ЛИНЕЙНАЯ АЛГЕБРА

Учебно-методическое пособие

Линейная алгебра : учеб.-метод. пособие [Текст]. – 149 с.

В пособии рассматриваются следующие разделы линейной алгебры: теория множеств, матрицы и определители, системы линейных алгебраических уравнений, векторные пространства, линейные операторы, комплексные числа. Теоретический материал изложен в доступной форме, но с сохранением необходимого уровня строгости изложения, сопровождается большим количеством примеров и решением типовых задач.

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

В настоящем учебном пособии рассматриваются элементы следующих разделов алгебры: теория множеств, матрицы и определители, системы линейных алгебраических уравнений, комплексные числа, алгебраические системы, векторные пространства, линейные операторы.

Преподавание высшей математики студентам-бакалаврам имеет свои особенности в силу, прежде всего, других по количеству часов учебных планов. Существующие учебники по линейной алгебре очень объемны, и поэтому студент (бакалавр) не всегда может найти в них нужную информацию, вычленить главное. Данное учебно-методическое пособие поможет обучающемся в лучшем усвоении и закреплении теоретического и практического материала по линейной алгебре.

Для краткой записи утверждений будем использовать следующие обозначения символов:

$\forall$ (квантор общности) читается «для любого», «для каждого», «для всех»;

$\exists$ (квантор существования) – «найдется», «существует», «хотя бы для одного»;

$\Rightarrow$ (импликация, знак логического следования) – «если …, то …», «следует»;

$\Leftrightarrow$ (эквивалентность, знак логической равносильности) – «тогда и только тогда».

1. Множества

1.1. Множества и их элементы. Способы задания множеств

Первичным понятием теории множества является понятие самого множества. Данный термин был введен в математику создателем теории множеств Г. Кантором[1]. Следуя ему, под *множеством* понимается совокупность объектов произвольной природы, которая рассматривается как единое целое. Объекты, входящие в состав множества, называются его *элементами*.

Это описание понятия множества нельзя считать логическим определением, а всего лишь пояснением. Понятие множества принимается как исходное, первичное, то есть не сводимое к другим понятиям.

Примерами множеств могут служить множество всех книг, составляющих данную библиотеку, множество всех точек данной линии, множество всех решений данного уравнения, множество всех одноклеточных организмов и т. п.

Множества принято обозначать прописными буквами латинского алфавита: A, B, C, … Обозначается множество скобками $\{\ldots\}$, внутри которых либо просто перечисляются элементы, либо описываются их свойства. Для числовых множеств будем использовать следующие обозначения:

$\mathbb{N}$ – множество натуральных чисел;

$\mathbb{N}_0$ – множество неотрицательных целых чисел;

$\mathbb{Z}$ – множество целых чисел;

$\mathbb{Q}$ – множество рациональных чисел;

$\mathbb{I}$ – множество иррациональных чисел;

$\mathbb{R}$ – множество действительных чисел;

$\mathbb{C}$ – множество комплексных чисел.

---

[1] Георг Кантор (1845–1918) – немецкий математик.

Элементы множества будем обозначать строчными латинскими буквами: a, b, c, …

Предложения вида «объект a есть элемент множества A», «объект a принадлежит множеству A», имеющие один и тот же смысл, кратко записывают в виде $a \in A$. Если элемент a не принадлежит множеству A, то пишут $a \notin A$. Символ $\in$ называется *знаком принадлежности*.

Множества могут содержать как конечное число элементов, так и бесконечное. Например, множество всех корней уравнения $x^2 - 3x + 2 = 0$ конечно (два элемента), а множество всех точек прямой бесконечно. Рассматривают в математике и множество, не содержащее ни одного элемента.

Определение 1.1. Множество, не содержащее ни одного элемента, называется *пустым* и обозначается символом $\varnothing$.

Число элементов конечного множества называется его *мощностью*. Если множество A содержит n элементов, то будем писать $|A| = n$. Если множество $A = \varnothing$, то $|A| = 0$. Мощность бесконечного множества является более сложным понятием и изучается в дискретной математике и в числовых системах.

Замечание 1.1. Элементами множества могут быть множества. Например, можно говорить о множестве групп некоторого факультета университета. Элементы этого множества – группы, являющиеся в свою очередь множествами студентов. Но конкретный студент одной из групп уже не является элементом множества групп факультета.

Определение 1.2. Множество, элементами которого являются другие множества, называется *семейством* (или *классом*).

Определение 1.3. Если все элементы данной совокупности множеств принадлежат некоторому одному множеству, то такое множество называется *универсальным множеством*, и обозначается U.

Множество считают заданным, если о любом объекте можно сказать, принадлежит он этому множеству или не принадлежит. *Множество* *можно задать следующими способами:*

1. перечислением всех его элементов;
2. характеристическим свойством элементов множества;
3. порождающей процедурой.

Первый способ задания множеств применим только для конечных множеств, да и то при условии, что число элементов множества невелико. Если a, b, c, d – обозначения различных объектов, то множество A этих объектов записывают так: $A = \{a, b, c, d\}$. Запись читают: «A – множество, элементы которого a, b, c, d».

Замечание 1.2. Порядок перечисления элементов множества не имеет значения. Например, множества $\{m, n, k, r\}$ и $\{n, m, r, k\}$ совпадают.

Вторым способом можно задавать как конечные, так и бесконечные множества. *Характеристическое свойство* – это такое свойство, которым обладает каждый элемент, входящий в данное множество, и не обладает ни один элемент, который ему не принадлежит. Если обозначить символом $P(a)$ характеристическое свойство элементов множества A, то тогда используется следующая запись: $A = \{a \mid P(a)\}$.

Порождающая процедура описывает способ получения элементов нового множества из уже полученных элементов или из других объектов. Тогда элементами множества считаются все объекты, которые могут быть получены с помощью этой процедуры. Другими словами, порождающая процедура – это процесс, который будучи запущен, порождает все элементы данного множества. С помощью порождающей процедуры можно задавать множества, содержащие любое число элементов.

Пример 1.1. Определим различными способами множество M всех нечетных натуральных чисел, не превышающих 10:

1. $M = \{1, 3, 5, 7, 9\}$;
2. $M = \{m \mid m \in \mathbb{N}, m < 10, m$ – нечетное число$\}$ или

   $M = \{2n - 1 \mid n \in \mathbb{N}, n \leq 5\}$;
3. порождающая процедура определяется правилами:

a) $1 \in M$;

b) если $m \in M$, то $(m + 2) \in M$, где $m \leq 7$.

1.2. Подмножества. Диаграммы Эйлера – Венна

Определение 1.4. Множество B называется *подмножеством* множества A, если каждый элемент множества B принадлежит множеству A.

Пример 1.2. Пусть $A = \{1, 2, 3, 4, 5, 6, 7\}$, а $B = \{2, 3, 5, 7\}$. Множество B является подмножеством множества A, поскольку каждый элемент множества B принадлежит множеству A.

Если множество B является подмножеством множества A, то говорят также, что B содержится в A или B включено в A, при этом пишут $B \subseteq A$ или $A \supseteq B$. Символ $\subseteq$ называется *знаком включения* (точнее, нестрого включения).

Согласно данному определению 1.4 подмножества, каждое множество является подмножеством самого себя, то есть $(\forall A)\ A \subseteq A$. Кроме того, считается, что пустое множество есть подмножество любого множества A: $(\forall A)\ \varnothing \subseteq A$.

Различают два вида подмножеств множества A.

Определение 1.5. Пустое множество $\varnothing$ и множество A называются *несобственными подмножествами* множества A.

Определение 1.6. Любые подмножества множества A, отличные от A и $\varnothing$, называются *собственными подмножествами* множества A.

Определение 1.7. Множества A и B, состоящие из одних и тех же элементов, называются *равными*. При этом пишут $A = B$, в противном случае $A \neq B$.

Справедливо следующее утверждение, которое также можно рассматривать в качестве определения равных множеств.

Утверждение 1.1. $A = B \Leftrightarrow A \subseteq B$ и $B \subseteq A$.

Замечание 1.3. Из утверждения 1.1 вытекает *способ доказательства равенства двух множеств*: если доказать, что каждый элемент из множества A является элементом множества B и каждый элемент из множества B является элементом множества A, то делают вывод, что $A = B$.

Говорят, что множество B *строго включено* в множество A или, по-другому, A строго включает B, если $B \subseteq A$ и $B \neq A$. В этом случае пишут $B \subset A$. Символ $\subset$ называется *знаком строгого включения*.

Пример 1.3. Имеют место следующие строгие включения числовых множеств: $\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ и $\mathbb{I} \subset \mathbb{R} \subset \mathbb{C}$.

Определение 1.8. Совокупность всех подмножеств множества A называется его *булеаном* (или *множеством-степенью*), и обозначается через $P(A)$ (или 2^A).

Пример 1.4. Если $A = \{a, b, c\}$, то булеан множества A это множество $P(A) = \{\varnothing, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}$.

Для наглядного изображения множеств и их свойств используют так называемые диаграммы Эйлера[2] – Венна[3]. Множество отождествляется с множеством точек на плоскости, лежащих внутри некоторых замкнутых кривых, например окружностей (так называемые круги Эйлера). В частности, универсальное множество U изображается множеством точек некоторого прямоугольника или всей плоскости (рис. 1.1).

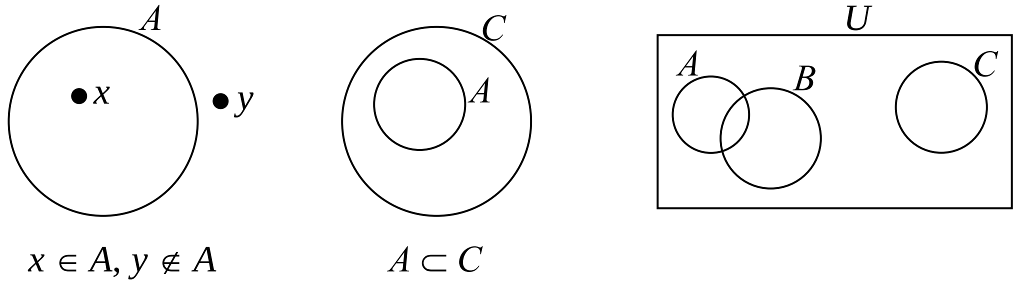

Рис. 1.1

[2] Леонард Эйлер (1707–1783) – швейцарский, немецкий и российский математик и механик.

[3] Джон Венн (1834–1923) – английский логик и философ.

1.3. Операции над множествами и их свойства

Определим операции над множествами, с помощью которых можно получать из любых имеющихся множеств новые множества.

1. Объединение (или сумма).

Определение 1.9. *Объединением* множеств A и B называется множество $A \cup B$, состоящее из тех и только тех элементов, которые принадлежат хотя бы одному из этих множеств.

То есть, по определению 1.9, $A \cup B = \{x \mid x \in A$ <u>или</u> $x \in B\}$.

Все операции над множествами можно иллюстрировать с помощью диаграмм Эйлера – Венна. Объединение множеств A и B заштриховано и изображено на рис. 1.2.

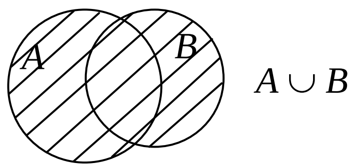

Рис. 1.2

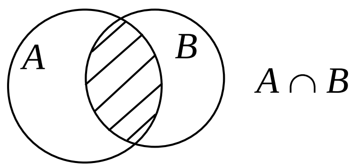

Рис. 1.3

Заметим, что в объединение двух множеств A и B могут входить элементы из A, не принадлежащие множеству B, элементы из B, не принадлежащие множеству A, и элементы, принадлежащие множествам A и B одновременно. Следовательно, $(\forall\, A, B)\ A \subseteq A \cup B$ и $B \subseteq A \cup B$.

2. Пересечение (или произведение).

Определение 1.10. *Пересечением* множеств A и B называется множество $A \cap B$, которое состоит из тех и только тех элементов, принадлежащих одновременно и множеству A и множеству B.

Таким образом, по определению 1. 10, $A \cap B = \{x \mid x \in A$ <u>и</u> $x \in B\}$. Пересечение множеств A и B заштриховано и изображено на рис. 1.3.

Замечание 1.4. Если $A \cap B \neq \varnothing$, то говорят, что множества A и B *пересекаются*. Если $A \cap B = \varnothing$, то в этом случае множества A и B называются *непересекающимися*.

Из определения пересечения следует, что $(\forall\, A, B)\ A \cap B \subseteq A$ и

$A \cap B \subseteq B$.

3. Разность.

Определение 1.11. *Разностью* множеств A и B называется множество $A \setminus B$, состоящее из тех и только тех элементов, которые принадлежат множеству A и одновременно не принадлежат множеству B.

Таким образом, по определению 1.11, $A \setminus B = \{x \mid x \in A \text{ и } x \notin B\}$. Разность множеств A и B заштриховано и изображено на рис. 1.4.

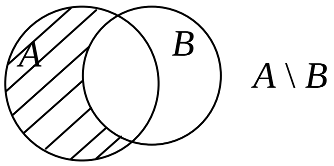

Рис. 1.4

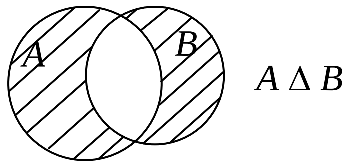

Рис. 1.5

Замечание 1.5. Если $B \subseteq A$, то в этом случае разность $A \setminus B$ называют *дополнением B до A*.

Определим частные случаи разности.

Определение 1.12. *Симметрической разностью* (или *кольцевой суммой*) множеств A и B называется множество $A \Delta B$ (или $A \oplus B$), состоящее из элементов объединения этих множеств, но не входящих в пересечение этих множеств (рис. 1.5).

Таким образом, по определению, $A \Delta B = (A \cup B) \setminus (A \cap B) =$ $= (A \setminus B) \cup (B \setminus A) = \{x \mid (x \in A \text{ и } x \notin B) \text{ или } (x \in B \text{ и } x \notin A)\}$.

Определение 1.13. *Дополнением* множества A (до универсального множества U) называется множество $\overline{A}$ (или A') равное разности $U \setminus A$.

Дополнение $\overline{A}$ множества A до универсального множества U заштриховано и изображено на рис. 1.6.

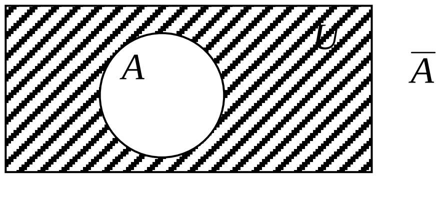

Рис. 1.6

Таким образом, по определению, $\overline{A} = U \setminus A = \{x \mid x \in U \text{ и } x \notin A\}$ или

$\overline{A} = \{x \mid x \notin A\}$.

Пример 1.5. Пусть $A = \{m, n, p, k, l\}$, $B = \{p, r, s, n\}$. Найти: $A \cup B$, $A \cap B$, $A \setminus B$, $B \setminus A$, $A \Delta B$.

Решение.

$A \cup B = \{m, n, p, k, l, r, s\}$; $A \cap B = \{p, n\}$; $A \setminus B = \{m, k, l\}$; $B \setminus A = \{r, s\}$;

$A \Delta B = \{ m, k, l, r, s\}$.

Пример 1.6. Пусть $A = \{x \mid x \in \mathbb{R}, -4 \le x < 1\}$, $B = \{x \mid x \in \mathbb{R}, 0 \le x \le 4\}$. Найти: $A \cup B$, $A \cap B$, $A \setminus B$, $B \setminus A$, $A \Delta B$, $\overline{A}$, $\overline{B}$.

Решение.

$A \cup B = \{x \mid x \in \mathbb{R}, -4 \le x \le 4\}$;

$A \cap B = \{x \mid x \in \mathbb{R}, 0 \le x < 1\}$;

$A \setminus B = \{x \mid x \in \mathbb{R}, -4 \le x < 0\}$;

$B \setminus A = \{x \mid x \in \mathbb{R}, 1 \le x \le 4\}$;

$A \Delta B = \{x \mid x \in \mathbb{R}, -4 \le x < 0, 1 \le x \le 4\}$;

$\overline{A} = \{x \mid x \in \mathbb{R}, x < -4, x \ge 1\}$;

$\overline{B} = \{x \mid x \in \mathbb{R}, x < 0, x > 4\}$.

4. Декартовое произведение (или прямое произведение).

Определение 1.14. *Упорядоченной парой* (или *парой*) (a, b) называется два элемента a, b взятые в определенном порядке. Пары (a_1, b_1) и (a_2, b_2) *равны* тогда и только тогда, когда $a_1 = a_2$ и $b_1 = b_2$.

Определение 1.15. *Декартовым произведение* множеств A и B называется множество $A \times B$, состоящее из всех упорядоченных пар (a, b), где $a \in A$ и $b \in B$.

То есть, по определению, $A \times B = \{(x, y) \mid x \in A \text{ и } y \in B\}$.

Определение 1.16. Произведение $A \times A$ называется *декартовым квадратом*.

Пример 1.7. Пусть $A = \{a, b, c\}$ и $B = \{5, 6\}$. Тогда

$A \times B = \{(a, 5), (a, 6), (b, 5), (b, 6), (c, 5), (c, 6)\}$;

$B \times A = \{(5, a), (6, a), (5, b), (6, b), (5, c), (6, c)\}$;

$A \times A = \{(a, a), (a, b), (a, c), (b, a), (b, b), (b, c), (c, a), (c, b) (c, c),\}$;

$B \times B = \{(5, 5), (6, 6), (5, 6), (6, 5)\}$.

Рассмотрим *геометрическую интерпретацию прямого произведения двух числовых множеств* A и B – множество всех точек координатной плоскости Oxy с координатами (x, y) такими, что $x \in A$, а $y \in B$. Тогда для двух заданных числовых множеств можно наглядно изображать их прямое произведение и, обратно, по изображению прямого произведения двух множеств определять их элементы.

Пример 1.8. Изобразить на координатной плоскости Oxy множество $A \times B$, если:

а) $A = \{3, 5, 7\}$, $B = \{2, 4\}$;

б) $A = \{3, 5, 7\}$, $B = [2; 4]$, то есть $B = \{x \mid x \in \mathbb{R}, 2 \le x \le 4\}$;

в) $A = [3, 7]$, $B = [2; 4]$.

Решение. Множество $A \times B$ изображено на рис. 1.7.

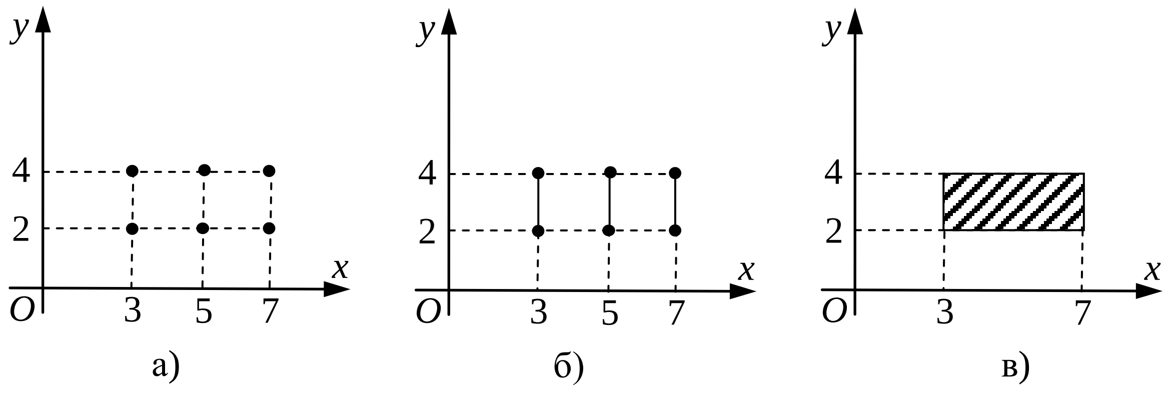

Рис. 1.7

Пример 1.9. По изображению прямого произведения $A \times B$ (рис. 1.8) найти множества A и B.

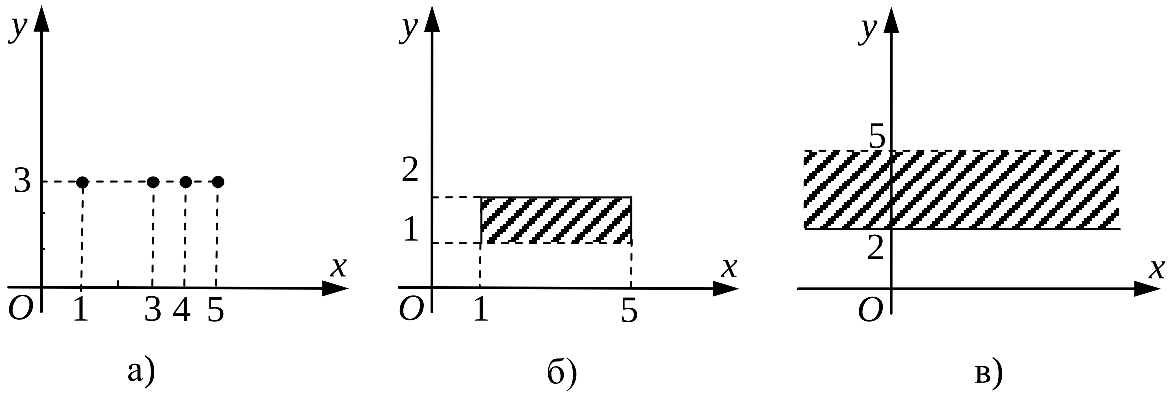

Рис. 1.8

Решение.

а) $A = \{1, 3, 4, 5\}$, $B = \{3\}$; б) $A = [1;5]$, $B = (1;2]$; в) $A = \mathbb{R}$, $B = [2;5)$.

Пример 1.10. Пусть $A = \{x \mid x \in \mathbb{N},\ 2 < x \le 6\}$, $B = \{x \mid x \in \mathbb{N},\ 1 < x < 4\}$, $C = \{x \mid x \in \mathbb{N},\ x^2 - 4 = 0\}$. Из каких элементов состоят множества $(A \cap B) \cup (B \cup C)$, $C \times B$, $B \times C$?

Решение. Перепишем множества A, B и C, перечислив их элементы: $A = \{3, 4, 5, 6\}$, $B = \{2, 3\}$, $C = \{2\}$. Тогда $A \cap B = \{3\}$, $B \cup C = \{2, 3\}$, а значит $(A \cap B) \cup (B \cup C) = \{2, 3\}$; $C \times B = \{(2, 2), (2, 3)\}$ и $B \times C = \{(2, 2), (3, 2)\}$.

Свойства операций над множествами

Свойства операций объединения, пересечения и дополнения иногда называют законами алгебры множеств. Перечислим основные свойства операций над множествами. Пусть задано универсальное множество U. Тогда $(\forall\ A, B, C)\ A, B, C \subseteq U$ выполняются следующие свойства:

1. *идемпотентность*: $A \cup A = A$, $A \cap A = A$;

2. *коммутативность*: $A \cup B = B \cup A$, $A \cap B = B \cap A$;

3. *ассоциативность*: $A \cup (B \cup C) = (A \cup B) \cup C$,

$$A \cap (B \cap C) = (A \cap B) \cap C;$$

4. *дистрибутивность*:

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (дистрибутивность $\cup$ относительно $\cap$),

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (дистрибутивность $\cap$ относительно $\cup$);

5. *свойства нуля*: $A \cup \varnothing = A, \quad A \cap \varnothing = \varnothing$;

6. *свойства единицы*: $A \cup U = U, \quad A \cap U = A$;

7. *поглощение*: $(A \cap B) \cup A = A$, $(A \cup B) \cap A = A$;

8. *инволютивность* (свойство двойного дополнения): $\overline{\overline{A}} = A$;

9. *законы де Моргана*[4]: $\overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \overline{A \cap B} = \overline{A} \cup \overline{B}$;

10. *закон включения*: $A \subset B \Leftrightarrow \overline{B} \subset \overline{A}$;

11. *свойства дополнения*: $A \cup \overline{A} = U, \quad A \cap \overline{A} = \varnothing$;

12. *выражение для разности*: $A \setminus B = A \cap \overline{B}$.

Другие соотношения между множествами могут быть выведены на основе вышеприведенных свойств по правилам алгебры логики.

Справедливость каждого из этих свойств можно доказать, используя утверждение 1.1 и замечание 1.3. В качестве примера приведем доказательство дистрибутивности объединения относительно пересечения: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Пусть $X = A \cup (B \cap C)$, $Y = (A \cup B) \cap (A \cup C)$. Надо доказать, что множества X и Y равны, то есть $X \subseteq Y$ и $Y \subseteq X$. Множество $X \subseteq Y$, если каждый элемент множества X принадлежит множеству Y. Пусть $x \in X \Rightarrow$ $(x \in A)$ или $(x \in B \cap C)$ тогда

- если $x \in A$, то $x \in A \cup B$ и $x \in A \cup C$, следовательно, $x \in Y$;
- если $x \in B \cap C$, то $x \in B$ и $x \in C$ $\Rightarrow$ $x \in A \cup B$ и $x \in A \cup C$, следовательно $x \in Y$.

Из произвольности элемента x следует, что $X \subseteq Y$.

Предложим теперь, что $y \in Y$; то есть $y \in (A \cup B) \cap (A \cup C)$, тогда $y \in A \cup B$ и $y \in A \cup C$. Возможны два случая:

- если $y \notin A$, то $y \in B$ и $y \in C$, значит $y \in B \cap C$; следовательно, $y \in A \cup (B \cap C) \Rightarrow y \in X$;

[4] Огастес (Август) де Морган (1806–1871) – шотланский математик и логик.

- если $y \in A$, то $y \in A \cup (B \cap C) = X$.

Из произвольности элемента y вытекает, что $Y \subseteq X$.

Таким образом, получили равенство множеств $X = Y$.

1.4. Метод математической индукции

Метод математической индукции используется для доказательства утверждений, в формулировке которых участвует натуральный параметр n. Метод математической индукции – метод доказательства математических утверждений, основанный на *принципе математической индукции*. Сформулируем этот принцип: утверждение $A(n)$, зависящее от натурального параметра n, считается доказанным, если доказано $A(1)$ и для каждого натурального числа k из предположения, что верно $A(k)$ выведено, что верно $A(k + 1)$.

Доказательство методом математической индукции состоит из трех этапов.

База индукции: проверяем, что $A(n)$ верно при $n = 1$.

Предположение индукции: предполагаем, что $A(k)$ истинно.

Шаг индукции: доказываем, используя предположение, что истинно $A(k + 1)$.

Замечание 1.6. Если требуется доказать утверждение $A(n)$, где $n \in \mathbb{N}_0$, то база индукции начинается с $n = 0$.

Замечание 1.7. Доказательство методом математической индукции можно начинать не с 1, а с любого натурального m. В этом случае утверждение считается истинным при $n \geq m$.

Замечание 1.8. С помощью принципа математической индукции можно давать индукционные определения. При этом для определения понятия $A(n)$ ($n \in \mathbb{N}$), во-первых, задается значение $A(1)$; во-вторых, для любого натурального числа k задается правило получения значения $A(k + 1)$ по числу k и значению $A(k)$.

Пример 1.11. Доказать равенство: 1 + 2 + … + n = $\frac{n(n+1)}{2}$.

Доказательство. Пусть $A(n)$ = 1 + 2 + … + n.

База индукции: для n = 1 имеем верное равенство 1 = $\frac{1(1+1)}{2}$.

Предположение индукции: пусть для $n = k$ имеем верное равенство: 1 + 2 + … + k = $\frac{k(k+1)}{2}$.

Шаг индукции: докажем, что при $n = k + 1$ будет верно равенство: 1 + 2 + … + k + (k + 1) = $\frac{(k+1)((k+1)+1)}{2}$. Преобразуем левую часть этого равенства 1 + 2 + … + k + (k + 1) = $A(k)$ + (k + 1) = $\frac{k(k+1)}{2}$ + (k + 1) = $= \frac{k(k+1)+2(k+1)}{2} = \frac{k^2+3k+2}{2} = \frac{(k+1)(k+2)}{2}$. Получили, что из истинности равенства при $n = k$ (k – произвольное натуральное число) следует его истинность при $n = k + 1$.

По принципу математической индукции утверждение $A(n)$ верно при любом натуральном n.

1.5. Комплексные числа

Понятие числа является одним из основных завоеваний человеческой культуры. Сначала появились натуральные числа $\mathbb{N} = \{1, 2, 3, \ldots, n, \ldots\}$ затем целые $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, рациональные $\mathbb{Q} = \{\frac{m}{n} \mid m \in \mathbb{Z}, n \in \mathbb{N}\}$ (для того чтобы всякое уравнение вида $a \cdot x = b$, где $a \neq 0$ имело решение); потом появились иррациональные числа это было связано с решением квадратных уравнений, например $x^2 = 2$, на множестве рациональных чисел это уравнение не имеет решения. Иррациональные числа – это бесконечные непериодические десятичные дроби, например π,

e, $\sqrt{2}$, $\sqrt[3]{5}$, …. Рациональные числа можно представить конечными или бесконечными периодическими десятичными дробями, например $\frac{1}{5}$ = 0,2; $\frac{1}{3}$ = 0,(3). Рациональные и иррациональные числа образуют множество $\mathbb{R}$ действительных чисел. На этом множестве уравнение $x^2 = 2$ уже имеет два корня $x_1 = \sqrt{2}$ и $x_2 = -\sqrt{2}$. Но действительных чисел оказалось недостаточно для того, чтобы, например решить квадратное уравнение вида $x^2 + 1 = 0$ (т. к. на множестве действительных чисел нет такого числа, квадрат которого отрицателен). Поэтому ввели комплексные числа $\mathbb{C}$. Впервые упоминание о комплексных числах появилось в работах итальянского ученого Кардано[5] в 1545 г., когда он пришел к выражению $\sqrt{-243}$, решая кубическое уравнение $x^3 - 12x + 16 = 0$. Термин «комплексное число» ввел немецкий математики Гаусс[6] в 1831 г. Первоначально комплексные числа называли мнимыми. И только когда датчанин Вессель[7](1799 г.) (независимо от него француз Арган[8] (1806 г.) и немец Гаусс (1832 г.)) дал геометрическое истолкование комплексного числа, они получили признание и нашли широкое применение.

Уравнение вида $x^2 + 1 = 0$ приводит к понятию мнимой единицы. Решая это уравнение, получаем $x^2 = -1$ или $x = \sqrt{-1}$; $\sqrt{-1}$ назвали *мнимой единицей* и обозначили $i = \sqrt{-1}$ или $i^2 = -1$.

Определение 1.17. *Комплексным числом* называется выражение вида $a + b \cdot i$, где $a, b \in \mathbb{R}$, $i = \sqrt{-1}$ – мнимая единица.

---

[5] Джероламо Кардано (1501–1576) – итальянский математик, инженер, философ, медик и астролог.

[6] Иоганн Карл Фридрих Гаусс (1777–1855) – немецкий математик, механик, физик и астроном.

[7] Каспар Вессель (1745–1818) – датско-норвежский математик.

[8] Жан Робер Арган (1768 – 1822) – непрофессиональный математик.

Комплексное число $z = a + b \cdot i$ состоит из двух частей: число $a = \mathrm{Re}z$– называется *действительной частью* z, $b = \mathrm{Im}z$ – *мнимой частью* комплексного числа z.

Используют следующие термины: если $b = 0$, то $a + 0 \cdot i = a$ – действительное число (точнее отождествляют с действительным числом), в частности $0 + 0 \cdot i = 0$; если $a = 0$, то числа вида $b \cdot i$ $(b \neq 0)$ называют *чисто мнимым*. Множество всех комплексных чисел обозначают $\mathbb{C} = \{a + b \cdot i \mid a, b \in \mathbb{R}, i = \sqrt{-1}\}$ при этом $\mathbb{R} \subset \mathbb{C}$.

Определение 1.18. Комплексное число, записанное в виде $z = a + b \cdot i$, называется *алгебраической формой* записи комплексного числа.

Определение 1.18. Два комплексных числа *равны* тогда и только тогда, когда равны их действительные и мнимые части равны, т. е. $a + b \cdot i = c + d \cdot i \Leftrightarrow a + c$ и $b = d$.

Определение 1.19. Комплексные числа вида $a + b \cdot i$ и $a - b \cdot i$ называются *сопряженными*.

Определение 1.20. Комплексные числа вида $a + b \cdot i$ и $-a - b \cdot i$ называются *противоположными*.

Операции над комплексными числами

1. *Сложение*: $(a + bi) + (c + di) = (a + c) + (b + d)i$, при сложении двух комплексных чисел их действительные части и мнимые части складываются.
2. *Вычитание*: $(a + bi) - (c + di) = (a - c) + (b - d)i$.
3. *Умножение*: $(a + bi) \cdot (c + di) = (ac - bd) + (ad + bc)i$, это простое умножение двучлена $a + bi$ на $c + di$ с последующей заменой i^2 на -1.
4. *Деление*: $\dfrac{a+bi}{c+di} = \dfrac{(a+bi)(c-di)}{(c+di)(c-di)} = \dfrac{(ac+bd)+(bc-ad)i}{c^2+d^2} = \dfrac{ac+bd}{c^2+d^2}$ $+ \dfrac{bc-ad}{c^2+d^2} \cdot i$, если $c + di \neq 0$.

Геометрическая интерпретация комплексных чисел

Известно, что отрицательные числа были введены в связи с решением линейных уравнений с одной переменной. В конкретных задачах отрицательный ответ истолковывался как значение направленной величины (положительные или отрицательные температуры, передвижение в противоположных направлениях, прибыль и долг и т. д.). Однако еще в 16 веке многие математики не признавали отрицательных чисел. Только с введением координатной прямой и координатной плоскости отчетливо проявился смысл отрицательных чисел, и они стали такими же «равноправными» и понятными, как и натуральные числа. Аналогично обстоит дело и комплексными числами. Смысл их отчетливо проявляется при введении их геометрической интерпретации.

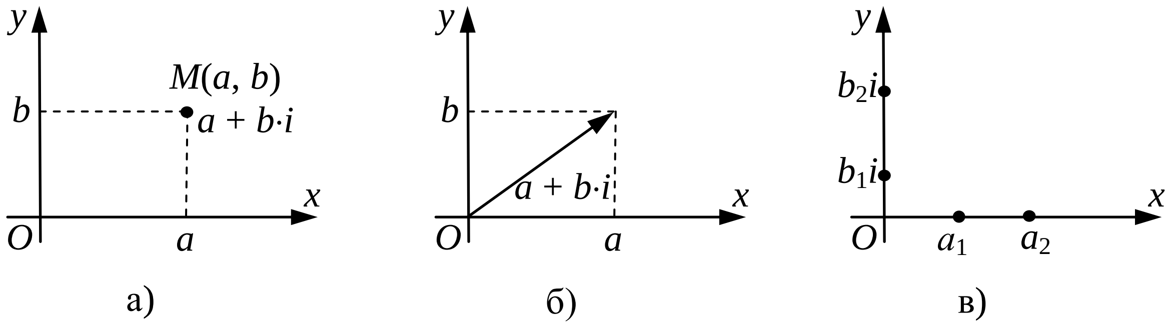

Рис. 1.9

Геометрическое представление комплексных чисел состоит в том, что каждому комплексному числу $z = a + b \cdot i$ ставится в соответствие точка $M(a, b)$ плоскости в прямоугольной системе координат (рис. 1.9, а), таким образом, что действительная часть комплексного числа представляет собой абсциссу, а мнимая – ординату точки. Или вместо точки можно поставить в соответствие вектор с координатами (a, b) (рис. 1.9, б). Таким образом, устанавливается взаимно однозначное соответствие между множеством комплексных чисел и множеством точек координатной плоскости.

Ясно, что действительное число $a + 0 \cdot i = a$ изображается точкой на оси Ox, а чисто мнимое число $0 + b \cdot i$, где $b \neq 0$ – точкой на оси Oy. Поэтому ось Oy называется *мнимой*, а ось Ox – *действительной*.

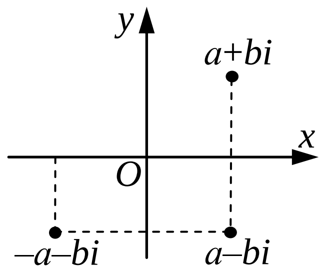

Рис. 1.10

Сопряженные комплексные числа $a + b \cdot i$ и $a - b \cdot i$ изображаются точками симметричными относительно оси Ox, а противоположные комплексные числа $a + b \cdot i$ и $-a - b \cdot i$ – точками симметричными относительно начала координат (рис. 1.10).

Тригонометрическая форма комплексного числа

Вектор $\overrightarrow{OM}$ можно задать не только координатами в прямоугольной системе координат, но и длиной и углом, который он образует с некоторым фиксированным направлением (полярная система координат), т. е. задать вектор $\overrightarrow{OM}$ полярными координатами (рис. 1.11).

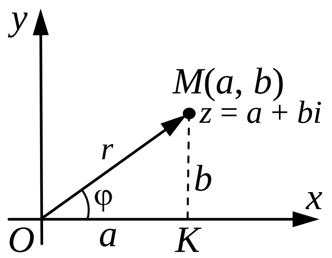

Рис. 1.11

Определение 1.21. Длина вектора, соответствующего комплексному числу z (или расстояние от начала системы координат до точки, изображающей комплексное число) называется *модулем* комплексного числа и обозначается $|z| = r$.

Определение 1.22. Радианная мера угла, образованного этим вектором с положительным направлением действительной оси Ox называется *аргументом* комплексного числа z и обозначается $Arg\ z = \varphi$.

Другими словами, аргумент комплексного числа – это угол между положительной полуосью Ox и лучом Oz.

Число ноль изображается нуль-вектором, для него модуль равен 0, аргумент нуля не определен. Для ненулевого комплексного числа z аргумент определяется с точностью до $2\pi k$, где k – любое целое число.

Главным значение аргумента называется такое значение φ, что $\varphi \in (-\pi, \pi]$. Часто главное значение аргумента обозначается $arg\ z$. Главное значение аргумента обратного комплексного числа отличается знаком аргумента исходного, т. е. $arg\left(\frac{1}{z}\right) = -arg(z)$.

Для комплексного числа $z = a + bi$ установим связь между числами a, b и r, φ, т. е. между декартовыми и полярными координатами.

Рассмотрим прямоугольный треугольник OMK (рис. 1.11), по теореме Пифагора имеем: $OM^2 = OK^2 + MK^2 \Rightarrow r^2 = a^2 + b^2 \Rightarrow r = \sqrt{a^2+b^2}$, $a = r\cdot cos\varphi$, $b = r\cdot sin\varphi$. Тогда $z = a + b\cdot i = r\cdot cos\varphi + r\cdot sin\varphi\cdot i = r(cos\varphi + i\cdot sin\varphi)$.

Определение 1.23. Выражение $z = r(cos\varphi + i\cdot sin\varphi)$ называется *тригонометрической формой* комплексного числа.

Формулы перехода от алгебраической формы комплексного числа $z = a + b\cdot i$ к тригонометрической следующие:

$$r = \sqrt{a^2+b^2},\ sin\varphi = \frac{b}{r} = \frac{b}{\sqrt{a^2+b^2}},\ cos\varphi = \frac{a}{r} = \frac{a}{\sqrt{a^2+b^2}}.$$

Если $a \neq 0$, то $\mathrm{tg}\varphi = \frac{b}{a}$. Можно найти аргумент числа z, пользуясь правилом:

$$a > 0 \Rightarrow \varphi = arctg\frac{b}{a},$$

$$a < 0 \Rightarrow \varphi = arctg\frac{b}{a} + \pi,$$

$$a = 0 \Rightarrow 1)\ b > 0 \Rightarrow \varphi = \frac{\pi}{2},\ 2)\ b < 0 \Rightarrow \varphi = -\frac{\pi}{2}.$$

Определение 1.24. $r_1(cos\varphi_1 + i\cdot sin\varphi_1) = r_2(cos\varphi_2 + i\cdot sin\varphi_2) \Leftrightarrow r_1 = r_2$ и $\varphi_1 = \varphi_2 + 2\pi k$, $k \in \mathbb{Z}$.

Пример 1.12. Представить следующие комплексные числа в тригонометрической форме: $z_1 = \sqrt{3} - i$, $z_2 = -3i$, $z_3 = -8$, $z_4 = -2(cos\frac{\pi}{5} - i\cdot sin\frac{\pi}{5})$, *Решение*. а) $z_1 = \sqrt{3} - i \Rightarrow a = \sqrt{3}$, $b = -1$. Найдем модуль и аргумент данного числа: $r = \sqrt{a^2 + b^2} = \sqrt{(\sqrt{3})^2 + (-1)^2} = 2$; т. к. $a = \sqrt{3} > 0$, то $\varphi = arctg\,\frac{b}{a} = arctg\,\frac{-1}{\sqrt{3}} = -\frac{\pi}{6}$.

Окончательно получаем тригонометрическую форму $z_1 = \sqrt{3} - i =$ $= r(cos\varphi + i{\cdot}sin\varphi) = 2(cos\left(\frac{-\pi}{6}\right) + i{\cdot}sin\left(\frac{-\pi}{6}\right))$.

б) $z_2 = -3i \Rightarrow a = 0$, $b = -3$. Найдем модуль и аргумент данного числа: $r = \sqrt{a^2 + b^2} = \sqrt{(0)^2 + (-3)^2} = 3$; т. к. $a = 0$ и $b < 0$, то $\varphi = -\frac{\pi}{2}$.

Получаем $z_2 = -3i = r(cos\varphi + i{\cdot}sin\varphi) = 3(cos\left(\frac{-\pi}{2}\right) + i{\cdot}sin\left(\frac{-\pi}{2}\right))$.

в) $z_3 = -8 \Rightarrow a = -8$, $b = 0$. Найдем модуль и аргумент данного числа: $r = \sqrt{a^2 + b^2} = \sqrt{(-8)^2 + (0)^2} = 8$; т. к. $a = -8 < 0$, то $\varphi = arctg\,\frac{b}{a} + \pi =$ $= arctg\,\frac{0}{-8} + \pi = arctg\,0 + \pi = 0 + \pi = \pi$.

Тогда $z_3 = -8 = r(cos\varphi + i{\cdot}sin\varphi) = 8(cos\pi + i{\cdot}sin\pi)$.

г) $z_4 = -2(cos\frac{\pi}{5} - i{\cdot}sin\frac{\pi}{5}) = -2cos\frac{\pi}{5} + 2i{\cdot}sin\frac{\pi}{5} \Rightarrow a = -2cos\frac{\pi}{5}$, $b = 2sin\frac{\pi}{5}$. Найдем модуль и аргумент данного числа: $r = \sqrt{a^2 + b^2} = \sqrt{\left(-2\cos\frac{\pi}{5}\right)^2 + \left(2\sin\frac{\pi}{5}\right)^2} = 2$; т. к. $a = -2cos\frac{\pi}{5} < 0$, то

$$\varphi = arctg\,\frac{b}{a} + \pi = arctg\,\frac{2\sin\frac{\pi}{5}}{-2\cos\frac{\pi}{5}} + \pi = -arctg\,(tg\,\frac{\pi}{5}) + \pi = -\frac{\pi}{5} + \pi = \frac{4\pi}{5}.$$

Тогда $z_4 = -2(cos\frac{\pi}{5} - i{\cdot}sin\frac{\pi}{5}) = r(cos\varphi + i{\cdot}sin\varphi) = 2(cos\frac{4\pi}{5} + i{\cdot}sin\frac{4\pi}{5})$.

Действия над комплексными числами в тригонометрической форме

Сложение и вычитание удобнее производить над комплексными числами в алгебраической форме, а умножение и деление – в тригонометрической форме.

1. Умножений. Пусть даны два комплексных числа, записанных в тригонометрической форме: $z_1 = r_1(\cos\varphi_1 + i\cdot\sin\varphi_1)$ $z_2 = r_2(\cos\varphi_2 + i\cdot\sin\varphi_2)$.

$z_1\cdot z_2 = r_1\cdot r_2(\cos\varphi_1\cdot\cos\varphi_2 - \sin\varphi_1\cdot\sin\varphi_2) + i\cdot(\cos\varphi_1\cdot\sin\varphi_2 + \sin\varphi_1\cdot\cos\varphi_2) =$

$= r_1\cdot r_2(\cos(\varphi_1 + \varphi_2) + i\cdot\sin(\varphi_1 + \varphi_2))$.

Итак, модуль $|z_1\cdot z_2| = r_1\cdot r_2$, аргумент $arg(z_1\cdot z_2) = arg\, z_1 + arg\, z_2$.

Пример 1.13. Для $z_1 = 2(\cos\frac{\pi}{6} + i\cdot\sin\frac{\pi}{6})$ и $z_2 = 3(\cos\frac{\pi}{3} + i\cdot\sin\frac{\pi}{3})$ найти их произведение $z_1\cdot z_2$.

Решение. Применяем формулу для нахождения произведения двух комплексных чисел, записанных в тригонометрической форме. $z_1\cdot z_2 = 2\cdot 3(\cos(\frac{\pi}{6} + \frac{\pi}{3}) + i\cdot\sin(\frac{\pi}{6} + \frac{\pi}{3})) = 6(\cos\frac{\pi}{2} + i\cdot\sin\frac{\pi}{2})$ – тригонометрическая форма произведения чисел z_1 и z_2 или в алгебраической форме $z_1\cdot z_2 = 6i$.

2. Деление.

$$\frac{z_1}{z_2} = \frac{r_1(\cos\varphi_1 + i\cdot\sin\varphi_1)}{r_2(\cos\varphi_2 + i\cdot\sin\varphi_2)} = \frac{r_1}{r_2}\cdot\frac{(\cos\varphi_1 + i\cdot\sin\varphi_1)(\cos\varphi_2 - i\cdot\sin\varphi_2)}{\cos^2\varphi_2 - i^2\cdot\sin^2\varphi_2} =$$

$$= \frac{r_1}{r_2}\cdot\frac{(\cos\varphi_1\cdot\cos\varphi_2 + \sin\varphi_1\sin\varphi_2) + i\cdot(-\cos\varphi_1\cdot\sin\varphi_2 + \cos\varphi_2\sin\varphi_1)}{\cos^2\varphi_2 + \sin^2\varphi_2} =$$

$$= \frac{r_1}{r_2}\cdot(\cos(\varphi_1 - \varphi_2) + i\cdot\sin(\varphi_1 - \varphi_2))$$

Итак, модуль $|\frac{z_1}{z_2}| = \frac{r_1}{r_2}$, аргумент $arg(\frac{z_1}{z_2}) = arg\, z_1 - arg\, z_2$.

Пример 1.14. Для $z_1 = 10(cos45° + i{\cdot}sin45°)$ и $z_2 = 5(cos60° + i{\cdot}sin60°)$ найти их частное от деления $\frac{z_1}{z_2}$.

Решение.

$\frac{z_1}{z_2} = \frac{10}{5}(cos(45° - 60°) + i{\cdot}sin(45° - 60°)) = 2(cos(-15°) + i{\cdot}sin(-15°))$ – тригонометрическая форма частного чисел z_1 и z_2. Заметим, что если данное выражение записать в виде равносильного выражения $2(cos15° - i{\cdot}sin15°)$, то это не будет уже тригонометрической формой записи комплексного числа.

3. Возведение в степень.

Если $z = r(cos\varphi + i{\cdot}sin\varphi)$, то $z^n = r^n(cos(n\varphi) + i{\cdot}sin(n\varphi))$, где $n \in \mathbb{Z}$. Данная формула называется формулой Муавра[9].

Пример 1.15. Для $z = \sqrt{3} - i$, найти z^4.

Решение. Воспользуемся формулой Муавра, но для начала надо это комплексное число записать в тригонометрической форме. В примере 1.12 мы это уже находили $z = \sqrt{3} - i = 2(cos\frac{-\pi}{6} + i{\cdot}sin\frac{-\pi}{6})$. Тогда

$z^4 = (\sqrt{3} - i)^4 = (2(cos\frac{-\pi}{6} + i{\cdot}sin\frac{-\pi}{6}))^4 = 2^4(cos\frac{-4\pi}{6} + i{\cdot}sin\frac{-4\pi}{6}) =$

$= 16(cos\frac{-2\pi}{3} + i{\cdot}sin\frac{-2\pi}{3})$ – тригонометрическая форма результата возведения в четвертую степень данного комплексного числа. Найдем также и алгебраическую форму записи числа z^4.

$z^4 = 16(cos\frac{-2\pi}{3} + i{\cdot}sin\frac{-2\pi}{3}) = 16(cos\frac{2\pi}{3} - i{\cdot}sin\frac{2\pi}{3}) = 16(-\frac{1}{2} - i{\cdot}\frac{\sqrt{3}}{2}) =$

$= -8 - 8\sqrt{3}{\cdot}i.$

4. Извлечение корня *n*-ой степени.

[9] Абрахам де Муавр (1667–1754) – английский математик.

Можно показать, что каждое комплексное число, отличное от нуля, имеет ровно n корней n-й степени.

Если $z = r(cos\varphi + i\cdot sin\varphi)$, то

$$\sqrt[n]{z} = \sqrt[n]{r(\cos\varphi + i\cdot\sin\varphi)} =$$

$$\sqrt[n]{r}\,(cos\frac{\varphi+2\pi k}{n} + i\cdot sin\frac{\varphi+2\pi k}{n}),\text{ где } k = 0, 1, \ldots, n-1.$$

Пример 1.16. Найти $\sqrt[4]{16}$.

Решение. Пусть $z = 16$, найдем сначала тригонометрическую форму данного комплексного числа. Имеем $z = 16 \Rightarrow a = 16,\ b = 0 \Rightarrow$ $r = \sqrt{a^2+b^2} = \sqrt{16^2+0^2} = 16$; т. к. $a = 16 > 0$, то $\varphi = arctg\,\frac{b}{a} = arctg\,\frac{0}{16} =$ $= arctg\,0 = 0$. Тогда $z = 16 = r(cos\varphi + i\cdot sin\varphi) = 16(cos0 + i\cdot sin0)$.

Применяем формулу для нахождения корня n-ой степени.

$$\sqrt[4]{z} = \sqrt[4]{16} = \sqrt[4]{16(\cos 0 + i\cdot\sin 0)} = \sqrt[4]{16}\,(cos\frac{0+2\pi k}{4} + i\cdot sin\frac{0+2\pi k}{4}) =$$

$= 2(cos\frac{2\pi k}{4} + i\cdot sin\frac{2\pi k}{4})$, где $k = 0, 1, 2, 3$. Найдем все четыре корня:

$$k = 0 \Rightarrow \omega_0 = 2(cos\frac{2\pi\cdot 0}{4} + i\cdot sin\frac{2\pi\cdot 0}{4}) = 2\,(cos0 + i\cdot sin0) = 2,$$

$$k = 1 \Rightarrow \omega_1 = 2(cos\frac{2\pi\cdot 1}{4} + i\cdot sin\frac{2\pi\cdot 1}{4}) = 2(cos\frac{\pi}{2} + i\cdot sin\frac{\pi}{2}) = 2(0 + i\cdot 1) = 2i,$$

$$k = 2 \Rightarrow \omega_2 = 2(cos\frac{2\pi\cdot 2}{4} + i\cdot sin\frac{2\pi\cdot 2}{4}) = 2(cos\pi + i\cdot sin\pi) = 2(-1 + i\cdot 0) = -2,$$

$$k = 3 \Rightarrow \omega_3 = 2(cos\frac{2\pi 3}{4} + i\cdot sin\frac{2\pi 3}{4}) = 2(cos\frac{3\pi}{2} + i\cdot sin\frac{3\pi}{2}) = 2(0 - i)) = -2i.$$

Замечание. Геометрически все n значений корней n-ой степени из комплексного числа $r(cos\varphi + i\cdot sin\varphi)$ изображаются точками, лежащими на окружности с центром в начале координат, радиус которой равен $\sqrt[n]{r}$. Если эти точки соединить, то в результате получится правильный n-угольник.

Показательная форма комплексного числа

Из математического анализа известно, что $e = \lim\limits_{n\to\infty}(1+\frac{1}{n})^n$, e – иррациональное число. Эйлер в 1740 г. опубликовал формулу, которая дает возможность записать комплексное число в показательной форме:

$$e^{i\varphi} = cos\varphi + i\cdot sin\varphi \text{ – формула Эйлера.}$$

Тогда $z = r(cos\varphi + i\cdot sin\varphi) = r\cdot e^{i\varphi}$ – *показательная форма* записи комплексного числа, где $r = |z|$, $\varphi = arg\ z$.

К комплексным числам в показательной форме применимы все правила действия над степенями. Пусть $z_1 = r_1\cdot e^{i\varphi_1}$, $z_2 = r_2\cdot e^{i\varphi_2}$, тогда

$$z_1\cdot z_2 = r_1\cdot r_2\cdot e^{i(\varphi_1+\varphi_2)},$$

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}\,e^{i(\varphi_1-\varphi_2)},$$

$$z^n = r^n\cdot e^{in\varphi},$$

$$\sqrt[n]{z} = \sqrt[n]{r}\,e^{i\frac{\varphi+2\pi k}{n}}, \text{ где } k = 0, 1, \dots, n-1.$$

Пример 1.17. Представить комплексное число $z = \sqrt{3} + i$ в показательной форме.

Решение. Имеем $z = \sqrt{3} + i \Rightarrow a = \sqrt{3}$, $b = 1 \Rightarrow r = \sqrt{a^2+b^2} = \sqrt{3+1} = 2$; т. к. $a > 0$, то $\varphi = arctg\,\frac{b}{a} = arctg\,\frac{1}{\sqrt{3}} = \frac{\pi}{6}$. Тогда $z = \sqrt{3} + i = r\cdot e^{i\varphi} = 2\cdot e^{i\cdot\frac{\pi}{6}}$.

2. Бинарные отношения

2.1. Понятие отношения

Определение 2.1. *n-арным* (или *n-местным*) *отношением* P на множествах A_1, A_2, …, A_n называется любое подмножество прямого произведения $A_1 \times A_2 \times \dots \times A_n$.

Обозначение n-местного отношения: $P(x_1, x_2, \dots, x_n)$.

В случае $n = 1$ отношение P называется *унарным* (*одноместным*) и является подмножеством множества A_1.

При $n = 2$ P называется *бинарным* (*двуместным*) *отношением* или *соответствием*. Если $P \subseteq A_1 \times A_2$, то также говорят, что P есть отношение между множествами A_1 и A_2 (между элементами множеств A_1 и A_2) или что P задано (определено) на паре множеств A_1 и A_2. Если $A_1 = A_2 = A$ ($P \subseteq A \times A$), то говорят, что P есть *бинарное отношение на множестве* A.

Пусть P – бинарное отношение и $(x, y) \in P$, тогда говорят, что элемент x находится в отношении P к элементу y, или что x и y связаны отношением P. Вместо записи $(x, y) \in P$ часто пишут xPy.

Определение 2.2. Элементы x_1, x_2, …, x_n.называются *координатами*, или *компонентами*, отношения P.

Определение 2.3. Пусть $P \subseteq A \times B$, $S \subseteq A \times B$. Бинарные отношения P и S называются *равными* (пишут $P = S$), если для любых $x \in A$ и $y \in B$: $(x, y) \in P \Leftrightarrow (x, y) \in S$.

Другими словами, отношения P и S равны, если P и S равны как множества.

Определение 2.4. Для любого множества A отношение $id_A = \{(x, x) \mid x \in A\}$ называется *тождественным отношением* (или *диагональю*), а $U_A = A^2 = A \times A = \{(x, y) \mid x, y \in A\}$ – *полным отношением* (или *универсальным отношением* или *полным квадратом*).

Пусть P – некоторое бинарное отношение, т. е. $P \subseteq A_1 \times A_2$.

Определение 2.5. *Областью определения* бинарного отношения P называется множество $DomP = \{x \mid \exists\, y : (x, y) \in P\}$.

Определение 2.6. *Областью значений* бинарного отношения P называется множество $ImP = \{y \mid \exists\, x : (x, y) \in P\}$.

Пример 2.1. Задано множество $P = \{(1, y), (2, y), (3, x)\}$ на множествах $A = \{1, 2, 3\}$ и $B = \{x, y\}$. Покажем, что это действительно отношение, т. е.

$P \subseteq A \times B$. Найдем декартовое произведение множеств A и B: $A \times B = \{(1, x), (1, y), (2, x), (2, y), (3, x), (3, y)\}$, следовательно, $P \subset A \times B$.

Найдем область определения и область значений бинарного отношения P.

$DomP = \{1, 2, 3\} = A$; $ImP = \{x, y\} = B$.

Пример 2.2. Пусть $P \subseteq \mathbb{R} \times \mathbb{R}$: $P = \{(x, y) \mid y = x^2\}$. Найдем область определения и область значений бинарного отношения P.

$DomP = \mathbb{R}$; $ImP = [0, +\infty)$.

Способы задания бинарных отношений

Бинарные отношения можно задать одним из перечисленных способов.

1. Перечислением (см. пример 2.1). Такой способ задания применим только для конечных множеств.

2. Характеристическим свойством (см. пример 2.2).

3. Диаграммой. Пусть $P \subseteq A \times B$ – бинарное отношение. На диаграмме множества A и B изображаются с помощью кругов (или любых других связных фигур) на плоскости, а элементы множеств – точками внутри соответствующих кругов. Каждой упорядоченной паре (a, b) из бинарного отношения P сопоставляется отрезок прямой (или любая другая линия без самопересечений), соединяющий точки a и b и имеющий направление, указываемое стрелкой, от первого элемента упорядоченной пары ко второму.

Пример 2.3. Пусть бинарное отношение P задано диаграммой на рис. 2.1. Определим множества A, B и отношение P зададим перечислением.

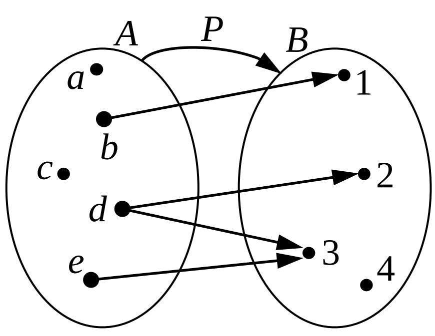

Рис. 2.1

Решение. $A = \{a, b, c, d, e\}$, $B = \{1, 2, 3\}$, $P = \{(b, 1), (d, 2), (d, 3), (e, 3)\}$.

4. Графом. Если $A = B$, то диаграмма

станет графом. Бинарному отношению P ставим в соответствие следующую геометрическую фигуру на плоскости: точки, являются элементами множества $DomP$ и ImP и ориентированные ребра (линии) – каждой паре $(a, b) \in P$ поставим в соответствие ориентированное ребро, идущее от a к b (если $a \neq b$) и петлю (если $a = b$) с фиксированным направлением обхода. Такую фигуру будем называть *ориентированным графом* отношения P.

Каждое бинарное отношение на конечном множестве можно представить ориентированным графом. Обратно, каждый ориентированный граф представляет бинарное отношение на множестве его вершин.

Пример 2.4. Граф, изображенный на рис. 2.2, задает отношение $P = \{(a, a), (a, c), (a, d), (b, e), (b, c), (d, c), (c, c)\}$ на множестве $A = \{a, b, c, d, e\}$.

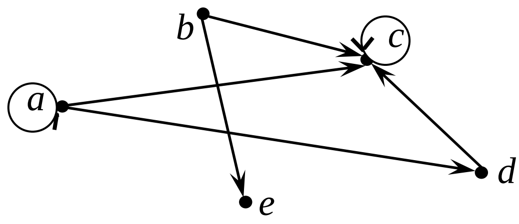

Рис. 2.2

5. Графиком. Этот способ применяется, если отношение задано на числовых множествах. *Графиком* бинарного отношения P называется множество точек плоскости Oxy с координатами (x, y) такие, что пара $(x, y) \in P$.

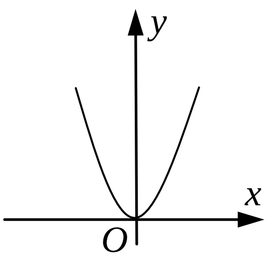

Рис. 2.3

Пример 2.5. График, изображенный на рис. 2.3, задает бинарное отношение $P = \{(x, y) \mid y = x^2, x, y \in \mathbb{R}\}$ на множестве $\mathbb{R}$, т. е. $P \subseteq \mathbb{R} \times \mathbb{R}$.

6. Таблицей. Например, таблица дежурств, синусов, логарифмов и др.

Операции над бинарными отношениями

Бинарные отношения – это множества упорядоченных пар. Следовательно, над ними можно выполнять любые теоретико-множественные операции, в частности, операции объединения и пересечения. Определим еще две операции над отношениями.

Определение 2.7. *Обратным* к отношению $P \subseteq A \times B$ (или *инверсией*) называется множество P^{-1}, подмножество прямого произведения $B \times A$ такое, что $P^{-1} = \{(y, x) \mid (x, y) \in P\}$.

Пример 2.6. Пусть $P = \{(a, 1), (b, 2), (c, 3), (d, 4), (e, 5)\}$. Тогда $P^{-1} = \{(1, a), (2, b), (3, c), (4, d), (5, e)\}$.

Определение 2.8. *Композицией* (или *суперпозицией*) отношений $P \subseteq A \times B$ и $Q \subseteq B \times C$ называется множество $P \circ Q = \{(x, y) \mid x \in A, y \in C, (\exists\, z \in B) : (x, z) \in P, (z, y) \in Q\}$, рис. 2.4.

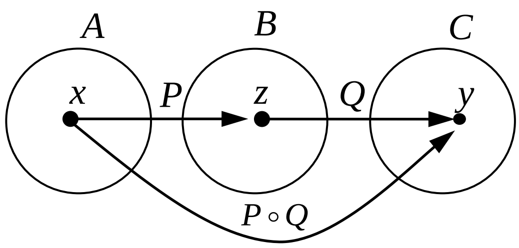

Рис. 2.4

Пример 2.7. Если $P = \{(a, b), (b, c), (b, d), (a, d), (c, a)\}$, $Q = \{(b, d), (c, a), (d, c)\}$, то $P \circ Q = \{(a, d), (b, a), (b, c), (a, c)\}$ и $Q \circ P = \{(c, b), (c, d), (d, a)\}$.

Утверждение 2.1. Для любых бинарных отношений P, Q и R выполняются следующие свойства:

1) $(P^{-1})^{-1} = P$;
2) $(P \circ Q)^{-1} = Q^{-1} \circ P^{-1}$;
3) $(P \circ Q) \circ R = P \circ (Q \circ R)$.

2.2. Свойства бинарных отношений

Пусть бинарное отношение P задано на непустом множестве A, т. е. $P \subseteq A^2$.

Определение 2.9. Бинарное отношение P на множестве A называется *рефлексивным*, если для любого элемента x из множества A пара $(x, x) \in P$.

Другими словами, отношение P рефлексивно тогда и только тогда, когда каждая вершина графа имеет петлю.

Примерами рефлексивных отношений являются отношение делимости на множестве целых чисел; отношение включения на булеане непустого множества; отношение быть одинаково направленными на множестве всех направленных лучей.

Пример 2.8. Выяснить обладает ли отношение $R = \{(m, n), (m, k), (m, m), (n, n), (k, n)\}$, заданный на множестве $A = \{m, n, k\}$, свойством рефлексивности.

Решение. Данное бинарное отношение $R \subseteq A^2$, не будет рефлескивным так как для $k \in A$ пара $(k, k) \notin R$.

Определение 2.10. Бинарное отношение P на множестве A называется *антирефлексивным*, если $\forall\, x \in A : (x, x) \notin P$.

Отношение антирефлексивно тогда и только тогда, когда ни одна вершина графа не имеет петли.

Например, отношение неравенства на некотором числовом множестве, отношение перпендикулярности на множестве всех прямых евклидовой плоскости являются антирефлексивными.

Пример 2.9. Выяснить обладает ли отношение $R = \{(x, y) \mid x, y \in \mathbb{R} : x + y = 2\}$, свойствами рефлексивности и антирефлексивности.

Решение. Имеем пара $(1, 1) \in R$ и $(2, 2) \notin R$, следовательно данное бинарное отношение R, не является рефлескивным и не является антирефлексивным.

Определение 2.11. Бинарное отношение P на множестве A называется *симметричным*, если $\forall\, x, y \in A$ из того что пара $(x, y) \in P$ следует $(y, x) \in P$.

Отношение симметрично тогда и только тогда, когда всякий раз вместе с ребром (x, y) граф содержит ребро (y, x).

Примерами симметричных отношений являются отношение равенства на некотором числовом множестве, отношение параллельности на множестве всех прямых евклидовой плоскости, отношение перпендикулярности на множестве всех прямых евклидовой плоскости.

Пример 2.10. Выяснить обладает ли отношение $R = \{(1, 2), (1, 3), (1, 1), (2, 3), (3, 1)\}$, заданное на множестве $A = \{1, 2, 3\}$, свойством симметричности.

Решение. Данное бинарное отношение $R \subseteq A^2$, не будет симметричным так как для пары $(2, 3) \in R$ пара $(3, 2) \notin R$.

Определение 2.12. Бинарное отношение P на множестве A называется *антисимметричным*, если $\forall\ x, y \in A$ из того, что пара $(x, y) \in P$ и $(y, x) \in P$ следует что $x = y$.

Отношение антисимметрично тогда и только тогда, когда вместе с каждым ребром (x, y) граф не содержит ребро (y, x). Граф антисимметричного отношения может содержать петли.

Примерами антисимметричных отношений являются отношение меньше (<) на множестве действительных чисел, отношение включения на булеане непустого множества.

Замечание 2.1. Если отношение не является симметричным, то это не означает, что оно антисимметрично. Например, отношение $R = \{(a, b), (b, a), (a, c)\}$ на множестве $A = \{a, b, c\}$ не симметрично, поскольку $(a, c) \in R$, а $(c, a) \notin R$, и не антисимметрично, так как $(a, b) \in R$ и $(b, a) \in R$, но $a \neq b$. Диагональ непустого множества A (id_A) является примером симметричного и антисимметричного отношения. Вообще, любое подмножество id_A обладает одновременно свойствами симметричности и антисимметричности.

Определение 2.13. Бинарное отношение P на множестве A называется асимметричным, если $\forall\ x, y \in A$ если пара $(x, y) \in P$, то $(y, x) \notin P$.

Отношение асимметрично тогда и только тогда, когда если граф содержит ребро (x, y), то он не содержит ребра (y, x).

Примерами асимметричных отношений являются отношение быть меньше (<) на числовом множестве. Отношение параллельности (||) на множестве прямых плоскости не является асимметричным, так как если $a \parallel b$, то $b \parallel a$ и оно является симметричным.

Определение 2.14. Бинарное отношение P на множестве A называется *транзитивным*, если $\forall\, x, y, z \in A$ из того, что пара $(x, y) \in P$ и $(y, z) \in P$ следует что $(x, z) \in P$.

Отношение транзитивно тогда и только тогда, когда вместе с каждой парой ребер (x, y) и (y, z) граф содержит ребро (x, z).

Например, отношение параллельности на множестве всех прямых евклидовой плоскости, отношение включения на булеане непустого множества являются транзитивными.

Утверждение 2.2. Пусть $A \neq \varnothing$ и $P \subseteq A^2$. Тогда справедливы следующие соотношения:

1) P – рефлексивно $\Leftrightarrow id_A \subseteq P$;
2) P – антирефлексивно $\Leftrightarrow P \cap id_A = \varnothing$;
3) P – симметрично $\Leftrightarrow P = P^{-1}$;
4) P – антисимметрично $\Leftrightarrow P \cap P^{-1} \subseteq id_A$;
5) P – транзитивно $\Leftrightarrow P \circ P \subseteq P$.

2.3. Отношение эквивалентности

Определение 2.15. Бинарное отношение на множестве A называется *отношением эквивалентности*, если оно рефлексивно, симметрично и транзитивно.

Отношение эквивалентности обычно обозначают символами $\sim$ или $\equiv$.

Примерами отношения эквивалентности являются отношение равенства на множестве действительных чисел, отношение параллельности на множестве прямых евклидовой плоскости, подобие треугольников на плоскости.

Определение 2.16. Пусть R – отношение эквивалентности на множестве A и $a \in A$. *Классом эквивалентности, порожденным элементом* a, называется множество $a/R = [a]_R = \{x \in A \mid (x, a) \in R\}$.

Совокупность всех классов эквивалентности отношения R на множестве A обозначается через A/R.

Определение 2.17. *Представителем класса эквивалентности* называется любой элемент этого класса.

Определение 2.18. Пусть A – непустое множество. *Фактор-множеством* множества A по отношению эквивалентности R называется множество A/R всех классов эквивалентности.

Определение 2.19. *Разбиением* непустого множества A называется совокупность его непустых подмножеств, таких что объединение всех подмножеств есть множество A, а пересечение любых двух различных подмножеств есть пустое множество.

Теорема 2.1. Пусть R – отношение эквивалентности на непустом множестве A. Тогда фактор-множество A/R является разбиением множества A.

*Доказательство***.** Так как отношение R рефлексивно, то для любого $a \in A$ имеем $(a, a) \in R$. Это значит, что каждый элемент a множества A принадлежит классу эквивалентности a/R. Итак, имеем семейство непустых классов a/R (a/R содержит по крайней мере один элемент a) и $\bigcup\limits_{a \in A} a/R = A$. Осталось доказать, что пересечение любых двух различных классов пусто. Для этого достаточно показать, что классы эквивалентности, имеющие хотя бы один общий элемент, совпадают. Пусть a/R и b/R – классы эквивалентности, имеющие общий элемент c. Тогда $(c, a) \in R$ и $(c, b) \in R$. В силу симметричности отношения R из $(c, a) \in R$ следует $(a, c) \in R$. Пусть x – любой элемент из a/R , тогда

$(x, a) \in R$. Имеем, $(x, a) \in R$ и $(a, c) \in R$. Следовательно, в силу транзитивности отношения R $(x, c) \in R$. Имеем, $(x, c) \in R$ и $(c, b) \in R$. Тогда $(x, b) \in R$, так как отношение R транзитивно. Следовательно, $x \in b/R$. Таким образом, $a/R \subseteq b/R$. Аналогично доказывается, что $b/R \subseteq a/R$. Следовательно, $a/R = b/R$.

Из теоремы 2.1 непосредственно вытекает следующее следствие.

Следствие. Пусть R – отношение эквивалентности на множестве A. Тогда

1) $\forall\ a \in A, a \in a/R$;

2) $\bigcup\limits_{a \in A} a/R = A$;

3) $\forall\ a, b \in A, a/R = b/R \Leftrightarrow (a, b) \in R$;

4) $a/R \neq b/R \Leftrightarrow a/R \cap b/R = \varnothing$.

Пусть S – разбиение непустого множества A и R_S – бинарное отношение, определяемое следующим образом: $(x, y) \in R_S$ тогда и только тогда, когда x и y принадлежат одному и тому же подмножеству семейства S.

Теорема 2.2. Отношение R_S, соответствующее разбиению S непустого множества A, является отношением эквивалентности на A, причем фактор-множество A/R_S совпадает с разбиением S.

Доказательство. 1. Так как S есть разбиение, то $\forall\ a \in A$, $\exists\ M_i \subseteq \mathrm{S}$: $a \in M_i$. Следовательно, по определению отношения R_S, $(a, a) \in R_S$, а значит R_S – рефлексивно.

2. Пусть a, b – произвольные элементы из A такие, что $(a, b) \in R_S$. Тогда, по определению отношения R_S, $\exists\ M_j \subseteq \mathrm{S}$: $a, b \in M_j$. Следовательно, $(b, a) \in R_S$. Получили, что R_S – симметрично.

3. Пусть a, b, c – произвольные элементы из A такие, что $(a, b) \in R_S$ и $(b, c) \in R_S$. Следовательно, по определению отношения R_S, $\exists\ M_i$, $M_j \subseteq S$: a, $b \in M_i$ и b, $c \in M_j$. Отсюда $b \in M_i \cap M_j$. Но тогда, по определению

разбиения, $M_i = M_j$, а значит, $a, c \in M_i$, и, по определению отношения R_S, $(a, c) \in R_S$. Получили, что R_S – транзитивно.

Из п. 1 – 3 следует, что R_S – отношение эквивалентности. Фактор-множество A/R_S совпадает с разбиением S по определению отношения R_S.

Замечание 2.2. Частным случаем отношения эквивалентности $\sim$ является отношение равенства элементов некоторого множества A, которое определяет разбиение множества на одноэлементные классы эквивалентности: $\forall\, x \in A$, $x/\sim\, = \{x\}$. В этом случае классов эквивалентности оказывается столько же, сколько элементов содержится в множестве A, так как каждый элемент из A эквивалентен только самому себе.

В другом частном случае все элементы множества A эквивалентны друг другу. При этом фактор-множество $A/\sim$ состоит всего из одного класса – самого множества A.

В любом другом случае среди классов эквивалентности имеется хотя бы один класс, который содержит больше одного элемента и в то же время не совпадает с самим множеством A.

Замечание 2.3. Понятие отношения эквивалентности имеет большое значение в математике. Дело в том, что элементы, входящие в один класс эквивалентности неразличимы с точки зрения рассматриваемого отношения эквивалентности. Поэтому считают, что класс эквивалентности определяется любым своим представителем (произвольным элементом этого класса). Это позволяет вместо всех элементов множества изучать совокупность представителей каждого класса эквивалентности. Свойства, которыми обладают все элементы некоторого класса эквивалентности, изучаются на одном его представителе.

Отношения эквивалентности играют важную роль в определении математических понятий.

2.4. Функции

Определение 2.20. Бинарное отношение $f \subseteq A \times B$ называется *функцией* из множества A в множество B, если для любого $x \in A$ существует единственный элемент $y \in B$ такой, что $(x, y) \in f$. При этом элемент y обозначается через $f(x)$ и называется *значением* функции f для *аргумента* x. Функция f из A в B обозначается через f: $A \to B$ или $A \xrightarrow{f} B$. Если $(x, y) \in f$, то используется общепринятая запись $y = f(x)$, а также запись f: $x \mapsto y$ (означает, что функция f ставит в соответствие элементу x элемент y).

Область определения и область значений функции, равные функции определяются так же, как и для бинарных отношений.

Аргументами функции могут являться элементы произвольной природы, в частности, кортежи длины n $(x_1, x_2, \ldots, x_n)$. Функцию f: $A^n \to B$ называют *n-местной функцией* из A в B. Тогда пишут $y = f(x_1, x_2, \ldots, x_n)$ и говорят, что y есть значение функции f при значении аргументов $x_1, x_2, \ldots, x_n$.

Функции называются также *отображениями*. Пусть f – функция из A в B. Если $A = Dom f$ и $Im f \subseteq B$, то говорят, что f есть *отображение множества A в множество B*. Если $A = Dom f$ и $B = Im f$, то говорят, что f есть *отображение множества A на множество B*.

Определение 2.21. Функция $f \subseteq A \times B$ называется *инъективной*, или *инъекцией*, если $\forall\, x, y \in A$, $f(x) = f(y) \Rightarrow x = y$.

Определение 2.22. Функция $f \subseteq A \times B$ называется *сюръективной*, или *сюръекцией*, если для каждого элемента $y \in B$ существует хотя бы один элемент $x \in A$ такой, что $y = f(x)$.

Заметим, что сюръективная функция $f \subseteq A \times B$ является отображением A на B.

Определение 2.23. Функция $f \subseteq A \times B$ называется *биективной* (*биекцией*) или *взаимно однозначным соответствием между множествами A и B*, если она одновременно инъективна и сюръективна.

Определение 2.24. Если соответствие, обратное к функции $f \subseteq A \times B$, является функциональным и полностью определенным, то оно называется *функцией*, *обратной* к f и обозначается f^{-1}.

Так как в обратном соответствии образы и прообразы меняются местами, то для существования функции, обратной к функции $f \subseteq A \times B$, необходимо и достаточно, чтобы $Imf = B$ и каждый элемент $y \in Imf$ имел единственный прообраз.

Утверждение 2.3. Для функции $f: A \to B$ существует обратная к ней функция $f^{-1}: B \to A$ тогда и только тогда, когда f – биекция.

Определение 2.25. Пусть даны функции f: $A \to B$ и g: $B \to C$. Функция h: $A \to C$ называется *композицией* (*суперпозицией*) *функций* f и g, если $\forall\ x \in A$, $h(x) = g(f(x))$.

Композиция функций f и g обозначается через $g \circ f$, при этом знак $\circ$ часто опускается.

3. Матрицы и действия над ними

3.1. Общие понятия

Определение 3.1. *Матрицей* называется прямоугольная таблица из чисел, содержащая m строк и n столбцов. Числа m и n называют *порядком* (или *типом*, или *размерностью*) матрицы; обозначают $m \times n$ или (m, n).

Матрицы обозначаются прописными буквами латинского алфавита: A, B, C, …, а числа, их составляющие – элементы матрицы – строчными буквами с двумя индексами: a_{ij}, где i – номер строки, j – номер столбца.

Например, матрица размерности $m \times n$ имеет вид:

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

коротко записывают $A = \|a_{ij}\|$ или $A = (a_{ij})$, где $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$.

Пусть $B = \begin{pmatrix} 3 & -4 \\ 0 & 6 \\ 8 & 1 \end{pmatrix}$ – конкретная матрица размерности 3×2, тогда пишут $B_{3\times 2}$.

Определение 3.2. Матрица называется *квадратной* порядка n, если число строк равно числу столбцов и равно n (то есть $m = n$).

Определение 3.3. Две матрицы $A = (a_{ij})$ и $B = (b_{ij})$ называются *однотипными*, если они имеют одинаковое число строк и одинаковое число столбцов.

Определение 3.4. Две однотипные матрицы $A = (a_{ij})$ и $B = (b_{ij})$ называют *равными* и пишут $A = B$, если $a_{ij} = b_{ij}$ для любых индексов i, j.

Определение 3.5. Матрица, состоящая из одной строки, называется *матрицей-строкой* или *вектором-строкой*.

Определение 3.6. Матрица, состоящая из одного столбца, называется *матрицей-столбцом* или *вектором-столбцом*.

Например, матрица-строка $A_{1\times 3} = (-2\ 5\ 1)$, матрица-столбец $B_{3\times 1} = \begin{pmatrix} 4 \\ 0 \\ -8 \end{pmatrix}$.

В квадратных матрицах $A = (a_{ij})$ порядка n вводится понятие *главной диагонали*: $a_{11}, a_{22}, \ldots, a_{nn}$ и *побочной диагонали*: $a_{1n}, a_{2n-1}, \ldots, a_{n1}$.

Определение 3.7. Квадратная матрица, все элементы которой вне главной диагонали равны нулю называется *диагональной*, т. е. матрица вида $\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$

Определение 3.8. Квадратная матрица вида $\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$ называется *верхнетреугольной*. Аналогично вводится понятие *нижнетреугольной* матрицы.

Определение 3.9. Квадратная матрица, у которой все элементы главной диагонали равны единице, а все остальные элементы равны нулю, называется *единичной* матрицей порядка n, т. е. $E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$.

Определение 3.10. Матрица называется *нулевой* и обозначается O, если все ее элементы равны нулю, т. е. $O = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}$.

Определение 3.11. Квадратная матрица называется *симметрической*, если ее элементы подчиняются условию: $a_{ij} = a_{ji}$ для любых индексов i, j.

3.2. Основные операции над матрицами и их свойства

3.2.1. Сложение однотипных матриц

Складывать можно только однотипные матрицы.

Определение 3.12. *Суммой двух матриц* $A = (a_{ij})$ и $B = (b_{ij})$, где $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$ называется матрица $C = (c_{ij})$ для которой $c_{ij} = a_{ij} + b_{ij}$, где $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$. Обозначение: $C = A + B$.

Иначе говоря: сложение матриц производится поэлементно.

Пример 3.1. Для матриц $A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 5 & 7 \end{pmatrix}$ и $B = \begin{pmatrix} 8 & 4 & 3 \\ 2 & 1 & 0 \end{pmatrix}$ найти их сумму.

Решение. $C = A + B = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 5 & 7 \end{pmatrix} + \begin{pmatrix} 8 & 4 & 3 \\ 2 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 9 & 6 & 6 \\ 1 & 6 & 7 \end{pmatrix}$.

Свойства сложения матриц

1) коммутативность: $\forall\, A, B : A + B = B + A$;
2) ассоциативность: $\forall\, A, B, C : (A + B) + C = A + (B + C)$;
3) $\forall\, A$, $A + O = A$, где O – нулевая матрица;
4) $\forall\, A$, $\exists\, -A : A + (-A) = O$, $(-A)$ – матрица, противоположная матрице A.

Замечание 3.1. Пусть $A = (a_{ij})$ тогда $-A = (-a_{ij})$, где элемент $-a_{ij}$ – противоположный элементу a_{ij} для любых индексов i и j, при этом матрица $-A$ называется *противоположной* к матрице A.

Пример 3.2. Пусть A = $\begin{pmatrix} 1 & 2 & -3 \\ 0 & 5 & 4 \end{pmatrix}$ тогда $-A = \begin{pmatrix} -1 & -2 & 3 \\ 0 & -5 & -4 \end{pmatrix}$.

3.2.2. Умножение матрицы на число

Определение 3.13. *Произведением матрицы* $A = (a_{ij})$ *на действительной число* k называется матрица $C = (c_{ij})$, для которой $c_{ij} = k \cdot a_{ij}$, где $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$. Обозначение: $C = k \cdot A = kA$, т. о. каждый элемент матрицы A умножают на действительное число k.

Пример 3.3. Пусть дано число $k = -2$ и матрица $A = \begin{pmatrix} 1 & 3 \\ 4 & -5 \end{pmatrix}$, тогда

$k \cdot A = (-2) \cdot \begin{pmatrix} 1 & 3 \\ 4 & -5 \end{pmatrix} = \begin{pmatrix} -2 & -6 \\ -8 & 10 \end{pmatrix}$.

Свойства умножения матрицы на число

1) $\forall A : 1 \cdot A = A$;

2) $\forall \alpha, \beta \in \mathbb{R}, \forall A : (\alpha\beta) \cdot A = \alpha \cdot (\beta \cdot A) = \beta \cdot (\alpha \cdot A)$;

3) $\forall \alpha \in \mathbb{R}, \forall A, B : \alpha \cdot (A + B) = \alpha \cdot A + \alpha \cdot B$;

4) $\forall \alpha, \beta \in \mathbb{R}, \forall A : (\alpha + \beta) \cdot A = \alpha \cdot A + \beta \cdot A$.

3.2.3. Умножение матриц

Определим умножение двух матриц; для этого необходимо ввести некоторые дополнительные понятия.

Определение 3.14. Матрицы A и B называются *согласованными*, если число столбцов матрицы A равно числу строк матрицы B.

Например матрицы A размерности $m \times n$ и B размерности $n \times p$ будут согласованными.

Обозначим строки матрицы A как $A_1, A_2, \ldots, A_m$, а столбцы матрицы B как $B^1, B^2, \ldots, B^p$. При этом в строке матрицы A столько же элементов, сколько в столбце матрицы B. Это условие позволяет умножать строку матрицы A на столбец матрицы B.

Умножим, например, A_1 на столбец B^1. Пусть $A_1 = (a_1\ a_2\ \ldots\ a_n)$,

$B^1 = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ тогда $A_1 \cdot B^1 = (a_1\ a_2\ \ldots\ a_n) \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1 \cdot b_1 + a_2 \cdot b_2 + \ldots + a_n \cdot b_n$.

Пример 3.4. Умножим строку (1 2 3 4) на столбец $\begin{pmatrix} 3 \\ 2 \\ 1 \\ 0 \end{pmatrix}$.

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

Теперь можно определить умножение матриц.

Определение 3.15. *Произведением* согласованных матриц A размерности $m \times n$ и B размерности $n \times p$ называется матрица $C = (c_{ij})$ размерности $m \times p$, для которой $c_{ij} = A_i \cdot B^j$, где $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, p$.

Обозначение: $C = A \cdot B = \begin{pmatrix} A_1 \cdot B^1 & A_1 \cdot B^2 & \ldots & A_1 \cdot B^p \\ A_2 \cdot B^1 & A_2 \cdot B^2 & \ldots & A_2 \cdot B^p \\ \ldots & \ldots & \ldots & \ldots \\ A_m \cdot B^1 & A_m \cdot B^2 & \ldots & A_m \cdot B^p \end{pmatrix}$.

Пример 3.5. Найти произведение матриц $A \cdot B$, где $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$,

$B = \begin{pmatrix} -1 & 5 & 3 & 1 \\ 1 & 2 & 3 & 1 \end{pmatrix}$.

Решение. Матрицы $A_{3\times 2}$ и $B_{2\times 4}$ согласованы, значит можно найти произведение этих матриц $A \cdot B$, результатом будет матрица размерности 3×4. $A \cdot B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} -1 & 5 & 3 & 1 \\ 1 & 2 & -4 & -1 \end{pmatrix} =$

$$= \begin{pmatrix} 1\cdot(-1)+2\cdot 1 & 1\cdot 5+2\cdot 2 & 1\cdot 3+2\cdot(-4) & 1\cdot 1+2\cdot(-1) \\ 3\cdot(-1)+4\cdot 1 & 3\cdot 5+4\cdot 2 & 3\cdot 3+4\cdot(-4) & 3\cdot 1+4\cdot(-1) \\ 5\cdot(-1)+6\cdot 1 & 5\cdot 5+6\cdot 2 & 5\cdot 3+6\cdot(-4) & 5\cdot 1+6\cdot(-1) \end{pmatrix} = \begin{pmatrix} 1 & 9 & -5 & -1 \\ 1 & 23 & -7 & -1 \\ 1 & 37 & -9 & -1 \end{pmatrix}$$

Свойства умножения матриц

1) Умножение матриц не коммутативно: $A \cdot B \neq B \cdot A$.

Продемонстрировать данное свойство можно на примерах.

Пример 3.6. **а)** Пусть даны две матрицы: $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ и $B = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

Перемножим матрицы $A \cdot B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1\cdot(-1)+2\cdot 2 \\ 3\cdot(-1)+4\cdot 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$,

получим матрицу размерности 2×1. Умножить матрицу $B = B_{2\times 1}$ на матрицу $A = A_{2\times 2}$ нельзя, так как эти матрицы не согласованные. Т. о. свойство коммутативности для умножения двух матриц не выполняется.

б) Возьмем две матрицы так, чтобы A и B были согласованы и чтобы также B и A были согласованные. Проверим, что при данных условиях свойство коммутативности также не выполняется. Пусть $A = A_{2\times 3} = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 4 & 0 \end{pmatrix}$ и $B = B_{3\times 2} = \begin{pmatrix} 5 & 6 \\ 7 & 8 \\ 1 & 1 \end{pmatrix}$, найдем их произведения.

$$A\cdot B = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 4 & 0 \end{pmatrix}\cdot\begin{pmatrix} 5 & 6 \\ 7 & 8 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 18 & 21 \\ 43 & 50 \end{pmatrix} = C_{2\times 2};$$

$$B\cdot A = \begin{pmatrix} 5 & 6 \\ 7 & 8 \\ 1 & 1 \end{pmatrix}\cdot\begin{pmatrix} 1 & 2 & -1 \\ 3 & 4 & 0 \end{pmatrix} = \begin{pmatrix} 23 & 34 & -5 \\ 31 & 46 & -7 \\ 4 & 6 & -1 \end{pmatrix} = C_{3\times 3}.$$

2) Ассоциативность: $(A\cdot B)\cdot C = A\cdot(B\cdot C)$.
3) Для любой квадратной матрицы A и согласованной с ней единичной матрицы E справедливо равенство: $A\cdot E = E\cdot A$.
4) Дистрибутивный закон умножения матриц относительно сложения матриц: $\forall\, A, B, C : (A + B)\cdot C = (A\cdot C) + (B\cdot C)$ и $A\cdot(B + C) = (A\cdot B) + (A\cdot C)$.

Пример 3.7. Пусть даны матрицы $A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 3 \end{pmatrix}$ и $C = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ проверим справедливость свойства 4.

$$(A + B)\cdot C = \left(\begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 3 \end{pmatrix}\right)\cdot\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 5 \\ 0 & 1 & 6 \end{pmatrix}\cdot\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 22 \\ 20 \end{pmatrix};$$

$$(A\cdot C) + (B\cdot C) = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \end{pmatrix}\cdot\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 3 \end{pmatrix}\cdot\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 8 \\ 12 \end{pmatrix} + \begin{pmatrix} 14 \\ 8 \end{pmatrix} = \begin{pmatrix} 22 \\ 20 \end{pmatrix}.$$

5) $\forall\, k \in \mathbb{R},\ \forall\, A, B : k(A\cdot B) = (kA)\cdot B = A\cdot(kB)$.

3.3. Транспонирование матриц

Определение 3.16. Матрица A^t, полученная из данной заменой каждой ее строки столбцом с тем же номером, называется *транспонированной* к данной матрице A.

Если A – матрица размерности $m \times n$, то A^t – матрица размерности $n \times m$.

При транспонировании верны следующие равенства:

1) $(A + B)^t = A^t + B^t$;
2) $(kA)^t = kA^t$;
3) $(A\cdot B)^t = B^t\cdot A^t$.

4. Определители квадратных матриц

4.1. Определители матриц второго и третьего порядка

Каждой квадратной матрице A порядка n ставится в соответствие число, которое называется *определителем* этой матрицы. Обозначение: Δ, $|A|$, $\det A$, $\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$.

Определение 4.1. *Определителем матрицы первого порядка* $A = (a_{11})$ называется число a_{11}.

Пример 4.1. Например: если дана матрица первого порядка $A = (3)$, то определитель этой матрицы $|A| = 3$.

Определение 4.2. *Определителем матрицы второго порядка* $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ *называется число, которой находится по формуле:* $|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$.

Пример 4.2. Если дана матрица второго порядка $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, то определитель этой матрицы $|A| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 \cdot 4 - 2 \cdot 3 = 4 - 6 = -2$.

Определение 4.3. *Определителем матрицы третьего порядка* $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ *называется число, которой находится по формуле:*

$$|A| = a_{11}a_{22}a_{33} + a_{12}a_{31}a_{23} + a_{21}a_{13}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{32}a_{23} - a_{33}a_{21}a_{12}.$$

Это число состоит из шести слагаемых, в каждое слагаемое в качестве множителей входит по одному элементу из каждой строки и каждого столбца. Для запоминания формулы можно воспользоваться наглядным правилом знаков для выписывания произведений, входящих в разложение определителя третьего порядка. Схема на рис. 4.1 называется *правилом треугольника* или *правилом Саррюса*[10].

Правило составления выражения для определителя третьего порядка строится следующим образом. Из членов, входящих со знаком «+», один будет произведением элементов главной диагонали, каждый из двух других – произведением элементов, лежащих на параллели к этой диагонали, с добавлением третьего множителя из противоположного угла

[10] Пьер Фредерик Саррюс (1798–1861) – французский математик.

матрицы (рис. 4.1). Члены, входящие со знаком «–», строятся таким же, образом относительно другой диагонали.

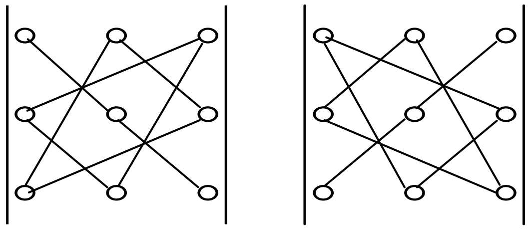

Рис. 4.1

Существует еще вторая схема правила Саррюса: к определителю приписывают справа два первых столбца и вычисляют сумму произведений элементов расположенных на главной диагонали и «прямых», параллельных ей, со знаком минус вычисляют сумму произведений элементов, расположенных на побочной диагонали, и «прямых», параллельных ей.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}\begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} -$$

$$a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Пример 4.3. Если дана матрица третьего порядка $A = \begin{pmatrix} 1 & 0 & -5 \\ -2 & 3 & 2 \\ 1 & -2 & 0 \end{pmatrix}$, то определитель этой матрицы

$$|A| = \begin{vmatrix} 1 & 0 & -5 \\ -2 & 3 & 2 \\ 1 & -2 & 0 \end{vmatrix} = 1\cdot 3\cdot 0 + 1\cdot 0\cdot 2 + (-5)\cdot(-2)\cdot(-2) - (-5)\cdot 3\cdot 1 - 1\cdot 2\cdot(-2) -$$

$$- 0\cdot 0\cdot(-2) = -20 + 15 + 4 = -1.$$

4.2. Определитель матрицы *n*-го порядка

Для того чтобы дать определение определителя произвольного порядка, введем некоторые понятия. Пусть a_{ij} – элемент определителя порядка n, где i, $j = 1, 2, \ldots, n$.

Определение 4.4. *Минором* элемента a_{ij} называется определитель M_{ij}, полученный из данного определителя вычеркиванием i-ой строки и j-го столбца.

Из определения следует, что минор элемента – это определитель $(n – 1)$ порядка.

Определение 4.5. *Алгебраическим дополнением* элемента a_{ij} называется его минор, взятый со знаком $(–1)^{i+j}$, т. е. $A_{ij} = (–1)^{i+j} \cdot M_{ij}$.

Для определения понятия определителя n-го порядка воспользуемся индукцией по n, где n – порядок матрицы A.

Определение 4.6.

1. При $n = 1$ матрица A состоит из одного числа: $|A| = a_{11}$.

2. Пусть для матрицы порядка $(n – 1)$ определитель известен.

3. *Определителем* матрицы A произвольного порядка n называется число, находящееся по формуле: $|A| = \dfrac{1}{n}\sum\limits_{i,j=1}^{n} a_{ij} \cdot (-1)^{i+j} M_{ij}$, где суммирование распространяется на все элементы матрицы A.

Эта формула сводит вычисление определителей порядка n к вычислению определителей порядка $(n – 1)$.

4.3. Свойства определителей

Для того чтобы вычислять определители порядков, больших, чем 3, используют свойства определителей и теорему Лапласа.

Теорема 4.1 (Лапласа). Определитель квадратной матрицы равен сумме произведений элементов любой строки (столбца) на их алгебраические дополнения, т. е.

$\Delta = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}$, где $i = 1, 2, \ldots, n$ (разложение определителя по элементам i-ой строки);

$\Delta = a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj}$, где $j = 1, 2, \ldots, n$ (разложение определителя по элементам j-го столбца).

Перечислим основные свойства определителей.

1. Если какая-либо строка (столбец) матрицы состоит из одних нулей, то определитель этой матрицы равен нулю.
2. Определитель диагональной матрицы равен произведению ее диагональных элементов.
3. Определитель треугольной (верхнетреугольной или нижнетреугольной) матрицы равен произведению элементов ее главной диагонали.
4. При перестановке двух строк (столбцов) матрицы ее определитель меняет знак на противоположный.
5. Если квадратная матрица содержит две одинаковые строки (столбца), то ее определитель равен нулю.
6. Если все элементы какой- либо строки (столбца) определителя умножить на некоторое число k, то определитель умножается на это число k. (Общий множитель элементов строки (столбца) можно выносить за знак определителя.)
7. Если квадратная матрица содержит две пропорциональные строки (столбца), то ее определитель равен нулю.
8. Определитель матрицы не изменится, если к элементам какой-либо строки (столбца) матрицы прибавить элементы другой строки (столбца) матрицы, предварительно умноженные на одно и то же число.
9. При транспонировании матрицы ее определитель не меняется.
10. Определитель произведения двух квадратных матриц равен произведению их определителей.
11. Сумма произведений элементов какой-либо строки (столбца) матрицы на алгебраические дополнения элементов другой строки (столбца) этой матрицы равна нулю.

4.4. Практическое вычисление определителей

Один из способов вычисления определителей порядка выше трех – разложение его по какому-либо столбцу или строке.

Пример 4.4. Вычислить определитель $\Delta = \begin{vmatrix} 3 & 1 & -1 & 3 \\ -5 & 1 & -3 & -4 \\ 2 & 0 & 1 & -1 \\ 1 & 5 & 3 & 3 \end{vmatrix}$.

Решение. Разложим данный определитель по третьей строке:

$$\begin{vmatrix} 3 & 1 & -1 & 3 \\ -5 & 1 & -3 & -4 \\ 2 & 0 & 1 & -1 \\ 1 & 5 & 3 & 3 \end{vmatrix} = a_{31}A_{31} + a_{32}A_{32} + a_{33}A_{33} + a_{34}A_{34} =$$

$$= 2\cdot(-1)^{3+1}M_{31} ++ 0\cdot(-1)^{3+2}M_{32} + 1\cdot(-1)^{3+3}M_{33} + (-1)\cdot(-1)^{3+4}M_{34} =$$

$$= 2\cdot1\cdot\begin{vmatrix} 1 & -1 & 2 \\ 1 & 3 & -4 \\ 5 & 3 & 3 \end{vmatrix} + 0 + 1\cdot1\cdot\begin{vmatrix} 3 & 1 & 2 \\ -5 & 1 & -4 \\ 1 & 5 & 3 \end{vmatrix} + (-1)\cdot(-1)\cdot\begin{vmatrix} 3 & 1 & -1 \\ -5 & 1 & 3 \\ 1 & 5 & 3 \end{vmatrix} =$$

$= 2(9 + 20 + 6 – 30 + 12 + 3) + (9 – 4 – 50 – 2 + 60 + 15) + (9 + 3 + 25 + 1 –$
$– 45 + 15) = 40 + 28 + 8 = 76.$

При вычислении определителей целесообразно так преобразовать исходную матрицу с помощью свойств определителей, чтобы в преобразованной матрице получилась строка (столбец), содержащая максимальное число нулей («обнулить» строку), а потом найти определитель разложением по этой строке (столбцу). В ходе преобразований необходимо следить за тем, чтобы значение определителя не менялось.

Пример 4.5. Вычислить определитель четвертого порядка:
$$\Delta = \begin{vmatrix} 4 & 6 & -2 & 4 \\ 1 & 2 & -3 & 1 \\ 4 & -2 & 1 & 0 \\ 6 & 4 & 4 & 6 \end{vmatrix}.$$

Решение. В третьей строке уже есть один ноль. Если к 1-ому столбцу прибавить 3-ий, умноженный на (-4), а ко 2-му столбцу прибавить 3-ий, умноженный на 2, то получим следующий определитель, который разложим по элементам 3-ей строки (теорема Лапласа):
$$\begin{vmatrix} 12 & 2 & -2 & 4 \\ 13 & -4 & -3 & 1 \\ 0 & 0 & 1 & 0 \\ -10 & 12 & 4 & 6 \end{vmatrix} = a_{33}(-1)^{3+3} \cdot M_{33} = 1 \cdot (-1)^{3+3} \cdot \begin{vmatrix} 12 & 2 & 4 \\ 13 & -4 & 1 \\ -10 & 12 & 6 \end{vmatrix}.$$

Полученный определитель можно вычислить по правилу треугольника или продолжить упрощение матрицы с последующим применением теоремы Лапласа. Прибавим к 1-ой строке 2-ую, умноженную на (–4), а к 3-ей строке 2-ую, умноженную на (–6), и получим такой определитель: $\begin{vmatrix} -40 & 18 & 0 \\ 13 & -4 & 1 \\ -88 & 36 & 0 \end{vmatrix} = 1 \cdot (-1)^{2+3} \begin{vmatrix} -40 & 18 \\ -88 & 36 \end{vmatrix} = -144.$

5. Ранг матрицы. Обратная матрица

5.1. Понятие ранга матрицы

Пусть A – матрица размерности $m \times n$. Выберем в этой матрице произвольно k строк и k столбцов, где $1 \le k \le min(m, n)$.

Определение 5.1. *Минором k-го порядка* матрицы A называется определитель матрицы, стоящей на пересечении этих k строк и k столбцов.

Другими словами, если в матрице A размерности $m \times n$ вычеркнуть $(m - k)$ строк и $(n - k)$ столбцов, а из оставшихся элементов составить матрицу, сохраняя расположение элементов матрицы A , то определитель полученной матрицы есть минор порядка k матрицы A.

Пример 5.1. Проиллюстрируем определение минором k-го порядка матрицы A. Рассмотрим матрицу $A = \begin{pmatrix} 1 & -8 & 5 \\ 4 & 2 & -3 \\ 0 & 7 & 2 \\ -2 & 3 & 7 \end{pmatrix}$. Запишем минор первого порядка этой матрицы. Например, если выбрать третью строку и второй столбец матрицы A, то данному выбору соответствует минор первого порядка $M_1 = det(7) = 7$. Иными словами, для получения этого минора надо вычеркнуть первую, вторую и четвертую строки, а также первый и третий столбцы из матрицы A, а из оставшегося элемента составить определитель. Таким образом, минорами первого порядка матрицы являются сами элементы матрицы.

Приведем пример минора второго порядка матрицы A. Выберем две строки, например, первую и вторую, и два столбца, например, первый и третий. Вычислим определитель, стоящий на их пересечении $M_2 = \begin{vmatrix} 1 & 5 \\ 4 & -3 \end{vmatrix} = -23$. Этот минор также можно было составить вычеркиванием из матрицы A третьей и четвертой строки, второго столбца.

Аналогично могут быть найдены миноры третьего порядка матрицы A. Так как в матрице A всего три столбца, то берем их все. Если к этим столбцам добавить три строки, например первую, третью и четвертую, то получим минор третьего порядка $M_3 = \begin{vmatrix} 1 & -8 & 5 \\ 0 & 7 & 2 \\ -2 & 3 & 7 \end{vmatrix} = 145$. Данный минор также может быть построен вычеркиванием второй строки матрицы A.

Можно получить другой минор третьего порядка, если вычеркивать третью строку матрицы A.

Для данной матрицы A миноров порядка выше третьего не существует, так как $k \le min(m, n) = min(4, 3) = 3$.

Замечание 5.1. Число миноров порядка k матрицы A размерности $m \times n$ может быть вычислено по формуле: $C_m^k \cdot C_n^k$, где C_m^k и C_n^k – число сочетаний из m по k и из n по k соответственно.

Определение 5.2. *Рангом* матрицы называется наибольший из порядков ее миноров, отличных от нуля. Обозначение: $rang\ A$, r или $r(A)$.

Из определения следует, что

1) для матрицы A размерности $m \times n$ имеем $0 \le rang\ A \le min(m, n)$;
2) ранг нулевой матрицы равен нулю;
3) ранг ненулевой матрицы не меньше единицы;
4) ранг квадратной матрицы порядка n равен n только тогда, когда ее определитель не равен нулю;
5) ранг матрицы не меняется при транспонировании.

5.2. Нахождение ранга матрицы методом окаймления миноров

Один из методов методом нахождения ранга матрицы является *метод перебора миноров*. Этот способ основан на определении ранга матрицы. Суть метода в следующем. Если есть хотя бы один элемент матрицы размерности $m \times n$, отличный от нуля, то ранг матрицы как минимум равен единице (так как есть минор первого порядка, не равный нулю). Далее перебираем миноры второго порядка. Если все миноры второго порядка равны нулю, то ранг матрицы равен единице. Если существует хотя бы один ненулевой минор второго порядка, то переходим к перебору миноров третьего порядка, а ранг матрицы как минимум равен двум. Аналогично, если все миноры третьего порядка равны нулю, то ранг матрицы равен двум. Если существует хотя бы один минор третьего порядка, отличный от нуля, то ранг матрицы как минимум равен трем, и преступаем к перебору миноров четвертого порядка. И так далее.

Следует учесть, что ранг матрицы не может превышать наименьшего из чисел m и n.

Опишем более рациональный метод нахождения ранга матрицы – это метод окаймляющих миноров.

Определение 5.3. Минор M' $(k+1)$-ого порядка матрицы A *окаймляет* минор M порядка k матрицы A, если матрица, соответствующая минору M', «содержит» матрицу, соответствующую минору M.

Другими словами, матрица, соответствующая окаймляемому минору M, получается из матрицы, соответствующей окаймляющему минору M', вычеркиванием элементов одной строки и одного столбца.

Теорема 5.1. Если в матрице A имеется минор M порядка r, отличный от нуля, а все миноры матрицы A, окаймляющие минор M (если они существуют) равны нулю, то ранг матрицы A равен r.

Замечание 5.2. Данную теорему можно переформулировать иначе. Если все миноры, окаймляющие минор k-ого порядка матрицы A размерности $m \times n$, равны нулю, то все миноры порядка $(k+1)$ матрицы A равны нулю.

Таким образом, для нахождения ранга матрицы не обязательно перебирать все миноры, достаточно окаймляющих. Количество миноров, окаймляющих минор k-ого порядка матрицы A размерности $m \times n$, находится по формуле $(m-k)\cdot(n-k)$. Отметим, что миноров, окаймляющих минор k-ого порядка матрицы A, не больше, чем миноров $(k+1)$-ого порядка матрицы A. Поэтому, в большинстве случаев использование метода окаймляющих миноров выгоднее простого перебора всех миноров.

Опишем алгоритм данного метода.

Если матрица A ненулевая, то в качестве минора первого порядка берем любой элемент матрицы A, отличный от нуля. Рассматриваем его окаймляющие миноры. Если все они равны нулю, то ранг матрицы равен единице. Если же есть хотя бы один ненулевой окаймляющий минор (его порядок равен двум), то переходим к рассмотрению его окаймляющих миноров. Если все они равны нулю, то $rang\,A = 2$. Если хотя бы один окаймляющий минор отличен от нуля (его порядок равен трем), то рассматриваем его окаймляющие миноры. И так далее. В итоге $rang\,A = k$, если все окаймляющие миноры $(k + 1)$-ого порядка матрицы A равны нулю, либо $rang\,A = min(m, n)$, если существует ненулевой минор, окаймляющий минор порядка $(min(m, n) - 1)$.

Пример 5.2. Найти ранг матрицы $A = \begin{pmatrix} 3 & 2 & 3 & 1 \\ -1 & 4 & -1 & 2 \\ 3 & 2 & 3 & 1 \end{pmatrix}$ методом окаймляющих миноров.

Решение. В этой матрице есть ненулевые элементы, значит ее ранг больше нуля. Так как элемент $a_{11} = 3$ матрицы A отличен от нуля, то возьмем его в качестве минора первого порядка. Начнем поиск окаймляющего минора, отличного от нуля: $\begin{vmatrix} 3 & 2 \\ -1 & 4 \end{vmatrix} = 14 \neq 0.$

Находим миноры третьего порядка, окаймляющие данный (их $(3 - 2)\cdot(4 - 2) = 2$ штуки) $\begin{vmatrix} 3 & 2 & 3 \\ -1 & 4 & -1 \\ 3 & 2 & 3 \end{vmatrix} = 0$, $\begin{vmatrix} 3 & 2 & 1 \\ -1 & 4 & 2 \\ 3 & 2 & 1 \end{vmatrix} = 0$. Все они равны нулю, следовательно ранг матрицы A равен двум, $rang\,A = 2$.

5.3. Нахождение ранга матрицы с помощью элементарных преобразований

Рассмотрим еще один способ нахождения ранга матрицы.

Определение 5.4. *Элементарными преобразованиями матрицы* называются следующие преобразования:

1. умножение строки матрицы на число, отличное от нуля;
2. прибавление к элементам какой-либо строки (столбца) матрицы соответствующих элементов другой строки (столбца), умноженных на произвольное число;
3. вычеркивание нулевой строки.

Замечание 5.3. С помощью преобразований 1 и 2 можно поменять местами любые две строки (столбца) матрицы.

Определение 5.5. Матрица A называется *ступенчатой*, если она имеет вид: $\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1r} & \cdots & a_{1k} \\ 0 & a_{22} & \cdots & a_{2r} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{rr} & \cdots & a_{rk} \end{pmatrix}$, $a_{ii} \neq 0$, $i = 1, 2, \ldots, r$, $r \leq k$.

Замечание 5.4. Условие $r \leq k$ всегда может быть достигнуто транспонированием матрицы.

Теорема 5.2. Применение к произвольной матрице цепочки элементарных преобразований не меняет ее ранга.

Теорема 5.3. Любую матрицу с помощью элементарных преобразований можно привести к ступенчатому виду.

Теорема 5.4. Ранг ступенчатой матрицы равен количеству ее ненулевых строк.

Определение 5.6. Первый ненулевой элемент строки называется ее *ведущим элементом*.

Из этих теорем следует практический способ нахождения ранга матрицы: *с помощью элементарных преобразований привести матрицу к ступенчатому виду и определить количество ее ненулевых строк.*·

Пример 5.3. Найти ранг матрицы $A = \begin{pmatrix} -1 & 0 & 3 & -2 \\ 2 & 3 & -1 & -3 \\ 3 & 6 & 1 & -8 \end{pmatrix}$ с помощью элементарных преобразований.

Решение. Приводим матрицу с помощью элементарных преобразований к ступенчатому виду. Выберем в 1-ой строке ведущий элемент. Это (–1). В столбце под этим элементом следует получить нули. Для этого к 2-ой строке прибавим 1-ю, умноженную на 2, а к 3-ей строке прибавим 1-ую, умноженную на 3; получим матрицу: $\begin{pmatrix} -1 & 0 & 3 & -2 \\ 0 & 3 & 5 & -7 \\ 0 & 6 & 10 & -14 \end{pmatrix}$. Выбираем ведущий элемент во второй строке и получим нули в столбце под ним: к 3-ей строке прибавим 2-ую, умноженную на (–2), в результате получим следующую матрицу: $\begin{pmatrix} -1 & 0 & 3 & -2 \\ 0 & 3 & 5 & -7 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Получена матрица ступенчатого вида, в которой две ненулевые строки, следовательно, ранг исходной равен 2, т. е. $rang\ A = 2$.

5.4. Понятие обратной матрицы и способы ее нахождения

Пусть дана квадратная матрица A.

Определение 5.7. Матрица A^{-1} называется *обратной* для матрицы A, если $A \cdot A^{-1} = A^{-1} \cdot A = E$.

Определение 5.8. Квадратная матрица A называется *невырожденной*, если ее определитель не равен нулю.

Заметит, что ранг невырожденной матрицы порядка n равен n.

Определение 5.9. Квадратная матрица A называется *вырожденной*, если ее определитель равен нулю.

Определение 5.10. Матрицей, *присоединенной* к матрице A, называется матрица A^*, где $A^* = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$, A_{ij} – алгебраическое дополнение элемента a_{ij} для всех индексов $i, j = 1, 2, \ldots n$.

Теорема 5.5. Для вырожденной матрицы не существует обратной матрицы.

Теорема 5.6. Для невырожденной матрицы A существует обратная матрица, причем только одна. Обратная матрица может быть найдена по формуле: $A^{-1} = \frac{1}{|A|} \cdot A^*$.

Алгоритм нахождения обратной матрицы

Рассмотрим один из способов нахождения обратной матрицы к данной с помощью алгебраических дополнений. Пусть дана квадратная матрица A.

1. Находим определитель матрицы $|A|$. Если $|A| = 0$, то у матрицы A нет обратной (теорема 5.5). Если $|A| \neq 0$, то обратная матрица существует, и переходим к пункту 2.
2. Находим алгебраические дополнения всех элементов матрицы A.
3. Составляем присоединенную матрицу A^*.
4. Находим A^{-1} по указанной формуле (теорема 5.6).

Пример 5.4. Найти матрицу, обратную для матрицы $A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$.

Решение. Определитель матрицы A равен 1, то есть не равен нулю. Тогда находим алгебраические дополнения элементов матрицы. $A_{11} = 3$, $A_{21} = -5$, $A_{12} = -1$, $A_{22} = 2$. Составляем присоединенную матрицу A^*, получаем $A^* = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$. С учетом формулы $A^{-1} = \frac{1}{|A|} \cdot A^*$ находим обратную матрицу A^{-1}, $A^{-1} = \frac{1}{1}\begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$.

Нахождение обратной матрицы с помощью элементарных преобразований

Рассмотрим еще способ нахождения обратной матицы с помощью элементарных преобразований. Сформулируем необходимые понятия и теоремы.

Определение 5.11. Матрица B называется *эквивалентной* матрице A, если B получена из A с помощью конечного числа элементарных преобразований. Обозначение $B \sim A$.

Теорема 5.7. Всякую невырожденную квадратную матрицу с помощью элементарных преобразований можно привести к единичной матрице.

Теорема 5.8. Если к единичной матрице применить те же самые элементарные преобразования, которые матрицу A переводят в единичную, то полученная матрица будет обратной для матрицы A.

Схематично преобразования выглядят так: $(A|E) \sim \ldots \sim (E|A^{-1})$.

Замечание 5.5. При нахождении матрицы A^{-1} нет необходимости проверять невырожденность матрицы A, т. к. сама возможность привести матрицу A к единичной матрице E будет означать, что A – невырожденная.

Пример 5.5. Найти матрицу, обратную для матрицы $A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$ с помощью элементарных преобразований.

Решение. Приписываем к матрице A (справа, а можно и слева) единичную матрицу E. Далее, с помощью элементарных преобразований над всей составной матрицей приводим матрицу A к единичной E. Тогда на месте первоначально приписанной матрицы E оказывается матрица A^{-1}.

$$\left(\begin{array}{cc|cc} 2 & 5 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 3 & 0 & 1 \\ 2 & 5 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 3 & 0 & 1 \\ 0 & -1 & 1 & -2 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 0 & 3 & -5 \\ 0 & -1 & 1 & -2 \end{array}\right) \sim$$

$$\left(\begin{array}{cc|cc} 1 & 0 & 3 & -5 \\ 0 & 1 & -1 & 2 \end{array}\right); \text{ т. о. } A^{-1} = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}.$$

6. Системы линейных уравнений

6.1. Основные понятия и определения

Определение 6.1. *Системой m линейных уравнений с n* неизвестными называется система вида:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases} \tag{1}$$

Здесь x_1, …, x_n – неизвестные (или переменные), числа a_{ij} – коэффициенты при неизвестных, i – номер уравнения, j – номер неизвестного, b_1, …, b_m – свободные члены.

Короче систему (1) можно записать в виде: $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i$, где $i = 1, 2, \ldots, m$.

С каждой системой вида (1) связаны следующие матрицы: A – *основная* матрица системы, составленная из коэффициентов при неизвестных; B – матрица-столбец свободных членов, X – матрица-столбец неизвестных; $(A|B)$ – *расширенная* матрица системы.

$$A_{m\times n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, B_{m\times 1} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}, X_{n\times 1} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix},$$

$$(A|B)_{m\times(n+1)} = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right).$$

Из определения 6.1 видно, что матрицы A и X согласованы, следовательно можно найти их произведение:

$$A\cdot X = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix}.$$

Если воспользоваться определением 3.4 равенства матриц, то равенство

$$A\cdot X = B \tag{2}$$

записывается в виде системы линейных уравнений (1).

Определение 6.2. Уравнение (2) называют *матричной формой записи системы* (1).

Определение 6.3. *Решением системы* линейных уравнений (1) называется любой упорядоченный набор (кортеж, вектор) $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ из чисел, который при подстановке в систему каждое уравнение обращает в верное равенство.

Таким образом, если $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ решение системы, то следующие равенства верны:

$$\begin{aligned} &a_{11}\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n = b_1, \\ &a_{21}\alpha_1 + a_{22}\alpha_2 + \ldots + a_{2n}\varepsilon_n = b_2, \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ &a_{m1}\alpha_1 + a_{m2}\alpha_2 + \ldots + a_{mn}\alpha_n = b_m. \end{aligned}$$

Определение 6.4. Система линейных уравнений называется *совместной*, если она имеет хотя бы одно решение, и *несовместной*, если она не имеет ни одного решения.

Определение 6.5. Система линейных уравнений называется *определенной*, если она имеет единственное решение, и *неопределенной,* если она имеет более одного решения.

Определение 6.6. Две системы линейных уравнений называются *равносильными*, если множества их решений совпадают.

То есть, если упорядоченный набор чисел $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ является решением первой системы, то он является решением второй и наоборот, если упорядоченный набор чисел является решением второй системы, то он является решением первой системы.

6.2. Методы решения систем линейных уравнений

6.2.1. Метод Крамера

Рассмотрим систему линейных уравнений, в которой число уравнений равно числу неизвестных, то есть $m = n$ и система имеет вид:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n. \end{cases}$$

В этом случае основная матрица системы – квадратная матрица порядка n.

Теорема 6.1 (Крамера[11]). Пусть в системе линейных уравнений число уравнений равно числу неизвестных и определитель основной матрицы не равен нулю. Тогда решение этой системы находят по формулам:

$$x_1 = \frac{|A_1|}{|A|},\ x_2 = \frac{|A_2|}{|A|},\ \ldots,\ x_n = \frac{|A_n|}{|A|}, \qquad (3)$$

[11] Габриэль Крамер (1704–1752) – швейцарский математик.

где матрица A_i получена из матрицы A заменой i-ого столбца столбцом свободных членов.

Формулы (3) называются *формулами Крамера*.

Пример 6.1. Решить систему линейных уравнений по формулам Крамера. Дана система $\begin{cases} 2x_1 + 3x_2 + x_3 = 6, \\ x_1 + 2x_2 - x_3 = 2, \\ -3x_1 - 5x_2 + x_3 = -7. \end{cases}$

Решение. Данной системе соответствует основная матрица $A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & -1 \\ -3 & -5 & 1 \end{pmatrix}$, столбец свободных членов $B = \begin{pmatrix} 6 \\ 2 \\ -7 \end{pmatrix}$ и столбец неизвестных $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. Проверим, можно ли решать данную систему по формулам Крамера. В этой системе число уравнений равно числу неизвестных. Найдем определитель основной матрицы этой системы: $|A| = \begin{vmatrix} 2 & 3 & 1 \\ 1 & 2 & -1 \\ -3 & -5 & 1 \end{vmatrix} = 1$. Этот определитель не равен нулю, поэтому систему можно решать по формулам Крамера. Вычислим определители матриц A_1, A_2, A_3:

$$|A_1| = \begin{vmatrix} 6 & 3 & 1 \\ 2 & 2 & -1 \\ -7 & -5 & 1 \end{vmatrix} = 1,\ |A_2| = \begin{vmatrix} 2 & 6 & 1 \\ 1 & 2 & -1 \\ -3 & -7 & 1 \end{vmatrix} = 1,\ |A_3| = \begin{vmatrix} 2 & 3 & 6 \\ 1 & 2 & 2 \\ -3 & -5 & -7 \end{vmatrix} = 1.$$

По формулам (3) находим неизвестные:

$$x_1 = \frac{|A_1|}{|A|} = \frac{1}{1} = 1,\ x_2 = \frac{|A_2|}{|A|} = \frac{1}{1} = 1,\ x_3 = \frac{|A_3|}{|A|} = \frac{1}{1} = 1.$$

Ответ: (1; 1, 1).

6.2.2. Метод обратной матрицы

Метод обратной матрицы применим для систем линейных уравнений, в которых число уравнений равно числу неизвестных и определитель основной матрицы не равен нулю.

Матричная форма записи системы линейных уравнений представляется в виде следующего матричного равенства: $A \cdot X = B$.

В силу условия матрица A – квадратная матрица порядка n с определителем не равным нулю. Это означает, что для матрицы A существует обратная матрица A^{-1}. Умножим обе части матричного равенства на матрицу A^{-1} слева. Получим $A^{-1} \cdot (A \cdot X) = A^{-1} \cdot B$. Преобразуем данной выражение:

$(A^{-1} \cdot A) \cdot X = A^{-1} \cdot B$;

$E \cdot X = A^{-1} \cdot B$;

$X = A^{-1} \cdot B$.

Вывод: если для системы n линейных уравнений с n неизвестными определитель основной матрицы не равен нулю, то система имеет единственное решение, которое находится по формуле $X = A^{-1} \cdot B$, где A – основная матрица данной системы, B – столбец свободных членов.

Пример 6.2. Решить систему линейных уравнений $\begin{cases} 2x_1 - 3x_2 = -1, \\ 4x_1 + 7x_2 = 11 \end{cases}$ методом обратной матрицы.

Решение. Здесь $A = \begin{pmatrix} 2 & -3 \\ 4 & 7 \end{pmatrix}$, $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $B = \begin{pmatrix} -1 \\ 11 \end{pmatrix}$. Найдем матрицу A^{-1} любым способом. Имеем $A^{-1} = \dfrac{1}{26}\begin{pmatrix} 7 & 3 \\ -4 & 2 \end{pmatrix}$. Теперь можно вычислить столбец неизвестных X.

$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \dfrac{1}{26}\begin{pmatrix} 7 & 3 \\ -4 & 2 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 11 \end{pmatrix} = \dfrac{1}{26}\begin{pmatrix} 26 \\ 26 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Значит $x_1 = 1$, $x_2 = 1$.

Ответ: (1; 1).

Очевидно, что применение этих методов связано с выполнением определенных условий и решить с их помощью произвольную систему невозможно.

6.2.3. Метод Гаусса

Для описания этого метода, который годится для решения произвольных систем линейных уравнений, необходимы некоторые новые понятия.

Определение 6.7. Уравнение вида $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = 0$ называется *нулевым*.

Решением такого уравнения является любой вектор.

Определение 6.8. Уравнение вида $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = b$, где $b \neq 0$, называется *несовместным* (или *противоречивым*).

У несовместного уравнения решений нет.

Определение 6.9. *Элементарными преобразованиями* системы линейных уравнений называются следующие ее преобразования:

1. умножение любого уравнения на число, не равное нулю;
2. прибавление к одному уравнению системы любого другого, умноженного на произвольное число;
3. вычеркивание нулевого уравнения.

Замечание 6.1. С помощью преобразований 1 и 2 уравнения системы можно поменять местами.

Теорема 6.2. Цепочка элементарных преобразований переводит исходную систему линейных уравнений в равносильную ей систему.

Описание метода Гаусса

Метод Гаусса – метод последовательного исключения неизвестных – заключается в том, что с помощью элементарных преобразований исходная система приводится к равносильной ей системе ступенчатого или треугольного вида, из которой последовательно, начиная с последних (по номеру), неизвестных находятся все остальные неизвестные. Дана система (1)

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases} \qquad (1)$$

Начинаем осуществлять *прямой ход*. Считаем, что коэффициент $a_{11} \neq 0$; если же это не так, меняем местами уравнения.

Первый шаг состоит в том, чтобы исключить неизвестное x_1 из всех уравнений, кроме первого. Для этого ко второму уравнению прибавим первое уравнение, умноженное на число $\left(-\frac{a_{21}}{a_{11}}\right)$, к третьему уравнению прибавим первое уравнение, умноженное на число $\left(-\frac{a_{31}}{a_{11}}\right)$, и так далее до последнего уравнения. После первого шага получим систему:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ \qquad a'_{22}x_2 + \ldots + a'_{2n}x_n = b'_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad a'_{m2}x_2 + \ldots + a'_{mn}x_n = b'_m. \end{cases}$$

Полученная система равносильна исходной системе.

Вторым шагом исключают неизвестное из всех уравнений, кроме первого и второго. Для этого повторяем все действия первого шага для второго и последующих уравнений, а именно: считаем, что коэффициент $a'_{22} \neq 0$ и так далее. Если в результате преобразований получается нулевое уравнение, то его удаляют, если же получается несовместное уравнение, то решение системы закончено – она несовместна. Процесс исключения неизвестных продолжаем до тех пор, пока это возможно. Обозначим количество уравнений, оставшихся после прямого хода, через r. Это число равно рангу основной матрицы системы и может быть меньше или равно n. Рассмотрим оба случая.

1) Если $r = n$, то система после прямого хода принимает вид:

$$\begin{cases} c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n = d_1, \\ \qquad\quad c_{22}x_2 + \ldots + c_{2n}x_n = d_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad\qquad\qquad c_{nn}x_n = d_n, \end{cases}$$

где $c_{11} \neq 0$, $c_{22} \neq 0$, …, $c_{nn} \neq 0$.

Обратным ходом, начиная с последнего уравнения, последовательно найдем значения x_n, (где $x_n = \dfrac{d_n}{c_{nn}}$), x_{n-1}, ..., x_1. В этом случае система линейных уравнений имеет единственное решение, то есть является определенной.

2) Если $r < n$, то система после прямого хода принимает вид:

$$\begin{cases} c_{11}x_1 + c_{12}x_2 + \ldots + c_{1r}x_r + \ldots + c_{1n}x_n = d_1, \\ \qquad\quad c_{22}x_2 + \ldots + c_{2r}x_r + \ldots + c_{2n}x_n = d_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad\qquad\quad c_{rr}x_r + \ldots + c_{rn}x_n = d_n, \end{cases}$$

где $c_{11} \neq 0$, $c_{22} \neq 0$, …, $c_{rr} \neq 0$. Неизвестные x_1, x_2, …, x_r, с которых начинаются уравнения, называются *главными неизвестными*, а остальные x_{r+1}, x_{r+2}, …, x_n – *свободными*. В этом случае обратным ходом, начиная с последнего уравнения, выражают главные неизвестные через свободные неизвестные. Получают следующие равенства:

$$x_1 = k_{1,r+1}x_{r+1} + \ldots + k_{1,n}x_n + t_1,$$

$$x_2 = k_{2,r+1}x_{r+1} + \ldots + k_{2,n}x_n + t_2,$$

……………………………………..

$$x_r = k_{r,r+1}x_{r+1} + \ldots + k_{r,n}x_n + t_r.$$

Определение 6.10. *Общим решением системы* называется выражение главных неизвестных через свободные.

Если свободным неизвестным придать какие-нибудь числовые значения, то из общего решения получим значения главных неизвестных. Таким образом, получают *частное решение* системы. Из способа его получения следует, что система имеет более одного решения, то есть является неопределенной.

Пример 6.3. Решить методом Гаусса систему линейных уравнений:

$$\begin{cases} x_1 - 2x_2 + x_3 + x_4 = 1, \\ x_1 - 2x_2 - x_3 - 2x_4 = 2, \\ 3x_1 - 6x_2 - x_3 - 3x_4 = 5. \end{cases}$$

Решение. Преобразования с системой линейных удобнее производить не с самими уравнениями, а с матрицей их коэффициентов. Расширенная матрица этой системы имеет вид: $(A|B) = \left(\begin{array}{cccc|c} 1 & -2 & 1 & 1 & 1 \\ 1 & -2 & -1 & -2 & 2 \\ 3 & -6 & -1 & -3 & 5 \end{array}\right)$.

Осуществляем прямой ход. Первым шагом исключаем неизвестное x_1 из всех уравнений, кроме первого. Так как $a_{11} = 1 \neq 0$, то переставлять уравнения местами не нужно. Прибавим ко второму уравнению системы первое уравнение, умноженное на (–1), к третьему уравнению – первое, умноженное на (–3). Получим после преобразований следующую матрицу: $\left(\begin{array}{cccc|c} 1 & -2 & 1 & 1 & 1 \\ 0 & 0 & -2 & -3 & 1 \\ 0 & 0 & -4 & -6 & 2 \end{array}\right)$, в которой элемент $a_{22} = 1$. Перестановка местами уравнений (первое уравнение трогать не следует) не поможет, поэтому переходим к следующему неизвестному x_3 и исключаем его из всех уравнений, кроме первого и второго. Для этого к третьему уравнению прибавим второе, умноженное на (–2) и вычеркнем получившееся нулевое уравнение. После прямого хода получаем следующую систему: $\left(\begin{array}{cccc|c} 1 & -2 & 1 & 1 & 1 \\ 0 & 0 & -2 & -3 & 1 \end{array}\right)$. Прямой ход завершен. В этом случае $n = 4$, $r = 2$,

$r < n$, и, следовательно, система неопределенная. Главные неизвестные – это те неизвестные, с которых начинаются уравнения, в нашем случае это x_1 и x_3. Неизвестные x_2 и x_4 – свободные.

Обратным ходом надо выразить главные неизвестные через свободные. Для этого в столбцах, содержащих ведущие элементы строк, следует получить нули. Здесь это элемент a_{13}. Прибавим к первому уравнению, умноженному на 2, второе и выпишем получившуюся матрицу коэффициентов: $\left(\begin{array}{cccc|c} 2 & -4 & 0 & -1 & 3 \\ 0 & 0 & -2 & -3 & 1 \end{array}\right)$, а затем и сами уравнения: $\begin{cases} 2x_1 - 4x_2 - x_4 = 3, \\ -2x_3 - 3x_4 = 1. \end{cases}$ Из этих уравнений получаем общее решение:

$$\begin{cases} x_1 = \frac{1}{2}(2x_2 + 4x_4 + 3), \\ x_3 = \frac{1}{2}(-3x_4 - 1). \end{cases}$$

Найдем какое-нибудь частное решение; пусть $x_2 = 3$, $x_4 = 1$, тогда из общего решения получим значения $x_1 = \frac{13}{2}$, и $x_1 = -2$. Таким образом, частное решение – вектор $a = (\frac{13}{2}, 3, -2, 1)$.

Ответ: общее решение $\{(x_2 + 2x_4 + \frac{3}{2}, x_2, -\frac{3}{2}x_4 - \frac{1}{2}, x_4)\}$, где $x_2, x_4 \in \mathbb{R}$;

частное решение, если $x_2 = 3$, $x_4 = 1$, то $(\frac{13}{2}, 3, -2, 1)$.

6.3. Исследование системы линейных уравнений

Исследовать систему линейных уравнений – это значит, не решая систему, ответить на вопрос: совместна система или нет, а если совместна, то, сколько у нее решений. Ответить на этот вопрос позволяет следующая теорема.

Теорема 6.3 (Кронекера[12]–Капелли[13]). Система линейных уравнений совместна тогда и только тогда, когда ранг основной матрицы системы равен рангу ее расширенной матрицы, то есть $r(A) = r(A|B)$.

Для совместных систем линейных уравнений верны следующие утверждения.

Теорема 6.4. Если ранг матрицы совместной системы равен числу неизвестных, то есть $r(A) = r(A|B) = n$, то эта система определенная, имеет единственное решение.

Теорема 6.5. Если ранг матрицы совместной системы меньше числа неизвестных, то есть $r(A) = r(A|B) < n$, то система неопределенная и имеет бесконечно много решений; причем количество свободных неизвестных равно $(n – r)$.

Пример 6.4. Исследовать систему линейных уравнений на совместность: $\begin{cases} 2x_1 - x_2 + x_3 + x_4 = 1, \\ x_1 + 2x_2 - x_3 + 4x_4 = 2, \\ x_1 + 7x_2 - 4x_3 + 11x_4 = 3. \end{cases}$

Решение. Так как в системе число уравнений меньше числа неизвестных, то применим только метод Гаусса. Приводим расширенную матрицу этой системы к ступенчатому виду с помощью элементарных преобразований.

$$\left(\begin{array}{cccc|c} 2 & -1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 4 & 2 \\ 1 & 7 & -4 & 11 & 3 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 2 & -1 & 4 & 2 \\ 2 & -1 & 1 & 1 & 1 \\ 1 & 7 & -4 & 11 & 3 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 2 & -1 & 4 & 2 \\ 0 & -5 & 3 & -7 & -3 \\ 0 & 5 & -3 & 7 & 1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{cccc|c} 1 & 2 & -1 & 4 & 2 \\ 0 & -5 & 3 & -7 & -3 \\ 0 & 0 & 0 & 0 & -2 \end{array}\right).$$

[12] Леопольд Кронекер (1823–1891) – немецкий математик.
[13] Альфредо Капелли (1855–1910) – итальянский математик.

Получили, что ранг основной матрицы равен двум, а расширенной матрицы – трем, следовательно, данная система линейных уравнений несовместна.

6.4. Однородные системы линейных уравнений

Определение 6.11. Система линейных уравнений называется *однородной*, если ее свободные члены равны нулю.

Однородная система m линейных уравнений с n неизвестными имеет вид:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0. \end{cases} \qquad (4)$$

Однородная система линейных уравнений всегда совместна, так как нулевой вектор $0 = (0,\ 0,\ \ldots,\ 0)$ является решением этой системы. Если однородная система линейных уравнений определенная ($r(A) = n$), то есть имеет единственное решение, то это решение нулевое (не нулевых решений нет). Если же однородная система неопределенная ($r(A) < n$), то кроме нулевого решения у этой системы есть ненулевые решения.

Свойства решений однородной системы линейных уравнений

1. Если вектор $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ является решением однородной системы, то вектор $k \cdot a = (k \cdot \alpha_1, k \cdot \alpha_2, \ldots, k \cdot \alpha_n)$ также является решением этой системы, где k – любое число.
2. Если векторы $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ и $b = (\beta_1, \beta_2, \ldots, \beta_n)$ являются решениями однородной системы, то вектор $a + b = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n)$ также является решением этой системы.

Следствие. Если a_1, a_2, …, a_p – решения однородной системы (4), то $k_1 \cdot a_1 + k_2 \cdot a_2 + \ldots + k_p \cdot a_p$ – тоже решение системы (4), где k_1, k_2, …, k_p – любые числа.

Фундаментальный набор решений однородной системы линейных уравнений

Пусть M_0 – множество решений однородной системы (4) линейных уравнений.

Определение 6.12. Векторы c_1, c_2, …, c_p, являющиеся решениями однородной системы линейных уравнений называются *фундаментальным набором решений* (сокращенно ФНР), если

1) векторы c_1, c_2, …, c_p линейно независимы (т. е. ни один из них нельзя выразить через другие);

2) любое другое решение однородной системы линейных уравнений можно выразить через решения c_1, c_2, …, c_p.

Заметим, что если c_1, c_2, …, c_p – какой-либо ф.н.р., то выражением $k_1 \cdot c_1 + k_2 \cdot c_2 + \ldots + k_p \cdot c_p$ можно описать все множество M_0 решений системы (4), поэтому его называют *общим видом решения системы* (4).

Теорема 6.6. Любая неопределенная однородная система линейных уравнений обладает фундаментальным набором решений.

Способ нахождения фундаментального набора решений состоит в следующем:

• найти общее решение однородной системы линейных уравнений;

• построить $(n - r)$ частных решений этой системы, при этом значения свободных неизвестных должны образовывать единичную матрицу;

• выписать общий вид решения, входящего в M_0.

Пример 6.5. Найти фундаментальный набор решений следующей системы: $\begin{cases} 3x_1 + 2x_2 + x_3 + 3x_4 + 5x_5 = 0, \\ 6x_1 + 4x_2 + 3x_3 + 5x_4 + 7x_5 = 0, \\ 9x_1 + 6x_2 + 5x_3 + 7x_4 + 9x_5 = 0, \\ 3x_1 + 2x_2 + 4x_4 + 8x_5 = 0. \end{cases}$

Решение. Найдем общее решение этой системы.

$$\begin{pmatrix} 3 & 2 & 1 & 3 & 5 \\ 6 & 4 & 3 & 5 & 7 \\ 9 & 6 & 5 & 7 & 9 \\ 3 & 2 & 0 & 4 & 8 \end{pmatrix} \sim \begin{pmatrix} 3 & 2 & 1 & 3 & 5 \\ 0 & 0 & 1 & -1 & -3 \\ 0 & 0 & 2 & -2 & -6 \\ 0 & 0 & -1 & 1 & 3 \end{pmatrix} \sim \begin{pmatrix} 3 & 2 & 1 & 3 & 5 \\ 0 & 0 & 1 & -1 & -3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 3 & 2 & 1 & 3 & 5 \\ 0 & 0 & 1 & -1 & -3 \end{pmatrix} \Rightarrow \begin{cases} 3x_1 + 2x_2 + x_3 + 3x_4 + 5x_5 = 0, \\ \qquad\qquad x_3 - 4x_4 - 3x_5 = 0. \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} x_1 = \frac{1}{3}(-2x_2 - 4x_4 - 8x_5, \\ x_3 = 4x_4 + 3x_5. \end{cases}$$ В этой системе пять неизвестных ($n = 5$), из них главных неизвестных два ($r = 2$), свободных неизвестных три ($n - r$), то есть в фундаментальном наборе решений содержится три вектора решения. Построим их. Имеем x_1 и x_3 – главные неизвестные, x_2, x_4, x_5 – свободные неизвестные

	x_1	$\boldsymbol{x_2}$	x_3	$\boldsymbol{x_4}$	$\boldsymbol{x_5}$
c_1	$-\frac{2}{3}$	**1**	0	**0**	**0**
c_2	$-\frac{4}{3}$	**0**	1	**1**	**0**
c_3	$-\frac{8}{3}$	**0**	3	**0**	**1**

Значения свободных неизвестных x_2, x_4, x_5 образуют единичную матрицу E третьего порядка. Получили, что векторы c_1, c_2, c_3 образуют ф.н.р. данной системы. Тогда множество решений данной однородной системы будет $M_0 = \{k_1 \cdot c_1 + k_2 \cdot c_2 + k_3 \cdot c_3,\ k_1, k_2, k_3 \in \mathbb{R}\}$.

Выясним теперь условия существования ненулевых решений однородной системы линейных уравнений, другими словами условия существования фундаментального набора решений.

Однородная система линейных уравнений имеет ненулевые решения, то есть является неопределенной, если

1) ранг основной матрицы системы меньше числа неизвестных;
2) в однородной системе линейных уравнений число уравнений меньше числа неизвестных;
3) если в однородной системе линейных уравнений число уравнений равно числу неизвестных, и определитель основной матрицы равен нулю (т. е. $|A| = 0$).

Пример 6.6. При каком значении параметра a однородная система линейных уравнений $\begin{cases} x_1 - x_2 + 2x_3 = 0, \\ ax_2 + x_3 = 0, \\ x_1 + 3x_2 + x_3 = 0 \end{cases}$ имеет ненулевые решения?

Решение. Составим основную матрицу этой системы и найдем ее определитель: $\begin{vmatrix} 1 & -1 & 2 \\ 0 & a & 1 \\ 1 & 3 & 1 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 2 \\ 0 & a & 1 \\ 0 & 4 & -1 \end{vmatrix} = 1\cdot(-1)^{1+1}\cdot\begin{vmatrix} a & 1 \\ 4 & -1 \end{vmatrix} = -a - 4.$

Определитель этой матрицы равен нулю при $a = -4$.

Ответ: –4.

7. Арифметическое *n*-мерное векторное пространство

7.1. Основные понятия

В предыдущих разделах уже встречалось понятие о наборе из действительных чисел, расположенных в определенном порядке. Это матрица-строка (или матрица-столбец) и решение системы линейных уравнений с n неизвестными. Эти сведения можно обобщить.

Определение 7.1. *n-мерным арифметическим вектором* называется упорядоченный набор из n действительных чисел.

Значит $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, где $\alpha_i \in \mathbb{R}$, $i = 1, 2, \ldots, n$ – общий вид вектора. Число n называется *размерностью* вектора, а числа α_i называются его *координатами*.

Например: $a = (1, -8, 7, 4, \frac{1}{3})$ – пятимерный вектор.

Все множество n-мерных векторов принято обозначать как R^n.

Определение 7.2. Два вектора $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ и $b = (\beta_1, \beta_2, \ldots, \beta_n)$ одинаковой размерности *равны* тогда и только тогда, когда равны их соответствующие координаты, т. е. $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_n = \beta_n$.

Определение 7.3. *Суммой* двух n-мерных векторов $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ и $b = (\beta_1, \beta_2, \ldots, \beta_n)$ называется вектор $a + b = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n)$.

Определение 7.4. *Произведением* действительного числа k на вектор $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ называется вектор $k\cdot a = (k\cdot\alpha_1, k\cdot\alpha_2, \ldots, k\cdot\alpha_n)$

Определение 7.5. Вектор $o = (0, 0, \ldots, 0)$ называется *нулевым* (или *нуль–вектором*).

Легко проверить, что действия (операции) сложения векторов и умножения их на действительное число обладают следующими свойствами: $\forall\, a, b, c \in R^n$, $\forall\, k, l \in \mathbb{R}$:

1) $a + b = b + a$;
2) $a + (b + c) = (a + b) + c$;
3) $a + o = a$;
4) $a + (-a) = o$;
5) $1\cdot a = a$, $1 \in \mathbb{R}$;
6) $k\cdot(l\cdot a) = l\cdot(k\cdot a) = (l\cdot k)\cdot a$;
7) $(k + l)\cdot a = k\cdot a + l\cdot a$;
8) $k\cdot(a + b) = k\cdot a + k\cdot b$.

Определение 7.6. Множество R^n с заданными на нем операциями сложения векторов и умножения их на действительное число называется *арифметическим n-мерным векторным пространством*.

7.2. Линейная зависимость и независимость системы векторов

Пусть a_1, a_2, …, a_m множество из m штук n-мерных векторов, о котором принято говорить – *система векторов*, и k_1, k_2, …, k_m – произвольные действительные числа.

Определение 7.7. *Линейной комбинацией* системы векторов $a_1, a_2, \ldots, a_m$ с коэффициентами $k_1, k_2, \ldots, k_m$ называется вектор $b = k_1 \cdot a_1 + k_2 \cdot a_2 + \ldots + k_m \cdot a_m$.

Принято говорить: вектор b линейно выражается через векторы $a_1, a_2, \ldots, a_m$ или вектор b разложен (раскладывается) по векторам $a_1, a_2, \ldots, a_m$.

Пример 7.1. Даны векторы $a_1 = (3, 2, -1, 0)$, $a_2 = (-1, 0, 4, 1)$, $a_3 = (-2, -2, -3, -1)$. Найти вектор $b = 2a_1 - a_2 - a_3$.

Решение. $b = 2a_1 - a_2 - a_3 = 2(3, 2, -1, 0) + (-1)(-1, 0, 4, 1) + (-1)(-2, -2, -3, -1) = (6, 4, -2, 0) + (1, 0, -4, -1) + (2, 2, 3, 1) = (9, 6, -3, 0)$.

Пример 7.2. Даны векторы $a_1 = (6, 4, -2)$, $a_2 = (-1, 0, 4)$, $a_3 = (-2, -2, -3)$. Найти вектор $b = a_1 + 2a_2 + 2a_3$.

Решение. $b = a_1 + 2a_2 + 2a_3 = (6, 4, -2) + 2(-1, 0, 4) + 2(-2, -2, -3) =$
$= (6, 4, -2) + (-2, 0, 8) + (-4, -4, -6) = (0, 0, 0) = o$.

Определение 7.8. *Линейной оболочкой* системы векторов a_1, a_2, …, a_m называется множество всех линейных комбинаций этих векторов. Принятое обозначение: $L(a_1, a_2, \ldots, a_m)$.

Из определения следует, что

$$L(a_1, a_2, \ldots, a_m) = \{k_1 \cdot a_1 + k_2 \cdot a_2 + \ldots + k_m \cdot a_m, k_i \in \mathbb{R}\}.$$

Если вектор b линейно выражается через векторы a_1, a_2, …, a_m, то в этих обозначениях можно записать, что $b \in L(a_1, a_2, \ldots, a_m)$.

Определение 7.9. Линейная комбинация системы векторов a_1, a_2, …, a_m вида $0 \cdot a_1 + 0 \cdot a_2 + \ldots + 0 \cdot a_m$ называется *нулевой*. Нулевая линейная комбинация векторов равна нулевому вектору.

Определение 7.10. Система векторов a_1, a_2, …, a_m называется *линейно независимой*, если линейная комбинация этих векторов равна нулевому вектору тогда и только тогда, когда эта комбинация нулевая, то есть

$$k_1 \cdot a_1 + k_2 \cdot a_2 + \ldots + k_m \cdot a_m = o \Leftrightarrow k_1 = 0,\ k_2 = 0,\ \ldots,\ k_m = 0.$$

Равносильное определение линейно независимой системы векторов звучит следующим образом: система векторов линейно независима тогда и только тогда, когда ни один вектор нельзя выразить через остальные векторы.

Определение 7.11. Система векторов a_1, a_2, …, a_m называется *линейно зависимой*, если существуют коэффициенты k_1, k_2,…, k_m, не все одновременно равные нулю и такие, что $k_1 \cdot a_1 + k_2 \cdot a_2 + \ldots + k_m \cdot a_m = o$.

Равносильное определение линейно зависимой системы векторов звучит следующим образом: система векторов линейно зависима тогда и только тогда, когда в этой системе существует хотя бы один вектор, который линейно выражается через остальные.

Если система векторов состоит только из одного вектора, то эта система линейно зависима, если этот вектор нулевой, и линейно независима, если он ненулевой.

Система векторов, содержащая два вектора, линейно зависима в случае пропорциональности координат этих векторов, и линейно независима в противном случае.

Свойства линейной зависимости системы векторов

1) Система векторов, содержащая нулевой вектор, линейно зависима.

2) Система векторов линейно зависима, если какая-нибудь ее подсистема линейно зависима.

Следствие. Если система векторов линейно независима, то и любая ее подсистема линейно независима.

3) Если к линейно независимой системе векторов a_1, a_2, …, a_m добавить вектор b и при этом система векторов a_1, a_2, …, a_m, b станет линейно зависимой, то вектор b линейно выражается через остальные.

Определение7.12. Система векторов a_1, a_2, …, a_m называется *ступенчатой* (или *лестничной*), если матрица, составленная из координат этих векторов, является ступенчатой, т. е.

$$a_1 = (a_{11}, a_{12}, \ldots, a_{1r}, \ldots, a_{1n}),\ a_{11} \neq 0;$$
$$a_2 = (0, a_{22}, \ldots, a_{2r}, \ldots, a_{2n}),\ a_{22} \neq 0;$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_m = (0, 0, \ldots, a_{rr}, \ldots, a_{rn}),\ a_{rr} \neq 0.$$

4) Ступенчатая система векторов линейно независима.

Пример 7.3. Выяснить является ли система векторов $a_1 = (2, 2, 7, -1)$, $a_2 = (3, -1, 2, 4)$, $a_3 = (1, 1, 3, 1)$ линейно зависимой.

Решение. По определениям 7.10 и 7.11 составим линейную комбинацию данных векторов $k_1 \cdot a_1 + k_2 \cdot a_2 + k_3 \cdot a_3 = 0$; подставим вместо векторов их координаты:

$$k_1 \cdot (2, 2, 7, -1) + k_2 \cdot (3, -1, 2, 4) + k_3 \cdot (1, 1, 3, 1) = (0, 0, 0, 0).$$

Данное векторное равенство (согласно определению 7.2 о равенстве векторов) запишем для каждой координаты, и получим систему:

$$\begin{cases} 2k_1 + 3k_2 + k_3 = 0, \\ 2k_1 + (-1)k_2 + k_3 = 0, \\ 7k_1 + 2k_2 + 3k_3 = 0, \\ (-1)k_1 + 4k_2 + k_3 = 0. \end{cases}$$

Решим систему методом Гаусса.

$$\begin{pmatrix} 2 & 3 & 1 \\ 3 & -1 & 1 \\ 7 & 2 & 3 \\ -1 & 4 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -4 & -1 \\ 2 & 3 & 1 \\ 2 & -1 & 1 \\ 7 & 2 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & -4 & -1 \\ 0 & 11 & 3 \\ 0 & 7 & 3 \\ 0 & 30 & 10 \end{pmatrix} \sim \begin{pmatrix} 1 & -4 & -1 \\ 0 & 11 & 3 \\ 0 & 0 & 12 \\ 0 & 0 & 3 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & -4 & -1 \\ 0 & 11 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -4 & -1 \\ 0 & 11 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -4 & -1 \\ 0 & 11 & 3 \\ 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow \begin{cases} k_1 = 0, \\ k_2 = 0, \\ k_3 = 0. \end{cases}$$

Следовательно, данная система векторов линейно независима.

Единичная система векторов

Определение 7.13. *Системой единичных векторов* пространства R^n называется система векторов e_1, e_2, …, e_n, где $e_1 = (1, 0, …, 0)$, $e_2 = (0, 1, …, 0)$, …, $e_n = (0, 0, …, 1)$.

Выпишем единичные векторы для пространств R^3 и R^4:

$e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$;

$e_1 = (1, 0, 0, 0)$, $e_2 = (0, 1, 0, 0)$, $e_3 = (0, 0, 1, 0)$, $e_4 = (0, 0, 0, 1)$.

Из свойств линейной зависимости следует, что единичная система векторов линейно независима. Любой вектор $a = (\alpha_1, \alpha_2, …, \alpha_n)$ пространства R^n может быть представлен в виде линейной комбинации единичных векторов. Например, если $a = (–2, 4, 7)$, то $a = (–2)e_1 + 4e_2 + 7e_3$.

Вывод. В пространстве R^n существует n линейно независимых векторов, через которые линейно выражаются все векторы пространства R^n.

Две теоремы о линейной зависимости

Теорема 7.1. Если большая система векторов линейно выражается через меньшую, то большая система линейно зависима.

Сформулируем эту теорему подробнее: пусть a_1, a_2, …, a_m и b_1, b_2, …, b_k две системы векторов и $m > k$, то есть первая система большая. Если a_1, a_2, …, $a_m \in L(b_1, b_2, …, b_k)$, то система векторов a_1, a_2, …, a_m линейно зависима.

Теорема 7.2. В пространстве R^n любая система, состоящая более чем из n векторов, линейно зависима.

Это следует из того, что в R^n любая система выражается через систему e_1, e_2, …, e_n.

Следствие. Если a_1, a_2, …, $a_m \in L(b_1, b_2, …, b_k)$ и система a_1, a_2, …, a_m линейно независима, то $m \le k$.

7.3. Базис и ранг системы векторов

Пусть S – система векторов пространства R^n; она может быть как конечной, так и бесконечной. S' – подсистема системы S, $S' \subset S$. Дадим два равносильных определения.

Определение 7.14. *Базисом* системы S называется такая ее подсистема S', что

1) система S' линейно независима;
2) каждый вектор системы S линейно выражается через векторы системы S'.

Определение 7.15. *Базисом* системы S называется максимальная линейно независимая ее подсистема S', то есть

1) система S' линейно независима;
2) если к S' добавить любой вектор из системы S, то получится линейно зависимая система.

Рассмотрим линейно независимую систему векторов; она совпадает со своей максимальной линейно независимой подсистемой. Это означает, что базис такой системы совпадает с ней самой. Базис ступенчатой системы векторов тоже совпадает с ней самой, в силу ее линейной независимости.

Теорема 7.3. Два различных базиса одной и той же системы векторов содержат одинаковое количество векторов.

Доказательство. Пусть S –данная система векторов. Векторы a_1, a_2, …, a_m – базис S' системы S, векторы b_1, b_2, …, b_k – базис S'' системы S. Так как a_1, a_2, …, a_m – базис, то $b_1, b_2, \ldots, b_k \in L(a_1, a_2, \ldots, a_m)$ и b_1, b_2, …, b_k линейно независимы, тогда по следствию из двух терем 7.1 и 7.2 $k \leq m$.

Так как b_1, b_2, …, b_k – базис, то $a_1, a_2, \ldots, a_m \in L(b_1, b_2, \ldots, b_k)$ и a_1, a_2, …, a_m – линейно независимы, тогда по тому же следствию $m \leq k$, и окончательно получаем, что $m = k$. Теорема доказана.

Базис пространства R^n

Векторы e_1, e_2, …, e_n образуют базис пространства R^n, так как e_1, e_2, …, e_n – линейно независимы, и каждый вектор из R^n линейно выражается через эти векторы. По предыдущей теореме 7.3 в другом базисе R^n должно быть столько же векторов, сколько и в этом, то есть n. Сформулируем этот вывод в виде теоремы.

Теорема 7.4. Базисы пространства R^n – это в точности все линейно независимые системы, состоящие из n векторов.

Другими словами, система, состоящая из n линейно независимых векторов – это базис, и наоборот, базис R^n – это система, состоящая из n линейно независимых векторов.

Ранг системы векторов

Дадим два равносильных определения ранга системы векторов.

Определение 7.16. *Рангом* системы векторов называется количество векторов в любом базисе этой системы.

Определение 7.17. *Рангом* системы векторов называется максимальное число линейно независимых векторов в этой системе.

Если дана система векторов a_1, a_2, …, a_m пространства R^n, то ранг r этой системы не больше количества векторов в ней и не больше размерности пространства R^n, то есть $r \le m$ и $r \le n$, или $r \le min(m, n)$.

Ранг линейно независимой системы векторов равен количеству векторов в ней.

Ранг ступенчатой системы векторов равен количеству векторов в ней.

Лемма 7.1 (о ранге системы векторов). Ранг системы векторов не изменится, если к ней добавить (или удалить) вектор, являющийся линейной комбинацией остальных.

На основании этой леммы можно доказать теорему.

Теорема 7.5. Ранг системы векторов не меняется при следующих преобразованиях, называемых элементарными:

1) умножение вектора на число, не равное нулю;

2) прибавление к одному вектору другого, умноженного на произвольное число;
3) удаление нулевого вектора.

С помощью этих преобразований векторы системы можно поменять местами.

Практическое нахождение ранга и базиса системы векторов

Из данной системы векторов составляем матрицу, расположив векторы как строки этой матрицы. Приводим матрицу к ступенчатому виду с помощью элементарных преобразований над строками этой матрицы. При этом не меняется ни ранг матрицы, ни ранг системы векторов-строк. Ранг полученной ступенчатой матрицы, а также полученной ступенчатой системы векторов равен количеству оставшихся ненулевых строк. Базисом системы векторов являются те векторы, на месте которых остались ненулевые строки.

Пример 7.4. Найти ранг и базис системы векторов $a_1 = (1, 3, 0, 5)$, $a_2 = (1, 2, 0, 4)$, $a_3 = (1, 1, 1, 3)$ $a_4 = (1, 0, -1, 0)$, $a_5 = (1, -3, 3, -1)$.

Решение. Действуем по описанной схеме.

$$\begin{pmatrix} 1 & 3 & 0 & 5 \\ 1 & 2 & 0 & 4 \\ 1 & 1 & 1 & 3 \\ 1 & 0 & -1 & 0 \\ 1 & -3 & 3 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & 0 & 5 \\ 0 & -1 & 0 & -1 \\ 0 & -2 & 1 & -2 \\ 0 & -3 & -1 & -1 \\ 0 & -6 & 3 & -6 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & 0 & 5 \\ 0 & 1 & 0 & 1 \\ 0 & 2 & -1 & 2 \\ 0 & 3 & 1 & 1 \\ 0 & 2 & -1 & 2 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 3 & 0 & 5 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & 0 & 5 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & 0 & 5 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Ранг системы векторов равен трем (по количеству оставшихся ненулевых строк), один из базисов образуют векторы a_1, a_2, a_3.

Нахождение ранга системы векторов позволяет решать вопрос о линейной зависимости системы векторов. Если ранг системы векторов равен количеству векторов в системе, то эта система линейно независима, если же ранг системы векторов меньше количества векторов в системе, то эта система векторов линейно зависима.

Так как ранг рассмотренной системы (пример 7.4) векторов a_1, a_2, a_3, a_4, a_5 равен трем и меньше числа векторов, то есть пяти, то система векторов a_1, a_2, a_3, a_4, a_5 линейно зависима.

8. Векторные (линейные) пространства

8.1. Определение векторного пространства над произвольным полем.

Пусть P – произвольное поле. Известные нам примеры полей – поле рациональных, действительных, комплексных чисел.

Определение 8.1. Множество V называется *векторным* (или *линейным*) *пространством над полем* P, если для каждых двух элементов $a, b \in V$ определена *сумма* $a + b \in V$,и для каждого $k \in P$ и для каждого $a \in V$ определено *произведение* $k \cdot a \in V$, причем справедливы следующие равенства: для любых $a, b, c \in V$ и любых $k, l \in P$

1) $a + b = b + a$;
2) $a + (b + c) = (a + b) + c$;
3) $\exists\, o \in V : a + o = a$;
4) $\forall\, a, \exists\, (-a) : a + (-a) = o$;
5) $1 \cdot a = a,\ 1 \in P$;
6) $k \cdot (l \cdot a) = l \cdot (k \cdot a) = (l \cdot k) \cdot a$;
7) $(k + l) \cdot a = k \cdot a + l \cdot a$;
8) $k \cdot (a + b) = k \cdot a + k \cdot b$.

Элементы векторного пространства принято называть *векторами*, o - *нулевой* вектор; $(-a)$ – вектор, *противоположный* вектору a; $1 \in P$ – единица поля P.

Примеры 8.1. Приведем примеры векторных пространств.

1) R^n – арифметическое n-мерное векторное пространство.

2) Множество матриц одного итого же размера с действительными коэффициентами $R^{m \times n}$, сложение матриц и умножение их на действительное число определены.

3) $R[x]$ – множество многочленов с действительными коэффициентами, сложение многочленов и умножение их на действительное число известны.

4) $R[x]_{(\leq n)}$ – множество многочленов с действительными коэффициентами степени, не превосходящей n.

5) Множество направленных отрезков плоскости или пространства с общим началом в начале координат. Сложение таких отрезков осуществляется по правилу параллелограмма, умножение по известному правилу.

6) $R(a, b)$ – множество функций определенных, дифференцируемых на отрезке $[a, b]$.

Если числа в определении 8.1 k, l … брать из поля действительных (вещественных) чисел $\mathbb{R}$, т. е. $P = \mathbb{R}$, то пространство называется *вещественным векторным* (*линейным*) *пространством*; если же из поля комплексных чисел, то приходим к понятию *комплексного линейного пространства*.

Простейшие свойства векторных пространств

1) o – нулевой вектор (элемент), определен единственным образом в произвольном векторном пространстве над полем.

2) Для любого вектора $a \in V$ существует единственный противоположный элемент $(-a) \in V$.

3) $a,\ b \in V$ уравнение $a + x = b$ разрешимо единственным образом $x = b + (–a)$ и обозначается как $x = b – a$, и называется *разностью*.

4) операция сложения сократима: если $a + b = a + c$, то $b = c$ для любых $a, b, c \in V$.

5) если $a + b = a$, то $b = o$.

6) если $a + b = o$, то $a = –b$ и $b = –a$.

7) $–(–a) = a$.

8) $0{\cdot}a = o$, где 0 элемент поля P, а o – нулевой вектор пространства V.

9) $k{\cdot}o = o$, здесь $k \in P, o \in V$.

10) если $k{\cdot}a = o$, то $k = 0$ или $a =o$.

11) $(–1){\cdot}a = –a$.

12) $(k – l){\cdot}a = k{\cdot}a – l{\cdot}a$, где $k, l \in P, a \in V$.

13) $k{\cdot}(a – b) = k{\cdot}a – k{\cdot}b$, где $k \in P, a, b \in V$.

14) $(–k){\cdot}a = –k{\cdot}a$.

Линейная зависимость и независимость системы векторов

Для произвольного векторного пространства понятия линейной комбинации, линейной оболочки системы векторов, линейной зависимости и независимости системы векторов определяется точно так, как и для n-мерного арифметического векторного пространства. Выполняются все свойства линейной зависимости (кроме свойства, связанного со ступенчатой системой векторов).

8.2. Подпространства. Линейные многообразия

Пусть V – векторное пространство, $L \subset V$ (L подмножество V).

Определение 8.2. Подмножество L векторного пространства V называется *подпространством* пространства V, если

1. $\forall\ a, b \in L : a + b \in L;$

2. $\forall\ a \in L, k \in P: k{\cdot}a \in L.$

Обозначение $L \prec V$. Принято говорить, что подмножество L замкнуто относительно сложения векторов и относительно умножения их на элемент из поля P.

Пример 8.2.

1) В каждом векторном пространстве есть два подпространства, называемых *несобственными*: $L = \{0\}$ – *нулевое подпространство*, $L = V$ – подпространство, совпадающее со всем пространством .

Приведем примеры *собственных* подпространств.

2) Пусть $V = R^4$, $L = \{((\alpha_1, \alpha_2, \alpha_3, 0), \alpha_i \in \mathbb{R}\}$ – подпространство, так как для произвольных векторов $a = (\alpha_1, \alpha_2, \alpha_3, 0) \in L$ и $b = (\beta_1, \beta_2, \beta_3, 0) \in L$ и $k \in P$:

- $a + b = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \alpha_3 + \beta_3, 0 + 0) \in L$;
- $k \cdot a = (k \cdot \alpha_1, k \cdot \alpha_2, k \cdot \alpha_3, 0) \in L$.

3) В пространстве квадратных матриц подпространство образует подмножество диагональных матриц.

4) В пространстве направленных отрезков подпространством является множество отрезков, лежащих на прямой, проходящей через начало координат.

Теорема 8.1. Линейная оболочка $L(a_1, a_2, \ldots, a_m)$ системы векторов $a_1, a_2, \ldots, a_m$ образует подпространство пространства V.

В этом случае принято говорить, что $L(a_1, a_2, \ldots, a_m)$ подпространство, *натянутое* на векторы $a_1, a_2, \ldots, a_m$, или что $L(a_1, a_2, \ldots, a_m)$ – подпространство, *порожденное* векторами $a_1, a_2, \ldots, a_m$. Система векторов $a_1, a_2, \ldots, a_m$ называется *системой образующих* подпространства $L(a_1, a_2, \ldots, a_m)$.

Пересечение и сумма подпространств

Пусть V – векторное пространство над полем P, L_1 и L_2 – его подпространства.

Определение 8.3. *Пересечением* подпространств называется множество $L_1 \cap L_2 = \{x \mid x \in L_1$ и $x \in L_2\}$.

Теорема 8.2. Пересечение подпространств является подпространством.

Определение 8.4. *Суммой* подпространств L_1 и L_2 называется множество $L_1 + L_2 = \{x = x_1 + x_2 \mid x_1 \in L_1$ и $x_2 \in L_2\}$.

Теорема 8.3. Сумма подпространств является подпространством.

Определение 8.5. Сумма подпространств L_1 и L_2 называется *прямой*, если каждый вектор из суммы $L_1 + L_2$ может быть единственным образом представлен в виде суммы векторов из L_1 и L_2.

Прямая сумма подпространств обозначается символом $L_1 \oplus L_2$.

Теорема 8.4. Сумма подпространств L_1 и L_2 является прямой тогда и только тогда, когда их пересечение состоит только из нулевого вектора, т. е. $L_1 \cap L_2 = \{o\}$.

Линейные многообразия

Пусть V – векторное пространство, L – подпространство, a – произвольный вектор из пространства V.

Определение 8.6. *Линейным многообразием* пространства V с направлением L, порожденным вектором a, называется множество $a + L = \{a + l, l \in L\}$. Вектор a называется *вектором сдвига*.

Пример 8.3. В пространстве $R^{2\times 2}$ выберем подпространство $L = \{\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}, \ a, \ b \in \mathbb{R}\}$ и вектор сдвига $A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$, тогда соответствующее линейное многообразие – это множество $A + L = \{\begin{pmatrix} c & d \\ 1 & 2 \end{pmatrix}, c, d \in \mathbb{R}\}$.

8.3. Базис и размерность векторного пространства

8.3.1. Конечномерные векторные пространства

Определение 8.7. Векторное пространство V называется *n-мерным,* если в нем существует линейно независимая система векторов, состоящая из n векторов, и при этом любая система, состоящая более чем из n векторов, линейно зависима.

В этом случае говорят, что размерность V равна n ($dimV = n$).

Определение 8.8. Векторное пространство, имеющее размерность, называется *конечномерным*.

Определение 8.9. Если в векторном пространстве V можно указать линейно независимую систему векторов с каким угодно количеством векторов, то пространство V называется *бесконечномерным*.

Пример 8.4. 1) Пространство R^4 четырехмерно, так как в нем есть 4 линейно независимых вектора e_1, e_2, e_3, e_4 и любая другая система векторов, в которой более 4 векторов, линейно зависима.

2) Пространство $R[x]$ бесконечномерно, поскольку система векторов 1, x, x^2, …, x^n линейно независима при любом n.

Базис конечномерного векторного пространства

V – конечномерное векторное пространство над полем P, S – система векторов (конечная или бесконечная).

Определение 8.10. *Базисом* системы S называется ее подсистема S' такая, что

• подсистема S' линейно независима;

• любой вектор системы S линейно выражается через векторы системы S'.

Как и для арифметических n-мерных векторных пространств верна следующая теорема:

Теорема 8.5. Любые два базиса одной и той же системы векторов состоят из одинакового количества векторов.

Теорема 8.6. В конечномерном векторном пространстве любая система, содержащая хотя бы один ненулевой вектор, имеет базис.

Определение 8.11. *Базисом* конечномерного векторного пространства V называется система векторов этого пространства, такая что

• эта система линейно независима;

• каждый вектор из V линейно выражается через векторы базиса.

Теорема 8.7. В n-мерном векторном пространстве есть базис. При этом базисы – это в точности все системы, состоящие из n линейно независимых векторов, то есть

- базис – это система, состоящая n из линейно независимых векторов, и

- любая система, состоящая из n линейно независимых векторов, является базисом.

Определение 8.12. *Размерностью* пространства называется количество векторов в любом его базисе.

Пример 8.5. Приведем примеры базисов конечномерных векторных пространств.

1) Пространство $V = R^n$. Система единичных векторов e_1, e_2, …, e_n образует базис этого пространства, что следует из свойств этой системы векторов. Размерность R^n равна n.

2) Пространство $V = R^3$. Система векторов e_1, e_2, e_3 – базис R^3. Любой другой базис этого пространства состоит из трех линейно независимых векторов, это, например, векторы a_1, a_2, a_3 : $a_1 = (2, 3, -1)$, $a_2 = (0, 4, 7)$, $a_3 = (0, 0, -4)$. Эти векторы образуют лестничную систему, поэтому они линейно независимы. Размерность R^3 равна 3.

3) Пространство $V = R^{2\times2}$. В качестве базиса этого пространства можно выбрать векторы $E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Можно показать, что они линейно независимы. Как выражаются элементы пространства через векторы базиса, видно из следующего примера: если вектор $A = \begin{pmatrix} 2 & -4 \\ 3 & 5 \end{pmatrix}$, то $A = 2E_1 + (-4)E_2 + 3E_3 + 5E_4$. Размерность $R^{2\times 2}$ равна 4.

4) Пространство $V = R[x]_{(\le 3)}$. Векторы $f_1 = 1$, $f_2 = x$, $f_3 = x^2$, $f_4 = x^3$ образуют базис пространства $R[x]_{(\le 3)}$. Линейная независимость этих векторов легко доказывается, произвольный вектор $f = a + bx + cx^2 + d\ x^3$ выражается через векторы следующим способом $\mathrm{f} = a\cdot f_1 + b\cdot f_2 + c\cdot f_3 + d\cdot f_4$. Размерность этого пространства равна 4.

5) Пространство направленных отрезков. На плоскости базис состоит из любых двух неколлинеарных векторов, в пространстве – из любых трех некомпланарных векторов.

8.3.2. Базисы и размерности подпространств

1. Пусть подпространство $L = L(a_1, a_2, \ldots, a_m)$, то есть L – линейная оболочка системы $a_1, a_2, \ldots, a_m$; векторы $a_1, a_2, \ldots, a_m$ – система образующих этого подпространства. Тогда базисом L является базис системы векторов $a_1, a_2, \ldots, a_m$, то есть базис системы образующих. Размерность L равна рангу системы образующих.

2. Пусть подпространство L является суммой подпространств L_1 и L_2. Систему образующих суммы подпространств можно получить объединением систем образующих подпространств, после чего находится базис суммы. Размерность суммы находится по следующей формуле:

$$dim(L_1 + L_2) = dimL_1 + dimL_2 - dim(L_1 \cap L_2).$$

3. Пусть сумма подпространств L_1 и L_2 прямая, то есть $L = L_1 \oplus L_2$. При этом $L_1 \cap L_2 = \{o\}$ и $dim(L_1 \cap L_2) = 0$. Базис прямой суммы равен объединению базисов слагаемых. Размерность прямой суммы равна сумме размерностей слагаемых.

4. Приведем важный пример подпространства и линейного многообразия.

Рассмотрим однородную систему m линейных уравнений с n неизвестными. Множество решений M_0 этой системы является подмножеством множества R^n и замкнуто относительно сложения векторов и умножения их на действительное число. Это означает, что это множество M_0 – подпространство пространства R^n. Базисом подпространства является фундаментальный набор решений однородной системы, размерность подпространства равна количеству векторов в фундаментальном наборе решений системы.

Множество M решений общей системы m линейных уравнений с n неизвестными так же является подмножеством множества R^n и равно сумме множества M_0 и вектора a, где a – некоторое частное решение исходной системы, а множество M_0 – множество решений однородной системы линейных уравнений, сопутствующей данной системе (она отличается от исходной только свободными членами),

$$M = a + M_0 = \{a = m,\, m \in M_0\}.$$

Это означает, что множество M является линейным многообразием пространства R^n с вектором сдвига a и направлением M_0.

Пример 8.6. Найти базис и размерность подпространства, заданного однородной системой линейных уравнений: $\begin{cases} x_1 + 2x_2 - 3x_3 + 4x_4 + 5x_5 = 0, \\ 2x_1 + 3x_2 + 6x_3 + 2x_5 = 0. \end{cases}$

Решение. Найдем общее решение этой системы и ее фундаментальный набор решений: $\begin{cases} x_1 = 21x_3 + 12x_4 + 11x_5, \\ x_2 = 12x_3 - 8x_4 - 8x_5, \end{cases}$ $c_1 = (–21, 12, 1, 0, 0)$, $c_2 = (12, –8, 0, 1, 0)$, $c_3 = (11, –8, 0, 0, 1)$.

Базис подпространства образуют векторы c_1, c_2, c_3, его размерность равна трем.

8.3.3. Координаты вектора относительно данного базиса

Рассмотрим конечномерное векторное пространство V размерности n, векторы e_1, e_2, …, e_n образуют его базис. Пусть a – произвольный вектор пространства V, тогда вектор линейно выражается через векторы базиса, $a = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n$.

Теорема 8.8. Разложение вектора a по векторам базиса производится единственным образом.

Доказательство. Предположим, что вектор a можно разложить по векторам базиса двумя способами:

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n.$$

$$a = \alpha'_1 e_1 + \alpha'_2 e_2 + \ldots + \alpha'_n e_n.$$

После вычитания из одного равенства другого, получим

$$(\alpha_1 - \alpha'_1)\, e_1 + (\alpha_2 - \alpha'_2) e_2 + \ldots + (\alpha_n - \alpha'_n) e_n = 0,$$

из чего в силу линейной независимости базисных векторов e_1, e_2, …, e_n следует, что $\alpha_1 - \alpha'_1 = 0$, $\alpha_2 - \alpha'_2 = 0$, …, $\alpha_n - \alpha'_n = 0$, а затем что $\alpha_1 = \alpha'_1$, $\alpha_2 = \alpha'_2$, …, $\alpha_n = \alpha'_n$. Таким образом, коэффициенты разложения определяются однозначно. Теорема доказана.

Определение 8.13. Координатами вектора a относительно базиса e_1, e_2, …, e_n называются *коэффициенты разложения* вектора a по базисным векторам.

Координаты вектора принято записывать или в виде строки координат (координатной строки) – $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, или в виде координатного столбца:

$$[a] = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Пример 8.7. **1)** В пространстве $R^{2\times 2}$ вектор $A = \begin{pmatrix} 2 & -1 \\ 4 & 7 \end{pmatrix}$ раскладывается по векторам базиса E_1, E_2, E_3, E_4 следующим образом: $A = 2E_1 - E_2 + 4E_3 + 7E_4$, следовательно, координатная строка этого вектора равна $(2, -1, 4, 7)$.

2) В пространстве выбран базис $a_1 = (1,\ 3,\ -1)$, $a_2 = (-2,\ 1,\ 1)$, $a_3 = (2,\ -2,\ -1)$. Найти координаты вектора $a = (3,\ 0,\ -2)$ относительно базиса a_1, a_2, a_3. Векторное равенство $a = x_1a_1 + x_2a_2 + x_3a_3$ перепишем в виде системы линейных уравнений $\begin{cases} x_1 - 2x_2 + 2x_3 = 3, \\ 3x_1 + x_2 - 2x_3 = 0, \\ -x_1 + x_2 - x_3 = -2. \end{cases}$ Решая эту систему, получим $x_1 = 1$, $x_2 = 1$, $x_3 = 2$, следовательно, координатная строка вектора a равна (1, 1, 2).

Каждому вектору a из произвольного векторного пространства V, в котором задан базис e_1, e_2, …, e_n, сопоставляется строка (или столбец) координат $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, причем единственным образом. Если V пространство размерности n, то строка координат принадлежит пространству R^n, то есть возникает отображение: $V \to R^n$. Обратно, по строке координат $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, (по вектору из R^n) единственным образом можно построить вектор $a = \alpha_1e_1 + \alpha_2e_2 + \ldots + \alpha_ne_n$. Для этого отображения верна следующая теорема.

Теорема 8.9. Если векторы a_1, a_2, …, a_m из произвольного пространства V образуют линейно независимую систему векторов, то их строки (или столбцы) координат тоже линейно независимы.

8.3.4. Координаты вектора в различных базисах

Пусть V – n-мерное векторное пространство, в котором заданы два базиса: e_1, e_2, …, e_n – старый базис, e'_1, e'_2, …, e'_n – новый базис. У произвольного вектора a есть координаты в каждом из них:

$$a = \alpha_1e_1 + \alpha_2e_2 + \ldots + \alpha_ne_n;$$

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

Для того чтобы установить связь между столбцами координат вектора a в старом и новом базисах, надо разложить векторы нового базиса по векторам старого базиса:

$$e'_1 = \alpha_{11}e_1 + \alpha_{21}e_2 + \ldots + \alpha_{n1}e_n,$$

$e'_2 = \alpha_{12}e_1 + \alpha_{22}e_2 + \ldots + \alpha_{n2}e_n$,

……………………………..

$e'_n = \alpha_{1n}e_1 + \alpha_{2n}e_2 + \ldots + \alpha_{nn}e_n$.

Определение 8.14. *Матрицей перехода от старого базиса к новому базису* называется матрица, составленная из координат векторов нового базиса относительно старого базиса, записанных в столбцы, т. е.

$$T = \begin{pmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}.$$

Столбцы матрицы T – это координаты базисных, а значит, линейно независимых, векторов, следовательно, эти столбцы линейно независимы. Матрица с линейно независимыми столбцами является невырожденной, ее определитель не равен нулю и для матрицы T существует обратная матрица T^{-1}.

Обозначим столбцы координат вектора a в старом и новом базисах, соответственно, как $[a]$ и $[a]'$. С помощью матрицы перехода устанавливается связь между $[a]$ и $[a]'$.

Теорема 8.10. Столбец координат вектора a в старом базисе равен произведению матрицы перехода на столбец координат вектора a в новом базисе, то есть $[a] = T\,[a]'$.

Следствие. Столбец координат вектора a в новом базисе равен произведению матрицы, обратной матрице перехода, на столбец координат вектора a в старом базисе, то есть $[a]' = T^{-1}[a]$.

Пример 8.8. Составить матрицу перехода от базиса e_1, e_2, к базису e'_1, e'_2, где $e'_1 = 3e_1 + e_2$, $e'_2 = 5e_1 + 2e_2$, и найти координаты вектора $a = 2e'_1 - 4e'_2$ в старом базисе.

Решение. Координатами новых базисных векторов относительно старого базиса являются строки (3, 1) и (5, 2), тогда матрица T примет вид $\begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$. Так как $[a]' = \begin{pmatrix} 2 \\ -4 \end{pmatrix}$, то $[a] = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -4 \end{pmatrix} = \begin{pmatrix} 14 \\ -8 \end{pmatrix}$.

Пример 8.9. Даны два базиса e_1, e_2 – старый базис, e'_1, e'_2 – новый базис, причем $e'_1 = 3e_1 + e_2$, $e'_2 = 5e_1 + 2e_2$. Найти координаты вектора $a = 2e_1 - e_2$ в новом базисе.

Решение. 1 способ. По условию даны координаты вектора a в старом базисе: $[a] = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$. Найдем матрицу перехода от старого базиса e_1, e_2 к новому базису e'_1, e'_2. Получим матрицу $T = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$ для нее найдем обратную матрицу $T^{-1} = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}$. Тогда согласно следствию из теоремы 8.10 имеем $[a]' = T^{-1}[a] = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 9 \\ -5 \end{pmatrix}$.

2 способ. Так как e'_1, e'_2 базис, то вектор a раскладывается по базисным векторам следующим образом $a = k_1e'_1 - k_2e'_2$. Найдем числа k_1 и k_2 – это и будут координаты вектора a в новом базисе.

$$a = k_1e'_1 - k_2e'_2 = k_1(3e_1 + e_2) - k_2(5e_1 + 2e_2) =$$

$$= e_1(3k_1 + 5k_2) + e_2(k_1 + 2k_2) = 2e_1 - e_2.$$

Так как координаты одного и того же вектора в данном базисе определяется однозначно, то имеем систему: $\begin{cases} 3k_1 + 5k_2 = 2, \\ k_1 + 2k_2 = -1. \end{cases}$ Решая данную систему, получим $k_1 = 9$ и $k_2 = -5$, т. о. $[a]' = \begin{pmatrix} 9 \\ -5 \end{pmatrix}$.

8.4 Евклидовы векторные пространства

Дано векторное пространство V над полем действительных чисел. Это пространство может быть как конечномерным векторным пространством размерности n, так и бесконечномерным.

Определение 8.15. Векторное пространство V называется *евклидовым векторным пространством*, если задано правило, по которому любой паре векторов ставится в соответствие единственное действительное число, обозначаемое (x, y) и называемое *скалярным произведением векторов* x и y (другими словами, задано отображение $V \times V \to \mathbb{R}$). При этом указанное правило подчинено 4 *аксиомам*:

1) $(x, y) = (y, x)$;
2) $(x + y, z) = (x, z) + (y, z)$;
3) $(\alpha x, y) = \alpha(x, y)$, где $\alpha \in \mathbb{R}$;
4) $(x, x) \geq 0$, причем $(x, x) = 0 \Leftrightarrow x = o$; при этом (x, x) называют *скалярным квадратом* элемента x.

Пример 8.10. Приведем примеры евклидовых пространств.

1. $V = R^n$ – арифметическое n-мерное векторное пространство. Если векторам $x = (x_1, x_2, \ldots, x_n)$ и $y = (y_1, y_2, \ldots, y_n)$ поставлено в соответствие число $(x, y) = x_1y_1 + x_2y_2 + \ldots + x_ny_n$, то аксиомы выполняются. Проверим последнюю из них. Найдем (x, x): $(x, x) = x_1^2 + x_2^2 + \ldots + x_n^2$. Сумма квадратов во множестве действительных чисел неотрицательна. Полученное евклидово векторное пространство называется *стандартным евклидовым векторным пространством*.

2. V – пространство направленных отрезков с общим началом в начале координат. Скалярным произведением двух векторов назовем число, равное произведению длин этих векторов на косинус угла между ними. Все требуемые свойства выполняются, что известно еще из школы.

3. $V = R(a, b)$ – множество функций, заданных и непрерывных на промежутке $[a, b]$. Зададим скалярное умножение векторов из $R(a, b)$

следующим способом: $(f,g)=\int\limits_a^b f(x)g(x)dx$. Из свойств определенного интеграла получаются все свойства, требуемые в определении скалярного произведения.

Скалярное произведение в координатах

В евклидовом векторном пространстве V размерности n задан базис e_1, e_2, …, e_n. Векторы x и y разложены по векторам базиса: $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$, $y = y_1e_1 + y_2e_2 + \ldots + y_ne_n$. Найдем скалярное произведение этих векторов

$(x, y) = (x_1e_1 + x_2e_2 + \ldots + x_ne_n)(y_1e_1 + y_2e_2 + \ldots + y_ne_n) =$

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

$= \sum_{i,j=1}^{n} x_i y_j (e_i, e_j)$.

Очевидно, что для задания скалярного произведения, необходимо задать матрицу $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$, где $a_{ij} = (e_i, e_j)$. Если $(x_1, x_2, \ldots, x_n)$ – строка координат вектора x, а $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ – столбец координат вектора y, то

$$(x, y) = (x_1, x_2, \ldots, x_n)\cdot\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\cdot\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Матрица A не может быть произвольной, иначе не станут выполняться аксиомы евклидова векторного пространства. Эта матрица должна быть симметрической и должна задавать положительно определенную квадратичную форму. Простейшим примером такой матрицы является матрица E. При этом $(x, y) = x_1y_1 + x_2y_2 + \ldots + x_ny_n$, то есть скалярное произведение равно сумме произведений соответствующих координат.

Метрические понятия

В евклидовых векторных пространствах от введенного скалярного произведения можно перейти к понятиям нормы вектора и угла между векторами.

Определение 8.16. *Нормой* (*длиной*, *модулем*) вектора a называется число, равное корню из скалярного квадрата вектора a: $||a|| = \sqrt{(a,a)}$.

Поскольку $(a, a) \geq 0$, то норма вектора определена.

Определение 8.17. Вектор a называется *нормированным*, если его норма равна единице, т. е. $||a|| = 1$.

Свойства нормы

1) $||a|| = 0 \Leftrightarrow a = o$.

2) $||\lambda a|| = |\lambda|\cdot||a||$, т. к. $||\lambda a|| = \sqrt{(\lambda a, \lambda a)} = \sqrt{\lambda^2(a,a)} = |\lambda|\cdot||a||$.

3) Неравенство Коши – Буняковского: $|(a, b)| \leq ||a||\cdot||b||$.

Доказательство. Для любого числа λ и любых векторов a, $b \neq 0$ выполняется условие $(a - \lambda b,\ a - \lambda b) \geq 0 \Rightarrow (a,\ a) - 2\lambda(a,\ b) + \lambda^2(b,\ b) \geq 0$. Квадратный трехчлен относительно λ неотрицателен при любом λ, если его дискриминант неположителен: $D = 4(a, b)^2 - 4(a, a)(b, b) \leq 0 \Rightarrow (a, b)^2 \leq ||a||^2\cdot||b||^2 \Rightarrow |(a, b)| \leq ||a||\cdot||b||$.

4) Неравенство треугольника: $||a + b|| \leq ||a|| + ||b||$.

Пример 8.11. Будем считать (если нет специальных оговорок), что скалярное произведение равно сумме произведений соответствующих координат.

1. Найти норму вектора $a = (\frac{2}{3},\ \frac{2}{3},\ -\frac{1}{3})$. Найдем $(a,\ a)$: $(a, a) = \frac{2}{3}\cdot\frac{2}{3}+\frac{2}{3}\cdot\frac{2}{3}+\left(-\frac{1}{3}\right)\cdot\left(-\frac{1}{3}\right) = 1 \Rightarrow ||a|| = 1$.

2. Нормировать вектор $b = (-4,\ 2,\ 2,\ -1)$. Найдем норму вектора $\|b\|$: $(b, b) = (-4)^2 + 2^2 + 2^2 + (-1)^2 = 25 \Rightarrow \|b\| = \sqrt{(b,b)} = 5$. Вектор $e = \frac{1}{\|b\|} \cdot b$ нормирован, это можно проверить, используя свойство 2, тогда $e = \frac{1}{5} \cdot b = (-\frac{4}{5}, \frac{2}{5}, \frac{2}{5}, -\frac{1}{5})$.

Определение 8.18. *Углом* между ненулевыми векторами a и b называется угол, меняющийся в пределах от 0 до π и определенный условием $\cos\varphi = \frac{(a,b)}{\|a\| \cdot \|b\|}$.

В силу неравенства Коши – Буняковского $\cos\varphi$ принимает значения от (–1) до 1 и, следовательно, угол между ненулевыми векторами определен.

Определение 8.19. Векторы a и b называются *ортогональными*, если их скалярное произведение равно нулю. Обозначение: $a \perp b$.

Это определение согласуется с определением угла между векторами.

Определение 8.20. Нулевой вектор считается ортогональным любому вектору.

Ортонормированный базис евклидова векторного пространства

Определение 8.21. Базис евклидова векторного пространства называется *ортогональным*, если векторы базиса попарно ортогональны, то есть если $a_1, a_2, \ldots, a_n$ – ортогональный базис пространства, то $(a_i, a_j) = 0$ при $i \neq j$, $i, j = 1, 2, \ldots, n$.

Определение 8.22. Ортогональный базис называется *ортонормированным*, если каждый вектор базиса нормирован, то есть если $e_1, e_2, \ldots, e_n$ – ортонормированный базис, то $(e_i, e_j) = 0$ при $i \neq j$ и $(e_i, e_i) = 1$, $i, j = 1, 2, \ldots, n$.

Докажем возможность существования ортонормированного базиса.

Теорема 8.11. Ортогональная система ненулевых векторов линейно независима.

Доказательство. Пусть система векторов $a_1, a_2, \ldots, a_k$ – ненулевая и ортогональная, то есть $a_i \neq o$, $(a_i, a_j) = 0$ при $i \neq j$, $i, j = 1, 2, \ldots, k$. Покажем, что равенство $\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = o$ возможно лишь тогда, когда $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$. Умножим это равенство скалярно на a_1:

$$(\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k, a_1) = (o, a_1) \Rightarrow$$

$$\alpha_1(a_1, a_1) + \alpha_2(a_2, a_1) + \ldots + \alpha_k(a_k, a_1) = 0.$$

В силу ортогональности системы (т. е. $(a_2, a_1) = 0, \ldots, (a_k, a_1) = 0$) получим $\alpha_1(a_1, a_1) = 0$ и, так как $a_1 \neq o$ (т. е. $(a_1, a_1) \neq 0$), то $\alpha_1 = 0$. Аналогично доказывается, что $\alpha_2 = 0, \ldots, \alpha_k = 0$ (умножая исходное равенство по очереди на $\alpha_2, \alpha_3, \ldots, \alpha_k$), следовательно, система векторов $a_1, a_2, \ldots, a_k$ линейно независима. Теорема доказана.

Процесс ортогонализации

Теорема 8.12. Во всяком n-мерном евклидовом пространстве существует ортонормированный базис.

Доказательство. Пусть $a_1, a_2, \ldots, a_n$ – произвольный базис евклидова пространства E. Доказательство заключатся в описании алгоритма построения ортогонального базиса по данному базису. Этот алгоритм называется процессом ортогонализации. Пусть $b_1 = a_1$, $b_1 \neq 0$ (т. к. $a_1 \neq 0$). Положим $b_2 = a_2 + \alpha_1 b_1$. Подберем коэффициент α_1 так, чтобы $b_2 \neq 0$ стал ортогонален b_1;

$(b_1, b_2) = 0 \quad \Rightarrow \quad (b_1, a_2 + \alpha_1 b_1) = 0 \quad \Rightarrow \quad (a_2 + \alpha_1 b_1, b_1) = 0 \quad \Rightarrow$ $(a_2, b_1) + \alpha_1(b_1, b_1) = 0$, т. к. $b_1 \neq 0$, то $(b_1, b_1) \neq 0 \Rightarrow \alpha_1 = -\dfrac{(b_1, a_2)}{(b_1, b_1)}$. Вектор b_2 не равен нулю, поскольку он является ненулевой линейной комбинацией линейно независимых векторов a_1 и a_2.

Положим, далее $b_3 = a_3 + \beta_1 b_1 + \beta_2 b_2$. Подберем β_1 и β_2 так, чтобы $b_3 \neq 0$ оказался ортогонален b_1 и b_2, для чего должны выполняться условия

$(b_1, b_3) = 0$, $(b_2, b_3) = 0$. Выполняя преобразования, получим, что $\beta_1 = -\dfrac{(a_3, b_1)}{(b_1, b_1)}$, $\beta_2 = -\dfrac{(a_3, b_2)}{(b_2, b_2)}$. Вектор b_3 не равен нулю, поскольку он является ненулевой линейной комбинацией векторов a_1, a_2, a_3.

Продолжая этот процесс, получим систему векторов $b_1, b_2, \ldots, b_n$, и так как эти векторы ненулевые и попарно ортогональны, то по теореме 8.11 они линейно независимы, а значит образуют ортогональный базис.

Нормируя ортогональный базис $b_1, b_2, \ldots, b_n$, получим ортонормированный базис n-мерного евклидова пространства:

$$e_1 = \frac{1}{\| b_1 \|} \cdot b_1,\ e_2 = \frac{1}{\| b_2 \|} \cdot b_2,\ \ldots,\ e_n = \frac{1}{\| b_n \|} \cdot b_n.$$

Пример 8.12. Применить процесс ортогонализации к векторам $a_1 = (2, -2, -2, 2)$, $a_2 = (3, -1, -1, 3)$, $a_3 = (2, -2, 0, 4)$.

Решение. Это задание можно сформулировать так: по данному базису подпространства построить ортогональный базис.

$b_1 = a_1$, $b_1 = (2, -2, -2, 2)$;

$b_2 = a_2 + \alpha_1 b_1$,

$\alpha_1 = -\dfrac{(b_1, a_2)}{(b_1, b_1)} = -\dfrac{3 \cdot 2 - 1 \cdot (-2) + 3 \cdot 2}{2 \cdot 2 + (-2) \cdot (-2) + (-2) \cdot (-2) + 2 \cdot 2} = -\dfrac{16}{16} = -1$. Тогда

$b_2 = a_2 - b_1 = (1, 1, 1, 1)$.

$b_3 = a_3 + \beta_1 b_1 + \beta_2 b_2$, $\beta_1 = -\dfrac{(a_3, b_1)}{(b_1, b_1)} = -\dfrac{16}{16} = -1$, $\beta_2 = -\dfrac{(a_3, b_2)}{(b_2, b_2)} = -\dfrac{4}{4} = -1$.

Тогда $b_3 = a_3 - b_1 - b_2 = (-1, -1, 1, 1)$.

Скалярное произведение в ортонормированном базисе

Дан ортонормированный базис $e_1, e_2, \ldots, e_n$ евклидова пространства V. Поскольку $(e_i, e_j) = 0$ при $i \neq j$ и $(e_i, e_i) = 1$, то

$$(x, y) = \sum_{i,j=1}^{n} x_i y_j\,(e_i, e_j) = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Вывод: скалярное произведение векторов в ортонормированном базисе равно сумме произведений соответствующих координат.

Ортогональное дополнение подпространства

V – евклидово векторное пространство, L – его подпространство.

Определение 8.23. Говорят, что вектор a *ортогонален подпространству* L , если вектор a ортогонален любому вектору из подпространства L, т. е.

$$a \perp L \Leftrightarrow a \perp x, \forall\, x \in L.$$

Определение 8.24. *Ортогональным дополнением* подпространства L называется множество L^* всех векторов, ортогональных подпространству L, то есть $L^* = \{x \mid x \perp L\}$.

Теорема 8.13. Ортогональное дополнение подпространства является подпространством.

Теорема 8.14. Прямая сумма подпространства L и его ортогонального дополнения L^* равна пространству V, т. е. $L \oplus L^* = V$.

Пример 8.13. Найти ортогональное дополнение подпространства L, натянутого на векторы $a_1 = (1, 1, 1, 1)$, $a_2 = (–1, 1, –1, 1)$, $a_3 = (2, 0, 2, 0)$.

Решение. Для того чтобы вектор x был ортогонален подпространству, необходимо и достаточно, чтобы он был ортогонален векторам системы образующих этого подпространства. Пусть $x = (x_1, x_2, x_3, x_4)$, запишем условие ортогональности этого вектора векторам a_1, a_2, a_3: $(x, a_1) = 0$, $(x, a_2) = 0$, $(x, a_3) = 0$. В координатной форме эти условия представляют собою однородную систему линейных уравнений: $\begin{cases} x_1 + x_2 + x_3 + x_4 = 0, \\ -x_1 + x_2 - x_3 + x_4 = 0, \\ 2x_1 + 2x_3 = 0. \end{cases}$

Множество решений этой системы представляет собою подпространство L^*, ортогональное подпространству L.

Решая систему, получим фундаментальный набор решений: $c_1 = (–1, 0, 1, 0)$, $c_2 = (0, –1, 0, 1)$. Эти векторы образуют базис множества решений системы, то есть базис L^*, т. о. $L^* = L(c_1,c_2)$, $dim\, L^* = 2$.

9. Линейные операторы

9.1. Основные понятия и способы задания линейных операторов

Дано V – векторное пространство над полем P, $\dim V = n$.

Определение 9.1. Говорят, что задано отображение φ множества V в себя, если каждому элементу x из V поставлен в соответствие единственный элемент y, тоже принадлежащий V. При этом приняты следующие обозначения и термины: φ: $V \to V$, φ: $x \to y$, $\varphi(x) = y$; элемент x – прообраз элемента y, y – образ x.

Определение 9.2. *Линейным оператором* пространства V называется отображение φ: $V \to V$ такое, что $\forall\ a, b \in V$, $\forall\ \lambda \in P$

1) $\varphi(a + b) = \varphi(a) + \varphi(b)$,
2) $\varphi(\lambda a) = \lambda\varphi(a)$.

Вместо «линейный оператор» говорят также «линейное отображение» или «отображение, сохраняющее операции сложения и умножения на элемент поля»

Пример 9.1. 1) В произвольном векторном пространстве V зададим отображение следующей формулой: $\varphi(x) = kx$. Это отображение является линейным оператором и называется оператором *гомотетии*.

Если $k = 1$, то отображение примет вид: $\varphi(x) = 1$. Его называют *тождественным оператором* и обозначают буквой ε: $\varepsilon(x) = x$.

Если $k = 0$, то получают *нулевой оператор* θ: $\theta(x) = o$.

2) В пространстве $V = R^{2\times2}$ оператор *транспонирования* задают формулой $\varphi(A) = A^t$, где A^t – матрица, транспонированная для матрицы A.

3) В пространстве $V = R[x]_{(\leq n)}$ (многочленов степени, не превосходящей n) можно задать отображение φ, ставящее в соответствие произвольному многочлену его производную, т. е. φ: $f(x) \to f'(x)$, $\varphi(f) = f'$. Покажем, что это отображение линейно:

$$\varphi(f + g) = (f + g)' = f' + g' = \varphi(f) + \varphi(g),$$

$$\varphi(\lambda f) = (\lambda f)' = \lambda(f)' = \lambda\varphi(f).$$

Способы задания линейных операторов

Приведенные примеры не показывают, сколько в каждом векторном пространстве существует линейных операторов, каким способом их можно задавать. Ответом на эти вопросы является следующая теорема.

Теорема 9.1. Для того чтобы задать линейный оператор, достаточно задать образы базисных векторов.

Доказательство. Другими словами, если e_1, e_2, …, e_n – некоторый базис векторного пространства V над полем P, а b_1, b_2, …, b_n – произвольные векторы этого же пространства, то существует единственный линейный оператор, такой, что $\varphi(e_1) = b_1$, $\varphi(e_2) = b_2$, …, $\varphi(e_n) = b_n$.

Покажем, как найти образ произвольного вектора x. Разложим вектор x по базисным векторам и найдем его образ, используя свойства линейного отображения:

$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$, где $x_i \in P$ $\Rightarrow$ $\varphi(x) = \varphi(x_1e_1 + x_2e_2 + \ldots + x_ne_n) =$
$= \varphi(x_1e_1) + \varphi(x_2e_2) + \ldots + \varphi(x_ne_n) = x_1\varphi(e_1) + x_2\varphi(e_2) + \ldots + x_n\varphi(e_n) =$
$= x_1b_1 + x_2b_2 + \ldots + x_nb_n$. Теорема доказана.

Пусть в векторном пространстве V задан линейный оператор φ, т. е. указаны образы базисных векторов $\varphi(e_1)$, $\varphi(e_2)$, …, $\varphi(e_n)$. Разложим эти векторы по векторам базиса:

$$\varphi(e_1) = \alpha_{11}e_1 + \alpha_{21}e_2 + \ldots + \alpha_{n1}e_n,$$

$$\varphi(e_2) = \alpha_{12}e_1 + \alpha_{22}e_2 + \ldots + \alpha_{n2}e_n,$$

……………………………………..…..

$$\varphi(e_n) = \alpha_{1n}e_1 + \alpha_{2n}e_2 + \ldots + \alpha_{nn}e_n.$$

Используем координаты образов базисных векторов.

Определение 9.3. *Матрицей линейного оператора* в данном базисе называется матрица, составленная из координат образов базисных векторов, записанных в столбцы.

$$M(\varphi) = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix}, M(\varphi) \in P^{n\times n}.$$

Если зафиксировать базис пространства V, то каждому линейному оператору φ ставится в соответствие единственная квадратная матрица порядка: $\varphi \to M(\varphi)$.

Верно и обратное: по произвольной квадратной матрице A единственным образом можно задать линейный оператор φ, взяв за координаты образов базисных векторов столбцы матрицы A.

Пример 9.2. 1) В пространстве V размерности 3 найти матрицу оператора гомотетии.

Решение. Выбираем произвольный базис e_1, e_2, e_3, находим образы базисных векторов $\varphi(e_1) = ke_1$, $\varphi(e_2) = ke_2$, $\varphi(e_3) = ke_3$, а затем их координаты.

$$[\varphi(e_1)] = \begin{pmatrix} k \\ 0 \\ 0 \end{pmatrix}, [\varphi(e_2)] = \begin{pmatrix} 0 \\ k \\ 0 \end{pmatrix}, [\varphi(e_3)] = \begin{pmatrix} 0 \\ 0 \\ k \end{pmatrix}.$$

Составляем матрицу : $M(\varphi) = \begin{pmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{pmatrix}$. В частности, получаем матрицы тождественного и нулевого операторов: $M(\varepsilon) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = E$, $M(\theta) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = O$.

2) Найти матрицу оператора дифференцирования в пространстве $R[x]_{(\le 3)}$ в базисе $e_1 = 1$, $e_2 = x$, $e_3 = x^2$, $e_4 = x^3$.

Решение. Найдем образы базисных векторов, их координаты и составим матрицу линейного оператора.

$$\varphi(e_1) = 0,\ \varphi(e_2) = 1,\ \varphi(e_3) = 2x,\ \varphi(e_4) = 3x^2;$$

$$[\varphi(e_1)] = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},\ [\varphi(e_2)] = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},\ [\varphi(e_3)] = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix},\ [\varphi(e_4)] = \begin{pmatrix} 0 \\ 0 \\ 3 \\ 0 \end{pmatrix};$$

$$M(\varphi) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

9.2. Матрица линейного оператора

Связь между координатами вектора и координатами его образа

В пространстве V задан линейный оператор φ, а также в некотором базисе $e_1, e_2, \ldots, e_n$ найдена его матрица $M(\varphi)$. Пусть в этом базисе найдены координаты векторов x и $\varphi(x)$: $[x] = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $[\varphi(x)] = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$. Установим связь между столбцами $[x]$ и $[\varphi(x)]$.

$\varphi(x) = y_1e_1 + y_2e_2 + \ldots + y_ne_n$;

$\varphi(x) = \varphi(x_1e_1 + x_2e_2 + \ldots + x_ne_n) = x_1\varphi(e_1) + x_2\varphi(e_2) + \ldots + x_n\varphi(e_n) =$
$= x_1(\alpha_{11}e_1 + \alpha_{21}e_2 + \ldots + \alpha_{n1}e_n) + x_2(\alpha_{12}e_1 + \alpha_{22}e_2 + \ldots + \alpha_{n2}e_n) + \ldots$
$\ldots + x_n(\alpha_{1n}e_1 + \alpha_{2n}e_2 + \ldots + \alpha_{nn}e_n) = (x_1\alpha_{11} + x_2\alpha_{12} + \ldots + x_n\alpha_{1n})e_1 +$
$(x_1\alpha_{21} + x_2\alpha_{22} + \ldots + x_n\alpha_{2n})e_2 + \ldots + (x_1\alpha_{n1} + x_2\alpha_{n2} + \ldots + x_n\alpha_{nn})e_n$.

Вектор $\varphi(x)$ разложен по векторам базиса $e_1, e_2, \ldots, e_n$ двумя способами, но в силу единственности такого разложения коэффициенты при одинаковых базисных векторах можно приравнять:

$$y_1 = x_1\alpha_{11} + x_2\alpha_{12} + \ldots + x_n\alpha_{1n},$$
$$y_2 = x_1\alpha_{21} + x_2\alpha_{22} + \ldots + x_n\alpha_{2n},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..$$
$$y_n = x_1\alpha_{n1} + x_2\alpha_{n2} + \ldots + x_n\alpha_{nn}.$$

Полученные равенства можно записать в матричной форме:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ или } [\varphi(x)] = M(\varphi)\cdot[x].$$

Теорема 9.2 (о матрице линейного оператора). Если для любого вектора x из пространства V выполняется матричное равенство $[\varphi(x)] = B\cdot[x]$, то матрица B является матрицей линейного оператора φ.

Матрицы линейного оператора в различных базисах

Зададим в пространстве V два базиса $e_1, e_2, \ldots, e_n$ и $e'_1, e'_2, \ldots, e'_n$ (старый и новый). Связь между двумя базисами выражается матрицей перехода T . В пространстве V действует линейный оператор φ. В каждом из этих базисов для линейного оператора найдены матрицы. Обозначим их, соответственно, $M(\varphi)$и $M'(\varphi)$ и установим, как одна из них выражается через другую.

Пусть $[x]$ и $[x]'$ столбцы координат произвольного вектора x в старом и новом базисах соответственно, связь между которыми дает формула: $[x] = T\cdot[x]'$. Вектор $\varphi(x)$ – образ вектора x, пусть $[\varphi(x)]$ и $[\varphi(x)]'$ – столбцы координат вектора $\varphi(x)$ в старом и новом базисах соответственно. Имеет место формула $[\varphi(x)] = T\cdot[\varphi(x)]'$.

Вставим в соотношение $[\varphi(x)] = M(\varphi)\cdot[x]$ выражение старых координат векторов x и $\varphi(x)$ через новые: $T\cdot[\varphi(x)]' = M(\varphi)\cdot T\cdot[x]'$. Умножим полученное равенство слева на матрицу T^{-1} и получим $[\varphi(x)]' = (T^{-1}\cdot M(\varphi)\cdot T)\cdot[x]'$.

Из теоремы 9.2 о матрице линейного оператора следует, что

$$M'(\varphi) = T^{-1}\cdot M(\varphi)\cdot T.$$

Пример 9.3. 1) Линейный оператор φ в базисе e_1, e_2 задан формулой $\varphi(x) = (3x_1 - x_2, x_1 + x_2)$. Найти матрицу этого линейного оператора в базисе e'_1, e'_2 если $e'_1 = 3e_1 + 2e_2$, $e'_2 = 4e_1 + 3e_2$.

Решение. Сначала составим матрицу линейного оператора в старом базисе, для чего нужны координаты образов базисных векторов:

$\varphi(e_1) = \varphi(1, 0) = (3\cdot1 - 0, 1 + 0) = (3, 1)$,

$\varphi(e_2) = \varphi(0, 1) = (3\cdot0 - 1, 0 + 1) = (-1, 1)$,

$$M(\varphi) = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}.$$

Затем находим матрицу перехода T и обратную к ней матрицу T^{-1}:

$e'_1 = 3e_1 + 2e_2 \Rightarrow e'_1 = (3, 2)$,

$e'_2 = 4e_1 + 3e_2 \Rightarrow e'_2 = (4, 3)$,

$$T = \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix} \text{ тогда } T^{-1} = \begin{pmatrix} 3 & -4 \\ -2 & 3 \end{pmatrix}.$$

Используем формулу и находим $M'(\varphi) = T^{-1}\cdot M(\varphi)\cdot T$:

$$M'(\varphi) = \begin{pmatrix} 3 & -4 \\ -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 5 & -7 \\ -3 & 5 \end{pmatrix} \cdot \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 3 \end{pmatrix}.$$

Ответ: $M'(\varphi) = \begin{pmatrix} 1 & -1 \\ 1 & 3 \end{pmatrix}$.

2) Линейный оператор φ в базисе e_1, e_2 задан формулой $\varphi(x) = (2x_1 + 4x_2,\ -x_1 - 3x_2)$. Найти матрицу этого линейного оператора в базисе e'_1, e'_2 если $e'_1 = -4e_1 + e_2$, $e'_2 = -e_1 + e_2$.

Решение. По рассмотренному алгоритму найдем $M(\varphi)$ и $M'(\varphi)$.

Ответ: $M(\varphi) = \begin{pmatrix} 2 & 4 \\ -1 & -3 \end{pmatrix}$, $M'(\varphi) = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}$.

Отметим, что в новом базисе матрица линейного оператора приняла диагональный вид.

9.3. Подобные матрицы

Рассмотрим множество $P^{n\times n}$ квадратных матриц порядка n с элементами из произвольного поля P.

Введем на этом множестве отношение между матрицами – отношение подобия.

Определение 9.4. Матрица A называется *подобной* матрице B, если существует обратимая матрица T такая, что $A = T^{-1}\cdot B\cdot T$. Обозначение $A \sim B$.

Свойства отношения подобия матриц

1. Рефлексивность. Любая матрица подобна сама себе, т. е. $A \sim A$.
2. Симметричность. Если матрица A подобна B, то и B подобна A, т. е. $A \sim B \Rightarrow B \sim A$.
3. Транзитивность. Если матрица A подобна B и матрица B подобна C, то матрица A подобна матрице C, т. е. $A \sim B$ и $B \sim C \Rightarrow$ то $A \sim C$.

Множество $P^{n\times n}$ рассмотренным отношением разбивается на подмножества (классы), в каждое из которых входят матрицы, подобные между собою. Общих элементов у полученных подмножеств нет.

Вспомним связь между матрицами линейного оператора в различных базисах: $M'(\varphi) = T^{-1}\cdot M(\varphi)\cdot T$. Очевидно, что $M'(\varphi) \sim M(\varphi)$. Сформулируем правило: *матрицы одного и того же линейного оператора в различных базисах подобны между собою*.

9.4. Действия над линейными операторами

В векторном пространстве V над произвольным полем P заданы линейные операторы φ и ψ.

1. Сложение линейных операторов.

Определение 9.5. *Суммой* линейных операторов φ и ψ называется отображение, обозначаемое $\varphi + \psi$ и действующее по правилу: $(\varphi + \psi)(x) = \varphi(x) + \psi(x)$, для $\forall\ x \in V$.

Теорема 9.3. Сумма линейных операторов является линейным оператором.

Доказательство. Проверим два свойства.

1. $(\varphi + \psi)(x + y) = \varphi(x + y) + \psi(x + y) =\varphi(x) + \varphi(y) + \psi(x) + \psi(y) = \varphi(x) +$ $+ \psi(x) + \varphi(y) + \psi(y) = (\varphi +\psi)(x) + (\varphi + \psi)(y)$.

2. $(\varphi + \psi)(\lambda x) = \varphi(\lambda x) + \psi(\lambda x) = \lambda\varphi(x) + \lambda\psi(x) = \lambda(\varphi + \psi)(x)$.

В ходе доказательства использовались определение линейного оператора и определение суммы линейных операторов. Теорема доказана.

Свойства сложения линейных операторов

1. $\varphi + \psi = \psi + \varphi$.
2. $(\varphi + \psi) + \eta = \varphi + (\psi + \eta)$.
3. $\exists\ \theta : \forall\ \varphi,\ \varphi + \theta =\varphi$, где θ – нулевой оператор.
4. $\forall\ \varphi,\ \exists\ (-\varphi) : \varphi + (-\varphi) = \theta$, где $(-\varphi)$ – противоположный оператор.

2. Умножение линейного оператора на элемент поля.

Определение 9.6. *Произведением* линейного оператора φ на элемент λ поля P называется отображение, обозначаемое $\lambda\varphi$, действующее по правилу $(\lambda\cdot\varphi)(x) = \lambda\varphi(x)$, для $\forall\ x \in V$.

Теорема 9.4. Произведение линейного оператора φ на элемент λ поля P является линейным оператором.

Свойства умножения линейного оператора на элемент λ поля *P*

1. $1\cdot\varphi = \varphi$.
2. $(\lambda\cdot\mu)\cdot\varphi = \lambda\cdot(\mu\cdot\varphi) = \mu\cdot(\lambda\cdot\varphi)$.
3. $(\lambda + \mu)\cdot\varphi = \lambda\cdot\varphi + \mu\cdot\varphi$.
4. $\lambda\cdot(\varphi + \psi) = \lambda\cdot\varphi + \lambda\cdot\psi$.

3. Умножение линейных операторов.

Определение 9.7. *Произведением* линейных операторов φ и ψ называется отображение, обозначаемое $\varphi\cdot\psi$ и действующее по правилу: $(\varphi\cdot\psi)(x) = \varphi(\psi(x))$, для $\forall\ x \in V$.

Теорема 9.5. Произведение линейных операторов является линейным оператором.

Свойства умножения линейных операторов

1. $\varphi\cdot\psi \neq \psi\cdot\varphi$.
2. $(\varphi\cdot\psi)\cdot\eta = \varphi\cdot(\psi\cdot\eta)$.
3. $(\varphi + \psi)\cdot\eta = \varphi\cdot\eta + \psi\cdot\eta$.
4. $\eta\cdot(\varphi + \psi) = \eta\cdot\varphi + \eta\cdot\psi$.
5. $(\lambda\cdot\varphi)\cdot\eta = \varphi\cdot(\lambda\cdot\eta)$.

Связь между действиями над линейными операторами и действиями над их матрицами

В векторном пространстве V над произвольным полем P выбран произвольный базис $e_1,\ e_2,\ \ldots,\ e_n$. В пространстве V заданы линейные операторы φ и ψ, для которых в данном базисе найдены матрицы: $M(\varphi)$, $M(\psi)$.

Теорема 9.6. а) Матрица суммы линейных операторов равна сумме их матриц, то есть $M(\varphi + \psi) = M(\varphi) + M(\psi)$.

б) Матрица произведения линейного оператора на элемент λ равна произведению его матрицы на этот элемент λ, то есть $M(\lambda\cdot\varphi) = \lambda\cdot M(\varphi)$.

в) Матрица произведения линейных операторов равна произведению их матриц, то есть $M(\varphi\cdot\psi) = M(\varphi)\cdot M(\psi)$.

9.5. Ядро и образ линейного оператора

В векторном пространстве V над произвольным полем P задан линейный оператор φ.

Определение9.8. *Ядром* линейного оператора φ называется множество векторов пространства V , образом которых является нулевой вектор. Принятое обозначение для этого множества: $Ker\varphi$, т. е.

$$Ker\varphi = \{x \mid \varphi(x) = o\}.$$

Теорема 9.7. Ядро линейного оператора является подпространством пространства V.

Определение 9.9. Размерность ядра линейного оператора называется *дефектом* линейного оператора. $dim\ Ker\varphi = d$.

Определение 9.10. *Образом* линейного оператора φ называется множество образов векторов пространства V . Обозначение для этого множества $Im\varphi$, т. е. $Im\varphi = \{\varphi(x) \mid x \in V\}$.

Теорема 9.8. Образ линейного оператора является подпространством пространства V.

Определение 9.11. Размерность образа линейного оператора называется *рангом* линейного оператора. $dim\ Im\varphi = r$.

Теорема 9.9. Пространство V является прямой суммой ядра и образа заданного в нем линейного оператора. Сумма ранга и дефекта линейного оператора равна размерности пространства V.

Пример 9.3. 1) В пространстве $R[x]_{(\leq 3)}$ найти ранг и дефект оператора дифференцирования. Найдем те многочлены, производная которых равна нулю. Это многочлены нулевой степени, следовательно, $Ker\varphi = \{f \mid f = c\}$ и $d = 1$. Производные многочленов, степень которых не превосходит трех, образуют множество многочленов, степень которых не превосходит двух, следовательно, $Im\varphi = R[x]_{(\leq 2)}$ и $r = 3$.

2) Если линейный оператор задан матрицей $M(\varphi)$, то для нахождения его ядра надо решить уравнение $\varphi(x) = o$, которое в матричной форме выглядит так: $M(\varphi)[x] = [o]$. Из этого следует, что базисом ядра линейного оператора является фундаментальный набор решений однородной системы линейных уравнений с основной матрицей $M(\varphi)$. Систему образующих образа линейного оператора составляют векторы $\varphi(e_1)$, $\varphi(e_2)$, …, $\varphi(e_n)$. Базис этой системы векторов дает базис образа линейного оператора.

9.6. Обратимые линейные операторы

Определение 9.12. Линейный оператор φ называется *обратимым*, если существует линейный оператор ψ такой что выполняется равенство $\varphi\cdot\psi = \psi\cdot\varphi = \varepsilon$, где ε – тождественный оператор.

Теорема 9.10. Если линейный оператор φ обратим, то оператор ψ определяется единственным образом и называется обратным для оператора φ.

В этом случае оператор, обратный для оператора φ, обозначается φ^{-1}.

Теорема 9.11. Линейный оператор φ обратим тогда и только тогда, когда обратима его матрица $M(\varphi)$, при этом $M(\varphi^{-1}) = (M(\varphi))^{-1}$.

Из этой теоремы следует, что ранг обратимого линейного оператора равен размерности пространства, а дефект равен нулю.

Пример 9.4 1) Определить, обратим ли линейный оператор φ, если $\varphi(x) = (2x_1 - x_2, -4x_1 + 2x_2)$.

Решение. Составим матрицу этого линейного оператора: $M(\varphi) = \begin{pmatrix} 2 & -1 \\ -4 & 2 \end{pmatrix}$. Так как $\begin{vmatrix} 2 & -1 \\ -4 & 2 \end{vmatrix} = 0$ то матрица $M(\varphi)$ необратима, а значит, необратим и линейный оператор φ.

2) Найти линейный оператор, обратный оператору φ, если $\varphi(x) = (2x_1 + x_2, 3x_1 + 2x_2)$.

Решение. Матрица этого линейного оператора, равная $M(\varphi) = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$, обратима, так как $|M(\varphi)| \neq 0$. $(M(\varphi))^{-1} = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}$, поэтому $\varphi^{-1} = (2x_1 - x_2, -3x_1 + 2x_2)$.

9.7. Собственные векторы линейного оператора

В векторном пространстве V над произвольным полем P задан линейный оператор φ.

Определение 9.13. Ненулевой вектор x называется *собственным вектором* линейного оператора φ с собственным значением λ, если $\varphi(x) = \lambda x$.

Говорят, что вектор x принадлежит собственному значению λ.

При этом λ называется не только собственным значением вектора x, но и *собственным значением линейного оператора* φ.

Пример 9.5. **1)** Любой ненулевой вектор является собственным вектором оператора гомотетии.

2) Рассмотрим оператор дифференцирования в пространстве дифференцируемых функций. Вектор $f = e^{3x}$ является собственным вектором этого оператора с собственным значением 3, так как $f' = 3e^{3x} = 3f$.

3) Для линейного оператора, заданного матрицей $M(\varphi) = \begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$

собственным является вектор $c = (1, 2, 0)$, так как $\varphi(c) = 2c$. Проверим это:

$$[\varphi(c)] = M(\varphi)[c] = \begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = 2[c].$$

9.7.1. Свойства собственных векторов

1. Каждый собственный вектор принадлежит только одному собственному значению.

Доказательство. Пусть x собственный вектор с двумя собственными значениями λ_1 и λ_2. Тогда $\varphi(x) = \lambda_1 x$ и $\varphi(x) = \lambda_2 x$. Отсюда $\lambda_1 x = \lambda_2 x \Rightarrow$ $(\lambda_1 - \lambda_2)x = 0$ и так как вектор x ненулевой, то $(\lambda_1 - \lambda_2) = 0 \Rightarrow \lambda_1 = \lambda_2$.

2. Если вектор x – собственный вектор линейного оператора φ с собственным значением λ, то вектор $y = kx$ ($k \neq 0$) тоже собственный с тем же собственным значением λ.

Доказательство. $\varphi(y) = \varphi(kx) = k\varphi(x) = k(\lambda x) = \lambda(kx) = \lambda y$. Следовательно, вектор y – собственный вектор оператора φ с собственным значением λ.

3. Множество собственных векторов с одним и тем собственным значением λ при добавлении нулевого вектора образует подпространство пространства V.

Доказательство. Обозначим это множество символом $L^{(\lambda)}$. Докажем, что множество $L^{(\lambda)} \cup \{o\}$ образует подпространство пространства V, для чего проверим его замкнутость относительно сложения векторов и умножения их на элемент поля.

Пусть x, $y \in L^{(\lambda)} \cup \{o\}$, тогда $\varphi(x) = \lambda x$, $\varphi(y) = \lambda y$. Найдем $\varphi(x + y)$: $\varphi(x + y) = \varphi(x) + \varphi(y) = \lambda x + \lambda y = \lambda(x + y)$, значит $(x + y) \in L^{(\lambda)} \cup \{o\}$.

Пусть $x \in L^{(\lambda)} \cup \{o\}$, $k \in P$, $\varphi(x) = \lambda x$, тогда $\varphi(kx) = k\varphi(x) = k(\lambda x) = \lambda(kx)$, значит $kx \in L^{(\lambda)} \cup \{o\}$.

4. Собственные векторы с попарно различными собственными значениями линейно независимы.

Следствие. Линейный оператор, заданный в n-мерном линейном пространстве, не может иметь более чем n различных собственных значений.

9.7.2. Характеристический многочлен матрицы

Дана матрица $A \in P^{n\times n}$ (или $A \in R^{n\times n}$).

Определение 9.14. *Характеристическим многочленом* матрицы A называется многочлен, зависящий от λ и равный $|A - \lambda E|$, т. е. многочлен $f(\lambda) = |A - \lambda E|$, где E – единичная матрица таково же порядка что и A.

Пример 9.6. Найти характеристический многочлен матрицы $A = \begin{pmatrix} 1 & 1 & 8 \\ 0 & 2 & 0 \\ 1 & 0 & -1 \end{pmatrix}$.

Решение. Составим определитель матрицы $A - \lambda E$ и вычислим его

$$|A - \lambda E| = \left| \begin{pmatrix} 1 & 1 & 8 \\ 0 & 2 & 0 \\ 1 & 0 & -1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right| = \left| \begin{pmatrix} 1-\lambda & 1 & 8 \\ 0 & 2-\lambda & 0 \\ 1 & 0 & -1-\lambda \end{pmatrix} \right| =$$

$$= (2 - \lambda)\cdot(-1)^{2+2}\cdot (-1)^{2+2} \cdot \begin{vmatrix} 1-\lambda & 8 \\ 1 & -1-\lambda \end{vmatrix} = (2 - \lambda)((1 - \lambda)(-1 - \lambda) - 8) =$$

$$= (2 - \lambda)(\lambda - 3)(\lambda + 3).$$

Определение 9.15. *Характеристическим уравнением* матрицы A называется уравнение $|A - \lambda E| = 0$.

Определение 9.16. *Собственными значениями* (собственными числами) матрицы A называются корни ее характеристического уравнения.

Найдем собственные значения матрицы из примера 9.6: $(2 - \lambda)(\lambda - 3)(\lambda + 3) = 0$. Получим: $\lambda_1 = 2$, $\lambda_2 = 3$, $\lambda_3 = -3$.

Теорема 9.12. Характеристические многочлены подобных матриц равны.

Доказательство. Пусть матрицы A и B подобны, то есть $A = T^{-1}\cdot B\cdot T$. Тогда $|A - \lambda E| = |T^{-1}\cdot B\cdot T - \lambda T^{-1}\cdot E\cdot T| = |T^{-1}(B - \lambda E)\cdot T| = |T^{-1}|\cdot|B - \lambda E|\cdot|T| =$
$= |T^{-1}|\cdot|T|\cdot|B - \lambda E| = |T^{-1}\cdot T|\cdot|B - \lambda E| = |E|\cdot|B - \lambda E| = 1\cdot|B - \lambda E| = |B - \lambda E|$.
Так как характеристические многочлены равны, то совпадают и множества собственных значений подобных матриц A и B.

Следствие. Характеристический многочлен матрицы линейного оператора не зависит от базиса, в котором найдена эта матрица (матрицы линейного оператора, найденные в различных базисах, подобны).

9.7.3. Нахождение собственных векторов линейного оператора

Для нахождения собственных векторов линейного оператора φ надо найти решения уравнения $\varphi(x) = \lambda x$, в котором неизвестными величинами являются собственные значения λ линейного оператора φ и ненулевые векторы x.

Для нахождения собственных значений линейного оператора φ используется следующая теорема.

Теорема 9.13. Множество собственных значений линейного оператора φ совпадает с множеством собственных значений его матрицы.

Доказательство. Пусть λ – собственное значение линейного оператора φ. Это означает, что существует ненулевой вектор x, такой что $\varphi(x) = \lambda x$. Из этого векторного равенства вытекают следующие матричные равенства:

$[\varphi(x)] = [\lambda x] \Rightarrow M(\varphi)[x] = \lambda[x] \Rightarrow M(\varphi)[x] - \lambda E[x] = [0] \Rightarrow$

$$(M(\varphi) - \lambda E)[x] = [0]. \qquad (*)$$

Равенство $(*)$ является матричной формой записи однородной системы линейных уравнений с основной матрицей $M(\varphi) - \lambda E$, причем эта система по условию имеет ненулевые решения. Условие существования ненулевых решений – равенство нулю определителя основной матрицы системы, следовательно, выполнятся равенство: $|M(\varphi) - \lambda E| = 0$, из которого следует, что λ – собственное значение матрицы $M(\varphi)$ линейного оператора $\varphi(x) = \lambda x$.

Обратно, пусть λ – собственное значение матрицы $M(\varphi)$ линейного оператора φ, то есть $|M(\varphi) - \lambda E| = 0$. Из этого равенства следует, что система однородных линейных уравнений (∗) имеет ненулевые решения. Следовательно, существует ненулевой вектор x, такой что $\varphi(x) = \lambda x$ и λ – собственное значение линейного оператора φ.

9.7.4. Алгоритм нахождения собственных векторов линейного оператора

1. Найти собственные значения линейного оператора как собственные значения его матрицы

2. Для каждого из найденных собственных значений λ_0 находим собственные векторы, решая однородную систему линейных уравнений (∗) с основной матрицей $M(\varphi) - \lambda_0 E$.

3. Множество $L^{(\lambda_0)}$ равно линейной оболочке фундаментального набора решений этой системы за исключением нулевого вектора.

Пример 9.7. Найти собственные векторы линейного оператора с матрицей $M(\varphi) = \begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$.

Решение. Находим собственные значения матрицы линейного оператора, для чего решаем уравнение $|M(\varphi) - \lambda E| = 0$.

$$\begin{vmatrix} -\lambda & 1 & 0 \\ -4 & 4-\lambda & 0 \\ -2 & 1 & 2-\lambda \end{vmatrix} = (2-\lambda)(-1)^{3+3}\begin{vmatrix} -\lambda & 1 \\ -4 & 4-\lambda \end{vmatrix} = (2-\lambda)((-\lambda)(4-\lambda) - (-4)) =$$

$$= (2-\lambda)(\lambda^2 - 4\lambda + 4) = (2-\lambda)(\lambda-2)^2 = (2-\lambda)^3 = 0 \Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = 2.$$

Итак, получили $f(\lambda) = (2 - \lambda)^3$ – характеристический многочлен матрицы $M(\varphi)$; $(2 - \lambda)^3 = 0$ – характеристическое уравнение матрицы $M(\varphi)$; $\lambda_1 = \lambda_2 = \lambda_3 = 2$ – собственные значения матрицы $M(\varphi)$, т. е. это собственные значения линейного оператора φ.

Собственное значение у этого линейного оператора только одно, поэтому решаем только одну однородную систему линейных уравнений с матрицей $\begin{pmatrix} -2 & 1 & 0 \\ -4 & 2 & 0 \\ -2 & 1 & 0 \end{pmatrix}$.

$$\begin{pmatrix} -2 & 1 & 0 \\ -4 & 2 & 0 \\ -2 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} -2 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} -2 & 1 & 0 \end{pmatrix}.$$

Выпишем общее решение этой системы $x_1 = \dfrac{1}{2}(x_2 + 0x_3)$ и составим фундаментальный набор решений

	x_1	x_2	x_3
c_1	1	**2**	**0**
c_1	0	**0**	**1**

$c_1 = (1, 2, 0)$, $c_2 = (0, 0, 1)$.

Ответ. Множество собственных векторов с собственным значением $\lambda = 2$ это множество $L = L(c_1, c_2)\backslash\{o\} = \{k_1c_1 + k_2c_2,\ k_1^2 + k_2^2 \neq 0\}$.

9.7.5. Условия, при которых матрица подобна диагональной матрице

Пусть A – квадратная матрица. Можно считать, что это матрица некоторого линейного оператора, заданного в каком-то базисе. Известно, что в другом базисе матрица линейного оператора примет другой вид, в частности, как в одном из предыдущих примеров 9.3, диагональный. Это значит, что исходная матрица подобна диагональной матрице. Возникает вопрос: всегда ли данная матрица подобна диагональной? Как это установить? Как найти соответствующий базис?

Теорема 9.14. Матрица A подобна диагональной матрице тогда и только тогда, когда линейный оператор φ, заданный этой матрицей, имеет n линейно независимых собственных векторов.

Доказательство. Пусть матрица A подобна диагональной матрице, то есть у линейного оператора φ с матрицей $A = M(\varphi)$ в некотором базисе c_1, c_2, …, c_n матрица примет следующий вид $M'(\varphi) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$.

Используя матрицу, найдем образы базисных векторов: $\varphi(c_1) = \lambda_1 c_1$, $\varphi(c_2) = \lambda_2 c_2$, …, $\varphi(c_n) = \lambda_n c_n$. Получены n линейно независимых собственных векторов.

У линейного оператора φ есть n линейно независимых собственных векторов c_1, c_2, …, c_n с собственными значениями λ_1, λ_2, …, λ_n. Выберем векторы c_1, c_2, …, c_n в качестве базисных векторов и найдем матрицу оператора φ в этом базисе. Используя равенства $\varphi(c_1) = \lambda_1 c_1$, $\varphi(c_2) = \lambda_2 c_2$, …, $\varphi(c_n) = \lambda_n c_n$ составим матрицу $M'(\varphi)$: $M'(\varphi) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$.

Теорема 9.15. Если матрица A имеет n попарно различных собственных значений, то она подобна диагональной матрице.

Это утверждение основано на свойстве собственных векторов: попарно различным собственным значениям соответствуют линейно независимые собственные векторы.

Пример 9.8. Привести матрицу A к диагональному виду, если это возможно, указать базис и матрицу перехода.

1) $A = \begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$. Для этого случая собственные векторы уже найдены (пример 9.7), линейно независимых векторов оказалось только 2, а в базисе должно быть 3. *Вывод*: матрица A к диагональному виду не приводится. Другими словами: матрица А не подобна диагональной.

2) $A = \begin{pmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{pmatrix}$. Находим собственные значения матрицы A. Вычислим определитель

$$|A - \lambda E| = \begin{vmatrix} 7-\lambda & -12 & 6 \\ 10 & -19-\lambda & 10 \\ 12 & -24 & 13-\lambda \end{vmatrix} = \begin{vmatrix} 7-\lambda & 0 & 6 \\ 10 & 1-\lambda & 10 \\ 12 & 2-2\lambda & 13-\lambda \end{vmatrix} =$$

$$=(1-\lambda)\begin{vmatrix} 7-\lambda & 0 & 6 \\ 10 & 1 & 10 \\ 12 & 2 & 13-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} 7-\lambda & 0 & 6 \\ 10 & 1 & 10 \\ -8 & 0 & -7-\lambda \end{vmatrix} =$$

$$= (1-\lambda)\cdot 1\cdot(-1)^{2+2}\cdot\begin{vmatrix} 7-\lambda & 6 \\ -8 & -7-\lambda \end{vmatrix} = (1-\lambda)((7-\lambda)(-7-\lambda) - 6(-8)) =$$

$= (1-\lambda)(\lambda^2 - 1) = -(\lambda + 1)(\lambda - 1)^2 = 0.$ Тогда $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = -1$ – собственные значения матрицы A.

Находим собственные векторы, соответствующие этим собственным значениям. Рассмотрим случай $\lambda_1 = \lambda_2 = 1$. Решаем однородную систему линейных уравнений $\begin{pmatrix} 6 & -12 & 6 \\ 10 & -20 & 10 \\ 12 & -24 & 12 \end{pmatrix} \sim (1\ \ -2\ \ 1)$, тогда $x_1 = 2x_2 - x_3$ –общее решение системы, векторы $c_1 = (2, 1, 0)$ $c_2 = (1, 0, -1)$ линейно независимые собственные векторы с собственным значением $\lambda_1 = \lambda_2 = 1$.

Рассмотрим случай $\lambda_3 = -1$. Получаем систему $\begin{pmatrix} 8 & -12 & 6 \\ 10 & -18 & 10 \\ 12 & -24 & 14 \end{pmatrix}$. Решая ее, получим только один линейно независимый собственный вектор $c_3 = (3, 5, 6)$.

Найдены три линейно независимых собственных вектора c_1, c_2, c_3. Выберем их в качестве нового базиса и найдем матрицу линейного оператора в этом базисе.

Поскольку $\varphi(c_1) = 1c_1$, $\varphi(c_2) = 1c_2$, $\varphi(c_3) = (-1)c_3$, то матрица линейного оператора $M'(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ и $T = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 0 & 5 \\ 0 & -1 & 6 \end{pmatrix}$ –матрица перехода от старого базиса к новому.

10. Жорданова нормальная форма матрицы линейного оператора

10.1. Понятие λ-матрицы

Известно, что матрица линейного оператора в базисе из собственных векторов приводится к диагональному виду. Однако над множеством действительных чисел линейный оператор может не иметь собственных значений, а значит и собственных векторов. Над множеством комплексных чисел любой линейный оператор имеет собственные векторы, но их может быть недостаточно для базиса. Есть другая каноническая форма матрицы линейного оператора, к которой можно привести любую матрицу над множеством комплексных чисел.

Теорема 10.1. Всякая матрица с комплексными элементами приводится во множестве комплексных чисел $\mathbb{C}$ к жордановой[14] нормальной форме.

Дадим необходимые определения:

Определение 10.1. Квадратная матрица порядка n, элементами которой служат многочлены произвольной степени от переменной λ с коэффициентами из множества комплексных чисел $\mathbb{C}$, называется *λ-матрицей* (или *многочленной матрицей*, или *полиномиальной матрицей*).

Примером многочленной матрицы служит характеристическая матрица $A - \lambda E$ произвольной квадратной матрицы A. На главной диагонали стоят многочлены первой степени, вне ее – многочлены нулевой степени или нули. Обозначим такую матрицу как $A(\lambda)$.

Пример 10.1. Пусть дана матрица $A = \begin{pmatrix} 3 & 2 & -5 \\ 2 & 6 & -10 \\ 1 & 2 & -3 \end{pmatrix}$, тогда $A - \lambda E =$

$$= \begin{pmatrix} 3-\lambda & 2 & -5 \\ 2 & 6-\lambda & -10 \\ 1 & 2 & -3-\lambda \end{pmatrix} = A(\lambda).$$

Определение 10.2. *Элементарными преобразованиями* λ-матрицы называют следующие преобразования:

1. умножение любой строки (столбца) матрицы $A(\lambda)$ на любое число, не равное нулю;
2. прибавление к любой i-той строке (i-ому столбцу) матрицы $A(\lambda)$ любой другой j-ой строки (j-ого столбца), умноженной на произвольный многочлен $\varphi(\lambda)$.

Свойства λ-матрицы

1) С помощью этих преобразований в матрице $A(\lambda)$ можно переставить любые две строки или любые два столбца.

[14] Камиль Жордан (1838–1922) – французский математик.

2) С помощью этих преобразований в диагональной матрице $A(\lambda)$ можно менять местами диагональные элементы.

Пример 10.2. 1) $\begin{pmatrix} i \\ j \end{pmatrix} \sim \begin{pmatrix} i+j \\ j \end{pmatrix} \sim \begin{pmatrix} i+j \\ -j \end{pmatrix} \sim \begin{pmatrix} j \\ -i \end{pmatrix} \sim \begin{pmatrix} j \\ i \end{pmatrix}$.

2) $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \sim \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix} \sim \begin{pmatrix} b & 0 \\ 0 & a \end{pmatrix}$.

Определение 10.3. Матрицы $A(\lambda)$ и $B(\lambda)$ называются *эквивалентными*, если от $A(\lambda)$ можно перейти к $B(\lambda)$ при помощи конечного числа элементарных преобразований.

Задача заключается в том, чтобы по возможности упростить матрицу $A(\lambda)$.

Определение 10.4. *Канонической λ-матрицей* называется λ-матрица, обладающая следующими свойствами:

1) матрица $A(\lambda)$ диагональная;
2) всякий многочлен $e_i(\lambda)$, $i = 1, 2, \ldots, n$ нацело делится на $e_{i-1}(\lambda)$;
3) старший коэффициент каждого многочлена $e_i(\lambda)$, $i = 1, 2, \ldots, n$ равен 1, или этот многочлен равен нулю.

$$A(\lambda) = \begin{pmatrix} e_1(\lambda) & 0 & \ldots & 0 \\ 0 & e_2(\lambda) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & e_n(\lambda) \end{pmatrix}.$$

Замечание. Если среди многочленов $e_i(\lambda)$ встречаются нули, то они занимают на главной диагонали последние места (по свойству 2), если есть многочлены нулевой степени, то они равны 1 и занимают на главной диагонали первые места.

Нулевая и единичная матрицы являются каноническими λ-матрицами.

Теорема 10.2. Всякая λ-матрица эквивалентна некоторой канонической λ-матрице (то есть она приводится элементарными преобразованиями к каноническому виду)

Пример 10.3. Привести матрицу $A(\lambda) = \begin{pmatrix} \lambda-2 & -1 & 0 \\ 0 & \lambda-2 & -1 \\ 0 & 0 & \lambda-2 \end{pmatrix}$ к каноническому виду.

Решение. Ход преобразований аналогичен преобразованиям в методе Гаусса, при этом левый верхний элемент матрицы при приведении ее к каноническому виду отличен от нуля и имеет наименьшую степень.

$A(\lambda) = \begin{pmatrix} \lambda-2 & -1 & 0 \\ 0 & \lambda-2 & -1 \\ 0 & 0 & \lambda-2 \end{pmatrix}$ ~ (меняем местами первый и второй столбцы) ~ $\begin{pmatrix} -1 & \lambda-2 & 0 \\ \lambda-2 & 0 & -1 \\ 0 & 0 & \lambda-2 \end{pmatrix}$ ~ (к второму столбцу прибавляем первый столбец, умноженный на $(\lambda - 2)$) ~ $\begin{pmatrix} -1 & 0 & 0 \\ \lambda-2 & (\lambda-2)^2 & -1 \\ 0 & 0 & \lambda-2 \end{pmatrix}$ ~ (ко второй строке прибавляем первую строку, умноженную на $(\lambda - 2)$) ~ $\begin{pmatrix} -1 & 0 & 0 \\ 0 & (\lambda-2)^2 & -1 \\ 0 & 0 & \lambda-2 \end{pmatrix}$ ~ (меняем местами второй и третий столбцы) ~ $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & (\lambda-2)^2 \\ 0 & \lambda-2 & 0 \end{pmatrix}$ ~ (к третьему столбцу прибавляем второй столбец умноженный на $(\lambda - 2)^3$) ~ $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & \lambda-2 & (\lambda-2)^2 \end{pmatrix}$ ~ (к третьей строке прибавляем вторую строку, умноженную на $(\lambda - 2)$) ~ $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & (\lambda-2)^3 \end{pmatrix}$.

10.2. Жорданова нормальная форма

Определение 10.5. *Жордановой клеткой порядка* k, относящейся к числу λ_0, называется матрица порядка k, $1 \le k \le n$, имеющая вид

$$\begin{pmatrix} \lambda_0 & 1 & 0 & \cdots & 0 \\ 0 & \lambda_0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_0 & 1 \\ 0 & 0 & \cdots & 0 & \lambda_0 \end{pmatrix},$$

на ее главной диагонали стоит одно и то же число λ_0, а на параллельной ей сверху диагонали стоят единицы, все же остальные элементы равны нулю. Например: (λ_0), $\begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix}$, $\begin{pmatrix} \lambda_0 & 1 & 0 \\ 0 & \lambda_0 & 1 \\ 0 & 0 & \lambda_0 \end{pmatrix}$ – жордановы клетки первого, второго и третьего порядков соответственно.

Определение 10.6. *Жордановой матрицей порядка* n называется матрица порядка n, имеющая вид: $J = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & J_s \end{pmatrix}$. В ней вдоль главной диагонали идут жордановы клетки J_1, J_2, …, J_s некоторых порядков, не обязательно различных, и относящиеся к некоторым числам, тоже не обязательно различным. Все места вне этих клеток заняты нулями. При этом $s \ge 1$, то есть одна жорданова клетка порядка n так же считается жордановой матрицей и $s \le n$.

Замечание. Говорят, что матрица J имеет нормальную жорданову форму. Диагональная матрица является частным случаем жордановой матрицы, у нее все клетки имеют порядок 1.

10.3.Приведение матрицы к жордановой (нормальной) форме

Теорема 10.3. Жорданова нормальная форма определяется для матрицы однозначно с точностью до порядка расположения жордановых клеток на главной диагонали.

Приведем матрицу $A(\lambda) = A - \lambda E$ к каноническому виду с помощью элементарных преобразований.

$$A - \lambda E = \begin{pmatrix} 1 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 1 & 0 & 0 & 0 & 0 \\ 0 & \cdots & 0 & e_{n-j+1}(\lambda) & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & 0 & e_{n-1}(\lambda) & 0 \\ 0 & \cdots & 0 & 0 & 0 & 0 & e_n(\lambda) \end{pmatrix}.$$

Отличные от единицы многочлены $e_{n-j+1}(\lambda)$, …, $e_{n-1}(\lambda)$, $e_n(\lambda)$ называют *инвариантными* множителями матрицы $A(\lambda)$. Среди них нет многочленов равных нулю, сумма степеней всех этих многочленов равна n, и все они раскладываются на линейные множители над множеством комплексных чисел. Пусть $e_{n-j+1}(\lambda)$ раскладывается в произведение следующих множителей: $(\lambda - \lambda_1)^{k_{1j}}$, $(\lambda - \lambda_2)^{k_{2j}}$, …, $(\lambda - \lambda_t)^{k_{tj}}$. Назовем эти множители *элементарными делителями* многочлена $e_{n-j+1}(\lambda)$.

Определение 10.7. *Элементарными делителями матрицы* $A(\lambda)$ называются элементарные делители всех многочленов $e_{n-j+1}(\lambda)$, …, $e_{n-1}(\lambda)$, $e_n(\lambda)$.

Выпишем жорданову матрицу J порядка n, составленную из жордановых клеток определяемых следующим образом: каждому элементарному делителю $(\lambda - \lambda_i)^{k_{ij}}$ матрицы $A(\lambda)$ ставим в соответствие жорданову клетку порядка k_{ij}, относящуюся к числу λ_i.

Пусть для некоторой матрицы порядка 9 характеристическая матрица $A - \lambda E$ приведена к каноническому виду.

$$A-\lambda E=\begin{pmatrix}1&0&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0&0\\0&0&1&0&0&0&0&0&0\\0&0&0&1&0&0&0&0&0\\0&0&0&0&1&0&0&0&0\\0&0&0&0&0&1&0&0&0\\0&0&0&0&0&0&\lambda-2&0&0\\0&0&0&0&0&0&0&(\lambda-2)(\lambda-5)^2&0\\0&0&0&0&0&0&0&0&(\lambda-2)^3(\lambda-5)^2\end{pmatrix}.$$

$e_1 = e_2 = e_3 = e_4 = e_5 = e_6 = 1$, $e_7 = \lambda - 2$, $e_8 = (\lambda - 2)(\lambda - 5)^2$, $e_9 = (\lambda - 2)^3(\lambda - 5)^2$ – инвариантные множители матрицы $A - \lambda E$, $(\lambda - 2)$, $(\lambda - 2)$, $(\lambda - 5)^2$, $(\lambda - 2)^3$, $(\lambda - 5)^2$ – элементарные делители матрицы $A - \lambda E$.

Получаем: две клетки порядка 1, относящиеся к числу 2, две клетки порядка 2, относящиеся к числу 5 , одну клетку порядка 3, относящуюся к числу 2, Выпишем жорданову форму матрицы A

$$\begin{pmatrix}2&0&0&0&0&0&0&0&0\\0&2&0&0&0&0&0&0&0\\0&0&5&1&0&0&0&0&0\\0&0&0&5&0&0&0&0&0\\0&0&0&0&5&1&0&0&0\\0&0&0&0&0&5&0&0&0\\0&0&0&0&0&0&2&1&0\\0&0&0&0&0&0&0&2&1\\0&0&0&0&0&0&0&0&2\end{pmatrix}$$

Алгоритм приведения матрицы *A* к жордановой форме

1. Составить характеристическую матрицу $A - \lambda E$.
2. Привести эту матрицу к канонической форме с помощью элементарных преобразований.
3. Разложить диагональные многочлены на линейные множители.
4. Найти элементарные делители и по ним выписать жорданову форму матрицы A.

Для того чтобы заданная матрица была подобна диагональной матрице, необходимо и достаточно, чтобы все элементарные делители ее характеристической матрицы были первой степени.

Пример 10.4. Привести к жордановой форме матрицу $A = \begin{pmatrix} 3 & 1 & -3 \\ -7 & -2 & 9 \\ -2 & -1 & 4 \end{pmatrix}$.

Решение. с помощью элементарных преобразований приводим матрицу $A - \lambda E$ к следующему виду: $A - \lambda E = \begin{pmatrix} 3-\lambda & 1 & -3 \\ -7 & -2-\lambda & 9 \\ -2 & -1 & 4-\lambda \end{pmatrix} \sim$

$\sim \ldots \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (\lambda-1)(\lambda-2)^2 \end{pmatrix}$. Инвариантные множители этой матрицы: $e_1 = 1$, $e_2 = 1$, $e_3 = (\lambda - 1)(\lambda - 2)^2$; элементарные делители будут $(\lambda - 1)$, $(\lambda - 2)^2$.

По найденным элементарным делителям выписываем жорданову форму исходной матрицы $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

11. Билинейные и квадратичные формы

11.1. Билинейные формы

Определение 11.1. *Билинейной формой* называется функция (отображение) f: $V \times V \to \mathbb{R}$ (или $\mathbb{C}$), где V – произвольное векторное пространство, и для любых векторов $x, y \in V$ и любого числа $\lambda \in \mathbb{R}$ (или $\mathbb{C}$) выполняются соотношения

$$f(x + y, z) = f(x, y) + f(z, y),$$
$$f(x, y + z) = f(x, y) + f(x, z),$$
$$f(\lambda x, y) = \lambda f(x, y),$$
$$f(x, \lambda y) = \lambda f(x, y).$$

Обозначим число, получаемое в результате отображения пары векторов x и y, как $A(x, y)$.

Определение 11.2. Билинейная форма $A(x, y)$ называется *симметрической*, если для любых $x, y \in V$ выполняется: $A(x, y) = A(y, x)$.

Определение 11.3. Билинейная форма $A(x, y)$ называется *кососимметрической*, если для любых x, $y \in V$ выполняется: $A(x, y) = –A(y, x)$.

Свойства билинейных форм

Любую билинейную форму можно представить в виде суммы симметричной кососимметричной форм.

При выбранном базисе $e_1, e_2, \ldots, e_n$ в векторном пространстве V любая билинейная форма A однозначно определяется матрицей

$$A(e) = \begin{pmatrix} A(e_1, e_1) & A(e_1, e_2) & \ldots & A(e_1, e_n) \\ A(e_2, e_1) & A(e_2, e_2) & \ldots & A(e_2, e_n) \\ \ldots & \ldots & \ldots & \ldots \\ A(e_n, e_1) & A(e_n, e_2) & \ldots & A(e_n, e_n) \end{pmatrix},$$

так, что для любых $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$, $y = y_1e_1 + y_2e_2 + \ldots + y_ne_n$;

$$A(x, y) = (x_1\ x_2\ \ldots\ x_n)\cdot\begin{pmatrix} A(e_1, e_1) & A(e_1, e_2) & \ldots & A(e_1, e_n) \\ A(e_2, e_1) & A(e_2, e_2) & \ldots & A(e_2, e_n) \\ \ldots & \ldots & \ldots & \ldots \\ A(e_n, e_1) & A(e_n, e_2) & \ldots & A(e_n, e_n) \end{pmatrix}\cdot\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \text{ то есть}$$

$$A(x, y) = \sum\nolimits_{i,j=1}^{n} A_{ij} x_i y_j \text{ , где } A_{ij} = A(e_i, e_j). \qquad (1)$$

Вид (1) назовем *общим видом билинейной формы* в n-мерном векторном пространстве.

Замечание. Если билинейная форма $A(x, y)$ симметрическая, то и матрица (A_{ij}) будет симметрической, то есть $A_{ij} = A_{ji}$ для $\forall i, j = 1, 2, \ldots, n$. Если билинейная форма $A(x, y)$ кососимметрическая, то и матрица (A_{ij}) будет кососимметрической, то есть $A_{ij} = -A_{ji}$ для $\forall i, j = 1, 2, \ldots, n$.

Преобразование матрицы билинейной формы при переходе к новому базису. Ранг билинейной формы

Пусть в векторном пространстве V заданы два базиса $e = \{e_1, e_2, \ldots, e_n\}$ и $f = \{f_1, f_2, \ldots, f_n\}$. Пусть $A(e) = (a_{ij})$ и $A(f) = (b_{ij})$ – матрицы билинейной формы A в указанных базисах.

Теорема 11.1. Матрицы $A(e)$ и $A(f)$ билинейной формы $A(x, y)$ в базисах $\{e\}$ и $\{f\}$ связаны соотношением

$$A(f) = C^t \cdot A(e) \cdot C, \qquad (*),$$

где C – матрица перехода от базиса $\{e\}$ к базису $\{f\}$, а C^t – транспонированная матрица C.

Следствие. Ранг матрицы $A(f)$ равен рангу матрицы $A(e)$.

Это утверждение следует из равенства $(*)$: так как C – матрица перехода от одного базиса к другому, то матрица C и матрица C^t– невырожденные, поэтому умножение на них матрицы $A(e)$ не меняет ее ранга.

Определение 11.4. *Рангом билинейной формы*, заданной в конечномерном векторном пространстве V, называется ранг матрицы этой формы в произвольном базисе пространства V.

Определение 11.5. Билинейная форма называется *невырожденной*, если ее ранг равен размерности пространства V и *вырожденной*, если ее ранг меньше размерности пространства V.

11.2. Квадратичные формы

Пусть $A(x, y)$ – симметрическая билинейная форма, заданная на векторном пространстве V.

Определение 11.6. *Квадратичной формой* называется числовая функция одного векторного аргумента x, которая получается из билинейной формы $A(x, y)$ при $x = y$.

Определение 11.7. Симметрическая билинейная форма $A(x, y)$ называется *полярной* квадратичной форме $A(x, x)$.

Пусть дана билинейная форма $A(x, y) = \sum_{i,j=1}^{n} a_{ij} x_i y_j$, положим в ней $x_i = y_j$, тогда мы получим представление квадратичной формы $A(x, x)$ в конечномерном векторном пространстве V с заданным базисом $\{e\}$:

$$A(x, x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j \ . \qquad (2)$$

Определение 11.8. Матрица (a_{ij}) называется *матрицей квадратичной формы* $A(x, x)$ в заданном базисе $\{e\}$.

При переходе к новому базису матрица квадратичной формы преобразуется по формуле $A(f) = C^t \cdot A(e) \cdot C$ и ранг этой матрицы не меняется при переходе к новому базису.

Определение 11.9. Ранг матрицы квадратичной формы $A(x, x)$ называется *рангом* квадратичной формы.

Определение 11.10. Квадратичная форма называется *невырожденной*, если ее ранг равен размерности пространства V и *вырожденной*, если ее ранг меньше размерности пространства V.

Определение 11.11. Квадратичная форма $A(x, x)$ называется

1. *Положительно определенной*, если для любого ненулевого вектора x выполняется неравенство $A(x, x) > 0$.
2. *Отрицательно определенной*, если для любого ненулевого вектора x выполняется неравенство $A(x, x) < 0$.
3. *Знакопеременной*, если существуют такие x и y, что $A(x, x) > 0$ и $A(y, y) < 0$.
4. *Квазизнакоопределенной*, если для всех x, $A(x, x) \geq 0$ (или $A(x, x) \leq 0$) и имеется вектор $x \neq 0$, для которого $A(x, x) = 0$.

Замечание. Если $A(x, y)$ – билинейная форма, полярная положительно определенной квадратичной форме $A(x, x)$, то $A(x, y)$ удовлетворяет аксиомам скалярного произведения векторов в евклидовом пространстве:

1) $A(x, y) = A(y, x)$ – в силу симметричности $A(x, x)$.
2) $A(x + y, z) = A(x, z) + A(y, z)$ – в силу определения билинейной формы.
3) $A(\lambda x, y) = \lambda A(x, y)$ – в силу определения билинейной формы.
4) $A(x, x) \geq 0$ и $A(x, x) > 0$ при $x \neq 0$, т. к. $A(x, x)$ положительно определена.

Вывод. Скалярное произведение в векторных пространствах может быть задано с помощью билинейной формы:

$$(x, y) = \sum_{i,j=1}^{n} A_{ij} x_i y_j = A(x, y).$$

Приведение квадратичной формы к каноническому виду

Дана квадратичная форма (2) $A(x, x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$, где $x = (x_1, x_2, \ldots, x_n)$. Рассмотрим квадратичную форму в пространстве R^3, то есть $x = (x_1, x_2, x_3)$, $A(x, x) = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{21}x_2x_1 + a_{31}x_3x_1 + a_{23}x_2x_3 + a_{32}x_3x_2 = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3$ (использовали условие симметричности формы, а именно $a_{12} = a_{21}$, $a_{13} = a_{31}$, $a_{23} = a_{32}$). Выпишем матрицу квадратичной формы A в базисе $\{e\}$, $A(e) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. При изменении базиса матрица квадратичной формы меняется по формуле $A(f) = C^t \cdot A(e) \cdot C$, где C – матрица перехода от базиса $\{e\}$ к базису $\{f\}$, а C^t – транспонированная матрица C.

Определение 11.12. Вид квадратичной формы с диагональной матрицей называется *каноническим*.

Итак, пусть $A(f) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$, тогда $A'(x,\ x) = \lambda_1 x'^2_1 + \lambda_2 x'^2_2 + \lambda_3 x'^2_3$,

где x'_1, x'_2, x'_3 – координаты вектора x в новом базисе $\{f\}$.

Определение 11.13. Пусть в n-мерном векторном пространстве V выбран такой базис $f = \{f_1, f_2, \ldots, f_n\}$, в котором квадратичная форма имеет вид

$$A(x, x) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2, \tag{3}$$

где y_1, y_2, …, y_n – координаты вектора x в базисе $\{f\}$. Выражение (3) называется *каноническим видом* квадратичной формы. Коэффициенты λ_1, λ_2, …, λ_n называются *каноническими*; базис, в котором квадратичная форма имеет канонический вид, называется *каноническим базисом*.

Замечание. Если квадратичная форма $A(x, x)$ приведена к каноническому виду, то, вообще говоря, не все коэффициенты λ_i отличны от нуля. Ранг квадратичной формы равен рангу ее матрицы в любом базисе.

Пусть ранг квадратичной формы $A(x, x)$ равен r, где $r \le n$. Матрица квадратичной формы в каноническом виде имеет диагональный вид. $A(f) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$, поскольку ее ранг равен r, то среди коэффициентов λ_i должно быть r, не равных нулю. Отсюда следует, что число отличных от нуля канонических коэффициентов равно рангу квадратичной формы.

Замечание. Линейным преобразованием координат называется переход от переменных x_1, x_2, …, x_n к переменным y_1, y_2, …, y_n, при котором старые переменные выражаются через новые переменные с некоторыми числовыми коэффициентами.

$$x_1 = \alpha_{11}y_1 + \alpha_{12}y_2 + \ldots + \alpha_{1n}y_n,$$
$$x_2 = \alpha_{21}y_1 + \alpha_{22}y_2 + \ldots + \alpha_{2n}y_n,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$x_1 = \alpha_{n1}y_1 + \alpha_{n2}y_2 + \ldots + \alpha_{nn}y_n.$$

Так как каждому преобразованию базиса отвечает невырожденное линейное преобразование координат, то вопрос о приведении квадратичной формы к каноническому виду можно решать путем выбора соответствующего невырожденного преобразования координат.

Теорема 11.2 (основная теорема о квадратичных формах). Всякая квадратичная форма $A(x, x)$, заданная в n-мерном векторном пространстве V, с помощью невырожденного линейного преобразования координат может быть приведена к каноническому виду.

Доказательство. (Метод Лагранжа) Идея этого метода состоит в последовательном дополнении квадратного трехчлена по каждой переменной до полного квадрата. Будем считать, что $A(x, x) \neq 0$ и в базисе $e = \{e_1, e_2, \ldots, e_n\}$ имеет вид (2):

$$A(x, x) = \sum\nolimits_{i,j=1}^{n} a_{ij}x_ix_j .$$

Если $A(x, x) = 0$, то $(a_{ij}) = 0$, то есть форма уже каноническая. Формулу $A(x, x)$ можно преобразовать так, чтобы коэффициент $a_{11} \neq 0$. Если $a_{11} = 0$, то коэффициент при квадрате другой переменной отличен от нуля, тогда при помощи перенумерации переменных можно добиться, чтобы $a_{11} \neq 0$. Перенумерация переменных является невырожденным линейным преобразованием. Если же все коэффициенты при квадратах переменных равны нулю, то нужные преобразования получаются следующим образом. Пусть, например, $a_{12} \neq 0$ ($A(x, x) \neq 0$, поэтому хотя бы один коэффициент $a_{ij} \neq 0$). Рассмотрим преобразование

$x_1 = y_1 - y_2,$

$x_2 = y_1 + y_2,$

$x_i = y_i$, при $i = 3, 4, \ldots, n$.

Это преобразование невырожденное, так как определитель его матрицы отличен от нуля $\begin{vmatrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} = 2 \neq 0.$

Тогда $2a_{12}x_1x_2 = 2\, a_{12}(y_1 - y_2)(y_1 + y_2) = 2a_{12}y_1^2 - 2a_{12}y_2^2$, то есть в форме $A(x, x)$ появятся квадраты сразу двух переменных.

Итак, станем считать, что в равенстве (2) $a_{11} \neq 0$. Выделим в выражении (2) группу слагаемых, которые содержат x_1. Получим:

$$A(x, x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{1n}x_1x_n + \sum_{i,j=2}^{n} a_{ij}x_ix_j. \qquad (4)$$

Преобразуем выделенную сумму к виду:

$$A(x, x) = a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \ldots + \frac{a_{1n}}{a_{11}}x_n\right)^2 + \sum_{i,j=2}^{n} a_{ij}^{*}x_ix_j, \qquad (5)$$

при этом коэффициенты a_{ij} меняются на a_{ij}^{*}. Рассмотрим невырожденное преобразование

$y_1 = x_1 + \dfrac{a_{12}}{a_{11}}x_2 + \ldots + \dfrac{a_{1n}}{a_{11}}x_n$,

$y_2 = x_2$,

………

$y_n = x_n$.

Тогда получим

$$A(x, x) = a_{11}y_1^2 + \sum_{i,j=2}^{n} a_{ij}^{*}y_iy_j. \qquad (6).$$

Если квадратичная форма $\sum_{i,j=2}^{n} a_{ij}^{*}y_iy_j = 0$, то вопрос о приведении $A(x, x)$ к каноническому виду решен.

Если эта форма не равна нулю, то повторяем рассуждения, рассматривая преобразования координат $y_2, \ldots, y_n$ и не меняя при этом координату y_1. Очевидно, что эти преобразования будут невырожденными.

За конечное число шагов квадратичная форма $A(x, x)$ будет приведена к каноническому виду (3).

Замечание 1. Нужное преобразование исходных координат $x_1, x_2, \ldots, x_n$ можно получить путем перемножения найденных в процессе рассуждений невырожденных преобразований: $[x] = A[y]$, $[y] = B[z]$, $[z] = C[t]$, тогда $[x] = A\cdot B\cdot[z] = A\cdot B\cdot C\cdot[t]$, то есть $[x] = M\cdot[t]$, где $M = A\cdot B\cdot C$.

Замечание 2. Пусть $A(x, x) = A(x, x) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_r y_r^2$, где $\lambda_i \neq 0$, $i = 1, 2, \ldots, r$, причем $\lambda_1 > 0, \lambda_2 > 0, \ldots, \lambda_q > 0, \lambda_{q+1} < 0, \ldots, \lambda_r < 0$.

Рассмотрим невырожденное преобразование

$$y_1 = \frac{1}{\sqrt{\lambda_1}} z_1, \quad y_2 = \frac{1}{\sqrt{\lambda_2}} z_2, \quad \ldots, \quad y_q = \frac{1}{\sqrt{\lambda_q}} z_q, \quad y_{q+1} = \frac{1}{\sqrt{-\lambda_{q+1}}} z_{q+1}, \quad \ldots,$$

$y_r = \frac{1}{\sqrt{-\lambda_r}} z_r$, $y_{r+1} = z_{r+1}$, $\ldots$, $y_n = z_n$. В результате $A(x, x)$ примет вид: $A(x, x) = z_1^2 + z_2^2 + \ldots + z_q^2 - z_{q+1}^2 - \ldots - z_r^2$, который называется *нормальным видом квадратичной формы*.

Пример 11.1. Привести к каноническому виду квадратичную форму $A(x, x) = 2x_1x_2 - 6x_2x_3 + 2x_3x_1$.

Решение. Поскольку $a_{11} = 0$, используем преобразование

$$x_1 = y_1 - y_2,$$

$$x_2 = y_1 + y_2,$$

$$x_3 = y_3.$$

Это преобразование имеет матрицу $A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, то есть $[x] = A[y]$

получим $A(x, x) = 2(y_1 - y_2)(y_1 + y_2) - 6(y_1 + y_2)y_3 + 2y_3(y_1 - y_2) =$

$= 2y_1^2 - 2y_2^2 - 6y_1y_3 - 6y_2y_3 + 2y_3y_1 - 2y_3y_2 = 2y_1^2 - 2y_2^2 - 4y_1y_3 - 8y_3y_2$.

Поскольку коэффициент при y_1^2 не равен нулю, можно выделить квадрат одного неизвестного, пусть это будет y_1. Выделим все члены, содержащие y_1.

$A(x,\ \ x) = 2(y_1^2 - 2\ \ y_1y_3) - 2y_2^2 - 8y_3y_2 = 2(y_1^2 - 2\ \ y_1y_3 + y_3^2) - 2y_3^2 - 2y_2^2 -$ $8y_3y_2 = 2(y_1 - y_3)^2 - 2y_3^2 - 2y_2^2 - 8y_3y_2.$

Выполним преобразование, матрица которого равна B.

$z_1 = y_1 - y_3, \quad \Rightarrow \quad y_1 = z_1 + z_3,$

$z_2 = y_2, \quad \Rightarrow \quad y_2 = z_2,$

$z_3 = y_3; \quad \Rightarrow \quad y_3 = z_3.$

$$B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ [y] = B\cdot[z].$$

Получим $A(x, x) = 2z_1^2 - 2z_3^2 - z_2^2 - 8z_2z_3$. Выделим члены, содержащие z_2. Имеем $A(x, x) = 2z_1^2 - 2(z_2^2 + 4z_2z_3) - 2z_3^2 = 2z_1^2 - 2(z_2^2 + 4z_2z_3 + 4z_3^2) +$ $+ 8z_3^2 - 2z_3^2 = 2z_1^2 - 2(z_2 + 2z_3)^2 + 6z_3^2$.

Выполняем преобразование с матрицей C:

$t_1 = z_1, \quad \Rightarrow \quad z_1 = t_1,$

$t_2 = z_2 + 2z_3, \quad \Rightarrow \quad z_2 = t_2 - 2t_3,$

$t_3 = z_3; \quad \Rightarrow \quad z_3 = t_3.$

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}, \ [z] = C\cdot[t].$$

Получили: $A(x,\ x) = 2t_1^2 - 2t_2^2 + 6t_3^2$ канонический вид квадратичной формы, при этом $[x] = A[y]$, $[y] = B[z]$, $[z] = C[t]$, отсюда $[x] = ABC[t]$;

$$A\cdot B\cdot C = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\cdot\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\cdot\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 3 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$ Формулы преобразований следующие

$x_1 = t_1 - t_2 + t_3,$

$x_2 = t_1 + t_2 - t_3,$

$x_3 = t_3.$

Закон инерции квадратичных форм

Установлено, что число отличных от нуля канонических коэффициентов квадратичной формы равно ее рангу и не зависит от выбора невырожденного преобразования, с помощью которого форма $A(x, x)$ приводится к каноническому виду. На самом деле не меняется и число положительных и отрицательных коэффициентов.

Теорема 11.3 (закон инерции квадратичных форм). Число положительных и отрицательных коэффициентов в нормальном виде квадратичной формы не зависит от способа приведения квадратичной формы к нормальному виду.

Пусть квадратичная форма f ранга r от n неизвестных $x_1, x_2, \ldots, x_n$ двумя способами приведена к нормальному виду, то есть

$f = y_1^2 + y_2^2 + \ldots + y_k^2 - y_{k+1}^2 - \ldots - y_r^2$,

$f = z_1^2 + z_2^2 + \ldots + z_l^2 - z_{l+1}^2 - \ldots - z_r^2$. Можно доказать, что $k = l$.

Определение 11.14. Число положительных квадратов в нормальной форме, к которой приводится действительная квадратичная форма, называется *положительным индексом инерции* этой формы; число отрицательных квадратов – *отрицательным индексом инерции*, а их сумма – *индексом инерции* квадратичной формы или *сигнатурой* формы f.

Если p – положительный индекс инерции; q – отрицательный индекс инерции; $k = r = p + q$ – индекс инерции.

Классификация квадратичных форм

Пусть у квадратичной формы $A(x, x)$ индекс инерции равен k, положительный индекс инерции равен р , отрицательный индекс инерции равен q, тогда $k = p + q$.

Было доказано, что в любом каноническом базисе $f = \{f_1, f_2, \ldots, f_n\}$ эта квадратичная форма $A(x, x)$ может быть приведена к нормальному виду $A(x, x) = \eta_1^2 + \eta_2^2 + \ldots + \eta_p^2 - \eta_{p+1}^2 - \ldots - \eta_k^2$, где $\eta_1, \eta_2, \ldots, \eta_n$ координаты вектора x в базисе $\{f\}$.

Необходимое и достаточное условие знакоопределенности квадратичной формы

Утверждение 11.1. Для того чтобы квадратичная форма $A(x, x)$, заданная в n-мерном векторном пространстве V, была *знакоопределенной,* необходимо и достаточно, чтобы либо положительный индекс инерции p, либо отрицательный индекс инерции q, был равен размерности n пространства V.

При этом если $p = n$, то форма *положительно* определена (то есть для любого $x \neq 0$ $A(x, x) > 0$).

Если же $q = n$, то форма *отрицательно* определена (то есть для любого $x \neq 0$ $A(x, x) < 0$).

Необходимое и достаточное условие знакопеременности квадратичной формы

Утверждение 11.2. Для того чтобы квадратичная форма $A(x, x)$, заданная в n-мерном векторном пространстве V, была *знакопеременной* (то есть существуют такие x, y что $A(x, x) > 0$ и $A(y, y) < 0$) необходимо и достаточно, чтобы как положительный, так и отрицательный индексы инерции этой формы были отличны от нуля.

Необходимое и достаточное условие квазизнакопеременности квадратичной формы

Утверждение 11.3. Для того чтобы квадратичная форма $A(x, x)$, заданная в n-мерном векторном пространстве V, была *квазизнакопеременной* (то есть для любого вектора x или $A(x, x) \geq 0$ или $A(x, x) \leq 0$ и найдется такой ненулевой вектор x, что $A(x, x) = 0$) необходимо и достаточно, чтобы выполнялось одно из двух соотношений: $p < n$, $q = 0$ или $p = 0$, $q < n$.

Замечание. Для того чтобы применять эти признаки, квадратичную форму надо привести к каноническому виду. В критерии знакоопределенности Сильвестра[15] этого не требуется.

Критерий Сильвестра знакоопределенности квадратичной формы

Пусть форма $A(x, x)$ в базисе $e = \{e_1, e_2, \ldots, e_n\}$ определяется матрицей $A(e) = (a_{ij})$,

$A(x, x) = \sum_{i,j=1}^{n} a_{ij}x_i x_j$, и пусть $\Delta_1 = a_{11}$, $\Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, …,

$\Delta_n = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$ угловые миноры и определители матрицы (a_{ij}).

Тогда справедливо утверждение:

Теорема 11.4 (критерий Сильвестра).

1. Для того чтобы квадратичная форма $A(x, x)$ была положительно определенной, необходимо и достаточно, чтобы были выполнены неравенства: $\Delta_1 > 0$, $\Delta_2 > 0$, …, $\Delta_n > 0$.
2. Для того чтобы квадратичная форма $A(x, x)$ была отрицательно определенной, необходимо и достаточно, чтобы знаки угловых миноров чередовались, причем $\Delta_1 < 0$.

Пример 11.2. Выяснить, является ли квадратичная форма $A(x, x) = 5x_1^2 + x_2^2 + 5x_3 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3$ положительно определенной?

Решение. Составим матрицу этой квадратичной формы:

$$M = \begin{pmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{pmatrix}$$

Вычислим ее угловые миноры:

[15] Джеймс Джозеф Сильвестр (1814–1897) – английский математик.

$\Delta_1 = 5 > 0,$

$$\Delta_2 = \begin{vmatrix} 5 & 2 \\ 2 & 1 \end{vmatrix} = 5 - 4 = 1 > 0,$$

$$\Delta_3 = \begin{vmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{vmatrix} = 25 + 16 + 16 - 16 - 20 - 20 = 1 > 0.$$

Все угловые миноры положительны, следовательно, квадратичная форма положительно определенна.

Пример 11.3. Выяснить, является ли квадратичная форма $A(x, x) = 3x_1^2 + x_2^2 + 5x_3 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3$ положительно определенной?

Решение. Составим матрицу этой квадратичной формы:

$$M = \begin{pmatrix} 3 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{pmatrix}$$

Вычислим ее угловые миноры:

$\Delta_1 = 3 > 0,$

$$\Delta_2 = \begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} = 3 - 4 = -1 < 0,$$

$$\Delta_3 = \begin{vmatrix} 3 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{vmatrix} = 15 + 16 + 16 - 16 - 12 - 20 = -1 < 0.$$

Вывод. Квадратичная форма $A(x, x)$ не является положительно определенной, т. к. $\Delta_2 < 0$, и отрицательно определенной не является, т. к. $\Delta_1 > 0$, т.о. она знакопеременная.

ЗАКЛЮЧЕНИЕ

Линейная алгебра является обязательной частью любой программы по высшей математике. Любой другой раздел предполагает наличие знаний, умений и навыков, заложенных во время преподавания этой дисциплины. Данное пособие в доступной, но математически строгой форме, дает возможность освоить все основные понятия и ознакомить с типовыми упражнениями.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

Бурмистрова Е.Б., Лобанов С.Г. Линейная алгебра с элементами аналитической геометрии. – М.: Изд-во ВШЭ, 2007.

Беклемишев Д.В. Курс аналитической геометрии и линейной алгебры. – М.: Наука, 2006.

Данко П.Е., Попов А.Г., Кожевникова Т.Я. Высшая математика в упражнениях и задачах. Часть I. – М.: ОНИКС 21 век, Мир и образование, 2003.

Ермаков В.И. Сборник задач по высшей математике для экономистов. – М.: Инфра - М, 2009.

Ермаков В.И. Общий курс высшей математики для экономистов. – М.: Инфра - М, 1999.

Ильин В.А., Позняк Э.Г. Линейная алгебра. – М.: Наука, 2002.

Кремер Н.Ш. Высшая математика для экономистов. – М.: Юнити, 2004.

Кудрявцев В.А., Демидович Б.П. Краткий курс высшей математики. – М.: Наука, 1989.

Крамор В. С. Алгебра и начало анализа / В. С. Крамор. – М.: Высш. школа, 1981.

Куликов Л. Я. Алгебра и теория чисел: учеб. пособие для педагогических институтов / Л. Я. Куликов. – М.: Высш. школа, 1979.

Куликов Л. Я. Сборник задач по алгебре и теории чисел / Л. Я. Куликов, А. И. Москаленко, А. А. Фомин. – М.: Просвещение, 1993.

Лунгу К. Н. Сборник задач по высшей математике / К. Н. Лунгу. – М.: Айрис-пресс, 2007.

Малыкин В.И. Математика в экономике. – М.: Инфра - М, 2002.

Солодовников А.С., Бабайцев В.А., Браилов А.В. Математика в экономике. Часть I. – М.: Финансы и кредит, 2000.

Фадеев Д.К., Соминский И.С. Задачи по высшей алгебре. – СПб.: Лань, 2001.

Чередникова А. В. Дискретная математика. Теория и практика / А. В. Чередникова, О. Б. Садовская, Л. А. Каминская. – Кострома: Изд-во Костром. гос. технол. ун-та, 2011.

Шапорев С. Д. Дискретная математика. Курс лекций и практических занятий / С. Д. Шапорев. – СПб.: БХВ-Петербург, 2007. – 400с.

Шепелев Ю. П. Дискретная математика: учеб. пособие / Ю. П. Шепелев. – СПб.: Лань, 2008.

Математический энциклопедический словарь / гл. ред. Ю. В. Прохоров; ред. кол.: С. И. Адян, Н. С. Бахвалов, В. И. Битюцков, А. П. Ершов, Л. Д. Кудрявцев, А. Л. Онищик, А. П. Юшкевич. – М.: Сов. энциклопедия, 1988.

Учебное издание

ЛИНЕЙНАЯ АЛГЕБРА

Учебно-методическое пособие

Уч.-изд. л. 4,0.