а A – матрица линейного отображения φ относительно базисов $\mathcal{E}$ и $\mathcal{F}$, то

$$\dim\ \mathrm{Im}\ \varphi = \mathrm{rk}\ A. \tag{8}$$

Доказательство. Для любого

$$x = x_1 e_1 + \ldots + x_n e_n \in V$$

имеем

$$\varphi(x) = x_1\varphi(e_1) + \ldots + x_n\varphi(e_n).$$

Следовательно, $\mathrm{Im}\ \varphi = \langle \varphi(e_1), \ldots, \varphi(e_n) \rangle$.

Если векторное пространство W также конечномерно, то определена матрица A линейного отображения φ относительно базисов $\mathcal{E}$ и $\mathcal{F}$. Столбцы матрицы A – это координаты векторов $\varphi(e_1), \ldots, \varphi(e_n)$ в базисе $\mathcal{F}$. Так как $\mathrm{Im}\ \varphi = \langle \varphi(e_1), \ldots, \varphi(e_n) \rangle$, а ранг матрицы равен рангу ее столбцов, получаем, что $\dim \mathrm{Im}\ \varphi = \mathrm{rk}\ A$. ■

Теорема 4.1. *Для любого линейного отображения* $\varphi : V \to W$ *конечномерного пространства* V *верно*

$$\dim \mathrm{Ker}\ \varphi + \dim \mathrm{Im}\ \varphi = \dim V.$$

Доказательство. Выберем базис пространства V специальным образом: сначала выберем базис $\{e_1, \ldots, e_k\}$ подпространства $\mathrm{Ker}\ \varphi$, а затем дополним его какими-нибудь векторами $e_{k+1}, \ldots, e_n$ до базиса пространства V (теорема 1.3). Так как по построению $\varphi(e_1) = \ldots = \varphi(e_k) = 0$, то из (7) следует, что

$$\mathrm{Im}\ \varphi = \langle \varphi(e_{k+1}), \ldots, \varphi(e_n) \rangle.$$

Докажем, что векторы $\varphi(e_{k+1}), \ldots, \varphi(e_n)$ линейно независимы, откуда и будет следовать утверждение теоремы.

Пусть

$$\lambda_1 \varphi(e_{k+1}) + \ldots + \lambda_{n-k} \varphi(e_n) = 0.$$

Рассмотрим вектор

$$x = \lambda_1 e_{k+1} + \ldots + \lambda_{n-k} e_n.$$

Предыдущее равенство означает, что $\varphi(x) = 0$, т.е.

$$x \in \mathrm{Ker}\ \varphi = \langle e_1, \ldots, e_k \rangle.$$

Так как $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ линейно независимы, то это возможно только при $\lambda_1 = \ldots = \lambda_{n-k} = 0$, что и требовалось доказать. ■

Пусть $\mathcal{E} = \{e_1, \ldots, e_n\}$ – произвольный базис пространства V, $\mathcal{F} = \{f_1, \ldots, f_m\}$ – произвольный базис пространства W, а A – матрица линейного отображения $\varphi : V \to W$ относительно базисов $\mathcal{E}$ и $\mathcal{F}$. Образом $\varphi(x)$ произвольного вектора x со столбцом координат X в базисе $\mathcal{E}$ является вектор, столбец координат которого в базисе $\mathcal{F}$ является AX. Поэтому

$$x \in \mathrm{Ker}\ \varphi \Leftrightarrow AX = 0.$$

Следовательно, $\mathrm{Ker}\ \varphi$ – это множество векторов, столбцы координат которых в базисе $\mathcal{E}$ составляют пространство решений однородной системы линейных уравнений с n неизвестными и матрицей коэффициентов A. Согласно доказанному в 1-ом семестре утверждению о размерности пространства решений однородной системы линейных уравнений,

$$\dim \mathrm{Ker}\ \varphi = n - \mathrm{rk}\ A.$$

Отсюда и из формулу (8), в частности, получаем другое доказательство теоремы 4.1 для конечномерных пространств V и W.

4.2. Определение линейного оператора и его матрицы.

Перейдем теперь к случаю, когда $V = W$.

Определение 4.3. Линейным оператором *(или* **линейным преобразованием***) векторного пространства V над полем F называется линейное отображение пространства V в себя.*

Более подробно, линейный оператор – это отображение $\mathcal{A} : V \to V$, удовлетворяющее условиям:

1) $\mathcal{A}(x + y) = \mathcal{A}x + \mathcal{A}y \ \ \forall x, y \in V$;

2) $\mathcal{A}(\lambda x) = \lambda \mathcal{A}x \ \ \forall \lambda \in F, \forall x \in V$.

(Мы будем обозначать линейные операторы рукописными заглавными латинскими буквами, а соответствующие им матрицы – курсивными заглавными буквами.)

В отличие от определения матрицы линейного отображения при определении матрицы линейного оператора принято брать лишь один базис.

Определение 4.4. Матрицей линейного оператора $\mathcal{A} : V \to V$ *в базисе* $\mathcal{E} = \{e_1, \ldots, e_n\}$ *пространства V называется квадратная матрица $A = (a_{i,j})$, определяемая из равенства*

$$\mathcal{A}e_j = \sum_{i=1}^{n} a_{i,j} e_i.$$

Иначе говоря, в j-м столбце матрицы A стоят координаты вектора $\mathcal{A}e_j$ в базисе $\mathcal{E}$.

Как и любое линейное отображение, линейный оператор $\mathcal{A} : V \to V$ однозначно определяется образами базисных векторов пространства V. В самом деле, для любого вектора $x = \sum_{i=1}^{n} x_i e_i$ имеем

$$\mathcal{A}x = \sum_{i=1}^{n} x_i \mathcal{A}e_i.$$

С другой стороны, для любых векторов $v_1, \ldots, v_n \in V$ существует единственный линейный оператор $\mathcal{A}$, переводящий базисные векторы $e_1, \ldots, e_n$ в $v_1, \ldots, v_n$ соответственно. Это оператор, переводящий вектор $x = \sum_{i=1}^{n} x_i e_i$ в вектор $\sum_{i=1}^{n} x_i v_i$. Следовательно, линейный оператор однозначно определяется своей матрицей, и любая квадратная матрица порядка n является матрицей некоторого линейного оператора (в данном базисе).

Как и для линейного отображения, получим явное выражение координат образа $y = \mathcal{A}x$ вектора $x \in V$ в базисе $\mathcal{E} = \{e_1, \ldots, e_n\}$. При $x = \sum_{j=1}^{n} x_j e_j$ имеем

$$y = \mathcal{A}x = \sum_{j=1}^{n} x_j \mathcal{A}e_j = \sum_{j=1}^{n} x_j (\sum_{i=1}^{n} a_{i,j} e_i) = \sum_{i=1}^{n} (\sum_{j=1}^{n} a_{i,j} x_j) e_i,$$

где

$$y = \sum_{i=1}^{n} y_i e_i.$$

В силу единственности разложения вектора по базису отсюда получаем, что $y_i = \sum_{j=1}^{n} a_{i,j} x_j$ для каждого $i = \overline{1,n}$ или в матричной форме

$$Y = AX, \tag{9}$$

где

$$A = (a_{i,j}) \text{ – матрица линейного оператора } \mathcal{A} \text{ в базисе } \mathcal{E},$$

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \text{ – столбцы координат векторов } x, y \text{ в базисе } \mathcal{E}.$$

Пусть $\widetilde{\mathcal{E}}$ – еще один базис пространства V, $T = T_{\mathcal{E}\to\widetilde{\mathcal{E}}}$ – матрица перехода от базиса $\mathcal{E}$ к базису $\widetilde{\mathcal{E}}$, $\widetilde{A}$ – матрица линейного оператора $\mathcal{A}$ в базисе $\mathcal{E}$. Тогда согласно формуле (6),

$$\widetilde{A} = T^{-1}AT. \tag{10}$$

Размерность подпространства $\mathrm{Im}\ \mathcal{A}$ называется **рангом** линейного оператора $\mathcal{A}$ и обозначается $\mathrm{rk}\ \mathcal{A}$. В силу утверждения 4.3 ранг линейного оператора совпадает с рангом матрицы этого оператора (в произвольном базисе). **Определителем** $\det\ \mathcal{A}$ линейного оператора $\mathcal{A}$ называется определитель матрицы этого оператора в произвольном базисе. Это определение корректно, так как учитывая формулу (10) и свойства определителя, имеем

$$\det\ \widetilde{A} = \det\ T^{-1}AT = \det\ T^{-1} \det\ A \det\ T = (\det\ T)^{-1} \det\ T \det\ A = \det\ A.$$

4.3. Алгебра линейных операторов.

На множестве $End\ V$ всех линейных операторов векторного пространства V над полем F имеются алгебраические операции: произведение (композиция) линейных операторов

$$(\mathcal{A}\mathcal{B})v = \mathcal{A}(\mathcal{B}v),$$

сумма линейных операторов

$$(\mathcal{A} + \mathcal{B})v = \mathcal{A}v + \mathcal{B}v$$

и умножение линейного оператора на число λ из F

$$(\lambda\mathcal{A})v = \lambda(\mathcal{A}v).$$

При этом удовлетворяются все аксиомы алгебры, то есть кольца и, в то же время, векторного пространства с аксиомой согласования $(\lambda x)y = \lambda(xy) = x(\lambda y)$.

Алгебра $End\ V$ является ассоциативной и обладает единицей. Единицей этой алгебры является тождественный линейный оператор, который мы будем обозначать буквой $\mathcal{I}$.

Как известно из первого семестра, другим примером алгебры является множество $M_n(F)$ всех квадратных матриц фиксированного произвольного порядка n с коэффициентами из поля F. Две алгебры $\mathfrak{A}$ и $\widetilde{\mathfrak{A}}$ называются изоморфными, если существует изоморфизм колец из первой алгебры во вторую, являющийся одновременно и изоморфизмом векторных пространств. Пишем $\mathfrak{A} \cong \widetilde{\mathfrak{A}}$.

Утверждение 4.4. *Для любого n-мерного векторного пространства V над полем F выполняется*

$$End\ V \cong M_n(F).$$

Доказательство. Зафиксируем базис $\mathcal{E} = \{e_1, \ldots, e_n\}$ в пространстве V. И рассмотрим отображение $\Phi : End\ V \to M_n(F)$, сопоставляющее каждому линейному оператору $\mathcal{A} : V \to V$ его матрицу A в базисе $\mathcal{E}$.

Так как любой линейный оператор однозначно определяется своей матрицей, и любая квадратная матрица порядка n является матрицей некоторого линейного оператора (в данном базисе), отображение Φ биективно.

Далее пусть $\Phi(\mathcal{A}) = A = (a_{i,j})$, $\Phi(\mathcal{B}) = (b_{i,j})$, $\Phi(\mathcal{C}) = C = (c_{i,j})$, где $\mathcal{C} = \mathcal{A}\mathcal{B}$. Тогда для любого $k = \overline{1, n}$

$$\mathcal{C}e_k = (\mathcal{A}\mathcal{B})e_k = \mathcal{A}(\sum_{j=1}^{n} b_{j,k}e_j) = \sum_{j=1}^{n} b_{j,k}\mathcal{A}e_j =$$

$$= \sum_{j=1}^{n} b_{j,k}(\sum_{i=1}^{n} a_{i,j}e_i) = \sum_{i=1}^{n}(\sum_{j=1}^{n} a_{i,j}b_{j,k})e_i.$$

С другой стороны,

$$\mathcal{C}e_k = \sum_{i=1}^{n} c_{i,k}e_i.$$

Откуда в силу единственности разложения вектора по базису получаем, что

$$c_{i,k} = \sum_{j=1}^{n} a_{i,j}b_{j,k} \ \ \forall i, k = \overline{1, n},$$

что в матричной форме означает

$$C = AB.$$

Таким образом, $\Phi(\mathcal{A}\mathcal{B}) = \Phi(\mathcal{A})\Phi(\mathcal{B})$.

Аналогично проверяется, что $\Phi(\lambda\mathcal{A}) = \lambda\Phi(\mathcal{A})$ и $\Phi(\mathcal{A} + \mathcal{B}) = \Phi(\mathcal{A}) + \Phi(\mathcal{B})$.
■

Следствие 4.1. *Для любого конечномерного векторного пространства V над полем F выполняется*

$$\dim End\ V = (\dim V)^2.$$

Определение 4.5. *Линейный оператор $\mathcal{B}$ называется* **обратным** *к линейному оператору $\mathcal{A}$, если*

$$\mathcal{AB} = \mathcal{BA} = \mathcal{I},$$

где $\mathcal{I}$ – тождественный оператор.

Оператор, обратный к линейному оператору $\mathcal{A}$, обозначается $\mathcal{A}^{-1}$.

Не для всякого линейного оператора существует обратный. Например, линейный оператор, проектирующий трехмерное пространство на плоскость, очевидно, не имеет обратного.

С понятием обратного линейного оператора связано понятие обратной матрицы. Матрица A^{-1} называется обратной к квадратной матрице A, если

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

где E – единичная матрица. Из первого семестра знаем, что квадратная матрица A имеет обратную тогда и только тогда, когда $\det A \neq 0$. Так как при заданном базисе между матрицами и линейными операторами имеется взаимно однозначное соответствие, сохраняющее операцию умножения, то *для того, чтобы линейный оператор $\mathcal{A}$ имел обратный, необходимо и достаточно, чтобы его матрица в каком-нибудь базисе имела бы определитель, отличный от нуля, т.е. имела бы ранг n.* Если линейный оператор имеет обратный, то он называется **невырожденным**. В противном случае линейный оператор называется **вырожденным**. Таким образом, в силу равенства (8) получаем, что *линейный оператор вырожден тогда и только тогда, когда* $\dim \operatorname{Im} \mathcal{A} < n$. Также в силу утверждения 4.1 имеем, что *линейный оператор вырожден тогда и только тогда, когда* $\operatorname{Ker} \mathcal{A} \neq \{0\}$.

Имея в виду возможность совершения над матрицами и линейными операторами всех естественных операций, дадим следующее

Определение 4.6. *Пусть $f(t) = \alpha_0 + \alpha_1 t + \ldots + \alpha_m t^m \in F[t]$ – многочлен от переменной t с коэффициентами из поля F, $\mathcal{A}$ – линейный оператор векторного пространства V. Тогда $f(\mathcal{A})$ определяется формулой*

$$f(\mathcal{A}) = \alpha_0\mathcal{I} + \alpha_1\mathcal{A} + \ldots + \alpha_m\mathcal{A}^m,$$

где $\mathcal{I}$ – тождественный оператор. В случае матрицы $A \in M_n(F)$ значение многочлена $f(t)$ от матрицы A определяется формулой

$$f(A) = \alpha_0 E + \alpha_1 A + \ldots + \alpha_m A^m,$$

где E – единичная матрица.

ПРИМЕР. Пусть $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$. Тогда

$$A^2 = \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix}, \ldots, A^m = \begin{pmatrix} a^m & 0 \\ 0 & b^m \end{pmatrix}.$$

Отсюда следует, что если $f(t) = \alpha_0 + \alpha_1 t + \ldots + \alpha_m t^m$, то

$$f(A) = \begin{pmatrix} f(a) & 0 \\ 0 & f(b) \end{pmatrix}.$$

Аналогичное равенство можно получить для любой диагональной матрицы.

4.4. Инвариантные подпространства.

Основным вопросом теории линейных операторов является вопрос о возможности приведения матрицы такого оператора к наиболее простому виду. При этом важную роль играет следующее понятие.

Определение 4.7. *Пусть $\mathcal{A} : V \to V$ – линейный оператор в векторном пространстве V. Подпространство $U \subset V$ называется* **инвариантным** *относительно оператора $\mathcal{A}$, если $\mathcal{A}u \in U$ для любого вектора $u \in U$.*

Пишут $\mathcal{A}U \subset U$.

Ограничение $\mathcal{A}|_U$ линейного оператора $\mathcal{A}$ на инвариантное подпространство U является линейным оператором в U.

Тривиальными инвариантными подпространствами являются подпространство, состоящее лишь из нуля, и все пространство.

Образ $\operatorname{Im} \mathcal{A}$ *и ядро* $\operatorname{Ker} \mathcal{A}$ *линейного оператора $\mathcal{A} : V \to V$ являются инвариантными подпространствами относительно* $\mathcal{A}$. Действительно, пусть $y \in \operatorname{Im} \mathcal{A}$. Тогда $\mathcal{A}y \in \operatorname{Im} \mathcal{A}$ в силу определения $\operatorname{Im} \mathcal{A}$. Точно так же, если $x \in \operatorname{Ker} \mathcal{A}$, то $\mathcal{A}x = 0 \in \operatorname{Ker} \mathcal{A}$. Этот простой факт будет использоваться в дальнейшем при приведении произвольного линейного оператора к простейшему виду.

Утверждение 4.5. *Пусть $\mathcal{E} = \{e_1, \ldots, e_m, e_{m+1}, \ldots, e_n\}$ – базис векторного пространства V такой, что $\mathcal{E}_1 = \{e_1, \ldots, e_m\}$ – базис инвариантного подпространства $U \subset V$ относительно линейного оператора $\mathcal{A} : V \to V$. Тогда в базисе $\mathcal{E}$ матрица для $\mathcal{A}$ имеет вид*

$$A = \begin{pmatrix} A_1 & C \\ 0 & B \end{pmatrix},$$

где A_1 – матрица ограничения оператора $\mathcal{A}$ на U.

Если, дополнительно, $W = \langle e_{m+1}, \ldots, e_n \rangle$ *– инвариантное подпространство относительно* $\mathcal{A}$*,* A_2 *– матрица ограничения оператора* $\mathcal{A}$ *на* W*, то* A *имеет вид*

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}.$$

Доказательство. Непосредственная проверка: по определению, j-й столбец матрицы линейного оператора есть набор координат образа j-го базисного вектора относительно этого же базиса. Остается применить определение инвариантного подпространства. ■

Более общо, если пространство V разложено в прямую сумму k инвариантных подпространств $V_1, \ldots, V_k$, то в базисе пространства V, составленного из базисов этих подпространств, матрица линейного оператора $\mathcal{A} : V \to V$ имеет вид

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_k \end{pmatrix},$$

где A_i – матрица ограничения $\mathcal{A}|_{V_i}$ линейного оператора $\mathcal{A}$ на V_i. В частности, если все $V_i = \langle e_i \rangle$ одномерны, то матрица A в базисе $\{e_1, \ldots, e_n\}$, составленном из базисов для $V_1, \ldots, V_n$, является диагональной

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

Определение 4.8. *Линейный оператор, матрица которого в некотором базисе диагональна, называется* **диагонализируемым.**

При этом базисные векторы $e_1, \ldots, e_n$ обладают следующим важным свойством: $\mathcal{A}e_i = \lambda_i e_i, i = \overline{1, n}$.

5. Собственные векторы и собственные значения.

5.1. Характеристический многочлен.

Как было уже отмечено, основная задача теории линейных операторов состоит в приведении матрицы линейного оператора к возможно более простому виду за счет выбора подходящего базиса. Для этого полезно знать инвариантные подпространства. Особую роль играют одномерные инвариантные подпространства. Их рассмотрение приводит к понятию собственного вектора.

Определение 5.1. *Пусть V – векторное пространство над полем F. Вектор $v \in V$, $v \neq 0$, называется* **собственным** *для линейного оператора $\mathcal{A} : V \to V$, если $\mathcal{A}v = \lambda v$ для подходящего числа $\lambda \in F$. При этом λ называется* **собственным значением** *для $\mathcal{A}$, отвечающим собственному вектору v.*

Нетрудно проверить, что если v – собственный вектор, то векторы αv, где $\alpha \in F$, образуют одномерное инвариантное подпространство. Обратно, все отличные от нуля векторы одномерного инвариантного подпространства являются собственными.

В базисе, состоящем из собственных векторов (если таковой существует), матрица оператора диагональна (причем диагональные элементы матрицы – собственные значения). Верно и обратное: если матрица линейного оператора в некотором базисе диагональна, то векторы этого базиса являются собственными для данного оператора.

Таким образом, *линейный оператор $\mathcal{A} : V \to V$ диагонализируем тогда и только тогда, когда в пространстве V есть базис из собственных векторов оператора $\mathcal{A}$.* Этим объясняется важность понятия собственного вектора и собственного значения для теории линейных операторов.

ПРИМЕР 1. Для оператора дифференцирования в пространстве многочленов единственным с точностью до пропорциональности собственным вектором является многочлен 1 (причем собственное значение равно 0). Таким образом, в этом случае из собственных векторов нельзя составить базис.

ПРИМЕР 2. Собственные векторы поворота на угол $\alpha \neq k\pi$ в трехмерном пространстве – это векторы, лежащие на оси поворота, причем соответствующее им собственное значение равно 1. При $\alpha = k\pi$ собственными (с собственным значением $(-1)^k$) являются также векторы, ортогональные оси поворота. Таким образом, базис из собственных векторов в этом примере существует только тогда, когда $\alpha = 0$ или π (если считать, что $0 \leqslant \alpha < 2\pi$).

Множество собственных значений линейного оператора называется его **спектром**. Для нахождения спектра линейного оператора необходимо ввести понятие характеристического многочлена.

Определение 5.2. *Пусть дан линейный оператор $\mathcal{A} : V \to V$ конечномерного векторного пространства V над полем F, $\mathcal{E}$ – некоторый базис в V, A – матрица для $\mathcal{A}$ в $\mathcal{E}$. Тогда многочлен степени n от переменной t вида*

$$\chi_A(t) = \det\,(A - tE)$$

называется **характеристическим многочленом матрицы** A, *а уравнение*

$$\chi_A(t) = 0$$

называется **характеристическим уравнением матрицы** A.

Утверждение 5.1. *Если A и $\widetilde{A}$ – матрицы линейного оператора $\mathcal{A} : V \to V$ в базисах $\mathcal{E}$ и $\widetilde{\mathcal{E}}$ соответственно, то характеристические многочлены этих матриц совпадают.*

Доказательство. Если $T = T_{\mathcal{E}\to\widetilde{\mathcal{E}}}$ – матрица перехода от базиса $\mathcal{E}$ к базису $\widetilde{\mathcal{E}}$, то учитывая формулу (10), имеем

$$\det\,(\widetilde{A} - tE) = \det\,(T^{-1}AT - tE) = \det\,(T^{-1}AT - tT^{-1}ET) =$$

$$= \det\,(T^{-1}(A - tE)T) = (\det\,T^{-1})\det\,(A - tE)(\det\,T) = \det\,(A - tE),$$

как и требовалось. ∎

Определение 5.3. Характеристическим многочленом $\chi_{\mathcal{A}}(t)$ линейного оператора $\mathcal{A} : V \to V$ *конечномерного векторного пространства V над полем F назовем характеристический многочлен $\chi_A(t)$ матрицы A линейного оператора $\mathcal{A}$ в некотором базисе, а уравнение*

$$\chi_{\mathcal{A}}(t) = 0$$

назовем **характеристическим уравнением линейного оператора** A.

Таким образом, утверждение 5.1 означает следующее:

Утверждение 5.2. *Определение характеристического многочлена линейного оператора корректно, то есть не зависит от выбора базиса.*

Замечательным следствием этого утверждения является то, что коэффициенты характеристического многочлена линейного оператора не зависят от выбора базиса, т. е. являются инвариантами линейного оператора. Например, коэффициент при t^{n-1} равен $(-1)^{n-1}\mathrm{tr}\,A$, откуда вытекает, что *след линейного оператора является его инвариантом.*

Теорема 5.1. *Число $\lambda \in F$ является собственным для линейного оператора $\mathcal{A} : V \to V$ тогда и только тогда, когда λ – корень характеристического многочлена $\chi_{\mathcal{A}}$.*

Доказательство. Пусть $\chi_{\mathcal{A}}(\lambda) = 0$. Рассмотрим какой-нибудь базис $\mathcal{E} = \{e_1, \ldots, e_n\}$ пространства V. И пусть A – матрица для $\mathcal{A}$ в этом базисе. Тогда $\det\,(A - \lambda E) = 0$, т.е. матрица $A - \lambda E$ является вырожденной. Из первого семестра известно, что в этом случае однородная система линейных уравнений $(A - \lambda E)X = 0$ имеет ненулевое решение $X = (x_1^0, \ldots, x_n^0)^t$. Значит, для ненулевого вектора $x = x_1^0 e_1 + \ldots + x_n^0 e_n$ выполняется $(\mathcal{A} - \lambda\mathcal{I})x = 0$ или $\mathcal{A}x = \lambda x$, так что x – собственный вектор для $\mathcal{A}$ с собственным значением λ. Для доказательства обратного утверждения достаточно прочитать только что написанное в обратном порядке. ■

По основной теореме алгебры комплексных чисел (из первого семестра) каждый многочлен ненулевой степени с комплексными коэффициентами имеет корень, поэтому из теоремы 5.1 получаем

Следствие 5.1. *Любой линейный оператор в конечномерном векторном пространстве над полем комплексных чисел имеет собственный вектор.*

Линейный оператор в вещественном векторном пространстве может не иметь собственных векторов, как показывает пример поворота плоскости на угол $\alpha \neq 0, \pi$. Однако использование комплексных чисел позволяет получить полезную информацию и о линейных операторах над полем вещественных чисел. Это достигается с помощью комплексификации.

5.2. Комплексификация вещественного пространства. Комплексные собственные значения линейного оператора в вещественном векторном пространстве.

Пусть V – некоторое векторное пространство над полем действительных чисел. Построим из него векторное пространство $V^{\mathbb{C}}$ над полем комплексных чисел аналогично тому, как из поля $\mathbb{R}$ строится поле $\mathbb{C}$. А именно, элементами пространства $V^{\mathbb{C}}$ будем считать пары (x, y), где $x, y \in V$. Определим сложение таких пар и умножение на комплексные числа по правилам

$$(x, y) + (z, w) = (x + z, y + w),$$

$$(\lambda + i\mu)(x, y) = (\lambda x - \mu y, \mu x + \lambda y).$$

Легко проверить, что при этом получается векторное пространство над полем $\mathbb{C}$. Полученное векторное пространство $V^{\mathbb{C}}$ называется **комплексификацией** пространства V.

Согласно данному определению, сложение пар вида $(x, 0)$ и умножение их на вещественное число сводится к соответствующим операциям над их первыми компонентами. Отождествим каждую пару вида $(x, 0)$ с вектором $x \in V$;

тогда пространство V окажется вложенным в $V^{\mathbb{C}}$ в виде вещественного подпространства. Так как $(x, y) = (x, 0) + (0, y) = (x, 0) + i(y, 0)$, каждый вектор (x, y) из $V^{\mathbb{C}}$ однозначно представим в виде

$$(x, y) = x + iy. \tag{11}$$

Любой базис $\{e_1, \ldots, e_n\}$ пространства V (над $\mathbb{R}$) является в то же время базисом пространства $V^{\mathbb{C}}$ (над $\mathbb{C}$). В самом деле, любой вектор $x + iy$ из $V^{\mathbb{C}}$ линейно выражается через $\{e_1, \ldots, e_n\}$ с комплексными коэффициентами:

$$x + iy = (x_1 + iy_1)e_1 + \ldots + (x_n + iy_n)e_n,$$

где $x = x_1e_1 + \ldots + x_ne_n$, $y = y_1e_1 + \ldots + y_ne_n \in V$ $(x_i, y_i \in \mathbb{R})$. Таким образом, система $\{e_1, \ldots, e_n\}$ является порождающей для $V^{\mathbb{C}}$ (над $\mathbb{C}$). Проверим, что система $\{e_1, \ldots, e_n\}$ является линейно независимой над $\mathbb{C}$. Если

$$(\lambda_1 + i\mu_1)e_1 + \ldots + (\lambda_n + i\mu_n)e_n = 0, \ \ \lambda_i, \mu_i \in \mathbb{R},$$

то

$$(\lambda_1e_1 + \ldots + \lambda_ne_n) + i(\mu_1e_1 + \ldots + \mu_ne_n) = 0,$$

т.е.

$$(\lambda_1e_1 + \ldots + \lambda_ne_n, \mu_1e_1 + \ldots + \mu_ne_n) = (0, 0).$$

Значит, $\lambda_1e_1 + \ldots + \lambda_ne_n = 0$ и $\mu_1e_1 + \ldots + \mu_ne_n = 0$. Отсюда в силу линейной независимости системы $\{e_1, \ldots, e_n\}$ над $\mathbb{R}$ получаем, что все коэффициенты λ_i, μ_j – нулевые.

Заметим, что хотя любой базис $\{e_1, \ldots, e_n\}$ пространства V (над $\mathbb{R}$) является в то же время базисом пространства $V^{\mathbb{C}}$ (над $\mathbb{C}$), в пространстве $V^{\mathbb{C}}$ существуют и другие базисы.

Любой линейный оператор $\mathcal{A}$ в пространстве V однозначно продолжается до линейного оператора $\mathcal{A}_{\mathbb{C}}$ в пространстве $V^{\mathbb{C}}$ по формуле:

$$\mathcal{A}_{\mathbb{C}}(x + iy) = \mathcal{A}x + i\mathcal{A}y.$$

При этом в базисе, составленном из вещественных векторов, оператор $\mathcal{A}_{\mathbb{C}}$ имеет такую же матрицу, как и оператор $\mathcal{A}$.

Оператор $\mathcal{A}_{\mathbb{C}}$ может иметь мнимые собственные значения и соответствующие им мнимые собственные векторы. Какой смысл они имеют в вещественных терминах?

Утверждение 5.3. *Пусть вектор $x + iy$ $(x, y \in V)$ является собственным вектором оператора $\mathcal{A}_{\mathbb{C}}$ с мнимым собственным значением $\lambda + i\mu$ $(\lambda, \mu \in \mathbb{R}, \mu \neq 0)$. Тогда $U = \langle x, y \rangle \subset V$ – двумерное инвариантное подпространство для $\mathcal{A}$, причем*

$$\begin{array}{rcrcr} \mathcal{A}x &=& \lambda x &-& \mu y, \\ \mathcal{A}y &=& \mu x &+& \lambda y. \end{array} \tag{12}$$

Доказательство. Имеем:

$$\mathcal{A}_{\mathbb{C}}(x+iy) = (\lambda + i\mu)(x+iy)$$

$$\Downarrow$$

$$\mathcal{A}x + i\mathcal{A}y = (\lambda x - \mu y) + i(\mu x + \lambda y)$$

$$\Downarrow$$

$$\begin{cases} \mathcal{A}x &= \lambda x - \mu y, \\ \mathcal{A}y &= \mu x + \lambda y, \end{cases}$$

т.е. $U = \langle x, y \rangle$ – инвариантное подпространство для $\mathcal{A}$. Докажем, что x, y – линейно независимые векторы в V. Действительно, если x, y линейно зависимы над $\mathbb{R}$, то $x+iy = \gamma z$, где $z \in V, \gamma \in \mathbb{C}$ и z является собственным вектором оператора $\mathcal{A}_{\mathbb{C}}$ с тем же собственным значением $\lambda + i\mu$, что невозможно при $\mu \neq 0$. ■

Равенство (12) означает, что в базисе $\{x, y\}$ пространства U оператор $\mathcal{A}|_U$ имеет матрицу

$$\begin{pmatrix} \lambda & \mu \\ -\mu & \lambda \end{pmatrix}.$$

ПРИМЕР 3. Оператор $\mathcal{A}$ поворота евклидовой плоскости на угол α в ортонормированном базисе $\{e_1, e_2\}$ имеет матрицу

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Следовательно, вектор $e_1 + ie_2$ является собственным вектором оператора $\mathcal{A}_{\mathbb{C}}$ с собственным значением $\cos\alpha - i\sin\alpha$, а вектор $e_1 - ie_2$ – собственным вектором с собственным значением $\cos\alpha + i\sin\alpha$. Таким образом, в комплексном пространстве матрица поворота может быть приведена к диагональному виду

$$\begin{pmatrix} \cos\alpha - i\sin\alpha & 0 \\ 0 & \cos\alpha + i\sin\alpha \end{pmatrix}.$$

В качестве следствия утверждения 5.3 получается важная

Теорема 5.2. *Для любого линейного оператора $\mathcal{A}$ конечномерного векторного пространства над полем вещественных чисел существует одномерное или двумерное инвариантное подпространство.*

Доказательство. Действительно, если характеристический многочлен линейного оператора $\mathcal{A}$ имеет хотя бы один вещественный корень, то этот оператор имеет хотя бы один собственный вектор, а следовательно, имеет и одномерное инвариантное подпространство (этим вектором порожденное).

Если же среди корней характеристического многочлена нет вещественных, то есть хотя бы один мнимый корень (по основной теореме алгебры комплексных чисел), а следовательно, по утверждению 5.3 оператор $\mathcal{A}$ имеет двумерное инвариантное подпространство. ■

6. Собственные векторы (продолжение).

6.1. Собственные подпространства.

Множество V_λ всех собственных векторов линейного оператора $\mathcal{A}$ с собственным значением λ, объединенных с нулевым вектором, является, очевидно, подпространством в векторном пространстве V.

Определение 6.1. *Подпространство V_λ называется* **собственным подпространством** *оператора $\mathcal{A}$, отвечающим собственному значению λ.*

Итак, собственное подпространство V_λ состоит из всех векторов $x \in V$ (в том числе нулевого вектора) таких, что $\mathcal{A}x = \lambda x$ или $(\mathcal{A} - \lambda\mathcal{I})x = 0$, т.е. $V_\lambda = \mathrm{Ker}\ (\mathcal{A} - \lambda\mathcal{I})$. Размерность собственного пространства V_λ равна $n - rk(\mathcal{A} - \lambda\mathcal{I})$, где $n = \dim V$.

Утверждение 6.1. *Если k – кратность собственного значения λ линейного оператора $\mathcal{A}$, то* $\dim V_\lambda \leqslant k$*.*

Доказательство. Допустим, $\dim V_\lambda = m > k$. Так как любое собственное подпространство является инвариантным относительно $\mathcal{A}$, применим к V_λ утверждение 4.5. Тогда в базисе, в котором первые m векторов $\{e_1, \ldots, e_m\}$ образуют базис подпространства V_λ, матрица линейного оператора $\mathcal{A}$ имеет вид

$$A = \begin{pmatrix} A_1 & C \\ 0 & B \end{pmatrix},$$

где A_1 – матрица ограничения оператора $\mathcal{A}$ на V_λ в базисе $\{e_1, \ldots, e_m\}$. Причем A_1 – квадратная матрица порядка m, равная λE в силу того, что все векторы в V_λ и, в частности, $e_1, \ldots, e_m$ – собственные векторы с собственным значением λ. Вычислим характеристический многочлен линейного оператора $\mathcal{A}$, используя теорему об определителе с углом нулей из 1-го семестра:

$$\det\ (A - tE) = \det\ \begin{pmatrix} A_1 - tE & C \\ 0 & B - tE \end{pmatrix} =$$

$$= \det\ (A_1 - tE)\det\ (B - tE) = (\lambda - t)^m \det\ (B - tE).$$

Отсюда по определению кратности следует, что кратность корня λ не меньше, чем m, что противоречит предположению о том, что кратность корня λ равна k. ■

Отметим, что собственному значению кратности k может отвечать собственное подпространство размерности, меньше, чем k. Например, линейный оператор в двумерном пространстве, задаваемый матрицей

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

имеет единственное собственное значение $\lambda = 1$, кратность этого собственного значения равна 2, а размерность отвечающего ему собственного подпространства равна одному. (Проверьте это!)

Важным утверждением о собственных векторах линейного оператора является следующее

Утверждение 6.2. *Пусть $\mathcal{A} : V \to V$ – линейный оператор векторного пространства V, $v_1, \dots, v_m$ – собственные векторы с попарно различными собственными значениями $\lambda_1, \dots, \lambda_m$. Тогда $v_1, \dots, v_m$ линейно независимы.*

Доказательство. Докажем утверждение индукцией по m. При $m = 1$ доказывать нечего. Пусть $m > 1$. Если

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_m v_m = 0$$

– нетривиальная линейная комбинация, то все ее коэффициенты α_i отличны от нуля в силу предположения индукции. Применим к обеим частям равенства линейный оператор $\mathcal{A}$. Получим

$$\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \ldots + \alpha_m \lambda_m v_m = 0.$$

Умножая первое равенство на $(-\lambda_1)$ и прибавляя ко второму, получим линейную комбинацию векторов $v_2, \dots, v_m$:

$$\alpha_2(\lambda_2 - \lambda_1)v_2 + \ldots + \alpha_m(\lambda_m - \lambda_1)v_m = 0.$$

Так как по предположению индукции векторы $v_2, \dots, v_m$ линейно независимы, отсюда следует, что $\alpha_i(\lambda_i - \lambda_1) = 0$ для всех $i = \overline{2, m}$. Так как все собственные значения $\lambda_1, \dots, \lambda_m$ – попарно различные, сокращаем на $(\lambda_i - \lambda_1)$ и получаем, что $\alpha_i = 0$ для всех $i = \overline{2, m}$, противоречие. ∎

Следствие 6.1. *Сумма собственных подпространств является прямой суммой.*

Заметим, что в общем случае сумма всех собственных подпространств, отвечающих всем различным собственным значениям линейного оператора, не обязана совпадать со всем векторным пространством V.

Следствие 6.2. *Если характеристический многочлен линейного оператора $\mathcal{A} : V \to V$ в n-мерном векторном пространстве V имеет n различных корней, то существует базис из собственных векторов оператора $\mathcal{A}$.*

Доказательство. Каждому корню λ_i характеристического многочлена отвечает хотя бы один собственный вектор. Так как соответствующие этим векторам собственные значения (корни характеристического уравнения) все различны, то согласно утверждению 6.2 мы имеем n линейно независимых собственных векторов, которые и будут образовывать искомый базис. ∎

Условие, указанное в следствии 6.2, не является необходимым для существования базиса из собственных векторов. Так, для тождественного оператора $\mathcal{I}$ все векторы являются собственными, и, стало быть, любой базис состоит из собственных векторов, однако его характеристический многочлен $\chi_{\mathcal{I}}(t) = (1-t)^n$ имеет единственный (но n-кратный) корень 1.

Теорема 6.1. *Пусть V – векторное пространство над полем F размерности n. Для существования базиса из собственных векторов линейного оператора $\mathcal{A} : V \to V$ необходимо и достаточно, чтобы выполнялись следующие условия:*

1) характеристический многочлен $\chi_{\mathcal{A}}(t)$ разлагается на линейные множители;

2) размерность каждого собственного подпространства равна кратности соответствующего корня характеристического многочлена $\chi_{\mathcal{A}}(t)$.

Доказательство. Пусть $\lambda_1, \ldots, \lambda_s$ – все различные корни характеристического многочлена $\chi_{\mathcal{A}}(t)$ и $k_1, \ldots, k_s$ – их кратности. Для каждого $i = \overline{1,s}$ через V_{λ_i} обозначим собственное подпространство, отвечающее собственному значению λ_i. В этих обозначениях условие 1) теоремы означает, что $\sum_{i=1}^{s} k_i = n$, а условие 2) означает, что $\dim V_{\lambda_i} = k_i$ для каждого $i = \overline{1,s}$.

Пусть существует базис $\mathcal{E}$ из собственных векторов линейного оператора $\mathcal{A}$. Для каждого $i = \overline{1,s}$ рассмотрим линейную оболочку V_i всех векторов из базиса $\mathcal{E}$, которые отвечают собственному значению λ_i. Имеем, $V = V_1 \oplus \ldots \oplus V_s$. Также, очевидно, $V_i \subset V_{\lambda_i}$. Учитывая, что сумма собственных подпространств является прямой (Следствие 6.1), получаем, что

$$V = V_1 \oplus \ldots \oplus V_s \subset V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_s} \subset V.$$

Таким образом, $V = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_s}$. По утверждению 6.1 $\dim V_{\lambda_i} \leqslant k_i$ для каждого $i = \overline{1,s}$, и, значит, учитывая, что размерность прямой суммы подпространств равна сумме размерностей слагаемых (Следствие 2.1),

$$n = \dim V = \sum_{i=1}^{s} \dim V_{\lambda_i} \leqslant \sum_{i=1}^{s} k_i \leqslant n.$$

Что возможно только в случае, когда $\sum_{i=1}^{s} k_i = n$ и $\dim V_{\lambda_i} = k_i$ для всех $i = \overline{1,s}$.

Обратно, пусть $\sum_{i=1}^{s} k_i = n$ и $\dim V_{\lambda_i} = k_i$. Учитывая следствие 6.1 и следствие 2.1, имеем

$$\dim V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_s} = \sum_{i=1}^{s} \dim V_{\lambda_i} = \sum_{i=1}^{s} k_i = n = \dim V.$$

Следовательно, по теореме 1.4 получаем, что $V = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_s}$. Значит, объединение базисов из собственных подпространств V_{λ_i}, даст базис для всего пространства V. Этот базис и будет искомым. ■

6.2. Приведение линейного оператора к треугольному виду. Теорема Гамильтона-Кэли.

При рассмотрении вопросов о приведении матриц линейных операторов к удобному виду нам придется накладывать ограничение на поле F, состоящее в алгебраической замкнутости поля.

Определение 6.2. *Поле называется* **алгебраически замкнутым**, *если всякий многочлен положительной степени над этим полем имеет хотя бы один корень.*

ПРИМЕР. Из основной теоремы алгебры комплексных чисел (из первого семестра) следует, что поле комплексных чисел $\mathbb{C}$ алгебраически замкнуто.

На примере поля действительных чисел видно, что не всякое поле является алгебраически замкнутым. Впрочем, приведем без доказательства следующий факт:

Каждое поле F может быть рассмотрено (с точностью до изоморфизма!) как подполе алгебраически замкнутого поля K.

Это позволяет любой линейный оператор, действующий в векторном пространстве размерности n над полем F рассмотреть как оператор, действующий в векторном пространстве той же размерности над полем K. Это соображение уже было нами использовано в разделе 5.2, когда линейный оператор $\mathcal{A}$ в вещественном пространстве V был продолжен до линейного оператора $\mathcal{A}_{\mathbb{C}}$ в комплексном пространстве $V^{\mathbb{C}}$, комплексификации пространства V. При этом в базисе, составленном из вещественных векторов, оператор $\mathcal{A}_{\mathbb{C}}$ имел такую же матрицу, как и оператор $\mathcal{A}$. Аналогичное имеем и в общем случае. Причем, в силу изоморфизма пространств одной и той же размерности, достаточно рассмотреть вложение векторных пространств $F^n \subset K^n$. Если $\mathcal{A} : F^n \to F^n$ – линейный оператор, A – его матрица в базисе из единичных строк, то те же векторы образуют базис и в K^n и можно продолжить действие линейного оператора $\mathcal{A}$ до оператора $\mathcal{A}_K$ на K^n, считая матрицей нового оператора все ту же матрицу A в указанном базисе векторного пространства K^n. Фактически действие обоих операторов на векторах - столбцах состоит в умножении на матрицу A.

Теорема 6.2. *Пусть $\mathcal{A} : V \to V$ – линейный оператор конечномерного векторного пространства V над алгебраически замкнутым полем F. Тогда существует базис $\mathcal{F} = \{f_1, \ldots, f_n\}$, в котором матрица линейного оператора $\mathcal{A}$ является треугольной.*

Доказательство. Докажем утверждение теоремы индукцией по $n = \dim V$. При $n = 1$ доказывать нечего. Пусть $n > 1$.

Отметим, что согласно формуле (10) на матричном языке утверждение теоремы эквивалентно тому, что для любой квадратной матрицы A порядка n над полем F существует невырожденная матрица T порядка n такая, что матрица $\widetilde{A} = T^{-1}AT$ – треугольная.

Рассмотрим характеристический многочлен $\chi_{\mathcal{A}}(t)$ для $\mathcal{A}$. В силу алгебраической замкнутости поля F у него есть корень λ. По теореме 5.1 это означает, что у линейного оператора $\mathcal{A}$ есть собственный вектор e_1 с собственным значением λ. Так как e_1 – ненулевой вектор, к нему можно применить теорему 1.3 и дополнить e_1 векторами $e_2, \ldots, e_n$ до базиса $\mathcal{E}$ в V. Так как $U = \langle e_1 \rangle$ является одномерным инвариантным подпространством, по утверждению 4.5 матрица A линейного оператора $\mathcal{A}$ в базисе $\mathcal{E}$ будет состоять из четырех блоков:

$$A = \begin{pmatrix} \lambda & u \\ 0 & B \end{pmatrix},$$

где u – строка длины $n-1$, B – квадратная матрица порядка $n-1$.

По предположению индукции найдется невырожденная матрица S порядка $n-1$ такая, что $D = S^{-1}BS$ – треугольная матрица. Построим квадратную матрицу T порядка n из четырех блоков

$$T = \begin{pmatrix} 1 & 0 \\ 0 & S \end{pmatrix}$$

с разбиением таким же, как и для A. Полученная матрица T – невырожденная, так как по теореме об определителе матрицы с углом нулей, доказанной в первом семестре, $\det T = \det S \neq 0$. Причем, используя правило сокращенного умножения блочных матриц, имеем

$$T^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & S^{-1} \end{pmatrix}$$

и далее

$$\widetilde{A} = T^{-1}AT = \begin{pmatrix} 1 & 0 \\ 0 & S^{-1} \end{pmatrix} \begin{pmatrix} \lambda & u \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & S \end{pmatrix} =$$

$$= \begin{pmatrix} \lambda & u \\ 0 & S^{-1}B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & S \end{pmatrix} = \begin{pmatrix} \lambda & uS \\ 0 & S^{-1}BS \end{pmatrix} = \begin{pmatrix} \lambda & uS \\ 0 & D \end{pmatrix}.$$

Очевидно, что полученная матрица $\widetilde{A}$ – треугольная. Теорема доказана в силу сделанного замечания о матрицах. Отметим, что при этом матрица T – матрица перехода из базиса $\mathcal{E}$ к искомому базису $\mathcal{F}$. ■

Заметим, что если поле не алгебраически замкнуто, то базиса, в котором матрица линейного оператора является треугольной, может и не быть, так как он предполагает наличие хотя бы одного собственного вектора. В качестве примера можно опять рассмотреть поворот евклидовой плоскости на угол $\alpha \neq 0, \pi$.

Также заметим, что у одного и того же линейного оператора может быть много треугольных видов:

ПРИМЕР. Рассмотрим трехмерное пространство V и базис $\{e_1, e_2, e_3\}$ в этом пространстве. Зададим линейный оператор по правилу

$$\mathcal{A}(x_1e_1 + x_2e_3 + x_3e_3) = x_1.$$

Тогда в базисе $\{e_1, e_2, e_3\}$ этот линейный оператор имеет одну треугольную матрицу

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

в базисе $\{e_1, e_1 + e_2, e_3\}$ – другую треугольную матрицу

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

а в базисе $\{e_1, e_1 + e_2, e_1 + e_3\}$ – треугольную матрицу третьего вида

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

ЗАДАЧА. Используя теорему 5.2, докажите, что над полем $\mathbb{R}$ любой линейный оператор приводим к блочно-треугольному виду, причем блоки имеют порядок не выше двух.

Одним из замечательных следствий теоремы 6.2 является следующая весьма важная теорема Гамильтона - Кэли.

Следствие 6.3. (Теорема Гамильтона - Кэли) *Всякий линейный оператор $\mathcal{A}$ в конечномерном векторном пространстве V является корнем своего характеристического многочлена. Иными словами,*

$$\chi_{\mathcal{A}}(\mathcal{A}) = 0.$$

Доказательство. Прежде всего отметим, что из доказательства утверждения 4.4 следует, что если $f(t)$ – произвольный многочлен с коэффициентами из F, то для матрицы A линейного оператора $\mathcal{A}$ в произвольном базисе выполняется

$$f(\mathcal{A}) = 0 \Leftrightarrow f(A) = 0.$$

1 случай. Пусть $\mathcal{A}$ – линейный оператор в векторном пространстве V над алгебраически замкнутым полем F. Как мы знаем (см. утверждение 5.2), характеристический многочлен не зависит от выбора базиса, так что в силу

теоремы 6.2 мы имеем право выбрать базис $\mathcal{F}$, в котором матрица для $\mathcal{A}$ имеет треугольный вид

$$\widetilde{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n,n} \end{pmatrix}.$$

В силу сделанного выше замечания достаточно доказать, что $\chi_{\mathcal{A}}(\widetilde{A}) = 0$. Вычислим характеристический многочлен в базисе $\mathcal{F}$:

$$\chi_{\mathcal{A}}(t) = det(\widetilde{A} - tE) = (a_{1,1} - t)(a_{2,2} - t) \ldots (a_{n,n} - t).$$

Таким образом, требуется доказать, что

$$(a_{1,1}E - \widetilde{A})(a_{2,2}E - \widetilde{A}) \ldots (a_{n,n}E - \widetilde{A}) = 0.$$

Это можно сделать индукцией по n. Разобъем матрицу $\widetilde{A}$ на четыре блока:

$$\widetilde{A} = \begin{pmatrix} a_{1,1} & u \\ 0 & B \end{pmatrix},$$

где u – строка длины $n-1$, а B – треугольная матрица порядка $n-1$:

$$B = \begin{pmatrix} a_{2,2} & \cdots & a_{2,n} \\ & \ddots & \vdots \\ 0 & & a_{n,n} \end{pmatrix}.$$

По предположению индукции имеем

$$(a_{2,2}E - B) \cdot \ldots \cdot (a_{n,n}E - B) = 0.$$

Используя правило сокращенного умножения блочных матриц, получаем

$$(a_{2,2}E - \widetilde{A}) \cdot \ldots \cdot (a_{n,n}E - \widetilde{A}) =$$

$$= \left(\begin{pmatrix} a_{2,2} & 0 \\ 0 & a_{2,2}E \end{pmatrix} - \begin{pmatrix} a_{1,1} & u \\ 0 & B \end{pmatrix} \right) \cdot \ldots \cdot \left(\begin{pmatrix} a_{n,n} & 0 \\ 0 & a_{n,n}E \end{pmatrix} - \begin{pmatrix} a_{1,1} & u \\ 0 & B \end{pmatrix} \right) =$$

$$= \begin{pmatrix} a_{2,2} - a_{1,1} & -u \\ 0 & a_{2,2}E - B \end{pmatrix} \cdot \ldots \cdot \begin{pmatrix} a_{n,n} - a_{1,1} & -u \\ 0 & a_{n,n}E - B \end{pmatrix} =$$

$$= \begin{pmatrix} (a_{2,2} - a_{1,1}) \ldots (a_{n,n} - a_{1,1}) & \widetilde{u} \\ 0 & (a_{2,2}E - B) \ldots (a_{n,n}E - B) \end{pmatrix} =$$

$$= \begin{pmatrix} (a_{2,2} - a_{1,1}) \ldots (a_{n,n} - a_{1,1}) & \widetilde{u} \\ 0 & 0 \end{pmatrix},$$

где $\widetilde{u}$ – некоторая новая строка длины $n-1$. Следовательно,

$$(a_{1,1}E - \widetilde{A})(a_{2,2}E - \widetilde{A}) \cdot \ldots \cdot (a_{n,n}E - \widetilde{A}) =$$

$$= \left(\begin{pmatrix} a_{1,1} & 0 \\ 0 & a_{1,1}E \end{pmatrix} - \begin{pmatrix} a_{1,1} & u \\ 0 & B \end{pmatrix} \right) \cdot \begin{pmatrix} (a_{2,2} - a_{1,1}) \ldots (a_{n,n} - a_{1,1}) & \widetilde{u} \\ 0 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & -u \\ 0 & a_{1,1}E - B \end{pmatrix} \cdot \begin{pmatrix} (a_{2,2} - a_{1,1}) \ldots (a_{n,n} - a_{1,1}) & \widetilde{u} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

как и требовалось.

2 случай. Пусть $\mathcal{A}$ – линейный оператор в векторном пространстве V над произвольным полем F, а A – матрица этого линейного оператора в некотором базисе $\mathcal{E}$. Как было отмечено в начале этого раздела, поле F можно рассматривать как подполе алгебраически замкнутого поля K, и можно продолжить действие линейного оператора $\mathcal{A}$ до оператора $\mathcal{A}_K$ в векторном пространстве V_K над полем K, считая матрицей нового оператора все ту же матрицу A в базисе $\mathcal{E}$, рассматриваемого уже как базис векторного пространства V_K. В силу только что доказанного для алгебраически замкнутого поля, мы можем утверждать, что матрица A – корень характеристического многочлена $\chi_{\mathcal{A}_K}(t) = \det\,(A - tE)$, который, очевидно, не зависит от того, рассматриваем ли мы матрицу A как матрицу с элементами из F или как матрицу с элементами из K, и равен $\chi_{\mathcal{A}}(t)$. ■

Пользуясь теоремой Гамильтона-Кэли, можно свести вычисление любого многочлена $f \in F[t]$ от линейного оператора $\mathcal{A}$ к вычислению многочлена степени меньше $n = \dim V$ от этого оператора. Действительно, разделим f на $\chi_{\mathcal{A}}$ с остатком:

$$f(t) = \chi_{\mathcal{A}}(t)q(t) + r(t), \quad \deg r(t) < n.$$

Тогда

$$f(\mathcal{A}) = r(\mathcal{A}).$$

7. Жорданова нормальная форма (начало).

7.1. Минимальный многочлен.

Пусть дан линейный оператор $\mathcal{A} : V \to V$ конечномерного векторного пространства V.

Определение 7.1. *Многочлен $f(t) \in F[t]$ называется* **аннулирующим многочленом** *для линейного оператора $\mathcal{A}$, если $f(\mathcal{A}) = 0$. Аналогично определяется аннулирующий многочлен для матрицы.*

По теореме Гамильтона-Кэли любой оператор $\mathcal{A}$ и соответствующая ему матрица A (в любом базисе) аннулируется своим характеристическим многочленом.

Определение 7.2. *Аннулирующий многочлен для линейного оператора $\mathcal{A}$ (соотв., для матрицы A) наименьшей степени со старшим коэффициентом 1 называется* **минимальным многочленом** *для оператора $\mathcal{A}$ (соотв., для матрицы A) и обозначается $\mu_{\mathcal{A}}(t)$ (соотв., $\mu_A(t)$).*

Согласно теореме Гамильтона - Кэли, $\chi_{\mathcal{A}}(\mathcal{A}) = 0$. Это означает, что степень минимального многочлена не превосходит $n = \dim V$. При этом, возможен разброс степени от 1 до n.

ПРИМЕР 1. Если $\mathcal{I}$ - тождественный оператор в произвольном n-мерном пространстве, то $\mu_{\mathcal{I}}(t) = t - 1$, т.е. $\deg \mu_{\mathcal{I}} = 1$.

ПРИМЕР 2. Если $\mathcal{A}$ – поворот евклидовой плоскости V^2 на $\pi/2$, то $\mu_{\mathcal{A}}(t) = t^2 + 1$, т.е. $\deg \mu_{\mathcal{A}} = 2 = \dim V^2$.

Простейшие свойства минимального многочлена описаны в следующей лемме.

Лемма 7.1. *Пусть $\mathcal{A} : V \to V$ – линейный оператор конечномерного векторного пространства V. Тогда*

1) минимальный многочлен линейного оператора равен минимальному многочлену его матрицы в любом базисе;

2) минимальный многочлен является делителем любого аннулирующего многочлена;

3) минимальный многочлен обладает свойством единственности;

4) наборы корней у минимального многочлена и характеристического многочлена совпадают.

Доказательство.

1) Из доказательства утверждения 4.4 следует, что если $f(t)$ – произвольный многочлен с коэффициентами из F, то для матрицы A линейного оператора $\mathcal{A}$ в произвольном базисе выполняется

$$f(\mathcal{A}) = 0 \Leftrightarrow f(A) = 0.$$

Откуда и получаем первое утверждение.

2) Второе утверждение вытекает из алгоритма деления с остатком. Пусть $f(t)$ – произвольный аннулирующий многочлен для $\mathcal{A}$. Поделим многочлен $f(t)$ на минимальный многочлен $\mu_{\mathcal{A}}(t)$ линейного оператора $\mathcal{A}$ с остатком:

$$f(t) = \mu_{\mathcal{A}}(t)q(t) + r(t),$$

при этом для остатка $r(t)$ должно выполняться $deg\, r(t) < deg\, \mu_{\mathcal{A}}(t)$. Подставляя $\mathcal{A}$ вместо t, получаем

$$r(\mathcal{A}) = f(\mathcal{A}) - \mu_{\mathcal{A}}(\mathcal{A})q(\mathcal{A}) = 0 - 0q(\mathcal{A}) = 0,$$

т.е. $r(t)$ – аннулирующий для $\mathcal{A}$. В силу минимальности степени многочлена $\mu_{\mathcal{A}}(t)$, получаем, что $r(t) \equiv 0$, т.е. деление происходит нацело.

3) Третье утверждение справедливо, так как любой из двух минимальных многочленов линейного оператора делит другой, а их старшие коэффициенты совпадают.

4) Так как характеристический многочлен $\chi_{\mathcal{A}}(t)$ линейного оператора $\mathcal{A}$ является его аннулирующим многочленом, из второго утверждения получаем, что минимальный многочлен $\mu_{\mathcal{A}}(t)$ делит $\chi_{\mathcal{A}}(t)$. Значит, корни минимального многочлена являются корнями характеристического многочлена. Обратно, пусть λ – корень характеристического многочлена. Тогда у $\mathcal{A}$ есть собственный вектор v с собственным значением λ. Применим к вектору v оператор $\mu_{\mathcal{A}}(\mathcal{A})$. С одной стороны, расписав минимальный многочлен явно:

$$\mu_{\mathcal{A}}(t) = t^k + b_{k-1}t^{k-1} + \ldots + b_1 t + b_0$$

и подставив линейный оператор $\mathcal{A}$ вместо переменной t, получаем

$$\begin{aligned} \mu_{\mathcal{A}}(\mathcal{A})v = (\mathcal{A}^k + b_{k-1}\mathcal{A}^{k-1} + \ldots + b_1\mathcal{A} + b_0\mathcal{I})v = \\ = \mathcal{A}^k v + b_{k-1}\mathcal{A}^{k-1}v + \ldots + b_1\mathcal{A}v + b_0\mathcal{I}v = \\ = \lambda^k v + b_{k-1}\lambda^{k-1}v + \ldots + b_1\lambda v + b_0 v = \\ = (\lambda^k + b_{k-1}\lambda^{k-1} + \ldots + b_1\lambda + b_0)v = \mu_{\mathcal{A}}(\lambda)v. \end{aligned}$$

С другой стороны, так как $\mu_{\mathcal{A}}(\mathcal{A}) = 0$, имеем

$$\mu_{\mathcal{A}}(\mathcal{A})v = 0v = 0.$$

Откуда

$$\mu_{\mathcal{A}}(\lambda)v = 0.$$

Так как $v \neq 0$, то число $\mu_{\mathcal{A}}(\lambda)$ равно нулю, т.е. λ – корень минимального многочлена, что и требовалось доказать. ■

В качестве примера применение минимального многочлена сформулируем утверждение, которое будет доказано в разделе 8:

Линейный оператор $\mathcal{A} : V \to V$ конечномерного векторного пространства V над алгебраически замкнутым полем F диагонализируем тогда и только тогда, когда его минимальный многочлен не имеет кратных корней.

Напомним, что линейный оператор $\mathcal{A}$ называется диагонализируемым, если в некотором базисе его матрица диагональна. Существование базиса из собственных векторов эквивалентно диагонализируемости.

7.2. Жорданова нормальная форма. Основные определения и вспомогательные утверждения.

При исследовании линейного оператора, действующего в n-мерном векторном пространстве над полем F, желательно выбрать базис так, чтобы матрица линейного оператора имела наиболее простой вид. Если линейный оператор имеет базис из собственных векторов, то его матрица в этом базисе является диагональной. В частности, это верно в случае, когда все корни характеристического уравнения линейного оператора принадлежат полю F и различны (см. следствие 6.2).

Однако число линейно независимых собственных векторов у линейного оператора может быть меньше, чем n. Такой оператор заведомо не может иметь диагональную матрицу ни в каком базисе. Возникает вопрос: каков простейший вид матрицы такого линейного оператора?

При рассмотрении этого вопроса нам придется накладывать ограничение на поле F, состоящее в алгебраической замкнутости поля. При таком ограничении на поле для произвольного линейного оператора в этой главе мы укажем базис, в котором матрица оператора имеет сравнительно простой вид (так называемую жорданову нормальную форму). В случае, когда число линейно независимых собственных векторов оператора равно размерности пространства, эта нормальная форма совпадает с диагональной.

Определение 7.3. *Квадратная матрица $J_k(\lambda)$ порядка k вида*

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

называется **жордановой клеткой** *порядка k с собственным значением λ на диагонали. У этой матрицы все диагональные элементы равны λ, над главной диагональю параллельно ей расположен один ряд из единиц, а все остальные элементы равны нулю. В случае $k = 1$ рассматриваемая матрица сводится к единственному числу λ.*

Определение 7.4. *Блочно - диагональная матрица J вида*

$$\begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{k_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{k_s}(\lambda_s) \end{pmatrix}, \tag{13}$$

где $J_{k_i}(\lambda_i)$ – жордановы клетки, называется **жордановой матрицей.**

Определение 7.5. *Базис векторного пространства V, в котором матрица J для линейного оператора $\mathcal{A} : V \to V$ является жордановой, называется* **жордановым базисом**. *При этом матрица J называется* **жордановой нормальной формой** *линейного оператора $\mathcal{A}$.*

Если матрица линейного оператора в некотором жордановом базисе $\mathcal{E}$ имеет вид (13), то $\mathcal{E} = \{e_{1,1}, \ldots, e_{1,k_1}; \ldots; e_{s,1}, \ldots e_{s,k_s}\}$, где для каждого $i = 1, \ldots, s$ подсистема $\{e_{i,1}, \ldots, e_{i,k_i}\}$ соответствует столбцам матрицы, которые образуют жорданову клетку $J_{k_i}(\lambda_i)$. Будем называть систему $\{e_{i,1}, \ldots, e_{i,k_i}\}$ *жордановым базисом для клетки $J_{k_i}(\lambda_i)$.*

Утверждение 7.1. *Пусть V – n-мерное векторное пространство, $\mathcal{A} : V \to V$ – линейный оператор в V, для которого существует жорданов базис, а J – (жорданова) матрица в этом базисе. Тогда жорданов базис для J является объединением жордановых базисов для жордановых клеток, а для одной клетки $J_k(\lambda)$ он имеет вид $\{e_1, e_2, \ldots, e_{k-1}, e_k\}$, где*

$$\begin{gathered} \mathcal{A}(e_1) = \lambda e_1, \\ \mathcal{A}(e_2) = \lambda e_2 + e_1, \\ \cdots \\ \mathcal{A}(e_{k-1}) = \lambda e_{k-1} + e_{k-2}, \\ \mathcal{A}(e_k) = \lambda e_k + e_{k-1}. \end{gathered}$$

В частности, если $\lambda = 0$,

$$\begin{gathered} \mathcal{A}(e_1) = 0, \\ \mathcal{A}(e_2) = e_1, \\ \cdots \\ \mathcal{A}(e_{k-1}) = e_{k-2}, \\ \mathcal{A}(e_k) = e_{k-1}. \end{gathered}$$

Доказательство. Следует из того, что по определению матрицы линейного оператора в некотором базисе столбцами этой матрицы являются столбцы координат образов базисных векторов при действии линейного оператора. ∎

Утверждение 7.2. *Для произвольной жордановой клетки порядка k с λ на диагонали*

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

и произвольной жордановой матрицы порядка n

$$J = \begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{k_2}(\lambda_2) & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & J_{k_s}(\lambda_s) \end{pmatrix},$$

где $J_{k_i}(\lambda_i)$ – жордановы клетки, верно следующее

1) характеристический многочлен $\chi_J(t)$ жордановой матрицы J равен $(-1)^n(t-\lambda_1)^{k_1}(t-\lambda_2)^{k_2}\dots(t-\lambda_s)^{k_s}$;

2) минимальный многочлен $\mu_{J_k(\lambda)}(t)$ жордановой клетки $J_k(\lambda)$ равен

$$(t-\lambda)^k;$$

3) минимальный многочлен $\mu_J(t)$ жордановой матрицы J равен

$$HOK\ (\mu_{J_{k_1}(\lambda_1)}(t), \dots, \mu_{J_{k_s}(\lambda_s)}(t)).$$

Доказательство.

1) Сначала посчитаем характеристический многочлен от жордановой клетки. По правилу вычисления определителя от верхней треугольной матрицы имеем:

$$\chi_{J_k(\lambda)}(t) = \det\ (J_k(\lambda) - tE) = \det \begin{pmatrix} \lambda-t & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda-t & 1 & \dots & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \dots & \lambda-t & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda-t \end{pmatrix} =$$

$$= (\lambda-t)^k = (-1)^k(t-\lambda)^k.$$

В силу теоремы об определителе матрицы с углом нулей характеристический многочлен от жордановой матрицы представляется в виде произведения характеристических многочленов от жордановых клеток:

$$\chi_J(t) = \det\ (J - tE) = \det \begin{pmatrix} J_{k_1}(\lambda_1) - tE & 0 & \dots & 0 \\ 0 & J_{k_2}(\lambda_2) - tE & \dots & 0 \\ . & . & & . \\ 0 & 0 & \dots & J_{k_s}(\lambda_s) - tE \end{pmatrix} =$$

$$= \det\ (J_{k_1}(\lambda_1) - tE) det(J_{k_2}(\lambda_2) - tE) \dots \det\ (J_{k_s}(\lambda_s) - tE) =$$
$$= [(-1)^{k_1}(t-\lambda_1)^{k_1}][(-1)^{k_2}(t-\lambda_2)^{k_2}] \dots [(-1)^{k_3}(t-\lambda_s)^{k_s}] =$$

$$= (-1)^{k_1+\ldots+k_s}(t-\lambda_1)^{k_1}(t-\lambda_2)^{k_2}\ldots(t-\lambda_s)^{k_s} =$$
$$= (-1)^n(t-\lambda_1)^{k_1}(t-\lambda_2)^{k_2}\ldots(t-\lambda_s)^{k_s}.$$

2) Так как характеристический многочлен $\chi_{J_k(\lambda)}(t)$ от жордановой клетки $J_k(\lambda)$ равен $(-1)^k(t-\lambda)^k$, а минимальный многочлен делит характеристический (по теореме Гамильтона-Кэли и лемме 7.1 п. 3), получаем, что $\mu_{J_k(\lambda)}(t) = (t-\lambda)^l$ для некоторого $l \leqslant k$. Осталось найти наименьшее натуральное l, при котором $(J_k(\lambda)-\lambda E)^l = 0$.

Рассмотрим, что происходит с матрицей $J_k(\lambda)-\lambda E = J_k(0)$ при возведении ее в степени.

ПРИМЕР для $k=4$:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix};$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix};$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}^4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Нетрудно понять, что в общем случае (для произвольного k) при возведении жордановой клетки $J_k(0)$ в степени мы будем иметь аналогичную ситуацию: при каждом последующем возведении ряд единиц, параллельный главной диагонали, смещается на один ряд в сторону правого верхнего угла матрицы. При $m<k$ в жордановой клетке $J_k(0)^m$ будет m нулевых строк и $(k-m)$ ненулевых ступенек. Когда же m станет равно k, жорданова клетка станет равной нулю. Таким образом, наименьшее l, при котором $(J_k(\lambda)-\lambda E)^l = 0$, равно k. Поэтому $\mu_{J_k(\lambda)}(t) = (t-\lambda)^k$.

3) Далее рассмотрим жорданову матрицу J. Из свойств сокращенного умножения блочных матриц получаем, что

$$\mu_J(J) = \mu_J(\begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \ldots & 0 \\ 0 & J_{k_2}(\lambda_2) & \ldots & 0 \\ . & . & . & . \\ 0 & 0 & \ldots & J_{k_s}(\lambda_s) \end{pmatrix}) =$$

$$= \begin{pmatrix} \mu_J(J_{k_1}(\lambda_1)) & 0 & \dots & 0 \\ 0 & \mu_J(J_{k_2}(\lambda_2)) & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & \mu_J(J_{k_s}(\lambda_s)) \end{pmatrix}.$$

Поэтому $\mu_J(J) = 0$ тогда и только тогда, когда $\mu_J(J_{k_i}(\lambda_i)) = 0$ для каждого $i = \overline{1, s}$. Учитывая лемму 7.1 п. 3, получаем, что μ_J – это многочлен минимальной степени со старшим коэффициентом, равным 1, который делится нацело на $\mu_{J_{k_i}(\lambda_i)}$ для каждого $i = \overline{1, s}$. Это означает, что

$$\mu_J(t) = HOK\ (\mu_{J_{k_1}(\lambda_1)}(t), \dots, \mu_{J_{k_s}(\lambda_s)}(t)).$$

■

Утверждение 7.3. *Пусть V – n-мерное векторное пространство, $\mathcal{A} : V \to V$ – линейный оператор в V, для которого существует жорданов базис, а J – (жорданова) матрица в этом базисе. И пусть V_μ – собственное подпространство для линейного оператора $\mathcal{A}$, отвечающее собственному значению μ. Тогда* $\dim V_\mu$ *равна числу жордановых клеток в матрице J с μ на диагонали.*

Доказательство. Размерность собственного подпространства V_μ совпадает с размерностью пространства решений системы $(J - \mu E)X = 0$, которая равняется $n - \mathrm{rk}\ (J - \mu E)$. В обозначениях (13) имеем, $J - \mu E$ – блочно-диагональная матрица с s блоками вида

$$J_{k_i}(\lambda_i) - \mu E = \begin{pmatrix} \lambda_i - \mu & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i - \mu & 1 & \dots & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \dots & \lambda_i - \mu & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_i - \mu \end{pmatrix}.$$

Поэтому $\mathrm{rk}\ (J - \mu E) = \sum_{i=1}^{s} \mathrm{rk}\ (J_{k_i}(\lambda_i) - \mu E)$, где

$$\mathrm{rk}\ (J_{k_i}(\lambda_i) - \mu E) = \begin{cases} k_i, & \text{если } \lambda_i \neq \mu, \\ k_i - 1, & \text{если } \lambda_i = \mu. \end{cases}$$

Так как $n = \sum_{i=1}^{s} k_i$, то $n - \mathrm{rk}\ (J - \mu E) = \sum_{i=1}^{s}(k_i - \mathrm{rk}\ (J_{k_i}(\lambda_i) - \mu E))$, где

$$k_i - \mathrm{rk}\ (J_{k_i}(\lambda_i) - \mu E) = \begin{cases} 0, & \text{если } \lambda_i \neq \mu, \\ 1, & \text{если } \lambda_i = \mu. \end{cases}$$

Отсюда получаем, что $n - \mathrm{rk}\ (J - \mu E)$ равно числу жордановых клеток в матрице J с μ на диагонали. ■

8. Жорданова нормальная форма (продолжение).

8.1. Существование жорданова базиса.

Теорема 8.1. *Пусть F – алгебраически замкнутое поле, V – конечномерное векторное пространство над F, $\mathcal{A} : V \to V$ – линейный оператор в V. Тогда в V имеется жорданов базис для $\mathcal{A}$.*

Доказательство. Будем доказывать теорему индукцией по $n = \dim V$ с очевидным основанием при $n = 1$.

Прежде всего заметим, что если к жордановой матрице (13) прибавить произвольную скалярную матрицу αE, где $\alpha \in F$, то она, очевидно, останется жордановой. Поэтому если найдется жорданов базис для оператора $\mathcal{A}$, то этот же базис будет являться жордановым и для оператора $\mathcal{A} + \alpha\mathcal{I}$.

Это замечание позволяют считать в ходе доказательства шага индукции, что мы имеет дело с вырожденным линейным оператором $\mathcal{A} : V \to V$. Если это не так, можно заменить $\mathcal{A}$ на $\mathcal{A} - \lambda\mathcal{I}$, где λ – любое собственное значение линейного оператора $\mathcal{A}$, которое существует в силу алгебраической замкнутости поля F.

Рассмотрим образ $\mathrm{Im}\ \mathcal{A}$ и ядро $\mathrm{Ker}\ \mathcal{A}$ линейного оператора $\mathcal{A}$. Поскольку $\mathcal{A}$ – вырожденный, $\mathrm{Ker}\ \mathcal{A} \neq \{0\}$. Если $\mathrm{Ker}\ \mathcal{A} = V$, то наличие жорданова базиса очевидно: любой базис жорданов и жорданова форма оператора равна нулевой матрице. Далее считаем, что $\mathrm{Ker}\ \mathcal{A} \neq V$, т.е. $0 < \dim \mathrm{Ker}\ \mathcal{A} < n$. Из утверждения 4.1 получаем, что $0 < \dim \mathrm{Im}\ \mathcal{A} < n$.

Так как ядро и образ инвариантны относительно действия оператора $\mathcal{A}$, мы можем рассмотреть ограничение действия оператора $\mathcal{A}$ на подпространства $\mathrm{Im}\ \mathcal{A}$ и $\mathrm{Ker}\ \mathcal{A}$. В подпространстве $\mathrm{Ker}\ \mathcal{A}$ все векторы – собственные, отвечающие собственному значению нуль, поэтому любой базис в $\mathrm{Ker}\ \mathcal{A}$ является жордановым, и в этом базисе матрица для ограничения $\mathcal{A}\,|_{\mathrm{Ker}\ \mathcal{A}}$ оператора $\mathcal{A}$ на $\mathrm{Ker}\ \mathcal{A}$ – нулевая. По предположению индукции в подпространстве $\mathrm{Im}\ \mathcal{A}$ тоже есть жорданов базис для ограничения $\mathcal{A}\,|_{\mathrm{Im}\ \mathcal{A}}$ оператора $\mathcal{A}$ на $\mathrm{Im}\ \mathcal{A}$.

Как уже упоминалось (утверждение 7.1), жорданов базис для линейного оператора является объединением жордановых базисов для клеток. Запишем жорданов базис для $\mathcal{A}\,|_{\mathrm{Im}\ \mathcal{A}}$, начиная с s клеток с собственным значением 0:

$$\begin{aligned}\mathcal{E} = \{e_{1,1}, \dots, e_{1,k_1};&\\ &\ddots\\ &\qquad e_{s,1}, \dots, e_{s,k_s};\\ &\qquad\qquad e_{s+1,1}, \dots, e_{s+1,k_{s+1}};\\ &\qquad\qquad\qquad\ddots\\ &\qquad\qquad\qquad\qquad e_{m,1}, \dots, e_{m,k_m}\}.\end{aligned}$$

В этом базисе матрица для $\mathcal{A}\,|_{\mathrm{Im}\,\mathcal{A}}$ имеет вид:

$$J_{\mathrm{Im}\,\mathcal{A}} = \begin{pmatrix} \begin{smallmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 \end{smallmatrix} & & & & \\ & \ddots & & & \\ & & \begin{smallmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 \end{smallmatrix} & & & \\ & & & \begin{smallmatrix} \lambda_{s+1} & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & \dots & \lambda_{s+1} \end{smallmatrix} & & \\ & & & & \ddots & \\ & & & & & \begin{smallmatrix} \lambda_m & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & \dots & \lambda_m \end{smallmatrix} \end{pmatrix}.$$

Так как собственное подпространство в $\mathrm{Im}\,\mathcal{A}$, отвечающее собственному значению 0, совпадает с пересечением $\mathrm{Im}\,\mathcal{A} \cap \mathrm{Ker}\,\mathcal{A}$, из утверждения 7.3 получаем, что $\dim\,(\mathrm{Im}\,\mathcal{A} \cap \mathrm{Ker}\,\mathcal{A}) = s$. Более того, s векторов $e_{1,1}, e_{2,1}, \dots, e_{s,1}$ линейно независимы и являются собственными с собственным значением 0, поэтому они и составляют базис в $\mathrm{Im}\,\mathcal{A} \cap \mathrm{Ker}\,\mathcal{A}$. Дополним $e_{1,1}, e_{2,1}, \dots, e_{s,1}$ векторами $\mathcal{F} = \{f_1, \dots, f_q\}$ до базиса подпространства $\mathrm{Ker}\,\mathcal{A}$.

Согласно формуле Грассмана (теорема 2.1):

$$\dim\,(\mathrm{Im}\,\mathcal{A} + \mathrm{Ker}\,\mathcal{A}) = \dim\,\mathrm{Im}\,\mathcal{A} + \dim\,\mathrm{Ker}\,\mathcal{A} - \dim\,(\mathrm{Im}\,\mathcal{A} \cap \mathrm{Ker}\,\mathcal{A}).$$

Так как $\dim\,(\mathrm{Im}\,\mathcal{A} \cap \mathrm{Ker}\,\mathcal{A}) = s$, а по утверждению 4.1

$$\dim\,\mathrm{Im}\,\mathcal{A} + \dim\,\mathrm{Ker}\,\mathcal{A} = \dim V = n,$$

получаем

$$\dim\,(\mathrm{Im}\,\mathcal{A} + \mathrm{Ker}\,\mathcal{A}) = n - s.$$

Значит, к базису $\mathcal{E} \sqcup \mathcal{F}$ суммы подпространств $\mathrm{Im}\,\mathcal{A} + \mathrm{Ker}\,\mathcal{A}$ надо добавить ровно s векторов, чтобы получить базис для V. В качестве этих векторов возьмем $g_1, g_2, \dots, g_s$ такие, что

$$\begin{array}{rcl} \mathcal{A}g_1 & = & e_{1,k_1} \\ \mathcal{A}g_2 & = & e_{2,k_2}, \\ & \cdots & \\ \mathcal{A}g_s & = & e_{s,k_s}. \end{array} \tag{14}$$

Такие векторы $g_1, g_2, \dots, g_s$ существует по той причине, что векторы e_{1,k_1}, $e_{2,k_2}, \dots,$ e_{s,k_s} в правых частях выписанных равенств лежат в образе $\mathrm{Im}\,\mathcal{A}$ линейного оператора $\mathcal{A}$.

Докажем, что система векторов

$$
\begin{array}{l}
\mathcal{H} = \{f_1; \\
\qquad \ddots \\
\qquad\qquad f_q; \\
\qquad\qquad\qquad e_{1,1}, \dots, e_{1,k_1}, g_1; \\
\qquad\qquad\qquad\qquad \ddots \\
\qquad\qquad\qquad\qquad\qquad e_{s,1}, \dots, e_{s,k_s}, g_s; \\
\qquad\qquad\qquad\qquad\qquad\qquad e_{s+1,1}, \dots, e_{s+1,k_{s+1}}; \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad \ddots \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad e_{m,1}, \dots e_{m,k_m}\}
\end{array}
$$

является жордановым базисом для действия линейного оператора $\mathcal{A}$ на всем пространстве V.

Сначала, докажем, что $\mathcal{H}$ – линейно независимая система. Заметим, что, исключив $g_1, \dots, g_s$, мы остаемся с базисом $\mathcal{E} \sqcup \mathcal{F}$ пространства $\mathrm{Im}\ \mathcal{A} + \mathrm{Ker}\ \mathcal{A}$, построенным именно так, как это делалось при доказательстве формулы Грассмана (теорема 2.1). Значит, система $\mathcal{E} \sqcup \mathcal{F}$ линейно независима. Поэтому в любой нетривиальной линейной комбинации векторов из $\mathcal{H}$ векторы $g_1, g_2, \dots, g_s$ входят с коэффициентами, среди которых хотя бы один не равен нулю. Итак, допустим,

$$\sum_{j=1}^{q} \alpha_j f_j + \sum_{i \leqslant s} \beta_i g_i + \sum_{i \leqslant s} (\sum_{j=1}^{k_i} \gamma_{i,j} e_{i,j}) + \sum_{i \geqslant s+1} (\sum_{j=1}^{k_i} \gamma_{i,j} e_{i,j}) = 0,$$

причем не все коэффициенты β_i равны нулю. Применим к этой линейной комбинации линейный оператор $\mathcal{A}$:

$$\sum_{j=1}^{q} \alpha_j 0 + \sum_{i \leqslant s} \beta_i e_{i,k_i} + \sum_{i \leqslant s} (\sum_{j=1}^{k_i - 1} \gamma_{i,j+1} e_{i,j}) + \sum_{i \geqslant s+1} (\lambda_i e_{i,1} + \sum_{j=2}^{k_i} \gamma_{i,j} (\lambda_i e_{i,j} + e_{i,j-1})) = 0$$

и получим равную нулю нетривиальную линейную комбинацию векторов базиса $\mathcal{E} \sqcup \mathcal{F}$ пространства $\mathrm{Im}\ \mathcal{A} + \mathrm{Ker}\ \mathcal{A}$. Это противоречие доказывает линейную независимость системы $\mathcal{H}$.

Выше было показано, что количество векторов в системе $\mathcal{H}$ равно размерности пространства V. Значит, эта линейно независимая система векторов образует базис в V.

Осталось доказать, что этот базис является жордановым. Действительно, чтобы получить базис $\mathcal{H}$ к базису $\mathcal{E}$ пространства $\mathrm{Im}\ \mathcal{A}$ были добавлены q векторов $f_1, \dots, f_q$ ядра, т.е. таких, что

$$\mathcal{A} f_i = 0,$$

и s векторов $g_1, \dots, g_s$, удовлетворяющих условию (14), которое согласно утверждению 7.1 говорит о том, что система $\{e_{i,1}, \dots, e_{i,k_i}, g_i\}$ соответствует жордановой клетке порядка k_i+1 с 0 на диагонали. Поэтому $\mathcal{H}$ есть объединение жордановых базисов жордановых клеток: если к матрице $J_{\mathrm{Im}\ \mathcal{A}}$ добавить q клеток размера 1 с нулем на диагонали и увеличить на 1 размер всех имевшихся в $J_{\mathrm{Im}\ \mathcal{A}}$ клеток с 0 на диагонали, то получится (жорданова) матрица оператора $\mathcal{A}$ в базисе $\mathcal{H}$. ■

8.2. Единственность жордановой нормальной формы.

Для доказательства единственности жордановой нормальной формы линейного оператора нам понадобится следующая лемма.

Лемма 8.1. *Пусть*

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ . & . & . & & . & . \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

– жорданова клетка порядка k с собственным значением λ на диагонали. Тогда

$$\mathrm{rk}\ (J_k(\lambda) - \alpha E)^m = \begin{cases} k, & \text{при } \lambda \neq \alpha, \quad \forall m, \\ k-m, & \text{при } \lambda = \alpha, \quad m < k, \\ 0, & \text{при } \lambda = \alpha, \quad m \geqslant k; \end{cases}$$

Доказательство. При $\lambda \neq \alpha$ матрица $J_k(\lambda) - \alpha E$ является невырожденной, каковой она останется и при возведении в произвольную степень. Поэтому ранг матрицы $(J_k(\lambda) - \alpha E)^m$ будет равняться ее порядку.

Пусть теперь $\lambda = \alpha$. При этом $J_k(\lambda) - \lambda E = J_k(0)$. Мы видели (смотри доказательство утверждения 7.2), что происходит при возведении жордановой клетки $J_k(0)$ в степени: при каждом последующем возведении ряд единиц, параллельный главной диагонали, смещается на один ряд в сторону правого верхнего угла матрицы. При $m < k$ в жордановой клетке $J_k(0)^m$ будет m нулевых строк и $(k-m)$ ненулевых ступенек. Когда же m станет больше или равно k, жорданова клетка станет равной нулю. Отсюда без труда выводятся требуемые утверждения о ранге матрицы. ■

Обозначим через $N_k(\alpha)$ количество жордановых клеток размера k с собственным значением α на диагонали в некоторой определенной жордановой матрице J. Вопрос о единственности жордановой нормальной форме полностью решается следующей теоремой.

Теорема 8.2. *Для любой жордановой матрицы порядка n*

$$J = \begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{k_2}(\lambda_2) & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & J_{k_s}(\lambda_s) \end{pmatrix},$$

где $J_{k_i}(\lambda_i)$ – жордановы клетки, справедлива следующая формула:

$$N_k(\alpha) = \mathrm{rk}\ (J - \alpha E)^{k-1} - 2\mathrm{rk}\ (J - \alpha E)^{k} + \mathrm{rk}\ (J - \alpha E)^{k+1}. \tag{15}$$

Здесь полагаем $\mathrm{rk}\ (J - \alpha E)^0 = \mathrm{rk}\ E = n$.

Доказательство. Замена матрицы J на $J - \alpha E$ и $N_k(\alpha)$ на $N_k(0)$ позволяет ограничиться случаем $\alpha = 0$. В этом случае формула (15), которую требуется доказать, перепишется следующим образом:

$$N_k(0) = \mathrm{rk}\ J^{k-1} - 2\mathrm{rk}\ J^k + \mathrm{rk}\ J^{k-1}.$$

Для доказательства этой формулы нам понадобятся следующие равенства

$$\begin{array}{ccccccccccccccc} N_1(0) & + & N_2(0) & + & N_3(0) & + & N_4(0) & + & \dots & + & N_n(0) & = & n & - & \mathrm{rk}\ J, \\ N_1(0) & + & 2N_2(0) & + & 2N_3(0) & + & 2N_4(0) & + & \dots & + & 2N_n(0) & = & n & - & \mathrm{rk}\ J^2, \\ N_1(0) & + & 2N_2(0) & + & 3N_3(0) & + & 3N_4(0) & + & \dots & + & 3N_n(0) & = & n & - & \mathrm{rk}\ J^3, \\ N_1(0) & + & 2N_2(0) & + & 3N_3(0) & + & 4N_4(0) & + & \dots & + & 4N_n(0) & = & n & - & \mathrm{rk}\ J^4, \\ & & & & . & & . & & . & & & & & & \\ N_1(0) & + & 2N_2(0) & + & 3N_3(0) & + & 4N_4(0) & + & \dots & + & nN_n(0) & = & n & - & \mathrm{rk}\ J^n. \end{array} \tag{16}$$

Докажем их.

Матрица J – блочно-диагональная. По свойству произведения блочных матриц для возведения блочно-диагональной матрицы в степень надо возвести в данную степень каждый блок в отдельности:

$$J^m = \begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{k_2}(\lambda_2) & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & J_{k_s}(\lambda_s) \end{pmatrix}^m =$$

$$= \begin{pmatrix} J_{k_1}(\lambda_1)^m & 0 & \dots & 0 \\ 0 & J_{k_2}(\lambda_2)^m & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & J_{k_s}(\lambda_s)^m \end{pmatrix}.$$

Очевидно, что ранг блочно-диагональной матрицы J^m равняется сумме рангов ее блоков. Так как $n = \sum_{i=1}^{s} k_i$, получаем, что

$$n - \mathrm{rk}\ J^m = \sum_{i=1}^{s} k_i - \sum_{i=1}^{s} \mathrm{rk}\ J_{k_i}(\lambda_i)^m = \sum_{i=1}^{s} (k_i - \mathrm{rk}\ J_{k_i}(\lambda_i)^m).$$

По лемме 8.1

$$k_i - \mathrm{rk}\ J_{k_i}(\lambda_i)^m = \begin{cases} 0, & \text{при} \quad \lambda \neq 0, \qquad \forall m, \\ m, & \text{при} \quad \lambda = 0, \quad m < k_i, \\ k_i, & \text{при} \quad \lambda = 0, \quad m \geqslant k_i; \end{cases}$$

Поэтому в сумме $\sum_{i=1}^{s}(k_i - \mathrm{rk}\ J_{k_i}(\lambda_i)^m)$ будут нулевыми все слагаемые, за исключением тех, которые соответствуют жордановым клеткам с нулем на диагонали.

В случае, когда $m = 1$, каждая жорданова клетка с нулем на диагонали даст слагаемое, равное 1. Всего же жордановых клеток с нулем на диагонали:

$$N_1(0) + N_2(0) + N_3(0) + \ldots + N_n(0),$$

и значит,

$$n - \mathrm{rk}\ J = \sum_{i=1}^{s}(k_i - \mathrm{rk}\ J_{k_i}(\lambda_i)) = N_1(0) + N_2(0) + N_3(0) + \ldots + N_n(0),$$

и мы получаем первое равенство из (16).

При $m = 2$ каждая жорданова клетка с нулем на диагонали порядка 1 даст слагаемое, равное своему порядку, т.е. 1. Всего же таких клеток $N_1(0)$. А каждая жорданова клетка с нулем на диагонали порядка $k \geqslant 2$ даст слагаемое, равное 2, и всего таких клеток $N_2(0) + N_3(0) + \ldots + N_n(0)$. Поэтому

$$n - \mathrm{rk}\ J^2 = \sum_{i=1}^{s}(k_i - \mathrm{rk}\ J_{k_i}(\lambda_i)^2) = N_1(0) + 2(N_2(0) + N_3(0) + \ldots + N_n(0)),$$

и мы получаем второе равенство из (16).

Аналогично для $m = 3$ каждая жорданова клетка с нулем на диагонали порядка $k \geqslant 3$ даст слагаемое, равное 3, и всего таких клеток $N_3(0) + \ldots + N_n(0)$; каждая такая жорданова клетка порядка $k \leqslant 2$ даст слагаемое, равное своему порядку, всего же жордановых клеток с нулем на диагонали порядка один – $N_1(0)$, а порядка два – $N_2(0)$. Поэтому

$$n - \mathrm{rk}\ J^3 = \sum_{i=1}^{s}(k_i - \mathrm{rk}\ J_{k_i}(\lambda_i)^3) = N_1(0) + 2N_2(0) + 3(N_3(0) + \ldots + N_n(0)),$$

и мы получаем третье равенство из (16).

Аналогично доказываются остальные равенства из (16).

Теперь начиная с предпоследнего равенства в (16), вычитаем из каждого равенства предыдущее и получаем треугольную систему:

$$\begin{array}{ccccccccccccccc}
N_1(0) & + & N_2(0) & + & N_3(0) & + & N_4(0) & + & \dots & + & N_n(0) & = & n & - & \mathrm{rk}\ J, \\
 & & N_2(0) & + & N_3(0) & + & N_4(0) & + & \dots & + & N_n(0) & = & \mathrm{rk}\ J & - & \mathrm{rk}\ J^2, \\
 & & & & N_3(0) & + & N_4(0) & + & \dots & + & N_n(0) & = & \mathrm{rk}\ J^2 & - & \mathrm{rk}\ J^3, \\
 & & & & & & N_4(0) & + & \dots & + & N_n(0) & = & \mathrm{rk}\ J^3 & - & \mathrm{rk}\ J^4, \\
 & & & & . & & . & & & & . & & & & \\
 & & & & & & & & & & N_n(0) & = & \mathrm{rk}\ J^{n-1} & - & \mathrm{rk}\ J^n.
\end{array}$$

Производя вычитание равенств на этот раз из верхних нижние, начиная с первого равенства, получим требуемые формулы (последняя формула $N_n(0) = \mathrm{rk}\ J^{n-1} - 2\mathrm{rk}\ J^n + \mathrm{rk}\ J^{n+1}$ следует из последнего равенства, если к нему прибавить тождество $0 = -\mathrm{rk}\ J^n + \mathrm{rk}\ J^{n+1}$). ■

Заметим, что в доказательстве предыдущей теоремы для получения равенств (16) можно было не применять лемму 8.1, а воспользоваться лишь информацией из доказательства этой леммы о том, что происходит с жордановой клеткой $J_k(0)$ на диагонали при возведении ее в степень: при каждом последующем возведении клетки $J_k(0)$ в степень ее ранг уменьшается на 1, пока не возведем в степень k. Тогда клетка зануляется, и дальше уменьшения ранга, связанного с данной клеткой нет. У клеток с ненулевым числом на диагонали ранг остается полным на протяжении всего процесса возведения в степень.

Из теоремы 8.2, поскольку ранги матриц линейного оператора не зависят от выбора базиса, получаем

Следствие 8.1. *Пусть V – n-мерное векторное пространство над полем F, $\mathcal{A} : V \to V$ – линейный оператор, который в некотором базисе имеет жорданову матрицу J, $N_k(\lambda)$ – количество жордановых клеток порядка k с собственным значением λ на диагонали в матрице J, A – матрица линейного оператора $\mathcal{A}$ в произвольном базисе. Тогда справедлива следующая формула:*

$$N_k(\lambda) = \mathrm{rk}\ (A - \lambda E)^{k-1} - 2\mathrm{rk}\ (A - \lambda E)^k + \mathrm{rk}\ (A - \lambda E)^{k+1}.$$

Здесь полагаем $\mathrm{rk}\ (A - \lambda E)^0 = \mathrm{rk}\ E = n$.

Следствие 8.2. *Всякий линейный оператор в конечномерном векторном пространстве над алгебраически замкнутым полем в некотором базисе записывается жордановой матрицей, причем эта матрица определена однозначно с точностью до перестановки жордановых клеток на диагонали.*

Напомним, что первое утверждение этого следствия есть теорема 8.1.

Отметим, что вместо алгебраической замкнутости поля, можно лишь потребовать, чтобы полю принадлежали все корни характеристического многочлена линейного оператора в конечномерном векторном пространстве. Тогда для линейного оператора найдется жорданов базис и его жорданова нормальная форма будет определена единственным образом с точностью до перестановки жордановых клеток на диагонали.

9. Жорданова нормальная форма (окончание).

9.1. Критерий диагонализируемости в терминах минимального многочлена.

Из теорем о существовании и единственности жордановой нормальной формы получим критерий диагонализируемости линейного оператора в терминах минимального многочлена:

Теорема 9.1. *Линейный оператор $\mathcal{A} : V \to V$ конечномерного векторного пространства V над алгебраически замкнутым полем F дигонализируем тогда и только тогда, когда его минимальный многочлен не имеет кратных корней.*

Доказательство. Пусть сначала $\mathcal{E} = \{e_1, \dots, e_n\}$ – базис из собственных векторов для $\mathcal{A}$, причем $\mathcal{A}(e_i) = \lambda_i e_i$, т.е. матрица оператора $\mathcal{A}$ в базисе $\mathcal{E}$ диагональна:

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

Так как минимальный многочлен линейного оператора равен минимальному многочлену его матрицы в произвольном базисе (лемма 7.1 п. 1), $\mu_{\mathcal{A}}(t) = \mu_A(t)$. В свою очередь, минимальный многочлен для матрицы A будет иметь вид

$$\mu_A(t) = (t - \alpha_1) \dots (t - \alpha_p),$$

где $\alpha_1, \dots \alpha_p$ – все попарно различные числа $\lambda_1, \dots, \lambda_n$. Это следует из того,что многочлен $(t - \alpha_1) \dots (t - \alpha_p)$ является аннулирующим для матрицы A, а минимальный многочлен делит любой аннулирующий многочлен и наборы корней у минимального и характеристического многочлена совпадают (лемма 7.1 п. 2, 4).

Пусть теперь минимальный многочлен не имеет кратных корней. И пусть J – жорданова нормальная форма оператора $\mathcal{A}$, которая существует в силу алгебраической замкнутости поля (по теореме 8.1). По лемме 7.1 и утверждению 7.2 п.2 имеем

$$\mu_{\mathcal{A}}(t) = \mu_J(t) = HOK\ (\mu_{J_{k_1}(\lambda_1)}(t), \dots, \mu_{J_{k_s}(\lambda_s)}(t)) =$$
$$= HOK\ ((t - \lambda_1)^{k_1}, \dots, (t - \lambda_s)^{k_s}) = (t - \alpha_1)^{n_1} \dots (t - \alpha_p)^{n_p},$$

где $\alpha_1, \dots, \alpha_p$ – все попарно различные собственные значения линейного оператора $\mathcal{A}$, n_i – максимальный порядок жордановой клетки с α_i на диагонали в жордановой нормальной форме J. Так как минимальный многочлен не имеет кратных корней, получаем, что $n_i = 1$ для всех i, т.е в J все клетки имеют порядок 1. Это и означает, что жорданова нормальная форма J диагональна. ■

9.2. Корневые подпространства.

При изучении недиагонализируемых линейных операторов в дополнение к собственным векторам принято рассматривать и более общие корневые вектора.

Определение 9.1. *Пусть дан линейный оператор $\mathcal{A} : V \to V$ векторного пространства V над полем F. Ненулевой вектор $v \in V$ называется* **корневым***, отвечающим корню λ, если для некоторого натурального m выполняется*

$$(\mathcal{A} - \lambda\mathcal{I})^m v = 0.$$

Наименьшее из таких m называется **высотой корневого вектора***.*

Отметим, что корневой вектор высоты 1 – это собственный вектор.

Если $e \in V$ – корневой вектор высоты $m > 0$, то

$$f = (\mathcal{A} - \lambda\mathcal{I})^{m-1} e$$

собственный вектор с собственным значением λ. Значит, λ – корень характеристического уравнения оператора $\mathcal{A}$.

Легко видеть, что все корневые векторы, отвечающие корню λ, с добавлением нулевого вектора, образуют подпространство в V. Это подпространство называется **корневым подпространством** и обозначается через V^λ.

Итак, V^λ содержит в себе подпространство V_λ собственных векторов, отвечающих характеристическому числу λ.

Кроме того, V^λ инвариантно относительно оператора $\mathcal{A}$. Действительно, если $x \in V^\lambda$, т.е. $(\mathcal{A} - \lambda\mathcal{I})^m x = 0$, тогда

$$(\mathcal{A} - \lambda\mathcal{I})^m (\mathcal{A}x) = \mathcal{A}(\mathcal{A} - \lambda\mathcal{I})^m x = \mathcal{A}0 = 0,$$

значит, $\mathcal{A}x \in V^\lambda$.

Заметим также, что из утверждения 7.1 видно, что если $\{e_1, e_2, \dots, e_{k-1}, e_k\}$ – жорданов базис для некоторой жордановой клетки $J_k(\lambda)$ из жордановой нормальной формы оператора $\mathcal{A}$, то

$$(\mathcal{A} - \lambda\mathcal{I})e_1 = 0,$$
$$(\mathcal{A} - \lambda\mathcal{I})e_2 = e_1,$$
$$\dots$$
$$(\mathcal{A} - \lambda\mathcal{I})e_{k-1} = e_{k-2},$$
$$(\mathcal{A} - \lambda\mathcal{I})e_k = e_{k-1}.$$

Отсюда следует, что e_i – корневой вектор высоты i, отвечающий корню λ.

Теорема 9.2. *Пусть дан линейный оператор $\mathcal{A} : V \to V$ векторного пространства V над алгебраически замкнутым полем F. Тогда V распадается в прямую сумму корневых подпространств, отвечающим всем (различным) собственным значениям оператора $\mathcal{A}$.*

Доказательство. Пусть $\mathcal{E}$ – некоторый жорданов базис для линейного оператора $\mathcal{A}$, который существует в силу теоремы 8.1. Если $\lambda_1, \ldots, \lambda_p$ – все различные собственные значения, то переупорядочим векторы базиса $\mathcal{E} = \mathcal{E}_1 \sqcup \ldots \sqcup \mathcal{E}_p$ так, чтобы базисные векторы $\mathcal{E}_i$ отвечали всем клеткам с λ_i на диагонали. Далее, пусть U_i – подпространство, порожденное векторами из $\mathcal{E}_i$. Понятно, что $V = U_1 \oplus \ldots \oplus U_p$. Из утверждения 7.1 получаем, что U_i – инвариантное подпространство относительно действия $\mathcal{A}$. Нужно показать, что каждое U_i и есть корневое подпространство V^{λ_i}, отвечающее λ_i.

Так как все векторы из жорданова базиса, отвечающие собственному значению λ_i, являются корневыми, отвечающими корню λ_i, то и любая их линейная комбинация будет корневым вектором, отвечающими корню λ_i. Значит, $U_i \subset V^{\lambda_i}$.

Докажем обратное включение. Пусть x – произвольный вектор из V^{λ_i}. Так как $V = U_1 \oplus \ldots \oplus U_p$, то

$$x = x_i + \tilde{x},$$

где

$$x_i \in U_i \subset V^{\lambda_i}, \;\; \tilde{x} \in U_1 \oplus \ldots \oplus U_{i-1} \oplus U_{i+1} \oplus \ldots \oplus U_p.$$

Тогда $\tilde{x} = x - x_i \in V^{\lambda_i}$, т.е. для некоторого натурального m выполняется $(\mathcal{A} - \lambda_i \mathcal{I})^m \tilde{x} = 0$.

Так как каждое подпространство U_i инвариантно относительно $\mathcal{A}$, то и

$$U_1 \oplus \ldots \oplus U_{i-1} \oplus U_{i+1} \oplus \ldots \oplus U_p$$

инвариантно относительно действия оператора $\mathcal{A}$, а значит, и относительно действия оператора $\mathcal{A} - \lambda_i \mathcal{I}$. Так как ограничение оператора $\mathcal{A} - \lambda_i \mathcal{I}$ на подпространство $U_1 \oplus \ldots \oplus U_{i-1} \oplus U_{i+1} \oplus \ldots \oplus U_p$, очевидно, имеет невырожденную матрицу в базисе $\mathcal{E}_1 \sqcup \ldots \sqcup \mathcal{E}_{i-1} \sqcup \mathcal{E}_{i+1} \sqcup \ldots \sqcup \mathcal{E}_p$, значит, на $U_1 \oplus \ldots \oplus U_{i-1} \oplus U_{i+1} \oplus \ldots \oplus U_p$ оператор $\mathcal{A} - \lambda_i \mathcal{I}$ действует невырожденно. Отсюда получаем, что равенство $(\mathcal{A} - \lambda_i \mathcal{I})^m \tilde{x} = 0$ может выполняться только для нулевого вектора $\tilde{x}$, т.е. $x = x_i \in U_i$, что доказывает включение $U_i \supset V^{\lambda_i}$.
∎

В частности, мы доказали:

Следствие 9.1. $\dim V^{\lambda} = m$, *где m – кратность корня λ характеристического многочлена линейного оператора $\mathcal{A} : V \to V$ векторного пространства V над алгебраически замкнутым полем F.*

9.3. Вычисление многочлена от матрицы.

Хотя жорданова нормальная форма выглядит сложнее, чем, например, диагональная матрица, однако и с ней можно достаточно просто производить алгебраические операции, в частности, используя правило сокращенного умножения блочных матриц. Так, если $f(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_m t^m$ – произвольный многочлен с коэффициентами из поля F характеристики ноль, а

$$J = \begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \ldots & 0 \\ 0 & J_{k_2}(\lambda_2) & \ldots & 0 \\ . & . & . & . \\ 0 & 0 & \ldots & J_{k_s}(\lambda_s) \end{pmatrix}$$

– жорданова матрица, то

$$f(J) = \begin{pmatrix} f(J_{k_1}(\lambda_1)) & 0 & \ldots & 0 \\ 0 & f(J_{k_2}(\lambda_2)) & \ldots & 0 \\ . & . & . & . \\ 0 & 0 & \ldots & f(J_{k_s}(\lambda_s)) \end{pmatrix}.$$

Покажем теперь, как вычислить многочлен от одной жордановой клетки.

Утверждение 9.1. *Пусть*

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & \ldots & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{pmatrix}$$

– жорданова клетка порядка k с собственным значением λ на диагонали. Тогда

$$f(J_k(\lambda)) = \begin{pmatrix} f(\lambda) & \frac{f'(\lambda)}{1!} & \frac{f''(\lambda)}{2!} & \ldots & \frac{f^{(k-2)}(\lambda)}{(k-2)!} & \frac{f^{(k-1)}(\lambda)}{(k-1)!} \\ 0 & f(\lambda) & \frac{f'(\lambda)}{1!} & \ldots & \frac{f^{(k-3)}(\lambda)}{(k-3)!} & \frac{f^{(k-2)}(\lambda)}{(k-2)!} \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \ldots & f(\lambda) & \frac{f'(\lambda)}{1!} \\ 0 & 0 & 0 & \ldots & 0 & f(\lambda) \end{pmatrix}.$$

(Таким образом, на диагонали стоят числа $f(\lambda)$, затем правее и выше диагонали параллельно диагонали идет ряд с числами $\frac{f'(\lambda)}{1!}$, еще выше идет ряд с числами $\frac{f''(\lambda)}{2!}$, далее – ряд с числами $\frac{f^{(3)}(\lambda)}{3!}$ и т.д.)

Доказательство. Представим многочлен $f(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_m t^m$ по формуле Тейлора в точке λ:

$$f(t) = f(\lambda) + \frac{f'(\lambda)}{1!}(t-\lambda) + \frac{f''(\lambda)}{2!}(t-\lambda)^2 + \ldots + \frac{f^{(m)}(\lambda)}{m!}(t-\lambda)^m.$$

Тогда

$$f(J_k(\lambda)) = f(\lambda)E + \frac{f'(\lambda)}{1!}(J_k(\lambda) - \lambda E) +$$

$$+\frac{f''(\lambda)}{2!}(J_k(\lambda) - \lambda E)^2 + \ldots + \frac{f^{(m)}(\lambda)}{m!}(J_k(\lambda) - \lambda E)^m.$$

Но $J_k(\lambda) - \lambda E = J_k(0)$, следовательно,

$$f(J_k(\lambda)) = f(\lambda)E + \frac{f'(\lambda)}{1!}J_k(0) + \frac{f''(\lambda)}{2!}J_k(0)^2 + \ldots + \frac{f^{(m)}(\lambda)}{m!}J_k(0)^m. \quad (17)$$

Как нам уже известно, i-я степень жордановой клетки $J_k(0)$ есть матрица, диагональ единиц в которой сдвинута на $i-1$ позицию вправо и вверх по сравнению с ее положением в $J_k(0)$. При вычислении $f(J_k(\lambda))$ по формуле (17) мы получим, что различные слагаемые суммы имеют ненулевые элементы на разных "диагоналях", причем эти элементы вычисляются по требуемой формуле $\frac{f^{(i)}(\lambda)}{i!}$. ■

Понятие значения многочлена от линейного оператора (матрицы) можно обобщить на случай функций, задаваемых степенными рядами $f(t) = \sum_{m=0}^{\infty} \alpha_m t^m$. Примерами степенных рядов являются ряды Тейлора. Например,

$$e^t = \sum_{m=0}^{\infty} \frac{t^m}{m!} \quad (\forall t \in \mathbb{R}),$$

$$\ln(1+t) = \sum_{m=1}^{\infty} (-1)^{m-1} \frac{t^m}{m} \quad (\forall t \in (-1, 1]),$$

$$\cos t = \sum_{m=0}^{\infty} (-1)^m \frac{t^m}{(2m)!} \quad (\forall t \in \mathbb{R}).$$

ПРИМЕР 1. Если $f(t) = e^t$, то для любого $\lambda \in \mathbb{R}$ имеем

$$e^{J_k(\lambda)} = \begin{pmatrix} e^\lambda & \frac{e^\lambda}{1!} & \frac{e^\lambda}{2!} & \cdots & \frac{e^\lambda}{(k-2)!} & \frac{e^\lambda}{(k-1)!} \\ 0 & e^\lambda & \frac{e^\lambda}{1!} & \cdots & \frac{e^\lambda}{(k-3)!} & \frac{e^\lambda}{(k-2)!} \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \ldots & e^\lambda & \frac{e^\lambda}{1!} \\ 0 & 0 & 0 & \ldots & 0 & e^\lambda \end{pmatrix}.$$

ПРИМЕР 2. Пусть требуется найти значение функции $f(t) = e^t$ от матрицы

$$A = \begin{pmatrix} 0 & -\varphi \\ \varphi & 0 \end{pmatrix}.$$

Находим жорданову нормальную форму матрицы A

$$J = \begin{pmatrix} i\varphi & 0 \\ 0 & -i\varphi \end{pmatrix}$$

(на диагонали стоят собственные значения матрицы A) и матрицу перехода в соответствующий жорданов базис

$$T = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$$

(столбцы – координаты собственных векторов для A, отвечающие $i\varphi$ и $-i\varphi$). Тогда $A = TJT^{-1}$ и, следовательно,

$$e^A = Te^J T^{-1} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

10. Билинейные функции.

10.1. Билинейные функции и их матрицы.

Пусть V – векторное пространство опять над произвольным полем F.

Определение 10.1. *Функция от двух аргументов $\beta : V \times V \to F$ называется билинейной функцией (или билинейной формой), если при любом фиксированном аргументе она является линейной по другому аргументу, т.е. для любых $x, y, z \in V, \lambda \in F$ она удовлетворяет равенствам:*

$$\beta(x + y, z) = \beta(x, z) + \beta(y, z);$$

$$\beta(\lambda x, y) = \lambda\beta(x, y);$$

$$\beta(x, y + z) = \beta(x, y) + \beta(x, z);$$

$$\beta(x, \lambda y) = \lambda\beta(x, y).$$

ПРИМЕР 1. Скалярное произведение $\beta(x, y) = |x| \cdot |y| \cdot \cos\alpha$ в пространстве геометрических векторов евклидовой плоскости V^2 (трехмерного евклидова пространства V^3) является билинейной функцией над $\mathbb{R}$.

ПРИМЕР 2. Если $\varphi, \psi \in V^*$ – две линейные функции, то $\beta(x, y) = \varphi(x)\psi(y)$ является билинейной функцией.

ПРИМЕР 3. Рассмотрим пространство $F[x]$ многочленов и фиксируем число $x_0 \in F$. Положим $\beta(f, g) = f'(x_0)g(x_0)$. Тогда β является билинейной функцией на $F[x]$.

ПРИМЕР 4. В пространстве $M_n(F)$ квадратных матриц с элементами из поля F функция $\beta(A, B) = \mathrm{tr}\ AB$ является билинейной функцией на $M_n(F)$.

ПРИМЕР 5. Рассмотрим пространство F^2 строк длины два. Тогда

$$\beta((a_1, a_2), (b_1, b_2)) = \det \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}$$

является билинейной функцией на F^2.

ПРИМЕР 6. В пространстве, векторами которого являются непрерывные функции, функция

$$\beta(f, g) = \int_a^b \int_a^b K(s, t) f(s) g(t) ds dt$$

является билинейной функцией векторов f и g ($K(s, t)$ – некоторая непрерывная функция переменных s и t). Если $K(s, t) \equiv 1$, то $\beta(f, g)$ есть произведение линейных функций $\int_a^b f(s)ds$ и $\int_a^b g(t)dt$.

Для практической работы с билинейными функциями важно перейти на матричный язык.

Пусть $\mathcal{E} = \{e_1, \ldots, e_n\}$ – произвольный базис n-мерного векторного пространства V. Выразим билинейную функцию $\beta(x, y)$ через координаты $x_1, \ldots, x_n$ и $y_1, \ldots, y_n$ векторов x и y в этом базисе. Имеем:

$$\beta(x, y) = \beta(x_1 e_1 + \ldots + x_n e_n, y_1 e_1 + \ldots + y_n e_n) = \sum_{i,j=1}^{n} x_i y_j \beta(e_i, e_j).$$

Обозначим числа $\beta(e_i, e_j)$ через $b_{i,j}$. Тогда

$$\beta(x, y) = \sum_{i,j=1}^{n} b_{i,j} x_i y_j. \qquad (18)$$

Определение 10.2. *Квадратная матрица $B = (b_{i,j})$ порядка n называется* **матрицей билинейной функции** β **в базисе** $\mathcal{E}$. *Ранг матрицы B называется* **рангом билинейной функции** β.

Как легко проверить, перемножая матрицы, равенство (18) можно записать в матричном виде:

$$\beta(x, y) = X^t B Y, \qquad (19)$$

где $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$.

При замене базиса матрица билинейной функции, разумеется, меняется. Получим закон ее изменения.

Утверждение 10.1. *Пусть в n-мерном пространстве V даны два базиса $\mathcal{E} = \{e_1, \ldots, e_n\}$ и $\widetilde{\mathcal{E}} = \{\tilde{e}_1, \ldots, \tilde{e}_n\}$. Пусть β – билинейная функция на векторном пространстве V, B и $\widetilde{B}$ – ее матрицы в базисах $\mathcal{E}$ и $\widetilde{\mathcal{E}}$, соответственно. Тогда*

$$\widetilde{B} = T^t B T, \qquad (20)$$

где $T = T_{\mathcal{E} \to \widetilde{\mathcal{E}}}$ – матрица перехода от базиса $\mathcal{E}$ к базису $\widetilde{\mathcal{E}}$.

Доказательство. Пусть даны два произвольных вектора x и y. И пусть X, Y – их столбцы координат в базисе $\mathcal{E}$, $\widetilde{X}$, $\widetilde{Y}$ – столбцы координат в базисе $\widetilde{\mathcal{E}}$. Тогда по формуле (1)

$$X = T\widetilde{X} \text{ и } Y = T\widetilde{Y}.$$

Имеем,

$$\widetilde{X}^t \widetilde{B} \widetilde{Y} = \beta(x, y) = X^t B Y = (T\widetilde{X})^t B (T\widetilde{Y}) = \widetilde{X}^t T^t B T \widetilde{Y}.$$

Взяв $x = \tilde{e}_i$, $y = \tilde{e}_j$, мы получим равенство (i, j)-х элементов матрицы $\widetilde{B}$ и матрицы $T^t B T$ для произвольных i, j. ■

Следствие 10.1. *Ранг билинейной функции не зависит от выбора базиса.*

Доказательство. Так как по утверждению 10.1 матрицы билинейной функции в разных базисах связаны формулой $\widetilde{B} = T^t BT$, где T – невырожденная матрица, то для доказательства достаточно воспользоваться леммой 4.1. ∎

10.2. Симметрические билинейные функции.

Определение 10.3. *Билинейная функция $\beta : V \times V \to F$ называется* **симметрической***, если для любых векторов $x, y \in V$ выполняется*

$$\beta(x, y) = \beta(y, x).$$

Если билинейная функция в конечномерном векторном пространстве V является симметрической, а $\mathcal{E} = \{e_1, \ldots, e_n\}$ – произвольный базис пространства V, то $\beta(e_i, e_j) = \beta(e_j, e_i)$ для любых i, j, т.е. $b_{i,j} = b_{j,i}$. Таким образом матрица B билинейной функции β в базисе $\mathcal{E}$ является симметрической.

Обратно, пусть билинейная функция β имеет симметрическую матрицу B в некотором базисе $\mathcal{E}$. Тогда, поскольку матрица размера 1×1 не меняется при транспонировании, по формуле (19) получаем

$$\beta(x, y) = X^t BY = (X^t BY)^t = Y^t B^t (X^t)^t = Y^t BX = \beta(y, x).$$

Мы доказали

Утверждение 10.2. *Билинейная функция в конечномерном векторном пространстве V является симметрической тогда и только тогда, когда ее матрица в каком-нибудь базисе является симметрической.*

Скалярное произведение в V^2 и V^3 является примером симметрической билинейной функции.

ЗАДАЧА. Проверьте, что в пространстве $M_n(F)$ квадратных матриц билинейная функция $\beta(A, B) = \operatorname{tr} AB$ является симметрической.

Нашей ближайшей задачей будет определить, к какому наиболее простому виду можно привести матрицу билинейной функции за счет выбора подходящего базиса. Геометрические ассоциации, связанные со скалярным произведением векторов, могут быть полезны при изучении произвольных билинейных функций. Этим объясняется вводимая ниже терминология.

Определение 10.4. *Пусть $\beta : V \times V \to F$ – симметрическая билинейная функция. Для произвольного подмножества $S \subset V$ множество*

$$S^{\perp} = \{v \in V | \beta(v, s) = 0 \ \ \forall s \in S\}$$

называется **ортогональным дополнением** *к S (относительно β).*

В силу линейности β по первому аргументу получаем, что $S^{\perp}$ является подпространством в пространстве V.

Основное свойство симметрических билинейных функций – следующее.

Теорема 10.1. *Для произвольной симметрической билинейной функции*

$$\beta : V \times V \to F$$

на конечномерном векторном пространстве V над полем F, $\mathrm{char}\, F \neq 2$, существует базис $\mathcal{E} = \{e_1, \ldots, e_n\}$ пространства V, в котором матрица функции β диагональна.

Доказательство. Докажем теорему индукцией по $n = \dim V$ с очевидным основанием при $n = 1$.

Если $\beta \equiv 0$, то матрица билинейной функции в любом базисе – нулевая, а значит, диагональная. Далее предполагаем, что β – ненулевая функция. Тогда в V обязательно есть вектор e_1, для которого $\beta(e_1, e_1) \neq 0$. Действительно, если это не так, то $\beta(x + y, x + y) = 0$ для всех $x, y \in V$. Но тогда

$$0 = \beta(x + y, x + y) = \beta(x, x) + \beta(x, y) + \beta(y, x) + \beta(y, y) = 2\beta(x, y).$$

Так как $\mathrm{char}\, F \neq 2$, получаем, что $\beta(x, y) = 0$ для всех $x, y \in V$, т.е. $\beta \equiv 0$, что противоречит предположению.

Теперь рассмотрим ортогональное дополнение $U = \{e_1\}^\perp$. Это – собственное подпространство в V размерности $n-1$. Действительно, рассмотрим произвольный базис $\mathcal{F} = \{f_1, \ldots, f_n\}$ пространства V. Пусть $\lambda_1, \ldots, \lambda_n$ – координаты вектора e_1 в базисе $\mathcal{F}$, $B = (b_{i,j})$ – матрица билинейной функции β в базисе $\mathcal{F}$. Вектор x принадлежит U тогда и только тогда, когда $\beta(x, e_1) = 0$. Таким образом, согласно формуле (18) для определения $x \in U$ нужно решить однородную систему из одного уравнения $\sum_{i,j=1}^n b_{i,j}\lambda_i x_j = 0$ с n неизвестными $x_1, \ldots, x_n$, координатами вектора x в базисе $\mathcal{F}$. Размерность пространства решений такой системы либо n, либо $n - 1$. Поскольку $\beta(e_1, e_1) \neq 0$, вектор e_1 не принадлежит U. Значит, $\dim U = n - 1$.

Ограничение функции β на U – это по-прежнему симметрическая билинейная функция. По предположению индукции, в U найдется базис $\{e_2, \ldots, e_n\}$, в котором матрица для этого ограничения имеет диагональный вид, т.е.

$$\beta(e_i, e_j) = 0 \text{ при } i \neq j, \ 2 \leqslant i, j \leqslant n.$$

Так как $U = \{e_1\}^\perp$, то и $\beta(e_1, e_j) = 0$ при $j \geqslant 2$. Таким образом, $\mathcal{E} = \{e_1, \ldots, e_n\}$ – искомый базис. ∎

Определение 10.5. *Говорят, что симметрическая билинейная функция $\beta : V \times V \to F$ имеет в базисе $\mathcal{E} = \{e_1, \ldots, e_n\}$ пространства V* **канонический** *(или* **диагональный***)* **вид***, если для всех векторов $x = x_1 e_1 + \ldots + x_n e_n, y = y_1 e_1 + \ldots + y_n e_n \in V$ значение $\beta(x, y)$ вычисляется по формуле $\beta(x, y) = \sum_{i=1}^n \lambda_i x_i y_i$. При этом базис $\mathcal{E}$ называется* **каноническим базисом** *для β.*

ЗАДАЧА. Можно ли билинейную функцию $\beta((x_1, x_2), (y_1, y_2)) = x_1y_2 + x_2y_1$ на пространстве $\mathbb{Z}_2^2$ привести к каноническому виду?

ЗАДАЧА. Пусть V – конечномерное векторное пространство над полем F, $\mathrm{char}\ F \neq 2$. Назовем билинейную функцию $\beta : V \times V \to F$ *кососимметрической*, если

$$\beta(x, y) = -\beta(y, x)$$

для любых векторов $x, y \in V$. Доказать, что

1) пространство всех билинейных функций является прямой суммой подпространств симметрических и кососимметрических билинейных функций.

2) билинейна функция β является кососимметрической тогда и только тогда, когда ее матрица B является кососимметрической, т.е. $B^t = -B$;

3) для любой кососимметрической билинейной функции $\beta : V \times V \to F$ существует базис, в котором ее матрица имеет блочно-диагональный вид

$$B = \begin{pmatrix} B_1 & & & & & \\ & \ddots & & & & \\ & & B_m & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

где $B_i = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, i = \overline{1, m}$;

4) ранг кососимметрической билинейной функции является четным числом.

11. Квадратичные функции.

11.1. Определение квадратичной функции.

Определение 11.1. *Пусть $\beta : V \times V \to F$ – симметрическая билинейная функция. Функция $q : V \to F$, которая для любого $x \in V$ определяется равенством*

$$q(x) = \beta(x, x),$$

называется **квадратичной функцией** *(или* **квадратичной формой***), ассоциированной с функцией β.*

Согласно (19) в матричном виде квадратичная функция на конечномерном векторном пространстве V записывается так:

$$q(x) = X^t B X. \tag{21}$$

А согласно (18) в координатах квадратичная функция записывается так:

$$q(x) = \sum_{i,j=1}^{n} b_{i,j} x_i x_j. \tag{22}$$

Правая часть равенства (22) – однородный многочлен второй степени относительно $x_1, \ldots, x_n$. (Собственно, слово "форма", когда-то употреблявшееся значительно шире, означает "однородный многочлен".)

Определение 11.2. *Матрица билинейной функции β в заданном базисе называется также* **матрицей квадратичной функции** *q в этом базисе. Ранг матрицы квадратичной функции в заданном базисе называется* **рангом квадратичной функции** *q.*

Замечание 11.1. *В случае поля F характеристики, отличной от 2, квадратичная функция $q(x)$ однозначно определяет симметрическую билинейную функцию $\beta(x, y)$, из которой она получена:*

$$\beta(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y)).$$

Процедура восстановления $\beta(x, y)$ по $q(x)$ называется **процедурой поляризации**.

Таким образом, имеется взаимно однозначное соответствие между симметрическими билинейными и квадратичными функциями, если $\operatorname{char} F \neq 2$. Имея в виду это соответствие, в дальнейшем изложении мы будем из соображений удобства иногда говорить о симметрических билиненых, иногда – о квадратичных функциях. Далее в этой лекции предполагаем, что $\operatorname{char} F \neq 2$ и $\dim V = n$.

Из теоремы 10.1 получаем

Следствие 11.1. *Для любой квадратичной функции* $q : V \to F$ *существует базис, в котором квадратичная функция* q *записывается в виде суммы квадратов:*

$$q(x) = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2. \tag{23}$$

Выражение (23) называется **каноническим** (или **диагональным**) **видом** квадратичной функции q.

11.2. Метод Лагранжа.

На самом деле, для практического приведения квадратичной функции к сумме квадратов используется **алгоритм Лагранжа** выделения полного квадрата.

Пусть дана координатная запись $q(x) = \sum_{i,j=1}^{n} b_{i,j} x_i x_j$ квадратичной функции в каком-нибудь базисе $\mathcal{E}$. После приведения подобных членов запись принимает вид

$$q(x) = \sum_{i=1}^{n} b_{i,i} x_i^2 + 2 \sum_{1 \leqslant i < j \leqslant n} b_{i,j} x_i x_j. \tag{24}$$

Если хотя бы один из коэффициентов при квадратах (не ограничивая общности, ниже считаем, что это $b_{1,1}$) отличен от нуля, то применяем следующую процедуру. Собираем вместе все члены содержащие x_1 и выносим общий множитель $b_{1,1}$ за скобки:

$$b_{1,1}x_1^2+2b_{1,2}x_1x_2+\ldots+2b_{1,n}x_1x_n+\ldots = b_{1,1}(x_1^2+2\frac{b_{1,2}}{b_{1,1}}x_1x_2+\ldots+2\frac{b_{1,n}}{b_{1,1}}x_1x_n)+\ldots.$$

Дополняем выражение в квадратных скобках до квадрата суммы:

$$b_{1,1}(x_1 + \frac{b_{1,2}}{b_{1,1}}x_2 + \ldots + \frac{b_{1,n}}{b_{1,1}}x_n)^2 + \ldots$$

И производя линейную замену переменных

$$\begin{cases} y_1 = x_1 + \frac{b_{1,2}}{b_{1,1}}x_2 + \ldots + \frac{b_{1,n}}{b_{1,1}}x_n, \\ y_2 = x_2, \\ \quad . \quad . \quad . \\ y_n = x_n, \end{cases}$$

получаем

$$b_{1,1}y_1^2 + q(y_2, \ldots, y_n),$$

где квадратичная форма $q(y_2, \ldots, y_n)$ не содержит переменную y_1. Далее применяется индукция по n к $q(y_2, \ldots, y_n)$. Указанная процедура работает, лишь если один из коэффициентов при квадратах (здесь $b_{1,1}$) отличен от нуля. Если ненулевых коэффициентов при квадратах нет, но есть, допустим, $2b_{1,2}x_1x_2$, $b_{1,2} \neq 0$, то полагаем

$$\begin{cases} x_1 = y_1 - y_2, \\ x_2 = y_1 + y_2, \\ x_3 = y_3, \\ \quad . \quad . \quad . \\ x_n = y_n, \end{cases}$$

в результате чего получаем коэффициент при y_1^2 равный $2b_{1,2} \neq 0$. После этого применяется процедура, описанная ранее. (Выпишите преобразования базиса, отвечающие каждому из произведенных преобразований координат, и убедитесь, что эти преобразования действительно переводят базис снова в базис!)

Исходя из замечания 11.1, мы видим, что алгоритм Лагранжа дает другое доказательство теоремы о приведении матрицы симметрической билинейной функции к диагональному виду.

11.2. Приведение квадратичной формы к сумме квадратов унитреугольным преобразованием.

Угловым минором порядка m квадратной матрицы B будем называть определитель матрицы, состоящей из элементов, стоящих на пересечении первых m строк и первых m столбцов матрицы B.

Теорема 11.1. *Пусть $\beta : V \times V \to F$ – симметрическая билинейная функция на векторном пространстве V, B – ее матрица в некотором базисе $\mathcal{E} = \{e_1, \ldots, e_n\}$, а $\Delta_1, \ldots, \Delta_n$ – угловые миноры матрицы B, $\Delta_0 = 1$.*

Если все угловые миноры $\Delta_1, \ldots, \Delta_{n-1}$ отличны от нуля, то существует базис $\mathcal{F} = \{f_1, \ldots, f_n\}$ в векторном пространстве V, в котором квадратичная функция q, ассоциированная с β, имеет вид

$$q(y_1f_1 + \ldots + y_nf_n) = \frac{\Delta_1}{\Delta_0}y_1^2 + \ldots + \frac{\Delta_n}{\Delta_{n-1}}y_n^2, \tag{25}$$

причем матрица перехода от $\mathcal{E}$ к $\mathcal{F}$ является (верхней) унитреугольной, т.е. треугольной с единицами на главной диагонали.

Формула (25) называется **формулой Якоби**.

Доказательство. Докажем теорему индукцией по n.

При $n = 1$ имеем $f_1 = e_1$ и $q(f_1) = \beta(e_1, e_1) = \Delta_1 = \frac{\Delta_1}{\Delta_0}$.

При $n > 1$ рассмотрим подпространство $U = \langle e_1, \ldots, e_{n-1}\rangle$. Ограничение $\beta\mid_U$ симметрической билинейной функции β на подпространство U является симметрической билинейной функцией. Матрица B' функции $\beta\mid_U$ в базисе

$\mathcal{E}' = \{e_1, \ldots, e_{n-1}\}$ совпадает с матрицей, полученной из матрицы B вычеркиванием последней строки и последнего столбца, поэтому угловые миноры матрицы B' совпадают с Δ_i, $i \leqslant n-1$ и, в частности, отличны от нуля. Отсюда следует, что к функции $\beta\mid_U$ применимо предположение индукции. Значит, найдется базис $\mathcal{F}' = \{f_1, \ldots, f_{n-1}\}$ в подпространстве U, в котором

$$q\mid_U (y_1 f_1 + \ldots + y_{n-1} f_{n-1}) = \frac{\Delta_1}{\Delta_0} y_1^2 + \ldots + \frac{\Delta_{n-1}}{\Delta_{n-2}} y_{n-1}^2,$$

причем матрица перехода T' от $\mathcal{E}'$ к $\mathcal{F}'$ является (верхней) унитреугольной. Следовательно, матрица $\widetilde{B}'$ билинейной функции $\beta\mid_U$ в базисе $\mathcal{F}'$ имеет вид

$$\widetilde{B}' = \begin{pmatrix} \frac{\Delta_1}{\Delta_0} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{\Delta_{n-1}}{\Delta_{n-2}} \end{pmatrix}.$$

т.е. для любых $i, j = \overline{1, n-1}$ имеем $\beta(f_i, f_j) = 0$ при $i \neq j$ и $\beta(f_i, f_i) = \frac{\Delta_i}{\Delta_{i-1}}$.

Положим

$$f_n = e_n - \frac{\beta(e_n, f_1)}{\beta(f_1, f_1)} f_1 - \ldots - \frac{\beta(e_n, f_{n-1})}{\beta(f_{n-1}, f_{n-1})} f_{n-1}. \tag{26}$$

В этой формуле мы имеем право делить на $\beta(f_i, f_i)$ при $i = \overline{1, n-1}$, так как

$$\beta(f_i, f_i) = \frac{\Delta_i}{\Delta_{i-1}} \neq 0.$$

Так как f_n не принадлежит U, то $\{f_1, \ldots, f_{n-1}, f_n\}$ – базис пространства V. Осталось проверим, что базис $\{f_1, \ldots, f_{n-1}, f_n\}$ – искомый.

Непосредственно проверяется, что матрица T перехода от базиса $\{e_1, \ldots, e_n\}$ к базису $\{f_1, \ldots, f_{n-1}, f_n\}$ является (верхней) унитреугольной, $T = \begin{pmatrix} T' & * \\ 0 & 1 \end{pmatrix}$.

Так как $\beta(f_i, f_j) = 0$ для любых $i, j = \overline{1, n-1}$ $(i \neq j)$, то, учитывая формулу (26) построение вектора f_n, для любого $k = \overline{1, n-1}$ имеем:

$$\beta(f_n, f_k) = \beta(e_n - \frac{\beta(e_n, f_1)}{\beta(f_1, f_1)} f_1 - \ldots - \frac{\beta(e_n, f_{n-1})}{\beta(f_{n-1}, f_{n-1})} f_{n-1}, f_k) =$$

$$= \beta(e_n, f_k) - \frac{\beta(e_n, f_1)}{\beta(f_1, f_1)} \beta(f_1, f_k) - \ldots - \frac{\beta(e_n, f_{n-1})}{\beta(f_{n-1}, f_{n-1})} \beta(f_{n-1}, f_k) =$$

$$= \beta(e_n, f_k) - \frac{\beta(e_n, f_k)}{\beta(f_k, f_k)} \beta(f_k, f_k) = 0.$$

Отсюда получаем, что в базисе $\{f_1, \ldots, f_{n-1}, f_n\}$ матрица $\widetilde{B}$ билинейной функции β имеет диагональный вид

$$\widetilde{B} = \begin{pmatrix} \widetilde{B}' & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{\Delta_1}{\Delta_0} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \frac{\Delta_{n-1}}{\Delta_{n-2}} & 0 \\ 0 & \cdots & 0 & \widetilde{b}_{n,n} \end{pmatrix}.$$

При этом

$$\det\, \widetilde{B} = \frac{\Delta_1}{\Delta_0} \cdot \frac{\Delta_2}{\Delta_1} \cdot \ldots \cdot \frac{\Delta_{n-1}}{\Delta_{n-2}} \cdot \widetilde{b}_{n,n} = \Delta_{n-1} \widetilde{b}_{n,n}.$$

С другой стороны, из формулы (20) получаем, что

$$\det\, \widetilde{B} = \det\, T^t B T = \det\, T^t \det\, B \det\, T = (\det\, T)^2 \det\, B = \det\, B = \Delta_n.$$

Так как $\Delta_{n-1} \neq 0$, получаем, что $\widetilde{b}_{n,n} = \frac{\Delta_n}{\Delta_{n-1}}$. ■

Отметим, что в доказательстве этой теоремы процедура приведения симметрической билинейной формы к каноническому виду с использованием формулы (26) называется *процессом ортогонализации.*

12. Квадратичные формы над $\mathbb{R}$ и над $\mathbb{C}$.

12.1. Нормальный вид квадратичных форм над $\mathbb{R}$ и над $\mathbb{C}$. Закон инерции.

Мы рассмотрим частный случай, особенно важный для приложений, когда основное поле – это $\mathbb{R}$ или $\mathbb{C}$.

Для начала рассмотрим конечномерное векторное пространство V над полем комплексных чисел $\mathbb{C}$. Если $\beta : V \times V \to \mathbb{C}$ – произвольная симметрическая билинейная функция, то по теореме 10.1 ее матрица $D = (d_{i,j})$ в некотором базисе $\mathcal{E} = \{e_1, \ldots, e_n\}$ является диагональной, причем можно считать, что лишь первые r диагональных элементов $d_{1,1} = \beta(e_1, e_1), \ldots, d_{r,r} = \beta(e_r, e_r)$ отличны от нуля, т.е. в этом базисе

$$\beta(x, y) = d_{1,1}x_1y_1 + d_{2,2}x_2y_2 + \ldots + d_{r,r}x_ry_y.$$

Значит, в базисе $\widetilde{\mathcal{E}} = \{\tilde{e}_1, \ldots, \tilde{e}_r, e_{r+1}, \ldots, e_n\}$, где $\tilde{e}_i = \frac{1}{\sqrt{\beta(e_i,e_i)}}e_i$, $i = \overline{1,r}$, матрица билинейной функции β будет диагональной с диагональными элементами, равными либо 1, либо 0. Таким образом, в базисе $\widetilde{\mathcal{E}}$

$$\beta(x, y) = \tilde{x}_1\tilde{y}_1 + \tilde{x}_2\tilde{y}_2 + \ldots + \tilde{x}_r\tilde{y}_r.$$

Так как число r равно рангу билинейной функции β, по следствию 10.1 число r не зависит от базиса, т.е. является инвариантом.

Поскольку матрицы для β и для ассоциированной с ней квадратичной формы q одинаковы, то получаем следующее

Утверждение 12.1. *Над полем комплексных чисел для любой квадратичной функции q существует базис, в котором*

$$q(x) = x_1^2 + \ldots + x_r^2. \tag{27}$$

При этом число r является инвариантом квадратичной функции q.

Выражение (27) для квадратичной функции q над полем комплексных чисел называется ее **нормальным видом**.

В случае конечномерного векторного пространства V над полем действительных чисел $\mathbb{R}$ абсолютно то же самое рассуждение доказывает

Утверждение 12.2. *Над полем действительных чисел для любой квадратичной функции q существует базис, в котором*

$$q(x) = x_1^2 + \ldots + x_k^2 - x_{k+1}^2 - \ldots - x_{k+l}^2, \tag{28}$$

причем сумма $k + l = r$ – инвариант квадратичной функции q.

Выражение (28) для квадратичной функции q над полем действительных чисел называется ее **нормальным видом**. Числа k и l в выражении (28) называются, соответственно, **положительным** и **отрицательным индексами инерции** квадратичной функции q.

Так как мы можем по-разному выбрать тот базис, в котором квадратичная функция записывается в виде суммы квадратов, то возникает вопрос, зависит ли от выбора базиса количество коэффициентов, равных $+1$ и -1, в нормальном виде квадратичной функции q над полем действительных чисел $\mathbb{R}$ или же эти числа зависят лишь от самой квадратичной функции q (т.е. являются ее инвариантами). Оказывается верно следующее.

Теорема 12.1. *(Закон инерции). Положительный и отрицательный индексы инерции являются инвариантами квадратичной функции (т.е. не зависят от выбора базиса).*

Доказательство. Пусть $\mathcal{E} = \{e_1, \ldots, e_n\}$ и $\widetilde{\mathcal{E}} = \{\tilde{e}_1, \ldots, \tilde{e}_n\}$ – два базиса векторного пространства V над полем $\mathbb{R}$ размерности n, в которых квадратичная функция q имеет виды

$$q(x) = x_1^2 + \ldots + x_k^2 - x_{k+1}^2 - \ldots - x_{k+l}^2, \tag{29}$$

$$q(x) = \tilde{x}_1^2 + \ldots + \tilde{x}_s^2 - \tilde{x}_{s+1}^2 - \ldots - \tilde{x}_{s+t}^2, \tag{30}$$

где $x_i, \tilde{x}_i$ – это координаты вектора x в базисах $\mathcal{E}, \widetilde{\mathcal{E}}$ соответственно.

Так как $k+l = s+t = r$, где r – ранг квадратичной функции q, достаточно доказать, что $k = s$. Допустим, что это не так. Пусть $k > s$. Рассмотрим подпространства $U = \langle e_1, \ldots, e_k\rangle$ и $W = \langle \tilde{e}_{s+1}, \ldots, \tilde{e}_{s+t}, \tilde{e}_{s+t+1} \ldots, \tilde{e}_n\rangle$. Согласно формуле Грассмана (Теорема 2.1),

$$\dim\,(U \cap W) = \dim\,U + \dim\,W - \dim\,(U+W) = k + (n-s) - \dim\,(U+W) =$$

$$= (n - \dim\,(U+W)) + (k-s) \geqslant k - s > 0.$$

Предпоследнее неравенство справедливо, так как размерность любого подпространства в V не превосходит размерности всего пространства (Теорема 1.4). Итак, мы видим, что подпространство $U \cap W$ – ненулевое. Значит, в $U \cap W$ существует ненулевой вектор u. Тогда $u = \lambda_1 e_1 + \ldots + \lambda_k e_k$, где не все λ_i равны нулю, и $u = \mu_{s+1}\tilde{e}_{s+1} + \ldots + \mu_n \tilde{e}_n$, где не все μ_j равны нулю. Подстановка в формулу (29) координат вектора u в базисе $\mathcal{E}$ дает

$$q(u) = \lambda_1^2 + \ldots + \lambda_k^2 > 0.$$

С другой стороны, подстановка в формулу (30) координат вектора u в базисе $\widetilde{\mathcal{E}}$ дает

$$q(u) = -\mu_{s+1}^2 - \ldots - \mu_{s+t}^2 \leqslant 0.$$

Полученное противоречие доказывает, что $k = s$. ∎

12.2. Положительно определенные квадратичные формы.

Пусть V – векторное пространство над $\mathbb{R}$.

Определение 12.1. *Квадратичная функция $q : V \to \mathbb{R}$ называется* **положительно определенной***, если $q(x) > 0$ для любого ненулевого вектора $x \in V$. Симметрическая билинейная функция $\beta : V \times V \to \mathbb{R}$ называется* **положительно определенной***, если соответствующая ей квадратичная функция положительно определена.*

Аналогично определяются **отрицательно определенные** квадратичные и симметрические билинейные функции.

ЗАДАЧА. Докажите, что в пространстве $M_n(\mathbb{R})$ билинейной функцией $\mathrm{tr}\ AB^t$ является симметрической и положительно определенной.

Очевидно, что нормальный вид положительно определенной квадратичной функции в n-мерном пространстве есть $x_1^2 + \ldots + x_n^2$.

Для решения вопроса о положительной определенности квадратичной функции используется так называемый критерий Сильвестра.

Теорема 12.2. *(Критерий Сильвестра). Симметрическая билинейная функция $\beta : V \times V \to \mathbb{R}$ является положительно определенной тогда и только тогда, когда все угловые миноры Δ_i ее матрицы B в произвольном базисе $\mathcal{E} = \{e_1, \ldots, e_n\}$ являются положительными числами.*

Доказательство. Пусть симметрическая билинейная функция β является положительно определенной. Утверждение будем доказывать индукцией по $n = \dim V$ с очевидным основанием при $n = 1$. Допустим, что утверждение теоремы верно для размерностей, меньших n. Пусть $\mathcal{E} = \{e_1, \ldots, e_n\}$ – произвольный базис пространства V, B – матрица билинейной функции β в этом базисе, а Δ_i – ее угловые миноры ($1 \leqslant i \leqslant n$). Так как функция β – положительно определенная, то ее ограничение на подпространство $U = \langle e_1, \ldots, e_{n-1} \rangle$ также положительно определено, откуда по предположению индукции вытекает, что угловые миноры Δ_i, $1 \leqslant i \leqslant n-1$, положительны. Далее осталось только доказать, что $\Delta_n > 0$. По теореме 10.1 существует базис $\widetilde{\mathcal{E}} = \{\tilde{e}_1, \ldots, \tilde{e}_n\}$, в котором матрица $\widetilde{B}$ билинейной функции β диагональна, причем ее диагональные элементы $\tilde{b}_{i,i} = \beta(\tilde{e}_i, \tilde{e}_i)$ положительны в силу положительной определенности билинейной функции β. Следовательно,

$$\det\ \widetilde{B} = \tilde{b}_{1,1}\tilde{b}_{2,2} \ldots \tilde{b}_{n,n} > 0.$$

С другой стороны, согласно формуле (20),

$$\det\ \widetilde{B} = (\det\ T^t)(\det\ B)(\det\ T) = (\det\ T)^2 \Delta_n,$$

где T – матрица перехода от базиса $\mathcal{E}$ к базису $\widetilde{\mathcal{E}}$. Отсюда $\Delta_n > 0$, как и требовалось.

Докажем теперь обратное утверждение. Пусть все угловые миноры $\Delta_1, \ldots, \Delta_n$ матрицы билинейной функции β в некотором базисе $\mathcal{E}$ положительны. В частности, они отличны от нуля. Значит, по теореме 11.1 о формуле Якоби существует базис, в котором квадратичная функция q, ассоциированная с β, имеет вид

$$q(x) = \frac{\Delta_1}{\Delta_0}x_1^2 + \ldots + \frac{\Delta_n}{\Delta_{n-1}}x_n^2.$$

Из этой формулы видно, что q является положительно определенной. ■

ЗАДАЧА. Докажите, что для того, чтобы симметрическая билинейная функция была отрицательно определена, необходимо и достаточно, чтобы знаки угловых миноров чередовались, начиная с минуса, т.е. $(-1)^i\Delta_i > 0$, $i = \overline{1, n}$.

13. Евклидовы пространства.

Векторное пространство существенно отличается от пространства геометрических векторов тем, что в векторном пространстве не определены понятия длины вектора и угла между векторами. В пространстве геометрических векторов, используя длину вектора и угол, определялось скалярное произведение. Здесь удобнее поступить наоборот. Мы аксиоматически определим операцию скалярного произведения, а длину и угол определим с ее помощью. Определение скалярного произведения для векторного пространства над $\mathbb{R}$ и для векторного пространства над $\mathbb{C}$ формулируется различно. В этом и следующем параграфе будем рассматривать вещественное пространство.

13.1. Скалярное произведение.

Определение 13.1. Евклидовым (векторным) пространством *называется векторное пространство над $\mathbb{R}$ с фиксированной положительно определенной симметрической билинейной функцией.*

Эта фиксированная билинейная функция называется **скалярным произведением** *и обозначается* $(*,*)$.

Итак, в соответствии с определением евклидово (векторное) пространство – это пара $(V,(*,*))$, где V – векторное пространство над $\mathbb{R}$, $(*,*)$ – фиксированная положительно определенная симметрическая билинейная функция. Отметим еще раз основные свойства (*аксиомы*) скалярного произведения:

1) $(x,y) = (y,x),\ \forall x,y \in V$;
2) $(\lambda x,y) = \lambda(x,y),\ \forall \lambda \in \mathbb{R}, \forall x,y \in V$;
3) $(x+y,z) = (x,z) + (y,z),\ \forall x,y,z \in V$;
4) $(x,x) > 0, \forall x \neq 0, x \in V$.

Из свойств симметричности (аксиома 1)) и линейности по первому аргументу (аксиомы 2)-3)), очевидно, следует линейность по второму аргументу. Также из линейности следует, что $(0,0) = 0$. Поэтому $(x,x) \geqslant 0$ для любого $x \in V$, и $(x,x) = 0$ тогда и только тогда, когда $x = 0$.

ПРИМЕР 1. В V^2 (или V^3) скалярное произведение двух геометрических векторов определено как произведение их длин на косинус угла между ними.

ПРИМЕР 2. В пространстве $\mathbb{R}^n$ скалярное произведение (так называемое *стандартное скалярное произведение*) можно определить по формуле

$$(x,y) = x_1y_1 + x_2y_2 + \ldots + x_ny_n. \tag{31}$$

ПРИМЕР 3. В пространстве $C[a,b]$ функций, непрерывных на отрезке $[a,b]$, можно ввести скалярное произведение по формуле

$$(f,g) = \int_a^b f(t)g(t)dt.$$

Определим с помощью введенного понятия скалярного произведения длину вектора и угол между векторами таким образом, чтобы в случае геометрических векторов они совпадали с обычной длиной и обычным углом.

Определение 13.2. Длиной $|x|$ *вектора* x *в евклидовом пространстве* V *называется число*

$$|x| = \sqrt{(x, x)}.$$

Углом *между ненулевыми векторами* x *и* y *называется число* α $(0 \leqslant \alpha \leqslant \pi)$ *такое, что*

$$\cos\alpha = \frac{(x, y)}{|x||y|}. \tag{32}$$

В силу положительной определенности скалярного произведения длина вектора – вещественное неотрицательное число, причем она равна нулю тогда и только тогда, когда вектор нулевой. С определением угла дело обстоит несколько сложнее. Для того, чтобы можно было определить α из равенства (32), нужно доказать, что

$$-1 \leqslant \frac{(x, y)}{|x||y|} \leqslant 1$$

или, что то же самое, что

$$\frac{(x, y)^2}{|x|^2|y|^2} \leqslant 1.$$

Последнее неравенство выполняется в силу следующего утверждения.

Утверждение 13.1. *Для любых векторов* x, y *евклидова пространства*

$$(x, y)^2 \leqslant (x, x)(y, y), \tag{33}$$

причем равенство имеет место тогда и только тогда, когда векторы x *и* y *пропорциональны.*

Неравенство (33) называется **неравенством Коши-Буняковского**.

Доказательство. Если $y = 0$, то неравенство (33), очевидно, выполняется. Если $y \neq 0$, рассмотрим вектор $x - ty$, где t – произвольное действительное число. Согласно положительной определенности скалярного произведения

$$(x - ty, x - ty) \geqslant 0,$$

т.е. для любого t

$$t^2(y, y) - 2t(x, y) + (x, x) \geqslant 0.$$

Так как стоящий слева квадратный относительно t трехчлен принимает лишь неотрицательные значения, дискриминант D уравнения

$$t^2(y, y) - 2t(x, y) + (x, x) = 0$$

не может быть положительным, т.е.

$$\frac{D}{4} = (x, y)^2 - (x, x)(y, y) \leqslant 0,$$

что и доказывает неравенство (33).

Если $(x, y)^2 - (x, x)(y, y) = 0$, т.е. $D = 0$, то либо $y = 0$ и, следовательно, $y = 0x$, либо $y \neq 0$ и для некоторого действительного числа t верно $t^2(y, y) - 2t(x, y) + (x, x) = 0$, т.е. $(x - ty, x - ty) = 0$, что возможно только при $x = ty$. Обратно, если x и y пропорциональны, то не ограничивая общности, можно считать, что $x = ty$ и

$$(x, y)^2 = (ty, y)^2 = t^2(y, y)^2 = t^2(y, y)(y, y) = (ty, ty)(y, y) = (x, x)(y, y).$$

■

Рассмотрим, как выглядит неравенство Коши-Буняковского в приведенных выше примерах евклидовых пространств.

ПРИМЕР 1. В V^2 (V^3) неравенство (33) не означает ничего нового.

ПРИМЕР 2. В $\mathbb{R}^n$ со скалярным произведением

$$(x, y) = \sum_{i=1}^{n} x_i y_i$$

имеем

$$(x, x) = \sum_{i=1}^{n} x_i^2, \ (y, y) = \sum_{i=1}^{n} y_i^2.$$

Поэтому неравенство (33) имеет здесь вид

$$(\sum_{i=1}^{n} x_i y_i)^2 \leqslant (\sum_{i=1}^{n} x_i^2)(\sum_{i=1}^{n} y_i^2).$$

ПРИМЕР 3. В $C[a, b]$ со скалярным произведением, заданным интегралом $(f, g) = \int_a^b f(t)g(t)dt$ неравенство (33) имеет вид

$$(\int_a^b f(t)g(t)dt)^2 \leqslant \int_a^b f^2(t)dt \cdot \int_a^b g^2(t)dt.$$

Это неравенство играет важную роль в математическом анализе.

Приведем пример неравенства, являющегося следствием неравенства Коши-Буняковского.

Утверждение 13.2. *Для любых векторов x и y в евклидовом пространстве имеет место неравенство, называемое* **неравенством треугольника***,*

$$|x + y| \leqslant |x| + |y|.$$

Доказательство.

$$|x+y|^2 = (x+y, x+y) = (x,x) + 2(x,y) + (y,y).$$

Так как в силу неравенства Коши-Буняковского $2(x,y) \leqslant 2|x||y|$, то

$$|x+y|^2 \leqslant (x,x) + 2|x||y| + (y,y) \leqslant (|x|+|y|)^2,$$

т.е. $|x+y| \leqslant |x| + |y|$, что и требовалось доказать.■

В следующей лекции мы увидим, что в евклидовом пространстве можно измерять и объемы параллелепипедов, ребрами которых является произвольный набор линейно независимых векторов $\{v_1, \ldots, v_k\}$. Для этого можно использовать матрицу порядка k

$$G = G(v_1, \ldots, v_k) = \begin{pmatrix} (v_1,v_1) & (v_1,v_2) & \ldots & (v_1,v_k) \\ (v_2,v_1) & (v_2,v_2) & \ldots & (v_2,v_k) \\ . & . & . & . \\ (v_k,v_1) & (v_k,v_2) & \ldots & (v_k,v_k) \end{pmatrix},$$

называемую **матрицей Грама**. Заметим, что матрица Грама является симметрической.

Утверждение 13.3. *Пусть* $v_1, \ldots, v_k$ *– произвольные векторы евклидова пространства. Тогда* $\det G(v_1, \ldots, v_k) > 0$ *тогда и только тогда, когда* $\{v_1, \ldots, v_k\}$ *– линейно независимая система. Иначе* $\det G(v_1, \ldots, v_k) = 0$.

Доказательство. Пусть $\det G(v_1, \ldots, v_k) > 0$. Допустим, что $\{v_1, \ldots, v_k\}$ линейно зависимы, т.е. существуют такие числа $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$, не все равные нулю, что

$$\lambda_1 v_1 + \ldots + \lambda_k v_k = 0.$$

Тогда для любого v_i выполняется

$$0 = (\lambda_1 v_1 + \ldots + \lambda_k v_k, v_i) = \lambda_1 (v_1, v_i) + \ldots + \lambda_k (v_k, v_i) = \lambda_1 g_{1,i} + \ldots + \lambda_k g_{k,i},$$

где $g_{i,j}$ – элементы матрицы Грама $G(v_1, \ldots, v_k)$. Отсюда вытекает линейная зависимость строк матрицы Грама $G(v_1, \ldots, v_k)$, откуда $\det G(v_1, \ldots, v_k) = 0$, что противоречит предположению и доказывает линейную независимость системы $\{v_1, \ldots, v_k\}$.

Обратно, пусть $\{v_1, \ldots, v_k\}$ – линейно независимая система. Дополним систему $\{v_1, \ldots, v_k\}$ до базиса $\mathcal{E} = \{v_1, \ldots, v_k, e_{k+1}, \ldots, e_n\}$ пространства V (это возможно по теореме 1.3). Рассмотрим симметрическую билинейную функцию

$$\beta(x,y) \equiv (x,y),$$

где (x,y) – скалярное произведение векторов x и y. Тогда $\det G(v_1, \ldots, v_k)$ равняется угловому минору Δ_k порядка k матрицы этой билинейной функции в базисе $\mathcal{E}$. Согласно критерию Сильвестра (теорема 12.2), в силу положительной определенности билинейной функции β, т.е. в силу положительной определенности скалярного произведения, имеем $\Delta_k > 0$. ∎

Заметим, что из этого утверждения следует, что для любых векторов $v_1, \dots, v_k$ евклидова пространства верно неравенство $\det G(v_1, \dots, v_k) \geqslant 0$. Причем неравенство Коши-Буняковского – частный случай этого утверждения при $k = 2$:

$$\det \begin{pmatrix} (x,x) & (x,y) \\ (y,x) & (y,y) \end{pmatrix} = (x,x)(y,y) - (x,y)^2 \geqslant 0.$$

Утверждение 13.4. *Пусть $\mathcal{E} = \{e_1, \dots, e_n\}$ – произвольный базис конечномерного евклидова пространства V, $x = \sum_{i=1}^n x_i e_i$, $y = \sum_{j=1}^n y_j e_j$ – произвольные векторы в V, X, Y – их столбцы координат в базисе $\mathcal{E}$. Тогда*

$$(x,y) = \sum_{i,j=1}^n x_i y_j g_{i,j} = X^t G Y,$$

где $g_{i,j} = (e_i, e_j)$ – элементы матрицы Грама $G = G(e_1, \dots, e_n)$.

Доказательство. Так как скалярное произведение – это билинейная функция, а матрица Грама $G = G(e_1, \dots, e_n)$ – это матрица данной билинейной функции в базисе $\mathcal{E}$, то достаточно воспользоваться записями билинейной функции в координатном виде (18) и в матричном виде (19). ∎

13.2. Ортогональные базисы.

Определение 13.3. *Векторы x и y евклидова пространства V называются* **ортогональными***, если $(x, y) = 0$.*

Таким образом, x и y ортогональны тогда и только тогда, когда либо угол между x и y равен $\frac{\pi}{2}$, либо один из этих векторов нулевой.

В частности, если x и y ортогональны, то $|x + y|^2 = |x|^2 + |y|^2$ (теорема Пифагора). Действительно,

$$|x+y|^2 = (x+y, x+y) = (x,x) + 2(x,y) + (y,y) = (x,x) + (y,y) = |x|^2 + |y|^2.$$

Эту теорему можно обобщить: если $x, y, z, \dots$ – конечное число попарно ортогональных векторов, то $|x + y + z + \dots|^2 = |x|^2 + |y|^2 + |z|^2 + \dots$

Полезно следующее наблюдение:

Утверждение 13.5. *Любая ортогональная система $\{v_1, \dots, v_k\}$ ненулевых векторов линейно независима.*

Доказательство. Допустим, существует нетривиальная линейная комбинация, равная нулю:

$$\lambda_1 v_1 + \ldots + \lambda_k v_k = 0.$$

Для каждого $i = 1, \ldots, k$ скалярно умножим обе части этого равенства на v_i. Получим $\lambda_i(v_i, v_i) = 0$. Так как v_i – ненулевой вектор, $(v_i, v_i) \neq 0$. Значит, $\lambda_i = 0$. ■

В векторном пространстве у нас нет оснований предпочитать один базис другому – там все базисы равноправны. В евклидовом пространстве существует наиболее удобные базисы, а именно ортонормированные базисы. Они играют здесь ту же роль, что и прямоугольная система координат в аналитической геометрии.

Определение 13.4. *Базис $\{e_1, \ldots, e_n\}$ евклидова пространства называется* **ортогональным***, если векторы базиса попарно ортогональны.*

Базис $\{e_1, \ldots, e_n\}$ евклидова пространства называется **ортонормированным***, если он является ортогональным и каждый вектор базиса имеет единичную длину, т.е.*

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

Из утверждения 13.5 следует, что в n-мерном евклидовом пространстве любые n ненулевых попарно ортогональных векторов образуют базис.

Непосредственно проверяется следующее

Утверждение 13.6. *Пусть $\{e_1, \ldots, e_n\}$ – ортонормированный базис евклидова пространства V, $x = \sum_{i=1}^n x_i e_i$, $y = \sum_{j=1}^n y_j e_j \in V$. Тогда*

1) матрица Грамма $G(e_1, \ldots, e_n)$ базисных векторов является единичной матрицей E;

2) $(x, y) = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n = X^t Y$, где $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$;

3) для координат вектора x верно $x_i = (x, e_i)$ (x_i естественно называть проекцией на e_i), т.е.

$$x = (x, e_1)e_1 + \ldots + (x, e_n)e_n.$$

Чтобы доказать существование ортогональных базисов, воспользуемся так называемым **процессом ортогонализации Грама-Шмидта.**

Процесс ортогонализации Грама - Шмидта позволяет по любой линейно независимой системе векторов $\{f_1, \ldots, f_k\}$ построить такую ортогональную систему $\{e_1, \ldots, e_k\}$, что для любого i имеем: $< f_1, \ldots, f_i > = < e_1, \ldots, e_i >$. Сначала, положим

$$e_1 = f_1.$$

Применяя индукцию, будем считать, что $\{e_1, \ldots, e_m\}$ уже построены, $(e_i, e_j) = 0$ при $1 \leqslant i, j \leqslant m, i \neq j$. Положим

$$e_{m+1} = f_{m+1} - \frac{(f_{m+1}, e_1)}{(e_1, e_1)} e_1 - \ldots - \frac{(f_{m+1}, e_m)}{(e_m, e_m)} e_m.$$

Непосредственная проверка показывает, что новый вектор e_{m+1} ортогонален к предыдущим, а также выполняется $< f_1, \ldots, f_{m+1} > = < e_1, \ldots, e_{m+1} >$. В силу линейной независимости системы $\{f_1, \ldots, f_k\}$ мы видим, что e_{m+1} – ненулевой вектор, так что $(e_{m+1}, e_{m+1}) \neq 0$ и процесс может быть продолжен далее и доведен до конца с получением ортогональной системы $\{e_1, \ldots, e_k\}$.

Чтобы получить не просто ортогональную, но ортонормированную систему векторов $\{\tilde{e}_1, \ldots, \tilde{e}_m\}$ с теми же свойствами достаточно нормировать каждый вектор e_i, т.е. поделить каждый вектор e_i на его длину: $\tilde{e}_i = \frac{1}{|e_i|} e_i$.

Теорема 13.1. *Во всяком n-мерном евклидовом пространстве существует ортогональный базис.*

Доказательство. По определению n-мерного пространства в нем существует какой-то базис $\{f_1, \ldots, f_n\}$. С помощью процесса ортогонализации Грама-Шмидта из него можно построить ортогональный базис $\{e_1, \ldots, e_n\}$, что и доказывает теорему. ■

Нетрудно видеть, что таких базисов существует много. Действительно, уже из данного базиса $\{f_1, \ldots, f_n\}$ можно построить разные ортогональные базисы, если начинать построение с разных векторов f_i. Позднее мы рассмотрим вопрос о связи между различными ортонормированными базисами.

Непосредственно из процесса ортогонализации Грама-Шмидта получаем

Утверждение 13.7. *Пусть $\mathcal{E} = \{e_1, \ldots, e_n\}$ – ортогональный базис, полученный из базиса $\mathcal{F} = \{f_1, \ldots, f_n\}$ с помощью процесса ортогонализации Грама-Шмадта, а $\widetilde{\mathcal{E}} = \{\tilde{e}_1, \ldots, \tilde{e}_n\}$ – ортонормированный базис, полученный из базиса $\mathcal{E}$ делением каждого вектора из $\mathcal{E}$ на его длину. Тогда матрица перехода от базиса $\mathcal{E}$ к базису $\mathcal{F}$ будет верхнетреугольной с единицами на диагонали, а матрица перехода от базиса $\widetilde{\mathcal{E}}$ к базису $\mathcal{F}$ будет верхнетреугольной с положительными диагональными элементами.*

ПРИМЕР 4. Рассмотрим пространство $\mathbb{R}_n[x]$ многочленов степени $\leqslant n$ со скалярным произведением

$$(f, g) = \int_{-1}^{1} f(x) g(x) dx.$$

Возьмем базис $1, x, x^2, \dots, x^n$. Процесс ортогонализации приведет к последовательности многочленов $1, x, x^2 - \frac{1}{3}, x^3 - \frac{3}{5}x, \dots$ Эти многочлены с точностью до множителей совпадают с многочленами

$$p_0(x) = 1, \;\; p_k(x) = \frac{1}{2^k \cdot k!}\frac{d^k}{dx^k}(x^2-1)^k \;\; (k = \overline{1, n}),$$

которые называются многочленами Лежандра. Многочлены Лежандра образуют ортогональный, но не ортонормированный базис в $\mathbb{R}_n[x]$. Умножая каждый из этих многочленов на соответствующий множитель, мы можем построить ортонормированный базис. Пусть $r_0(x), r_1(x), \dots, r_n(x)$ – нормированные многочлены Лежандра нулевой, первой, …, n-й степени. Так как они образуют ортонормированный базис в $\mathbb{R}_n[x]$, произвольный многочлен $q(x)$ из $\mathbb{R}_n[x]$ представим в виде

$$q(x) = c_0 r_0(x) + c_1 r_1(x) + \ldots + c_n r_n(x),$$

где коэффициенты c_i, как это следует из утверждения 13.6 п.3, вычисляются по формуле

$$c_i = (q(x), r_i(x)) = \int_{-1}^{1} q(x) r_i(x) dx.$$

ПРИМЕР 5. Рассмотрим на интервале $(0, 2\pi)$ систему функций

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \cos nx, \sin nx. \tag{34}$$

Их линейная комбинация

$$P(x) = \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \ldots + a_n \cos nx + b_n \sin nx$$

называется тригонометрическим многочленом n-го порядка. Совокупность тригонометрических многочленов n-го порядка образует $(2n+1)$-мерное пространство $\mathcal{R}_n$. Определим в $\mathcal{R}_n$ скалярное произведение, как обычно, т.е. положим

$$(P(x), Q(x)) = \int_0^{2\pi} P(x) Q(x) dx.$$

Легко проверить, что система (34) будет ортогональным базисом. Действительно,

$$\int_0^{2\pi} \cos kx \cos lx dx = 0, \;\; \text{если } k \neq l,$$

$$\int_0^{2\pi} \sin kx \cos lx dx = 0,$$

$$\int_0^{2\pi} \sin kx \sin lx dx = 0, \;\; \text{если } k \neq l.$$

Так как
$$\int_0^{2\pi} 1dx = 2\pi, \int_0^{2\pi} \cos^2 kx = \pi, \int_0^{2\pi} \sin^2 kx = \pi, \text{ если } k \neq 0,$$
то функции
$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos x, \frac{1}{\sqrt{\pi}}\sin x, \frac{1}{\sqrt{\pi}}\cos 2x, \frac{1}{\sqrt{\pi}}\sin 2x, \ldots, \frac{1}{\sqrt{\pi}}\cos nx, \frac{1}{\sqrt{\pi}}\sin nx \tag{35}$$
образуют в $\mathcal{R}_n$ ортонормированный базис.

13.3. Матрицы перехода от одного ортонормированного базиса к другому.

В силу особой важности ортонормированных базисов для евклидова пространства полезно указать вид матриц перехода от одного ортонормированного базиса к другому.

Пусть даны два ортонормированных базиса $\mathcal{E} = \{e_1, \ldots, e_n\}$ и $\mathcal{F} = \{f_1, \ldots, f_n\}$. И пусть $T = T_{\mathcal{E}\to\mathcal{F}} = (t_{i,j})$ – матрица перехода от базиса $\mathcal{E}$ к базису $\mathcal{F}$. Имеем, $f_i = \sum_{k=1}^n t_{k,i}e_k$, $f_j = \sum_{l=1}^n t_{l,j}e_l$ для любой пары индексов $i, j = \overline{1, n}$. Вычисляя скалярное произведение в ортонормированном базисе $\mathcal{E}$ (утверждение 13.6 п.2), получаем
$$(f_i, f_j) = \sum_{k=1}^n t_{k,i}t_{k,j}.$$
Кроме того, так как $\mathcal{F}$ – ортонормированный базис, должно выполняться
$$(f_i, f_j) = \delta_i^j.$$
Следовательно,
$$\sum_{k=1}^n t_{k,i}t_{k,j} = \delta_i^j.$$

Это равенство говорит о том, что столбцы матрицы T образуют ортонормированный базис в пространстве $\mathbb{R}^n$ последовательностей длины n относительно стандартного скалярного произведения. На языке матриц имеем $T^tT = E$. Отсюда $T^t = T^{-1}$, так что и $TT^t = E$. Последнее равенство показывает, что система строк матрицы T тоже является ортонормированным базисом в $\mathbb{R}^n$.

Определение 13.5. *Матрица T, удовлетворяющая условию $T^tT = E$, называется* **ортогональной.**

Условие $T^tT = E$ ортогональности матрицы T эквивалентно любому из следующих условий:

1) $\sum_{k=1}^n t_{k,i}t_{k,j} = \delta_i^j$ для всех i, j;
2) $\sum_{k=1}^n t_{i,k}t_{j,k} = \delta_i^j$ для всех i, j;

3) $T^t = T^{-1}$;

4) $TT^t = E$.

ПРИМЕР. Матрица

$$T = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

является ортогональной, что можно проверить непосредственно.

Свойства ортогональных матриц:

1) если T – ортогональная матрица, то матрица T^{-1}, обратная к матрице T, является ортогональной матрицей;

2) если T_1, T_2 – ортогональные матрицы, то произведение T_1T_2 матриц T_1 и T_2 является ортогональной матрицей;

3) если T – ортогональная матрица, то $\det T = \pm 1$ (но не наоборот!).

4) множество O_n ортогональных матриц образует подгруппу в группе $\mathrm{GL}_n(\mathbb{R})$ всех невырожденных квадратных матриц порядка n над полем $\mathbb{R}$.

Доказательство свойств. Из свойств операций над матрицами получаем следующее.

1) Так как T – ортогональная матрица, т.е. $TT^t = E$, то $(T^{-1})^t T^{-1} = (T^t)^{-1}T^{-1} = (TT^t)^{-1} = E^{-1} = E$. Следовательно, T^{-1} – ортогональная матрица.

2) Так как T_1, T_2 – ортогональные матрицы, т.е. $T_1^tT_1 = E$ и $T_2^tT_2 = E$, получаем, что $(T_1T_2)^t(T_1T_2) = T_2^tT_1^tT_1T_2 = T_2^{-1}ET_2 = T_2^{-1}T_2 = E$. Следовательно, T_1T_2 – ортогональная матрица.

3) Так как T – ортогональная матрица, то $T^tT = E$. Посчитаем определитель от правой и левой части этого равенства:

$$\det E = 1$$

и

$$\det T^tT = \det T^t \det T = \det T \det T = det^2\ T$$

(здесь мы воспользовались тем, что определитель произведения равен произведению определителей и определитель от транспонированной матрицы равен определителю от самой матрицы). Следовательно, $\det T = \pm 1$.

4) Множество O_n образует подгруппу в группе $\mathrm{GL}_n(\mathbb{R})$ в силу первых двух свойств и того, что единичная матрица является ортогональной. ■

Утверждение 13.8. *Матрица перехода от одного ортонормированного базиса к другому является ортогональной. Обратно, если матрица T перехода от базиса $\mathcal{E} = \{e_1, \ldots, e_n\}$ к базису $\mathcal{F} = \{f_1, \ldots, f_n\}$ является ортогональной и один из базисов – ортонормированный, то и другой – ортонормированный.*

Доказательство. Первая часть этого утверждения уже доказана выше.

Для доказательства второй части сперва предположим, что базис $\mathcal{E}$ – ортонормированный. Так как $T = (t_{i,j})$ – ортогональная матрица,

$$\sum_{k=1}^{n} t_{k,i}t_{k,j} = \delta_i^j.$$

Вычислим скалярное произведение с учетом утверждения 13.6 п.2:

$$(f_i, f_j) = (\sum_{k=1}^{n} t_{k,i}e_k, \sum_{l=1}^{n} t_{l,j}e_l) = \sum_{k=1}^{n} t_{k,i}t_{k,j}.$$

Значит, $(f_i, f_j) = \delta_i^j$ для любой пары индексов i, j, т.е. $\mathcal{F}$ является ортонормированным базисом.

Предположим теперь, что изначально известна ортонормированность базиса $\mathcal{F}$. Так как матрица, обратная к ортогональной, сама ортогональна, получаем, что матрица перехода от $\mathcal{F}$ к $\mathcal{E}$, т.е. T^{-1}, является ортогональной. Это сводит задачу к предыдущему случаю. ∎

Утверждение 13.9. *(о QR-разложении)*
Если матрица A невырождена, то она может быть представлена в виде произведения $A = QR$, где Q – ортогональная, а R – верхняя треугольная матрица, причем диагональные элементы матрицы R положительны.

Доказательство. Рассмотрим столбцы матрицы A как столбцы координат векторов $f_1, \ldots, f_n$ в ортонормированном базисе $\mathcal{G} = \{g_1, \ldots, g_n\}$. Так как матрица A невырождена, векторы $\mathcal{F} = \{f_1, \ldots, f_n\}$ составляют базис. При этом A является матрицей $T_{\mathcal{G}\to\mathcal{F}}$ перехода от $\mathcal{G}$ к $\mathcal{F}$. Пусть $\mathcal{E}$ – ортонормированный базис, полученный из базиса $\mathcal{F}$ с помощью процесса ортогонализации Грама-Шмидта и нормировки. Тогда по утверждению 13.7 матрица $T_{\mathcal{E}\to\mathcal{F}}$ перехода от $\mathcal{E}$ к $\mathcal{F}$ является верхней треугольной с положительными элементами на главной диагонали. Кроме того, так как базисы $\mathcal{E}$ и $\mathcal{G}$ ортонормированы, матрица $T_{\mathcal{G}\to\mathcal{E}}$ перехода от $\mathcal{G}$ к $\mathcal{E}$ является ортогональной. В силу свойств матриц перехода получаем

$$T_{\mathcal{G}\to\mathcal{F}} = T_{\mathcal{G}\to\mathcal{E}}T_{\mathcal{E}\to\mathcal{F}},$$

т.е $A = QR$, где $Q = T_{\mathcal{G}\to\mathcal{E}}$, $R = T_{\mathcal{E}\to\mathcal{F}}$. ∎

14. Евклидовы пространства (продолжение).

14.1. Ортогональное дополнение.

Пусть V – евклидово пространство.

Определение 14.1. *Говорят, что* **вектор** $x \in V$ **ортогонален подпространству** U *евклидова пространства* V*, если он ортогонален любому вектору* $u \in U$*.*

Лемма 14.1. *Пусть* $\{e_1, \ldots, e_k\}$ *– некоторый базис в подпространстве* U *евклидова пространства* V*. Тогда вектор* $x \in V$ *ортогонален подпространству* U *тогда и только тогда, когда* $(x, e_1) = \ldots = (x, e_k) = 0$*.*

Доказательство. Если x ортогонален подпространству U, то он ортогонален всем векторам из U. В частности, x ортогонален базисным векторам подпространства U.

Обратно, пусть $(x, e_1) = \ldots = (x, e_k) = 0$. Любой вектор $u \in U$ выражается через базисные векторы: $u = \lambda_1 e_1 + \ldots + \lambda_k e_k$. Поэтому

$$(x, u) = (x, \lambda_1 e_1 + \ldots + \lambda_k e_k) = \lambda_1 (x, e_1) + \ldots + \lambda_k (x, e_k) = 0.$$

Это означает, что x ортогонален подпространству U. ■

Определение 14.2. *Пусть* S *– некоторое подмножество в евклидовом пространстве* V*.* **Ортогональным дополнением** $S^\perp$ *к* S *называется множество*

$$S^\perp = \{x \in V | (x, s) = 0 \ \ \forall s \in S\}.$$

Для любого $S \subset V$ ортогональное дополнение $S^\perp$ является подпространством в V. Действительно, $\forall x, y \in S^\perp$, $(x, s) = 0$, $(y, s) = 0$ для любого $s \in S$. Поэтому $(x + y, s) = (x, s) + (y, s) = 0 + 0 = 0$, т.е. $x + y \in S^\perp$. Также для любого $\lambda \in \mathbb{R}$, $(\lambda x, s) = \lambda(x, s) = \lambda \cdot 0 = 0$, т.е. $\lambda x \in S^\perp$.

Отметим, что ортогональное дополнение к подпространству U состоит из всех векторов, ортогональных подпространству U.

Теорема 14.1. *Для любого подпространства* U *конечномерного евклидова пространства* V *имеет место равенство:*

$$V = U \oplus U^\perp.$$

В частности, $U \cap U^\perp = \{0\}$ *и* $\dim U^\perp = n - \dim U$*, где* $n = \dim V$*.*

Доказательство. Пусть $k = \dim U$. Выберем какой-нибудь базис $\{f_1, \ldots, f_k\}$ в U и дополним его до базиса $\mathcal{F} = \{f_1, \ldots, f_k, f_{k+1}, \ldots, f_n\}$ в V (это возможно по теореме 1.3). Применим к $\mathcal{F}$ процесс ортогонализации Грама - Шмидта. Тогда получим ортонормированный базис $\mathcal{E} = \{e_1, \ldots, e_n\}$ в V. Причем начальный отрезок $\{e_1, \ldots, e_k\}$ базиса $\mathcal{E}$ будет являться базисом в U, так как процесс ортогонализации Грама-Шмидта таков, что $\langle f_1, \ldots, f_k \rangle = \langle e_1, \ldots, e_k \rangle$.

По лемме 14.1 вектор $x \in V$ лежит в $U^\perp$ тогда и только тогда, когда $(x, e_1) = \ldots = (x, e_k) = 0$. Но, согласно утверждению 13.6 в ортонормированном базисе $\mathcal{E}$ произвольный вектор $x \in V$ представляется в виде

$$x = (x, e_1)e_1 + \ldots + (x, e_k)e_k + (x, e_{k+1})e_{k+1} + \ldots + (x, e_n)e_n.$$

Следовательно, x лежит в $U^\perp$ тогда и только тогда, когда $x \in \langle e_{k+1}, \ldots, e_n \rangle$. Таким образом, $U^\perp = \langle e_{k+1}, \ldots, e_n \rangle$.

Теперь из того, что $U = \langle e_1, \ldots, e_k \rangle$, $U^\perp = \langle e_{k+1}, \ldots, e_n \rangle$ и $\{e_1, \ldots, e_n\}$ – базис в V, все утверждения теоремы следуют автоматически. ∎

Из теоремы 14.1 следует, что для любого подпространства U конечномерного евклидова пространства V любой вектор $x \in V$ имеет единственное разложение $x = u + u^\perp$, где $u \in U, u^\perp \in U^\perp$. Вектор u называется **ортогональной проекцией** вектора x на подпространство U и обозначается через $\mathrm{pr}_U x$; вектор $u^\perp$ называется **ортогональной составляющей** вектора x относительно подпространства U и обозначается через $\mathrm{ort}_U x$.

Если $\{e_1, \ldots, e_k\}$ – ортонормированный базис подпространства U, то ортогональная проекция $\mathrm{pr}_U x$ вектора x может быть найдена по формуле

$$\mathrm{pr}_U x = (x, e_1)e_1 + \ldots + (x, e_k)e_k.$$

Более общо, если $\{e_1, \ldots, e_k\}$ – ортогональный (но не обязательно ортонормированный) базис подпространства U, то

$$\mathrm{pr}_U x = \frac{(x, e_1)}{(e_1, e_1)}e_1 + \ldots + \frac{(x, e_k)}{(e_k, e_k)}e_k.$$

При описании процесса ортогонализации Грама-Шмидта для построения ортогональной системы векторов $\{e_1, \ldots, e_n\}$ по произвольной линейно независимой системе векторов $\{f_1, \ldots, f_n\}$ использовалась формула

$$e_{m+1} = f_{m+1} - \frac{(f_{m+1}, e_1)}{(e_1, e_1)}e_1 - \ldots - \frac{(f_{m+1}, e_m)}{(e_m, e_m)}e_m.$$

Таким образом

$$e_{m+1} = f_{m+1} - \mathrm{pr}_{U_m} f_{m+1} = \mathrm{ort}_{U_m} f_{m+1},$$

где $U_m = \langle e_1, \ldots, e_m \rangle$.

Утверждение 14.1. *Для любых подпространств U, W n-мерного евклидова пространства V справедливы соотношения:*

(i) $(U^\perp)^\perp = U$;
(ii) $(U + W)^\perp = U^\perp \cap W^\perp$;
(iii) $(U \cap W)^\perp = U^\perp + W^\perp$.

Доказательство. (i) Включение $U \subset (U^{\perp})^{\perp}$ очевидно. В силу совпадения размерностей: $\dim\,(U^{\perp})^{\perp} = n - \dim\,(U^{\perp}) = n - (n - \dim\,U) = \dim\,U$, имеет место равенство, что и доказывает (i).

(ii) Для доказательства будем использовать метод двух включений.

Пусть $x \in (U+W)^{\perp}$. Это означает, что $\forall y \in U+W, (x,y)=0$. Так как $U \subset U+W$, $\forall u \in U, (x,u)=0$, т.е. $x \in U^{\perp}$. Так как $W \subset U+W$, $\forall w \in W, (x,w)=0$, т.е. $x \in W^{\perp}$. Следовательно, $x \in U^{\perp} \cap W^{\perp}$, что доказывает включение $(U+W)^{\perp} \subset U^{\perp} \cap W^{\perp}$.

Теперь пусть $x \in U^{\perp} \cap W^{\perp}$. Значит, $x \in U^{\perp}$, т.е. $\forall u \in U, (x,u)=0$. Также $x \in W^{\perp}$, т.е. $\forall w \in W, (x,w)=0$. Так как любой вектор $y \in U+W$ представим в виде $y=u+w, u \in U, w \in W$, имеем $(x,y)=(x,u+w)=(x,u)+(x,w)=0$, что означает $x \in (U+W)^{\perp}$ и доказывает обратное включение $U^{\perp} \cap W^{\perp} \subset (U+W)^{\perp}$.

(iii) Равенство $(U \cap W)^{\perp} = U^{\perp} + W^{\perp}$ двойственно к равенству (ii). Действительно, для ортогональных дополнений к U и W верно равенство (ii):

$$(U^{\perp}+W^{\perp})^{\perp} = (U^{\perp})^{\perp} \cap (W^{\perp})^{\perp},$$

Если есть равенство подпространств, то есть равенство их ортогональных дополнений:

$$((U^{\perp}+W^{\perp})^{\perp})^{\perp} = ((U^{\perp})^{\perp} \cap (W^{\perp})^{\perp})^{\perp},$$

к которому применяем (i). ■

14.2. Расстояние между вектором и подпространством.

В качестве приложения рассмотренных понятий изучим вопрос о расстоянии между вектором x и подпространством U евклидова пространства V.

Определим **расстояние** ρ между векторами x и y евклидова пространства V по формуле

$$\rho(x,y) = |x-y|.$$

Это расстояние удовлетворяет аксиомам метрического пространства. (Докажите это!)

Напомним, что *метрика*, расстояние, на множестве Z – это определенная на декартовом произведении $Z \times Z$ функция ρ с неотрицательными действительными значениями, удовлетворяющая при любых $x,y,z \in Z$ условиям:

1) $\rho(x,y)=0 \Leftrightarrow x=y$ (аксиома тождества);

2) $\rho(x,z) \leqslant \rho(x,y)+\rho(y,z)$ (аксиома треугольника);

3) $\rho(x,y)=\rho(y,x)$ (аксиома симметрии).

Множество Z, наделенное некоторой метрикой, называется *метрическим пространством.*

Определение 14.3. Расстояние *между вектором x и подпространством U евклидова пространства определяется по формуле*

$$\rho(x, U) = \inf_{u \in U} \rho(x, u).$$

Углом *между ненулевым вектором x и ненулевым подпространством U называется наименьший из углов между x и ненулевыми векторами из U.*

Заметим, что в определении угла между векторами был определен угол $\alpha \in [0, \pi]$. Таким образом, в определении 14.3 речь идет об угле с максимальным косинусом.

Теорема 14.2. *1) Расстояние от вектора x конечномерного евклидова пространства V до подпространства $U \subset V$ равно $|\mathrm{ort}_U x|$, причем единственным ближайшим к x вектором подпространства U является $\mathrm{pr}_U x$.*

2) Угол между ненулевым вектором x и ненулевым подпространством U равен углу между x и $\mathrm{pr}_U x$, если $\mathrm{pr}_U x \neq 0$, иначе $\frac{\pi}{2}$.

Доказательство. Представим x в виде $x = \mathrm{pr}_U x + \mathrm{ort}_U x$.

1) Для произвольного вектора $u \in U$ имеем

$$\rho^2(x, u) = |x - u|^2 = |(\mathrm{pr}_U x + \mathrm{ort}_U x) - u|^2 = |(\mathrm{pr}_U x - u) + \mathrm{ort}_U x|^2.$$

Так как векторы $\mathrm{pr}_U x - u \in U$ и $\mathrm{ort}_U x \in U^\perp$ ортогональны, по теореме Пифагора далее получаем

$$|(\mathrm{pr}_U x - u) + \mathrm{ort}_U x|^2 = |\mathrm{pr}_U x - u|^2 + |\mathrm{ort}_U x|^2 \geqslant |\mathrm{ort}_U x|^2 = \rho^2(x, \mathrm{pr}_U x),$$

так что наименьшее значение для $\rho(x, u)$ достигается при $u = \mathrm{pr}_U x$ и равно $|\mathrm{ort}_U x|$, а при $u \neq \mathrm{pr}_U x$ неравенство строгое.

2) Если $\mathrm{pr}_U x = 0$, то вектор x ортогонален подпространству U.

Если $\mathrm{pr}_U x \neq 0$, для произвольного ненулевого вектора u из U требуется доказать, что

$$cos(x, u) = \frac{(x, u)}{|x||u|} \leqslant \frac{(x, \mathrm{pr}_U x)}{|x||\mathrm{pr}_U x|} = cos(x, \mathrm{pr}_U x).$$

Умножая на положительные числа обе части неравенства, получим

$$(x, u)|\mathrm{pr}_U x| \leqslant (x, \mathrm{pr}_U x)|u|$$
$$\Updownarrow$$
$$(\mathrm{pr}_U x + \mathrm{ort}_U x, u)|\mathrm{pr}_U x| \leqslant (\mathrm{pr}_U x + \mathrm{ort}_U x, \mathrm{pr}_U x)|u|$$
$$\Updownarrow$$
$$(\mathrm{pr}_U x, u)|\mathrm{pr}_U x| \leqslant (\mathrm{pr}_U x, \mathrm{pr}_U x)|u|$$
$$\Updownarrow$$
$$(\mathrm{pr}_U x, u) \leqslant |\mathrm{pr}_U x||u|.$$

Последнее неравенство верно, потому что это неравенство есть неравенство Коши - Буняковского, и равенство в нем достигается, лишь когда $\mathrm{pr}_U x$ и u линейно зависимы. Это и доказывает, что наибольший косинус получается для угла между x и его проекцией $\mathrm{pr}_U x \in U$. ■

Используя определитель матрицы Грама, можно получить явную формулу для расстояния от вектора до подпространства, заданного произвольным базисом.

Теорема 14.3. *Пусть U – ненулевое подпространство евклидова пространства V, $\{e_1, \dots, e_k\}$ – произвольный базис подпространства U. Тогда для произвольного вектора $x \in V$ верна формула*

$$\rho(x, U) = \frac{\det\ G(e_1, \dots, e_k, x)}{\det\ G(e_1, \dots, e_k)}.$$

Доказательство. Если $x \in U$, то $\rho(x, U) = 0$ и $\det\ G(e_1, \dots, e_k, x) = 0$.

Если x не принадлежит подпространству U, то система $\{e_1, \dots, e_k, x\}$ образует базис пространства $U \oplus \langle x \rangle$. Скалярное произведение на пространстве $U \oplus \langle x \rangle$ – это симметрическая билинейная функция, $G(e_1, \dots, e_k, x)$ – ее матрица в базисе $\{e_1, \dots, e_k, x\}$, причем по свойству матрицы Грама все ее угловые миноры отличны от нуля. Поэтому к скалярному произведению на пространстве $U \oplus \langle x \rangle$ можно применить теорему 11.1 о формуле Якоби. В этой теореме построение канонического базиса $\{f_1, \dots, f_k, f_{k+1}\}$ соответствует ортогонализации Грама-Шмидта и

$$f_{k+1} = e_{k+1} - \frac{(e_{k+1}, f_1)}{(f_1, f_1)} f_1 - \ldots - \frac{(e_{k+1}, f_k)}{(f_k, f_k)} f_k,$$

т.е.

$$f_{k+1} = e_{k+1} - \mathrm{pr}_U e_{k+1} = \mathrm{ort}_U e_{k+1},$$

где $e_{k+1} = x$.

По формуле Якоби имеем

$$|y_1 f_1 + \ldots + y_k f_k + y_{k+1} f_{k+1}|^2 = \frac{\Delta_1}{\Delta_0} y_1^2 + \ldots + \frac{\Delta_k}{\Delta_{k-1}} y_k^2 + \frac{\Delta_{k+1}}{\Delta_k} y_{k+1}^2.$$

Следовательно, $|\mathrm{ort}_U x|^2 = |f_{k+1}|^2 = \frac{\Delta_{k+1}}{\Delta_k} = \frac{\det\ G(e_1, \dots, e_k, x)}{\det\ G(e_1, \dots, e_k)}$. ■

Полученная формула может быть применена к вычислению объема параллелепипеда в евклидовом пространстве.

k-мерным *параллелепипедом*, натянутым на векторы $a_1, \dots, a_k$ евклидова пространства, будем называть множество

$$P(a_1, \dots, a_k) = \{\sum_i x_i a_i | 0 \leqslant x_i \leqslant 1\}.$$

Основанием этого k-мерного параллелепипеда будем называть $(k-1)$-мерный параллелепипед $P(a_1, \dots, a_{k-1})$, а его *высотой* будем называть длину вектора

$\mathrm{ort}_{\langle a_1,\ldots,a_{k-1}\rangle} a_k$. *Объемом* k-мерного ($k > 1$) параллелепипеда будем называть произведение его основания на высоту. Объемом одномерного параллелепипеда $P(a_1)$ будем называть длину вектора a_1. Будем обозначать объем параллелепипеда P через vol P.

Теорема 14.4. vol $P(a_1, \ldots, a_k)^2 = \det G(a_1, \ldots, a_k)$.

Доказательство. Докажем эту формулу индукцией по k.

При $k = 1$ она верна по определению.

При $k > 1$, согласно индуктивному определению объема, имеем

$$\mathrm{vol}\ P(a_1, \ldots, a_k) = \mathrm{vol}\ P(a_1, \ldots, a_{k-1}) \cdot h,$$

где h – длина вектора $\mathrm{ort}_{\langle a_1,\ldots,a_{k-1}\rangle} a_k$, т.е. расстояние от вектора a_k до подпространства $\langle a_1, \ldots, a_{k-1}\rangle$. По предположению индукции

$$\mathrm{vol}\ P(a_1, \ldots, a_{k-1})^2 = \det G(a_1, \ldots, a_{k-1}).$$

Если система $\{a_1, \ldots, a_{k-1}\}$ линейно зависима, то $\det G(a_1, \ldots, a_{k-1}) = 0$. Следовательно, vol $P(a_1, \ldots, a_k) = 0$ и $\det G(a_1, \ldots, a_k) = 0$.

Если система $\{a_1, \ldots, a_{k-1}\}$ линейно независима, то $\det G(a_1, \ldots, a_{k-1}) \neq 0$. Используя предположение индукции и теорему 14.3, получаем

$$\mathrm{vol}\ P(a_1, \ldots, a_k)^2 = \det G(a_1, \ldots, a_{k-1}) \cdot \frac{\det G(a_1, \ldots, a_k)}{\det G(a_1, \ldots, a_{k-1})} = \det G(a_1, \ldots, a_k).$$

■

ПРИМЕР. Рассмотрим пространство V^3 (или плоскость V^2) геометрических векторов. По определению скалярное произведение в V^3 (V^2) находится по формуле $(x, y) = |x||y| \cos\alpha$, где α – угол между векторами x и y. Поэтому

$$\det G(x, y) = \det \begin{pmatrix} (x,x) & (x,y) \\ (y,x) & (y,y) \end{pmatrix} = (x,x)(y,y) - (x,y)^2 =$$

$$= |x|^2|y|^2 - |x|^2|y|^2\cos^2\alpha = |x|^2|y|^2(1 - \cos^2\alpha) = |x|^2|y|^2\sin^2\alpha,$$

что равно квадрату площади параллелограмма, построенного на векторах x и y.

Следствие 14.1. *(о "геометрическом смысле" числа* $|\det A|$*)*

Пусть $a_1, \ldots, a_n$ *– произвольные векторы* n*-мерного евклидова пространства,* A *– квадратная матрица порядка* n*, столбцами которой являются столбцы координат векторов* $a_1, \ldots, a_n$ *в каком-нибудь ортонормированном базисе* $\{e_1, \ldots, e_n\}$*. Тогда*

$$\mathrm{vol}\ P(a_1, \ldots, a_n) = |\det A|.$$

Доказательство следует из того, что

$$G(a_1, \ldots, a_n) = A^t E A = A^t A$$

и, значит,

$$\det\, G(a_1, \ldots, a_n) = (\det\, A)^2.$$

■

Что касается знака числа $\det\, A$, то он может быть истолкован как ориентация (линейно независимых) векторов $\{a_1, \ldots, a_n\}$ по отношению к базису $\{e_1, \ldots, e_n\}$.

ПРИМЕР. В пространстве V^3 геометрических векторов ориентированный объем параллелепипеда, построенного на векторах x, y, z, как показано в аналитической геометрии, равен величине определителя

$$\upsilon = \det \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix},$$

где x_i, y_i, z_i – координаты векторов x, y, z в ортонормированном базисе $\{i, j, k\}$. Вычислим квадрат этого определителя:

$$\upsilon^2 = \det \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} \det \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} =$$

$$= \det \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} =$$

$$= \det \begin{pmatrix} x_1^2 + x_2^2 + x_3^2 & x_1y_1 + x_2y_2 + x_3y_3 & x_1z_1 + x_2z_2 + x_3z_3 \\ y_1x_1 + y_2x_2 + y_3x_3 & y_1^2 + y_2^2 + y_3^2 & y_1z_1 + y_2z_2 + y_3z_3 \\ z_1x_1 + z_2x_2 + z_3x_3 & z_1y_1 + z_2y_2 + z_3y_3 & z_1^2 + z_2^2 + z_3^2 \end{pmatrix} =$$

$$= \det \begin{pmatrix} (x, x) & (x, y) & (x, z) \\ (y, x) & (y, y) & (y, z) \\ (z, x) & (z, y) & (z, z) \end{pmatrix} = \det\, G(x, y, z).$$

Таким образом, как и в общем случае, определитель матрицы Грама векторов x, y, z равен квадрату объема параллелепипеда, построенного на этих векторах.

14.3. Изоморфизм евклидовых пространств.

Определение 14.4. *Евклидовы пространства V и $\widetilde{V}$ называются* **изоморфными**, *если существует отображение $\Phi : V \to \widetilde{V}$, являющееся изоморфизмом векторных пространств и удовлетворяющее условию*

$$(\Phi(x), \Phi(y)) = (x, y) \ \ \forall x, y \in V.$$

Само отображение Φ называется при этом **изоморфизмом** *евклидовых пространств V и $\widetilde{V}$.*

Из теоремы 1.5 ясно, что изоморфными могут быть только евклидовы пространства одинаковой размерности. Оказывается, верно и обратное.

Теорема 14.5. *Любые два евклидовых пространства одинаковой (конечной) размерности изоморфны.*

Доказательство. Пусть V и $\widetilde{V}$ – произвольные евклидовы пространства размерности n. Выберем ортонормированный базис $\mathcal{E} = \{e_1, \ldots, e_n\}$ в V и ортонормированный базис $\widetilde{\mathcal{E}} = \{\tilde{e}_1, \ldots, \tilde{e}_n\}$ в $\widetilde{V}$. Пусть $\Phi : V \to \widetilde{V}$ – изоморфизм векторных пространств, переводящий e_i в $\tilde{e}_i$ $(i = \overline{1, n})$. Рассмотрим произвольные векторы $x, y \in V$, $x = x_1 e_1 + \ldots + x_n e_n$, $y = y_1 e_1 + \ldots + x_n e_n$. Используя формулу вычисления скалярного произведения в координатах в ортонормированном базисе (утверждение 13.6 п.2) и то, что изоморфизм Φ сохраняет операции векторного пространства, получаем

$$(\Phi(x), \Phi(y)) = (x_1\Phi(e_1) + \ldots + x_n\Phi(e_n), y_1\Phi(e_1) + \ldots + x_n\Phi(e_n)) =$$
$$= (x_1\tilde{e}_1 + \ldots + x_n\tilde{e}_n, y_1\tilde{e}_1 + \ldots + x_n\tilde{e}_n) = x_1y_1 + \ldots + x_ny_n =$$
$$= (x_1e_1 + \ldots + x_ne_n, y_1e_1 + \ldots + x_ne_n) = (x, y).$$

■

Из этой теоремы получаем, что любое n-мерное евклидово пространство изоморфно пространству $\mathbb{R}^n$ со стандартным скалярным произведением

$$(x, y) = \sum_{i=1}^{n} x_i y_i.$$

Также из этой теоремы можно вывести интересное следствие: любое геометрическое утверждение о двух или трех векторах евклидова пространства достаточно проверить в известном из элементарной геометрии трехмерном пространстве. Действительно, линейная комбинация этих векторов образует подпространство, размерности не выше трех. В силу теоремы 14.5 это подпространство изоморфно обычному трехмерному пространству V^3 (либо его подпространству), а значит, это подпространство устроено совершенно так же, как V^3 (соотв., как подпространство пространства V^3). Следовательно, геометрическое утверждение достаточно проверить в последнем пространстве.

Например, таким способом можно доказать неравенство Коши-Буняковского и неравенство треугольника.

Мы завершим рассмотрение евклидовых пространств рассмотрением линейных функций. Основной результат состоит в том, что линейные функции на евклидовом пространстве могут быть вычислены с помощью скалярного произведения. Именно, справедливо следующее.

Теорема 14.6. *Для любого вектора x конечномерного евклидова пространства V определим линейную функцию $\varphi_x : V \to \mathbb{R}$, полагая $\varphi_x(v) = (v, x)$. Тогда отображение $\Phi : x \mapsto \varphi_x$ – изоморфизм векторных пространств V и V^*.*

Доказательство. Во-первых, φ_x действительно является линейной функцией в силу линейности скалярного произведения по первому аргументу.

Далее, Φ сохраняет операции векторного пространства. Действительно, для любого вектора $v \in V$ имеем

$$\varphi_{x+y}(v) = (v, x+y) = (v, x) + (v, y) = \varphi_x(v) + \varphi_y(v) = (\varphi_x + \varphi_y)(v).$$

Поэтому $\varphi_{x+y} = \varphi_x + \varphi_y$. Значит, $\Phi(x+y) = \Phi(x) + \Phi(y)$. Аналогично можно получить, что $\Phi(\lambda x) = \lambda \Phi(x)$.

Осталось проверить биективность отображения Φ.

Допустим для каких-то различных векторов $x, y \in V$ оказывается, что $\Phi(x) = \Phi(y)$, т.е. $\varphi_x = \varphi_y$. Это означает, что для любого вектора $v \in V$ выполняется $(v, x) = (v, y)$, отсюда $(v, x - y) = 0$. Взяв в качестве v вектор $x - y$, получим $(x-y, x-y) = 0$, откуда в силу положительной определенности скалярного произведения следует, что $x - y = 0$, т.е. $x = y$. Это противоречие доказывает инъективность отображения Φ.

Рассмотрим теперь какой-нибудь базис $\{e_1, \ldots, e_n\}$ векторного пространства V. В силу инъективности векторы $\{\Phi(e_1), \ldots, \Phi(e_n)\}$ будут линейно независимы (проверьте это!). Так как $\dim V = \dim V^*$ (следствие 3.1), векторы $\{\Phi(e_1), \ldots, \Phi(e_n)\}$ будут образовывать базис в V^*. Значит, для любой линейной функции $\psi \in V^*$ существуют $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ такие, что

$$\psi = \lambda_1 \Phi(e_1) + \ldots + \lambda_n \Phi(e_n) = \Phi(\lambda_1 e_1 + \ldots + \lambda_n e_n) = \Phi(x),$$

где $x = \lambda_1 e_1 + \ldots + \lambda_n e_n \in V$. Это доказывает сюръективность Φ. ■

15. Линейные операторы в евклидовом пространстве.

Все сказанное выше о линейных операторах в произвольном векторном пространстве остается, конечно, в силе и в евклидовом пространстве, т.е. в векторном пространстве над $\mathbb{R}$ с заданным на нем скалярным произведением. С введением же скалярного произведения операторы приобретают новые свойства подобно тому, как векторы приобретают длину.

Для начала рассмотрим, какая связь возникает между линейными операторами и билинейными функциями в случае евклидова пространства.

15.1. Связь между линейными операторами и билинейными функциями в евклидовом пространстве.

Пусть V – конечномерное евклидово пространство (над полем действительных чисел).

Как мы знаем (Теорема 14.6), каждому вектору $v \in V$ соответствует линейная функция $\varphi_v \in V^*$, заданная по правилу

$$\varphi_v(x) = (x, v),$$

и это соответствие $v \mapsto \varphi_v$ является изоморфизмом векторных пространств V и V^*. Причем, если $\mathcal{E} = \{e_1, \dots, e_n\}$ – произвольный ортонормированный базис в V, то коэффициенты $\varphi_v(e_i) = (e_i, v)$ линейной функции φ_v в этом базисе равны координатам вектора v в этом базисе.

Аналогично, каждому линейному оператору $\mathcal{A} : V \to V$ соответствует билинейная функция

$$\beta_{\mathcal{A}}(x, y) = (x, \mathcal{A}y). \tag{36}$$

При этом матрица билинейной функции $\beta_{\mathcal{A}}$ в ортонормированном базисе $\mathcal{E}$ совпадает с матрицей оператора $\mathcal{A}$ в этом базисе. Проверим это.

Для начала докажем, что правило (36) действительно задает билинейную функцию. Имеем:

1) $\beta_{\mathcal{A}}(x, y+z) = (x, \mathcal{A}(y+z)) = (x, \mathcal{A}y + \mathcal{A}z) = (x, \mathcal{A}y) + (x, \mathcal{A}z) =$

$$= \beta_{\mathcal{A}}(x, y) + \beta_{\mathcal{A}}(x, z);$$

$$\beta_{\mathcal{A}}(x, \lambda y) = (x, \mathcal{A}(\lambda y)) = (x, \lambda\mathcal{A}y) = \lambda(x, \mathcal{A}y) = \lambda\beta_{\mathcal{A}}(x, y).$$

2) $\beta_{\mathcal{A}}(x+y, z) = (x+y, \mathcal{A}z) = (x, \mathcal{A}z) + (y, \mathcal{A}z) = \beta_{\mathcal{A}}(x, z) + \beta_{\mathcal{A}}(y, z);$

$$\beta_{\mathcal{A}}(\lambda x, y) = (\lambda x, \mathcal{A}y) = \lambda(x, \mathcal{A}y) = \lambda\beta_{\mathcal{A}}(x, y).$$

Пусть теперь $B = (b_{i,j})$ – матрица билинейной функции $\beta_{\mathcal{A}}$ в ортонормированном базисе $\mathcal{E} = \{e_1, \dots, e_n\}$, а $A = (a_{i,j})$ – матрица оператора $\mathcal{A}$ в этом базисе. Тогда для всех i, j имеем

$$b_{i,j} = \beta_{\mathcal{A}}(e_i, e_j) = (e_i, \mathcal{A}e_j) = (e_i, \sum_{k=1}^n a_{k,j}e_k) = \sum_{k=1}^n a_{k,j}(e_i, e_k) = \sum_{k=1}^n a_{k,j}\delta_i^k = a_{i,j},$$

т.е. $A = B$. В частности, отсюда следует, что отображение $\mathcal{A} \mapsto \beta_{\mathcal{A}}$, определяемое формулой (36), инъективно.

Более того, для любой билинейной функции β в пространстве V найдется линейный оператор $\mathcal{A}$ такой, что $\beta = \beta_{\mathcal{A}}$. Действительно, пусть теперь $B = (b_{i,j})$ – матрица произвольной билинейной функции β в ортонормированном базисе $\mathcal{E} = \{e_1, \ldots, e_n\}$, X, Y – столбцы координат произвольных векторов $x, y \in V$ в данном базисе. Как мы знаем (смотри формулу (19)), билинейная функция в матричном виде запишется так

$$\beta(x, y) = X^t BY.$$

Рассмотрим линейный оператор $\mathcal{A}$, который в базисе $\mathcal{E}$ имеет матрицу B. И пусть Z – столбец координат вектора $z = \mathcal{A}y$ в базисе $\mathcal{E}$, таким образом, $Z = BY$. По правилу вычисления скалярного произведения в координатах в ортонормированном базисе (Утверждение 13.6 п.2) имеем

$$(x, \mathcal{A}y) = (x, z) = X^t Z = X^t BY.$$

Значит, $\beta(x, y) = \beta_{\mathcal{A}}(x, y)$ для любых $x, y \in V$, т.е. $\beta = \beta_{\mathcal{A}}$ для данного оператора $\mathcal{A}$.

Итак, нами доказана следующая теорема

Теорема 15.1. *Формула*

$$\beta_{\mathcal{A}}(x, y) = (x, \mathcal{A}y)$$

устанавливает в конечномерном евклидовом пространстве взаимно однозначное соответствие между билинейными функциями и линейными операторами.

Подчеркнем, что установленное соответствие не зависит от базиса.

И хотя в произвольном ортонормированном базисе матрицы линейного оператора и соответствующей ему билинейной функции совпадают, в не ортонормированном базисе они не обязаны совпадать.

ЗАДАЧА. Проверьте, что в конечномерном евклидовом пространстве V отображение $\mathcal{A} \mapsto \beta_{\mathcal{A}}$, определяемое формулой (36), является изоморфизмом пространства линейных операторов на пространство билинейных функций.

15.2. Операция перехода к сопряженному оператору.

Пусть V – конечномерное евклидово пространство.

Определение 15.1. *Пусть $\mathcal{A}$ – линейный оператор в V. Линейный оператор $\mathcal{A}^* : V \to V$ называется* **сопряженным** *к $\mathcal{A}$, если для любых $x, y \in V$*

$$(\mathcal{A}x, y) = (x, \mathcal{A}^* y).$$

Теорема 15.2. *В конечномерном евклидовом пространстве V каждому линейному оператору $\mathcal{A}$ отвечает сопряженный оператор и притом только один.*

Доказательство. Пусть $\mathcal{E} = \{e_1, \ldots, e_n\}$ – произвольный ортонормированный базис в V. И пусть A – матрица линейного оператора $\mathcal{A}$ в этом базисе. Докажем, что линейный оператор $\mathcal{C}$ с матрицей A^t в базисе $\mathcal{E}$ является сопряженным к $\mathcal{A}$. Надо проверить, что $(\mathcal{A}x, y) = (x, \mathcal{C}y)$ для любых $x, y \in V$. В самом деле, используя матричную форму записи скалярного произведения в ортонормированном базисе (Утверждение 13.6 п.2), имеем

$$(\mathcal{A}x, y) = (AX)^t Y = X^t A^t Y = X^t (A^t Y) = (x, \mathcal{C}y),$$

где X, Y – столбцы координат произвольных векторов $x, y \in V$ в базисе $\mathcal{E}$.

Осталось проверить единственность сопряженного оператора. Допустим, существуют два различных линейных оператора $\mathcal{A}_1^*$ и $\mathcal{A}_2^*$, сопряженных к $\mathcal{A}$. Тогда для любых $x, y \in V$

$$(x, \mathcal{A}_1^* y) = (\mathcal{A}x, y) = (x, \mathcal{A}_2^* y).$$

Значит, $\mathcal{A}_1^*$ и $\mathcal{A}_2^*$ определяют одну и ту же билинейную функцию по правилу (36), что противоречит теореме 15.1. ∎

В частности, мы доказали

Утверждение 15.1. *В любом ортонормированном базисе матрица линейного оператора $\mathcal{A}^*$, сопряженного к линейному оператору $\mathcal{A}$, является транспонированной к матрице оператора $\mathcal{A}$.*

Заметим, что в не ортонормированном базисе связь между матрицами операторов $\mathcal{A}$ и $\mathcal{A}^*$ более сложна.

Получим связь операции перехода от $\mathcal{A}$ к $\mathcal{A}^*$ с операциями сложения и умножения линейных операторов.

Утверждение 15.2. *Пусть $\mathcal{A}$, $\mathcal{B}$ – произвольные линейные операторы в конечномерном евклидовом пространстве V, λ – произвольное вещественное число. Тогда верны следующие соотношения:*

1) $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$,

2) $(\mathcal{A}^*)^* = \mathcal{A}$,

3) $(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$,

4) $(\lambda\mathcal{A})^* = \lambda\mathcal{A}^*$,

5) $\mathcal{I}^* = \mathcal{I}$.

Доказательство.

1) $(\mathcal{A}\mathcal{B}x, y) = (\mathcal{B}x, \mathcal{A}^* y) = (x, \mathcal{B}^*\mathcal{A}^* y)$.

Но, с другой стороны, по определению $(\mathcal{A}\mathcal{B})^*$ имеем:

$$(\mathcal{A}\mathcal{B}x, y) = (x, (\mathcal{A}\mathcal{B})^* y).$$

Сравнивая первые части этих двух равенств и вспомнив, что линейный оператор однозначно определяется соответствующей билинейной функцией (теорема 15.1), получаем:
$$(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*.$$

2) Пользуясь определением сопряженного оператора и симметричностью скалярного произведения, получаем:
$$(x, (\mathcal{A}^*)^* y) = (\mathcal{A}^* x, y) = (y, \mathcal{A}^* x) = (\mathcal{A}y, x) = (x, \mathcal{A}y).$$
Далее, как и в п.1, используем теорему 15.1.

Пункты 3)-4) доказываются аналогично. Проверьте это! ■

15.3. Самосопряженные линейные операторы.

Пусть V – конечномерное евклидово пространство.

Определение 15.2. *Линейный оператор $\mathcal{A} : V \to V$ евклидова пространства V называется* **самосопряженным** *(или* **симметрическим***), если*
$$\mathcal{A} = \mathcal{A}^*,$$
т.е. для любых $x, y \in V$ выполняется
$$(\mathcal{A}x, y) = (x, \mathcal{A}y).$$

Утверждение 15.3. *Линейный оператор $\mathcal{A} : V \to V$ евклидова пространства V является самосопряженным тогда и только тогда, когда его матрица в произвольном ортонормированном базисе является симметрической.*

Доказательство. Пусть $\mathcal{E} = \{e_1, \ldots, e_n\}$ – произвольный ортонормированный базис в V, а $A = (a_{i,j})$ – матрица линейного оператора $\mathcal{A}$ в этом базисе. По определению матрицы линейного оператора и правилу вычисления скалярного произведения в координатах в ортонормированном базисе для всех i, j имеем
$$(\mathcal{A}e_i, e_j) = (\sum_{k=1}^{n} a_{k,i} e_k, e_j) = a_{j,i}$$
и
$$(e_i, \mathcal{A}e_j) = (e_i, \sum_{k=1}^{n} a_{k,j} e_k) = a_{i,j}.$$

Отсюда получаем, что если $\mathcal{A}$ – самосопряженный оператор, то первые части этих двух равенств равны. Тогда и последние части этих двух равенств равны: $a_{i,j} = a_{j,i}$ для всех i, j, что означает, что матрица A является симметрической.

Обратно, если матрица A является симметрической, то $a_{i,j} = a_{j,i}$ для всех i, j, т.е. последние части этих двух равенств равны, откуда следует, что и первые части этих двух равенств равны: $(\mathcal{A}e_i, e_j) = (e_i, \mathcal{A}e_j)$ для всех i, j.

В силу линейности скалярного произведения по каждому аргументу и по определению линейного оператора отсюда следует, что и $(\mathcal{A}x, y) = (x, \mathcal{A}y)$ для любых векторов $x, y \in V$, т.е. $\mathcal{A}$ – самосопряженный оператор. ■

ЗАДАЧА 1. Доказать, что линейный оператор $\mathcal{A}$ является кососимметрическим (т.е. $\mathcal{A}^* = -\mathcal{A}$) тогда и только тогда, когда в произвольном ортонормированном базисе его матрица кососимметрична (т.е. $A^t = -A$).

Далее нашей целью будет приведение матрицы самосопряженного оператора к наиболее простому виду. Для этого нам понадобится следующая лемма.

Лемма 15.1. *Пусть V – евклидово пространство, $\mathcal{A} : V \to V$ – самосопряженный линейный оператор пространства V, U – подпространство в V, инвариантное относительно $\mathcal{A}$, $U^\perp$ – ортогональное дополнение к U. Тогда $U^\perp$ является подпространством, инвариантным относительно $\mathcal{A}$.*

Доказательство. Пусть w – произвольный вектор из $U^\perp$. Требуется, доказать, что $\mathcal{A}w \in U^\perp$.

Пусть u – произвольный вектор из U. Тогда по определению сопряженного линейного оператора имеем $(\mathcal{A}w, u) = (w, \mathcal{A}u)$. В свою очередь, $(w, \mathcal{A}u) = 0$, так как $\mathcal{A}u \in U$ в силу инвариантности подпространства U относительно $\mathcal{A}$, а $w \in U^\perp$. Таким образом, $(\mathcal{A}w, u) = 0$ для любого $u \in U$, что и означает, $\mathcal{A}w \in U^\perp$. ■

С помощью этой леммы получим **канонический вид** для матрицы самосопряженного линейного оператора.

Теорема 15.3. *Для любого самосопряженного линейного оператора $\mathcal{A}$ конечномерного евклидова пространства V найдется ортонормированный базис, в котором матрица этого оператора диагональна.*

Доказательство. Прежде всего заметим, что утверждение теоремы эквивалентно тому, что для самосопряженного линейного оператора $\mathcal{A}$ конечномерного евклидова пространства V найдется ортонормированный базис из собственных векторов линейного оператора $\mathcal{A}$. Существование такого базиса будем доказывать индукцией по $n = \dim V$.

Случай, когда $n = 1$, тривиален. Пусть далее $n > 1$.

В силу леммы 15.1 достаточно доказать существование хотя бы одного собственного вектора оператора $\mathcal{A}$, т.е. хотя бы одного одномерного подпространства, инвариантного относительно $\mathcal{A}$. Действительно, если U – одномерное инвариантное подпространство, порожденное вектором u, то по лемме 15.1 подпространство $U^\perp$ также инвариантно относительно $\mathcal{A}$. Очевидно, что ограничение самосопряженного линейного оператора на инвариантное подпространство – снова самосопряженный линейный оператор. Кроме того, размерность подпространства $U^\perp$ меньше n (по теореме 14.1). Поэтому к

$U^\perp$ можно применить предположение индукции и найти в $U^\perp$ ортонормированный базис $\mathcal{E}_{U^\perp} = \{e_2, \dots, e_n\}$ из собственных векторов ограничения $\mathcal{A}|_{U^\perp}$ линейного оператора $\mathcal{A}$ на подпространство $U^\perp$. В силу равенства $V = U \oplus U^\perp$ (см. теорему 14.1) система векторов $\{e_1 = \frac{1}{|u|}u, e_2, \dots, e_n\}$ и будет искомым ортонормированным базисом всего пространства V.

Итак, докажем, что найдется хотя бы одно одномерное подпространство, инвариантного относительно $\mathcal{A}$. Согласно теореме 5.2, $\mathcal{A}$, как линейный оператор над $\mathbb{R}$, имеет одномерное или двумерное инвариантное подпространство W. Если $\dim W = 2$, рассмотрим $\mathcal{A}|_W$ ограничение оператора $\mathcal{A}$ на подпространство W. Пусть A – матрица оператора $\mathcal{A}|_W$ в некотором ортонормированном базисе подпространства W. Так как $\mathcal{A}|_W$ – самосопряженный оператор, по утверждению 15.3 матрица A симметрична:

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Тогда характеристический многочлен оператора $\mathcal{A}|_W$ равен

$$\chi_{\mathcal{A}|_W}(t) = \det\,(A - tE) = \begin{vmatrix} a-t & b \\ b & c-t \end{vmatrix} = (a-t)(c-t) - b^2 = t^2 - (a+c)t + ac - b^2.$$

Дискриминант этого квадратного многочлена

$$D = (a+c)^2 - 4(ac - b^2) = (a-c)^2 + 4b^2$$

всегда неотрицателен, так что $\chi_{\mathcal{A}|_W}(t)$ имеет вещественный корень и, значит, $\mathcal{A}$ имеет собственный вектор v. Линейная оболочка $\langle v \rangle$ этого вектора и будет одномерным инвариантным подпространством для $\mathcal{A}$. ∎

ЗАДАЧА 2. Доказать, что для любого кососимметрического линейного оператора $\mathcal{A}$ (т.е. оператора, для которого $\mathcal{A}^* = -\mathcal{A}$) существует ортонормированный базис, в котором его матрица является блочно-диагональной с блоками порядка 1 или 2

$$A = \begin{pmatrix} H(a_1) & & & & & \\ & \ddots & & & & \\ & & H(a_k) & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

где блоки порядка один – это (0), а блоки порядка два имеют вид

$$H(a) = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}.$$

Следствие 15.1. *Для любого самосопряженного линейного оператора $\mathcal{A}$ конечномерного евклидова пространства V выполняется:*

1) Характеристический многочлен $\chi_{\mathcal{A}}(t)$ оператора $\mathcal{A}$ разлагается на линейные множители (над $\mathbb{R}$), т.е. все собственные числа самосопряженного оператора $\mathcal{A}$ вещественны.

2) Размерность каждого собственного подпространства V_λ равна кратности соответствующего корня λ характеристического многочлена $\chi_{\mathcal{A}}(t)$.

3) Собственные подпространства, отвечающие различным собственным значениям, ортогональны друг другу.

Доказательство. Первые два утверждения немедленно следуют из предыдущей теоремы и теоремы 6.1.

Для вывода последнего утверждения из предыдущей теоремы достаточно заметить, что если $\{e_1, \ldots, e_n\}$ – ортонормированный базис из собственных векторов оператора $\mathcal{A}$, причем $\mathcal{A}e_i = \lambda_i e_i$, то V_λ есть линейная оболочка тех e_i, для которых $\lambda_i = \lambda$. Впрочем, это утверждение легко можно доказать и непосредственно. В самом деле, пусть $x \in V_\lambda$, $y \in V_\mu$, $\lambda \neq \mu$. Тогда имеем

$$(\mathcal{A}x, y) = (\lambda x, y) = \lambda(x, y)$$

и

$$(x, \mathcal{A}y) = (x, \mu y) = \mu(x, y).$$

По определению самосопряженного оператора, первые части этих двух равенств равны, следовательно, равны и их последние части: $\lambda(x, y) = \mu(x, y)$. Откуда получаем, что $(\lambda - \mu)(x, y) = 0$. Так как $\lambda - \mu \neq 0$, сокращаем на $(\lambda - \mu)$ и получаем, что $(x, y) = 0$. ■

Следствие 15.2. *Для любой симметрической матрицы A найдется ортогональная матрица T и диагональная матрица D такие, что*

$$A = TDT^t.$$

Доказательство. Рассмотрим матрицу A как матрицу некоторого самосопряженного оператора $\mathcal{A}$ евклидова пространства в ортонормированном базисе $\mathcal{E}_1$. По теореме 15.3 для оператора $\mathcal{A}$ найдется ортонормированный базис $\mathcal{E}_2$, в котором его матрица D диагональна. Тогда из формулы (10) получаем

$$A = TDT^{-1},$$

где T – матрица перехода от базиса $\mathcal{E}_1$ к базису $\mathcal{E}_2$. Теперь осталось только заметить, что по утверждению 13.8 матрица T, как матрица перехода от одного ортонормированного базиса к другому ортонормированному базису, ортогональна и, следовательно, $T^{-1} = T^t$. ■

Заметим, что для теоремы 15.3 справедлива обратная теорема.

Утверждение 15.4. *Если для линейного оператора в конечномерном евклидовом пространстве найдется ортонормированный базис, в котором его матрица диагональна, то этот оператор является самосопряженным.*

Действительно, так как диагональная матрица симметрична, то по утверждению 15.3 оператор будет самосопряженным.

Замечание 15.1. *Из теоремы 15.3 следует наглядно-геометрический смысл действия произвольного самосопряженного оператора: в n-мерном евклидовом пространстве выделяются n попарно ортогональных направлений (соответствующих направлениям собственных векторов, образующих ортонормированный базис). Каждому из этих направлений ставится в соответствие вещественное число (собственное значение λ_i). Если $\lambda_i > 0$, то по соответствующему направлению производится растяжение (сжатие) пространства в λ_i раз. Если $\lambda_i < 0$, то по соответствующему направлению производится растяжение (сжатие) пространства в $|\lambda_i|$ раз и симметрия относительно подпространства, ортогонального данному направлению. Нулевому собственному значению соответствует уже не сжатие, а проектирование на подпространство, ортогональное данному направлению.*

Утверждение 15.5. *Для любого конечномерного евклидова пространства верно следующее.*

1) Сумма самосопряженных линейных операторов, а также произведение самосопряженного оператора на вещественное число является снова самосопряженным оператором.

2) Композиция двух самосопряженных операторов является самосопряженным оператором тогда и только тогда, когда эти операторы перестановочны.

Доказательство. Используя свойства операции сопряжения (см. утверждение 15.2), получаем следующее.

1) Если $\mathcal{A}$ и $\mathcal{B}$ – самосопряженные операторы, $\alpha \in \mathbb{R}$, то

$$(\mathcal{A}+\mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^* = \mathcal{A}+\mathcal{B},$$

$$(\alpha\mathcal{A})^* = \alpha\mathcal{A}^* = \alpha\mathcal{A}.$$

2) Из $\mathcal{AB} = \mathcal{BA}$, $\mathcal{A}^* = \mathcal{A}$, $\mathcal{B}^* = \mathcal{B}$ следует

$$(\mathcal{AB})^* = \mathcal{B}^*\mathcal{A}^* = \mathcal{BA} = \mathcal{AB}.$$

Обратно, если $(\mathcal{AB})^* = \mathcal{AB}$, $\mathcal{A}^* = \mathcal{A}$, $\mathcal{B}^* = \mathcal{B}$, то

$$\mathcal{AB} = (\mathcal{AB})^* = \mathcal{B}^*\mathcal{A}^* = \mathcal{BA}.$$

■

15.4. Приведение к главным осям.

Уже знаем, что формула

$$\beta_{\mathcal{A}}(x, y) = (x, \mathcal{A}y)$$

устанавливает в конечномерном евклидовом пространстве взаимно однозначное соответствие между билинейными функциями и линейными операторами (теорема 15.1).

Покажем, что *для того, чтобы линейный оператор $\mathcal{A}$ был самосопряженным, необходимо и достаточно, чтобы билинейная функция $\beta_{\mathcal{A}}$ была симметрической.*

В самом деле, симметричность билинейной функции $\beta_{\mathcal{A}}$ означает, что

$$\beta_{\mathcal{A}}(x, y) = \beta_{\mathcal{A}}(y, x),$$

что эквивалентно равенству

$$(x, \mathcal{A}y) = (y, \mathcal{A}x),$$

которое, в свою очередь, эквивалентно определению самосопряженного линейного оператора:

$$(x, \mathcal{A}y) = (\mathcal{A}x, y).$$

Из теоремы 15.3 вытекает важная

Теорема 15.4. *Пусть V – n-мерное евклидово пространство, $\beta : V \times V \to \mathbb{R}$ – симметрическая билинейная функция. Тогда в V существует ортонормированный базис $\mathcal{E}$, в котором квадратичная функция $q : V \to \mathbb{R}$, соответствующая β, записывается в виде суммы квадратов:*

$$q(x) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2, \tag{37}$$

где $\lambda_i \in \mathbb{R}$, а x_i – координаты вектора x в базисе $\mathcal{E}$.

Доказательство. Как было показано выше, если β симметрическая билинейная функция, то по теореме 15.1 существует такой самосопряженный линейный оператор $\mathcal{A}$, что

$$\beta(x, y) = (x, \mathcal{A}y).$$

Выберем в V в качестве векторов ортонормированного базиса систему $\{e_1, \ldots, e_n\}$ попарно ортогональных собственных векторов самосопряженного оператора $\mathcal{A}$ (это возможно в силу теоремы 15.3). Тогда

$$\mathcal{A}e_1 = \lambda_1 e_1,\ \mathcal{A}e_2 = \lambda_2 e_2, \ldots, \mathcal{A}e_n = \lambda_n e_n.$$

Пусть $x = \sum_{i=1}^{n} x_i e_i, y = \sum_{j=1}^{n} y_j e_j$ – произвольные векторы из V. Так как $(e_i, e_j) = \delta_i^j$, то

$$\beta(x,y) = (x, \mathcal{A}y) = (\sum_{i=1}^{n} x_i e_i, \mathcal{A}(\sum_{j=1}^{n} y_j e_j)) = (\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j \mathcal{A} e_j) =$$

$$= (\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j \lambda_j e_j) = \lambda_1 x_1 y_1 + \lambda_2 x_2 y_2 + \ldots + \lambda_n x_n y_n.$$

В частности,

$$q(x) = \beta(x,x) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2.$$

■

Нахождение в евклидовом пространстве ортонормированного базиса, в котором данная квадратичная формы записывается в виде суммы квадратов, называется **приведением этой формы к главным осям**.

Отметим, что числа $\lambda_1, \ldots, \lambda_n$ в выражении (37) определены однозначно с точностью до перестановки, так как это собственные значения соответствующего самосопряженного оператора.

15.5. Одновременное приведение пары квадратичных форм к сумме квадратов.

Следующая теорема является по существу другой формулировкой теоремы 15.4.

Теорема 15.5. *Пусть V – произвольное n-мерное векторное пространство над полем действительных чисел. И пусть в V заданы две квадратичные функции $q(x)$ и $h(x)$, причем $h(x)$ положительно определена. Тогда существует базис, в котором обе эти функции записываются в виде суммы квадратов.*

Доказательство. Пусть $\beta(x,y)$ – симметрическая билинейная функция, соответствующая квадратичной функции $h(x)$. Введем в V скалярное произведение, положив

$$(x,y) = \beta(x,y).$$

Это законно, так как β является положительно определенной симметрической билинейной функцией. Пространство V станет, таким образом, евклидовым. По теореме 15.4 в V существует ортонормированный (в смысле определенного нами скалярного произведения) базис $\mathcal{E}$, в котором квадратичная функция $q(x)$ приводится к сумме квадратов, т.е. к виду

$$q(x) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2, \tag{38}$$

где x_i – координаты вектора x в базисе $\mathcal{E}$. В то же время скалярное произведение в ортонормированном базисе $\mathcal{E}$ имеет вид

$$(x, y) = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n,$$

т.е. для $h(x) = \beta(x, x) = (x, x)$ выполняется:

$$h(x) = x_1^2 + x_2^2 + \ldots + x_n^2. \tag{39}$$

Таким образом, мы нашли базис, в котором обе квадратичные функции $q(x)$ и $h(x)$ одновременно записываются в виде суммы квадратов, что и требовалось. ■

В теореме 15.5 доказано, что в V существует базис, в котором квадратичные функции $q(x)$ и $h(x)$ имеют вид (38) и (39). Покажем, как найти числа $\lambda_1, \lambda_2, \ldots, \lambda_n$.

В базисе $\mathcal{E}$ из доказательства теоремы 15.5 матрицы квадратичных функций $q(x)$ и $h(x)$ имеют диагональный вид

$$Q = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \; H = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}.$$

Следовательно,

$$\det\,(Q - tH) = (\lambda_1 - t)(\lambda_2 - t)\ldots(\lambda_n - t).$$

Пусть $\widetilde{Q}$ и $\widetilde{H}$ – матрицы квадратичных функций $q(x)$ и $h(x)$ в произвольном базисе $\widetilde{\mathcal{E}}$. Тогда по формуле (20) имеем

$$\widetilde{Q} = T^t Q T, \;\; \widetilde{H} = T^t H T,$$

где T – матрица перехода от базиса $\mathcal{E}$ к базису $\widetilde{\mathcal{E}}$. Поэтому

$$\det\,(\widetilde{Q} - t\widetilde{H}) = \det\,T^t \det\,(Q - tH) \det\,T = (\det\,T^2)(\lambda_1 - t)(\lambda_2 - t)\ldots(\lambda_n - t).$$

Отсюда следует, что числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ являются корнями уравнения

$$\det\,(\widetilde{Q} - t\widetilde{H}) = 0,$$

где $\widetilde{Q}$ и $\widetilde{H}$ – матрицы функций $q(x)$ и $h(x)$ в каком-нибудь базисе $\widetilde{\mathcal{E}}$.

Заметим, что требование положительной определенности одной из функций является существенным в теореме 15.5. В качестве примера, показывающего это, рассмотрим следующие квадратичные функции

$$q(x) = x_1^2 - x_2^2, \;\; h(x) = 2x_1 x_2,$$

из которых ни одна не является положительно определенной. Функции $q(x)$ соответствует матрица

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

а $h(x)$ – матрица

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

Так как определитель

$$\det\,(Q - tH) = \begin{vmatrix} 1 & -t \\ -t & -1 \end{vmatrix} = -(t^2 + 1)$$

не имеет вещественных корней, то, согласно сказанному выше, обе функции не могут быть приведены одновременно к сумме квадратов.

16. Линейные операторы в евклидовом пространстве (продолжение).

Кроме самосопряженных операторов, со скалярным произведением также естественно связаны ортогональные операторы.

16.1. Ортогональные операторы.

Пусть V – конечномерное евклидово пространство.

Определение 16.1. *Линейный оператор $\mathcal{A} : V \to V$ евклидова пространства V называется* **ортогональным***, если он сохраняет значение скалярного произведения, т.е. для любых $x, y \in V$ выполняется*

$$(\mathcal{A}x, \mathcal{A}y) = (x, y).$$

Отметим, что равенство $(\mathcal{A}x, \mathcal{A}y) = (x, y)$ для любых $x, y \in V$ эквивалентно тому, что $\mathcal{A}^*\mathcal{A} = \mathcal{I}$. Действительно, из равенства $(\mathcal{A}x, \mathcal{A}y) = (x, y)$ для любых $x, y \in V$ имеем $(\mathcal{A}^*\mathcal{A}x, y) = (x, y)$, или $(\mathcal{A}^*\mathcal{A}x - x, y) = 0$. Это означает, что вектор $\mathcal{A}^*\mathcal{A}x-x$ ортогонален любому вектору пространства и, следовательно, является нулевым. Поскольку равенство $\mathcal{A}^*\mathcal{A}x = x$ выполняется для всех x, преобразование $\mathcal{A}^*\mathcal{A}$ является тождественным. Обратно, из равенства $\mathcal{A}^*\mathcal{A} = \mathcal{I}$ получаем

$$(\mathcal{A}x, \mathcal{A}y) = (\mathcal{A}^*\mathcal{A}x, y) = (\mathcal{I}x, y) = (x, y).$$

В силу свойств операции сопряжения имеем, что равенство $\mathcal{A}^*\mathcal{A} = \mathcal{I}$ равносильно равенству $\mathcal{A}\mathcal{A}^* = \mathcal{I}$. Следовательно, линейный оператор $\mathcal{A}$ ортогонален тогда и только тогда, когда $\mathcal{A}^* = \mathcal{A}^{-1}$. В частности, ортогональный линейный оператор является невырожденным.

Утверждение 16.1. *Линейный оператор $\mathcal{A} : V \to V$ конечномерного евклидова пространства V является ортогональным тогда и только тогда, когда его матрица в произвольном ортонормированном базисе является ортогональной.*

Доказательство. Пусть $\mathcal{E} = \{e_1, \ldots, e_n\}$ – произвольный ортонормированный базис в V, а $A = (a_{i,j})$ – матрица линейного оператора $\mathcal{A}$ в этом базисе. По определению матрицы линейного оператора и правилу вычисления скалярного произведения в координатах в ортонормированном базисе для всех i, j имеем два равенства

$$\begin{aligned} (\mathcal{A}e_i, \mathcal{A}e_j) = (\textstyle\sum_{k=1}^n a_{k,i}e_k, \sum_{l=1}^n a_{l,j}e_l) &= \textstyle\sum_{k=1}^n a_{k,i}a_{k,j}, \\ (e_i, e_j) &= \delta_i^j. \end{aligned} \tag{40}$$

Если оператор $\mathcal{A}$ ортогонален, то первые части равенств (40) равны, значит, равны и последние части этих равенств: $\sum_{k=1}^n a_{k,i}a_{k,i} = \delta_i^j$ для всех i, j, т.е. $A^tA = E$, что означает, что матрица A ортогональна.

Обратно, если матрица A является ортогональной, то $A^tA = E$, что в координатной форме означает: $\sum_{k=1}^n a_{k,i}a_{k,i} = \delta_i^j$ для всех i, j, т.е. последние части в равенствах (40) равны, откуда следует, что и первые части этих двух равенств равны: $(\mathcal{A}e_i, \mathcal{A}e_j) = (e_i, e_j)$ для всех i, j. В силу линейности скалярного произведения по каждому аргументу и по определению линейного оператора отсюда следует, что и $(\mathcal{A}x, \mathcal{A}y) = (x, y)$ для любых векторов $x, y \in V$, т.е. $\mathcal{A}$ – ортогональный оператор. ■

Утверждение 16.2. *Линейный оператор ортогонален тогда и только тогда, когда он сохраняет длины векторов.*

В частности, ортогональный оператор сохраняет углы между векторами.

Доказательство. Если оператор $\mathcal{A}$ ортогонален, то

$$|\mathcal{A}x| = \sqrt{(\mathcal{A}x, \mathcal{A}x)} = \sqrt{(x, x)} = |x|.$$

Обратное утверждение следует из тождества

$$(x, y) = \frac{1}{2}(|x + y|^2 - |x|^2 - |y|^2).$$

Действительно, если оператор $\mathcal{A}$ сохраняет длины векторов, то

$$(\mathcal{A}x, \mathcal{A}y) = \frac{1}{2}(|\mathcal{A}x + \mathcal{A}y|^2 - |\mathcal{A}x|^2 - |\mathcal{A}y|^2) =$$

$$= \frac{1}{2}(|\mathcal{A}(x + y)|^2 - |\mathcal{A}x|^2 - |\mathcal{A}y|^2) = \frac{1}{2}(|x + y|^2 - |x|^2 - |y|^2) = (x, y).$$

Для доказательства того, что ортогональный оператор сохраняет углы между векторами, достаточно заметить, что раз ортогональный оператор сохраняет скалярное произведение и длины векторов, то он сохраняет и косинус угла между векторами. ■

Утверждение 16.3. *Линейный оператор конечномерного евклидова пространства V ортогонален тогда и только тогда, когда он переводит ортонормированный базис в ортонормированный базис.*

Доказательство. Если оператор $\mathcal{A}$ ортогонален, то по предыдущему утверждению он сохраняет как длины векторов, так и углы между векторами, а значит, $\mathcal{A}$ переводит ортонормированный базис в ортонормированный базис.

Обратно, допустим, некоторый линейный оператор $\mathcal{A}$ в евклидовом пространстве V переводит ортонормированный базис $\{e_1, \ldots, e_n\}$ снова в ортонормированный базис $\{\mathcal{A}e_1, \ldots, \mathcal{A}e_n\}$. Рассмотрим в V произвольные векторы

$$x = \sum_{i=1}^n x_ie_i, \quad y = \sum_{j=1}^n y_je_j.$$

Имеем
$$(x, y) = (\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j) = \sum_{i=1}^{n} x_i y_i,$$
$$(\mathcal{A}x, \mathcal{A}y) = (\mathcal{A}(\sum_{i=1}^{n} x_i e_i), \mathcal{A}(\sum_{j=1}^{n} y_j e_j)) = (\sum_{i=1}^{n} x_i \mathcal{A}e_i, \sum_{j=1}^{n} y_j \mathcal{A}e_j) = \sum_{i=1}^{n} x_i y_i,$$
т.е. $(x, y) = (\mathcal{A}x, \mathcal{A}y)$. Следовательно, оператор $\mathcal{A}$ ортогонален. ■

Утверждение 16.4. *Собственные значения ортогонального линейного оператора равны* ± 1.

Доказательство. Пусть x – собственный вектор ортогонального оператора $\mathcal{A}$ с собственным значением λ. Тогда
$$(\mathcal{A}x, \mathcal{A}x) = (\lambda x, \lambda x) = \lambda^2 (x, x)$$
и
$$(\mathcal{A}x, \mathcal{A}x) = (x, x).$$
Отсюда $\lambda^2(x, x) = (x, x)$. Так как $x \neq 0$, то $(x, x) \neq 0$ в силу положительной определенности скалярного произведения. Сокращая на число (x, x), получаем $\lambda^2 = 1$. Следовательно, $\lambda = \pm 1$. ■

Далее нашей целью будет установить простейший вид, который может иметь матрица ортогонального оператора в евклидовом пространстве. Для этого мы будем использовать следующую лемму, аналогичную лемме 15.1 о самосопряженном операторе.

Лемма 16.1. *Пусть V – конечномерное евклидово пространство, $\mathcal{A} : V \to V$ – ортогональный линейный оператор пространства V, U – подпространство в V, инвариантное относительно $\mathcal{A}$, $U^\perp$ – ортогональное дополнение к U. Тогда $U^\perp$ является подпространством, инвариантным относительно $\mathcal{A}$.*

Доказательство. Пусть w – произвольный вектор из $U^\perp$. Требуется, доказать, что $\mathcal{A}w \in U^\perp$.

Пусть $\{e_1, \ldots, e_k\}$ – какой-нибудь ортонормированный базис в пространстве U. Так как U – инвариантное относительно $\mathcal{A}$ подпространство, векторы $f_1 = \mathcal{A}e_1, \ldots, f_k = \mathcal{A}e_k$ лежат в U, а в силу утверждения 16.2 эти векторы взаимно ортогональны и имеют единичную длину. Кроме того, число векторов $f_1, \ldots, f_k$ равно k, т.е. размерности подпространства U. Следовательно, $f_1, \ldots, f_k$ образуют ортонормированный базис в U.

Пусть $u = \sum_{i=1}^k u_i f_i$ – произвольный вектор из U. Тогда вектор $\tilde{u} = \sum_{i=1}^k u_i e_i$ также лежит в U и $\mathcal{A}\tilde{u} = u$.

Отсюда по определению ортогонального линейного оператора имеем
$$(\mathcal{A}w, u) = (\mathcal{A}w, \mathcal{A}\tilde{u}) = (w, \tilde{u}).$$

В свою очередь, $(w, \tilde{u}) = 0$, так как $\tilde{u} \in U$, а $w \in U^{\perp}$. Следовательно, $(\mathcal{A}w, u) = 0$ для любого $u \in U$, что и означает, $\mathcal{A}w \in U^{\perp}$. ■

С помощью этой леммы получим **канонический вид** для матрицы ортогонального линейного оператора.

Теорема 16.1. *Для любого ортогонального линейного оператора $\mathcal{A}$ конечномерного евклидова пространства V найдется ортонормированный базис, в котором матрица этого оператора имеет блочно-диагональный вид*

$$A = \begin{pmatrix} \Pi(\alpha_1) & & & & & & & & \\ & \ddots & & & & & & & \\ & & \Pi(\alpha_s) & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix},$$

где $\Pi(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$.

Заметим, что, используя матрицы $\Pi(\pi) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ и $\Pi(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, можно при желании оставить не более одного свободного диагонального элемента, равного -1, и не более одного, равного 1.

Доказательство. Доказательство будем вести индукцией по размерности n евклидова пространства V.

Пусть $n = 1$ и e – произвольный единичный вектор пространства V. Тогда $\{e\}$ образует ортонормированный базис в V и $\mathcal{A}e = \lambda e$, причем $\lambda = \pm 1$ по утверждению 16.4. Таким образом матрица оператора $\mathcal{A}$ в этом ортонормированном базисе имеет указанный вид.

Пусть $n = 2$ и $\{e_1, e_2\}$ – произвольный ортонормированный базис пространства V. Тогда $\{\mathcal{A}e_1, \mathcal{A}e_2\}$ – также ортонормированный базис по утверждению 16.3. Пусть вектор $\mathcal{A}e_1$ получается из вектора e_1 поворотом на угол α. Для вектора $\mathcal{A}e_2$ возможны два случая: либо он получается из e_2 также поворотом на угол α, либо он противоположен получаемому таким образом вектору.

В первом случае оператор $\mathcal{A}$ есть поворот на угол α, и его матрица в базисе $\{e_1, e_2\}$ имеет вид

$$\Pi(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Во втором случае оператор $\mathcal{A}$ есть отражение относительно биссектрисы угла между e_1 и $\mathcal{A}e_1$, и его матрица в подходящем ортонормированном базисе имеет вид

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Итак, при $n \leqslant 2$ теорема доказана.

Предположим она доказана для всех евклидовых пространств размерности меньше n, и докажем ее для евклидова пространства V размерности n.

Согласно теореме 5.2, $\mathcal{A}$, как линейный оператор над $\mathbb{R}$, имеет одномерное или двумерное инвариантное подпространство W. По лемме 16.1 подпространство $W^{\perp}$ также инвариантно относительно $\mathcal{A}$. Так как ограничение оператора $\mathcal{A}$ на инвариантное подпространство $W^{\perp}$ продолжает быть ортогональным, и размерность подпространства $W^{\perp}$ меньше n (по теореме 14.1), то к $W^{\perp}$ можно применить предположение индукции и найти в $W^{\perp}$ ортонормированный базис $\mathcal{E}_{W^{\perp}}$, в котором матрица $A_{W^{\perp}}$ ограничения линейного оператора $\mathcal{A}$ на подпространство $W^{\perp}$ имеет канонический вид.

Если W одномерно, то в W существует собственный вектор e_1 длины 1 с собственным значением ± 1. Тогда $\{e_1\} \sqcup \mathcal{E}_{W^{\perp}}$ – ортонормированный базис пространства V. По утверждению 4.5 в этом базисе матрица A оператора $\mathcal{A}$ имеет вид

$$A = \begin{pmatrix} \pm 1 & 0 \\ 0 & A_{W^{\perp}} \end{pmatrix},$$

который является каноническим видом, или превращается в него после изменения нумерации векторов $e_1, e_2, \ldots, e_n$.

Если $\dim W = 2$, то в W существует ортонормированный базис $\{e_1, e_2\}$, в котором матрица A_W оператора $\mathcal{A}$ (рассматриваемого лишь в W) имеет вид

$$A_W = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \text{ или } A_W = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Тогда $\{e_1, e_2\} \sqcup \mathcal{E}_{W^{\perp}}$ – ортонормированный базис пространства V, в котором по утверждению 4.5 матрица A оператора $\mathcal{A}$ имеет вид

$$A = \begin{pmatrix} A_W & 0 \\ 0 & A_{W^{\perp}} \end{pmatrix},$$

который является каноническим видом, или превращается в него после изменения нумерации векторов $e_1, e_2, \ldots, e_n$. ■

Такое представление матрицы ортогонального оператора, иногда, называют *разложение оператора на плоские вращения*, так как каждому двумерному подпространству соответствует поворот, и эти повороты могут осуществляться последовательно. Надо однако помнить, что в общем случае имеются собственные подпространства с собственными значениями 1 и -1.

Для трехмерного евклидова пространства теорема 16.1 означает, что матрица любого ортогонального оператора $\mathcal{A}$ в подходящем ортонормированном базисе имеет один из двух видов:

$$\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

В первом случае оператор $\mathcal{A}$ представляет собой поворот на угол α вокруг некоторой оси, во втором – зеркальный поворот, т.е. поворот, совмещенный с отражением относительно плоскости, ортогональной оси поворота.

Ясно, что зеркальный поворот не может быть результатом непрерывного движения, так как он изменяет ориентацию пространства. Следовательно, конечный результат сколь угодно сложного реального движения твердого тела с закрепленной точкой – такой же, как при простом повороте вокруг подходящей оси на подходящий угол. Это совершенно не тривиальное утверждение называется *теоремой Эйлера.*

Множество невырожденных линейных операторов n-мерного пространства V образует группу, называемую **полной линейной группой** пространства V и обозначаемую $\mathrm{GL}(V)$.

Утверждение 16.5. *Ортогональные операторы в евклидовом пространстве V образуют подгруппу группы $\mathrm{GL}(V)$, называемую* **ортогональной группой** *и обозначаемую $\mathrm{O}(V)$.*

Доказательство. Как мы знаем, ортогональные операторы являются невырожденными. Проверим, что множество ортогональных операторов замкнуто относительно групповых операций.

Пусть $\mathcal{O}$ – ортогональный оператор, т.е. $\mathcal{O}\mathcal{O}^* = \mathcal{I}$. Тогда по свойствам операции сопряжения имеем

$$\mathcal{O}^{-1}(\mathcal{O}^{-1})^* = \mathcal{O}^*(\mathcal{O}^*)^* = \mathcal{O}^*\mathcal{O} = \mathcal{I},$$

откуда следует, что $\mathcal{O}^{-1}$ – ортогональный оператор.

Пусть $\mathcal{O}_1$ и $\mathcal{O}_2$ – ортогональные операторы, т.е. $\mathcal{O}_1\mathcal{O}_1^* = \mathcal{I}$ и $\mathcal{O}_2\mathcal{O}_2^* = \mathcal{I}$. Тогда по свойствам операции сопряжения имеем

$$\mathcal{O}_1\mathcal{O}_2(\mathcal{O}_1\mathcal{O}_2)^* = \mathcal{O}_1\mathcal{O}_2\mathcal{O}_2^*\mathcal{O}_1^* = \mathcal{O}_1\mathcal{I}\mathcal{O}_1^* = \mathcal{O}_1\mathcal{O}_1^* = \mathcal{I},$$

что означает ортогональность оператора $\mathcal{O}_1\mathcal{O}_2$.

Из равенств $\mathcal{I}\mathcal{I}^* = \mathcal{I}\mathcal{I} = \mathcal{I}$ следует, что тождественный оператор также является ортогональным. ■

Отметим, что группа $\mathrm{Gl}_n(\mathbb{R})$ изоморфна группе $\mathrm{GL}(V)$, а группа O_n изоморфна группе $\mathrm{O}(V)$.

16.2. Полярное разложение.

Определение 16.2. *Самосопряженный линейный оператор $\mathcal{C}$ в евклидовом пространстве V называется* **положительно определенным**, *если соответствующая ему симметрическая билинейная функция положительно определена, т.е. $(x, \mathcal{C}x) > 0$ для любого ненулевого вектора $x \in V$.*

Лемма 16.2. *Самосопряженный оператор $\mathcal{C}$ в конечномерном евклидовом пространстве V положительно определен тогда и только тогда, когда все собственные значения оператора $\mathcal{C}$ положительны.*

Доказательство. Пусть $\mathcal{C}$ положительно определен и e – собственный вектор, соответствующий произвольному собственному значению λ оператора $\mathcal{C}$. С одной стороны, по определению положительно определенного оператора имеем

$$(e, \mathcal{C}e) > 0.$$

С другой стороны,

$$(e, \mathcal{C}e) = (e, \lambda e) = \lambda(e, e).$$

Следовательно,

$$\lambda(e, e) > 0.$$

Так как $(e, e) > 0$ в силу положительной определенности скалярного произведения, отсюда получаем, что $\lambda > 0$.

Обратно, пусть все собственные значения оператора $\mathcal{C}$ положительны. По теореме 15.3 для $\mathcal{C}$ найдется ортонормированный базис $\{e_1, \ldots, e_n\}$ из собственных векторов. Пусть $\lambda_1, \ldots, \lambda_n$ – (положительные) собственные значения, соответствующие векторам $e_1, \ldots, e_n$. Тогда для любого ненулевого вектора $x = \sum_{i=1}^n x_i e_i$ пространства V будет выполняться

$$(x, \mathcal{C}x) = (\sum_{i=1}^n x_i e_i, \mathcal{C}(\sum_{i=1}^n x_i e_i)) = (\sum_{i=1}^n x_i e_i, \sum_{i=1}^n \lambda_i x_i e_i) = \sum_{i=1}^n \lambda_i x_i^2 > 0.$$

Последнее верно, так как все $\lambda_i > 0$ и хотя бы одна координата x_i ненулевого вектора x отлична от нуля. ∎

В частности, если самосопряженный линейный оператор положительно определен, то он не вырожден, так как его матрица в ортонормированном базисе из собственных векторов диагональна с положительными элементами на диагонали, а значит, не вырождена.

Лемма 16.3. *Всякий положительно определенный самосопряженный оператор $\mathcal{B}$ в конечномерном евклидовом пространстве V единственным образом представляется в виде*

$$\mathcal{B} = \mathcal{C}^2,$$

где $\mathcal{C}$ – также положительно определенный самосопряженный оператор.

Доказательство. Пусть $\lambda_1, \ldots, \lambda_s$ – все различные собственные значения оператора $\mathcal{B}$ и $V_1, \ldots, V_s$ – соответствующие собственные подпространства. Пусть $\mathcal{E}_i$ – произвольный ортонормированный базис в V_i для каждого $i = \overline{1, n}$. Тогда $\mathcal{E} = \mathcal{E}_1 \sqcup \ldots \sqcup \mathcal{E}_s$ – ортонормированный базис всего пространства V, в котором матрица оператора $\mathcal{B}$ диагональна:

$$B = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_1 \end{matrix}} & & \\ & \ddots & \\ & & \boxed{\begin{matrix} \lambda_s & & \\ & \ddots & \\ & & \lambda_s \end{matrix}} \end{pmatrix}.$$

По лемме 16.2 $\lambda_i > 0$. Положим $\mu_i = \sqrt{\lambda_i}$ (числа $\sqrt{\lambda_i}$ выбираются неотрицательными). Тогда линейный оператор $\mathcal{C}$, действующий в V_i как умножение на μ_i, удовлетворяет условиям леммы. Действительно, в ортонормированном базисе $\mathcal{E}$ матрица оператора $\mathcal{C}$ диагональна:

$$C = \begin{pmatrix} \boxed{\begin{matrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_1} \end{matrix}} & & \\ & \ddots & \\ & & \boxed{\begin{matrix} \sqrt{\lambda_s} & & \\ & \ddots & \\ & & \sqrt{\lambda_s} \end{matrix}} \end{pmatrix}.$$

Поэтому оператор $\mathcal{C}$ самосопряжен (по утверждению 15.4). А так как все его собственные значения $\mu_i = \sqrt{\lambda_i}$ положительны, то оператор $\mathcal{C}$ положительно определен (по лемме 16.2). Осталось только заметить, что $\mathcal{B} = \mathcal{C}^2$, так как это равенство верно для их матриц в базисе $\mathcal{E}$.

Докажем, что оператор $\mathcal{C}$ определен однозначно. Допустим, существует еще один положительно определенный самосопряженный оператор $\widetilde{\mathcal{C}}$ такой, что $\mathcal{B} = \widetilde{\mathcal{C}}^2$. Пусть $\widetilde{\mu}_1, \ldots, \widetilde{\mu}_{\tilde{s}}$ – его различные собственные значения ($\widetilde{\mu}_i > 0$ по лемме 16.2), а $W_1, .., W_{\tilde{s}}$ – соответствующие собственные подпространства. Тогда оператор $\mathcal{C}^2 = \mathcal{B}$ действует на W_i как умножение на $\widetilde{\mu}_i^2$. Следовательно, $\tilde{s} = s$ и при подходящей нумерации $\widetilde{\mu}_i^2 = \lambda_i$ и $W_i = V_i$. Таким образом, оператор $\mathcal{C}$ и оператор $\widetilde{\mathcal{C}}$ действуют одинаково на каждом подпространстве V_i, а значит, и на всем пространстве V. Следовательно, $\mathcal{C} = \widetilde{\mathcal{C}}$. ∎

Теорема 16.2. *Всякий невырожденный линейный оператор $\mathcal{A}$ в конечномерном евклидовом пространстве единственным образом представляется в виде произведения*

$$\mathcal{A} = \mathcal{C}\mathcal{O},$$

где $\mathcal{C}$ – положительно определенный самосопряженный оператор, а $\mathcal{O}$ – ортогональный оператор.

Такое представление линейного оператора называется его **полярным разложением**.

Доказательство. Рассмотрим линейный оператор $\mathcal{A}\mathcal{A}^*$. Этот оператор является положительно определенным самосопряженным оператором. Действительно, из свойств операции сопряжения получаем

$$(\mathcal{A}\mathcal{A}^*)^* = (\mathcal{A}^*)^*\mathcal{A}^* = \mathcal{A}\mathcal{A}^*,$$

что доказывает самосопряженность оператора $\mathcal{A}\mathcal{A}^*$. Далее так как $\mathcal{A}$ не вырожден, то его матрица A в любом ортонормированном базисе не вырождена. Поэтому и матрица A^t оператора $\mathcal{A}^*$ в том же базисе не вырождена. А значит, оператор $\mathcal{A}^*$ тоже не вырожден. Следовательно, $\mathcal{A}^*x \neq 0$ для любого ненулевого вектора $x \in V$. Поэтому из положительной определенности скалярного произведения получаем

$$(x, \mathcal{A}\mathcal{A}^*x) = (\mathcal{A}^*x, \mathcal{A}^*x) > 0,$$

что означает положительную определенность самосопряженного оператора $\mathcal{A}\mathcal{A}^*$.

Пользуясь леммой 16.3, найдем положительно определенный самосопряженный оператор $\mathcal{C}$ такой, что $\mathcal{A}\mathcal{A}^* = \mathcal{C}^2$, и положим $\mathcal{O} = \mathcal{C}^{-1}\mathcal{A}$. Тогда $\mathcal{A} = \mathcal{C}\mathcal{O}$ и

$$\mathcal{A}\mathcal{A}^* = \mathcal{C}\mathcal{O}(\mathcal{C}\mathcal{O})^* = \mathcal{C}\mathcal{O}\mathcal{O}^*\mathcal{C}^* = \mathcal{C}\mathcal{O}\mathcal{O}^*\mathcal{C} = \mathcal{C}^2,$$

откуда после сокращения на $\mathcal{C}$ получаем, что $\mathcal{O}\mathcal{O}^* = \mathcal{I}$, т.е. $\mathcal{O}$ – ортогональный оператор.

Предположим, что существует другое разложение $\mathcal{A} = \widetilde{\mathcal{C}}\widetilde{\mathcal{O}}$, где $\widetilde{\mathcal{C}}$ – положительно определенный самосопряженный оператор, а $\widetilde{\mathcal{O}}$ – ортогональный оператор. Тогда

$$\mathcal{A}\mathcal{A}^* = \widetilde{\mathcal{C}}\widetilde{\mathcal{O}}\widetilde{\mathcal{O}}^*\widetilde{\mathcal{C}}^* = \widetilde{\mathcal{C}}^2.$$

Так как $\mathcal{A}\mathcal{A}^* = \mathcal{C}^2$, то ввиду леммы 16.3 получаем, что $\mathcal{C} = \widetilde{\mathcal{C}}$ и, следовательно, $\mathcal{O} = \mathcal{C}^{-1}\mathcal{A} = \widetilde{\mathcal{C}}^{-1}\mathcal{A} = \widetilde{\mathcal{O}}$. ∎

Аналогично можно получить, что всякий невырожденный линейный оператор $\mathcal{A}$ в конечномерном евклидовом пространстве единственным образом представляется в виде $\mathcal{A} = \mathcal{O}'\mathcal{C}'$, где $\mathcal{C}'$ – положительно определенный самосопряженный оператор, а $\mathcal{O}'$ – ортогональный оператор.

17. Унитарные пространства.

Рассмотрим векторное пространство V над полем комплексных чисел. Оказываются, что в этом пространстве естественные аксиомы, определяющие скалярное произведение в евклидовом пространстве, выполнены быть не могут. В самом деле, пусть x – ненулевой вектор из V. При этом ix тоже принадлежит V, где i – мнимая единица. Для билинейного скалярного произведения должно иметь место равенство

$$(ix, ix) = -(x, x),$$

которое противоречит положительной определенности скалярного произведения.

Таким образом билинейные функции не представляются удовлетворительными, если мы хотим развить теорию, аналогичную метрической теории в евклидовом пространстве, для случая пространств над полем комплексных чисел.

17.1. Полуторалинейные функции.

В векторном пространстве V над полем комплексных чисел чаще всего используются не билинейные, а полуторалинейные функции.

Определение 17.1. *Функция $\beta : V \times V \to \mathbb{C}$ называется* **полуторалинейной функцией**, *если*

1) β линейна по первому аргументу, т.е.

$$\beta(x + y, z) = \beta(x, z) + \beta(y, z) \ \ \forall x, y, z \in V,$$

$$\beta(\lambda x, y) = \lambda\beta(x, y) \ \ \forall \lambda \in \mathbb{C}, \forall x, y \in V;$$

2) β антилинейна по второму аргументу, т.е.

$$\beta(x, y + z) = \beta(x, y) + \beta(x, z) \ \ \forall x, y, z \in V,$$

$$\beta(x, \lambda y) = \overline{\lambda}\beta(x, y) \ \ \forall \lambda \in \mathbb{C}, \forall x, y \in V.$$

Замечание. *Иногда требуется, чтобы полуторалинейная функция была, наоборот, линейна по второму аргументу и антилинейна по первому.*

Теория полуторалиненых функций аналогична теории билинейных функций. Поэтому мы изложим ее кратко, останавливаясь более подробно лишь в тех местах, где имеется существенное различие.

Пусть $\mathcal{E} = \{e_1, \ldots, e_n\}$ – произвольный базис векторного пространства V. Полуторалинейная функция β определяется числами $b_{i,j} = \beta(e_i, e_j)$. А именно, если $x_1, \ldots, x_n$ и $y_1, \ldots, y_n$ – координаты векторов x и y в базисе $\mathcal{E}$, то:

$$\beta(x, y) = \sum_{i,j=1}^{n} b_{i,j} x_i \bar{y}_j. \tag{41}$$

Квадратная матрица $B = (b_{i,j})$ порядка n называется **матрицей полуторалинейной функции β в базисе $\mathcal{E}$**. В матричном виде равенство (41) запишется так

$$\beta(x,y) = X^t B\overline{Y}, \tag{42}$$

где $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$.

При замене базиса матрица полуторалинейной функции преобразуется по правилу

$$\widetilde{B} = T^t B\overline{T}, \tag{43}$$

где $\widetilde{B}$ – матрица полуторалинейной функции β в базисе $\widetilde{\mathcal{E}}$, а $T = T_{\mathcal{E}\to\widetilde{\mathcal{E}}}$ – матрица перехода от базиса $\mathcal{E}$ к базису $\widetilde{\mathcal{E}}$.

Определение 17.2. *Полуторалинейная функция β называется* **эрмитовой**, *если для любых векторов $x, y \in V$ выполняется*

$$\beta(x,y) = \overline{\beta(y,x)}.$$

Функция β является эрмитовой тогда и только тогда, когда ее матрица B в произвольном базисе удовлетворяла условию

$$B^t = \overline{B}.$$

Такие матрицы называются **эрмитовыми**.

Если $B = (b_{i,j})$ – эрмитова матрица, то ее элементы, симметричные относительно главной диагонали, комплексно сопряжены: $b_{i,j} = \overline{b}_{j,i}$, в частности, ее элементы на главной диагонали вещественны: $b_{i,i} = \overline{b}_{i,i}$. Также стоит отметить, что определитель эрмитовой матрицы всегда веществен, так как:

$$\det B = \det B^t = \det \overline{B} = \overline{\det B}.$$

Каждой эрмитовой полуторалинейной функции β соответствует **эрмитова квадратичная функция**

$$q(x) = \beta(x,x).$$

Утверждение 17.1. *Всякая эрмитова квадратичная функция q принимает только вещественные значения.*

Доказательство. Пусть β – эрмитова полуторалинейная функция, соответствующая q. Тогда для любого вектора $x \in V$ имеем

$$q(x) = \beta(x,x) = \overline{\beta(x,x)} = \overline{q(x)}.$$

Значит, $q(x) \in \mathbb{R}$. ∎

Утверждение 17.2. *Всякая эрмитова полуторалинейная функция β однозначно определяется своей эрмитовой квадратичной функцией q.*

Доказательство. Пусть x, y – произвольные векторы пространства V. Тогда

$$q(x+y) = \beta(x+y, x+y) = \beta(x,x) + \beta(y,y) + \beta(x,y) + \beta(y,x) =$$
$$= q(x) + q(y) + \beta(x,y) + \beta(y,x)$$

и

$$q(x+iy) = \beta(x+iy, x+iy) = \beta(x,x) + \beta(y,y) - i\beta(x,y) + i\beta(y,x) =$$
$$= q(x) + q(y) - i\beta(x,y) + i\beta(y,x).$$

Аналогично получаем, что

$$q(x-y) = q(x) + q(y) - \beta(x,y) - \beta(y,x)$$

и

$$q(x-iy) = q(x) + q(y) + i\beta(x,y) - i\beta(y,x).$$

Из полученных соотношений следует, что

$$\beta(x,y) = \frac{1}{4}(q(x+y) + iq(x+iy) - q(x-y) - iq(x-iy).$$

■

Как и в случае симметрической билинейной функции, доказывается аналог теоремы 10.1 для эрмитовой полуторалинейной функции и аналог следствия 11.1 для эрмитовой квадратичной функции, из которых аналогично вещественному случаю выводится следующая

Теорема 17.1. *В n-мерном векторном пространстве V над полем комплексных чисел для всякой эрмитовой полуторалинейной функции β существует базис, в котором функция β и соответствующая ей эрмитова квадратичная функция q имеют* **нормальный вид**

$$\beta(x,y) = x_1\bar{y}_1 + \ldots + x_k\bar{y}_k - x_{k+1}\bar{y}_{k+1} - \ldots - x_{k+l}\bar{y}_{k+l},$$
$$q(x) = |x_1|^2 + \ldots + |x_k|^2 - |x_{k+1}|^2 - \ldots - |x_{k+l}|^2. \qquad (44)$$

При этом для эрмитовых квадратичных форм справедлив **закон инерции**: числа k, l являются инвариантами, т.е. не зависят от базиса, в котором функция q имеет нормальный вид. Числа k и l называются **положительным** и **отрицательным индексами инерции** функции q.

Определение 17.3. *Эрмитова квадратичная функция q (и соответствующая ей эрмитова полуторалинейная функция) называется* **положительно определенной**, *если $q(x) > 0$ для любого ненулевого вектора x.*

В силу утверждения 17.1 это определение корректно.

Для того, чтобы эрмитова квадратичная функция была положительно определенной, необходимо и достаточно, чтобы $k = n$, $l = 0$ в нормальном виде (44).

Если все угловые миноры эрмитовой полуторалинейной функции отличны от нуля, то можно так же, как в случае билинейной функции получить формулу Якоби. В частности, имеет место аналог критерия Сильвестра:

Теорема 17.2. *Эрмитова квадратичная функция в конечномерном комплексном векторном пространстве положительно определена тогда и только тогда, когда все угловые миноры ее матрицы в произвольном базисе положительны.*

Отметим, что мы имеем право сравнивать угловые миноры с нулем, так как матрица эрмитовой квадратичной функции (т.е. соответствующей ей эрмитовой полуторалинейной функции) является эрмитовой, угловые миноры эрмитовой матрицы – это определители от эрмитовых матриц, а определители от эрмитовых матриц всегда вещественны.

17.2. Понятие унитарного пространства.

Комплексным аналогом евклидова пространства является унитарное пространство

Определение 17.4. *Векторное пространство V над полем комплексных чисел $\mathbb{C}$ называется* **унитарным** *(или* **эрмитовым***), если на V зафиксирована некоторая положительно определенная эрмитова полуторалинейная функция, называемая* **(эрмитовым) скалярным произведением** *и обозначаемая $(*, *)$.*

Итак, для эрмитова скалярного произведения выполняются следующие основные свойства (аксиомы):
1) $(x, y) = \overline{(y, x)}\ \forall x, y \in V$;
2) $(x + y, z) = (x, z) + (y, z)\ \forall x, y, z \in V$;
3) $(\lambda x, y) = \lambda(x, y)\ \forall \lambda \in \mathbb{C}, \forall x, y \in V$;
4) $(x, x) > 0\ \forall x \neq 0, x \in V$.

Антилинейность по второму аргументу уже следует из этих свойств.

ПРИМЕР 1. В пространстве строк $\mathbb{C}^n$ длины n с комплексными элементами эрмитово скалярное произведение можно задать так

$$((a_1, \ldots, a_n), (b_1, \ldots, b_n)) = a_1\overline{b}_1 + \ldots + a_n\overline{b}_n.$$

ПРИМЕР 2. В пространстве квадратных матриц $M_n(\mathbb{C})$ порядка n с комплексными элементами эрмитово скалярное произведение можно задать так

$$(A, B) = tr\ (A\overline{B}^t).$$

В унитарном пространстве можно определить понятие **длины** вектора по той же самой формуле, что и для евклидова пространства

$$|x| = \sqrt{(x;x)}.$$

Утверждение 17.3. *Для любых векторов x, y унитарного пространства*

$$(x,x)(y,y) \geqslant (x,y)(y,x) = |(x,y)|^2, \tag{45}$$

причем равенство достигается в точности тогда, когда векторы x и y пропорциональны.

Неравенство (45) называется **неравенством Коши-Буняковского** (или *неравенством Шварца*).

Доказательство. Представим комплексное число (x, y) в тригонометрической форме $(x,y) = |(x,y)|e^{i\varphi}, \varphi \in \mathbb{R}$. Тогда при любом $t \in \mathbb{R}$ выполняется неравенство

$$(x,x)t^2 + ((x,y)e^{-i\varphi} + \overline{(x,y)}e^{i\varphi})t + (y,y) = (xt + ye^{i\varphi}, xt + ye^{i\varphi}) \geqslant 0.$$

Так как

$$(x,y)e^{-i\varphi} = |(x,y)|e^{i\varphi}e^{-i\varphi} = |(x,y)|$$

и

$$\overline{(x,y)}e^{i\varphi} = |(x,y)|e^{-i\varphi}e^{i\varphi} = |(x,y)|,$$

то неравенство имеет вид

$$(x,x)t^2 + 2|(x,y)|t + (y,y) \geqslant 0.$$

Получающееся условие на дискриминант приводит к неравенству (45). Оно превращается в равенство тогда, когда $xt_0 + ye^{i\varphi} = 0$ при подходящем $t_0 \in \mathbb{R}$, т.е. при пропорциональных x, y. ■

Из неравенства (45) следует **неравенство треугольника**

$$|x + y| \leqslant |x| + |y|$$

(докажите это!)

Так как скалярное произведение двух векторов, вообще говоря, комплексно, то мы не будем определять угол между векторами, а будем лишь использовать, определяемое, как и ранее, понятие ортогональности двух векторов: векторы x и y называются **ортогональными**, если $(x, y) = 0$. Дословно повторяется описание процесса ортогонализации Грама - Шмидта, позволяющего находить по линейно независимым системам ортонормированные системы с сохранением линейных оболочек. Отсюда получается доказательство того факта, что в конечномерном унитарном пространстве существует ортогональный базис.

Базис $\{e_1, \ldots, e_n\}$ унитарного пространства V называется **ортонормированным**, если $(e_i, e_j) = \delta_i^j$. Полезно отметить, что если векторы $x = x_1e_1 +$

$\ldots + x_n e_n$, $y = y_1 e_1 + \ldots + y_n e_n$ даны своими координатами в ортонормированном базисе $\{e_1, \ldots, e_n\}$, то

$$(x, y) = x_1 \bar{y}_1 + \ldots + x_n \bar{y}_n. \tag{46}$$

Из этой формулы вытекает, что любые два унитарных пространства одной и той же размерности изоморфны как унитарные пространства, т.е. при изоморфизме Φ сохраняется значение эрмитова скалярного произведения: $(\Phi(x), \Phi(y)) = (x, y)$.

В отличии от евклидова пространства эрмитово пространство V не отождествляется с V^*: линейная функция $\varphi \in V^*$ представляется в виде $\varphi(x) = (x, a)$ для некоторого однозначно определенного вектора $a \in V$, но соответствие между φ и a не является линейным.

Матрицы перехода от одного ортонормированного базиса конечномерного унитарного пространства к другому ортонормированному базису удовлетворяют условию $T^t\overline{T} = E$. Такие комплексные матрицы называются **унитарными**. Действительно, если даны два ортонормированных базиса $\mathcal{E} = \{e_1, \ldots, e_n\}$ и $\widetilde{\mathcal{E}} = \{\tilde{e}_1, \ldots, \tilde{e}_n\}$, $T = (t_{i,j})$ – матрица перехода от $\mathcal{E}$ к $\widetilde{\mathcal{E}}$, то для любой пары индексов i, j должно выполняться $(e_i, e_j) = \delta_i^j$ и $(\tilde{e}_k, \tilde{e}_l) = \delta_k^l$. Однако $\tilde{e}_k = \sum_{i=1}^n t_{i,k} e_i$, $\tilde{e}_l = \sum_{i=1}^n t_{i,l} e_i$. Вычисляя скалярное произведение с учетом предыдущих формул и определения эрмитова скалярного произведения, получим

$$\delta_k^l = (\tilde{e}_k, \tilde{e}_l) = \sum_{i=1}^n t_{i,k} \bar{t}_{i,l}.$$

Это равенство говорит о том, что столбцы матрицы T образуют ортонормированный базис в пространстве $\mathbb{C}^n$. На языке же матриц имеем $T^t\overline{T} = E$. Снова, как и прежде, можно заметить, что совокупность U_n унитарных матриц порядка n образует группу.

Отметим, что определитель унитарной матрицы T равен по модулю 1. В самом деле, взяв определитель от обеих частей равенства $T^t\overline{T} = E$, получаем

$$\det T \overline{\det T} = 1,$$

а это и означает, что $|\det T| = 1$.

Как и ранее, можно ввести понятие ортогонального дополнения. Так же, как и в случае евклидова пространства, для любого подпространства U конечномерного унитарного пространства V получаем разложение $V = U \oplus U^\perp$. Как и ранее, определяются понятия ортогональной проекции $\mathrm{pr}_U x$ и ортогональной составляющей $\mathrm{ort}_U x$ вектора x относительно подпространства U. В унитарном пространстве также справедливо, что расстояние от вектора x унитарного пространства V до подпространства $U \subset V$ равно $|\mathrm{ort}_U x|$, причем единственным ближайшим к x вектором подпространства U является $\mathrm{pr}_U x$.

18. Линейные операторы в унитарном пространстве.

Для линейных операторов в унитарных пространствах имеется теория, аналогичная рассмотренной для линейных операторов в евклидовых пространствах, причем в унитарном пространстве все оказывается даже проще, так как в унитарном пространстве всякий линейный оператор имеет собственный вектор. Изложим эту теорию вкратце, опуская большинство доказательств, так как они аналогичны приведенным выше для евклидовых пространств.

ЗАДАЧА. Провести доказательства всех теорем, утверждений и лемм этого раздела.

18.1. Сопряженный линейный оператор в унитарном пространстве.

Пусть V – конечномерное унитарное пространство.

Также как и для евклидова пространства можно доказать, что формула

$$\beta_{\mathcal{A}}(x,y) = (x, \mathcal{A}y)$$

устанавливает в конечномерном унитарном пространстве взаимно однозначное соответствие между полуторалинейными функциями и линейными операторами. В произвольном ортонормированном базисе для матрицы A линейного оператора и матрицы B соответствующей ему полуторалинейной функции выполняется

$$B = \overline{A},$$

где черта над матрицей означает комплексное сопряжение всех ее элементов.

Определение 18.1. *Пусть $\mathcal{A}$ – линейный оператор в V. Линейный оператор $\mathcal{A}^* : V \to V$ называется* **сопряженным** *к $\mathcal{A}$, если для любых $x, y \in V$*

$$(\mathcal{A}x, y) = (x, \mathcal{A}^*y).$$

Повторяя рассуждения, приведенные для случая евклидова пространства, получаем, что справедливо следующее.

Теорема 18.1. *В конечномерном унитарном пространстве V каждому линейному оператору $\mathcal{A}$ отвечает сопряженный оператор и притом только один.*

Утверждение 18.1. *Если оператора $\mathcal{A}$ в каком-нибудь ортонормированном базисе $\{e_1, \ldots, e_n\}$ имеет матрицу A, то оператор $\mathcal{A}^*$ в том же базисе имеет матрицу $\overline{A}^t$.*

Связь операции перехода от $\mathcal{A}$ к $\mathcal{A}^*$ с операциями сложения и умножения линейных операторов аналогична как в случае евклидова пространства (смотри утверждение 15.2), отличие только в четвертом утверждении: в унитарном пространстве выполняется равенство

$$(\lambda\mathcal{A})^* = \overline{\lambda}\mathcal{A}^*.$$

18.2. Эрмитовый линейный оператор.

Пусть V – конечномерное унитарное пространство.

Определение 18.2. *Линейный оператор $\mathcal{A} : V \to V$ унитарного пространства V называется* **эрмитовым** *(или* **самосопряженным***), если*

$$\mathcal{A} = \mathcal{A}^*,$$

т.е. для любых $x, y \in V$ выполняется

$$(\mathcal{A}x, y) = (x, \mathcal{A}y).$$

Утверждение 18.2. *Линейный оператор $\mathcal{A} : V \to V$ конечномерного унитарного пространства V является эрмитовым тогда и только тогда, когда его матрица A в произвольном ортонормированном базисе является эрмитовой (т.е. $A^t = \overline{A}$). В частности, на диагонали матрицы A стоят действительные числа.*

Утверждение 18.3. *Собственные значения эрмитова оператора вещественны.*

Доказательство. Пусть x – собственный вектор эрмитова оператора $\mathcal{A}$ и λ – соответствующее собственное значение, т.е.

$$\mathcal{A}x = \lambda x, \quad x \neq 0.$$

По определению эрмитова оператора

$$(\mathcal{A}x, x) = (x, \mathcal{A}x),$$

т.е.

$$(\lambda x, x) = (x, \lambda x).$$

Вынося λ за скобки, получим

$$\lambda(x, x) = \overline{\lambda}(x, x),$$

и так как $(x, x) \neq 0$, то $\lambda = \overline{\lambda}$, что и требовалось доказать. ■

Утверждение 18.4. *Собственные векторы эрмитова оператора, соответствующие различным собственным значениям, взаимно ортогональны.*

Утверждение 18.5. *Для любого конечномерного унитарного пространства выполняется следующее.*

1) Сумма эрмитовых линейных операторов, а также произведение эрмитова оператора на **вещественное** *число является снова эрмитовым оператором.*

2) Композиция двух эрмитовых операторов является эрмитовым оператором тогда и только тогда, когда эти операторы перестановочны.

18.3. Унитарный линейный оператор.

Пусть V – конечномерное унитарное пространство.

Определение 18.3. *Линейный оператор $\mathcal{A}: V \to V$ унитарного пространства V называется* **унитарным**, *если он сохраняет значение скалярного произведения, т.е. для любых $x, y \in V$ выполняется*

$$(\mathcal{A}x, \mathcal{A}y) = (x, y).$$

Отметим, что равенство $(\mathcal{A}x, \mathcal{A}y) = (x, y)$ для любых $x, y \in V$ эквивалентно тому, что $\mathcal{A}^*\mathcal{A} = \mathcal{I}$. В силу свойств операции сопряжения в конечномерном пространстве имеем, что условия $\mathcal{A}^*\mathcal{A} = \mathcal{I}$ и $\mathcal{A}\mathcal{A}^* = \mathcal{I}$ эквивалентны. Следовательно, линейный оператор $\mathcal{A}$ унитарен тогда и только тогда, когда $\mathcal{A}^* = \mathcal{A}^{-1}$. В частности, унитарный линейный оператор является невырожденным.

Множество невырожденных линейных операторов n-мерного пространства V образует группу, называемую **полной линейной группой** пространства V и обозначаемую $\mathrm{GL}(V)$.

Утверждение 18.6. *Унитарные операторы в унитарном пространстве V образуют подгруппу группы $\mathrm{GL}(V)$, называемую* **унитарной группой** *и обозначаемую $\mathrm{U}(V)$.*

Утверждение 18.7. *Линейный оператор $\mathcal{A}: V \to V$ конечномерного унитарного пространства V является унитарным тогда и только тогда, когда его матрица A в произвольном ортонормированном базисе является унитарной (т.е. $A^t\overline{A} = E$).*

Утверждение 18.8. *Линейный оператор унитарен тогда и только тогда, когда он сохраняет длины векторов.*

В частности, унитарный оператор сохраняет углы между векторами.

Утверждение 18.9. *Линейный оператор конечномерного унитарного пространства V унитарен тогда и только тогда, когда он переводит ортонормированный базис в ортонормированный базис.*

Утверждение 18.10. *Собственные значения унитарного линейного оператора по модулю равны* 1.

Доказательство. Пусть x – собственный вектор унитарного оператора $\mathcal{A}$ с собственным значением λ. Тогда

$$(x, x) = (\mathcal{A}x, \mathcal{A}x) = (\lambda x, \lambda x) = \lambda\overline{\lambda}(x, x).$$

Так как $(x, x) \neq 0$, получаем $\lambda\overline{\lambda} = 1$, значит, $|\lambda| = 1$, что и требовалось. ■

18.4. Канонические виды для эрмитовых и унитарных операторов.

Лемма 18.1. *Пусть V – конечномерное унитарное пространство, $\mathcal{A} : V \to V$ – эрмитовый или унитарный линейный оператор пространства V, U – подпространство в V, инвариантное относительно $\mathcal{A}$, $U^\perp$ – ортогональное дополнение к U. Тогда $U^\perp$ является подпространством, инвариантным относительно $\mathcal{A}$.*

С помощью этой леммы получим канонический вид для матриц эрмитова и унитарного линейных операторов.

Теорема 18.2. *Для любого эрмитова или унитарного линейного оператора $\mathcal{A}$ конечномерного унитарного пространства V найдется ортонормированный базис $\mathcal{E}$ из собственных векторов.*

Доказательство. У любого линейного оператора над полем комплексных чисел есть одномерное инвариантное подпространство U (следствие 5.1). По лемме 18.1 ортогональное дополнение $U^\perp$ к U также будет инвариантным. Применив к $U^\perp$ предположение индукции, получим требуемое. ■

Утверждение теоремы 18.2 эквивалентно следующему: *Для любого эрмитова или унитарного линейного оператора $\mathcal{A}$ конечномерного унитарного пространства V найдется ортонормированный базис, в котором матрица оператора $\mathcal{A}$ диагональна, т.е. имеет вид*

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

где $\lambda_1, \ldots, \lambda_n$ – собственные значения оператора $\mathcal{A}$. Причем если $\mathcal{A}$ – эрмитов, то $\lambda_1, \ldots, \lambda_n$ вещественны (в силу утверждения 18.3), а если $\mathcal{A}$ унитарен, то $\lambda_1, \ldots, \lambda_n$ – числа, по модулю равные единице (в силу утверждения 18.10).

Тем, что каждый унитарный оператор имеет ортонормированный базис из собственных вектор, унитарные операторы существенно отличаются от ортогональных операторов евклидова пространства.

Также заметим, что прямое обращение теоремы 15.3 – утверждение 15.4 – на эрмитовы операторы не переносится: эрмитова матрица должна иметь вещественные числа на главной диагонали, а потому не всякая диагональная матрица эрмитова. Однако, используя утверждение 18.2, можно доказать следующее.

Утверждение 18.11. *Если для линейного оператора в конечномерном унитарном пространстве найдется ортонормированный базис, в котором его матрица диагональна с вещественными числами на главной диагонали, то этот оператор является эрмитовым.*

Так как матрица перехода от одного ортонормированного базиса к другому задается унитарной матрицей, то основной результат данного раздела можно в матричных терминах сформулировать следующим образом:

Пусть A – заданная эрмитова матрица. Тогда A может быть представлена в виде

$$A = T^{-1}DT,$$

где T – унитарная матрица, а D – диагональная матрица, у которой на диагонали стоят вещественные числа.

Пусть U – заданная унитарная матрица. Тогда существует такая унитарная матрица T, что U представляется в виде

$$U = T^{-1}DT,$$

где D – диагональная матрица, у которой по диагонали стоят числа, по модулю равные 1.

Замечание 18.1. *Комплексификация $V^{\mathbb{C}}$ евклидова пространства V каноническим образом превращается в унитарное пространство, если определить эрмитово скалярное умножение по формуле*

$$(x + iu, y + iv) = [(x, y) + (u, v)] + i[(x, v) - (y, u)].$$

При этом комплексное продолжение $\mathcal{A}_{\mathbb{C}}$ самосопряженного (соотв., ортогонального) оператора $\mathcal{A}$ будет эрмитовым (соотв., унитарным) оператором.

Формула

$$\beta_{\mathcal{A}}(x, y) = (x, \mathcal{A}y)$$

устанавливает в конечномерном унитарном пространстве взаимно однозначное соответствие между эрмитовыми полуторалинейными функциями и эрмитовыми линейными операторами.

Применяя теорему 18.2 о существовании ортонормированного базиса из собственных векторов для эрмитова оператора, мы получаем, что *для любой эрмитовой квадратичной функции q в конечномерном унитарном пространстве существует ортонормированный базис, в котором ее матрица диагональна, т.е.*

$$q(x) = \lambda_1|x_1|^2 + \ldots + \lambda_n|x_n|^2. \tag{47}$$

Числа $\lambda_1, \ldots, \lambda_n$ определены однозначно с точностью до перестановки, так как это собственные значения соответствующего эрмитова оператора.

18.5. Аналогия с комплексными числами.

Операция * в известной мере аналогична операции перехода от данного комплексного числа z к комплексно сопряженному $\overline{z}$. Эта аналогия не случайна. Действительно, для матриц первого порядка над комплексным полем операция * как раз и состоит в замене данного числа комплексно сопряженным. Среди всех комплексных чисел вещественные числа характеризуются тем свойством, что $\overline{z} = z$. Для линейных операторов аналогичным объектом являются эрмитовы операторы, которые, как мы знаем, характеризуются равенством $\mathcal{A}^* = \mathcal{A}$.

Всякое комплексное число z можно однозначно представить в алгебраической форме: $z = a + ib$, где a и b – вещественные числа. Аналогично этому:

Всякий линейный оператор $\mathcal{A}$ может быть единственным образом представлен в виде
$$\mathcal{A} = \mathcal{A}_1 + i\mathcal{A}_2,$$
где $\mathcal{A}_1$ и $\mathcal{A}_2$ – эрмитовы операторы.

Действительно,
$$\mathcal{A} = \frac{\mathcal{A} + \mathcal{A}^*}{2} + i\frac{\mathcal{A} - \mathcal{A}^*}{2i}.$$
Введем обозначения
$$\frac{\mathcal{A} + \mathcal{A}^*}{2} = \mathcal{A}_1, \quad \frac{\mathcal{A} - \mathcal{A}^*}{2i} = \mathcal{A}_2.$$
Тогда
$$\mathcal{A}_1^* = (\frac{\mathcal{A} + \mathcal{A}^*}{2})^* = \frac{1}{2}(\mathcal{A} + \mathcal{A}^*)^* = \frac{1}{2}(\mathcal{A}^* + (\mathcal{A}^*)^*) =$$
$$= \frac{1}{2}(\mathcal{A}^* + \mathcal{A}) = \mathcal{A}_1$$
и
$$\mathcal{A}_2^* = (\frac{\mathcal{A} - \mathcal{A}^*}{2i})^* = -\frac{1}{2i}(\mathcal{A} - \mathcal{A}^*)^* = -\frac{1}{2i}(\mathcal{A}^* - (\mathcal{A}^*)^*) =$$
$$= -\frac{1}{2i}(\mathcal{A}^* - \mathcal{A}) = \mathcal{A}_2,$$
т.е. $\mathcal{A}_1$ и $\mathcal{A}_2$ – эрмитовы операторы.

Докажем единственность представления в виде $\mathcal{A} = \mathcal{A}_1 + i\mathcal{A}_2$, где $\mathcal{A}_1$ и $\mathcal{A}_2$ – эрмитовы операторы. Допустим, $\mathcal{A} = \mathcal{A}_1' + i\mathcal{A}_2'$ – другое такое представление. Тогда
$$\mathcal{A}_1 + i\mathcal{A}_2 = \mathcal{A}_1' + i\mathcal{A}_2'$$
или
$$\mathcal{A}_1 - \mathcal{A}_1' = i(\mathcal{A}_2' - \mathcal{A}_2).$$

Применяя операцию сопряжения к левой части равенства, получим:

$$(\mathcal{A}_1 - \mathcal{A}'_1)^* = \mathcal{A}_1^* - {\mathcal{A}'_1}^* = \mathcal{A}_1 - \mathcal{A}'_1.$$

А применяя операцию сопряжения к правой части равенства,

$$(i(\mathcal{A}'_2 - \mathcal{A}_2))^* = -i(\mathcal{A}'_2 - \mathcal{A}_2)^* = -i({\mathcal{A}'_2}^* - \mathcal{A}_2^*) = -i(\mathcal{A}'_2 - \mathcal{A}_2) = -(\mathcal{A}_1 - \mathcal{A}'_1).$$

Следовательно, $\mathcal{A}_1 - \mathcal{A}'_1 = -(\mathcal{A}_1 - \mathcal{A}'_1)$. Откуда получаем, что $\mathcal{A}_1 = \mathcal{A}'_1$. Значит, и $\mathcal{A}_2 = \mathcal{A}'_2$, что и завершает доказательство.

Таким образом, эрмитовы операторы играют среди всех линейных операторов унитарного пространства роль, аналогичную роли вещественных чисел среди всех комплексных.

Своеобразным аналогом положительных вещественных чисел служат положительно определенные эрмитовы операторы.

Определение 18.4. *Эрмитов оператор $\mathcal{C}$ называется* **положительно определенным***, если соответствующая ему эрмитова полуторалинейная функция положительно определена, т.е. для любого ненулевого вектора x выполняется:*

$$(x, \mathcal{C}x) > 0.$$

Положительная определенность эрмитова оператора равносильна тому, что все его собственные значения положительны.

Теорема 18.3. *Всякий невырожденный линейный оператор $\mathcal{A}$ в конечномерном унитарном пространстве единственным образом представляется в виде произведения*

$$\mathcal{A} = \mathcal{C}\mathcal{O},$$

где $\mathcal{C}$ – положительно определенный эрмитовый оператор, а $\mathcal{O}$ – унитарный оператор.

Такое представление линейного оператора называется его **полярным разложением**.

В одномерном случае невырожденный линейный оператор есть просто ненулевое комплексное число, а его полярное разложение – тригонометрическая форма этого числа. Поскольку тригонометрическая форма комплексного числа связана с полярными координатами на плоскости, это объясняет термин "полярное разложение" в общем случае.

19. Аффинные пространства.

В элементарной геометрии мы имели дело не только с векторами, но и с точками. Подобно тому, как аксиоматика векторного пространства отражает в обобщенном виде свойства векторов элементарной геометрии, аксиоматика аффинного пространства отражает свойства точек и векторов элементарной геометрии в их взаимосвязи.

19.1. Основные понятия.

Пусть V – векторное пространство над полем F.

Определение 19.1. Аффинным пространством, *ассоциированным с векторным пространством* V, *называется множество* $\mathbb{A}$ *(а точнее, пара* $(\mathbb{A}, V)$*) вместе с операцией сложения* $\mathbb{A} \times V \to \mathbb{A}$, *удовлетворяющая следующим условиям:*

1) $a + (u + v) = (a + u) + v \quad (a \in \mathbb{A}, u, v \in V)$;

2) $a + 0 = a \quad (a \in \mathbb{A},\ 0$ *– нулевой вектор);*

3) для любых $a, b \in \mathbb{A}$ *существует единственный вектор* $x \in V$, *такой, что* $a + x = b$.

Будем также говорить, что $\mathbb{A}$ – аффинное пространство над V.

Элементы множества $\mathbb{A}$ называются **точками**. Вектор x из условия 3) называется **вектором, соединяющим точки** a **и** b, и обозначается через $\overline{ab}$. Таким образом, $b = a + \overline{ab}$ (неформально говоря, точка b получается откладыванием от точки a вектора $\overline{ab}$). Отметим, что для любой точки $a \in \mathbb{A}$ и любого вектора $v \in V$ существует единственная точка $b \in \mathbb{A}$ такая, что $\overline{ab} = v$.

Из аксиомы 3) также следует, что, фиксируя точку $a \in \mathbb{A}$, мы получаем биекцию $v \mapsto a + v$ между V и $\mathbb{A}$.

Из аксиомы 1) получаем, что для любых (не обязательно различных) трех точек $a, b, c \in \mathbb{A}$ выполнено равенство $(a + \overline{ab}) + \overline{bc} = a + \overline{ac}$ и равенство $\overline{ab} + \overline{bc} = \overline{ac}$ (называемое аксиомой треугольника).

Из аксиомы треугольника легко получить, что для любых точек $a, b \in \mathbb{A}$ выполняется $\overline{aa} = 0$ и $\overline{ab} = -\overline{ba}$.

Пример. Типичный пример аффинного пространства получается при $\mathbb{A} = V$. При этом векторы рассматриваются как точки. Сопоставление упорядоченной паре точек вектора, указанное в определении, в данном случае состоит в том, что упорядоченной паре точек $u, v \in V$ соответствует вектор $\overline{uv} = v - u$.

Размерностью $\dim \mathbb{A}$ **аффинного пространства** $\mathbb{A}$ над векторным пространством V называется размерность векторного пространства V.

Далее будем рассматривать аффинные пространства конечной размерности.

Определение 19.2. **Аффинной системой координат** *в n-мерном аффинном пространстве $\mathbb{A}$ над векторным пространством V называется пара $(o, \mathcal{E})$, где o – некоторая фиксированная точка в $\mathbb{A}$ (называемая* **началом координат***), $\mathcal{E} = \{e_1, ..., e_n\}$ – базис линейного пространства V.* **Координатами точки** *$a \in \mathbb{A}$ называются координаты вектора $\overline{oa}$ в базисе $\mathcal{E}$.*

Из равенства $\overline{ab} = \overline{ob} - \overline{oa}$ следует, что если даны координаты $(a_1, \ldots, a_n)$ точки a и координаты $(b_1, \ldots, b_n)$ точки b в системе координат $(o, \mathcal{E})$, то координаты вектора $\overline{ab}$ в базисе $\mathcal{E}$ равны $(b_1 - a_1, \ldots, b_n - a_n)$. Обратно, если $b = a + v$ и даны координаты $(a_1, ..., a_n)$ точки a в системе координат $(o, \mathcal{E})$ и координаты $(v_1, ..., v_n)$ вектора v в базисе $\mathcal{E}$, то координаты точки b в системе координат $(o, \mathcal{E})$ равны $(a_1 + v_1, ..., a_n + v_n)$.

Замечание 19.1. *Систему координат можно задать также $n+1$ точками $\{o_0, o_1, \ldots, o_n\}$ такими, что векторы $\overline{o_0o_1}, \ldots, \overline{o_0o_n}$ образуют базис пространства V.*

Пусть в аффинном пространстве $\mathbb{A}$ даны две системы координат $(o, \mathcal{E})$ и $(\widetilde{o}, \widetilde{\mathcal{E}})$. Рассмотрим произвольную точку $a \in \mathbb{A}$. Из равенства $\overline{oa} = \overline{\widetilde{o}a} + \overline{o\widetilde{o}}$, учитывая формулу (1) преобразования координат вектора при переходе к другому базису, получаем формулу преобразования координат точки при переходе к другой системе координат:

$$X = T_{\mathcal{E} \to \widetilde{\mathcal{E}}} \widetilde{X} + X_o, \tag{48}$$

где $T_{\mathcal{E} \to \widetilde{\mathcal{E}}}$ – матрица перехода от базиса $\mathcal{E}$ к базису $\widetilde{\mathcal{E}}$, X – столбец координат точки a в $(o, \mathcal{E})$, $\widetilde{X}$ столбец координат точки a в $(\widetilde{o}, \widetilde{\mathcal{E}})$, X_o – столбец координат точки $\widetilde{o}$ в $(o, \mathcal{E})$, .

Матрицей перехода от $(o, \mathcal{E})$ к $(\widetilde{o}, \widetilde{\mathcal{E}})$ называется блочная матрица

$$\widehat{T} = \begin{pmatrix} T_{\mathcal{E} \to \widetilde{\mathcal{E}}} & X_o \\ 0 & 1 \end{pmatrix}.$$

Имеем $\begin{pmatrix} X \\ 1 \end{pmatrix} = \widehat{T} \begin{pmatrix} \widetilde{X} \\ 1 \end{pmatrix}$.

19.2. Изоморфизм.

Пусть V и W – векторные пространства над одним и тем же полем

Определение 19.3. *Пусть даны два аффинных пространства $\mathbb{A}$ над V и $\mathbb{B}$ над W. Отображение $\Phi : \mathbb{A} \to \mathbb{B}$ называется* **аффинным** *(***аффинно-линейным***), если существует линейное отображение $\varphi : V \to W$, такое, что для любых $a \in \mathbb{A}$ и $v \in V$ выполняется $\Phi(a + v) = \Phi(a) + \varphi(v)$.*

Отображение $\varphi : V \to W$ называется иногда **линейной частью** *(или* **дифференциалом***) отображения Φ.*

Биективное аффинное отображение называется **(аффинным) изоморфизмом**.

Заметим, что если через b обозначить точку $a + v$, то условие

$$\Phi(a + v) = \Phi(a) + \varphi(v)$$

переписывается в виде

$$\varphi(\overline{ab}) = \overline{\Phi(a)\Phi(b)},$$

а также

$$\Phi(b) = \Phi(a) + \varphi(\overline{ab}).$$

Последнее, в частности, означает, что достаточно знать аффинное отображение в любой фиксированной точке и соответствующее ему линейное отображение.

Лемма 19.1. *Аффинное отображение $\Phi : \mathbb{A} \to \mathbb{B}$ является биективным тогда и только тогда, когда его линейная часть $\varphi : V \to W$ также биективна, т.е. является изоморфизмом векторных пространств.*

Доказательство. Пусть Φ – биекция. Для любой точки $a \in \mathbb{A}$ и любого вектора $v \in V$ имеем $\Phi(a+v) = \Phi(a)+\varphi(v)$. Если $v \neq 0$, то $a+v \neq a$. В силу инъективности отображения Φ получаем, что $\Phi(a+v) \neq \Phi(a)$. Следовательно, $\varphi(v) \neq 0$, т.е. φ инъективно. Далее, для любого $w \in W$ рассмотрим точку $b = \Phi(a) + w \in \mathbb{B}$. В силу сюръективности отображения Φ существует точка $\widetilde{a} \in \mathbb{A}$ такая, что $\Phi(\widetilde{a}) = b$. Тогда по определению $\varphi(\overline{a\widetilde{a}}) = \overline{\Phi(a)\Phi(\widetilde{a})} = w$, т.е. φ сюръективно.

Обратно, если $\varphi : V \to W$ является изоморфизмом векторных пространств и точки $a, \widetilde{a} \in \mathbb{A}$ различны, то $\overline{a\widetilde{a}} \neq 0$. Следовательно, $\varphi(\overline{a\widetilde{a}}) \neq 0$, а значит, $\Phi(\widetilde{a}) = \Phi(a)+\varphi(\overline{a\widetilde{a}}) \neq \Phi(a)$, т.е. Φ инъективно. Далее, для любой точки $b \in \mathbb{B}$ рассмотрим вектор $w = \overline{\Phi(a)b}$. В силу сюръективности φ существует вектор $v \in V$, такой, что $\varphi(v) = w$. Имеем $\Phi(a + v) = \Phi(a) + \varphi(v) = \Phi(a) + w = b$, т.е. Φ сюръективно. ∎

Утверждение 19.1. *Два конечномерных аффинных пространства изоморфны тогда и только тогда, когда их размерности равны.*

Доказательство следует из леммы 19.1 и теоремы 1.5 о том, что векторные пространства изоморфны тогда и только тогда, когда их размерности равны.

Действительно, если аффинные пространства $\mathbb{A}$ и $\mathbb{B}$ изоморфны, то векторные пространства V и W, ассоциированные с ними, изоморфны, а значит, $\dim V = \dim W$. Следовательно, $\dim \mathbb{A} = \dim \mathbb{B}$. Обратно, если $\dim \mathbb{A} = \dim \mathbb{B}$, т.е. $\dim V = \dim W$, то существует изоморфизм $\varphi : V \to W$. Тогда аффинный изоморфизм $\Phi : \mathbb{A} \to \mathbb{B}$ строится по формуле $\Phi(a + v) = b + \varphi(v)$ для любых $a \in \mathbb{A}, b \in \mathbb{B}$. ∎

Таким образом, каждое аффинное пространство размерности n изоморфно пространству строк длины n над соответствующим полем. Вектор, идущий от одной строки к другой, равен разности этих строк.

19.3. Аффинные подпространства.

Определение 19.4. *Пусть U – подпространство векторного пространства V, a – точка из аффинного пространства $\mathbb{A}$. Множество π всех точек из $\mathbb{A}$ вида $a + u$, где $u \in U$, называется* **аффинным подпространством** *или* **плоскостью, проходящей через точку** a **и имеющей направляющее подпространство** U. *Пишем $\pi = a + U$.*

Размерностью плоскости π называется размерность ее направляющего подпространства U, т.е. $\dim \pi = \dim U$. Нульмерная плоскость есть точка. Одномерная плоскость называется **прямой**, а $(n-1)$-мерная плоскость n-мерного аффинного пространства – **гиперплоскостью.**

Заметим, что для любой точки $b \in \pi = a+U$ верно $\pi = b+U$. Действительно, если $b \in \pi = a+U$, то $\overline{ab} \in U$. Следовательно, если $c \in b+U$, то для некоторого $u \in U$ выполняется $c = b+u = (a+\overline{ab})+u = a+(\overline{ab}+u) \in a+U = \pi$. Обратно, если $c' \in a + U = \pi$, то для некоторого $u' \in U$ выполняется $c' = a + u' = (b + \overline{ba}) + u' = b + (-\overline{ab} + u') \in b + U$.

Заметим, что направляющее подпространство U плоскости $\pi = a+U$ состоит из тех и только тех векторов, которые могут быть получены в виде $\overline{pq}$ для произвольных точек $p, q \in \pi$. В самом деле, если $p, q \in \pi$, то $\overline{pq} = \overline{aq}-\overline{ap} \in U$. Если же $u \in U$, то существуют точки $p, q \in \pi$ такие, что $u = \overline{pq}$, достаточно взять $p = a, q = a + u$. Направляющее подпространство U однозначно определяется плоскостью π как совокупность всех векторов, соединяющих точки плоскости π.

Утверждение 19.2. *Всякая плоскость $\pi = a+U$ в аффинном пространстве сама является аффинным пространством, ассоциированным с векторным пространством U.*

Доказательство. Для любой точки $p \in \pi$ и любого вектора $u \in U$ сумма $p+u$ принадлежит π. Относительно этой операции аксиомы 1) и 2) аффинного пространства (с заменой V на U, $\mathbb{A}$ на π) выполняются для π, так как они выполняются для $\mathbb{A}$. Далее, как показано выше, для любых точек $p, q \in \pi$ вектор $u = \overline{pq}$ принадлежит U. При этом $p = q + u$, причем вектор u определен однозначно в V, а значит, и в U. Следовательно, аксиома 3) аффинного пространства тоже выполняется. ■

20. Аффинные подпространства (продолжение)

Пусть V – (конечномерное) векторное пространство над полем F, $\mathbb{A}$ – аффинное пространство, ассоциированное с V

20.1. Взаимное расположение плоскостей.

Утверждение 20.1. *Пусть $\pi_1 = a_1 + U_1$ и $\pi_2 = a_2 + U_2$ – две плоскости аффинного пространства $\mathbb{A}$. Тогда их пересечение $\pi = \pi_1 \cap \pi_2$ либо пусто, либо является плоскостью, причем $\pi = c + U$, где c – произвольная точка из пересечения $\pi_1 \cap \pi_2$, а $U = U_1 \cap U_2$.*

Доказательство. Предположим, что пересечение $\pi_1 \cap \pi_2$ не пусто и $c \in \pi_1 \cap \pi_2$. Тогда $\pi_1 = c + U_1$ и $\pi_2 = c + U_2$. В этом случае $\pi_1 \cap \pi_2 = c + U_1 \cap U_2$.

Действительно, включение $c + U_1 \cap U_2 \subseteq \pi_1 \cap \pi_2$ очевидно.

Докажем обратное включение. Для любой точки d из пересечения $\pi_1 \cap \pi_2$

$$d = c + u_1$$

и

$$d = c + u_2,$$

где $u_i \in U_i$. Следовательно, $u_1 = u_2 \in U_1 \cap U_2$, а значит, $d \in c + U_1 \cap U_2$

■

Пересечение двух плоскостей (если оно не пусто) снова является плоскостью. В общем случае для объединения это не верно. Обозначим через $\langle \pi, \tau \rangle$ *аффинную оболочку* плоскостей π и τ, т.е. наименьшую плоскость, содержащую π и τ.

Теорема 20.1. *Пусть $\pi = a + U$, $\tau = b + W$ – две плоскости. Тогда*

$$\langle \pi, \tau \rangle = a + \langle \overline{ab}, U + W \rangle$$

и

$$\dim \langle \pi, \tau \rangle = \dim \langle \overline{ab}, U + W \rangle.$$

Если пересечение π и τ не пусто, то справедлива обычная формула Грассмана

$$\dim \langle \pi, \tau \rangle + \dim \pi \cap \tau = \dim \pi + \dim \tau.$$

Иначе

$$\dim \langle \pi, \tau \rangle + \dim U \cap W = \dim \pi + \dim \tau + 1.$$

Доказательство. Обозначим через σ плоскость $a + \langle \overline{ab}, U + W \rangle$. Понятно, что $\pi, \tau \subseteq \sigma$, а значит, $\langle \pi, \tau \rangle \subseteq \sigma$. Обратно, так как $a \in \langle \pi, \tau \rangle$, то $\langle \pi, \tau \rangle = a + X$ для некоторого подпространства X векторного пространства V. Поскольку $b \in \langle \pi, \tau \rangle$, имеем $\overline{ab} \in X$. Так как для любого $u \in U$ точка $a + u \in \pi \subseteq \langle \pi, \tau \rangle$, то $u \in X$. Далее, для любого $w \in W$ имеем $b + w \in \tau \subseteq \langle \pi, \tau \rangle$ и $b + w = (a + \overline{ab}) + w = a + (\overline{ab} + w)$, то $\overline{ab} + w \in X$. Поскольку уже $\overline{ab} \in X$, имеем $w \in X$. Получилось, что $\langle \overline{ab}, U+W\rangle \subseteq X$, т.е. $\sigma \subseteq \langle \pi, \tau\rangle$, а значит, σ – наименьшая плоскость, содержащая как π, так и τ, что и требовалось доказать.

Пусть теперь $\pi \cap \tau \neq \emptyset$ и c – произвольная точка из пересечения $\pi \cap \tau$. Тогда $\overline{ab} = \overline{cb} - \overline{ca} \in U+W$. Следовательно, $\dim\,\langle \overline{ab}, U+W\rangle = \dim\,(U+W)$. Из формулы Грассмана (теорема 2.1) получаем, что

$$\dim\,\langle \pi, \tau\rangle = \dim\,(U+W) = \dim\,U + \dim\,W - \dim\,U \cap W =$$

$$= \dim\,\pi + \dim\,\tau - \dim\,\pi \cap \tau,$$

что и требовалось.

Если же $\pi \cap \tau = \emptyset$, то вектор $\overline{ab}$ не лежит в $U+W$, так как если бы $\overline{ab} = u + w$, то по аксиоме 1) аффинного пространства выполнялось бы $b = a + \overline{ab} = a + (u+w) = (a+u) + w$, т.е. $a + u = b - w$ являлась бы общей точкой для π и τ. Следовательно, $\dim\,\langle \overline{ab}, U+W\rangle = \dim\,(U+W) + 1$ и

$$\dim\,\langle \pi, \tau\rangle = \dim\,(U+W) + 1 = \dim\,U + \dim\,W - \dim\,U \cap W + 1 =$$

$$= \dim\,\pi + \dim\,\tau - \dim\,U \cap W + 1,$$

как и утверждалось. ■

В частности, мы доказали, что плоскости $\pi = a + U$ и $\tau = b + W$ пересекаются тогда и только тогда, когда вектор $\overline{ab}$ лежит в $U+W$.

Определение 20.1. *Плоскости $\pi = a+U$, $\tau = b+W$ аффинного пространства $\mathbb{A}$ называются* **параллельными**, *если $U \subseteq W$ или $W \subseteq U$.*

Легко видеть, что если плоскости π и τ параллельны, то либо они не пересекаются, либо одно из них является подмножеством другого: скажем, если $U \subseteq W$ и $c \in \pi \cap \tau$, то $\pi = c + U \subseteq c + W = \tau$.

Определение 20.2. *Плоскости π и τ называются* **скрещивающимися**, *если они не параллельны и не пересекаются.*

Итак, для любых двух плоскостей имеется одна из трех возможностей: они пересекаются, они не пересекаются и параллельны, они скрещиваются.

ЗАДАЧА. Докажите, что если π – гиперплоскость в $\mathbb{A}$, т.е $\dim\,\pi = n-1$, то плоскость ρ, не имеющая с π общих точек, обязательно параллельна π.

20.2. Аффинно-линейные функции и аффинные подпространства.

Рассмотрим теперь класс функций на аффинном пространстве, соответствующий классу линейных функций на векторном пространстве.

Определение 20.3. *Аффинное отображение $f : \mathbb{A} \to F$ называется* **аффинной** *(или* **аффинно-линейной***)* **функцией**. *При этом F рассматривается как одномерное аффинное пространство над одномерным векторным пространством F.*

Таким образом, отображение $f : \mathbb{A} \to F$ называется аффинно-линейной функцией, если существует линейная функция $\varphi : V \to F$, такая, что для любой точки $a \in \mathbb{A}$ и любого вектора $v \in V$ выполняется

$$f(a+v) = f(a) + \varphi(v).$$

Пусть в $\mathbb{A}$ фиксирована система координат $(o, \mathcal{E})$, где $\mathcal{E} = \{e_1, \ldots, e_n\}$. И пусть $\varphi = \alpha_1 \varepsilon^1 + \ldots + \alpha_n \varepsilon^n$ – запись линейной функции φ в сопряженном базисе сопряженного пространства V^* (т.е. $(\alpha_1, \ldots, \alpha_n)$ – строка коэффициентов функции φ). Тогда если $(x_1, \ldots, x_n)$ – строка координат точки a в системе координат $(o, \mathcal{E})$, а $\alpha = f(o)$ – значение аффинной функции f в начале координат o, то получаем запись аффинной функции в координатах

$$f(a) = \alpha + \alpha_1 x_1 + \ldots + \alpha_n x_n, \tag{49}$$

так как $f(a) = f(o + \overline{oa}) = f(o) + \varphi(\overline{oa}) = f(o) + \varphi(x_1 e_1 + \ldots + x_n e_n)$.

Обратно, если некая функция $f : \mathbb{A} \to F$ задана по формуле (49), то она является аффинно-линейной функцией. Действительно, для произвольного вектора $v = v_1 e_1 + \ldots + v_n e_n \in V$ точка $a + v = o + (\overline{oa} + v)$ имеет координаты $(x_1 + v_1, \ldots, x_n + v_n)$ в системе координат $(o, \mathcal{E})$. Поэтому по формуле (49) значение функции f в точке $a + v$ равно

$$f(a+v) = \alpha + \alpha_1(x_1 + v_1) + \ldots + \alpha_n(x_n + v_n).$$

Следовательно,

$$f(a+v) = (\alpha + \alpha_1 x_1 + \ldots + \alpha_n x_n) + (\alpha_1 v_1 + \ldots + \alpha_n v_n) = f(a) + \varphi(v).$$

Выражение $f(a) = \alpha + \alpha_1 x_1 + \ldots + \alpha_n x_n$ будем обозначать $f(x_1, \ldots, x_n)$ и называть аффинно-линейной формой.

Частным случаем аффинно-линейных функций являются постоянные функции. Они характеризуются тем, что их линейная часть равна нулю.

Аффинно-линейные функции образуют $(n+1)$-мерное подпространство в пространстве всех функций на $\mathbb{A}$. Это ясно хотя бы из координатной записи (49).

Определение 20.4. *Множество точек, в которых аффинно-линейная функция $f : \mathbb{A} \to F$ обращается в ноль, называется* **ядром** *функции f и обозначается* $\mathrm{Ker}\ f$.

Утверждение 20.2. *Пусть $f : \mathbb{A} \to F$ – аффинно-линейная функция с линейной частью $\varphi : V \to F$. Тогда* $\mathrm{Ker}\ f$ *либо пусто, либо является плоскостью и равно $a + \mathrm{Ker}\ \varphi$, где a – произвольная точка с условием $f(a) = 0$.*

Доказательство. Пусть $\mathrm{Ker}\ f \neq \emptyset$ и a – произвольная точка с условием $a \in \mathrm{Ker}\ f$. Докажем, что $\mathrm{Ker}\ f = a + \mathrm{Ker}\ \varphi$.

Если b – произвольная точка из $a + \mathrm{Ker}\ \varphi$, т.е. $b = a + u$, где $u \in \mathrm{Ker}\ \varphi$, то $f(b) = f(a + u) = f(a) + \varphi(u) = 0 + 0 = 0$, т.е. $b \in \mathrm{Ker}\ f$.

Если же b – произвольная точка из $\mathrm{Ker}\ f$, т.е. $f(b) = 0$, то $f(b) = f(a) + \varphi(\overline{ab})$. Следовательно, $\varphi(\overline{ab}) = 0$, т.е. $\overline{ab} \in \mathrm{Ker}\ \varphi$, а значит, $b = a + \overline{ab} \in a + \mathrm{Ker}\ \varphi$. ■

В частности, получаем, что пересечение ядер конечного числа аффинно-линейных функций либо пусто, либо является плоскостью.

Оказывается, верно и обратное: любая плоскость является пересечением ядер некоторого конечного множества аффинно-линейных функций.

Теорема 20.2. *Пусть $\mathbb{A}$ – аффинное пространство размерности n. Тогда*

1) Множество точек из $\mathbb{A}$, координаты которых удовлетворяют совместной системе линейных уравнений ранга r, образуют $(n - r)$-мерную плоскость $\pi \subseteq \mathbb{A}$.

2) При фиксированной системе координат $(o, \mathcal{E})$ произвольная плоскость аффинного пространства $\mathbb{A}$ состоит из всех точек, координаты которых составляют множество решений некоторой системы линейных уравнений.

Доказательство. 1) Пусть дана совместная система линейных уравнений

$$\left\{ \begin{array}{ccccccc} a_{1,1}x_1 & + & \dots & + & a_{1,n}x_n & = & b_1, \\ & . & . & . & . & & \\ a_{m,1}x_1 & + & \dots & + & a_{m,n}x_n & = & b_m. \end{array} \right. \tag{50}$$

Рассмотрим аффинные функции $f_i : \mathbb{A} \to F$, которые относительно некоторой аффинной системы координат $(o, \mathcal{E})$ аффинного пространства $\mathbb{A}$ имеют координатную запись $f_i(a) = a_{i,1}x_1 + \ldots + a_{i,n}x_n - b_i$, где $(x_1, \ldots, x_n)$ – строка координат точки a в системе координат $(o, \mathcal{E})$, $i = 1, \ldots, m$. Тогда систему (50) можно записать в виде

$$\left\{ \begin{array}{ccc} f_1(a) & = & 0, \\ . & . & . \\ f_m(a) & = & 0. \end{array} \right. \tag{51}$$

Получаем, что числа $x_1, \ldots, x_n$ являются решением системы (50) тогда и только тогда, когда $(x_1, \ldots, x_n)$ – строка координат точки a в системе координат $(o, \mathcal{E})$, лежащей в пересечении ядер функций $f_1, \ldots, f_m$, т.е. лежащей в некоторой плоскости $\pi = \mathrm{Ker}\ f_1 \cap \ldots \cap \mathrm{Ker}\ f_m$.

Предположим, что $x_1^0, \ldots, x_n^0$ – одно из решений системы (50), а a_0 – точка с координатами $(x_1^0, \ldots, x_n^0)$ в системе координат $(o, \mathcal{E})$. Из 1-го семестра знаем, что общее решение неоднородной системы линейных уравнений есть сумма частного решения неоднородной системы и общего решения соответствующей однородной системы линейных уравнений, т.е.любое решение системы (50) или (51) имеет вид $a = a_0 + x$, где $x \in V$ удовлетворяет системе

$$\left\{\begin{array}{rcl} \varphi_1(x) & = & 0, \\ & . \ . \ . & \\ \varphi_m(x) & = & 0, \end{array}\right. \qquad (52)$$

где φ_i – линейные части функций f_i и $\varphi_i(x) = a_{i,1}x_1 + \ldots + a_{i,n}x_n$, $(x_1, \ldots, x_n)$ – координаты вектора x в базисе $\mathcal{E}$. Из 1-го семестра знаем, что множество решений системы (52) образует подпространство $U \subseteq V$ размерности $n-r$, где r – ранг матрицы коэффициентов системы (50). Таким образом, совокупность решений системы (50) будет плоскость $\pi = a_0 + U$ размерности $n - r$.

2) Докажем, что любую плоскость $\pi = a_0 + U$ можно задать системой линейных уравнений. Действительно, согласно теореме 3.1 подпространство U пространства V есть множество векторов, у которых координаты в некотором базисе $\mathcal{E}$ являются решениями подходящей однородной системы линейных уравнений

$$\left\{\begin{array}{rcccccl} a_{1,1}x_1 & + & ... & + & a_{1,n}x_n & = & 0, \\ & & . \ . \ . \ . & & & & \\ a_{m,1}x_1 & + & ... & + & a_{m,n}x_n & = & 0. \end{array}\right. \qquad (53)$$

Точка a принадлежит плоскости $\pi = a_0 + U$ тогда и только тогда, когда $\overline{a_0a} \in U$. Если $(x_1^0, \ldots, x_n^0)$ – координаты точки a_0, а $(x_1, \ldots, x_n)$– координаты точки a в системе координат $(o, \mathcal{E})$, то $(x_1 - x_1^0, \ldots, x_n - x_n^0)$ – координаты вектора $\overline{a_0a}$ в базисе $\mathcal{E}$, которые удовлетворяют системе

$$\left\{\begin{array}{rcccccl} a_{1,1}(x_1 - x_1^0) & + & ... & + & a_{1,n}(x_n - x_n^0) & = & 0, \\ & & . \ . \ . \ . & & & & \\ a_{m,1}(x_1 - x_1^0) & + & ... & + & a_{m,n}(x_n - x_n^0) & = & 0 \end{array}\right. \qquad (54)$$

Откуда

$$\left\{\begin{array}{rcccccl} a_{1,1}x_1 & + & ... & + & a_{1,n}x_n & = & b_1, \\ & & . \ . \ . \ . & & & & \\ a_{m,1}x_1 & + & ... & + & a_{m,n}x_n & = & b_m, \end{array}\right. \qquad (55)$$

где $b_i = \sum a_{i,j}x_j^0, i = 1, \ldots, m$. ■

21. Аффинные отображения.

21.1. Аффинные отображения в координатах.

Пусть даны аффинное пространство $\mathbb{A}$, ассоциированное с линейным пространством V над полем F, и аффинное пространство $\mathbb{B}$, ассоциированное с линейным пространством W над тем же полем F.

Напомним, что отображение $\Phi : \mathbb{A} \to \mathbb{B}$ называется **аффинным (аффинно-линейным)**, если существует линейное отображение $\varphi : V \to W$ такое, что для любых точек $a, b \in \mathbb{A}$ выполняется

$$\overline{\Phi(a)\Phi(b)} = \varphi(\overline{ab}).$$

При этом φ называется **линейной частью** аффинного отображения Φ.

Аффинное отображение можно записать в виде $\Phi(b) = \Phi(a) + \varphi(\overline{ab})$. Таким образом, достаточно знать аффинное отображение в любой фиксированной точке и соответствующее ему линейное отображение.

Замечание 21.1. *Можно доказать, что композиция аффинного отображения $\Phi_1 : \mathbb{A} \to \mathbb{A}_1$ с линейной частью $\varphi_1 : V \to V_1$ и аффинного отображения $\Phi_2 : \mathbb{A}_1 \to \mathbb{A}_2$ с линейной частью $\varphi_2 : V_1 \to V_2$ есть аффинное отображение $\Phi_2 \circ \Phi_1$ с линейной частью $\varphi_2 \circ \varphi_1$. Проверка непосредственна.*

Пусть $(a_0, \mathcal{E})$ – некоторая аффинная система координат в n-мерном аффинном пространстве $\mathbb{A}$, $(b_0, \mathcal{F})$ – некоторая аффинная система координат в m-мерном аффинном пространстве $\mathbb{B}$. Тогда **матрицей аффинно-линейного отображения** $\Phi : \mathbb{A} \to \mathbb{B}$ с линейной частью $\varphi : V \to W$ в системах координат $(a_0, \mathcal{E})$ и $(b_0, \mathcal{F})$ будем называть блочную матрицу

$$\widehat{A} = \begin{pmatrix} A & X_0 \\ 0 & 1 \end{pmatrix},$$

где A – матрица линейного отображения φ в базисах $\mathcal{E}$ и $\mathcal{F}$, X_0 – столбец координат точки $\Phi(a_0)$ в системе координат $(b_0, \mathcal{F})$. **Рангом аффинного отображения** будем называть ранг его линейной части.

Если $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ – столбец координат некоторой точки $a \in \mathbb{A}$ в $(a_0, \mathcal{E})$, то X – это столбец координат вектора $\overline{a_0a}$ в базисе $\mathcal{E}$. Из матричной формы записи (4) для линейного отображения AX – столбец координат вектора $\varphi(\overline{a_0a})$ в базисе $\mathcal{F}$. Имеем: $\Phi(a) = \Phi(a_0) + \varphi(\overline{a_0a}) = b_0 + \overline{b_0\Phi(a_0)} + \varphi(\overline{a_0a})$. Таким образом, если $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$ – столбец координат точки $\Phi(a)$ в системе координат $(b_0, \mathcal{F})$, то получаем запись аффинного отображения в матричной форме

$$Y = X_0 + AX$$

или в координатах

$$y_i = x_i^0 + \sum_{j=1}^{n} a_{i,j} x_j \qquad (i = 1, \dots, m),$$

где $A = (a_{i,j})$, $X_0 = \begin{pmatrix} x_1^0 \\ \vdots \\ x_m^0 \end{pmatrix}$.

Будем использовать такое обозначение: если X – столбец координат некоторой точки, то полагаем $\widehat{X} = \begin{pmatrix} X \\ 1 \end{pmatrix}$ (*увеличенный* столбец координат).

В предыдущих обозначениях получаем

Утверждение 21.1. *Увеличенный столбец $\widehat{Y}$ координат точки $\Phi(a)$ выражается через увеличенный столбец координат $\widehat{X}$ точки a по формуле*

$$\widehat{Y} = \widehat{A}\widehat{X}.$$

Доказательство. Перемножая блочные матрицы и используя равенство $Y = X_0 + AX$, получаем

$$\widehat{A}\widehat{X} = \begin{pmatrix} A & X_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ 1 \end{pmatrix} = \begin{pmatrix} AX + X_0 \\ 1 \end{pmatrix} = \begin{pmatrix} Y \\ 1 \end{pmatrix} = \widehat{Y}.$$

■

Как и в случае линейных отображений, справедливо следующее.

Утверждение 21.2. *При замене системы координат с матрицей $\widehat{T}$ в аффинном пространстве $\mathbb{A}$ и матрицей $\widehat{S}$ в аффинном пространстве $\mathbb{B}$ матрица аффинного отображения меняется по формуле $\widehat{S}^{-1}\widehat{A}\widehat{T}$.*

21.2. Аффинные операторы.

Более подробно рассмотрим аффинные отображения в себя.

Определение 21.1. Аффинным оператором *(или* **аффинным преобразованием***) называется аффинное отображение $\Phi : \mathbb{A} \to \mathbb{A}$ аффинного пространства $\mathbb{A}$ в себя.*

ПРИМЕР. *Гомотетия* с центром в точке o и коэффициентом $\lambda \in F$ есть аффинное преобразование, заданное формулой

$$\Phi(o + v) = o + \lambda v.$$

Ясно, что ее линейная часть равна $\lambda\mathcal{I}$. Гомотетия с $\lambda = -1$ называется центральной симметрией.

ЗАДАЧА. Докажите, что всякое аффинное преобразование Φ с линейной частью $\lambda\mathcal{I}$ ($\lambda \neq 1$) есть гомотетия с центром в некоторой точке.

Для получения матричной и координатной записи аффинного оператора используются не две, а одна система координат.

Пусть $(o, \mathcal{E})$ – некоторая аффинная система координат в аффинном пространстве $\mathbb{A}$, ассоциированном с n-мерным линейным пространством V над полем F. И пусть $X = (x_1, ..., x_n)^t$ – координаты произвольной точки $a \in \mathbb{A}$, $Y = (y_1, ..., y_n)^t$ – координаты ее образа $\Phi(a)$, $X^0 = (x_1^0, ..., x_n^0)^t$ – координаты образа начала координат $\Phi(o)$ в системе координат $(o, \mathcal{E})$.

Тогда аффинный оператор $\Phi : \mathbb{A} \to \mathbb{A}$ в матричной форме записывается следующим образом

$$Y = AX + X^0,$$

где $A = (a_{i,j})$ – это матрица линейной части $\varphi : V \to V$ в базисе $\mathcal{E}$. Это равенство в координатной форме перепишется так:

$$\begin{cases} y_1 = a_{1,1}x_1 + \ldots + a_{1,n}x_n + x_1^0, \\ \quad . \quad . \quad . \\ y_n = a_{n,1}x_1 + \ldots + a_{n,n}x_n + x_n^0. \end{cases}$$

Назовем аффинный оператор $\Phi : \mathbb{A} \to \mathbb{A}$ **невырожденным**, если его ранг равен $\dim \mathbb{A}$. Как и в случае линейных операторов, аффинный оператор является невырожденным тогда и только тогда, когда он обратим. Заметим, что для невырожденного аффинного оператора Φ такого, что $\Phi(a + v) = \Phi(a) + \varphi(v)$, обратным будет $\Phi^{-1}(\Phi(a) + v) = a + \varphi^{-1}(v)$.

Множество всех невырожденных аффинных операторов аффинного пространства $\mathbb{A}$ является группой, называемой **(полной) аффинной группой** пространства $\mathbb{A}$ и обозначаемой $GA(\mathbb{A})$.

Среди всех невырожденных аффинных операторов выделяются сдвиги и операторы с неподвижной точкой.

Определение 21.2. **Сдвигом на вектор** $v \in V$ *(или* **параллельным переносом***) называется отображение* $t_v : \mathbb{A} \to \mathbb{A}$*, заданное по правилу* $t_v(a) = a + v$ *для любого* $a \in \mathbb{A}$*.*

Утверждение 21.3. *1) Сдвиг на вектор – это аффинный оператор с тождественной линейной частью.*

2) Любой аффинный оператор с тождественной линейной частью есть сдвиг на некоторый вектор.

3) Множество $Tran(\mathbb{A})$ *всех сдвигов аффинного пространства* $\mathbb{A}$ *над* V *есть подгруппа в группе* $GA(\mathbb{A})$*, изоморфная аддитивной группе векторного пространства* V*.*

Доказательство.

1) Для доказательства аффинности сдвига рассмотрим

$$t_v(a) + \overline{t_v(a)t_v(b)} = (a+v) + \overline{(a+v)(b+v)} =$$

$$= b + v = a + \overline{ab} + v = a + v + \overline{ab} = t_v(a) + \overline{ab}.$$

Отсюда $\overline{t_v(a)t_v(b)} = \overline{ab} = \mathcal{I}(\overline{ab})$. Таким образом, t_v – аффинный оператор с тождественной линейной частью $\mathcal{I}$.

2) Пусть $\Phi : \mathbb{A} \to \mathbb{A}$ – аффинный оператор с тождественной линейной частью, $a \in \mathbb{A}$ – произвольная точка. Рассмотрим вектор $v = \overline{a\Phi(a)}$. Тогда $\Phi = t_v$. В самом деле, для любой точки $b \in \mathbb{A}$ имеем

$$\Phi(b) = \Phi(a) + \mathcal{I}(\overline{ab}) = a + v + \overline{ab} = b + v.$$

3) Из того, что линейная часть произведения аффинных преобразований равна произведению линейных частей, и равенства $\mathcal{I}\mathcal{I} = \mathcal{I}$ вытекает, что множество сдвигов замкнуто относительно взятия произведения и обратного элемента, т.е. является подгруппой.

В явном виде: $t_v t_u = t_{v+u}$, $t_v^{-1} = t_{-v}$, t_0 – тождественный аффинный оператор.

Сопоставляя сдвигу t_v вектор v, получаем отображение, сохраняющее операции и являющееся биекцией. Таким образом, $Tran(\mathbb{A}) \cong V$. ∎

Зафиксируем некоторую точку $a \in \mathbb{A}$ и рассмотрим множество G_a всех аффинных операторов из $GA(\mathbb{A})$, оставляющих a на месте. Тогда для любых аффинных операторов $\Phi, \Psi \in G_a$ имеем $(\Phi\Psi)(a) = \Phi(\Psi(a)) = a$ и $\Phi^{-1}(a) = a$. Таким образом, G_a – подгруппа в группе $GA(\mathbb{A})$. Заметим, что $Tran(\mathbb{A}) \cap G_a$ состоит только из тождественного аффинного оператора, так как это единственный сдвиг с неподвижной точкой.

Теорема 21.1. *Пусть $\mathbb{A}$ – аффинное пространство над векторным пространством V, a – произвольная фиксированная точка из $\mathbb{A}$. Тогда каждый аффинный оператор $\Phi \in GA(\mathbb{A})$ (однозначно) представляется в виде композиции*

$$\Phi = t_v \Psi_a \quad \text{и} \quad \Phi = \Psi_a t_w,$$

где $v, w \in V$ и $\Psi_a \in G_a$, т.е. имеет место разложение

$$GA(\mathbb{A}) = G_a Tran(\mathbb{A}) = Tran(\mathbb{A}) G_a.$$

Доказательство. Рассмотрим произвольный невырожденный аффинный оператор Φ с линейной частью φ.

Положим

$$v = \overline{a\Phi(a)}.$$

Определим аффинный оператор Ψ по формуле $\Psi = t_{-v}\Phi$. Аффинный оператор Ψ является невырожденным и оставляет точку a на месте:

$$\Psi(a) = t_{-v}(\Phi(a)) = \Phi(a) - v = a.$$

Следовательно, $\Psi \in G_a$ и $\Phi = (t_{-v})^{-1}\Psi = t_v\Psi$, как и требовалось. Отметим, что линейная часть для Ψ равна φ, так как Ψ есть композиция аффинных операторов t_{-v} и Φ, а значит, линейная часть для Ψ есть композиция линейной части сдвига t_{-v} и линейной части для Φ, т.е. композиция тождественного линейного оператора и линейного оператор φ.

Чтобы получить второе разложение для Φ, представим Φ в виде

$$\Phi = \Psi(\Psi^{-1}t_v\Psi).$$

Линейная часть аффинного оператора $\Psi^{-1}t_v\Psi$ равна произведению линейных частей сомножителей $\varphi^{-1}\mathcal{I}\varphi = \mathcal{I}$. Поэтому аффинный оператор $\Psi^{-1}t_v\Psi$ является сдвигом t_u на некоторый вектор u, т.е. имеем разложение

$$\Phi = \Psi t_u.$$

Для нахождения явного вида для u запишем

$$a + u = t_u(a) = (\Psi^{-1}t_v\Psi)(a) = \Psi^{-1}(t_v(\Psi(a))) = \Psi^{-1}(t_v(a)) =$$
$$= \Psi^{-1}(a+v) = \Psi^{-1}(a) + \varphi^{-1}(v) = a + \varphi^{-1}(v).$$

Отсюда получаем, что

$$u = \varphi^{-1}(v).$$

Осталось доказать однозначность разложения. Если бы существовало два разложения $\Phi = t_v\Psi$ и $\Phi = t_{v'}\Psi'$, то

$$t_v\Psi = t_{v'}\Psi'.$$

Умножая слева обе части равенства на обратное к $t_{v'}$, а справа на Ψ^{-1}, получим

$$t_{v-v'} = \Psi'\Psi^{-1},$$

т.е. отображение $\Psi'\Psi^{-1}$ с неподвижной точкой равно сдвигу $t_{v-v'}$. Как отмечалось, сдвиг с неподвижной точкой тождествен, значит,

$$t_v = t_{v'}, \Psi = \Psi'.$$

Аналогично доказывается однозначность разложения $\Phi = \Psi t_u$. ■

В матричном виде имеем:

$$\begin{pmatrix} E & X_0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} A & X_0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} E & A^{-1}X_0 \\ 0 & 1 \end{pmatrix}.$$

Определение 21.3. *Пусть $S \subseteq \mathbb{A}$ – некоторое множество точек в аффинном пространстве $\mathbb{A}$. Наименьшая плоскость, содержащая S, называется* **аффинной оболочкой** *S и обозначается $\langle S\rangle$.*

Если $S = \{s_0, s_1, ..., s_m\}$ – некоторое множество точек из $\mathbb{A}$, то $\langle S \rangle = s_0 + \langle \overline{s_0 s_1}, ..., \overline{s_0 s_m} \rangle$.

Определение 21.4. *Множество из $m+1$ точек $s_0, s_1, ..., s_m$ в аффинном пространстве $\mathbb{A}$ называется* **аффинно независимым**, *если они не лежат в плоскости размерности меньше, чем m; в этом случае говорят также, что точки $s_0, s_1, ..., s_m$ находятся* **в общем положении.**

Точки $s_0, s_1, ..., s_m \in \mathbb{A}$ являются аффинно независимыми тогда и только тогда, когда $\dim \langle s_0, s_1, ..., s_m \rangle = m$, т.е. векторы $\overline{s_0 s_1}, ..., \overline{s_0 s_m}$ линейно независимы.

Теорема 21.2. *Если $\{a_0, a_1, ..., a_n\}$ и $\{b_0, b_1, ..., b_n\}$ – две системы аффинно независимых точек в n-мерном аффинном пространстве $\mathbb{A}$, то существует единственное аффинное преобразование $\Phi : \mathbb{A} \to \mathbb{A}$, переводящее a_i в b_i ($i = 0, 1, ..., n$).*

Доказательство. Существует единственный линейный оператор $\varphi : V \to V$ пространства V, переводящее базис $\{\overline{a_0 a_1}, ..., \overline{a_0 a_n}\}$ в базис $\{\overline{b_0 b_1}, ..., \overline{b_0 b_n}\}$. Тогда искомое аффинное преобразование Φ получается по правилу

$$\Phi(a_0 + v) = b_0 + \varphi(v),$$

где v – произвольный вектор линейного пространства V. ■

Группа аффинных преобразований определяет аффинную геометрию в том смысле, что задачей аффинной геометрии является изучение свойств фигур, инвариантных при (биективных) аффинных преобразованиях. Например, так как при таких преобразованиях любая плоскость переходит в плоскость той же размерности, то понятие плоскости относится к числу понятий аффинной геометрии. Но, например, как показывает последняя теорема, в аффинной геометрии не существует понятия расстояния между точками, так как любую пару различных точек с помощью аффинного преобразования можно перевести в любую другую пару различных точек.

Список литературы

[1] **Винберг Э.Б.,** *Курс алгебры:* Научное издание М.: "Факториал"1999 528.

[2] **Кострикин А.И.,** *Введение в алгебру, часть II.* Физматлит Минобр 2004.

[3] **Гельфанд И.М.,** *Лекции по линейной алгебре:* Наука 1966 280.

[4] **Беклемишев Д.В.,** *Курс аналитической геометрии и линейной алгебры:* Учеб. для вузов. Физматлит Минобр 2002 374.

Содержание