

С.К. Росошек

И АЛГЕБРА ГЕОМЕТРИЯ

УЧЕБНОЕ ПОСОБИЕ

Часть I

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ



ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

С.К. Росошек

АЛГЕБРА И ГЕОМЕТРИЯ

Часть 1

Учебное пособие



Томск – 2008

УДК 512.64+514.12(075.8)
ББК 22.143я73+22.151.5я73
Р 757

Р 757 **Росошек С.К.** Алгебра и геометрия. Часть 1: Учебное пособие.
– Томск: Изд-во НТЛ, 2008. – 196 с.

ISBN 978-5-89503-381-4

В первой части учебного пособия «Алгебра и геометрия» наряду со стандартным набором тем линейной алгебры (матрицы, определители, системы линейных уравнений) дается введение в современную алгебру (группы, кольца, поля и многочлены), необходимое для специалистов по информационным технологиям.

Для студентов вузов, обучающихся по специальностям, относящимся к информационным технологиям и обеспечению их безопасности.

УДК 512.64+514.12(075.8)
ББК 22.143я73+22.151.5я73

Рецензент: доктор физ.-мат. наук, профессор **П.А. Крылов**

ISBN 978-5-89503-381-4

© Томский госуниверситет, 2008
© С.К. Росошек, 2008

Предисловие

Данное учебное пособие предназначено для факультетов информатики или информационных технологий университетов, в частности тех, которые начали обучение по программе бакалавриата.

Структура курса следует из методики контроля знаний, принятой на факультете информатики Томского государственного университета, согласно которой после каждых четырех недель занятий проводится мини-сессия на пятой неделе. Поэтому курс состоит из четырех модулей по четыре лекции каждый. Так как вопросы по четвертому модулю и выборочно по остальным трем модулям выносятся на экзаменационную сессию, а по результатам трех мини-сессий учитывается основное содержание первых трех модулей, то изменилась последовательность изложения материала. В первых трех модулях собран материал, существенной частью которого являются те или иные вычисления. А именно, в модуле I – это матрицы и определители, в модуле III – это векторные пространства, в модуле II – это системы линейных уравнений. Тот же материал, который обычно вызывал серьезные затруднения у студентов, когда он читался в начале курса, теперь помещен в четвертый модуль (речь идет об алгебраических структурах). Более того, объем лекций этого модуля существенно увеличен по сравнению с лекциями других модулей для свободы маневра лекторов и увеличения самостоятельной работы студентов, а также для организации проектной деятельности.

Следует заметить, что знакомство специалистов по информационным технологиям с алгебраическими структурами становится все более востребованным по мере перехода абстрактных алгебраических понятий в практическую плоскость. Например, математической основой наиболее часто применяемых на практике криптосистем с открытым ключом, таких, как криптосистемы RSA, Эль-Гамала и Рабина, являются кольца вычетов. Далее, в новом стандарте США шифрования данных AES, принятом в 2001 г., ключевую роль играет математический аппарат конечных полей, а в стандартах России и США электронной цифровой подписи используется математический аппарат групп точек эллиптических кривых над конечным полем и их циклических подгрупп.

МОДУЛЬ I

Лекция 1. МНОЖЕСТВА, ОТОБРАЖЕНИЯ И ПОДСТАНОВКИ

Основатель теории множеств Георг Кантор дал следующее интуитивное определение понятия множества: «множество есть нечто, мыслимое как единое». Расширительное толкование определения Кантора приводит к следующему интуитивному определению, которым мы будем пользоваться в дальнейшем.

Определение I.1.1. Совокупность различных объектов, объединенных некоторым общим признаком, называется **множеством**, а сами объекты – **элементами** множества.

Множества принято обозначать прописными буквами некоторого алфавита, обычно латинского: $A, B, C, \dots, X, Y, Z$. Элементы множеств принято обозначать строчными буквами некоторого алфавита, обычно латинского: $a, b, c, \dots, x, y, z$. В записи множеств часто используют фигурные скобки, внутри которых указываются элементы множества, например: $A = \{a, b, c\}$. Для данного множества M запись $t \in M$ означает, что t есть элемент множества M (читается « t принадлежит M »). Соответственно запись $t \notin M$ означает, что t не является элементом множества M (читается « t не принадлежит M »).

Пример I.1.2. Рассмотрим множество $A = \{1, 1, 2\}$. Это множество имеет два различных элемента.

Определение I.1.3. Будем говорить, что множество A есть **подмножество** множества B (обозначение $A \subseteq B$), если для любого элемента $a \in A$ имеет место $a \in B$. Если существует хотя бы один элемент $b \in B$, который не принадлежит A (т.е. $b \notin A$), то будем говорить, что A есть **собственное подмножество** множества B (обозначается $A \subset B$).

Запись $A \subseteq B$ (соответственно $A \subset B$) читается так: «множество A содержится во множестве B » (соответственно «множество A строго содержится во множестве B »). Иногда используется и следующая терми-

нология: запись $A \subseteq B$ читается как «включение A в B » (соответственно запись $A \subset B$ читается как «строгое включение A в B »).

Пример 1.1.4. Пусть A есть множество всех десятичных цифр, т.е. $A = \{0, 1, 2, \dots, 8, 9\}$, а B есть множество всех натуральных чисел, не превосходящих 8, т.е. $B = \{1, 2, 3, \dots, 7, 8\}$. Поскольку каждый элемент множества B является также и элементом множества A , то $B \subseteq A$. Поскольку существует элемент множества A , который не является элементом множества B (например, число 0), то имеет место строгое включение $B \subset A$, т.е. $B \subset A$.

Определение 1.1.5. Будем говорить, что множества A и B равны (обозначение $A = B$), если $A \subseteq B$ и $B \subseteq A$.

Пример 1.1.6. Пусть $A = \{0, 2, 4, 6, 8\}$ и B есть множество всех неотрицательных четных чисел, меньших 10. Тогда для любого $a \in A$ имеем также $a \in B$ и для любого $b \in B$ имеем также $b \in A$. Следовательно, имеем $A \subseteq B$ и $B \subseteq A$, откуда получаем $A = B$.

Обычно, если в некотором рассуждении идет речь о множествах A , B , C , ..., то предполагается, что существует некоторое универсальное множество U , подмножествами которого являются множества A , B , C , ... Множество U принято называть **универсумом** рассуждения.

Заметим, что U не является множеством всех множеств и его использование не приводит к противоречию.

Определение 1.1.7. Множество, которое не имеет ни одного элемента, называется **пустым множеством**.

Пустое множество обычно обозначается $\emptyset$.

Пример 1.1.8. Пусть M есть множество всех натуральных чисел, каждое из которых меньше нуля. Тогда очевидно, что $M = \emptyset$.

Рассмотрим теперь способы задания множеств.

1. **Задание множества списком.** Пример: множество букв алфавита, множество студентов группы.

2. **Задание множества характеристическим свойством.** Пример: множество B из примера 1.1.6.

3. **Задание множества посредством порождающей процедуры.** Пример: множество чисел Фибоначчи. Действительно, порождающей процедурой для чисел Фибоначчи является получение каждого следующего числа посредством суммирования двух предыдущих чисел.

В дальнейшем будут использоваться стандартные обозначения для следующих множеств. Множество всех натуральных чисел

$\mathbb{N} = \{1, 2, 3, \dots\}$, множество всех неотрицательных целых чисел $\mathbb{N}_0 = \{0, 1, 2, \dots\}$, множество всех целых чисел $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$, множество всех рациональных чисел $\mathbb{Q} = \{m/n | n \in \mathbb{N}, m \in \mathbb{Z}\}$, множество всех ненулевых рациональных чисел $\mathbb{Q}^* = \{m/n | n \in \mathbb{N}, 0 \neq m \in \mathbb{Z}\}$, множество всех действительных чисел $\mathbb{R}$ и множество всех ненулевых действительных чисел $\mathbb{R}^*$. Мощностью множества M называется число элементов множества M , обозначается мощность M так: $|M|$. Мощность множества $\mathbb{N}$ называется счетной мощностью, и все множества, мощность которых равна счетной мощности, называются счетными множествами. Примеры счетных множеств:

$$\mathbb{N}, \mathbb{N}_0, \mathbb{Z}, \mathbb{Q}, \mathbb{Q}^*.$$

Бесконечное множество, мощность которого не является счетной, называется несчетным множеством. Примеры несчетных множеств: $\mathbb{R}, \mathbb{R}^*$.

Имеет место следующая

Теорема I.1.9. Мощность множества $P(M)$ всех подмножеств конечного множества M равна $|P(M)| = 2^{|M|}$.

Доказательство. Используем метод математической индукции.

Пусть $|M| = 1$, т.е. $M = \{m_1\}$. Тогда $P(M) = \{\emptyset, M\}$, откуда $|P(M)| = 2 = 2^1 = 2^{|M|}$.

Пусть утверждение теоремы верно для всех множеств мощности $\leq k$, и рассмотрим множество M мощности $k + 1$: $M = \{m_1, \dots, m_{k+1}\}$. Обозначим через $K = \{m_1, \dots, m_k\}$, тогда по предположению индукции $|P(K)| = 2^k$. Заметим, что для любого подмножества L множества M имеют место две возможности:

- 1) $m_{k+1} \notin L$, и тогда L есть подмножество множества K ;
- 2) $m_{k+1} \in L$, и тогда L состоит из элементов некоторого подмножества K' множества K , к которым присоединен элемент m_{k+1} .

Но тогда число подмножеств вида «1» равно 2^k и число подмножеств вида «2» тоже равно 2^k , поскольку для каждого подмножества K' вида «1» существует единственное подмножество вида «2», полученное из K' присоединением m_{k+1} . Поскольку любое подмножество множества M имеет либо вид «1», либо вид «2», причем не существует под-

множества множества M одновременно одного и другого вида, то общее число подмножеств множества M равно сумме чисел подмножеств вида «1» и «2». Таким образом, $|P(M)| = 2^k + 2^k = 2 \cdot 2^k = 2^{k+1} = 2^{|M|}$.

Определение I.1.10. Декартовым произведением множеств A и B называется множество всех упорядоченных пар вида (a, b) , где $a \in A$, $b \in B$.

Декартово произведение множеств A и B принято обозначать $A \times B$.

Пример I.1.11. Если $A = B = \mathbb{R}$, то $A \times B = \mathbb{R} \times \mathbb{R}$.

Понятие декартова произведения естественным образом обобщается на случай любого конечного числа множеств.

Определение I.1.12. Декартовым произведением множеств $A_1, A_2, \dots, A_n$ называется множество всех упорядоченных последовательностей вида $(a_1, a_2, \dots, a_n)$, $a_i \in A_i$, $i = \overline{1, n}$.

В частном случае $A_1, A_2, \dots, A_n = A$ декартово произведение $\underbrace{A \times A \times \dots \times A}_{n \text{ раз}}$ принято называть n -й декартовой степенью множества A и обозначать ее A^n .

Теорема I.1.13. Для любого множества A и любого натурального числа n имеет место равенство

$$|A^n| = |A|^n.$$

Доказательство получается методом математической индукции, и его проведение предоставляется читателю.

Рассмотрим теперь операции над множествами.

Определение I.1.14. Пусть A, B — произвольные множества, U — универсум, тогда

а) назовем **пересечением** множеств A и B следующее множество:

$$A \cap B = \{x | x \in A \text{ и } x \in B\};$$

б) назовем **объединением** множеств A и B следующее множество:

$$A \cup B = \{x | x \in A \text{ или } x \in B\};$$

в) назовем **разностью** множеств A и B следующее множество:

$$A \setminus B = \{x | x \in A \text{ и } x \notin B\};$$

г) назовем **симметрической разностью** (по другой терминологии — суммой) множеств A и B следующее множество:

$$A + B = (A \setminus B) \cup (B \setminus A);$$

д) назовем (абсолютным) дополнением множества A следующее множество:

$$\bar{A} = U \setminus A.$$

Теорема I.1.15. Для произвольных множеств A , B и C имеют место следующие тождества:

- а) $A \cap B = B \cap A$;
- б) $A \cap (B \cap C) = (A \cap B) \cap C$;
- в) $A \cup B = B \cup A$;
- г) $A \cup (B \cup C) = (A \cup B) \cup C$;
- д) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;
- е) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;
- ж) $A \cup A = A$, $A \cup U = U$;
- з) $A \cap A = A$, $A \cap U = A$;
- и) $A \cup \emptyset = A$; $A \cap \emptyset = \emptyset$;
- к) $A \cup \bar{A} = U$, $A \cap \bar{A} = \emptyset$;
- л) $\overline{A \cup B} = \bar{A} \cap \bar{B}$, $\overline{A \cap B} = \bar{A} \cup \bar{B}$;
- м) $A \cup (A \cap B) = A$, $A \cap (A \cup B) = A$.

Докажем одно из этих тождеств, например «д», остальные доказываются аналогично.

Доказательство «д».

Покажем сначала, что левая часть содержится в правой части. Пусть $x \in A \cup (B \cap C)$, тогда $x \in A$ или $x \in B \cap C$. Отсюда ($x \in A$ или $x \in B$) и ($x \in A$ или $x \in C$). Так как $x \in A$ или $x \in B$ означает $x \in A \cup B$, $B \subseteq A \subseteq C$ или $x \in C$ означает $x \in A \cup C$, то ($x \in A$ или $x \in B$) и ($x \in A$ или $x \in C$) означает $x \in [(A \cup B) \cap (A \cup C)]$.

Теперь покажем, что правая часть содержится в левой части. Пусть $x \in [(A \cup B) \cap (A \cup C)]$, тогда $x \in A \cup B$ и $x \in A \cup C$. Отсюда ($x \in A$ или $x \in B$) и ($x \in A$ или $x \in C$). Но тогда $x \in A$ или ($x \in B$ и $x \in C$), что

означает $x \in [A \cup (B \cap C)]$. Мы показали включения в обе стороны, следовательно, доказано равенство.

Определение I.1.16. Будем говорить, что задано отображение f с областью определения X и областью значений Y , если каждому $x \in X$ ставится в соответствие однозначно определенный $y \in Y$, который обозначается $y = f(x)$. Элемент y называется образом элемента x при действии отображения f , а элемент x называется прообразом элемента y при действии отображения f .

Символическая запись отображения f выглядит так:

$$f: X \rightarrow Y,$$

а действие f на элементах множества X записывается в виде

$$f: x \mapsto y = f(x).$$

Заметим, что соответствие f из множества X в множество Y является отображением, если выполняются следующие условия:

- 1) всюду определенность в X ;
- 2) замкнутость в Y ;
- 3) однозначность.

Пример I.1.17.

а) соответствие $f: \mathbb{N} \rightarrow \mathbb{N}$, $f: x \mapsto x^2$ является отображением из $\mathbb{N}$ в $\mathbb{N}$;

б) соответствие $f: \mathbb{N} \rightarrow \mathbb{N}$, $f: x \mapsto \sqrt{x}$ отображением не является, так как нарушено условие замкнутости в $\mathbb{N}$;

в) соответствие $f: \mathbb{N} \rightarrow \mathbb{R}$, $f: x \mapsto \sqrt{x}$ является отображением;

г) соответствие $f: \mathbb{N} \rightarrow \mathbb{R}$, $f: x \mapsto \pm\sqrt{x}$ не является отображением, так как нарушено условие однозначности;

д) соответствие $f: \mathbb{Z} \rightarrow \mathbb{Q}$, $f: x \mapsto \frac{1}{x}$ не является отображением,

так как нарушено условие всюду определенности ($f(0)$ не определено).

Определение I.1.18. Пусть $f: X \rightarrow Y$ – отображение. Множество $\text{Im } f = \{y \in Y \mid \exists x \in X \text{ такой, что } y = f(x)\}$ называется образом отображения f .

Очевидно, что $\text{Im } f \subseteq Y$.

Определение I.1.19. Пусть $f: X \rightarrow Y$ – отображение. Множество $\Gamma(f) = \{(x, y) \in X \times Y \mid \forall x \in X \text{ выполнено } y = f(x)\}$ называется **графиком отображения** f .

Очевидно, что $\Gamma(f) \subseteq X \times Y$.

Пример I.1.20.

а) для отображения $f: \mathbb{R} \rightarrow \mathbb{R}$, $f: x \mapsto x^2$ имеем

$$\text{Im } f = [0, \infty), \Gamma(f) \text{ есть парабола } y = x^2;$$

б) для отображения $f: \mathbb{N} \rightarrow \mathbb{R}$, $f: x \mapsto \sqrt{x}$ имеем

$$\text{Im } f = \{1, \sqrt{2}, \sqrt{3}, \dots, \sqrt{n}, \dots\},$$

$\Gamma(f)$ есть множество точек на кривой $y = \sqrt{x}$ вида $(n, \sqrt{n})$ для $n \in \mathbb{N}$.

Определение I.1.21. Будем говорить, что два отображения $f: X \rightarrow Y$ и $g: X' \rightarrow Y'$ равны, если $X = X'$, $Y = Y'$, и для любого $x \in X$ имеем $f(x) = g(x)$.

Определение I.1.22. Отображение $f: X \rightarrow Y$ называется **сюръективным**, если для любого $y \in Y$ существует $x \in X$ такой, что $y = f(x)$.

Другими словами, отображение f сюръективно тогда и только тогда, когда $\text{Im } f = Y$.

Пример I.1.23.

а) отображение $f: \mathbb{N} \rightarrow \mathbb{N}$, $f: x \mapsto x^2$ не является сюръективным отображением, так как, например, $2 \notin \text{Im } f$;

б) отображение $f: \mathbb{R} \rightarrow \mathbb{R}$, $f: x \mapsto x^2$ не является сюръективным отображением, так как, например, $-1 \notin \text{Im } f$;

в) отображение $f: \mathbb{R} \rightarrow [0, \infty)$, $f: x \mapsto x^2$ является сюръективным отображением, так как для любого $y \in [0, \infty)$ существует $x \in \mathbb{R}$ (а именно $x = \sqrt{y}$) такой, что $y = f(x)$ ($f(x) = (\sqrt{y})^2 = y$).

Определение I.1.24. Отображение $f: X \rightarrow Y$ называется **инъективным**, если для любых $x, x' \in X$ из того, что $x \neq x'$ следует $f(x) \neq f(x')$.

Пример I.1.25. Отображение $f: \mathbb{Z} \rightarrow \mathbb{Q}$, $f: n \rightarrow \frac{n}{1}$ является инъективным отображением, так как для любых целых чисел n и m , если $n \neq m$, то $\frac{n}{1} \neq \frac{m}{1}$.

Предложение I.1.26. Отображение $f: X \rightarrow Y$ является инъективным тогда и только тогда, когда для любых $x, x' \in X$ из того, что $f(x) = f(x')$, следует $x = x'$.

Доказательство очевидно.

Примером сюръективного, но не инъективного отображения является отображение из примера I.1.23 «в», а примером инъективного, но не сюръективного отображения является отображение из примера I.1.25.

Определение I.1.27. Отображение $f: X \rightarrow Y$ называется биективным (или биекцией), если f является сюръективным и инъективным отображением.

Предложение I.1.28. Пусть $f_{a,b}: \mathbb{R} \rightarrow \mathbb{R}$, $f_{a,b}: x \mapsto ax + b$ – отображения, зависящие от действительных чисел a и b , как от параметров. Отображение $f_{a,b}$ является биекцией тогда и только тогда, когда $a \neq 0$.

Доказательство. Необходимость очевидна, так как если для биекции $f_{a,b}$ имеет место $a = 0$, то, так как $f_{a,b}: x \mapsto b$ не является биекцией, получаем противоречие.

Достаточность. Пусть $a \neq 0$, и покажем, что $f_{a,b}$ есть биекция. Для доказательства инъективности $f_{a,b}$ пусть x_1, x_2 – произвольные действительные числа, такие, что $f_{a,b}(x_1) = f_{a,b}(x_2)$. Это означает, что $ax_1 + b = ax_2 + b$. Отсюда следует, что $ax_1 = ax_2$, откуда ввиду $a \neq 0$ получаем $x_1 = x_2$. По предложению I.1.26 имеем инъективность отображения $f_{a,b}$. Покажем далее сюръективность отображения $f_{a,b}$. Пусть $y \in \mathbb{R}$ – произвольное число, нужно показать, что существует $x \in \mathbb{R}$ такой, что $y = f_{a,b}(x)$, т.е. $y = ax + b$. Очевидно, что в качестве x можно взять $x = \frac{y-b}{a}$. Таким образом, $f_{a,b}$ является инъективным и сюръективным отображением и, следовательно, $f_{a,b}$ – биекция.

Определение I.1.29. Для любого множества X определим отображение $I_X : X \rightarrow X$ по формуле $I_X(x) = x$ для любого $x \in X$. Отображение I_X называется тождественным отображением на множестве X .

Определение I.1.30. Композицией отображений $f : U \rightarrow V$ и $g : V \rightarrow W$ называется отображение $g \circ f : U \rightarrow W$, которое действует по формуле

$$(g \circ f)(u) = g(f(u)) \text{ для любого } u \in U.$$

Наглядно композиция $g \circ f$ представляется следующей диаграммой (рис. 1):

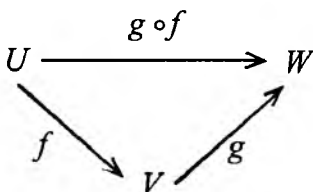


Рис. 1

Заметим, что для любых трех отображений вида $f : U \rightarrow V$, $g : V \rightarrow W$ и $h : W \rightarrow T$ имеет место равенство

$$(h \circ g) \circ f = h \circ (g \circ f).$$

Определение I.1.31. Пусть $f : X \rightarrow Y$ и $g : Y \rightarrow X$ – отображения. Если $g \circ f = I_X$ и $f \circ g = I_Y$, то отображение g называется обратным отображением к f (соответственно f называется обратным к g).

Обозначения: $g = f^{-1}$ ($f = g^{-1}$).

Пример I.1.32. Пусть $f_{a,b} : \mathbb{R} \rightarrow \mathbb{R}$, $f_{a,b} : x \mapsto ax + b$ – отображения из предложения I.1.28. Тогда обратным к $f_{a,b}$ является отображение

$f_{a,b}^{-1} : \mathbb{R} \rightarrow \mathbb{R}$, $f_{a,b}^{-1} : x \mapsto \frac{1}{a}(x - b)$. Действительно,

$$\begin{aligned} (f_{a,b} \circ f_{a,b}^{-1})(x) &= f_{a,b}(f_{a,b}^{-1}(x)) = f_{a,b}\left(\frac{1}{a}(x - b)\right) = \\ &= a\left(\frac{1}{a}(x - b)\right) + b = x - b + b = x, \end{aligned}$$

т.е. $f_{a,b} \circ f_{a,b}^{-1} = I_R$. С другой стороны,

$$\begin{aligned} (f_{a,b}^{-1} \circ f_{a,b})(x) &= f_{a,b}^{-1}(f_{a,b}(x)) = \\ &= f_{a,b}^{-1}(ax + b) = \frac{1}{a}((ax + b) - b) = x, \end{aligned}$$

т.е.

$$f_{a,b}^{-1} \circ f_{a,b} = I_R.$$

В примере I.1.32 обратное отображение найдено для биекции. То, что этот факт не случаен, показывает следующая теорема.

Теорема I.1.33. Отображение $f: X \rightarrow Y$ имеет обратное отображение $f^{-1}: Y \rightarrow X$ тогда и только тогда, когда f есть биекция.

Достаточность. Пусть f — биекция, тогда f — инъективное и сюръективное отображение. Пусть $y \in Y$ — произвольный элемент, тогда из сюръективности f вытекает то, что существует $x_0 \in X$ такой, что $y = f(x_0)$. Из инъективности f следует, что x_0 есть единственный элемент X со свойством $y = f(x_0)$. Действительно, пусть существует $x' \in X$ такой, что $f(x') = y$, тогда имеем $f(x_0) = f(x')$, и из инъективности f вытекает $x_0 = x'$. Построим соответствие $g: Y \rightarrow X$, $g: y \mapsto x_0$ для произвольного $y \in Y$ и единственного $x_0 \in X$ так, что $y = f(x_0)$, тогда g есть отображение из Y в X . Непосредственно проверяется, что $g \circ f = I_X$ и $f \circ g = I_Y$, откуда следует, что $g = f^{-1}$.

Необходимость. Пусть для отображения $f: X \rightarrow Y$ существует обратное отображение $f^{-1}: Y \rightarrow X$, тогда $f^{-1} \circ f = I_X$ и $f \circ f^{-1} = I_Y$. Проверим инъективность f . Пусть $x_1, x_2 \in X$ — произвольные элементы X такие, что $f(x_1) = f(x_2)$. Имеем

$$\begin{aligned} x_1 = I_X(x_1) &= (f^{-1} \circ f)(x_1) = f^{-1}(f(x_1)) = f^{-1}(f(x_2)) = \\ &= (f^{-1} \circ f)(x_2) = I_X(x_2) = x_2. \end{aligned}$$

Следовательно, f есть инъективное отображение. Проверим сюръективность f . Для любого $y \in Y$ имеем

$$y = I_Y(y) = (f \circ f^{-1})(y) = f(f^{-1}(y)).$$

Обозначим $x = f^{-1}(y) \in X$, тогда

$$f(x) = f(f^{-1}(y)) = (f \circ f^{-1})y = y.$$

Следовательно, y есть сюръективное отображение.

Имеет место следующая

Теорема I.1.34. Пусть X – конечное множество, $f: X \rightarrow X$ – отображение. Если f – инъективное отображение, то f есть биекция.

Доказательство. Достаточно убедиться в сюръективности отображения f . Пусть $y \in X$ – произвольный элемент, и рассмотрим следующую последовательность элементов множества X :

$$(*) f^{(0)}(y) = y, f^{(1)}(y) = f(y), f^{(2)}(y) = f(f(y)) \dots$$

Так как X – конечное множество, то для некоторых натуральных чисел $m \neq n$ выполняется $f^{(m)}(y) = f^{(n)}(y)$. Действительно, в противном случае последовательность $(*)$ бесконечна, что противоречит конечности множества X . Без ограничения общности можем считать, что $m > n$. Ввиду инъективности f , получаем $f^{(m-n)}(y) = f^{(n)}(y)$. Продолжая таким образом, получаем

$$f^{(m-n)}(y) = f^{(0)}(y) = y.$$

Тогда из $f(f^{(m-n-1)}(y)) = y$ следует, что существует $z \in X$ (а именно $z = f^{(m-n-1)}(y)$) так, что $f(z) = y$. Следовательно, f есть сюръективное отображение.

Определение I.1.35. Бинарным отношением между множествами X и Y называется произвольное подмножество декартового произведения X и Y , $\rho \subseteq X \times Y$.

Другими словами, любое бинарное отношение ρ между X и Y есть некоторое множество упорядоченных пар (x, y) , где $x \in X$, $y \in Y$. Тот факт, что $(x, y) \in \rho$, будем записывать как $x \rho y$ и читать – « x находится в отношении ρ к y ». Если $X = Y$, то будем говорить, что ρ есть бинарное отношение на множестве X .

Пример 1.1.36. Бинарное отношение « $<$ » на множестве $\mathbb{R}$ есть множество точек плоскости $M(x, y)$, координаты которых удовлетворяют соотношению $x < y$. Аналогично определяются бинарные отношения « $>$ » и « $=$ ». На рис. 2 даны эти бинарные отношения.

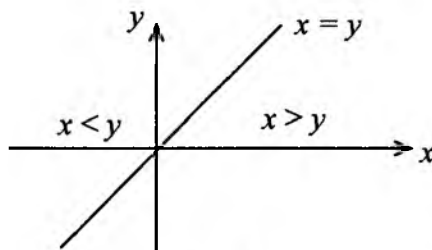


Рис. 2

Пример 1.1.37. Пусть $f: X \rightarrow Y$ – отображение. Тогда график $\Gamma(f)$ отображения f есть бинарное отношение между X и Y . Действительно,

$$\Gamma(f) = \{(x, y) \mid \forall x \in X, y = f(x)\} \subseteq X \times Y.$$

Определение 1.1.38. Бинарное отношение ρ на множестве X называется отношением эквивалентности, если выполнены следующие три условия для любых $x, y, z \in X$:

- 1) рефлексивность, т.е. $x\rho x$;
- 2) симметричность, т.е. если $x\rho y$, то $y\rho x$;
- 3) транзитивность, т.е. если $x\rho y$ и $y\rho z$, то $x\rho z$.

Другими словами, бинарное отношение ρ на множестве X есть отношение эквивалентности, если для $\forall x \in X$ имеем $(x, x) \in \rho$, для любых $x, y \in X$ имеем, что из $(y, x) \in \rho$ следует $(x, y) \in \rho$, и для любых $x, y, z \in X$ из того, что $(x, y) \in \rho$ и $(y, z) \in \rho$, следует $(x, z) \in \rho$.

Пример 1.1.39. Зададим на множестве $\mathbb{Z}$ бинарное отношение ρ следующим образом: для $x, y \in \mathbb{Z}$, если $x - y$ есть четное число, то $(x, y) \in \rho$, т.е. $x\rho y$. Проверим, что ρ есть отношение эквивалентности. Рефлексивность ρ следует из того, что для любого $x \in \mathbb{Z}$ из того, что

$x - x = 0$ есть четное число, получаем xrx . Симметричность ρ следует из того, что если для $x, y \in \mathbb{Z}$ имеем xry , то $x - y$ есть четное число и, следовательно, $y - x$ есть четное число, откуда вытекает $урх$. Транзитивность ρ следует из того, что если xry и $урz$ выполнены для $x, y, z \in \mathbb{Z}$, то $x - y$ и $y - z$ являются четными числами, но тогда $x - z = (x - y) + (y - z)$ есть четное число, откуда вытекает xrz .

Определение I.1.40. Пусть ρ есть отношение эквивалентности на множестве X . Для любого $x \in X$ обозначим $\bar{x} = \{y \in X \mid урх\}$ и назовем $\bar{x}$ классом эквивалентности элемента x .

Пример I.1.41. Пусть ρ есть отношение эквивалентности из примера I.1.39. В этом случае $\bar{0} = 2\mathbb{Z}$ – множество всех четных чисел и $\bar{1} = \mathbb{Z} \setminus 2\mathbb{Z}$ есть множество всех нечетных чисел. Заметим, что в этом примере выполняются следующие равенства: $\bar{0} \cup \bar{1} = \mathbb{Z}$ и $\bar{0} \cap \bar{1} = \emptyset$.

Лемма I.1.42. Для произвольного отношения эквивалентности ρ на множестве X и произвольного класса эквивалентности $\bar{x}$ для $x \in X$ имеют место следующие свойства:

а) для любого $x \in X$ имеем $x \in \bar{x}$;

б) для любого $y \in \bar{x}$ имеем $\bar{y} = \bar{x}$.

Доказательство. а) Пусть $x \in X$, тогда из рефлексивности отношения эквивалентности ρ следует xrx . По определению $\bar{x}$ для любого $y \in X$, если $урх$, то $y \in \bar{x}$. Но из xrx тогда следует, что $x \in \bar{x}$.

б) Пусть $z \in \bar{y}$ – произвольный элемент, тогда имеем $zру$. Так как $y \in \bar{x}$, то $урх$ выполнено и вместе с $zру$ по транзитивности ρ получаем $zрх$, откуда следует, что $z \in \bar{x}$. Мы показали, что $\bar{y} \subseteq \bar{x}$. Покажем теперь обратное включение. Пусть $z \in \bar{x}$, тогда имеем $zрх$. Учитывая, что $урх$, получаем по симметричности ρ , что $xру$. Но тогда из $zрх$ и $xру$ вытекает $zру$, откуда $z \in \bar{y}$. Следовательно, $\bar{x} \subseteq \bar{y}$. Из $\bar{y} \subseteq \bar{x}$ и $\bar{x} \subseteq \bar{y}$ получаем $\bar{y} = \bar{x}$.

Определение I.1.43. Обозначим через $\pi(X)$ множество, элементами которого являются некоторые непустые подмножества S_α множества X , где $\alpha \in I$, т.е.

$$\pi(X) = \{S_\alpha \mid \emptyset \neq S_\alpha \subseteq X, \alpha \in I\}.$$

Множество $\pi(X)$ называется разбиением множества X , если выполняются следующие условия:

- 1) $X = \bigcup_{\alpha \in I} S_\alpha$;
- 2) из $S_\alpha \neq S_\beta$ следует $S_\alpha \cap S_\beta = \emptyset$.

Пример I.1.44. Пусть ρ есть отношение эквивалентности из примера I.1.39. Как следует из примера I.1.41, $\pi(\mathbb{Z}) = \{\bar{0}, \bar{1}\}$ есть разбиение множества $\mathbb{Z}$.

Теорема I.1.45.

1. Пусть ρ есть отношение эквивалентности на множестве X . Тогда $\pi(X) = \{\bar{x} \mid x \in X\}$ есть разбиение множества X .

2. Пусть $\pi(X) = \{S_\alpha \mid \alpha \in I\}$ – произвольное разбиение множества X . Тогда существует отношение эквивалентности ρ на множестве X такое, что любой класс эквивалентности $\bar{x}$ относительно ρ совпадает с некоторым $S_\alpha \in \pi(X)$ и любой $S_\alpha \in \pi(X)$ совпадает с некоторым классом эквивалентности $\bar{x}$ относительно ρ .

Доказательство 1. Пусть ρ есть отношение эквивалентности на множестве X . Покажем, что $\pi(X) = \{\bar{x} \mid x \in X\}$ есть разбиение множества X . Покажем сначала, что $X = \bigcup_{x \in X} \bar{x}$. Так как $\bar{x} \subseteq X$ для любого $x \in X$, то имеем $\bigcup_{x \in X} \bar{x} \subseteq X$. Для доказательства обратного включения пусть $y \in X$ – произвольный элемент, тогда из $y \in \bar{y}$ и $\bar{y} \subseteq \bigcup_{x \in X} \bar{x}$ получаем $y \in \bigcup_{x \in X} \bar{x}$. Следовательно, $X \subseteq \bigcup_{x \in X} \bar{x}$ и тогда $X = \bigcup_{x \in X} \bar{x}$.

Для проверки второй части определения разбиения пусть $\bar{x} \neq \bar{y}$ – произвольные различные классы эквивалентности относительно ρ . Покажем, что $\bar{x} \cap \bar{y} = \emptyset$. Пусть $z \in \bar{x} \cap \bar{y}$, тогда $z \in \bar{x}$ и $z \in \bar{y}$, откуда имеем $z\rho x$ и $z\rho y$, тогда из $x\rho z$ и $z\rho y$ получаем, что имеет место $x\rho y$. Следовательно, $x \in \bar{y}$ и по Лемме I.1.42 «б» получаем $\bar{x} = \bar{y}$.

Доказательство 2. Пусть $\pi(X)$ есть произвольное разбиение множества X , $\pi(X) = \{S_\alpha \mid \emptyset \neq S_\alpha \subseteq X, \alpha \in I\}$. Зададим бинарное отношение ρ на множестве X следующим образом: для любых $x, y \in X$ пусть $x\rho y$ тогда и только тогда, когда x и y принадлежат одному и тому же подмножеству S_α . Непосредственно проверяется, что ρ есть отношение эквивалентности. Обозначим $\pi_1(X) = \{\bar{x} \mid x \in X\}$ – множество классов эквивалентности относительно ρ . Для доказательства утверждения 2 нужно показать, что $\pi(X) = \pi_1(X)$.

Покажем сначала, что $\pi(X) \subseteq \pi_1(X)$. Пусть $S_\alpha \in \pi(X)$ – произвольный элемент $\pi(X)$, и пусть x_0 есть произвольный элемент S_α . Докажем, что $S_\alpha = \bar{x}_0$, что и будет означать то, что $S_\alpha \in \pi_1(X)$ и, следовательно, $\pi(X) \subseteq \pi_1(X)$.

Пусть $y \in S_\alpha$ – произвольный элемент, тогда из $x_0 \in S_\alpha$ получаем $x_0\rho y$, откуда следует $y \in \bar{x}_0$, т.е. $S_\alpha \subseteq \bar{x}_0$. Пусть $y \in \bar{x}_0$, тогда имеем $x_0\rho y$, откуда по определению ρ получаем, так как $x_0 \in S_\alpha$, что $y \in S_\alpha$, т.е. $\bar{x}_0 \subseteq S_\alpha$. Следовательно, $S_\alpha = \bar{x}_0$.

Покажем теперь, что $\pi_1(X) \subseteq \pi(X)$. Пусть $\bar{x} \in \pi_1(X)$ – произвольный класс эквивалентности относительно ρ , тогда существует $S_\alpha \in \pi(X)$ такой, что $x \in S_\alpha$ ($\pi(X)$ есть разбиение множества X). Покажем, что $\bar{x} = S_\alpha$. Если $y \in \bar{x}$, то получаем $x\rho y$, откуда следует $y \in S_\alpha$. Если $y \in S_\alpha$, то, учитывая $x \in S_\alpha$, получаем $x\rho y$, откуда вытекает $y \in \bar{x}$. Таким образом, $\bar{x} = S_\alpha \in \pi(X)$, следовательно,

$$\pi_1(X) \subseteq \pi(X).$$

Определение I.1.46. Пусть $\mathbb{N}_k = \{1, 2, \dots, k\}$ – отрезок натурального ряда чисел. Назовем подстановкой на множестве $\mathbb{N}_k$ произвольную биекцию на этом множестве.

Пример I.1.47. Рассмотрим отображение $\varphi: \mathbb{N}_3 \rightarrow \mathbb{N}_3$, заданное в виде

$$\varphi(1) = 2, \quad \varphi(2) = 1, \quad \varphi(3) = 3.$$

Убедимся, что отображение φ является биекцией.

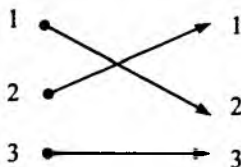
1. Сюръективность φ .

Для любого из чисел в $\mathbb{N}_3$ есть прообраз относительно φ :

для 1 прообразом является 2,
для 2 прообразом является 1,
для 3 прообразом является 3.

2. Инъективность φ .

Различные числа из $\mathbb{N}_3$ при действии отображения φ переходят в различные числа, наглядно это можно изобразить так:



Следовательно, φ есть биекция на $\mathbb{N}_3$, т.е. φ – подстановка.

Условимся записывать подстановку φ на множестве $\mathbb{N}_k$ следующим образом:

$$\varphi = \begin{pmatrix} 1 & 2 & \dots & k \\ \downarrow & \downarrow & & \downarrow \\ \varphi(1) & \varphi(2) & \dots & \varphi(k) \end{pmatrix},$$

где верхняя строка чисел из $\mathbb{N}_k$ упорядочена в естественном порядке, а нижняя строка состоит из образов чисел верхней строки при действии отображения φ . Заметим, что перестановка столбцов этой таблицы не меняет подстановку φ . Для упрощения впредь будем опускать стрелки в записи подстановки. Например, подстановка φ из предыдущего примера запишется как

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

Рассмотрим произвольный набор чисел из $\mathbb{N}_k$, скажем, $a_{i_1}, a_{i_2}, \dots, a_{i_k}$.

Будем говорить, что пара чисел a_{i_s} и a_{i_t} из этого набора образует порядок, если меньшее из этих двух чисел расположено левее большего числа, а если оно расположено правее большего числа, то эта пара чисел образует инверсию. Например, рассмотрим набор чисел 3, 1, 4, 5, 2. Пара чисел (3, 2) образует инверсию. Кроме того, инверсию в этом наборе также образуют пары (3, 1), (3, 2), (4, 2) и (5, 2).

Лекция 2. МАТРИЦЫ И ОПРЕДЕЛИТЕЛИ

Определение I.2.1. Любую прямоугольную таблицу A из m строк и n столбцов, элементы которой являются действительными числами, будем называть матрицей размера $m \times n$, а ее элементы обозначать a_{ij} , где i – номер строки, j – номер столбца, на пересечении которых находится элемент a_{ij} . Если $m = n$, то A называется квадратной матрицей порядка n .

Пример I. 2.2. Рассмотрим матрицу

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \text{ размера } 2 \times 3.$$

Элемент a_{13} находится на пересечении первой строки и третьего столбца, элемент a_{21} – на пересечении второй строки и первого столбца.

Обозначим через $M_{m,n}(\mathbb{R})$ множество всех матриц размера $m \times n$ с элементами из $\mathbb{R}$, а $M_n(\mathbb{R})$ – множество всех квадратных матриц порядка n с элементами из $\mathbb{R}$.

Определение I.2.3. Пусть $A, B \in M_{m,n}(\mathbb{R})$. Будем говорить, что матрицы $A = (a_{ij})$ и $B = (b_{ij})$ равны (обозначение: $A = B$), если для любых $i = \overline{1, m}$ и $j = \overline{1, n}$ имеем $a_{ij} = b_{ij}$.

Зададим для матриц операции сложения и умножения на число.

Определение I.2.4. Пусть $A, B \in M_{m,n}(\mathbb{R})$. Назовем суммой матриц $A = (a_{ij})$ и $B = (b_{ij})$ матрицу $C = (c_{ij}) \in M_{m,n}(\mathbb{R})$, для которой

$$c_{ij} = a_{ij} + b_{ij}$$

для всех $i = \overline{1, m}$, $j = \overline{1, n}$.

Пример I.2.5. Пусть

$$A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \\ 0 & 1 \\ -1 & 2 \end{pmatrix}, B = \begin{pmatrix} 3 & 4 \\ -3 & 1 \\ -2 & 0 \\ 4 & -3 \end{pmatrix}.$$

Тогда

$$A + B = \begin{pmatrix} 1+3 & -1+4 \\ 2-3 & 3+1 \\ 0-2 & 1+0 \\ -1+4 & 2-3 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ -1 & 4 \\ -2 & 1 \\ 3 & -1 \end{pmatrix}.$$

Определение I.2.6. Пусть $A \in M_{m,n}(\mathbb{R})$, $A = (a_{ij})$ и $\alpha \in \mathbb{R}$. Назовем произведением матрицы A на число α следующую матрицу:

$$C = (c_{ij}),$$

где $c_{ij} = \alpha a_{ij}$ для любых $i = \overline{1, m}$, $j = \overline{1, n}$.

Пример I.2.7. Пусть $\alpha = -2$ и

$$A = \begin{pmatrix} -1 & 1 & 2 & 3 \\ 3 & -2 & 0 & 1 \end{pmatrix},$$

тогда

$$\begin{aligned} \alpha A &= \begin{pmatrix} (-2) \cdot (-1) & (-2) \cdot 1 & (-2) \cdot 2 & (-2) \cdot 3 \\ (-2) \cdot 3 & (-2) \cdot (-2) & (-2) \cdot 0 & (-2) \cdot 1 \end{pmatrix} = \\ &= \begin{pmatrix} 2 & -2 & -4 & -6 \\ -6 & 4 & 0 & -2 \end{pmatrix}. \end{aligned}$$

Определим далее операцию умножения матриц. Кажется естественным определить умножение матриц аналогично сложению, т.е. перемножить элементы матриц на одинаковых позициях. Однако такое определение умножения матриц (оно называется произведением матриц по Адамару) не нашло применения в различных приложениях матриц, и поэтому рассмотрим определение умножения матриц, которое является общепринятым.

Определение 1.2.8. Пусть $A \in M_{m,n}(\mathbb{R})$, $B \in M_{n,k}(\mathbb{R})$, т.е. $A = (a_{ij})$, $a_{ij} \in \mathbb{R}$ для $i = \overline{1, m}$; $j = \overline{1, n}$ – матрица размера $m \times n$, $B = (b_{rs})$, $b_{rs} \in \mathbb{R}$ для $r = \overline{1, n}$; $s = \overline{1, k}$ – матрица размера $n \times k$. Тогда назовем произведением матрицы A на матрицу B следующую матрицу $C = (c_{is}) = A \cdot B$ размера $m \times k$:

$$c_{is} = \sum_{j=1}^n a_{ij} b_{js}$$

для $i = \overline{1, m}$; $s = \overline{1, k}$.

Другими словами, если число столбцов матрицы A равно числу строк матрицы B , то произведение матриц $AB = C$ определено и любой элемент $c_{is} \in C$ на пересечении i -й строки и s -го столбца матрицы C вычисляется по правилу «строка на столбец», а именно: первый элемент i -й строки матрицы A умножаем на первый элемент s -го столбца матрицы B плюс второй элемент i -й строки умножаем на второй элемент s -го столбца плюс и т.д. плюс n -й элемент i -й строки умножаем на n -й элемент s -го столбца.

Пример 1.2.9. Пусть

$$A = \begin{pmatrix} 1 & 2 & -3 \\ -2 & 1 & 2 \end{pmatrix} \text{ и } B = \begin{pmatrix} 0 & 1 & -1 & 2 \\ 1 & 1 & 0 & -1 \\ -1 & 2 & -2 & 1 \end{pmatrix},$$

тогда произведение $C = AB$ определено, так как число столбцов матрицы A равно числу строк матрицы B , причем матрица C имеет размер $m \times k$, т.е. в данном случае 2×4 . Имеем

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

Заметим, что в примере 1.2.9 произведение BA не определено, так как число столбцов матрицы B не равно числу строк матрицы A . Очевидно, что для квадратных матриц оба произведения AB и BA определены, но не обязательно $AB = BA$.

Пример 1.2.10. Пусть

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \text{ и } B = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix},$$

тогда

$$AB = \begin{pmatrix} -2 & 3 \\ -1 & 0 \end{pmatrix} \text{ и } BA = \begin{pmatrix} -1 & 1 \\ -2 & -1 \end{pmatrix}.$$

Определение 1.2.11. Пусть $A \in M_n(\mathbb{R})$ – квадратная матрица порядка n с элементами из $\mathbb{R}$. Назовем определителем n -го порядка квадратной матрицы $A = (a_{ij})$ число $\alpha \in \mathbb{R}$, являющееся алгебраической суммой $n!$ слагаемых вида

$$A_j = (-1)^{s_j} a_{1i_1} a_{2i_2} \cdots a_{ni_n}, j = \overline{1, n!},$$

где каждое такое слагаемое A_j является произведением элементов матрицы A , взятых по одному из каждой строки и каждого столбца этой матрицы, а s_j есть сумма общего количества инверсий в обоих строках подстановки σ_j индексов слагаемого A_j , а именно

$$\sigma_j = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}.$$

Определитель матрицы A обозначается $|A|$ (иногда $\det A$). Допуская вольность речи, говорят об элементах, строках или столбцах определителя, хотя это элементы, строки или столбцы матрицы данного определителя.

Вычислим определители 2-го и 3-го порядка, используя определение, данное выше.

Пример 1.2.12. Пусть $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ есть квадратная матрица 2-го

порядка. Согласно определению, имеем $2! = 2$ слагаемых:

$$|A| = (-1)^{s_1} a_{11} a_{22} + (-1)^{s_2} a_{12} a_{21}.$$

Имеем

$$\sigma_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},$$

$s_1 = 0$ и $s_2 = 1$, тогда

$$|A| = a_{11}a_{22} - a_{12}a_{21}.$$

Пример I.2.13. Пусть

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

есть квадратная матрица 3-го порядка. Согласно определению, имеем $3! = 6$ слагаемых:

$$\begin{aligned} |A| = & (-1)^{s_1} a_{11}a_{22}a_{33} + (-1)^{s_2} a_{11}a_{23}a_{32} + \\ & + (-1)^{s_3} a_{12}a_{21}a_{33} + (-1)^{s_4} a_{12}a_{23}a_{31} + \\ & + (-1)^{s_5} a_{13}a_{22}a_{31} + (-1)^{s_6} a_{13}a_{21}a_{32}. \end{aligned}$$

Имеем

$$\begin{aligned} \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \\ \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \sigma_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}. \end{aligned}$$

Отсюда

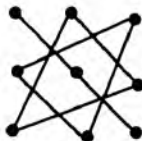
$$s_1 = 0, s_2 = 1, s_3 = 1, s_4 = 2, s_5 = 3, s_6 = 2.$$

Тогда

$$\begin{aligned} |A| = & a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + \\ & + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31} + a_{13}a_{21}a_{32}. \end{aligned}$$

Для облегчения запоминания этой формулы вычисления определителя 3-го порядка можно сформулировать правило треугольников.

Три слагаемых со знаком «плюс» изображаются как



а три слагаемых со знаком «минус» – в виде



Рассмотрим теперь свойства определителей.

Определение 1.2.14. Пусть $A = (a_{ij})$ – квадратная матрица порядка n , тогда транспонированной матрицей A^T называется матрица, полученная заменой строк матрицы A на ее столбцы с теми же номерами.

Пример 1.2.15. Пусть

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix},$$

тогда

$$A^T = \begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix}.$$

Свойство 1. $|A^T| = |A|$, т.е. при транспонировании матрицы A определитель не меняется.

Доказательство. Пусть произвольное слагаемое A_j определителя $|A|$ имеет вид

$$A_j = (-1)^{s_j} a_{i_1 i_1} a_{i_2 i_2} \dots a_{i_n i_n}.$$

Тогда соответствующее слагаемое определителя $|A^T|$ имеет вид

$$A'_j = (-1)^{s'_j} a_{i_1 1} a_{i_2 2} \dots a_{i_n n},$$

где подстановки индексов выглядят так:

$$\sigma_j = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}, \quad \sigma'_j = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ 1 & 2 & \dots & n \end{pmatrix}.$$

Ясно, что $s_j = s'_j$ для любого $j = \overline{1, n!}$. Таким образом, каждое слагаемое определителя $|A|$ является однозначно определенным слагаемым $|A^T|$, причем с тем же знаком, и наоборот, каждое слагаемое $|A^T|$ является однозначно определенным слагаемым $|A|$, причем с тем же знаком. Таким образом, множества слагаемых $|A|$ и $|A^T|$ совпадают, но тогда $|A| = |A^T|$.

Свойство 2. Если некоторая строка (или столбец) матрицы A состоит из нулей, то $|A| = 0$.

Доказательство. В этом случае произвольное слагаемое определителя $|A|$ содержит в качестве множителя нуль, следовательно, $|A| = 0$.

Свойство 3. При перестановке двух строк (столбцов) матрицы A определитель $|A|$ меняет знак.

Доказательство. Пусть матрица A_1 получается из матрицы A перестановкой строк с номерами t и r . Рассмотрим произвольное слагаемое определителя $|A|$:

$$(-1)^{s_j} a_{1i_1} \dots a_{ni_n},$$

где подстановка

$$\sigma_j = \begin{pmatrix} 1 & 2 \dots & r \dots & t \dots & n \\ i_1 & i_2 \dots & i_r \dots & i_t \dots & i_n \end{pmatrix}.$$

Это слагаемое входит и в определитель $|A_1|$, но с подстановкой

$$\sigma'_j = \begin{pmatrix} 1 & 2 \dots & r \dots & t \dots & n \\ i_1 & i_2 \dots & i_t \dots & i_r \dots & i_n \end{pmatrix}$$

и со знаком $(-1)^{s'_j}$.

Покажем, что четности s_j и s'_j различны. Понятно, что различие чисел s_j и s'_j зависит от инверсий чисел i_r и i_t с числами, расположенными между ними. Пусть в нижней строке подстановки σ_j между числами i_r и i_t находится p чисел и число i_r составляет ℓ инверсий с эти-

ми p числами, а число i_k составляет k инверсий с этими же числами. Тогда в нижней строке подстановки σ'_j число i_r составляет $p - \ell$ инверсий с числами между i_k и i_r , а число i_k составляет $p - k$ инверсий. Следовательно, общее число инверсий, которые числа i_r и i_k составляют с числами между ними в нижней строке подстановки σ_j , равно $\ell + k$, а общее число инверсий чисел i_r и i_k с промежуточными числами в нижней строке подстановки σ'_j равно $2p - (k + \ell)$. Обозначая количество всех остальных инверсий в нижней строке подстановки σ_j через w , получим, что

$$s_j = w + k + \ell,$$

а

$$s'_j = w + [2p - (k + \ell)] \pm 1,$$

поскольку здесь учитывается еще одно изменение числа инверсий от перестановки i_r и i_k между собой. Тогда

$$|s_j - s'_j| = |2p - 2(k + \ell) \pm 1|$$

есть нечетное число. Следовательно, четности чисел s_j и s'_j различны, и тогда знаки соответствующих слагаемых в $|A|$ и $|A_1|$ различны. Таким образом, $|A|$ и $|A_1|$ различаются только знаком.

Свойство 4. Если в матрице A две одинаковые строки (столбца), то $|A| = 0$.

Доказательство. Если в матрице A поменять местами две одинаковые строки, то от этого матрица не изменится. С другой стороны, по свойству 3 определитель матрицы A меняет знак. Тогда $|A| = -|A|$, откуда $2|A| = 0$ и $|A| = 0$.

Свойство 5. Если все элементы некоторой строки (столбца) матрицы A имеют общий множитель α , то α можно вынести за знак определителя.

Доказательство. По определению каждое слагаемое определителя матрицы A имеет общий множитель α , откуда и следует утверждение.

Свойство 6. Если в матрице $A = (a_{ij}) \in M_n(\mathbb{R})$ все элементы фиксированной строки i_0 представлены в виде суммы $a_{i_0 j} = b_{i_0 j} + c_{i_0 j}$ для любого $j = \overline{1, n}$, то $|A| = |B| + |C|$, где в матрицах B и C все строки, кроме строки i_0 , такие же, как в матрице A , а строка i_0 состоит в матрице B из элементов $b_{i_0 j}$, а в матрице C – из элементов $c_{i_0 j}$.

Доказательство. Рассмотрим произвольное слагаемое определителя $|A|$, скажем,

$$(-1)^{s_1} a_{1j_1} a_{2j_2} \dots a_{i_0 j_0} \dots a_{nj_n}.$$

Поскольку $a_{i_0 j_0} = b_{i_0 j_0} + c_{i_0 j_0}$, то это слагаемое $|A|$ разбивается на два слагаемых

$$(-1)^{s_1} a_{1j_1} a_{2j_2} \dots b_{i_0 j_0} \dots a_{nj_n}$$

и

$$(-1)^{s_1} a_{1j_1} a_{2j_2} \dots c_{i_0 j_0} \dots a_{nj_n}.$$

Собирая вместе $n!$ первых слагаемых и $n!$ вторых слагаемых, получим соответственно $|B|$ и $|C|$.

Определение 1.2.16. Будем говорить, что строка (столбец) матрицы A , скажем, с номером i_0 является линейной комбинацией остальных строк (столбцов), если каждый элемент строки i_0 является суммой соответствующих элементов остальных строк, предварительно умноженных на фиксированные числа для каждой позиции.

Пример 1.2.17. Пусть

$$A = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 1 & 1 \\ -1 & 8 & -1 \end{pmatrix}.$$

Тогда третья строка матрицы A является линейной комбинацией остальных строк, а именно, третья строка является суммой первой строки, предварительно умноженной на 3, и второй строки, предварительно умноженной на 2. Действительно, если все элементы первой строки умножить на 3, то получится набор чисел $(3, 6, -3)$, а если все элементы второй

строки умножить на 2, то получится набор чисел $(-4, 2, 2)$. Сумма этих наборов чисел (т.е. суммируем соответствующие элементы) является набором чисел $(-1, 8, -1)$, что совпадает с третьей строкой матрицы A .

Свойство 7. Если некоторая ℓ -я строка (столбец) матрицы A является линейной комбинацией остальных строк (столбцов), то $|A| = 0$.

Доказательство. Рассмотрим произвольное слагаемое $|A|$:

$$(-1)^{j_k} a_{1j_1} a_{2j_2} \dots a_{\ell j_\ell} \dots a_{nj_n}.$$

По условию имеем для каждого $j_\ell = \overline{1, n}$:

$$a_{\ell j_\ell} = \sum_{\substack{p=1 \\ p \neq \ell}}^n \lambda_p a_{pj_\ell}.$$

По свойству 6 определитель $|A|$ записывается в виде суммы определителей:

$$|A| = \sum_{\substack{p=1 \\ p \neq \ell}}^n |A_p|,$$

где в матрицах A_p все строки, кроме ℓ -й, такие же, как в матрице A , а ℓ -я строка имеет вид

$$(\lambda_p a_{p1} \lambda_p a_{p2} \dots \lambda_p a_{pn}).$$

По свойству 5 тогда имеем для любого $p = \overline{1, n}, p \neq \ell$:

$$|A_p| = \lambda_p |A'_p|,$$

где в матрицах A'_p все строки, кроме ℓ -й, такие же, как в матрице A , а ℓ -я строка имеет вид $(a_{p1} a_{p2} \dots a_{pn})$. Но тогда матрица $A'_p, p = \overline{1, n}, p \neq \ell$, содержит две одинаковые строки с различными номерами p и ℓ .

Следовательно, по свойству 4 все матрицы $A'_p, p = \overline{1, n}, p \neq \ell$, имеют определитель $|A'_p| = 0$. Но тогда и все $A_p = 0, p = \overline{1, n}, p \neq \ell$. Таким образом,

поскольку $|A| = \sum_{\substack{p=1 \\ p \neq \ell}}^n |A_p|$, то $|A| = 0$.

Свойство 8. Определитель матрицы A не меняется, если к любой строке (столбцу) матрицы A прибавить произвольную линейную комбинацию остальных строк (столбцов).

Доказательство. Рассмотрим матрицу A_1 , у которой фиксированная i -я строка является произвольной линейной комбинацией остальных строк, а все остальные строки – такие же, как у матрицы A . Пусть матрица A_2 имеет i -ю строку, которая является суммой i -х строк матриц A и A_1 , а все остальные строки – такие же, как у матрицы A . Тогда по свойству 6 имеем

$$|A_2| = |A| + |A_1|.$$

По свойству 7 определитель $|A_1| = 0$, следовательно,

$$|A_2| = |A|.$$

Определение I.2.18. Пусть $A = (a_{ij}) \in M_n(\mathbb{R})$. Выберем k строк и k столбцов в матрице A . Элементы матрицы A , находящиеся на их пересечениях, образуют квадратную матрицу порядка k , определитель которой $|L_k|$ называется минором k -го порядка.

Пример I.2.19. Пусть

$$A = \begin{pmatrix} 3 & 4 & 2 & 1 \\ -1 & 0 & 1 & 5 \\ 2 & -3 & 0 & 1 \\ 1 & -1 & 2 & 1 \end{pmatrix} \text{ и } k = 2.$$

Выберем первую и третью строки, второй и четвертый столбцы матрицы A . На их пересечениях находится матрица $L_2 = \begin{pmatrix} 4 & 1 \\ -3 & 1 \end{pmatrix}$, определитель которой $|L_2| = 7$ является минором 2-го порядка.

Определение I.2.20. Пусть $|L_k|$ есть минор k -го порядка в матрице $A = (a_{ij}) \in M_n(\mathbb{R})$. Обозначим через S_{L_k} сумму номеров строк и столбцов, на пересечениях которых расположены элементы матрицы L_k . Элементы матрицы A , находящиеся на пересечениях $n-k$ строк и $n-k$ столбцов, не вошедших в L_k , образуют квадратную матрицу L'_k порядка $n-k$, определитель которой $|L'_k|$ называется дополнительным

минором к минору $|L_k|$. Алгебраическим дополнением к минору $|L_k|$ называется число

$$A_{L_k} = (-1)^{s_{L_k}} |L'_k|.$$

Пример I.2.21. Найдем алгебраическое дополнение к минору $|L_2|$ из примера I.2.19. Элементы дополнительного минора $|L'_2|$ расположены на пересечениях второй и четвертой строк, первого и третьего столбцов матрицы A , а именно

$$L'_2 = \begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Тогда дополнительный минор $|L_2|$ есть $|L'_2| = -3$.

Для вычисления алгебраического дополнения A_{L_2} найдем сначала S_{L_2} . Имеем

$$S_{L_2} = \underset{\text{номера строк}}{(1+3)} + \underset{\text{номера столбцов}}{(2+4)} = 10.$$

Тогда

$$A_{L_2} = (-1)^{10} \cdot |L'_2| = -3.$$

Мы теперь можем сформулировать одну из основных теорем теории определителей.

Теорема I.2.22. (Теорема Лапласа). Пусть в матрице $A \in M_n(\mathbb{R})$ выбрано k строк (или k столбцов). Определитель $|A|$ матрицы A равен сумме произведений всевозможных миноров k -го порядка, расположенных в выбранных строках (столбцах), на их алгебраические дополнения.

Пример I.2.23. Пусть матрица

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

Пусть $k = 2$, и выбираем первые две строки. В этих двух строках имеется всего один ненулевой минор, а именно

$$|L_2| = \begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix} = -5.$$

Алгебраическое дополнение к минору $|L_2|$ равно

$$A_{L_2} = (-1)^{S_{L_2}} \cdot \begin{vmatrix} 1 & 1 & 4 \\ 0 & 1 & 0 \\ 2 & 2 & -2 \end{vmatrix}.$$

Находим S_{L_2} :

$$S_{L_2} = (1+2) + (1+3) = 7.$$

Вычисляем определитель по правилу треугольников:

$$\begin{vmatrix} 1 & 1 & 4 \\ 0 & 1 & 0 \\ 2 & 2 & -2 \end{vmatrix} = -2 + 0 + 0 - 8 + 0 + 0 = -10.$$

Тогда $A_{L_2} = (-1)^7 \cdot (-10) = 10$ и, следовательно,

$$|A| = |L_2| \cdot A_{L_2} = (-5) \cdot 10 = -50.$$

Следствие 1. Определитель матрицы A равен сумме произведений всех элементов некоторой строки (столбца) матрицы A на их алгебраические дополнения.

Доказательство. Применяем теорему Лапласа для случая $k = 1$.

Пример 1.2.24.

Пусть

$$A = \begin{pmatrix} 2 & 0 & -1 \\ 3 & 3 & -2 \\ 1 & 0 & 1 \end{pmatrix}.$$

Выбираем второй столбец, тогда

$$|A| = 3 \cdot (-1)^{2+2} \cdot \begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} = 9.$$

Следствие 2. Сумма произведений всех элементов некоторой строки (столбца) матрицы A на алгебраические дополнения соответствующих элементов любой другой строки (столбца) равна нулю.

Доказательство. Рассмотрим сумму произведений всех элементов произвольной k -й строки матрицы A на алгебраические дополнения соответствующих элементов любой другой, скажем i -й, строки матрицы A . Пусть A' – матрица, у которой все строки, кроме i -й, такие же, как у матрицы A , а элементами i -й строки матрицы A' являются соответствующие элементы k -й строки матрицы A . Тогда у матрицы A' две одинаковые строки и, следовательно, по свойству 4 имеем, что $|A'| = 0$.

С другой стороны, по следствию 1 определитель $|A'|$ равен сумме произведений всех элементов i -й строки матрицы A' на их алгебраические дополнения. Заметим, что алгебраические дополнения элементов i -й строки матрицы A' совпадают с алгебраическими дополнениями соответствующих элементов i -й строки матрицы A . Но элементами i -й строки матрицы A' являются соответствующие элементы k -й строки матрицы A . Таким образом, сумма произведений всех элементов i -й строки матрицы A' на их алгебраические дополнения с одной стороны равна нулю, а с другой стороны равна сумме произведений всех элементов k -й строки матрицы A на алгебраические дополнения соответствующих элементов i -й строки матрицы A .

Теорема I.2.25. Пусть матрицы $A, B \in M_n(\mathbb{R})$. Тогда $|A \cdot B| = |A| \cdot |B|$.

Лекция 3. РАНГ МАТРИЦЫ

Пусть $A = (a_{ij}), i = \overline{1, n}, j = \overline{1, m}$ – прямоугольная матрица над $\mathbb{R}$, и предположим, что $A \neq 0$, т.е. хотя бы один из $a_{ij} \neq 0$.

Определение I.3.1. Назовем рангом матрицы A и обозначим его $rk(A)$ наибольший из порядков отличных от нуля миноров матрицы A . Если $r = rk(A)$, то любой ненулевой минор порядка r назовем базисным минором матрицы A , а строки и столбцы матрицы A , на пересечении которых расположен базисный минор, назовем базисными строками и столбцами.

Пример I.3.2. Пусть матрица A имеет вид

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

В этой матрице отличен от нуля только один минор второго порядка, а именно

$$\begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} = 1.$$

Все миноры третьего порядка равны нулю, поскольку содержат нулевую строку или нулевой столбец. Тогда наибольший из порядков ненулевых миноров матрицы A равен 2, т.е. $rk(A) = 2$.

Первая и третья строки являются базисными строками, а первый и второй столбцы являются базисными столбцами.

Рассмотрим методы вычисления ранга матрицы.

Определение I.3.3. Пусть M_k – минор k -го порядка матрицы A . Выберем строку с номером i и столбец с номером j , которые не проходят через минор M_k . Тогда минор M_{k+1} порядка $k+1$, проходящий через k строк и столбцов минора M_k и также i -ю строку и j -й столбец,

назовем минором, окаймляющим минор M_k , или просто окаймляющим минором матрицы A .

Пример I.3.4. Пусть матрица A имеет вид

$$A = \begin{pmatrix} 1 & -1 & 0 & 1 & 0 \\ 2 & -2 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 1 \\ -1 & 1 & 0 & -1 & 0 \end{pmatrix}.$$

Рассмотрим минор M_2 , расположенный на пересечении первой и третьей строк, второго и пятого столбцов матрицы A , а именно

$$M_2 = \begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix}.$$

Выберем для окаймления вторую строку и первый столбец матрицы A , тогда окаймляющий минор M_3 расположен на пересечении первых трех строк и первого, второго и пятого столбцов матрицы A :

$$\begin{pmatrix} \boxed{1} & \boxed{-1} & 0 & 1 & \boxed{0} \\ \boxed{2} & \boxed{-2} & 0 & 0 & \boxed{0} \\ \boxed{1} & \boxed{0} & 0 & -1 & \boxed{1} \\ -1 & 1 & 0 & -1 & 0 \end{pmatrix},$$

то есть

$$M_3 = \begin{vmatrix} 1 & -1 & 0 \\ 2 & -2 & 0 \\ 1 & 0 & 1 \end{vmatrix}.$$

1. Метод окаймляющих миноров.

Поскольку $A \neq 0$, то $rk(A) \geq 1$, т.е. существует ненулевой минор первого порядка $a_{ij} \neq 0$.

Если все миноры второго порядка матрицы A , окаймляющие a_{ij} , равны нулю, то $rk(A) = 1$. В противном случае выберем ненулевой минор второго порядка M_2 , окаймляющий a_{ij} . Опять имеем две возможности. Если все миноры третьего порядка матрицы A , окаймляющие

минор M_2 , равны нулю, то $rk(A) = 2$. В противном случае выберем ненулевой минор третьего порядка M_3 , окаймляющий M_2 , и т.д.

Продолжая подобным образом, получим, что либо на некотором шаге для ненулевого минора M_k все окаймляющие миноры $k+1$ -го порядка равны нулю, и тогда $rk(A) = k$, либо дойдем до ненулевого минора M_ℓ максимально возможного порядка, равного $\ell = \min(n, m)$, и тогда $rk(A) = \ell$.

Пример I.3.5. Рассмотрим матрицу A из примера I.3.4:

$$A = \begin{pmatrix} 1 & -1 & 0 & 1 & 0 \\ 2 & -2 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 1 \\ -1 & 1 & 0 & -1 & 0 \end{pmatrix}.$$

На пересечении третьей и четвертой строк и первых двух столбцов матрицы A находится ненулевой минор второго порядка

$$M_2 = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1 \neq 0.$$

Среди миноров третьего порядка, окаймляющих минор M_2 , есть ненулевой минор, полученный добавлением третьей строки и четвертого столбца, а именно

$$M_3 = \begin{vmatrix} 2 & -2 & 0 \\ 1 & 0 & -1 \\ -1 & 1 & -1 \end{vmatrix} = -2 \neq 0.$$

Найдем два окаймляющих минора четвертого порядка, полученных добавлением первой строки и третьего столбца, а также первой строки и пятого столбца.

Имеем

$$M_4^{(1)} = \begin{vmatrix} 1 & -1 & 0 & 1 \\ 2 & -2 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & -1 \end{vmatrix} = 0,$$

так как третий столбец состоит из нулей,

$$M_4^{(2)} = \begin{vmatrix} 1 & -1 & 1 & 0 \\ 2 & -2 & 0 & 0 \\ 1 & 0 & -1 & 1 \\ -1 & 1 & -1 & 0 \end{vmatrix} = 0,$$

так как первая и четвертая строки пропорциональны с коэффициентом пропорциональности -1 .

Таким образом, все миноры четвертого порядка, окаймляющие минор M_3 , равны нулю, тогда $rk(A) = 3$.

Рассмотрим теперь другой метод вычисления ранга матрицы, а именно метод элементарных преобразований.

Определение 1.3.6. Будем называть элементарными преобразованиями матрицы A следующие ее преобразования:

- 1) транспонирование матрицы A ;
- 2) перестановка двух строк или двух столбцов матрицы A ;
- 3) умножение любого элемента некоторой строки или столбца на отличное от нуля число $\alpha \in \mathbb{R}$;
- 4) прибавление ко всем элементам некоторой строки или столбца соответствующих элементов другой строки или столбца, предварительно умноженной на отличное от нуля число $\alpha \in \mathbb{R}$.

Имеет место следующая теорема, которую приведем без доказательства.

Теорема 1.3.7. Любое из элементарных преобразований матрицы A не меняет ее ранг.

Теперь мы можем привести метод элементарных преобразований.

Шаг 1. Если $a_{11} \neq 0$, то умножаем первую строку матрицы A на $1/a_{11}$, в результате получим $a'_{11} = 1$. В противном случае вначале перестановкой строк и/или столбцов матрицы A добиваемся, чтобы $a_{11} \neq 0$, а затем применяем указанное выше преобразование.

Шаг 2. С помощью $a'_{11} = 1$ получаем нули в первом столбце посредством следующих элементарных преобразований:

- ко второй строке прибавляем первую строку, предварительно умноженную на $-a_{21}$;
- к третьей строке прибавляем первую строку, предварительно умноженную на $-a_{31}$, и так далее;

–;

– к m -й строке прибавляем первую строку, предварительно умноженную на $-a_{m1}$.

В результате шага 2 получаем матрицу

$$\left(\begin{array}{c|ccc} 1 & a'_{12} & \dots & a'_{1n} \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array} \right).$$

Шаг 3. Получаем нули в первой строке посредством следующих элементарных преобразований:

– из второго столбца вычитаем первый, предварительно умноженный на $-a'_{12}$;

– из третьего столбца вычитаем первый, предварительно умноженный на $-a'_{13}$;

–;

– из n -го столбца вычитаем первый, предварительно умноженный на $-a'_{1n}$.

В результате шага 3 получаем

$$\left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array} \right).$$

Повторяя указанные выше преобразования, в итоге приходим к матрице вида

$$\left(\begin{array}{cccccc} 1 & 0 & \dots & & 0 & 0 \\ 0 & 1 & \dots & & 0 & 0 \\ & & \ddots & & & \\ \dots & \dots & & 1 & \dots & 0 & 0 \\ 0 & 0 & \dots & & \ddots & 0 & 0 \\ 0 & 0 & \dots & & & 0 & 0 \end{array} \right),$$

и число единиц на главной диагонали равно рангу $rk(A)$.

Пример I.3.8. Применим метод элементарных преобразований к матрице

$$A = \begin{pmatrix} -1 & 2 & 1 & 0 \\ 1 & -2 & -1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Прибавляя первую строку ко второй и четвертой строкам, получим матрицу

$$\begin{pmatrix} -1 & 2 & 1 & 0 \\ 1 & -2 & -1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} -1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 3 & 2 & 1 \end{pmatrix}.$$

Далее переставляем вторую и четвертую строки:

$$\begin{pmatrix} -1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 3 & 2 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} -1 & 2 & 1 & 0 \\ 0 & 3 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Затем получаем нули в первой строке, прибавляя ко второму столбцу первый, предварительно умноженный на второй, и к третьему столбцу прибавляем первый:

$$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 3 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Далее поменяем местами вторую и третью строки и дополнительно умножаем первый столбец на -1 , получим матрицу

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

Подматрица A' имеет вид

$$A' = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 2 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Получаем нули в первом столбце:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \end{pmatrix},$$

а затем в первой строке:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Умножаем вторую строку на -1 и затем получаем нули во второй строке:

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

Итак, в итоге получаем матрицу

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

Следовательно, $rk(A) = 3$.

В заключение применим метод элементарных преобразований к матрице из примера 1.3.5 и сравним трудоемкость вычислений с методом окаймляющих миноров.

Пример 1.3.9. Найдём методом элементарных преобразований ранг матрицы

$$A = \begin{pmatrix} 1 & -1 & 0 & 1 & 0 \\ 2 & -2 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 1 \\ -1 & 1 & 0 & -1 & 0 \end{pmatrix}.$$

Имеем

$$\begin{aligned}
 & \begin{pmatrix} 1 & -1 & 0 & 1 & 0 \\ 2 & -2 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 1 \\ -1 & 1 & 0 & -1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 1 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \\
 & \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 1 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -2 & 1 \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \\
 & \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \\
 & \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.
 \end{aligned}$$

Получим тот же ответ, что и в примере I.3.5 методом окаймляющих миноров, но трудоемкость вычислений существенно меньше, чем в том примере.

Лекция 4. ОБРАТНАЯ МАТРИЦА И МАТРИЧНЫЕ УРАВНЕНИЯ

Определение I.4.1. Пусть $A = (a_{ij}) \in M_n(\mathbb{R})$ – квадратная матрица порядка n с числовыми элементами. Будем говорить, что квадратная матрица $B = (b_{ij}) \in M_n(\mathbb{R})$ порядка n с числовыми элементами является обратной матрицей к матрице A , если $A \cdot B = B \cdot A = E$, где

$$E = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 1 \end{pmatrix} -$$

единичная матрица порядка n . Условимся обозначать обратную матрицу B к матрице A через A^{-1} .

Следующая теорема дает критерий существования обратной матрицы.

Теорема I.4.2. Матрица $A = (a_{ij}) \in M_n(\mathbb{R})$ имеет обратную матрицу тогда и только тогда, когда $|A| \neq 0$.

Необходимость. Если матрица A^{-1} существует, то из $A^{-1} \cdot A = A \cdot A^{-1} = E$ следует по теореме I.2.24, что

$$|A^{-1}| \cdot |A| = |E| = 1.$$

Следовательно, $|A| \neq 0$.

Достаточность. Пусть $|A| = \alpha \neq 0$.

Рассмотрим алгебраические дополнения A_{ij} к элементам a_{ij} матрицы A и составим матрицу A^* , которая называется присоединенной к матрице следующим образом:

$$A^* = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}.$$

Другими словами, для составления матрицы A^* нужно сначала каждый элемент a_{ij} матрицы A заменить на его алгебраическое дополнение A_{ij} , а затем полученную матрицу транспонировать.

Теперь обратная матрица A^{-1} вычисляется следующим образом:

$$B = \frac{1}{\alpha} A^* = \begin{pmatrix} \frac{1}{\alpha} A_{11} & \frac{1}{\alpha} A_{21} & \dots & \frac{1}{\alpha} A_{n1} \\ \frac{1}{\alpha} A_{12} & \frac{1}{\alpha} A_{22} & \dots & \frac{1}{\alpha} A_{n2} \\ \dots & \dots & \dots & \dots \\ \frac{1}{\alpha} A_{1n} & \frac{1}{\alpha} A_{2n} & \dots & \frac{1}{\alpha} A_{nn} \end{pmatrix}.$$

Действительно, рассмотрим матрицы $C = AB$ и $D = BA$.

Любой элемент этих матриц $c_{ij} (d_{ij})$ вне главной диагонали, т.е. $i \neq j$, равен нулю, поскольку после вынесения общего множителя $\frac{1}{\alpha}$ за скобки элемент $c_{ij} (d_{ij})$ равен сумме произведений элементов i -й строки (j -го столбца) матрицы A на алгебраические дополнения к соответствующим элементам j -й строки (i -го столбца), которая по следствию 2 из теоремы Лапласа равна нулю.

Любой диагональный элемент этих матриц $c_{ii} (d_{ii})$ равен единице, поскольку после вынесения общего множителя $\frac{1}{\alpha}$ за скобки элемент $c_{ii} (d_{ii})$ равен сумме произведений элементов i -й строки (j -го столбца) матрицы A на их алгебраические дополнения, которая по следствию 1 из теоремы Лапласа равна определителю матрицы A , который равен α , т.е. $c_{ii} = d_{ii} = 1$ для любого $i = \overline{1, n}$. Следовательно, $C = D = E$, и таким образом $B = A^{-1}$.

Определение I.4.3. Будем называть матрицу $A = (a_{ij}) \in M_n(\mathbb{R})$ невырожденной, если $|A| \neq 0$.

Пример I.4.4. Пусть $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Шаг 1. Вычисляем $|A| = ad - bc = \alpha$.

Если $\alpha \neq 0$, то переходим к следующему шагу.

Шаг 2. Находим присоединенную матрицу

$$\begin{aligned} 2a) \quad A_{11} &= d, \quad A_{12} = -c, \\ A_{21} &= -b, \quad A_{22} = a; \end{aligned}$$

$$2b) \quad A^* = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Шаг 3. Находим обратную матрицу

$$A^{-1} = \frac{1}{\alpha} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Рассмотрим теперь методы вычисления обратной матрицы.

Метод 1. Этот метод изложен в доказательстве теоремы I.4.2, поэтому нет необходимости вновь излагать его.

Метод 2. Пусть $A \in M_n(\mathbb{R})$, и рассмотрим расширенную матрицу A' , полученную присоединением к матрице A единичной матрицы E порядка n , т.е. A' имеет вид $(A|E)$.

Шаг 1. Применяем к матрице A' метод элементарных преобразований вычисления ранга матрицы, причем разрешается использовать только элементарные преобразования строк матрицы A' .

Если ранг матрицы A' меньше n (т.е. матрица A приводится к диагональному виду с числом единиц, меньшим n), то $|A| = 0$ и обратной матрицы к матрице A не существует. В противном случае переходим к следующему шагу.

Шаг 2. Пусть ранг матрицы A' равен n , т.е. матрица A' приводится к виду

$$\left\langle \begin{array}{cccc|c} 1 & 0 & \dots & 0 & \\ 0 & 1 & \dots & 0 & \\ \dots & \dots & \dots & \dots & \\ 0 & 0 & \dots & 1 & \end{array} A'' \right\rangle.$$

Тогда $A^{-1} = A''$.

Метод 3 (метод Фаддеева). Пусть $A \in M_n(\mathbb{R})$, и рассмотрим последовательность матриц B_i , $i = \overline{1, n}$, вычисленных следующим образом:

$$\begin{aligned} A_1 &= A; \operatorname{Tr} A_1 = p_1; A_1 - p_1 E = B_1, \\ B_1 A &= A_2; \frac{1}{2} \operatorname{Tr} A_2 = p_2; A_2 - p_2 E = B_2, \\ B_2 A &= A_3; \frac{1}{3} \operatorname{Tr} A_3 = p_3; A_3 - p_3 E = B_3, \\ &\dots\dots\dots \\ B_{n-1} A &= A_n; \frac{1}{n} \operatorname{Tr} A_n = p_n; A_n - p_n E = B_n, \end{aligned}$$

где $\operatorname{Tr} A_i$ – след матрицы A_i , т.е. сумма диагональных элементов A_i , $i = \overline{1, n}$.

Тогда, если матрица A – невырожденная, то $A^{-1} = \frac{1}{p_n} B_{n-1}$.

Замечание. Контроль правильности вычислений ведется по $B_n = 0$.

Пример I.4.5. Пусть матрица

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ -4 & -2 & 1 \end{pmatrix}.$$

Вычислим A^{-1} согласно методу 1.

Шаг 1. Находим $|A| = -1 \neq 0$.

Шаг 2. а) Вычисляем A_{ij} для $i, j = \overline{1, n}$:

$$A_{11} = 3; A_{12} = -2; A_{13} = 8,$$

$$A_{21} = -2; A_{22} = 1; A_{23} = -6,$$

$$A_{31} = 0; A_{32} = 0; A_{33} = -1.$$

б) Составляем матрицу A^* :

$$A^* = \begin{pmatrix} 3 & -2 & 0 \\ -2 & 1 & 0 \\ 8 & -6 & -1 \end{pmatrix}.$$

Шаг 3. Находим обратную матрицу:

$$A^{-1} = -A^* = \begin{pmatrix} -3 & 2 & 0 \\ 2 & -1 & 0 \\ -8 & 6 & 1 \end{pmatrix}.$$

Проверка

$$A \cdot A^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Пример I.4.6. Пусть матрица A из примера I.4.5, т.е.

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ -4 & -2 & 1 \end{pmatrix}.$$

Вычислим A^{-1} согласно методу 2.

Шаг 1. Составляем матрицу

$$A' = \left\langle \begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 3 & 0 & 0 & 1 & 0 \\ -4 & -2 & 1 & 0 & 0 & 1 \end{array} \right\rangle.$$

Шаг 2. Выполняем следующую цепочку элементарных преобразований строк A' :

$$\begin{aligned}
 &\left\langle \begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 3 & 0 & 0 & 1 & 0 \\ -4 & -2 & 1 & 0 & 0 & 1 \end{array} \right\rangle \rightarrow \left\langle \begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -2 & 1 & 0 \\ 0 & 6 & 1 & 4 & 0 & 1 \end{array} \right\rangle \rightarrow \\
 &\rightarrow \left\langle \begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & -8 & 6 & 1 \end{array} \right\rangle \rightarrow \left\langle \begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 2 & -1 & 0 \\ 0 & 0 & 1 & -8 & 6 & 1 \end{array} \right\rangle \rightarrow \\
 &\rightarrow \left\langle \begin{array}{ccc|ccc} 1 & 0 & 0 & -3 & 2 & 0 \\ 0 & 1 & 0 & 2 & -1 & 0 \\ 0 & 0 & 1 & -8 & 6 & 1 \end{array} \right\rangle.
 \end{aligned}$$

Шаг 3. Таким образом,

$$A^{-1} = \begin{pmatrix} -3 & 2 & 0 \\ 2 & -1 & 0 \\ -8 & 6 & 1 \end{pmatrix}.$$

Проверка

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

Пример I.4.7. Пусть матрица A из примеров I.4.5 и I.4.6, т.е.

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ -4 & -2 & 1 \end{pmatrix}.$$

Шаг 1. Находим матрицу B_1 :

$$A_1 = A; \operatorname{Tr} A_1 = 5; A_1 - 5E = \begin{pmatrix} -4 & 2 & 0 \\ 2 & -2 & 0 \\ -4 & -2 & -4 \end{pmatrix} = B_1.$$

Шаг 2. Находим матрицу B_2 :

$$B_1 A = A_2 = \begin{pmatrix} 0 & -2 & 0 \\ -2 & -2 & 0 \\ 8 & -6 & -4 \end{pmatrix}; \frac{1}{3} \operatorname{Tr} A_2 = -3; A_2 + 3E = \begin{pmatrix} 3 & -2 & 0 \\ -2 & 1 & 0 \\ 8 & -6 & -1 \end{pmatrix} = B_2.$$

Шаг 3. Находим матрицу B_3 :

$$B_2 A = A_3 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}; \quad \frac{1}{3} \text{Tr } A_3 = -1; \quad A_3 + E = B_3 = 0.$$

Таким образом, $A^{-1} = -B_2 = \begin{pmatrix} -3 & 2 & 0 \\ 2 & -1 & 0 \\ -8 & 6 & 1 \end{pmatrix}.$

Проверка

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

Укажем применение обратной матрицы к решению матричных уравнений.

Рассмотрим матричные уравнения вида

$$AX = B \quad (1) \quad \text{и} \quad YA = B. \quad (2)$$

Если матрица A невырожденная, то решения этих уравнений находятся так: умножаем слева обе части уравнения (1) на матрицу A^{-1} , получаем

$$A^{-1}(AX) = A^{-1}B,$$

$$(A^{-1}A)X = A^{-1}B,$$

$$EX = A^{-1}B, \quad X = A^{-1}B.$$

Для решения уравнения (2) действуем аналогично справа, получаем

$$(YA)A^{-1} = BA^{-1},$$

$$Y(AA^{-1}) = BA^{-1},$$

$$YE = BA^{-1}, \quad Y = BA^{-1}.$$

Пример I.4.8. Пусть даны уравнения

$$AX = B \quad \text{и} \quad YA = B,$$

где $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, $B = \begin{pmatrix} -3 & 1 \\ -4 & 2 \end{pmatrix}$, X и Y – неизвестные матрицы из $M_2(\mathbb{R})$,
тогда

$$X = A^{-1}B = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} -3 & 1 \\ -4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -2 & 0 \end{pmatrix},$$

$$Y = BA^{-1} = \begin{pmatrix} -3 & 1 \\ -4 & 2 \end{pmatrix} \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 11 & -5 \\ 16 & -10 \end{pmatrix}.$$

Вопросы для самоконтроля

1. Верно ли, что произведение матриц $A = (a_{ij})$ и $B = (b_{ij})$, $i, j = \overline{1, n}$, состоит из произведений соответствующих элементов, т.е. $C = (c_{ij})$ есть произведение матриц $C = A \cdot B$, если $c_{ij} = a_{ij} \cdot b_{ij}$, $i, j = \overline{1, n}$?

2. Вычислите определитель матрицы

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 1 & 1 \end{pmatrix}.$$

3. Дайте словесную формулировку вычисления определителя второго порядка общего вида

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \dots$$

4. Как изменится определитель n -го порядка, если транспонировать его матрицу?
5. Как изменится определитель n -го порядка, если поменять местами его две строки?
6. Чему равен определитель n -го порядка с двумя пропорциональными столбцами?
7. Как изменится определитель n -го порядка, если к первой его строке прибавить сумму всех остальных строк?
8. Верно ли, что определитель суммы матриц равен сумме определителей слагаемых, т.е. $|A + B| = |A| + |B|$?
9. Верно ли, что определитель произведения матриц равен произведению определителей сомножителей, т.е. $|A \cdot B| = |A| \cdot |B|$?
10. Верно ли, что диагональные матрицы n -го порядка перестановочны, т.е. если $A = \text{diag}(a_1, \dots, a_n)$ и $B = \text{diag}(b_1, \dots, b_n)$, то $A \cdot B = B \cdot A$?
11. Найдите ошибку в «доказательстве» следующего утверждения: «любая ненулевая матрица порядка n имеет ранг n ».

Доказательство. По определению ранга матрицы ранг данной матрицы порядка n равен наибольшему из порядков ненулевых миноров этой матрицы. Поскольку матрица ненулевая, то она сама яв-

ляется ненулевым минором порядка n и, следовательно, ранг данной матрицы равен n как наибольшему из порядков ненулевых миноров этой матрицы.

12. Сформулируйте метод окаймляющих миноров вычисления ранга матрицы.
13. Сформулируйте метод элементарных преобразований вычисления ранга матрицы.
14. Чему равен ранг невырожденной матрицы порядка n ?
15. Если матрица порядка n имеет ранг n , то означает ли это, что матрица невырожденная?
16. Верно ли, что обратная матрица к невырожденной матрице также является невырожденной матрицей?
17. Сформулируйте метод вычисления обратной матрицы через вычисление присоединенной матрицы.
18. Сформулируйте метод вычисления обратной матрицы через присоединение единичной матрицы.
19. Сформулируйте метод Фаддеева вычисления обратной матрицы.
20. Верно ли, что если матричное уравнение $AX = B$ имеет решение $X = C$, то матрица A невырожденная?

МОДУЛЬ II

Лекция 1. ПОНЯТИЕ ВЕКТОРНОГО ПРОСТРАНСТВА И ПРОСТЕЙШИЕ СВОЙСТВА

Понятие векторного пространства в общем виде (забегая вперед, «для произвольного поля скаляров») мы сможем дать только после рассмотрения различных алгебраических структур, в частности поля, что и будет сделано позже. Здесь это понятие вводится для частного случая, когда скаляры – действительные числа.

Определение II.1.1. Будем говорить, что множество V является векторным пространством над $\mathbb{R}$, если выполняются следующие условия:

1) задано отображение $f_1: V^2 \rightarrow V$, т.е. для любой упорядоченной пары $x, y \in V$ существует однозначно определенный $z \in V$ такой, что $z = f_1(x, y)$, такой элемент z назовем суммой элементов x, y и обозначим $z = x + y$;

2) задано отображение $f_2: \mathbb{R} \times V \rightarrow V$, т.е. для любой пары элементов $\alpha \in \mathbb{R}$ и $x \in V$ существует однозначно определенный $y \in V$ такой, что $y = f_2(\alpha, x)$, такой элемент y назовем произведением α на x и обозначим $y = \alpha \cdot x$;

3) отображения f_1 и f_2 удовлетворяют следующим условиям:

3.1) $(x + y) + z = x + (y + z)$ для любых $x, y, z \in V$;

3.2) существует элемент $0 \in V$ такой, что $0 + x = x + 0$ для любого $x \in V$, такой элемент $0 \in V$ назовем нуль-вектором и обозначим $\bar{0}$;

3.3) для любого элемента $x \in V$ существует элемент $y \in V$ такой, что $x + y = y + x = \bar{0}$, такой элемент y назовем противоположным к x и обозначим $y = -x$;

3.4) $x + y = y + x$ для любых $x, y \in V$;

3.5) $(\alpha + \beta)x = \alpha x + \beta x$ для любых $\alpha, \beta \in \mathbb{R}$ и любого $x \in V$;

3.6) $\alpha(x+y) = \alpha x + \alpha y$ для любого $\alpha \in \mathbb{R}$ и любых $x, y \in V$;

3.7) $(\alpha\beta)x = \alpha(\beta x)$ для любых $\alpha, \beta \in \mathbb{R}$ и любого $x \in V$;

3.8) $1 \cdot x = x$ для любого $x \in V$.

Условимся называть элементы векторного пространства V векторами, а числа из $\mathbb{R}$ – скалярами.

Пример II.1.2. Рассмотрим множество V всех направленных векторов-отрезков на плоскости. Непосредственно проверяется, что относительно обычных операций сложения векторов-отрезков и их умножения на число множество V является векторным пространством над $\mathbb{R}$.

Пример II.1.3. Рассмотрим множество $\mathbb{R}^{(n)}$ всевозможных строк $(\alpha_1, \alpha_2, \dots, \alpha_n)$, где $\alpha_i \in \mathbb{R}, i = \overline{1, n}$. Зададим сложение строк $x = (\alpha_1, \dots, \alpha_n)$ и $y = (\beta_1, \dots, \beta_n)$ следующим образом:

$$x + y = (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n).$$

Далее для любого $\alpha \in \mathbb{R}$ и любой строки $x = (\alpha_1, \dots, \alpha_n)$ определим произведение α на x следующим образом:

$$\alpha \cdot x = (\alpha\alpha_1, \dots, \alpha\alpha_n).$$

Непосредственно проверяется, что для введенных операций выполняются свойства 3.1 – 3.8, и, следовательно, $\mathbb{R}^{(n)}$ есть векторное пространство, которое называется арифметическим векторным пространством n -мерных векторов-строк.

Пример II.1.4. Рассмотрим множество $\mathbb{R}_{(n)}$ всевозможных столбцов

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix},$$

где $\alpha_i \in \mathbb{R}, i = \overline{1, n}$. Задавая операции сложения столбцов и их умножения на число аналогично соответствующим операциям для строк, убеждаемся, что $\mathbb{R}_{(n)}$ есть векторное пространство, которое называется арифметическим векторным пространством n -мерных векторов-столбцов.

Пример II.1.5. Рассмотрим множество $M_{m,n}(\mathbb{R})$ всевозможных матриц $A = (a_{ij})$ размера $m \times n$ с элементами $a_{ij} \in \mathbb{R}$, $i = \overline{1, m}$; $j = \overline{1, n}$.

Зададим сложение матриц $A = (a_{ij})$ и $B = (b_{ij})$ так:

$$A + B = C = (c_{ij}),$$

где $c_{ij} = a_{ij} + b_{ij}$ для $i = \overline{1, m}$, $j = \overline{1, n}$.

Зададим умножение матрицы $A = (a_{ij})$ на число $\alpha \in \mathbb{R}$ так:

$$\alpha \cdot A = (\alpha \cdot a_{ij})$$

для $i = \overline{1, m}$, $j = \overline{1, n}$.

Поскольку операции над матрицами сводятся к операциям над числами для каждой позиции (ij) , то получаем векторное пространство действительных матриц, которое будем обозначать $\mathbb{R}_{m,n}$.

Рассмотрим теперь простейшие свойства векторных пространств.

Свойство 1. Пусть V – векторное пространство и $x \in V$ – произвольный вектор. Тогда

$$0 \cdot x = \bar{0}.$$

Другими словами, произведение числа 0 на произвольный вектор x есть нуль-вектор.

Доказательство. Имеем равенства

$$0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x.$$

Прибавляя к обеим частям равенства вектор $-(0 \cdot x)$, получим

$$-(0 \cdot x) + 0 \cdot x = -(0 \cdot x) + [0 \cdot x + 0 \cdot x].$$

Левая часть равна

$$-(0 \cdot x) + 0 \cdot x = \bar{0}.$$

Правая часть равна

$$-(0 \cdot x) + [0 \cdot x + 0 \cdot x] = [-(0 \cdot x) + 0 \cdot x] + 0 \cdot x = \bar{0} + 0 \cdot x = 0 \cdot x.$$

Следовательно, $0 \cdot x = \bar{0}$.

Свойство 2. Пусть V – векторное пространство и $\bar{0} \in V$ – нуль-вектор. Тогда для любого числа $\alpha \in \mathbb{R}$

$$\alpha \cdot \bar{0} = \bar{0}.$$

Доказательство. Имеем равенства

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

Прибавляя к обеим частям равенства вектор $-(\alpha \cdot \bar{0})$, получим

$$-(\alpha \cdot \bar{0}) + \alpha \cdot \bar{0} = -(\alpha \cdot \bar{0}) + [\alpha \cdot \bar{0} + \alpha \cdot \bar{0}].$$

Левая часть равна

$$-(\alpha \cdot \bar{0}) + \alpha \cdot \bar{0} = \bar{0}.$$

Правая часть равна

$$\begin{aligned} -(\alpha \cdot \bar{0}) + [\alpha \cdot \bar{0} + \alpha \cdot \bar{0}] &= [-(\alpha \cdot \bar{0}) + \alpha \cdot \bar{0}] + \alpha \cdot \bar{0} = \\ &= \bar{0} + \alpha \cdot \bar{0} = \alpha \cdot \bar{0}. \end{aligned}$$

Следовательно, $\alpha \cdot \bar{0} = \bar{0}$.

Свойство 3. Пусть V – векторное пространство, $x \in V$ – произвольный вектор и $\alpha \in \mathbb{R}$ – произвольное число. Тогда, если $\alpha \cdot x = \bar{0}$, то $x = \bar{0}$ или $\alpha = 0$.

Доказательство. Пусть $\alpha \neq 0$, тогда, умножая обе части равенства $\alpha \cdot x = \bar{0}$ на $1/\alpha$, получаем

$$1/\alpha (\alpha \cdot x) = 1/\alpha \cdot \bar{0}.$$

Левая часть равна

$$1/\alpha (\alpha \cdot x) = (1/\alpha \cdot \alpha) x = 1 \cdot x = x.$$

Правая часть по свойству 2 равна

$$1/\alpha \cdot \bar{0} = \bar{0}.$$

Следовательно, $\alpha = 0$ или $x = \bar{0}$.

Свойство 4. Пусть V – векторное пространство, $x \in V$ – произвольный вектор. Тогда $(-1) \cdot x = -x$.

Доказательство. Имеем

$$x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = [1 + (-1)] \cdot x = 0 \cdot x = \bar{0}.$$

Тогда вектор $(-1) \cdot x$ является противоположным к вектору x и, следовательно,

$$(-1)x = -x.$$

В дальнейшем условимся обозначать вектор $x + (-y)$ как $x - y$.

Определение II.1.6. Пусть V – векторное пространство, $x \in V$ – произвольный вектор и

$$x_1, x_2, \dots, x_k \quad (1)$$

произвольная система векторов из V . Будем говорить, что вектор x линейно выражается через систему (1), и обозначим это так: $x \rightarrow (1)$, если

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k,$$

где $\alpha_i \in \mathbb{R}$.

Правую часть этого равенства будем называть линейной комбинацией векторов $x_1, x_2, \dots, x_k$, а числа $\alpha_1, \dots, \alpha_k$ – коэффициентами этой линейной комбинации.

Пример II.1.7. В арифметическом векторном пространстве строк $\mathbb{R}^{(4)}$ рассмотрим систему векторов-строк:

$$x_1, x_2, x_3, \quad (2)$$

где $x_1 = (1, -1, 0, 1)$, $x_2 = (2, 1, -2, 0)$, $x_3 = (-1, 1, -1, 1)$ и пусть $x = (2, 4, -3, -4)$.

Легко проверить, что вектор x линейно выражается через систему (2) посредством следующей линейной комбинации:

$$x = -3x_1 + 2x_2 - x_3.$$

Приведем еще два примера векторных пространств.

Пример II.1.8. Рассмотрим множество всевозможных функций (отображений) из $\mathbb{R}$ в $\mathbb{R}$, обозначим его как

$$\mathbb{R}^{\mathbb{R}} = \{f \mid f: \mathbb{R} \rightarrow \mathbb{R}\}.$$

Определим для любых функций $f_1, f_2 \in \mathbb{R}^{\mathbb{R}}$ их сумму $f_1 + f_2$ следующим образом:

$$(f_1 + f_2)(x) = f_1(x) + f_2(x)$$

для любого $x \in \mathbb{R}$.

Далее определим $\alpha \cdot f$ для любой $f \in \mathbb{R}^{\mathbb{R}}$ и любого $\alpha \in \mathbb{R}$ так:

$$(\alpha \cdot f)(x) = \alpha \cdot f(x)$$

для любого $x \in \mathbb{R}$.

Непосредственно проверяется, что все пункты определения векторного пространства выполняются для $\mathbb{R}^{\mathbb{R}}$ относительно введенных выше операций.

Пример II.1.9. Рассмотрим множество всевозможных многочленов от одного неизвестного x с действительными коэффициентами, обозначим его $\mathbb{R}[x]$. Для любых многочленов $f_1, f_2 \in \mathbb{R}[x]$ определим сумму $f_1 + f_2$ следующим образом.

Если

$$f_1 = a_n x^n + \dots + a_1 x + a_0,$$

$$f_2 = b_m x^m + \dots + b_1 x + b_0,$$

и пусть для определенности $n \geq m$, то запишем

$$f_2 = b_n x^n + \dots + b_1 x + b_0,$$

определив $b_i = 0$ для $m < i \leq n$.

Тогда зададим

$$f_1 + f_2 = (a_n + b_n)x^n + \dots + (a_1 + b_1)x + a_0 + b_0.$$

Для любых $f \in \mathbb{R}[x]$ и $\alpha \in \mathbb{R}$ определим

$$\alpha \cdot f = \alpha a_n x^n + \dots + \alpha a_1 x + \alpha a_0,$$

если

$$f = a_n x^n + \dots + a_1 x + a_0.$$

Непосредственно проверяется, что все пункты определения векторного пространства выполняются для $\mathbb{R}[x]$ относительно введенных выше операций.

Пример II.1.10. Пусть $\mathbb{R}^{(n)}[x]$ есть множество всевозможных многочленов от неизвестной x с коэффициентами из $\mathbb{R}$ степени, не превосходящей n . Если в $\mathbb{R}^{(n)}[x]$ задать такие же операции сложения многочленов и их умножения на скаляр, то аналогично $\mathbb{R}[x]$ проверяется, что $\mathbb{R}^{(n)}[x]$ есть векторное пространство над $\mathbb{R}$.

Лекция 2. ЛИНЕЙНАЯ ЗАВИСИМОСТЬ ВЕКТОРОВ

Определение II.2.1. Пусть V – векторное пространство, и рассмотрим произвольную систему векторов из V :

$$(I) \ x_1, x_2, \dots, x_k.$$

Будем говорить, что система (I) линейно зависима, если выполняются следующие условия:

а) если $k = 1$ (т.е. система состоит из одного вектора), то единственный вектор системы (I) есть нулевой вектор;

б) если $k > 1$, то существует вектор x_i из системы (I) такой, что x_i линейно выражается через подсистему (II) системы (I), где

$$(II) = (I) \setminus \{x_i\}.$$

Другими словами, существует вектор данной системы векторов, который линейно выражается через остальные векторы этой системы векторов.

Пример II.2.2. В арифметическом векторном пространстве $\mathbb{R}^{(4)}$ рассмотрим систему векторов-строк:

$$(I) \ x_1, x_2, x_3,$$

где

$$x_1 = (-2, 1, -1, 3), \ x_2 = (1, -2, 2, -1), \ x_3 = (-1, -4, 4, 3).$$

Система (I) линейно зависима, поскольку существует вектор из системы (I), а именно вектор x_3 , который линейно выражается через остальные векторы x_1 и x_2 следующим образом: $x_3 = 2x_1 + 3x_2$.

Определение II.2.3. Система векторов из векторного пространства V называется линейно независимой, если эта система векторов не является линейно зависимой.

Пример II.2.4. Любая система векторов, состоящая из одного ненулевого вектора, будет линейно независимой согласно пункту «а» из определения II.2.1 линейной зависимости системы векторов.

Сформулируем теперь необходимое и достаточное условие линейной независимости системы векторов.

Теорема II.2.5 (критерий линейной независимости). Пусть V – векторное пространство. Система (I) $x_1, \dots, x_n$ векторов пространства V является линейно независимой тогда и только тогда, когда выполнено условие

(*) для любой линейной комбинации векторов из системы (I), если $\alpha_1 x_1 + \dots + \alpha_k x_k = \bar{0}$, то $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$.

Доказательство. Необходимость. Пусть система (I) линейно независима, и покажем, что тогда условие (*) выполняется.

Для $k=1$ утверждение очевидно, поскольку в этом случае из линейной независимости системы (I) следует, что $x_1 \neq \bar{0}$, и по свойству 3 (с. 56) из $\alpha_1 x_1 = \bar{0}$ следует, что $\alpha_1 = 0$. Поэтому можем считать, что $k \geq 2$.

Предположим, что существует линейно независимая система (I) векторов из V с $k \geq 2$, для которой условие (*) не выполняется, т.е. $\alpha_1 x_1 + \dots + \alpha_k x_k = \bar{0}$ и для некоторого $i \in \{1, \dots, k\}$ имеем $\alpha_i \neq 0$. Без ограничения общности можем считать, что $i=1$ (при необходимости можем перенумеровать векторы из системы (I)), т.е. $\alpha_1 \neq 0$. Тогда имеем

$$\alpha_1 x_1 = -\alpha_2 x_2 - \dots - \alpha_k x_k,$$

откуда получаем

$$x_1 = -\frac{\alpha_2}{\alpha_1} x_2 - \dots - \frac{\alpha_k}{\alpha_1} x_k,$$

и, следовательно, система (I) линейно зависима, что противоречит предположению. Необходимость условия (*) доказана.

Достаточность. Пусть для системы (I) выполняется условие (*), и покажем, что тогда (I) есть линейно независимая система векторов из V .

Предположим противное: пусть система (I) линейно зависима. Если $k=1$, то это означает, что $x_1 = \bar{0}$, и тогда из $\alpha_1 x_1 = \bar{0}$ не следует, что $\alpha_1 = 0$, т.е. условие (*) не выполняется, что противоречит предположению о том, что для системы (I) выполняется условие (*). Тогда можем считать, что $k \geq 2$.

Из линейной зависимости системы (I) в этом случае следует, что существует $i \in \{1, \dots, k\}$ такой, что вектор x_i линейно выражается через остальные векторы системы (I). Без ограничения общности можем считать, что $i = 1$, тогда

$$x_1 = \alpha_2 x_2 + \dots + \alpha_k x_k.$$

Отсюда имеем

$$1 \cdot x_1 - \alpha_2 x_2 - \dots - \alpha_k x_k = \bar{0}.$$

Получили линейную комбинацию векторов из системы (I), равную нуль-вектору, но, по крайней мере, один из коэффициентов этой линейной комбинации, а именно $\alpha_1 = 1 \neq 0$. Это означает, что условие (*) не выполняется, что противоречит предположению. Таким образом, доказана достаточность условия (*).

Следствие. Система (I) $x_1, \dots, x_k$ векторов из V является линейно зависимой тогда и только тогда, когда существует линейная комбинация векторов из системы (I), которая равна нуль-вектору, причем хотя бы один из коэффициентов этой линейной комбинации отличен от нуля.

Доказательство получается из записи отрицания условия (*).

Отметим ряд свойств векторных пространств, относящихся к линейной зависимости и линейной независимости систем векторов (нумерация этих свойств ведется с учетом свойств 1 – 4, приведенных ранее).

Свойство 5. Если система векторов (I) содержит линейно зависимую подсистему (II), то система (I) является линейно зависимой.

Доказательство. Пусть система (I) $x_1, x_2, \dots, x_k$ содержит линейно зависимую подсистему, которую без ограничения общности (перенумеровав, если необходимо) можем записать так:

$$(II) \quad x_1, x_2, \dots, x_\ell,$$

где $\ell \leq k$.

Тогда в подсистеме (II) существует вектор x_i , $i \leq \ell$, такой, что

$$x_i = \alpha_1 x_1 + \dots + \alpha_i x_{i-1} + \alpha_{i+1} + \dots + \alpha_\ell x_\ell.$$

Если в правой части этого равенства добавить слагаемые с нулевыми коэффициентами (т.е. слагаемые, равные нуль-вектору), то вектор x_i от этого не изменится. Следовательно,

$$x_i = \alpha_1 x_1 + \dots + \alpha_{i-1} x_{i-1} + \alpha_{i+1} x_{i+1} + \dots + \alpha_\ell x_\ell + \\ + 0 \cdot x_{\ell+1} + \dots + 0 \cdot x_k,$$

и тогда система (I) линейно зависима.

Свойство 6. Если в системе векторов (I) есть нуль-вектор, то эта система линейно зависима.

Доказательство. Если система (I) состоит из одного нуль-вектора, то (I) есть линейно зависимая система векторов согласно определению II.2.1. Если в системе (I), кроме нуль-вектора, есть и другие векторы, то, взяв линейную комбинацию этих ненулевых векторов с нулевыми коэффициентами, получим, что нуль-вектор является линейной комбинацией остальных векторов системы (I). Это означает, что система (I) линейно зависима.

Свойство 7. Если система векторов (I) линейно независима, то любая ее подсистема также линейно независима.

Доказательство. От противного, если существует подсистема (II) системы (I), которая линейно зависима, то по свойству 5 система (I) является линейно зависимой, противоречие.

Свойство 8. Если система векторов (I) $x_1, x_2, \dots, x_k$ линейно независима, а система векторов (II) $x_1, x_2, \dots, x_k, y$ линейно зависима, то вектор y линейно выражается через систему (I) (будем обозначать это так: $y \rightarrow (I)$).

Доказательство. Из линейной зависимости системы (II) следует, что существуют числа $\alpha_1, \alpha_2, \dots, \alpha_k, \alpha_{k+1}$, среди которых хотя бы одно отлично от нуля, такие, что

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k + \alpha_{k+1} y = \bar{0}.$$

Возможны два случая.

Случай 1. $\alpha_{k+1} = 0$.

Тогда

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k = \bar{0},$$

и из линейной независимости системы (II) следует, что

$$\alpha_1 = \alpha_2 = \dots = \alpha_k = 0,$$

т.е. все числа $\alpha_i, i = \overline{1, k+1}$, равны нулю, что противоречит наличию хотя бы одного ненулевого числа среди $\alpha_i, i = \overline{1, k+1}$. Таким образом, случай 1 не выполняется.

Случай 2. $\alpha_{k+1} \neq 0$. Тогда имеем

$$\alpha_{k+1}y = -\alpha_1x_1 - \alpha_2x_2 - \dots - \alpha_kx_k,$$

откуда получаем

$$y = -\frac{\alpha_1}{\alpha_{k+1}}x_1 - \frac{\alpha_2}{\alpha_{k+1}}x_2 - \dots - \frac{\alpha_k}{\alpha_{k+1}}x_k.$$

Рассмотрим произвольную систему векторов

$$(I) \ a_1, a_2, \dots, a_k$$

и укажем способ нахождения ответа на вопрос, является ли данная система векторов линейно независимой или нет.

Пусть

$$\alpha_1a_1 + \alpha_2a_2 + \dots + \alpha_k a_k = \vec{0}$$

есть произвольная линейная комбинация векторов системы (I). Рассматривая векторы системы (I) как упорядоченные наборы чисел, по каждой позиции этих наборов получаем слева линейные комбинации коэффициентов $\alpha_1, \alpha_2, \dots, \alpha_k$, а справа – нули. Отсюда, если рассматривать коэффициенты $\alpha_1, \alpha_2, \dots, \alpha_k$ как неизвестные, то получаем систему линейных уравнений относительно этих неизвестных. Если эта система уравнений имеет только нулевое решение, т.е. $\alpha_1 = 0, \alpha_2 = 0, \dots, \alpha_k = 0$, то по теореме II.2.5 это будет означать, что система (I) линейно независима, в противном случае – линейно зависима.

Пример II.2.6. Пусть в $\mathbb{R}^{(4)}$ задана система векторов

$$(I) \ a_1, a_2, a_3,$$

где $a_1 = (1, -1, 2, -2)$; $a_2 = (2, 0, 1, -1)$; $a_3 = (3, -2, 1, 2)$.

Рассмотрим произвольную линейную комбинацию этих векторов, равную нулю-вектору:

$$\alpha_1a_1 + \alpha_2a_2 + \alpha_3a_3 = \vec{0}.$$

По первой позиции имеем

$$1 \cdot \alpha_1 + 2 \cdot \alpha_2 + 3 \cdot \alpha_3 = 0,$$

аналогично по остальным позициям

$$(-1) \cdot \alpha_1 + 0 \cdot \alpha_2 + (-2) \cdot \alpha_3 = 0,$$

$$2 \cdot \alpha_1 + 1 \cdot \alpha_2 + 1 \cdot \alpha_3 = 0,$$

$$(-2) \cdot \alpha_1 + (-1) \cdot \alpha_2 + 2 \cdot \alpha_3 = 0.$$

Таким образом, получаем систему уравнений с неизвестными α_1, α_2 и α_3 :

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

Из второго уравнения получаем

$$\alpha_1 = -2\alpha_3.$$

Подставляя $-2\alpha_3$ вместо α_1 в остальные уравнения, получим

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

Из второго уравнения получаем

$$\alpha_1 = -2\alpha_3.$$

Подставляя $-2\alpha_3$ вместо α_1 в остальные уравнения, получим

$$\begin{cases} 2\alpha_2 + \alpha_3 = 0, \\ \alpha_2 - 3\alpha_3 = 0, \\ \alpha_2 - 4\alpha_3 = 0. \end{cases}$$

Из двух последних уравнений имеем

$$\alpha_3 = 0,$$

откуда $\alpha_2 = 0$. Но тогда и $\alpha_1 = 0$.

Таким образом, данная система уравнений имеет только нулевое решение и, следовательно, система (I) линейно независима.

Пример II.2.7. Пусть в $\mathbb{R}^{(4)}$ задана система векторов

$$(I) \quad a_1, a_2, a_3,$$

где $a_1 = (-2, 1, -1, 1)$; $a_2 = (-3, 2, -2, 1)$; $a_3 = (0, -1, 1, 1)$.

Действуя аналогично предыдущему примеру, получаем систему уравнений:

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

Из первого уравнения имеем

$$\alpha_1 = -\frac{3}{2}\alpha_2.$$

После подстановки вместо α_1 в остальные уравнения из второго уравнения получаем $\alpha_3 = -\frac{1}{2}\alpha_2$. Придавая значение $\alpha_2 = 2$, получаем ненулевое решение

$$\alpha_1 = -3, \alpha_2 = 2, \alpha_3 = 1,$$

т.е. система (I) линейно зависима.

Лекция 3. БАЗИСЫ

Определение II.3.1. Пусть V есть векторное пространство над $\mathbb{R}$, и рассмотрим систему векторов из V :

$$(I) \quad a_1, a_2, \dots, a_k.$$

Подсистема системы (I):

$$(II) \quad a_{i_1}, a_{i_2}, \dots, a_{i_r}, \quad r \leq k,$$

называется базисом системы (I), если (II) есть максимальная линейно независимая подсистема системы (I).

Другими словами, подсистема (II) является базисом системы (I), если

- 1) (II) является линейно независимой системой векторов;
- 2) присоединение к подсистеме (II) любого вектора $b \in [(I) \setminus (II)]$ дает линейно зависимую подсистему

$$(III) = (II) \cup \{b\}.$$

Замечание 1. Если система (I) линейно независима, то она совпадает со своим базисом.

Замечание 2. Если система (I) линейно зависима, то она может иметь несколько базисов.

Пример II.3.2. В пространстве $\mathbb{R}^{(4)}$ рассмотрим систему векторов-строк:

$$(I) \quad x_1, x_2, x_3,$$

где $x_1 = (-2, 1, -1, 1)$; $x_2 = (1, 2, 0, -1)$; $x_3 = (4, 3, 1, -3)$.

Подсистема

$$(II) \quad x_1, x_2$$

является базисом системы (I). Действительно, (II) является линейно независимой подсистемой, поскольку вектор-строки x_1 и x_2 не пропорциональны.

Далее, присоединение к системе (II) вектора x_3 дает линейно зависимую систему векторов, поскольку $x_3 = 2x_2 - x_1$.

Аналогично можно убедиться, что системы векторов

$$(III) \quad x_1, x_3 \quad \text{и} \quad (IV) \quad x_2, x_3$$

являются базисами системы (I).

Теорема II.3.3. Пусть в векторном пространстве V над $\mathbb{R}$ задана система

$$(I) \quad a_1, a_2, \dots, a_k$$

и пусть система

$$(II) \quad a_{i1}, a_{i2}, \dots, a_{ir}$$

является базисом системы (I). Тогда любой вектор системы (I) единственным образом линейно выражается через базис (II) системы (I).

Доказательство. Покажем сначала, что любой вектор x системы (I) линейно выражается через базис (II).

Случай 1. $x \in (II)$.

В этом случае утверждение очевидно.

Случай 2. $x \notin (II)$.

В этом случае рассмотрим систему

$$(III) \quad a_{i1}, a_{i2}, \dots, a_{ir}, x.$$

По определению базиса эта система линейно зависима. Поскольку система (II) линейно независима, то по свойству 8 (с. 62) вектор x линейно выражается через базис (II).

Покажем теперь, что вектор x единственным образом линейно выражается через базис (II). Пусть вектор x двумя способами линейно выражается через базис (II):

$$x = \alpha_{i1} a_{i1} + \dots + \alpha_{ir} a_{ir},$$

$$x = \beta_{i1} a_{i1} + \dots + \beta_{ir} a_{ir}.$$

Тогда имеем

$$0 = (\alpha_{i1} - \beta_{i1}) a_{i1} + \dots + (\alpha_{ir} - \beta_{ir}) a_{ir}.$$

Поскольку (II) есть линейно независимая система векторов, то по теореме II.2.5 получаем, что

$$\alpha_{i1} - \beta_{i1} = 0, \dots, \alpha_{ir} - \beta_{ir} = 0.$$

Но тогда $\alpha_{i1} = \beta_{i1}, \dots, \alpha_{ir} = \beta_{ir}$.

Определение II.3.4. Пусть в векторном пространстве V над $\mathbb{R}$ есть две системы векторов:

$$(I) \quad a_1, a_2, \dots, a_k$$

и $(II) \quad b_1, b_2, \dots, b_e.$

Будем говорить, что система (I) линейно выражается через (II), если любой вектор системы (I) линейно выражается через систему (II). Будем обозначать это так: $(I) \rightarrow (II)$.

Теорема II.3.5. В векторном пространстве V над $\mathbb{R}$ рассмотрим систему

$$(I) \quad a_1, a_2, \dots, a_k$$

и ее подсистему

$$(II) \quad a_{i1}, a_{i2}, \dots, a_{ir}.$$

Подсистема (II) является базисом системы (I) тогда и только тогда, когда выполняются следующие условия:

а) подсистема (II) линейно независима;

б) $(I) \rightarrow (II)$.

Доказательство. Необходимость. Пусть (II) есть базис системы (I), тогда по определению базиса, во-первых, система (II) линейно независима, во-вторых, для любого вектора $x \in (I) \setminus (II)$ система

$$(III) \quad a_{i1}, a_{i2}, \dots, a_{ir}, x$$

является линейно зависимой системой. Но тогда по свойству 8 (с. 62) имеем $x \rightarrow (II)$. Поскольку это имеет место также и для любого $x \in (II)$, то оба условия (а) и (б) выполняются.

Достаточность. Пусть оба условия (а) и (б) выполняются, тогда система (II) линейно независима по условию (а) и для любого вектора $x \in [(I)/(II)]$ система

$$(III) \quad a_{i1}, a_{i2}, \dots, a_{ir}, x$$

является линейно зависимой, поскольку $x \rightarrow (II)$ согласно условию (б). Но тогда подсистема (II) является базисом системы (I) согласно определению базиса.

Определение II.3.6. Две системы векторов

$$(I) \quad a_1, \dots, a_k$$

и $(II) \quad b_1, \dots, b_l.$

назовем эквивалентными, если $(I) \rightarrow (II)$ и $(II) \rightarrow (I)$. Эквивалентность систем (I) и (II) будем обозначать так: $(I) \sim (II)$.

Замечание 1. Непосредственно проверяется, что эквивалентность систем векторов пространства V над $\mathbb{R}$ является отношением эквивалентности.

Замечание 2. Легко видеть, что любая система векторов эквивалентна любому своему базису.

Предложение II.3.7. Любые два базиса данной системы векторов эквивалентны.

Доказательство. Пусть в системе

$$(I) \quad a_1, \dots, a_k$$

имеются два базиса:

$$(II) \quad a_{i_1}, a_{i_2}, \dots, a_{i_r}$$

и

$$(III) \quad a_{j_1}, a_{j_2}, \dots, a_{j_s}.$$

Очевидно, что любой базис системы векторов эквивалентен этой системе. Тогда имеем

$$(II) \sim (I) \quad \text{и} \quad (I) \sim (III).$$

По свойству транзитивности эквивалентности систем векторов получаем

$$(II) \sim (III).$$

Теорема II.3.8 (теорема Штейница о замене).

Пусть даны две системы в векторном пространстве V :

$$(I) \quad a_1, \dots, a_k$$

и

$$(II) \quad b_1, \dots, b_s.$$

Если система (I) является линейно независимой и $(I) \rightarrow (II)$, то справедливы следующие утверждения:

а) число векторов системы (I) не превосходит число векторов системы (II), т.е. $k \leq s$;

б) в системе (II) можно так выбрать k векторов, что, заменив их векторами из системы (I) и разместив на первых k местах, получим систему

$$(III) \quad a_1, \dots, a_k, b_{k+1}, \dots, b_s,$$

которая эквивалентна системе (II), т.е. $(III) \sim (II)$.

Рассмотрим теперь ряд следствий из теоремы о замене.

Следствие 1. Любые две линейно независимые и эквивалентные системы векторов имеют одинаковое число векторов.

Доказательство. Пусть системы (I) $a_1, \dots, a_k$ и (II) $b_1, \dots, b_s$ линейно независимые и эквивалентные. Тогда по теореме о замене имеем, что

$$k \leq s \text{ и } s \leq k,$$

следовательно, $k = s$.

Следствие 2. Любые два базиса данной системы векторов имеют одинаковое число векторов.

Доказательство. Пусть система (I) $a_1, \dots, a_k$ имеет два базиса:

$$(II) \quad a_{i_1}, a_{i_2}, \dots, a_{i_r}$$

и

$$(III) \quad a_{j_1}, a_{j_2}, \dots, a_{j_s}.$$

Согласно предложению II.3.7, базисы (II) и (III) эквивалентны. Тогда по следствию 1 из теоремы о замене получаем, что $r = s$.

Определение II.3.9. Число векторов в любом базисе системы векторов (I) назовем рангом данной системы векторов и обозначим $\text{rang}(I)$.

Пример II.3.10. В $\mathbb{R}^{(4)}$ рассмотрим систему векторов

$$(I) \quad x_1, x_2, x_3, x_4,$$

где $x_1 = (1, -1, 1, 2)$; $x_2 = (1, 1, -1, 2)$; $x_3 = (-1, -3, 3, -2)$; $x_4 = (1, -3, 3, 2)$.

Найдем ранг системы (I). Согласно следствию 2, из теоремы о замене достаточно найти хотя бы один базис системы (I) и тогда ранг системы (I) равен числу векторов в этом базисе.

Покажем, что подсистема (II) x_1, x_2 является базисом системы (I).

Действительно, во-первых, система (II) линейно независима, так как строки x_1 и x_2 не пропорциональны. Далее, во-вторых, любое расширение подсистемы (II) посредством вектор-строки из $(I) \setminus (II)$ дает линейно зависимую систему. Действительно, подсистема

$$(III) \quad x_1, x_2, x_3$$

линейно зависима, так как $x_3 = x_1 - 2x_2$.

Далее, подсистема

$$(IV) \quad x_1, x_2, x_4$$

линейно зависима, так как

$$x_4 = 2x_1 - x_2.$$

Следовательно, подсистема (II) есть базис системы (I), и тогда

$$\text{rank}(I) = 2.$$

Следствие 3. Пусть даны две системы

$$(I) \ a_1, \dots, a_k \text{ и } (II) \ b_1, \dots, b_\ell.$$

Если $(I) \rightarrow (II)$, то $\text{rank}(I) \leq \text{rank}(II)$.

Доказательство. Пусть система $(III) \ a_{i_1}, \dots, a_{i_r}$ является базисом системы (I) , а система $(IV) \ b_{j_1}, \dots, b_{j_\ell}$ — базисом системы (II) .

Поскольку $(I) \sim (III)$ и $(II) \sim (IV)$ согласно замечанию 2 к определению эквивалентности систем векторов, то с учетом условия $(I) \rightarrow (II)$ получаем цепочку:

$$(III) \rightarrow (I) \rightarrow (II) \rightarrow (IV).$$

По транзитивности отношения $\rightarrow$ получаем $(III) \rightarrow (IV)$. Поскольку система (III) линейно независима и $(III) \rightarrow (IV)$, то по теореме о замене имеем $r \leq \ell$. Поскольку $\text{rank}(I) = r$ и $\text{rank}(II) = \ell$, то получаем доказательство.

Следствие 4. Эквивалентные системы векторов имеют равные ранги.

Доказательство. Пусть системы (I) и (II) эквивалентны, тогда

$$(I) \rightarrow (II) \text{ и } (II) \rightarrow (I).$$

По следствию 3 имеем

$$\text{rank}(I) \leq \text{rank}(II) \text{ и } \text{rank}(II) \leq \text{rank}(I).$$

Отсюда получаем $\text{rank}(I) = \text{rank}(II)$.

Верно ли обратное утверждение к следствию 4, т.е. если две системы векторов имеют одинаковые ранги, то будут ли они эквивалентны? Следующий пример показывает, что ответ на этот вопрос отрицательный.

Пример II.3.11. В $\mathbb{R}^{(4)}$ рассмотрим две системы векторов:

$$(I) \ a_1,$$

где $a_1 = (1, -1, 1, 1)$, и

$$(II) \ a_2,$$

где $a_2 = (1, 2, 3, 1)$.

Обе системы имеют ранг 1, но не эквивалентны.

Следствие 5. Если две системы векторов имеют равные ранги и одна из них линейно выражается через другую, то эти системы эквивалентны.

Доказательство. Пусть две системы

$$(I) \ a_1, \dots, a_k \quad \text{и} \quad (II) \ b_1, \dots, b_\ell.$$

имеют равные ранги и $(I) \rightarrow (II)$.

Рассмотрим системы векторов

$$(III) \ a_{i1}, \dots, a_{ir} \quad \text{и} \quad (IV) \ b_{j1}, \dots, b_{je},$$

которые являются базисами систем (I) и (II) соответственно. Имеем

$$(III) \rightarrow (I) \rightarrow (II) \rightarrow (IV),$$

тогда по транзитивности отношения $\rightarrow$ получаем $(III) \rightarrow (IV)$. Поскольку система (III) линейно независима и $(III) \rightarrow (IV)$, то из теоремы о замене следует, с учетом равенства рангов $r = \ell$, что если заменить систему (IV) системой (III), то получим систему, эквивалентную (IV), т.е. $(III) \sim (IV)$. Поскольку базис системы векторов эквивалентен этой системе, то из эквивалентности базисов следует эквивалентность систем векторов.

Определение II.3.12. Назовем базисом векторного пространства V над $\mathbb{R}$ максимальную линейно независимую подсистему векторов из V (если такая подсистема существует).

Другими словами, подсистема (I) $a_1, \dots, a_k$ есть базис пространства V , если

- 1) система векторов (I) является линейно независимой системой;
- 2) для любого вектора $x \in [V \setminus (I)]$ подсистема $(II) = (I) \cup \{x\}$ является линейно зависимой системой.

Замечание. Векторное пространство может не иметь конечной максимальной линейно независимой подсистемы.

Пример II.3.13. Рассмотрим векторное пространство $\mathbb{R}[x]$ многочленов от неизвестной x с действительными коэффициентами (см. пример II.1.9). Предположим, что существует конечная максимальная линейно независимая подсистема в $\mathbb{R}[x]$

$$(I) \ f_1, \dots, f_k.$$

Обозначим через n максимальную из степеней многочленов $f_1, \dots, f_k$.

Рассмотрим систему

$$(II) = (I) \cup \{x^{n+1}\}.$$

Предположим, что система (II) линейно зависима. Тогда по свойству 8 линейной зависимости x^{n+1} должен линейно выражаться через многочлены $f_1, \dots, f_k$, степени которых не превосходят n , что невозможно.

Итак, пространство $\mathbb{R}[x]$ не имеет конечного базиса. В то же время существует бесконечное максимальное линейно независимое множество в $\mathbb{R}[x]$, а именно

$$1, x, \dots, x^m, \dots$$

Определение II.3.14. Векторное пространство V над $\mathbb{R}$ назовем конечномерным, если V имеет конечный базис (т.е. конечное максимальное линейно независимое множество). В противном случае V называется бесконечномерным пространством.

Условимся в дальнейшем рассматривать только конечномерные векторные пространства, поэтому в дальнейшем термин «векторное пространство», если не будет оговорено противное, будет по умолчанию означать «конечномерное векторное пространство», а «базис» — «конечный базис».

Мы хотим определить понятие размерности векторного пространства аналогично понятию ранга системы векторов. Для этого необходимо получить аналоги тех теорем и утверждений, на основании которых было введено понятие ранга системы векторов.

Теорема II.3.15. Пусть V есть векторное пространство над $\mathbb{R}$ и система

$$(I) \ a_1, \dots, a_k$$

— базис V . Тогда любой $x \in V$ единственным образом линейно выражается через базис (I).

Доказательство получается дословным повторением доказательства теоремы II.3.3.

Теорема II.3.16. Система векторов (I) $a_1, \dots, a_k$ векторного пространства V является базисом пространства V тогда и только тогда, когда выполнены условия:

- 1) система (I) линейно независима;
- 2) любой вектор $x \in V$ линейно выражается через систему (I).

Доказательство получается дословным повторением доказательства теоремы II.3.5.

А теперь необходимо расширить понятие эквивалентности систем векторов.

Определение II.3.17. Две системы векторов (возможно, бесконечные) назовем эквивалентными, если каждый вектор любой из этих систем линейно выражается через другую систему.

Замечание. «Система векторов линейно выражается через бесконечную систему векторов» означает, что система векторов линейно выражается через некоторую ее конечную подсистему.

Предложение II.3.18. Любые два базиса векторного пространства V эквивалентны.

Доказательство получается дословным повторением доказательства предложения II.3.7, с тем различием, что вместо конечной системы (I) рассматривается бесконечное пространство V , и тогда нужно использовать эквивалентность систем векторов в смысле определения II.3.17.

Теперь из теоремы о замене получается новое следствие.

Следствие 2+. Любые два базиса векторного пространства V имеют одинаковое число векторов.

Теперь, наконец, мы можем ввести понятие размерности.

Определение II.3.19. Число векторов в любом базисе векторного пространства V назовем размерностью пространства V и обозначим $\dim V$.

Пример II.3.20. Найдём базис и размерность пространства $\mathbb{R}^{(4)}$. Рассмотрим систему векторов (I) a_1, a_2, a_3, a_4 , где $a_1 = (1, 0, 0, 0)$, $a_2 = (0, 1, 0, 0)$, $a_3 = (0, 0, 1, 0)$ и $a_4 = (0, 0, 0, 1)$.

Покажем, что (I) есть базис $\mathbb{R}^{(4)}$.

Сначала проверим линейную независимость системы (I). Пусть

$$\alpha_1 a_1 + \alpha_2 a_2 + \alpha_3 a_3 + \alpha_4 a_4 = \bar{0},$$

тогда

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = \bar{0},$$

откуда

$$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.$$

Теперь покажем, что любой вектор

$$x = (x_1, x_2, x_3, x_4) \in \mathbb{R}^{(4)}$$

линейно выражается через систему (I). Действительно, имеем

$$x = x_1 a_1 + x_2 a_2 + x_3 a_3 + x_4 a_4.$$

Тогда по теореме II.3.16 система (I) является базисом в $\mathbb{R}^{(4)}$. Отсюда следует, что $\dim \mathbb{R}^{(4)} = 4$, поскольку система (I) состоит из четырех векторов.

Применим теперь полученные выше результаты к понятию ранга матрицы.

Определение II.3.21. Пусть A есть $m \times n$ -матрица над $\mathbb{R}$. Рассмотрим систему L_1 векторов-строк матрицы A и систему L_2 векторов-столбцов матрицы A . Ранг системы векторов L_1 назовем построчным рангом матрицы A и обозначим его через $rk_r(A)$. Ранг системы векторов L_2 назовем постолбцевым рангом матрицы A и обозначим его через $rk_c(A)$.

Пример II.3.22. Для матрицы

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

система L_1 состоит из векторов

$$a_1 = (1, 0, 0, 1), a_2 = (2, 0, 0, 3), a_3 = (0, 0, 0, 0),$$

а система L_2 — из векторов

$$b_1 = (1, 2, 0), b_2 = (0, 0, 0), b_3 = (0, 0, 0), b_4 = (1, 3, 0).$$

Очевидно, что ранги систем L_1 и L_2 равны 2, причем подсистема в L_1 из векторов a_1, a_2 является базисом L_1 , а подсистема в L_2 из векторов b_1, b_4 является базисом в L_2 . В то же время легко видеть, что $rk(A) = 2$ и минор

$$\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} = 1 \neq 0$$

является базисным минором матрицы A , причем базисными строками являются первые две строки, а базисными столбцами — первый и четвертый столбцы матрицы A . Заметим, что в этом примере

$$rk(A) = rk_r(A) = rk_c(A).$$

Теорема II.3.23 (теорема о базисном миноре). Базисные строки матрицы A составляют базис системы векторов-строк L_1 , а базисные столбцы матрицы A составляют базис системы векторов-столбцов L_2 .

Доказательство приведем для строк (для столбцов аналогично).

Покажем сначала, что любая система базисных строк является линейно независимой. Действительно, если предположить, что эта система линейно зависима, то одна из базисных строк линейно выражается через другие базисные строки из этой системы базисных строк. Но тогда по свойству 7 определителей (с. 30) базисный минор равен нулю, что противоречит определению базисного минора.

Покажем далее, что любая строка матрицы A является линейной комбинацией базисных строк. Поскольку перестановка строк и столбцов не меняет равенство нулю определителя, то без ограничения общности можем считать, что базисный минор Δ находится в левом верхнем углу матрица A , т.е. в первых r строках и столбцах матрицы A , где $r = rk(A)$ – ранг матрицы A .

Для любой строки с номером k , где $1 \leq k \leq m$, и любого столбца с номером j , где $1 \leq j \leq n$, обозначим через Δ_{kj} следующий определитель порядка $r+1$:

$$(*)\Delta_{kj} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} & a_{1j} \\ a_{21} & a_{22} & \dots & a_{2r} & a_{2j} \\ \dots & \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rr} & a_{rj} \\ a_{k1} & a_{k2} & \dots & a_{kr} & a_{kj} \end{vmatrix}.$$

Покажем, что для любых k и j имеем $\Delta_{kj} = 0$.

Действительно, для $k \leq r$ или $j \leq r$ имеем $\Delta_{kj} = 0$, так как в этом случае Δ_{kj} имеет две одинаковые строки или два одинаковых столбца. Если же $k > r$ и $j > r$, то Δ_{kj} является минором матрицы A порядка $r+1$, и по определению ранга матрицы такой минор должен равняться нулю.

Пусть теперь k -я строка матрицы A есть произвольная фиксированная строка матрицы A , которая не является базисной строкой, т.е. $k > r$.

Будем предполагать далее, что для Δ_{kj} имеем $r < k \leq m$ и $1 \leq j \leq n$. Обозначим алгебраические дополнения элементов j -го столбца в (*) через

$$A_{1j} = c_1, A_{2j} = c_2, \dots, A_{rj} = c_r, A_{kj} = c_{r+1}.$$

Тогда для любого $j = \overline{1, n}$ имеем

$$0 = \Delta_{kj} = c_1 a_{1j} + c_2 a_{2j} + \dots + c_r a_{rj} + c_{r+1} a_{kj}.$$

Действительно, для $j = \overline{1, r}$ это вытекает из следствия 2 к теореме I.2.22, а для $j > r$ из следствия 1 к теореме I.2.22. Учтем теперь, что $c_{r+1} = \Delta \neq 0$, и тогда для любого $j = \overline{1, n}$ имеем

$$a_{kj} = \left(-\frac{c_1}{c_{r+1}} \right) a_{1j} + \dots + \left(-\frac{c_r}{c_{r+1}} \right) a_{rj}.$$

Но это и означает, что k -я строка является линейной комбинацией первых r базисных строк. Поскольку для базисной строки это утверждение очевидно, то мы получили, что любая строка матрицы A линейно выражается через систему из r базисных строк. Таким образом, эта система является базисом в системе векторов L_1 .

Следствие 1. Для любой матрицы A имеем

$$rk(A) = rk_r(A) = rk_c(A).$$

Доказательство. По теореме II.3.23 базисные строки составляют базис системы L_1 . Но тогда число этих базисных строк есть ранг системы L_1 , т.е. $rk_r(A)$. С другой стороны, число базисных строк есть порядок базисного минора, т.е. $rk(A)$. Следовательно, $rk(A) = rk_r(A)$. Аналогично получаем $rk(A) = rk_c(A)$.

Следствие 2. Пусть A есть квадратная матрица порядка n . $|A| = 0$ тогда и только тогда, когда система векторов-строк L_1 (система векторов-столбцов L_2) линейно зависима.

Доказательство. Необходимость. Если $|A| = 0$, то $rk(A) < n$. По следствию 1 из теоремы II.3.23

$$rk(A) = rk_r(A) = rk_c(A) < n.$$

Но если ранг системы из n векторов меньше n , то системы векторов L_1 и L_2 линейно зависимы в нашем случае.

Достаточность. Если строки или столбцы матрицы A линейно зависимы, то строка или столбец линейно выражаются через остальные строки или столбцы соответственно. Но тогда по свойству 7 определителей (с. 30) имеем $|A| = 0$.

Замечание 1. Если для некоторого ненулевого минора k -го порядка M_k все окаймляющие его миноры $k+1$ -го порядка равны нулю, то M_k есть базисный минор и тогда $rk(A) = k$.

Действительно, строки минора M_k линейно независимы по следствию 2 из теоремы II.3.23. Добавление любой строки к этим строкам по условию даст линейно зависимую систему согласно следствию 2 из теоремы II.3.23. Таким образом, строки минора M_k образуют максимальную линейно независимую подсистему системы L_1 , т.е. базис системы L_1 . Это означает, что M_k есть базисный минор. Но тогда, так как число строк M_k равно k , имеем $k = rk_r(A)$, следовательно, $k = rk(A)$.

Это замечание дает обоснование методу окаймляющих миноров вычисления ранга матрицы.

Замечание 2. Используя теорему о базисном миноре и её следствия, получаем метод определения линейной зависимости или линейной независимости системы векторов посредством нахождения ранга матрицы, составленной из векторов этой системы.

Пример II.3.24. Рассмотрим систему векторов в $\mathbb{R}^{(5)}$:

$$a_1 = (1, -1, 0, 1, 0), a_2 = (2, -2, 0, 0, 0),$$

$$a_3 = (1, 0, 0, -1, 1), a_4 = (-1, 1, 0, -1, 0).$$

Составим матрицу, строки которой $a_i, i = \overline{1, 4}$:

$$A = \begin{pmatrix} 1 & -1 & 0 & 1 & 0 \\ 2 & -2 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 1 \\ -1 & 1 & 0 & -1 & 0 \end{pmatrix}.$$

Ранг этой матрицы вычислен в примере I.3.9 и равен $rk(A) = 3$. Но тогда $rk_r(A) = 3$, т.е. ранг системы векторов $a_1, \dots, a_4$ равен 3. Следова-

тельно, система линейно зависима. Более того, зная базисный минор матрицы A , а именно минор

$$M_3 = \begin{pmatrix} 2 & -2 & 0 \\ 1 & 0 & -1 \\ -1 & 1 & -1 \end{pmatrix}$$

из примера I.3.5, мы получаем базис нашей системы векторов из строк матрицы A , проходящих через минор M_3 , т.е. $\{a_2, a_3, a_4\}$.

Лекция 4. ПОДПРОСТРАНСТВА

Определение II.4.1. Пусть V есть векторное пространство над $\mathbb{R}$ и $V' \subseteq V$ – подмножество V . Будем говорить, что V' есть подпространство пространства V , если V' само является векторным пространством над $\mathbb{R}$ относительно тех же операций, что и в пространстве V .

Теорема II.4.2 (критерий подпространства). Подмножество V' векторного пространства V над $\mathbb{R}$ является подпространством пространства V тогда и только тогда, когда выполняется условие: для любых $a, b \in V'$ и любого $\alpha \in \mathbb{R}$, $a + b \in V'$ и $\alpha a \in V'$.

Доказательство. Необходимость очевидна.

Достаточность. Пусть выполнено условие теоремы, рассмотрим операции сложения векторов и умножение вектора на скаляр, заданные для V , относительно V' . Согласно условию, пункты 1 и 2 определения II.1.1 векторного пространства применительно к V' выполняются. Пункт 3.1 этого определения для V' выполняется, поскольку он выполняется для всех векторов из V , следовательно, и для $V' \subseteq V$. Для проверки выполнения пункта 3.2 пусть $a \in V'$ – произвольный элемент и $\alpha \in \mathbb{R}$ – произвольный скаляр. По условию $\alpha a \in V'$, но тогда $0 \cdot a = \bar{0} \in V'$, что означает выполнение пункта 3.2 для V' .

Для проверки пункта 3.3 пусть $a \in V'$ – произвольный элемент и $\alpha \in \mathbb{R}$ – произвольный скаляр. По условию $\alpha a \in V'$, но тогда $(-1) \cdot a = -a \in V'$, что означает выполнение пункта 3.3 для V' . Остальные пункты 3.4 – 3.8 определения II.1.1 выполняются для V' , поскольку они выполняются для V .

Пример II.4.3. Пусть $\mathbb{R}^{(k)}$ и $\mathbb{R}^{(\ell)}$ – векторные пространства вектор-строк и $k < \ell$. Для любой строки $x = (x_1, \dots, x_k) \in \mathbb{R}^{(k)}$ рассмотрим

$$\bar{x} = (x_1, \dots, x_k, 0, \dots, 0) \in \mathbb{R}^{(\ell)}.$$

Обозначим множество всех вектор-строк длины ℓ , полученных таким образом, через $\bar{\mathbb{R}}^{(k)}$. Понятно, что $\bar{\mathbb{R}}^{(k)} \subseteq \mathbb{R}^{(\ell)}$. Применяя критерий подпространства к $\bar{\mathbb{R}}^{(k)}$, получаем, что $\bar{\mathbb{R}}^{(k)}$ есть подпространство в $\mathbb{R}^{(\ell)}$.

Следующее определение даст универсальный способ построения подпространств.

Определение II.4.4. Пусть дана система векторов пространства V

$$(I) \ a_1, \dots, a_k.$$

Линейной оболочкой системы (I) назовем совокупность всевозможных линейных комбинаций векторов из системы (I). Будем обозначать линейную оболочку системы (I) через $\langle a_1, \dots, a_k \rangle$.

Замечание. В определении II.4.4 система векторов (I) может быть бесконечной.

Теорема II.4.5.

а) Линейная оболочка произвольной системы векторов из V является подпространством пространства V .

б) Любое подпространство $V' \subseteq V$ является линейной оболочкой некоторой системы векторов из V' .

Доказательство.

Утверждение «а» следует из критерия подпространства.

Непосредственно проверяется, что V' является линейной оболочкой своего базиса, откуда следует «б».

Определение II. 4.6. Пусть V_1 и V_2 есть подпространства векторного пространства V над $\mathbb{R}$. Определим сумму подпространств

$$V_1 + V_2 = \{x = x_1 + x_2 \mid x_1 \in V_1, x_2 \in V_2\},$$

и $V_1 \cap V_2$ – пересечение V_1 и V_2 .

Предложение II.4.7. Сумма и пересечение подпространств векторного пространства V над $\mathbb{R}$ являются подпространствами V .

Доказательство. Утверждение получается непосредственным применением критерия подпространства.

Определение II.4.8. Будем говорить, что подпространство V' векторного пространства V над $\mathbb{R}$ является прямой суммой подпространств V_1 и V_2 , если

$$1) \ V' = V_1 + V_2;$$

$$2) \ V_1 \cap V_2 = 0.$$

Будем обозначать прямую сумму подпространств V_1 и V_2 как $V_1 \oplus V_2$.

Теорема II.4.9. Подпространство V' векторного пространства V над $\mathbb{R}$ является прямой суммой подпространств V_1 и V_2 тогда и только тогда, когда любой $x \in V'$ можно единственным образом выразить в виде

$$x = x_1 + x_2,$$

где $x_1 \in V_1$, $x_2 \in V_2$.

Доказательство. Необходимость. Пусть $V' = V_1 \oplus V_2$ и $x \in V'$ есть произвольный элемент. Поскольку

$$V' = V_1 + V_2, \text{ то } x = x_1 + x_2,$$

где $x_1 \in V_1$, $x_2 \in V_2$. Докажем теперь единственность такого представления. Пусть

$$x = x_1 + x_2 \text{ и } x = x'_1 + x'_2,$$

где $x_1, x'_1 \in V_1$ и $x_2, x'_2 \in V_2$.

Тогда

$$x_1 + x_2 = x'_1 + x'_2,$$

откуда получаем

$$x_1 - x'_1 = x'_2 - x_2.$$

Поскольку $x_1 - x'_1 \in V_1$ и $x'_2 - x_2 \in V_2$, то

$$z = x_1 - x'_1 = x'_2 - x_2 \in V_1 \cap V_2 = 0.$$

Отсюда имеем $x_1 = x'_1$ и $x'_2 = x_2$.

Достаточность. Пусть произвольный элемент $x \in V'$ единственным образом представим в виде

$$x = x_1 + x_2,$$

где $x_1 \in V_1$, $x_2 \in V_2$. Тогда, во-первых, $V' = V_1 + V_2$. Покажем, что, во-вторых, $V_1 \cap V_2 = 0$. Пусть $z \in V_1 \cap V_2$ — произвольный элемент. Тогда имеем два представления элемента z :

$$z = 0 + z = z + 0.$$

Из единственности представления z в таком виде получаем, что $z = 0$. Поскольку $V' = V_1 + V_2$ и $V_1 \cap V_2 = 0$, то $V' = V_1 \oplus V_2$.

Теорема II.4.10 (теорема о достройке базиса).

1) Любая линейно независимая подсистема данной системы векторов из конечномерного векторного пространства V над $\mathbb{R}$ может быть дополнена до базиса этой системы векторов.

2) Любая линейно независимая система векторов конечномерного пространства V над $\mathbb{R}$ может быть дополнена до базиса пространства V .

Доказательство (1). Пусть имеется система

$$(I) \ a_1, \dots, a_k,$$

и в (I) рассмотрим произвольную линейно независимую подсистему

$$(II) \ a_{i_1}, \dots, a_{i_r}.$$

Можем считать, что $r < k$ (в противном случае доказывать нечего). Возможны два случая.

Случай 1. Присоединение любого вектора $b \in [(I) \setminus (II)]$ к подсистеме (II) приводит к линейно зависимой подсистеме $a_{i_1}, \dots, a_{i_r}, b$. В этом случае (II) является максимальной линейно независимой подсистемой системы (I), следовательно, (II) есть базис (I).

Случай 2. Существует вектор $b \in [(I) \setminus (II)]$, присоединение которого к подсистеме (II) дает линейно независимую подсистему. В этом случае рассмотрим расширенную подсистему (III) $a_{i_1}, \dots, a_{i_r}, b$ и применим к ней те же рассуждения, что и к подсистеме (II), т.е. опять получим случаи 1 и 2.

Продолжая этот процесс, заметим, что он оборвется не более чем за s шагов, где s есть ранг системы (I).

Доказательство (2) получается аналогично (1).

Следующая теорема позволяет не только вычислить размерность суммы подпространств, но и дает способ нахождения базиса суммы подпространств.

Теорема II.4.11. Размерность суммы двух подпространств равна сумме размерностей этих подпространств минус размерность их пересечения.

Доказательство. Пусть V_1 и V_2 – подпространства векторного пространства V над $\mathbb{R}$. Покажем, что

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

Обозначим $\dim(V_1 \cap V_2) = k_0$, $\dim V_1 = k_1$, $\dim V_2 = k_2$, и пусть система

$$(I) \ a_1, \dots, a_{k_0}$$

есть базис $V_1 \cap V_2$. Поскольку $V_1 \cap V_2 \subseteq V_1$ и $V_1 \cap V_2 \subseteq V_2$, то по теореме о достройке базиса (теорема II.4.10) систему (I) можно дополнить до базисов как V_1 , так и V_2 . Имеем системы

$$(II) \ a_1, \dots, a_{k_0}, a_{k_0+1}, \dots, a_{k_1} \text{ — базис } V_1,$$

$$(III) \ a_1, \dots, a_{k_0}, a_{k_0+1}, \dots, a_{k_2} \text{ — базис } V_2.$$

Рассмотрим следующую систему из $k_1 + k_2 - k_0$ векторов:

$$(IV) \ a_1, \dots, a_{k_0}, a_{k_0+1}, \dots, a_{k_1}, b_{k_0+1}, \dots, b_{k_2}.$$

Заметим, что система (IV) является объединением систем (II) и (III). Покажем, что система (IV) является базисом $V_1 + V_2$.

Покажем сначала, что (IV) является линейно независимой системой. Пусть имеется произвольная линейная комбинация векторов системы (IV), равная нулю:

$$\begin{aligned} a_1 + \dots + \alpha_{k_0} a_{k_0} + \alpha_{k_0+1} a_{k_0+1} + \dots + \alpha_{k_1} a_{k_1} + \\ + \beta_{k_0+1} b_{k_0+1} + \dots + \beta_{k_2} b_{k_2} = 0. \end{aligned} \quad (*)$$

Обозначим через

$$\begin{aligned} x &= \alpha_1 + \dots + \alpha_{k_0} a_{k_0} + \alpha_{k_0+1} a_{k_0+1} + \dots + \alpha_{k_1} a_{k_1} = \\ &= -\beta_{k_0+1} b_{k_0+1} - \dots - \beta_{k_2} b_{k_2}. \end{aligned}$$

С одной стороны, как линейная комбинация векторов системы (II), вектор $x \in V_1$. С другой стороны, как линейная комбинация векторов системы (III), вектор $x \in V_2$. Следовательно, $x \in V_1 \cap V_2$. Но тогда вектор x можно представить в виде линейной комбинации векторов из базиса подпространства $V_1 \cap V_2$, а именно через систему (I):

$$x = \gamma_1 a_1 + \dots + \gamma_{k_0} a_{k_0}.$$

Отсюда получаем линейную комбинацию векторов линейно независимой системы (II), равную нулю:

$$(\alpha_1 - \gamma_1)a_1 + \dots + (\alpha_{k_0} - \gamma_{k_0})a_{k_0} + \\ + \alpha_{k_0+1}a_{k_0+1} + \dots + \alpha_{k_1}a_{k_1} = 0.$$

Следовательно, все коэффициенты этой линейной комбинации нулю, в частности

$$\alpha_{k_0+1} = \dots = \alpha_{k_1} = 0.$$

Но тогда

$$x = \alpha_1 a_1 + \dots + \alpha_{k_0} a_{k_0} = -\beta_{k_0+1} b_{k_0+1} - \dots - \beta_{k_1} b_{k_1}.$$

Отсюда получаем равную нулю линейную комбинацию векторов линейно независимой системы (III):

$$\alpha_1 a_1 + \dots + \alpha_{k_0} a_{k_0} + \beta_{k_0+1} b_{k_0+1} + \beta_{k_1} b_{k_1} = 0.$$

Следовательно, имеем

$$\alpha_1 = \dots = \alpha_{k_0} = \beta_{k_0+1} = \dots = \beta_{k_1} = 0.$$

Поскольку равенства

$$\alpha_{k_0+1} = \dots = \alpha_{k_1} = 0$$

были установлены выше, то все коэффициенты линейной комбинации (*) равны нулю и, следовательно, система (IV) линейно независима.

Пусть $z \in V_1 + V_2$ — произвольный элемент, тогда $z = x_1 + x_2$, где $x_1 \in V_1$ и $x_2 \in V_2$.

Отсюда имеем

$$x_1 = \alpha_1 a_1 + \dots + \alpha_{k_0} a_{k_0} + \alpha_{k_0+1} a_{k_0+1} + \dots + \alpha_{k_1} a_{k_1}, \\ x_2 = \beta_1 a_1 + \dots + \beta_{k_0} a_{k_0} + \beta_{k_0+1} a_{k_0+1} + \dots + \beta_{k_2} a_{k_2}$$

и, следовательно,

$$x_1 + x_2 = (\alpha_1 + \beta_1)a_1 + \dots + (\alpha_{k_0} + \beta_{k_0})a_{k_0} + \\ + \alpha_{k_0+1}a_{k_0+1} + \dots + \alpha_{k_1}a_{k_1} + \beta_{k_0+1}b_{k_0+1} + \dots + \beta_{k_2}b_{k_2}.$$

Таким образом, вектор $x_1 + x_2$ является линейной комбинацией векторов системы (IV). Следовательно, система (IV) есть базис $V_1 + V_2$, и тогда

$$\dim(V_1 + V_2) = k_1 + k_2 - k_0 = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

Приведем теперь примеры нахождения базисов и размерностей подпространств, но сначала соберем вместе информацию о базисах и размерностях основных классов векторных пространств, с которыми мы имеем в основном дело.

Пример II.4.12. Пусть $V = \mathbb{R}^{(n)}[x]$ – пространство многочленов от неизвестной x с действительными коэффициентами степени, не превосходящей n . Стандартный базис этого пространства

$$1, x, \dots, x^n,$$

и, следовательно, его размерность равна

$$n+1.$$

Пример II.4.13. Пусть $V = M_n(\mathbb{R})$ – пространство квадратных матриц порядка n с действительными элементами. Стандартный базис этого пространства

$$E_{ij}, i, j = \overline{1, n},$$

где E_{ij} содержит единицу на позиции (i, j) и нули на остальных позициях.

$$\text{Следовательно, } \dim(M_n(\mathbb{R})) = n^2.$$

Пример II.4.14. Пусть $V = \mathbb{R}^{(n)}$ – векторное пространство векторов-строк с n элементами. Стандартный базис этого пространства

$$e_1 = (1, 0, \dots, 0), \dots, e_n = (0, 0, \dots, 0, 1),$$

следовательно, $\dim \mathbb{R}^{(n)} = n$.

Пусть $V_1 = \{x \in V \mid x = (x_1, \dots, x_n) \text{ удовлетворяет } x_1 + \dots + x_n = 0\}$.

По критерию подпространства V_1 есть подпространство. Рассмотрим следующую систему векторов из V_1 :

$$a_1 = (1, -1, 0, \dots, 0), a_{22} = (0, 1, -1, 0, \dots, 0), a_{n-1} = (0, \dots, 0, 1, -1).$$

Данная система векторов линейно независима. Действительно, пусть

$$\alpha_1 a_1 + \alpha_2 a_2 + \dots + \alpha_{n-1} a_{n-1} = 0,$$

тогда

$$\alpha_1 = 0, \alpha_2 - \alpha_1 = 0, \dots, \alpha_{n-1} - \alpha_{n-2} = 0$$

и, следовательно, $\alpha_1 = \alpha_2 = \dots = \alpha_{n-1} = 0$.

Пусть $x = (x_1, \dots, x_n) \in V$ – произвольный вектор, тогда

$$x_1 + x_2 + \dots + x_n = 0.$$

Выразим вектор x как линейную комбинацию $a_1, \dots, a_{n-1}$:

$$\begin{aligned} x = & x_1 (1, -1, 0, \dots, 0) + (x_1 + x_2) (0, 1, -1, 0, \dots, 0) + \dots + \\ & + (x_1 + x_2 + \dots + x_{n-1}) (0, 0, \dots, 0, 1, -1). \end{aligned}$$

Следовательно, $a_1, \dots, a_{n-1}$ есть базис подпространства V_1 , и тогда $\dim V_1 = n - 1$.

Вопросы для самоконтроля

1. Сформулируйте определение векторного пространства.
2. Сформулируйте простейшие свойства векторных пространств.
3. Когда система из двух векторов линейно зависима?
4. Верно ли, что если система векторов содержит линейно независимую подсистему, то и сама система векторов будет линейно независимой?
5. Верно ли, что если система векторов линейно зависима, то в этой системе есть нуль-вектор?
6. Сформулируйте определения линейно зависимой и линейно независимой систем векторов.
7. Сформулируйте критерии линейной независимости и линейной зависимости систем векторов.
8. Найдите ошибку в «доказательстве» следующего утверждения: «любая система ненулевых векторов линейно независима». Доказательство. Пусть имеется система ненулевых векторов $a_1, a_2, \dots, a_n$. Рассмотрим произвольную линейную комбинацию этих векторов, равную нуль-вектору: $\alpha_1 a_1 + \alpha_2 a_2 + \dots + \alpha_n a_n = 0$.
9. Сформулируйте определение и критерий базиса системы векторов.
10. Может ли линейно зависящая система векторов иметь единственный базис?
11. Сформулируйте теорему Штейница о замене.
12. Верно ли, что в системе векторов ранга r любая линейно независимая подсистема из r векторов будет базисом?
13. Верно ли, что если система векторов и некоторая ее подсистема имеют одинаковые ранги, то они эквивалентны?
14. Сформулируйте теорему о базисном миноре и ее следствия.
15. Верно ли, что если у квадратной матрицы система векторов-столбцов линейно независима, то матрица невырожденная?
16. Сформулируйте определение и критерий подпространства.
17. Верно ли, что размерность прямой суммы подпространств равна сумме размерностей этих подпространств?

-
18. Сформулируйте критерий представления подпространства векторного пространства в виде прямой суммы некоторых подпространств.
 19. Верно ли, что любое векторное пространство совпадает с линейной оболочкой своего базиса?
 20. Верно ли, что если размерности векторного пространства и его подпространства равны, то пространство совпадает с этим подпространством?

МОДУЛЬ III

Лекция 1. СОВМЕСТИНОСТЬ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ

Определение III.1.1. Рассмотрим систему из m линейных уравнений с n неизвестными над $\mathbb{R}$:

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

где $a_{ij}, b_i \in \mathbb{R}, i = \overline{1, m}; j = \overline{1, n}$.

Решением системы уравнений (1) будем называть упорядоченный набор n значений неизвестных

$$x_1 = \alpha_1, x_2 = \alpha_2, \dots, x_n = \alpha_n,$$

при подстановке которых все уравнения системы обращаются в верные равенства. Система уравнений (1), имеющая хотя бы одно решение, называется совместной, а система уравнений (1), не имеющая решений, называется несовместной. Совместная система линейных уравнений (СЛУ), которая имеет единственное решение, называется определенной системой линейных уравнений. Далее матрицу $A = (a_{ij}), i = \overline{1, m}, j = \overline{1, n}$, коэффициентов при неизвестных СЛУ (1) будем называть матрицей СЛУ(1), а матрицу

$$B = \left(A \left| \begin{array}{c} b_1 \\ \vdots \\ b_m \end{array} \right. \right)$$

будем называть расширенной матрицей СЛУ(1).

Следующая теорема дает ответ на вопрос: при каких условиях система уравнений (1) совместна?

Теорема III.1.2 (теорема Кронекера – Капелли). Система уравнений (1) совместна тогда и только тогда, когда ранг матрицы системы уравнений (1) равен рангу расширенной матрицы этой системы, т.е.

$$r(A) = r(B).$$

Доказательство. Необходимость. Пусть система уравнений (1) совместна, тогда существует решение СЛУ (1):

$$x_1 = \alpha_1, x_2 = \alpha_2, \dots, x_n = \alpha_n,$$

т.е. верны равенства

$$a_{11}\alpha_1 + a_{12}\alpha_2 + \dots + a_{1n}\alpha_n = b_1,$$

$$a_{21}\alpha_1 + a_{22}\alpha_2 + \dots + a_{2n}\alpha_n = b_2,$$

$$\dots \dots \dots \dots \dots \dots$$

$$a_{m1}\alpha_1 + a_{m2}\alpha_2 + \dots + a_{mn}\alpha_n = b_m.$$

Вычитая из последнего столбца матрицы B первый ее столбец, умноженный на α_1 , затем второй столбец, умноженный на α_2 , и так далее, наконец, n -й столбец, умноженный на α_n , получим матрицу

$$C = \left(\begin{array}{c|c} & \begin{matrix} 0 \\ 0 \\ \vdots \\ 0 \end{matrix} \\ A & \end{array} \right).$$

Заметим, что ранги $r(C) = r(B)$, поскольку матрица C получена из матрицы B посредством элементарных преобразований. С другой стороны, $r(C) = r(A)$, поскольку все ненулевые миноры матрицы C равны соответствующим минорам матрицы A и наоборот (т.е. множество всех ненулевых миноров матрицы C равно множеству всех ненулевых миноров матрицы A). Следовательно, $r(B) = r(A)$.

Достаточность. Пусть $r(B) = r(A) = r$, и предположим для определенности, что отличный от нуля определитель r -го порядка матрицы A расположен в левом верхнем углу (в противном случае сделаем перестановку уравнений системы уравнений (1), т.е. перестановку строк матрицы B , а значит, и матрицы A , и/или перенумеруем неизвестные в системе уравнений (1), т.е. сделаем перестановку столбцов матрицы A , а

значит и матрицы B). Из условия следует, что базисный минор матрицы A является базисным минором матрицы B . Следовательно, определитель

$$D = \begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \cdots & \cdots & \cdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix},$$

расположенный в левом верхнем углу матрицы A , отличен от нуля и является базисным минором матриц A и B . Тогда по теореме о базисном миноре первые r строк матрицы B являются базисными, следовательно, они линейно независимы и остальные строки матрицы B линейно выражаются через эти базисные строки. Но это означает, что первые r уравнений системы уравнений (1) линейно независимы, а остальные $m-r$ уравнений являются их линейными комбинациями. Поэтому достаточно ограничиться решениями подсистемы системы уравнений (1), состоящей из первых r уравнений системы уравнений (1). Поскольку $r \leq n$, то возможны два случая.

Случай I: $r = n$.

В этом случае подсистема системы уравнений (1), состоящая из первых r уравнений, принимает вид

$$\begin{cases} a_{11}x_1 + \cdots + a_{1r}x_r = b_1, \\ \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{r1}x_1 + \cdots + a_{rr}x_r = b_r, \end{cases} \quad (2)$$

причем определитель матрицы этой системы уравнений есть базисный минор $D \neq 0$ системы уравнений (1).

Обозначим через A_{ik} алгебраическое дополнение к элементам a_{ik} в определителе D . Умножим первое уравнение системы уравнений (2) на A_{11} , второе уравнение – на A_{21} и так далее, последнее r -е уравнение – на A_{r1} и сложим их все вместе. Тогда получим

$$\begin{aligned} & A_{11}(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1r}x_r) + \\ & + A_{21}(a_{21}x_1 + a_{22}x_2 + \cdots + a_{2r}x_r) + \cdots + \\ & + A_{r1}(a_{r1}x_1 + a_{r2}x_2 + \cdots + a_{rr}x_r) = \\ & = (A_{11}a_{11} + A_{21}a_{21} + \cdots + A_{r1}a_{r1})x_1 + \end{aligned}$$

$$\begin{aligned}
& + (A_{11}a_{12} + A_{21}a_{22} + \dots + A_{r1}a_{r2})x_2 + \\
& + \dots \dots \dots \dots \dots \dots \dots \dots + \\
& + (A_{11}a_{1r} + A_{21}a_{2r} + \dots + A_{r1}a_{rr})x_r = \\
& = A_{11}b_1 + A_{21}b_2 + \dots + A_{r1}b_r.
\end{aligned}$$

Заметим, что коэффициент при x_1 , т.е. $(A_{11}a_{11} + A_{21}a_{21} + \dots + A_{r1}a_{r1})$, есть сумма произведений элементов первого столбца определителя D на их алгебраические дополнения и равен D согласно следствию 1 из теоремы I.2.22. В то же время коэффициенты при $x_2, \dots, x_r$ равны нулю по следствию 2 из теоремы I.2.22 как суммы произведений элементов одного столбца определителя D на алгебраические дополнения к другому столбцу. Следовательно, получаем линейное уравнение относительно неизвестного x_1 :

$$Dx_1 = A_{11}b_1 + A_{21}b_2 + \dots + A_{r1}b_r.$$

Обозначим через D_1 определитель, полученный из определителя D заменой его первого столбца на столбец из свободных членов $b_1, \dots, b_r$ системы уравнений (2). Тогда из разложения D_1 относительно первого столбца по следствию 1 из теоремы I.2.22 получаем

$$D_1 = b_1A_{11} + b_2A_{21} + \dots + b_rA_{r1}.$$

Следовательно, имеем

$$Dx_1 = D_1.$$

Действуя аналогично для остальных неизвестных, получаем систему уравнений

$$\begin{cases} Dx_1 = D_1, \\ Dx_2 = D_2, \\ \dots \dots \dots \\ Dx_r = D_r, \end{cases} \quad (3)$$

где D_i $i = \overline{1, r}$ есть определитель, полученный из определителя D заменой i -го столбца на столбец свободных членов $b_1, \dots, b_r$.

Поскольку $D \neq 0$, то система уравнений (3) имеет единственное решение, а именно:

$$x_1 = \frac{D_1}{D}, x_2 = \frac{D_2}{D}, \dots, x_r = \frac{D_r}{D}. \quad (4)$$

(Заметим, что формулы (4) называются формулами Крамера.)

Система уравнений (3) является следствием системы уравнений (2), поскольку каждое уравнение системы (3) является линейной комбинацией уравнений системы (2). Но тогда любое решение системы уравнений (2) является также и решением системы уравнений (3). Поэтому из единственности решения системы (3) следует, что если предположить существование решения системы (2), то оно единственно.

Покажем теперь существование решения системы уравнений (2) непосредственной подстановкой решения (4) в уравнения системы (2), например, без ограничения общности, в первое уравнение. Имеем

$$\begin{aligned} & a_{11} \frac{D_1}{D} + a_{12} \frac{D_2}{D} + \dots + a_{1r} \frac{D_r}{D} = \\ &= \frac{1}{D} [a_{11} (A_{11}b_1 + A_{21}b_2 + \dots + A_{r1}b_r) + \\ & \quad + a_{12} (A_{12}b_1 + A_{22}b_2 + \dots + A_{r2}b_r) + \\ & \quad + \dots \dots \dots \dots \dots \dots + \\ & \quad + a_{1r} (A_{1r}b_1 + A_{2r}b_2 + \dots + A_{rr}b_r)] = \\ &= \frac{1}{D} [b_1 (a_{11}A_{11} + a_{12}A_{21} + \dots + a_{1r}A_{r1}) + \\ & \quad + b_2 (a_{11}A_{12} + a_{12}A_{22} + \dots + a_{1r}A_{r2}) + \\ & \quad + \dots \dots \dots \dots \dots \dots + \\ & \quad + b_r (a_{11}A_{1r} + a_{12}A_{2r} + \dots + a_{1r}A_{rr})] = \\ &= \frac{1}{D} \cdot (b_1 D) = b_1, \end{aligned}$$

поскольку скобки при всех $\overline{b_i}$, $i = \overline{1, r}$, $i \neq 1$, равны нулю по следствию 2 из теоремы I.2.22, а скобка при b_1 равна D по следствию 1 из теоремы I.2.22.

Таким образом, система уравнений (2) имеет единственное решение, задаваемое формулами Крамера (4).

Вернемся теперь к решениям системы уравнений (1). Очевидно, что системы уравнений (1) и (2) равносильны (т.е. множества их решений совпадают). Действительно, поскольку система (2) является подсистемой системы (1), то любое решение системы (1) является также и решением системы (2). С другой стороны, поскольку любое уравнение системы (1) линейно выражается через подсистему (2), то любое решение подсистемы (2) является решением системы (1).

Следовательно, система уравнений (1) совместна и имеет единственное решение, задаваемое формулами Крамера (4).

Случай II: $r < n$.

В этом случае подсистема системы уравнений (1), состоящая из первых r линейно независимых уравнений, принимает вид

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1, \\ \dots \dots \dots \dots \dots \\ a_{r1}x_1 + \dots + a_{rn}x_n = b_r. \end{cases} \quad (5)$$

Оставим в левых частях уравнений из системы (5) первые r неизвестных, а остальные слагаемые перенесем в правые части и остальные $n - r$ неизвестных объявим свободными неизвестными:

$$\begin{cases} a_{11}x_1 + \dots + a_{1r}x_r = b_1 - a_{1,r+1}x_{r+1} - \dots - a_{1n}x_n, \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ a_{r1}x_1 + \dots + a_{rr}x_r = b_r - a_{r,r+1}x_{r+1} - \dots - a_{rn}x_n, \end{cases} \quad (6)$$

где определитель системы уравнений (6) есть базисный минор $D \neq 0$ системы уравнений (1). Для системы (6) выполняется случай I и, согласно доказанному выше, тогда система (6) имеет единственное решение, зависящее от свободных неизвестных $x_{r+1}, \dots, x_n$. Найдём это решение.

Обозначим правые части уравнений системы (6) через $c_1, \dots, c_r$, в результате будем иметь систему

$$\begin{cases} a_{11}x_1 + \dots + a_{1r}x_r = c_1, \\ \dots \dots \dots \dots \dots \\ a_{r1}x_1 + \dots + a_{rr}x_r = c_r, \end{cases} \quad (7)$$

и ее единственное решение, зависящее от свободных неизвестных $x_{r+1}, \dots, x_n$, находится по формулам Крамера. Имеем

$$x_1 = \frac{D_1(x_{r+1}, \dots, x_n)}{D}, \dots, x_r = \frac{D_r(x_{r+1}, \dots, x_n)}{D}.$$

Придавая свободным неизвестным произвольные значения из $\mathbb{R}$, получаем бесконечное множество решений системы (7), а следовательно, и бесконечное множество решений системы (1) вида

$$x_1 = \frac{D_1(x_{r+1}, \dots, \alpha_n)}{D}, \dots, x_r = \frac{D_r(\alpha_{r+1}, \dots, \alpha_n)}{D},$$

$x_{r+1} = \alpha_{r+1}, \dots, x_n = \alpha_n$ — произвольные действительные числа.

Следовательно, система уравнений (1) в этом случае есть совместная система уравнений с бесконечным числом решений.

Приведем теперь примеры применения теоремы Кронекера – Капелли к решению вопроса о совместности системы уравнений.

Пример III.1.3. Рассмотрим систему линейных уравнений:

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

1. Составляем матрицы A и B :

$$A = \begin{pmatrix} 1 & -2 & -3 & 1 \\ 2 & -3 & 2 & -1 \\ 1 & 0 & 13 & -5 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -2 & -3 & 1 & 3 \\ 2 & -3 & 2 & -1 & 5 \\ 1 & 0 & 13 & -5 & 1 \end{pmatrix}.$$

2. Вычисляем ранги матриц A и B :

$$r(A) = 2 \text{ и } r(B) = 2.$$

Действительно, для матрицы B имеем

$$\begin{aligned} \left(\begin{array}{cccc|c} 1 & -2 & -3 & 1 & 3 \\ 2 & -3 & 2 & -1 & 5 \\ 1 & 0 & 13 & -5 & 1 \end{array} \right) &\rightarrow \left(\begin{array}{cccc|c} 1 & -2 & -3 & 1 & 3 \\ 0 & 1 & 8 & -3 & -1 \\ 0 & 2 & 16 & -6 & -2 \end{array} \right) \rightarrow \\ &\rightarrow \left(\begin{array}{cccc|c} 1 & -2 & -3 & 1 & 3 \\ 0 & 1 & 8 & -3 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right). \end{aligned}$$

3. По теореме Кронекера – Капелли данная система совместна, причем поскольку $r(A) = 2 < n = 4$, то система имеет бесконечное множество решений.

Пример III.1.4. Рассмотрим систему линейных уравнений:

$$\begin{cases} x_1 + 2x_2 - x_3 = 3, \\ 2x_1 + 3x_2 + 2x_3 = 4, \\ 3x_1 - 2x_2 - 3x_3 = 5, \\ -3x_1 + 9x_2 + 7x_3 = -3. \end{cases}$$

1. Составляем матрицы A и B :

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & 2 \\ 3 & -2 & -3 \\ -3 & 9 & 7 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & 3 & 2 & 4 \\ 3 & -2 & -3 & 5 \\ -3 & 9 & 7 & -3 \end{pmatrix}.$$

2. Вычисляем ранги матриц A и B :

$$r(A) = 3 \text{ и } r(B) = 3.$$

Действительно, для матрицы B имеем

$$\begin{aligned} \left(\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 2 & 3 & 2 & 4 \\ 3 & -2 & -3 & 5 \\ -3 & 9 & 7 & -3 \end{array} \right) &\rightarrow \left(\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & -1 & 4 & -2 \\ 0 & -8 & 0 & -4 \\ 0 & 7 & 4 & 2 \end{array} \right) \rightarrow \\ &\rightarrow \left(\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 1 & -4 & 2 \\ 0 & 0 & -32 & 12 \\ 0 & 0 & 32 & -12 \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 1 & -4 & 2 \\ 0 & 0 & 1 & -\frac{3}{8} \\ 0 & 0 & 0 & 0 \end{array} \right). \end{aligned}$$

3. По теореме Кронекера – Капелли данная система совместна, причем поскольку $r(A) = n = 3$, то система имеет единственное решение, которое может быть найдено по формулам Крамера. Имеем

$$x_1 = \frac{D_1}{D} = \frac{52}{32} = \frac{13}{8},$$

$$x_2 = \frac{D_2}{D} = \frac{16}{32} = \frac{1}{2},$$

$$x_3 = \frac{D_3}{D} = -\frac{12}{32} = -\frac{3}{8},$$

где

$$D = \begin{vmatrix} 1 & 2 & -1 \\ 2 & 3 & 2 \\ 3 & -2 & -3 \end{vmatrix} = 32, \quad D_1 = \begin{vmatrix} 3 & 2 & -1 \\ 4 & 3 & 2 \\ 5 & -2 & -3 \end{vmatrix} = 52,$$

$$D_2 = \begin{vmatrix} 1 & 3 & -1 \\ 2 & 4 & 2 \\ 3 & 5 & -3 \end{vmatrix} = 16, \quad D_3 = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & -2 & 5 \end{vmatrix} = -12.$$

Пример III.1.5. Рассмотрим систему линейных уравнений:

$$\begin{cases} x_1 + 3x_2 - 2x_3 + 2x_4 = 1, \\ 2x_1 - x_2 - 3x_3 + x_4 = 3, \\ x_1 - 11x_2 - 4x_4 = 5. \end{cases}$$

1. Составляем матрицы A и B :

$$A = \begin{pmatrix} 1 & 3 & -2 & 2 \\ 2 & -1 & -3 & 1 \\ 1 & -11 & 0 & -4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 3 & -2 & 2 & 1 \\ 2 & -1 & -3 & 1 & 3 \\ 1 & -11 & 0 & -4 & 5 \end{pmatrix}.$$

2. Вычисляем ранги матриц A и B :

$$r(A) = 2 \text{ и } r(B) = 3.$$

Действительно, для матрицы B имеем

$$\begin{pmatrix} 1 & 3 & -2 & 2 & 1 \\ 2 & -1 & -3 & 1 & 3 \\ 1 & -11 & 0 & -4 & 5 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 3 & -2 & 2 & 1 \\ 0 & -7 & 1 & -3 & 1 \\ 0 & -14 & 2 & -6 & 4 \end{pmatrix} \rightarrow$$

$$\rightarrow \begin{pmatrix} 1 & 3 & -2 & 2 & 1 \\ 0 & -7 & 1 & -3 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 3 & -2 & 2 & 1 \\ 0 & 1 & -\frac{1}{7} & \frac{3}{7} & -\frac{1}{7} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

3. По теореме Кронекера – Капелли данная система несовместна.

Лекция 2. ОДНОРОДНЫЕ СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

Определение III.2.1. Система линейных уравнений, у которой свободные члены равны нулю, называется однородной системой линейных уравнений.

Таким образом, однородная система линейных уравнений имеет вид

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0, \\ a_{21}x_1 + \dots + a_{2n}x_n = 0, \\ \dots \quad \dots \quad \dots \quad \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = 0. \end{cases} \quad (1)$$

Очевидно, что однородная система линейных уравнений всегда совместна, так как имеет нулевое решение $x_1 = x_2 = \dots = x_n = 0$.

Лемма III.2.2. Однородная система линейных уравнений (1) имеет ненулевое решение тогда и только тогда, когда матрица A системы (1) удовлетворяет условию

$$r(A) < n.$$

Доказательство. Необходимость. Пусть система (1) имеет ненулевое решение. Заметим, что ранг матрицы A удовлетворяет условию $r(A) \leq n$. Пусть $r(A) = n$, тогда в силу $r(A) = r(B) = n$ из доказательства достаточности теоремы Кронекера – Капелли (случай 1) следует, что система (1) имеет единственное решение. Следовательно, нулевое решение является единственным решением системы (1), что противоречит предположению. Тогда получаем $r(A) < n$.

Достаточность. Пусть $r(A) < n$, тогда из доказательства достаточности теоремы Кронекера – Капелли (случай 2) следует, что система (1) имеет бесконечное множество решений. Но тогда система (1) имеет ненулевое решение.

Лемма III.2.3. Однородная система линейных уравнений (1), для которой $m = n$, имеет ненулевое решение тогда и только тогда, когда

$|A| = 0$. Доказательство следует из леммы III.2.2 и того факта, что определитель квадратной матрицы A порядка n равен нулю тогда и только тогда, когда $r(A) < n$.

Пусть $x_1 = \alpha_1, x_2 = \alpha_2, \dots, x_n = \alpha_n$ – некоторое решение однородной системы линейных уравнений (1). Его можно рассматривать как элемент $(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{R}^{(n)}$.

Лемма III.2.4. Множество P всех решений однородной системы линейных уравнений (1) является подпространством в $\mathbb{R}^{(n)}$.

Доказательство. По критерию подпространства достаточно проверить, что для любых $x, y \in P$ и $\alpha \in \mathbb{R}$ выполняется $x - y \in P$ и $\alpha x \in P$. Имеем

$$x = (\alpha_1, \alpha_2, \dots, \alpha_n), y = (\beta_1, \beta_2, \dots, \beta_n) -$$

решения системы (1), тогда

$$x - y = (\alpha_1 - \beta_1, \alpha_2 - \beta_2, \dots, \alpha_n - \beta_n),$$

и, подставляя координаты $x - y$ в систему (1), получаем

$$a_{11}(\alpha_1 - \beta_1) + \dots + a_{1n}(\alpha_n - \beta_n) = (a_{11}\alpha_1 + \dots + a_{1n}\alpha_n) -$$

$$-(a_{11}\beta_1 + \dots + a_{1n}\beta_n) = 0 - 0 = 0,$$

... ..

$$a_{m1}(\alpha_1 - \beta_1) + \dots + a_{mn}(\alpha_n - \beta_n) = (a_{m1}\alpha_1 + \dots + a_{mn}\alpha_n) -$$

$$-(a_{m1}\beta_1 + \dots + a_{mn}\beta_n) = 0 - 0 = 0.$$

Следовательно, $x - y \in P$.

Аналогично проверяется, что αx есть решение (1), т.е. $\alpha x \in P$.

Поскольку $\dim \mathbb{R}^{(n)} = n$, то $\dim P \leq n$.

Определение III.2.5. Базис подпространства P всех решений однородной системы уравнений (1) называется фундаментальной системой решений (ФСР) системы (1). Обозначим через d размерность подпространства P , т.е. $d = \dim P$.

Теорема III.2.6 (теорема о фундаментальной системе решений). Пусть n – число неизвестных, r – ранг матрицы однородной системы линейных уравнений (1), $d = \dim P$ – размерность подпространства всех решений системы (1). Тогда

$$d = n - r.$$

Доказательство. Поскольку $r \leq n$, то рассмотрим сначала случай, когда $r = n$. По лемме III.2.2 в этом случае получаем, что система (1) имеет единственное решение, а именно нулевое. Тогда $P = 0$ и $d = \dim P = 0 = n - r$.

Рассмотрим теперь случай, когда $r < n$. Предположим, для определенности, что базисный минор $D \neq 0$ порядка r находится в левом верхнем углу матрицы A системы (1), т.е.

$$D = \begin{vmatrix} a_{11} & \dots & a_{1r} \\ \dots & \dots & \dots \\ a_{r1} & \dots & a_{rr} \end{vmatrix} \neq 0.$$

Перенесем слагаемые со свободными неизвестными $x_{r+1}, \dots, x_n$ в первых r уравнениях системы (1) в правую часть со сменой знака коэффициентов этих слагаемых, тогда получим систему уравнений:

$$\begin{cases} a_{11}x_1 + \dots + a_{1r}x_r = b_1(x_{r+1}, \dots, x_n), \\ a_{21}x_1 + \dots + a_{2r}x_r = b_2(x_{r+1}, \dots, x_n), \\ \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ a_{r1}x_1 + \dots + a_{rr}x_r = b_r(x_{r+1}, \dots, x_n), \end{cases} \quad (2)$$

где правые части уравнений системы (2) имеют вид

$$b_1(x_{r+1}, \dots, x_n) = -a_{1,r+1}x_{r+1} - \dots - a_{1n}x_n,$$

$$b_2(x_{r+1}, \dots, x_n) = -a_{2,r+1}x_{r+1} - \dots - a_{2n}x_n,$$

$$\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots$$

$$b_r(x_{r+1}, \dots, x_n) = -a_{r,r+1}x_{r+1} - \dots - a_{rn}x_n.$$

Придадим свободным неизвестным следующие значения:

$$x_{r+1} = 1, x_{r+2} = 0, \dots, x_n = 0,$$

тогда обозначим через

$$b_1(1, 0, \dots, 0) = c^{(1)}, b_2(1, 0, \dots, 0) = c^{(2)}, \dots, b_r(1, 0, \dots, 0) = c^{(r)}.$$

Имеем систему линейных уравнений:

$$\begin{cases} a_{11}x_1 + \dots + a_{1r}x_r = c^{(1)}, \\ a_{21}x_1 + \dots + a_{2r}x_r = c^{(2)}, \\ \dots \quad \dots \quad \dots \quad \dots \\ a_{r1}x_1 + \dots + a_{rr}x_r = c^{(r)}, \end{cases} \quad (3)$$

определитель матрицы которой равен $D \neq 0$. Тогда по формулам Крамера получаем решение системы (3):

$$x_1 = \frac{D_1}{D} = \alpha_1, x_2 = \frac{D_2}{D} = \alpha_2, \dots, x_r = \frac{D_r}{D} = \alpha_r.$$

Отсюда получаем следующее частное решение системы (1):

$$e_1 = (\alpha_1, \alpha_2, \dots, \alpha_r, 1, 0, \dots, 0).$$

Аналогично, придавая свободным неизвестным значения

$$x_{r+1} = 0, x_{r+2} = 1, x_{r+3} = \dots = x_n = 0,$$

получаем частное решение системы (1):

$$e_2 = (\beta_1, \dots, \beta_r, 0, 1, 0, \dots, 0).$$

Продолжая таким образом, найдем систему из $n-r$ векторов, являющихся частными решениями системы (1):

$$\begin{cases} e_1 = (\alpha_1, \alpha_2, \dots, \alpha_r, 1, 0, \dots, 0), \\ e_2 = (\beta_1, \beta_2, \dots, \beta_r, 0, 1, \dots, 0), \\ \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ e_{n-r} = (\gamma_1, \gamma_2, \dots, \gamma_r, 0, 0, \dots, 1). \end{cases} \quad (4)$$

Покажем, что система векторов (4) является базисом подпространства P , т.е. (4) есть фундаментальная система решений однородной системы линейных уравнений (1).

Необходимо доказать, что

а) система векторов (4) линейно независима;

б) любое решение системы уравнений (1) линейно выражается через систему векторов (4).

а) Система векторов (4) линейно независима, поскольку ранг матрицы, строки которой составлены из векторов системы (4), а именно

$$S = \begin{pmatrix} \alpha_1, \alpha_2, \dots, \alpha_r, 1, 0, \dots, 0 \\ \beta_1, \beta_2, \dots, \beta_r, 0, 1, \dots, 0 \\ \dots & \dots & \dots & \dots \\ \gamma_1, \gamma_2, \dots, \gamma_r, 0, 0, \dots, 1 \end{pmatrix},$$

равен $n - r$. Действительно, матрица S имеет $n - r$ строк, n столбцов и содержит отличный от нуля минор порядка $n - r$, а именно

$$\begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{vmatrix} = 1.$$

Из того, что ранг матрицы S равен $n - r$, следует, что построчный ранг матрицы S равен $n - r$. Но тогда система $n - r$ векторов-строк матрицы S совпадает со своим базисом и, следовательно, является линейно независимой. С другой стороны, система векторов-строк матрицы S совпадает с системой векторов (4). Таким образом, система векторов (4) линейно независима.

б) Пусть $e = (\eta_1, \eta_2, \dots, \eta_n)$ – произвольное решение системы уравнений (1).

Рассмотрим следующую линейную комбинацию решений системы уравнений (1):

$$e_0 = e - \eta_{r+1}e_1 - \dots - \eta_n e_{n-r}.$$

Найдем представление e_0 в виде вектор-строки:

$$\begin{aligned} e_0 &= (\eta_1, \eta_2, \dots, \eta_n) - \eta_{r+1}(\alpha_1, \dots, \alpha_r, 1, 0, \dots, 0) - \\ &\quad - \eta_{r+2}(\beta_1, \dots, \beta_r, 0, 1, \dots, 0) - \dots - \\ &\quad - \eta_n(\gamma_1, \dots, \gamma_r, 0, 0, \dots, 1) = (\eta_1, \eta_2, \dots, \eta_n) - \\ &\quad - (\eta_{r+1}\alpha_1, \dots, \eta_{r+1}\alpha_r, \eta_{r+1}, 0, \dots, 0) - \\ &\quad - (\eta_{r+2}\beta_1, \dots, \eta_{r+2}\beta_r, 0, \eta_{r+2}, \dots, 0) - \dots - \\ &\quad - (\eta_n\gamma_1, \dots, \eta_n\gamma_r, 0, 0, \dots, \eta_n) = (\rho_1, \rho_2, \dots, \rho_r, 0, 0, \dots, 0), \end{aligned}$$

где

$$\begin{aligned}\rho_1 &= \eta_1 - \eta_{r+1}\alpha_1 - \eta_{r+2}\beta_1 - \dots - \eta_n\gamma_1, \\ \rho_2 &= \eta_2 - \eta_{r+1}\alpha_2 - \eta_{r+2}\beta_2 - \dots - \eta_n\gamma_2, \\ &\dots \dots \dots \dots \dots \dots \\ \rho_r &= \eta_r - \eta_{r+1}\alpha_r - \eta_{r+2}\beta_r - \dots - \eta_n\gamma_r.\end{aligned}$$

Заметим, что $e_0 \in P$ как линейная комбинация векторов из подпространства P всех решений системы уравнений (1), т.е. вектор e_0 является решением системы уравнений (1). Далее, из $e_0 = (\rho_1, \rho_2, \dots, \rho_r, 0, 0, \dots, 0)$ следует, что решение e_0 получено из системы уравнений (2) для следующих значений свободных неизвестных:

$$x_{r+1} = 0, x_{r+2} = 0, \dots, x_n = 0.$$

Но тогда

$$\begin{aligned}b_1(0, 0, \dots, 0) &= 0, \\ b_2(0, 0, \dots, 0) &= 0, \\ b_r(0, 0, \dots, 0) &= 0\end{aligned}$$

и, следовательно,

$$x_1 = \rho_1, x_2 = \rho_2, \dots, x_r = \rho_r$$

является решением системы уравнений

$$\begin{cases} a_{11}x_1 + \dots + a_{1r}x_r = 0, \\ \dots \dots \dots \dots \dots \dots \\ a_{r1}x_1 + \dots + a_{rr}x_r = 0. \end{cases} \quad (5)$$

Но однородная система линейных уравнений (5) имеет определитель $D \neq 0$, следовательно, имеет единственное нулевое решение. Тогда

$$\rho_1 = \rho_2 = \dots = \rho_r = 0,$$

и получаем

$$e_0 = (0, 0, \dots, 0).$$

Отсюда следует, что

$$e = \eta_{r+1}e_1 + \dots + \eta_n e_{n-r}.$$

Таким образом, система векторов (4) является базисом подпространства P , и тогда $\dim P = n - r$.

Пример III.2.7. Найдем ФСР однородной системы линейных уравнений

$$\begin{cases} x_1 + 2x_2 - x_3 = 0, \\ 2x_1 + 4x_2 - 2x_3 = 0, \\ 3x_1 + 6x_2 - 3x_3 = 0. \end{cases}$$

1. Находим ранг матрицы A коэффициентов:

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ 3 & 6 & -3 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Следовательно, $r(A) = 1$.

2. Находим число векторов в ФСР:

$$d = \dim P = n - r = 3 - 1 = 2.$$

3. Находим общее решение системы:

$$x_1 = -2x_2 + x_3.$$

4. Находим ФСР:

	x_1	x_2	x_3
a_1	-2	1	0
a_2	1	0	1

5. Таким образом, любое решение системы имеет вид

$$a = \lambda_1 a_1 + \lambda_2 a_2,$$

где $\lambda_1, \lambda_2 \in \mathbb{R}$.

Найдем связь между решениями однородной и произвольной системами линейных уравнений.

Рассмотрим матричные формы записи этих систем уравнений:

$$(*) \quad AX = b,$$

где

$$A = (a_{ij}), i = \overline{1, m}, j = \overline{1, n} \text{ — матрица системы } (*),$$

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} - \text{столбец неизвестных,}$$

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} - \text{столбец свободных членов,}$$

$$(**) AX = 0 -$$

ассоциированная с (*) однородная система линейных уравнений, где A и X – такие же, как и в (*), и вместо b – столбец из m нулей.

Пусть далее P есть подпространство всех решений однородной системы (**), и Q – множество всех решений произвольной системы линейных уравнений (*). Следующая теорема устанавливает связь между множествами P и Q .

Теорема III.2.8. Пусть Q и P есть множества всех решений системы линейных уравнений (*) и соответствующей ассоциированной однородной системы линейных уравнений (**). Если x_0 – произвольный фиксированный элемент Q , то имеет место равенство

$$Q = x_0 + P = \left\{ y \in \mathbb{R}^{(n)} \mid \exists y_0 \in P, y = x_0 + y_0 \right\}.$$

Доказательство. Включение

$$x_0 + P \subseteq Q$$

имеет место, поскольку при подстановке в систему уравнений (*) произвольного элемента

$$y = x_0 + y_0 \in x_0 + P,$$

где $y_0 \in P$, $x_0 \in Q$, убеждаемся, что $y \in Q$.

Для доказательства включения

$$Q \subseteq x_0 + P$$

пусть $y \in Q$ – произвольный элемент. Тогда

$$A(y - x_0) = Ay - Ax_0 = b - b = 0$$

и, следовательно, $y_0 = y - x_0 \in P$. Отсюда

$$y = x_0 + y_0 \in x_0 + P.$$

Лекция 3. МЕТОД ГАУССА РЕШЕНИЯ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ

Рассмотрим систему линейных уравнений с действительными коэффициентами:

[illegible]

Очевидно, можем считать, что среди коэффициентов a_{ij} системы (1) есть ненулевые.

Шаг 1. Перестановкой уравнений или изменением нумерации неизвестных система (1) приводится к виду, где $a_{11} \neq 0$.

Шаг 2. Умножаем левую и правую части первого уравнения на $1/a_{11}$, в результате коэффициент при x_1 станет равным 1.

Шаг 3. Из каждого j -го уравнения, где $j = 2, 3, \dots, m$, вычитаем почленно первое уравнение, левая и правая части которого предварительно умножены на a_{1j} . В результате система (1) примет вид

[illegible]

Указанные выше три шага образуют I этап исключения неизвестных (в данном случае неизвестное x_1 исключено из всех уравнений, кроме первого).

Рассмотрим подсистему системы (2), состоящую из всех её уравнений, кроме первого, и применим к этой подсистеме три шага, аналогичные тем, которые применялись при исключении неизвестной x_1 . Они

образуют II этап исключения неизвестных, и в результате его получаем систему уравнений:

$$\left\{ \begin{array}{l} x_1 + a'_{12}x_2 + a'_{13}x_3 + \dots + a'_{1n}x_n = b'_1, \\ x_2 + a''_{23}x_3 + \dots + a''_{2n}x_n = b''_2, \\ a''_{33}x_3 + \dots + a''_{3n}x_n = b''_3, \\ \dots\dots\dots \\ a''_{m3}x_3 + \dots + a''_{mn}x_n = b''_m. \end{array} \right. \quad (3)$$

Продолжая процесс исключений неизвестных, в конечном счёте приходим к системе уравнений:

$$\left\{ \begin{array}{l} x_1 + a'_{12}x_2 + a'_{13}x_3 + \dots + a'_{1n}x_n = b'_1, \\ x_2 + a''_{23}x_3 + \dots + a''_{2n}x_n = b''_2, \\ \dots\dots\dots \\ x_r + \dots + a^{(r)}_{rn}x_n = b^{(r)}_r, \\ 0 = b^{(r)}_{r+1}, \\ \dots\dots\dots \\ 0 = b^{(r)}_m. \end{array} \right. \quad (4)$$

Говорят, что система (4) имеет ступенчатый вид.

Для системы (4) можно легко решить, совместна первоначальная система (1) или нет. Действительно, если хотя бы одно из чисел $b^{(r)}_{r+1}, \dots, b^{(r)}_m$ отлично от нуля, то получим противоречивое равенство (отличное от нуля число равно нулю), и в этом случае система (1) несовместна.

Пусть $b^{(r)}_{r+1} = \dots = b^{(r)}_m = 0$, и покажем в этом случае, что система (4), а следовательно и система (1), совместна, нахождением решения посредством процедуры, которая называется обратным ходом метода Гаусса, а процедура приведения системы уравнений к ступенчатому виду (4) называется прямым ходом метода Гаусса.

Из последнего ненулевого уравнения выражаем неизвестное x_r , через неизвестные $x_{r+1}, \dots, x_n$, которые объявляются свободными неизвестными:

$$x_r = b^{(r)}_r - a^{(r)}_{rr+1}x_{r+1} - \dots - a^{(r)}_{rn}x_n. \quad (5)$$

Затем поднимаемся на одну ступеньку выше и выражаем неизвестные x_{r-1} через свободные неизвестные $x_{r+1}, \dots, x_n$ (при этом вместо x_r подставляем правую часть из (5)).

Продолжаем подниматься по ступеньки вида (4) и в итоге получим общее решение системы (4), а следовательно и решение системы (1), в котором главные неизвестные $x_1, \dots, x_r$ выражены как линейные комбинации свободных неизвестных $x_{r+1}, \dots, x_n$.

Замечание 1. Число ступенек r в системе (4) равно рангу матрицы системы (1), поскольку все преобразования, которые применялись при переходе от системы (1) к системе (4), не меняют ранг матрицы системы.

Замечание 2. Из критерия совместности системы линейных уравнений и замечания 1 вытекает, что необходимым и достаточным условием совместности системы (1) является равенство нулю правых частей уравнений системы (4), которые располагаются ниже r ненулевых ступенек, т.е.

$$b_{r+1}^{(r)} = \dots = b_m^{(r)} = 0.$$

Замечание 3. Если $r = n$, то свободных неизвестных нет, и обратным ходом метода Гаусса получается единственное решение системы (1).

Рассмотрим теперь примеры решения методом Гаусса совместных для $r = n$ и $r < n$ систем линейных уравнений.

Пример III.3.1. Рассмотрим систему линейных уравнений:

$$\begin{cases} 2x_1 + x_2 + x_3 + x_4 = 1, \\ 3x_1 + 4x_2 - x_3 - x_4 = 0, \\ x_1 + 3x_2 - x_3 + x_4 = 2, \\ 5x_1 - 3x_2 + 6x_3 + 3x_4 = 3, \\ 2x_1 - 7x_2 + 7x_3 + 4x_4 = 3. \end{cases} \quad (*)$$

Если действовать строго по схеме метода Гаусса, то на шаге 2 этапа (I) исключения неизвестных обе части первого уравнения нужно разделить на $a_{11} = 2$ и дальше работать с дробями. Чтобы как можно дольше избегать появления дробей в «ручных» вычислениях, можем применять перестановки уравнений системы (*) или переставлять неизвестные вместе с коэффициентами так, чтобы в левом верхнем углу появилось неизвестное с коэффициентом ± 1 . Например, поменять местами первое

и третье уравнения или во всех уравнениях поменять местами x_1 и x_2 вместе со своими коэффициентами. Выберем первый способ, тогда получим систему уравнений:

$$\begin{cases} x_1 + 3x_2 - x_3 + x_4 = 2, \\ 3x_1 + 4x_2 - x_3 - x_4 = 0, \\ 2x_1 + x_2 + x_3 + x_4 = 1, \\ 5x_1 - 3x_2 + 6x_3 + 3x_4 = 3, \\ 2x_1 - 7x_2 + 7x_3 + 4x_4 = 3. \end{cases}$$

В результате исключения неизвестного x_1 из всех уравнений, кроме первого, получим систему уравнений:

$$\begin{cases} x_1 + 3x_2 - x_3 + x_4 = 2, \\ -5x_2 + 2x_3 - 4x_4 = -6, \\ -5x_2 + 3x_3 - x_4 = -3, \\ -18x_2 + 11x_3 - 2x_4 = -7, \\ -13x_2 + 9x_3 + 2x_4 = -1. \end{cases} \quad (*1)$$

Если на этапе (II) исключения неизвестных действовать строго по схеме метода Гаусса, то получатся дроби. Используем наличие в третьем уравнении слагаемого $-x_4$ и переставим во всех уравнениях системы (*1) слагаемые с неизвестными x_2 и x_4 , а затем поменяем местами второе и третье уравнения, тогда получим систему уравнений (дополнительно обе части уравнения со слагаемым $-x_4$ умножаем на -1):

$$\begin{cases} x_1 + x_4 - x_3 + 3x_2 = 2, \\ x_4 - 3x_3 + 5x_2 = 3, \\ -4x_4 + 2x_3 - 5x_2 = -6, \\ -2x_4 + 11x_3 - 18x_2 = -7, \\ 2x_4 + 9x_3 - 13x_2 = -1. \end{cases}$$

После исключения неизвестной x_4 из всех уравнений этой системы, кроме первых двух, получим систему уравнений:

$$\begin{cases} x_1 + x_4 - x_3 + 3x_2 = 2, \\ x_4 - 3x_3 + 5x_2 = 3, \\ -10x_3 + 15x_2 = 6, \\ 5x_3 - 8x_2 = -1, \\ 15x_3 - 23x_2 = -7. \end{cases}$$

На третьем этапе исключения неизвестных отсрочим появление дробей перестановкой третьего и четвертого уравнений:

$$\begin{cases} x_1 + x_4 - x_3 + 3x_2 = 2, \\ x_4 - 3x_3 + 5x_2 = 3, \\ 5x_3 - 8x_2 = -1, \\ -10x_3 + 15x_2 = 6, \\ 15x_3 - 23x_2 = -7. \end{cases} \quad (*)2$$

Исключаем неизвестное x_3 из всех уравнений системы $(*)2$, кроме первых трёх (можем не делить третье уравнение на коэффициент 5 при x_3 , поскольку в четвёртом и пятом уравнениях коэффициенты при x_3 кратны 5):

$$\begin{cases} x_1 + x_4 - x_3 + 3x_2 = 2, \\ x_4 - 3x_3 + 5x_2 = 3, \\ 5x_3 - 8x_2 = -1, \\ -x_2 = 4, \\ x_2 = -4. \end{cases} \quad (*)3$$

Меняя местами четвёртое и пятое уравнения и исключая неизвестное x_2 из последнего уравнения, окончательно получим систему уравнений ступенчатого вида:

$$\begin{cases} x_1 + x_4 - x_3 + 3x_2 = 2, \\ x_4 - 3x_3 + 5x_2 = 3, \\ 5x_3 - 8x_2 = -1, \\ x_2 = -4, \\ 0 = 0. \end{cases} \quad (*)4$$

Поскольку ниже ненулевых уравнений имеется равенство $0 = 0$, то система (*4), а следовательно и система (*), совместна, и можем обратным ходом Гаусса найти единственное решение этой системы ($r = n = 4$).

Итак, $x_2 = -4$, и подставляя это значение x_2 в следующее уравнение, получим

$$5x_3 + 32 = -1, \quad x_3 = -\frac{33}{5}.$$

Продолжая подобным образом, получим $x_4 = \frac{16}{5}$ и $x_1 = \frac{21}{5}$.

Пример III.3.2. Рассмотрим систему уравнений:

$$\begin{cases} x_1 + 15x_2 + 5x_3 + 9x_4 = 0, \\ x_1 - 3x_2 - x_3 - 2x_4 = 0, \\ 2x_1 + x_2 - 3x_3 + 4x_4 = 1, \\ 3x_1 + 2x_2 + 4x_3 - 3x_4 = -1, \\ 3x_1 - 2x_2 - 4x_3 + 2x_4 = 1. \end{cases}$$

Меняем местами первое и второе уравнения и исключаем неизвестное x_1 из всех уравнений, кроме первого:

$$\begin{cases} x_1 - 3x_2 - x_3 - 2x_4 = 0, \\ 18x_2 + 6x_3 + 11x_4 = 0, \\ 7x_2 - x_3 + 8x_4 = 1, \\ 11x_2 + 7x_3 + 3x_4 = -1. \end{cases}$$

Меняем местами неизвестные x_2 и x_3 , а также второе и третье уравнения:

$$\begin{cases} x_1 - x_3 - 3x_2 - 2x_4 = 0, \\ -x_3 + 7x_2 + 8x_4 = 1, \\ 6x_3 + 18x_2 + 11x_4 = 0, \\ 7x_3 + 11x_2 + 3x_4 = -1. \end{cases}$$

Исключаем неизвестную x_3 из третьего и четвёртого уравнений:

$$\begin{cases} x_1 - x_3 - 3x_2 - 2x_4 = 0, \\ -x_3 + 7x_2 + 8x_4 = 1, \\ 60x_2 + 59x_4 = 6, \\ 60x_2 + 59x_4 = 6. \end{cases}$$

Исключаем неизвестную x_2 из последнего уравнения:

$$\begin{cases} x_1 - x_3 - 3x_2 - 2x_4 = 0, \\ -x_3 + 7x_2 + 8x_4 = 1, \\ 60x_2 + 59x_4 = 6, \\ 0 = 0. \end{cases}$$

Обратным ходом метода Гаусса находим общее решение:

$$\begin{aligned} x_2 &= \frac{1}{10} - \frac{56}{60}x_4, \\ x_3 &= -\frac{3}{10} - \frac{67}{60}x_4, \\ x_1 &= \frac{1}{6}x_4. \end{aligned}$$

Пример III.3.3. Рассмотрим систему уравнений:

$$\begin{cases} -3x_1 + 2x_2 - x_3 + x_4 + 4x_5 = 1, \\ 2x_1 - x_2 + 3x_3 + 2x_4 - x_5 = 2, \\ -5x_1 + 4x_2 + 2x_3 + x_4 + 3x_5 = 3, \\ -19x_1 + 15x_2 + 6x_3 + x_4 + 9x_5 = 1. \end{cases}$$

Ставим неизвестную x_4 во всех уравнениях на первую позицию и исключаем её из всех уравнений, кроме первого:

$$\begin{cases} x_4 - 3x_1 + 2x_2 - x_3 + 4x_5 = 1, \\ 8x_1 - 5x_2 + 5x_3 - 9x_5 = 0, \\ -2x_1 + 2x_2 + 3x_3 - x_5 = 2, \\ -16x_1 + 13x_2 + 7x_3 + 5x_5 = 0. \end{cases}$$

Слагаемые с неизвестной x_5 ставим на вторую позицию во всех уравнениях и меняем местами второе и третье уравнения, затем исключаем неизвестную x_5 :

$$\begin{cases} x_4 + 4x_5 - 3x_1 + 2x_2 - x_3 = 1, \\ -x_5 - 2x_1 + 2x_2 + 3x_3 = 2, \\ -9x_5 + 8x_1 - 5x_2 + 5x_3 = 0, \\ 5x_5 - 16x_1 + 13x_2 + 7x_3 = 0; \end{cases}$$

$$\begin{cases} x_4 + 4x_5 - 3x_1 + 2x_2 - x_3 = 1, \\ -x_5 - 2x_1 + 2x_2 + 3x_3 = 2, \\ 26x_1 - 23x_2 - 22x_3 = -18, \\ -26x_1 + 23x_2 + 22x_3 = 10. \end{cases}$$

Исключаем неизвестную x_1 :

$$\begin{cases} x_4 + 4x_5 - 3x_1 + 2x_2 - x_3 = 1, \\ -x_5 - 2x_1 + 2x_2 + 3x_3 = 2, \\ 26x_1 - 23x_2 - 22x_3 = -18, \\ 0 = -8. \end{cases}$$

Следовательно, первоначальная система уравнений несовместна.

Лекция 4. ТРЕУГОЛЬНОЕ РАЗЛОЖЕНИЕ МАТРИЦ (матричная интерпретация метода Гаусса)

Получим теперь матричное представление метода Гаусса, а именно треугольное разложение матрицы системы уравнений, которое позволит дать обоснование метода Гаусса.

Определение III.4.1. Назовем две системы линейных уравнений равносильными, если их множества решений совпадают.

Метод Гаусса состоит из последовательности преобразований системы уравнений, и если мы сможем установить, что на каждом шаге метода Гаусса происходит переход к равносильной системе, то тем самым будет получено обоснование метода Гаусса.

Для упрощения изложения рассмотрим частный случай системы линейных уравнений с матрицей, ранг которой совпадает с числом неизвестных, т.е. $r = n$:

[illegible]

где $A = (a_{ij})$ – невырожденная матрица.

Запишем систему (1) в матричном виде

$$Ax = b, \quad (1')$$

где $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$.

Среди преобразований метода Гаусса есть перестановки уравнений или неизвестных, в матричном виде – это перестановки строк или столбцов матрицы A с соответствующей перестановкой компонент векторов x и b .

Пример III.4.2. Пусть система уравнений имеет вид

$$Ax = b,$$

$$\text{где } A = \begin{pmatrix} 3 & -2 & 3 \\ 5 & -4 & 1 \\ -2 & 3 & 2 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}.$$

При реализации первого шага метода Гаусса необходимо переставить местами первое и второе уравнение, а также неизвестные x_1 и x_3 вместе с их коэффициентами, тогда после этих преобразований $a_{11} = 1$. Для их осуществления нужно переставить в матрице A первую и вторую строки, а также первый и третий столбец, кроме того, в векторе x поменять местами x_1 и x_3 , а также в векторе b поменять местами первую и вторую компоненты.

В результате получим систему уравнений

$$A'x' = b',$$

$$\text{где } A' = \begin{pmatrix} 1 & -4 & 5 \\ 3 & -2 & 3 \\ 2 & 3 & -2 \end{pmatrix}; \quad x' = \begin{pmatrix} x_3 \\ x_2 \\ x_1 \end{pmatrix}; \quad b' = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Рассмотрим $n \times n$ -матрицу перестановок $P_{ij} = (p_{ij})$, где

$$p_{ii} = p_{jj} = 0, \quad p_{ij} = p_{ji} = 1, \quad p_{kk} = 1, \quad k \neq i, j,$$

а все остальные элементы равны нулю. Можно проверить, что умножение матрицы A на матрицу P_{ij} справа дает перестановку i -го и j -го столбцов матрицы A , а умножение матрицы A слева на P_{ij} дает перестановку i -й и j -й строк матрицы A .

Пример III.4.3. Пусть A – матрица из примера III.4.2. Рассмотрим матрицы P_{13} и P_{12} :

$$P_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad P_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Тогда матрицы

$$P_{12}A = \begin{pmatrix} 5 & -4 & 1 \\ 3 & -2 & 3 \\ -2 & 3 & 2 \end{pmatrix}, \quad AP_{13} = \begin{pmatrix} 3 & -2 & 3 \\ 1 & -4 & 5 \\ 2 & 3 & -2 \end{pmatrix} \quad \text{и} \quad P_{12}AP_{13} = \begin{pmatrix} 1 & -4 & 5 \\ 3 & -2 & 3 \\ 2 & 3 & -2 \end{pmatrix}$$

совпадают с матрицами, полученными из матрицы A соответственно перестановкой первой и второй строки, первого и третьего столбцов, а также первой и второй строки и первого и второго столбцов.

Заметим, что любая матрица перестановок содержит в точности одну единицу в каждой строке и каждом столбце, поэтому ее определитель равен ± 1 , следовательно, матрица перестановок есть невырожденная матрица. Более того, поскольку повторная перестановка той же пары строк или столбцов возвращает матрицу в исходное состояние, то для любой матрицы перестановок P_{ij} имеем $P_{ij}^2 = E_n$, где E_n – единичная матрица порядка n .

Следующее преобразование метода Гаусса – умножение левой и правой частей i -го уравнения системы на a_{ii}^{-1} . Это преобразование можно представить как умножение обеих частей матричного уравнения

$$Ax = b$$

слева на невырожденную матрицу

$$C_i = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & a_{ii}^{-1} & & \\ & 0 & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

Наконец, еще одно преобразование метода Гаусса – вычитание i -го уравнения системы из всех последующих, при этом необходимо предварительно умножить его на a_{ji} для каждого j -го уравнения соответственно, где $j = \overline{i+1, \dots, n}$. Это преобразование можно представить как умножение обеих частей матричного уравнения

$$Ax = b$$

слева на невырожденную матрицу

$$L_i = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & -a_{i+1,i} & 1 & \\ 0 & & \vdots & & \ddots \\ & & -a_{ni} & & 1 \end{pmatrix}.$$

Заметим, что обратной матрицей для L_i^{-1} будет следующая матрица:

$$L_i^{-1} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & 0 \\ & & a_{i+1,j} & 1 & & \\ 0 & & \vdots & & \ddots & \\ & a_{ni} & & & & 1 \end{pmatrix}.$$

Теперь можно дать матричную интерпретацию метода Гаусса.

Итак, рассмотрим систему уравнений в виде матричного уравнения $Ax = b$.

1 этап исключения

Шаг 1. Перестановку первого и j -го уравнения осуществляем умножением обеих частей матричного уравнения на матрицу P_{1j} слева:

$$P_{1j}Ax = P_{1j}b.$$

Перестановку первого и k -го столбцов матрицы A осуществляем совместно с перестановкой компонент x_1 и x_k вектора x :

$$(P_{1j}AP_{1k}) \cdot (P_{1k}x) = P_{1j}b.$$

Действительно, так как $P_{1k}^2 = E$, то имеем

$$(P_{1j}AP_{1k})(P_{1k}x) = (P_{1j}A)(P_{1k}P_{1k}x) = P_{1j}Ax.$$

Поскольку очевидно, что перестановка уравнений и/или слагаемых уравнений (т.е. коэффициентов совместно с неизвестными) не меняет множество решений системы уравнений, то в дальнейшем при изложении метода Гаусса преобразования этого вида для упрощения записи учитывать не будем.

Шаг 2. Согласно предположению, сделанному в шаге 1, в результате этого шага имеем

$$C_1Ax = C_1b,$$

где

$$C_1 = \begin{pmatrix} a_{11}^{-1} & & & & \\ & 1 & 0 & & \\ & 0 & \ddots & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix}.$$

Шаг 3. В результате этого шага имеем

$$L_1 C_1 A x = L_1 C_1 b,$$

где

$$L_1 = \begin{pmatrix} 1 & & & \\ -a_{21} & 1 & 0 & \\ \vdots & 0 & \ddots & \\ -a_{n1} & & & 1 \end{pmatrix}.$$

Обозначим

$$L_1 C_1 A = A_1 = (a_{ij}^{(1)}),$$

$$L_1 C_1 b = b^{(1)} = \begin{pmatrix} b_1^{(1)} \\ \vdots \\ b_n^{(1)} \end{pmatrix}.$$

В результате I этапа исключения получаем систему уравнений

$$A_1 x = b^{(1)},$$

в которой неизвестная x_1 (мы решили не учитывать перестановки неизвестных) исключена из всех уравнений, кроме первого.

II этап исключения

Шаг 1. Не учитываем.

Шаг 2. В результате этого шага получаем

$$C_2 A_1 x = C_2 b^{(1)},$$

где

$$C_2 = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & (a_{22}^{(1)})^{-1} & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Шаг 3. В результате этого шага получаем

$$L_2 C_2 A_1 x = L_2 C_2 b^{(1)},$$

где

$$L_2 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & -a_{32}^{(1)} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & -a_{n2}^{(1)} & 0 & \cdots & 1 \end{pmatrix}.$$

Обозначим $L_2 C_2 A_1 = A_2 = (a_{ij}^{(2)})$,

$$L_2 C_2 b^{(1)} = b^{(2)} = \begin{pmatrix} b_1^{(2)} \\ \vdots \\ b_n^{(2)} \end{pmatrix}.$$

И в результате II этапа исключения имеем систему уравнений

$$A_2 x = b^{(2)},$$

в которой проведено исключение для неизвестных x_1 и x_2 (напомним, что было решено не учитывать возможные перестановки неизвестных).

В итоге после всех этапов исключений прямой ход метода Гаусса дает систему уравнений

$$A_n x = b^{(n)},$$

где $A_n = L_n C_n \dots L_1 C_1 A$ есть верхнетреугольная матрица, и обозначим $A_n = R$.

Выразим теперь матрицу A через матрицу R и какие-то другие матрицы. Для этого нужно сначала вычислить обратные матрицы для L_i и C_i . Матрица L_i^{-1} была приведена выше, найдем C_i^{-1} .

Можно убедиться непосредственной проверкой, что

$$C_i^{-1} = \begin{pmatrix} 1 & & & & \\ & 1 & & & 0 \\ & & \ddots & & \\ & & & a_{ii} & \\ 0 & & & & \ddots & \\ & & & & & 1 \end{pmatrix}.$$

Рассмотрим матрицу

$$L = C_1^{-1} L_1^{-1} C_2^{-1} L_2^{-1} \dots C_n^{-1} L_n^{-1},$$

которая как произведение диагональных и нижнетреугольных матриц является нижнетреугольной матрицей.

Умножим теперь обе части равенства

$$A_n = L_n C_n \dots L_1 C_1 A$$

на матрицу L слева, получим

$$L A_n = (C_1^{-1} L_1^{-1} \dots C_n^{-1} L_n^{-1}) (L_n C_n \dots L_1 C_1 A) = A,$$

откуда с учетом обозначения $A_n = R$ получаем разложение $A = L \cdot R$, которое называется треугольным разложением матрицы A .

При наличии треугольного разложения матрицы A решение системы уравнений $Ax = b$ сводится к решению двух треугольных систем уравнений

$$Ly = b$$

и

$$Rx = y.$$

Вернемся теперь к обоснованию метода Гаусса. Поскольку на каждом шаге этого метода обе части матричного уравнения умножаются слева на одну и ту же невырожденную матрицу, то множество решений этого уравнения при таком преобразовании не меняется, так как при умножении на обратную матрицу возвращаемся к исходному матричному уравнению.

Вопросы для самоконтроля

1. Сформулируйте критерий совместности системы линейных уравнений (СЛУ).
2. Когда СЛУ из n уравнений с n неизвестными имеет единственное решение?
3. Может ли СЛУ из $n - 1$ уравнений с n неизвестными иметь единственное решение?
4. Может ли СЛУ из $n + 1$ уравнений с n неизвестными иметь единственное решение?
5. Пусть СЛУ имеет n уравнений с n неизвестными, и при попытке применить формулы Крамера для решения этой СЛУ выяснилось, что определитель этой системы $D = 0$. Означает ли это, что данная СЛУ несовместна?
6. Может ли однородная СЛУ из $n - 1$ уравнений с n неизвестными иметь единственное нулевое решение?
7. Когда однородная СЛУ из n уравнений с n неизвестными имеет ненулевое решение?
8. Сформулируйте теорему о фундаментальной системе решений однородной СЛУ. Какое отношение имеет размерность подпространства решений однородной СЛУ к ее фундаментальной системе решений?
9. Сформулируйте связь между решениями произвольной СЛУ и ассоциированной с ней однородной СЛУ.
10. Образует ли подпространство множество решений произвольной СЛУ?
11. Можно ли при реализации метода Гаусса решения СЛУ обойтись без операций деления?
12. Верно ли, что если при реализации метода Гаусса для СЛУ из n уравнений с n неизвестными ступенчатый вид СЛУ принимает треугольную форму, то СЛУ имеет единственное решение?
13. В каком случае при реализации метода Гаусса СЛУ несовместна?
14. Верно ли, что две системы двух линейных уравнений с двумя неизвестными равносильны, если обе системы имеют бесконечное множество решений?

15. Верно ли, что две системы двух линейных уравнений с двумя неизвестными равносильны, если графики соответствующих уравнений обеих систем попарно параллельны?
16. Верно ли, что две системы двух линейных уравнений с двумя неизвестными равносильны, если все коэффициенты (в том числе и свободные члены) соответствующих уравнений пропорциональны?
17. Верно ли, что две СЛУ из n уравнений с n неизвестными равносильны, если обе СЛУ являются определенными?
18. Верно ли, что две СЛУ из n уравнений с n неизвестными равносильны, если обе СЛУ несовместны?
19. Сформулируйте, что такое треугольное разложение невырожденной матрицы.
20. Можно ли получить аналог треугольного разложения для вырожденной матрицы?

МОДУЛЬ IV

Лекция 1. АЛГЕБРАИЧЕСКИЕ ОПЕРАЦИИ И ГРУППЫ

Определение IV.1.1. Будем говорить, что на множестве X задана алгебраическая операция $*$, если любой упорядоченной паре $x, y \in X$ поставлен в соответствие однозначно определенный элемент $z \in X$, который обозначается $z = x * y$. Множество X с заданной на нем алгебраической операцией $*$ будем называть алгебраической структурой и обозначать $(X; *)$.

Пример IV.1.2. Множество $\mathbb{Z}$ целых чисел образует алгебраические структуры $(\mathbb{Z}; +)$ и $(\mathbb{Z}; \cdot)$ относительно алгебраических операций сложения и умножения.

Пример IV.1.3. На множестве $\mathbb{N}$ натуральных чисел рассмотрим следующее соответствие Δ из $\mathbb{N} \times \mathbb{N}$ в $\mathbb{N}$:

$$(x, y) \mapsto x \Delta y = x : y.$$

Данное соответствие не является алгебраической операцией на множестве $\mathbb{N}$, так как, например, $2 \Delta 3 = 2/3 \notin \mathbb{N}$.

Пример IV.1.4. Множество ненулевых рациональных чисел $\mathbb{Q}^*$ образует алгебраические структуры $(\mathbb{Q}^*; \cdot)$ и $(\mathbb{Q}^*; :)$ относительно операций умножения и деления.

Пример IV.1.5. Алгебраическую операцию $*$ на множестве $X = \{0, 1, 2\}$ зададим таблицей:

$*$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

Как видно из определения алгебраической операции, характеристическими ее признаками являются:

- а) всюдуопределенность;
- б) однозначность;
- в) замкнутость.

Рассмотрим некоторые свойства алгебраических операций.

Определение IV.1.6. Будем говорить, что алгебраическая операция $*$, заданная на множестве X , имеет свойство ассоциативности (ассоциативная операция), если для любых $x, y, z \in X$ выполняется следующее:

$$x * (y * z) = (x * y) * z.$$

Алгебраическая структура $(X; *)$ с ассоциативной алгебраической операцией $*$ называется **полугруппой**.

Пример IV.1.7. Алгебраические структуры $(\mathbb{Z}; +)$, $(\mathbb{Z}; \cdot)$ и $(\mathbb{Q}^*; \cdot)$ являются полугруппами. Алгебраическая структура $(\mathbb{Q}^*; :)$ полугруппой не является, так как, например, $1:(2:3) \neq (1:2):3$. Можно проверить, что алгебраическая структура из примера IV.1.5 является полугруппой.

Определение IV.1.8. Пусть $(X; *)$ – алгебраическая структура. Элемент $e \in X$ называется **нейтральным элементом**, если для любого $x \in X$ имеет место

$$x * e = e * x = x.$$

Заметим, что нейтральный элемент единственный. Действительно, если $e' \in X$ – другой нейтральный элемент, то из $e' * e = e$ и $e' * e = e'$ получаем, что $e = e'$.

Пример IV.1.9. Алгебраическая структура $(\mathbb{Z}; +)$ имеет нейтральный элемент 0, $(\mathbb{Z}; \cdot)$ и $(\mathbb{Q}^*; \cdot)$ – нейтральный элемент 1. Алгебраическая структура $(\mathbb{Q}^*; :)$ не имеет нейтрального элемента. Действительно, если в $(\mathbb{Q}^*; :)$ есть нейтральный элемент $e \in \mathbb{Q}^*$, то, например, из $2 : e = 2$ следует $e = 1$. В то же время $e = 1$ не является нейтральным элементом в $(\mathbb{Q}^*; :)$, так как, например, $1 : 2 \neq 2$.

В алгебраической структуре из примера IV.1.5 нейтральным элементом является 0.

Определение IV.1.10. Алгебраическая структура $(X; *)$ называется **моноидом**, если:

- а) $(X; *)$ есть полугруппа;
- б) существует нейтральный элемент $e \in X$ алгебраической структуры $(X; *)$.

Пример IV.1.11. Алгебраические структуры $(\mathbb{Z}; +)$, $(\mathbb{Z}; \cdot)$ и $(\mathbb{Q}^*; \cdot)$ являются моноидами.

Приведем пример полугруппы, которая не является моноидом.

Пример IV.1.12. Пусть $(\mathbb{N}; *)$ – алгебраическая структура, где $a * b = \min(a, b)$ для любых $a, b \in \mathbb{N}$. Ясно, что $(\mathbb{N}; *)$ есть полугруппа, так как

$$\min(\min(a, b), c) = \min(a, \min(b, c)).$$

Пусть $e \in \mathbb{N}$ – нейтральный элемент в $(\mathbb{N}; *)$, тогда для любого $e \in \mathbb{N}$ имеем $\min(x, e) = x$, т.е. $x \leq e$. Это означает, что любое натуральное число не превосходит числа $e \in \mathbb{N}$, что очевидно невозможно, так как $e + 1 > e$. Следовательно, $(\mathbb{N}; *)$ не имеет нейтрального элемента.

Определение IV.1.13. Алгебраическая структура $(X; *)$ называется квазигруппой, если для любых $a, b \in X$ однозначно разрешимы уравнения $a * x = b$ и $y * a = b$.

Пример IV.1.14. Алгебраическая структура из примера IV.1.5 является квазигруппой, поскольку уравнения вида $a * x = b$ и $y * a = b$ для любых $a, b \in X$ однозначно разрешимы согласно таблице, задающей операцию $*$.

Приведем пример квазигруппы, которая не является полугруппой.

Пример IV.1.15. Пусть $X = \{0, 1, 2\}$, а алгебраическая операция $*$ задается таблицей:

$*$	0	1	2
0	1	0	2
1	2	1	0
2	0	2	1

$(X; *)$ является квазигруппой, поскольку уравнения вида $a * x = b$ и $y * a = b$ для любых $a, b \in X$ однозначно разрешимы согласно таблице, задающей операцию $*$. В то же время $(X; *)$ не является полугруппой, так как имеют место равенства

$$(0 * 1) * 2 = 0 * 2 = 2,$$

$$0 * (1 * 2) = 0 * 0 = 1.$$

Определение IV.1.16. Пусть $(X; *)$ есть алгебраическая структура с нейтральным элементом $e \in X$. Будем говорить, что элемент $y \in X$ является **обратным элементом** для элемента $x \in X$, если $x * y = y * x = e$. Будем обозначать обратный элемент для элемента $x \in X$ через x^{-1} , а элемент, для которого существует обратный, будем называть **обратимым**.

Пример IV.1.17. В $(\mathbb{Z}; +)$ любой элемент обратим (нейтральный элемент – 0, обратный элемент для целого числа n есть $-n$). В $(\mathbb{Z}; \cdot)$ обратимыми являются только 1 и -1 (нейтральный элемент есть 1, обратный для -1 равен -1).

Определение IV.1.18. Алгебраическая операция $*$ на множестве X называется коммутативной, если для любых элементов $a, b \in X$ справедливо равенство

$$a * b = b * a.$$

Пример IV.1.19. Алгебраические операции сложения и умножения на множествах $\mathbb{Z}$, $\mathbb{Q}$ и $\mathbb{R}$ являются коммутативными.

Пример IV.1.20. Следующая алгебраическая операция на множестве $\mathbb{N}$ не является коммутативной:

$$n * m = m, \quad n, m \in \mathbb{N}.$$

Каждая из алгебраических структур, которые рассматривались выше, а именно полугруппы, моноиды и квазигруппы, удовлетворяют некоторым из небольшого фиксированного набора свойств, т.е. ассоциативность, существование нейтрального элемента и однозначная разрешимость уравнений вида $ax = b$ и $ya = b$. Сейчас будет рассмотрена фундаментальная алгебраическая структура, которая нашла многочисленные применения не только в математике, но и в физике, химии, биологии и других науках.

Определение IV.1.21. Алгебраическая структура $(G; *)$ называется **группой**, если

- а) алгебраическая операция $*$ ассоциативна;
- б) для любых $a, b \in G$ однозначно разрешимы в G уравнения $a * x = b$ и $y * a = b$.

Другими словами, класс групп определяется как пересечение классов полугрупп и квазигрупп (см. рис. 1).



Рис. 1

Условимся, если $(G; *)$ – группа, для краткости будем говорить о группе G , а групповую операцию называть умножением, т.е. принимать мультипликативную форму записи.

Теорема IV.1.22. Алгебраическая структура $(G; \cdot)$ является группой тогда и только тогда, когда выполняются следующие условия:

1) для любых $a, b, c \in G$

$$a(bc) = (ab)c;$$

2) существует элемент $1 \in G$ такой, что для любого элемента $a \in G$

$$a \cdot 1 = 1 \cdot a = a;$$

3) для любого элемента $a \in G$ существует элемент $a^{-1} \in G$ такой, что

$$aa^{-1} = a^{-1}a = 1.$$

Определение IV.1.23. Группа G называется коммутативной или абелевой, если для любых $a, b \in G$ выполняется условие коммутативности

$$ab = ba.$$

Для абелевых групп принято групповую операцию записывать как операцию сложения, тогда нейтральный элемент группы обозначается как 0, обратный элемент для элемента a обозначается как $-a$, т.е. для абелевых групп принята аддитивная форма записи. Приведем теперь некоторые примеры групп.

Пример IV.1.24. $(\mathbb{Z}; +)$, $(\mathbb{Q}; +)$, $(\mathbb{R}; +)$, $(\mathbb{Q}^*; \cdot)$ и $(\mathbb{R}^*; \cdot)$ – абелевы группы. $(GL_2(\mathbb{R}); \cdot)$ – (неабелева) группа невырожденных действительных матриц второго порядка по умножению, $C_2(\mathbb{R})$ – абелева группа

скалярных действительных матриц второго порядка по умножению (т.е. матриц вида $\begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}$, где $\alpha \in \mathbb{R}$).

Определение IV.1.25. Подмножество H группы G называется подгруппой, если H само является группой относительно той же операции, что и в G . Условимся обозначать подгруппу так: $H \leq G$.

Пример IV.1.26. $(\mathbb{Z}; +) < (\mathbb{Q}; +) < (\mathbb{R}; +)$.

Теорема IV.1.27. Подмножество H группы G есть подгруппа в G тогда и только тогда, когда выполняется условие

$$(*) \text{ для любых } x, y \in H \text{ выполнено } xy^{-1} \in H.$$

Доказательство. Необходимость очевидна.

Достаточность. Пусть выполнено условие (*), покажем, что $H < G$. Для этого сначала проверим выполнение условий 1 – 3 теоремы IV.1.22.

Начнем с условия 2. Рассмотрим $x, y \in H$, где $x = y$, тогда $xy^{-1} = xx^{-1} = 1_G \in H$ по условию (*). Тогда 1_G , очевидно, является нейтральным элементом и в H .

Проверим условие 3. Возьмем пару $x, y \in H$, где $x = 1_G$, y – любой элемент H . Тогда

$$xy^{-1} = 1_G \cdot y^{-1} = y^{-1} \in H$$

по условию (*). Условие 1, очевидно, выполняется. Осталось убедиться, что алгебраическая операция группы G является алгебраической операцией и в H . Пусть $x, y \in H$, нужно проверить, что $x, y \in H$. По доказанному ранее $y^{-1} \in H$, тогда рассмотрим пару $x, y^{-1} \in H$. По условию (*) имеем

$$x(y^{-1}) \in H.$$

Но $(y^{-1}) = y$, следовательно, $x, y \in H$.

Теорема доказана.

Замечание. В аддитивной записи критерий подгруппы принимает следующий вид: $H \leq G$ тогда и только тогда, когда для любых $x, y \in H$ выполняется условие

$$(**) \quad x - y \in H.$$

Определение IV.1.28. Пусть G – группа, $g \in G$ – произвольный элемент. Для любого целого числа n определим n -ю степень элемента g следующим образом:

$$g^n = \begin{cases} \underbrace{g \cdot g \cdots g}_n, & \text{если } n > 0, \\ \underbrace{g^{-1} \cdot g^{-1} \cdots g^{-1}}_{|n|}, & \text{если } n < 0, \\ 1 & , \text{ если } n = 0. \end{cases}$$

Теорема IV.1.29. Пусть G – группа, $g \in G$ – произвольный элемент. Тогда $g^m \cdot g^n = g^{m+n}$ для любых целых чисел m и n .

Теорема IV.1.30. Пусть G – группа, $g \in G$ – произвольный элемент. Тогда $(g^m)^n = g^{mn}$ для любых целых чисел m и n .

Пусть G – группа, $g \in G$ – произвольный элемент. Обозначим через $\langle g \rangle$ множество всевозможных целочисленных степеней элемента g :

$$\langle g \rangle = \{ \dots, g^{-2}, g^{-1}, g^0 = 1, g, g^2, \dots \}.$$

По теореме IV.1.27 множество $\langle g \rangle$ является подгруппой группы G . Действительно, для любых элементов $x = g^m$ и $y = g^n$, m, n – целые числа, множества $\langle g \rangle$ имеет место по теоремам IV.1.29 и IV.1.30:

$$x \cdot y^{-1} = g^m \cdot (g^n)^{-1} = g^m \cdot g^{-n} = g^{m-n} \in \langle g \rangle.$$

Будем говорить, что $\langle g \rangle$ есть подгруппа группы G , порожденная элементом g , а сам элемент g будем называть образующим подгруппы $\langle g \rangle$.

Определение IV.1.31. Группа G называется циклической группой, если $G = \langle g \rangle$ для некоторого $g \in G$, который называется образующим циклической группы G .

Если все целочисленные степени элемента g в циклической группе $\langle g \rangle$ различны, то в этом случае говорят, что $\langle g \rangle$ есть бесконечная циклическая группа, а элемент g имеет бесконечный порядок, который обозначается так: $O(g) = \infty$.

Пусть некоторые целочисленные степени элемента g с различными показателями степени совпадают, т.е. $g^n = g^m$ для некоторых целых чисел $n \neq m$. Для определенности можем считать, что $n < m$, тогда, умножая обе части этого равенства на g^{-n} , получаем $g^{m-n} = 1$. Пусть k есть наименьшее натуральное число такое, что $g^k = 1$, тогда в этом случае говорят, что порядок элемента g равен k (обозначение: $O(g) = k$), а группа $\langle g \rangle$ называется циклической группой порядка k .

Определение IV.1.32. Пусть G есть произвольная группа. Назовем порядком группы G мощность $|G|$ группы G .

Имеет место следующая

Теорема IV.1.33. Пусть $G = \langle g \rangle$ – циклическая группа с образующим g . Тогда

$$|G| = O(g).$$

Доказательство. Рассмотрим два случая.

Случай 1. G – бесконечная циклическая группа. В этом случае $|G| = \infty$ и $O(g) = \infty$, следовательно,

$$|G| = O(g).$$

Случай 2. G – циклическая группа порядка n . В этом случае группа $G = \langle g \rangle$ имеет следующий вид:

$$\langle g \rangle = \{1, g, g^2, \dots, g^{n-1}\}.$$

Действительно, с одной стороны, так как $O(g) = g^n = 1$, имеем

$$g^{n+1} = g, \dots, g^{2n-1} = g^{n-1}, g^{2n} = 1, \dots$$

С другой стороны,

$$g^{-1} = 1 \cdot g^{-1} = g^n \cdot g^{-1} = g^{n-1},$$

$$g^{-2} = 1 \cdot g^{-2} = g^n \cdot g^{-2} = g^{n-2}, \dots, g^{-k} = g^{n-k}$$

для $k = 1, 2, \dots, n-1$. Далее имеем

$$g^{-n} = (g^n)^{-1} = 1^{-1} = 1,$$

$$g^{-n-1} = g^{-n} \cdot g^{-1} = g^{-1} = g^{n-1}, \dots, g^{-n-k} = g^{-k} = g^{n-k}$$

для любого $k = 1, 2, \dots, n-1$. Продолжая таким образом, получаем, что любой элемент из $\langle g \rangle$ является одним из элементов $1, g, g^2, \dots, g^{n-1}$. Но тогда $|G| = n = O(g)$, что и требовалось доказать.

Замечание. Для аддитивной записи групп целочисленные степени образующего заменяются на его целочисленные кратные, тогда бесконечная циклическая группа записывается так:

$$\langle g \rangle = \{ \dots, -2g, -g, 0, g, 2g, \dots \},$$

а конечная циклическая группа порядка n имеет вид

$$\langle g \rangle = \{ 0, g, 2g, \dots, (n-1)g \}.$$

Пример IV.1.34. Пусть G есть мультипликативная группа целочисленных матриц второго порядка вида

$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}, \text{ где } k \in \mathbb{Z}.$$

Проверка того, что G – группа:

а) умножение матриц – ассоциативно;

б) нейтральный элемент – $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$;

в) для матрицы $\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ обратным элементом является $\begin{pmatrix} 1 & -k \\ 0 & 1 \end{pmatrix}$.

Обозначим $g_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, тогда $G = \langle g_0 \rangle$.

Действительно, $g_0^2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$,

$$g_0^3 = g_0^2 \cdot g_0 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix},$$

$$g_0^\ell = \begin{pmatrix} 1 & \ell \\ 0 & 1 \end{pmatrix}, \ell = 1, 2, \dots$$

Это утверждение обосновывается с использованием принципа математической индукции. Для $\ell = 1$ утверждение очевидно.

Пусть для произвольного натурального числа ℓ имеем

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

Найдем $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{\ell+1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{\ell} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \ell \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \ell+1 \\ 0 & 1 \end{pmatrix}.$

Следовательно, согласно принципу математической индукции, для любого натурального числа ℓ имеем

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

Аналогично для отрицательных показателей степени: для любого $s < 0$

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

Заметим, что $s = -|s|$, и тогда

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^s = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-|s|} = \left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} \right)^{|s|} = \left(\begin{array}{c|c} 1 & -1 \\ \hline 0 & 1 \end{array} \right)^{|s|} = \left(\begin{array}{c|c} 1 & -|s| \\ \hline 0 & 1 \end{array} \right).$$

Таким образом, G есть бесконечная циклическая группа.

Пример IV.1.35. Аддитивная группа $(\mathbb{Z}; +)$ целых чисел является бесконечной циклической группой с образующим 1. Действительно, любое целое число n является целочисленным кратным 1, так как

$$n = n \cdot 1.$$

Приведем теперь примеры конечных циклических групп.

Пример IV.1.36. Пусть $G = \{1, -1\}$, тогда относительно умножения чисел G является циклической группой второго порядка, так как $(-1)^2 = 1$. Образующий группы G — это число -1 , имеет порядок 2. Порядок группы G (т.е. число ее элементов) тоже равен 2.

Пример IV.1.37. Рассмотрим циклическую группу G , состоящую из целочисленных степеней матрицы

$$g = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

которая является подгруппой мультипликативной группы действительных невырожденных матриц второго порядка. Заметим, что $O(g) = 4$. Действительно,

$$g^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$g^3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^3 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

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

Таким образом, $\langle g \rangle$ есть циклическая группа порядка 4.

Определение IV.1.38. Пусть G – группа, $g \in G$ – произвольный элемент. Назовем подгруппу $\langle g \rangle$ группы G циклической подгруппой, порожденной элементом g . Заметим, что по теореме IV.1.33 порядок циклической подгруппы, порожденной элементом g (т.е. число ее элементов), равен порядку элемента g .

Определение IV.1.39. Пусть G – группа, H – подгруппа группы G . Для произвольного $g \in G$ рассмотрим множество

$$gH = \{x \in G \mid x = gh \text{ для некоторого } h \in H\},$$

которое называется правым смежным классом относительно подгруппы H с представителем $g \in G$. Аналогично определяется левый смежный класс Hg , $g \in G$.

Для различных приложений имеет важное значение следующая

Теорема IV.1.40 (теорема Лагранжа). Пусть G есть конечная группа, $|G| = n$. Если H есть подгруппа группы G и $|H| = k$, то для некоторого натурального числа s имеем

$$n = k \cdot s.$$

Доказательство. Введем на группе G бинарное отношение ρ следующим образом: для любых $x, y \in G$ пусть $x\rho y$ тогда и только тогда, когда $x^{-1}y \in H$.

Покажем, что ρ есть отношение эквивалентности. Действительно, имеем

а) $x\rho x$, так как $x^{-1}x = 1 \in H$;

б) $x\rho y$ влечет $u\rho x$, так как из $x^{-1}y \in H$ следует $(x^{-1}y)^{-1} = y^{-1}x \in H$;

в) из $x\rho y$ и $u\rho z$ следует $x\rho z$, так как из $x^{-1}y \in H$ и $x^{-1}z \in H$ следует $(x^{-1}y) \cdot (y^{-1}z) = x^{-1}z \in H$.

По теореме I.1.45 классы эквивалентности относительно отношения эквивалентности ρ составляют разбиение группы G , скажем, $G = \overline{x_1} \cup \overline{x_2} \cup \dots \cup \overline{x_s}$. Рассмотрим произвольный класс эквивалентности $\overline{x_i}$, $i = 1, \dots, s$. Покажем, что $\overline{x_i} = x_i H$ для любого $i = 1, \dots, s$.

Пусть $g \in \overline{x_i}$, тогда $g\rho x_i$, откуда $g^{-1}x_i = h \in H$. Следовательно, $x_i = gh$, и тогда $g = x_i h^{-1} \in x_i H$. Пусть $g \in x_i H$, тогда $g = x_i h$, где $h \in H$, откуда получаем $g^{-1}x_i = (h^{-1}x_i^{-1}) \cdot x_i = h^{-1}(x_i^{-1}x_i) = h^{-1} \in H$, т.е. $g\rho x_i$ и, следовательно, $g \in \overline{x_i}$.

Тогда $G = x_1 H \cup \dots \cup x_s H$ – разбиение группы G . Поскольку $|H| = k$, то $H = \{h_1, \dots, h_k\}$, и тогда $x_i H = \{x_i h_1, \dots, x_i h_k\}$ для любого $i = 1, \dots, s$. Заметим, что $|x_i H| = k$ для любого $i = 1, \dots, s$. Действительно, если $x_i h_r = x_i h_t$, то тогда $x_i^{-1}(x_i h_r) = x_i^{-1}(x_i h_t)$ влечет $h_r = h_t$, т.е. для k различных элементов $h_1, \dots, h_k$ элементы $x_i h_1, \dots, x_i h_k$ тоже будут различными. Таким образом, в каждом классе эквивалентности по k элементов, всего этих классов s , следовательно, из определения разбиения получаем $n = |G| = k \cdot s$. Теорема доказана.

Следствие. Если G – конечная группа и $|G| = n$, то для любого элемента $g \in G$ имеет место $g^n = 1$.

Доказательство. Обозначим $H = \langle g \rangle$. Так как H – подгруппа группы G , то, если $|H| = k$, то по теореме Лагранжа $n = k \cdot s$ для некоторого натурального числа s . Поскольку по теореме IV.1.33 имеем

$$|H| = O(g),$$

то $O(g) = k$ и, следовательно, $g^k = 1$. Тогда имеем

$$g^n = g^{ks} = (g^k)^s = 1^s = 1.$$

Теорема IV.1.41. Пусть $G = \langle x \rangle$ есть циклическая группа порядка n . Элемент $g = x^k \in G$ является образующим группы G тогда и только тогда, когда выполняется условие

$$HOD(k, n) = 1.$$

Доказательство. Необходимость. Пусть g есть образующий группы G , тогда $G = \langle g \rangle$. Предположим, что $HOD(k, n) = d > 1$. Учтем, что $O(x) = n$, тогда имеем

$$g^{n/d} = (x^k)^{n/d} = x^{kn/d} = x^{n \cdot k/d} = (x^n)^{k/d} = 1^{k/d} = 1,$$

откуда следует, что $O(g) \leq n/d < n$, что противоречит тому, что $G = \langle g \rangle$.

Достаточность. Пусть для элемента $g = x^k \in G$ имеем $HOD(k, n) = 1$. Покажем, что тогда g есть образующий группы G , т.е. $\langle g \rangle = G$. Ясно, что $\langle g \rangle \subseteq G$. Проверим, что $G \subseteq \langle g \rangle$. Достаточно убедиться, что образующий x группы такой, что $x \in \langle g \rangle$, поскольку тогда из того, что любой элемент $y \in G$ имеет вид x^t , следует $y \in \langle g \rangle$.

По одному из свойств наибольшего общего делителя двух натуральных чисел, имеем, что если $d = HOD(k, n)$, то существуют целые числа s и t такие, что $sk + tn = d$. Тогда

$$g^s = g^s \cdot 1 = g^s \cdot (x^n)^t = x^{ks} \cdot x^{nt} = x^{ks+nt} = x^d = x^1 = x,$$

так как $d = 1$ по условию. Следовательно,

$$x = g^s \in \langle g \rangle,$$

что и требовалось доказать.

Теорема IV.1.42. Любая подгруппа конечной циклической группы G является циклической группой.

Доказательство. Пусть $G = \langle x \rangle$ и $|G| = n$, тогда $O(x) = n$. Предположим, что H есть произвольная собственная подгруппа группы G (т.е. $H \neq G$ и $H \neq \{1\}$) и $h \in H$ — произвольный элемент. Так как $h \in G$, то $h = x^k$ для некоторого целого числа k , где $0 \leq k < n$. Если $k = 0$, то $h = 1$, если $k = 1$, то $h = x$ и $\langle x \rangle = G \subseteq H$, т.е. $G = H$. Следовательно, можем считать, что $1 < k < n$.

Пусть L — множество натуральных чисел, которые являются показателями некоторой степени с основанием x для элемента $h \in H$ такого, что $h \neq 1$. Заметим, что L — непустое множество, поскольку из $H \neq \{1\}$ и $H \neq G$ следует, что существует $h = x^k \in H$ такой, что $1 < k < n$, т.е. $k \in L$. Далее ясно, что для любого $\ell \in L$ имеем $1 < \ell < n$. Поскольку любое непустое множество натуральных чисел имеет наименьший элемент, то в L есть наименьший элемент ℓ_0 . Обозначим $h_0 = x^{\ell_0} \in H$ и рассмотрим $\langle h_0 \rangle \subseteq H$. Покажем, что $H = \langle h_0 \rangle$.

Достаточно убедиться в том, что любой $h \in H$ также $h \in \langle h_0 \rangle$. Пусть $h \in H$ — произвольный элемент такой, что $h \neq 1$, тогда $h = x^k$, $1 < k < n$. Так как $k \in L$ и ℓ_0 — наименьший в L , то делим k на ℓ_0 с остатком и получаем

$$k = q \cdot \ell_0 + r, \quad 0 \leq r < \ell_0.$$

Имеем

$$\begin{aligned} x^r &= x^{k - q\ell_0} = x^k \cdot (x^{\ell_0})^{-q} = x^k \cdot \left((x^{\ell_0})^{-1} \right)^q = \\ &= x^k \cdot (h_0^{-1})^q = h \cdot (h_0^{-1})^q \in H, \end{aligned}$$

поскольку $h \in H$ и $h_0^{-1} \in H$, а, следовательно, $(h_0^{-1})^q \in H$.

Предположим, что $r > 0$ и $x^r = 1$. Тогда $O(x) \leq r < n$, что противоречит условию $O(x) = n$.

Предположим, что $r > 0$ и $x^r \neq 1$. Тогда $r \in L$, и так как $r < \ell_0$, то получили противоречие с тем, что ℓ_0 — наименьший элемент L . Итак, если $r > 0$, то получаем противоречие, следовательно, $r = 0$. Тогда имеем

$$k = q \cdot \ell_0,$$

и тогда $h = x^k = x^{\ell_0 q} = (x^{\ell_0})^q = h_0^q \in \langle h_0 \rangle$.

Теорема доказана.

Следующее утверждение показывает, что для циклических групп обратная теорема к теореме Лагранжа имеет место.

Теорема IV.1.43. Пусть $G = \langle x \rangle$ – циклическая группа порядка n . Если d есть делитель n , то существует и единственная подгруппа H группы G такая, что $|H| = d$.

Доказательство. Рассмотрим элемент $y = x^{n/d}$ и подгруппу $H = \langle y \rangle$. Тогда

$$y^d = \left(x^{n/d} \right)^d = x^n = 1.$$

Так как $O(x) = n$, то $O(y) = d$, следовательно, $|H| = d$.

Покажем единственность подгруппы порядка d . Пусть H' есть другая подгруппа группы G порядка d . Поскольку по теореме IV.1.42 имеем, что H' есть циклическая группа, то $H' = \langle z \rangle$ для некоторого элемента $z = x^k \in G$. Так как $O(z) = d$, то d есть наименьший показатель степени элемента z , такой, что $z^d = 1$. Имеем

$$1 = z^d = (x^k)^d = x^{kd}.$$

Поскольку $O(x) = n$, то n есть наименьший показатель степени элемента x , такой, что $x^n = 1$. Следовательно, $n \leq kd$. Заметим, что kd нацело делится на n . Действительно, если $kd = nq + r$, где $0 \leq r < n$, то из $x^{kd} = 1$ получаем

$$1 = x^{kd} = x^{nq+r} = x^{nq} \cdot x^r = (x^n)^q \cdot x^r = 1 \cdot x^r = x^r,$$

что противоречит $O(x) = n$, поскольку $r < n$. Таким образом, $kd = nt$ для некоторого натурального числа t . Отсюда

$$k = \frac{nt}{d},$$

и так как d есть делитель n , то можем записать число k в виде

$$k = \frac{nt}{d} = \frac{n}{d} \cdot t.$$

Следовательно, имеем

$$z = x^k = \left(x^{n/d}\right)^t = y' \in H.$$

Поскольку $O(z) = d$ и $z \in H$, то $\langle z \rangle = H$. Учитывая $H' = \langle z \rangle$, получаем $H = H'$, что и требовалось доказать.

Рассмотрим следующую задачу.

Пусть в группе G элемент x имеет порядок n , а элемент y имеет порядок m . Вычислить порядок элемента xy .

На первый взгляд можно предположить, что порядок xy как-то связан с порядками x и y . Однако следующий пример обескураживает.

Пример IV.1.44. Пусть G – мультипликативная группа невырожденных целочисленных матриц второго порядка, $x, y \in G$ такие, что

$$x = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}.$$

Непосредственными вычислениями убеждаемся, что $O(x) = 4$ и $O(y) = 3$. Рассмотрим элемент xy :

$$xy = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 0 & -1 \end{pmatrix}.$$

Найдем порядок элемента xy , для чего вычисляем его степени для любого натурального числа n :

$$\begin{pmatrix} -1 & -1 \\ 0 & -1 \end{pmatrix}^n = (-1)^n \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n = (-1)^n \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}.$$

Поскольку все степени xy различны для разных натуральных чисел n , то $O(xy) = \infty$.

В то же время для некоторых частных случаев порядок произведения двух элементов зависит от порядков сомножителей. В частности, это имеет место в следующем случае.

Теорема IV.1.45. Пусть G – произвольная группа, а элементы $x, y \in G$ имеют взаимно простые порядки $O(x) = s$, $O(y) = t$, т.е.

$HOD(s, t) = 1$. Если элементы x, y перестановочны (т.е. $xy = yx$), то $O(xy) = st$ и $\langle xy \rangle$ есть прямое произведение циклических подгрупп $\langle x \rangle$ и $\langle y \rangle$, т.е.

$$\langle xy \rangle = \langle x \rangle \times \langle y \rangle.$$

Доказательство. Для проверки представимости $\langle xy \rangle$ в виде прямого произведения $\langle x \rangle$ и $\langle y \rangle$ достаточно убедиться в выполнении следующих условий:

а) любой элемент $g \in \langle xy \rangle$ можно записать в виде $g = g_1 g_2$, где

$$g_1 \in \langle x \rangle, \quad g_2 \in \langle y \rangle;$$

б) $\langle x \rangle \cap \langle y \rangle = \{1\}$;

в) любые элементы $u \in \langle x \rangle, v \in \langle y \rangle$ перестановочны.

Условие «в» вытекает из перестановочности элементов x и y .

Условие «а» выполняется, так как для любого $k \in \mathbb{N}$

$$\begin{aligned} (xy)^k &= \underbrace{(xy)(xy)\dots(xy)}_{k \text{ раз}} = x(yx)(yx)\dots(xy) = \\ &= x(xy)(yx)\dots(xy) = x^2 y^2 (xy)\dots(xy) = \dots = x^k y^k, \end{aligned}$$

где использованы ассоциативность умножения в группе и перестановочность элементов x и y . Осталось убедиться в справедливости условия «б».

Пусть $z \in \langle x \rangle \cap \langle y \rangle$ – произвольный элемент и $O(z) = \ell$, тогда из $z = x^i = y^j$ следует

$$z^s = (x^i)^s = x^{is} = x^{si} = (x^s)^i = 1,$$

$$z^t = (y^j)^t = y^{jt} = y^{tj} = (y^t)^j = 1.$$

Но тогда ℓ , как порядок элемента z , удовлетворяет условиям $\ell | s$ и $\ell | t$, т.е. ℓ есть общий делитель s и t . Поскольку $HOD(s, t) = 1$, постольку $\ell = 1$, откуда $z^1 = 1$, влечет $z = 1$. Это означает, что $\langle x \rangle \cap \langle y \rangle = \{1\}$.

То, что $O(xy) = st$, следует из

$$O(xy) = |\langle xy \rangle| = |\langle x \rangle| \cdot |\langle y \rangle| = O(x) \cdot O(y) = st.$$

Теорема доказана.

В предыдущей теореме прямое произведение циклических групп взаимно простых порядков оказывается циклической группой. В общем случае прямое произведение циклических групп может не быть циклической группой.

Пример IV.1.46. Пусть $G = \langle x \rangle \times \langle y \rangle$ – прямое произведение циклических групп $\langle x \rangle$ и $\langle y \rangle$, где $O(x) = O(y) = 2$. Предположим, что $G = \langle z \rangle$, тогда $O(z) = |G| = 4$. Но в группе G нет элемента порядка 4. Действительно,

$$G = \{1, x, y, xy\} \text{ и } O(x) = O(y) = O(xy) = 2,$$

так как $(xy)(xy) = x^2 y^2 = 1$.

Следующая теорема показывает, что класс групп, разложимых в прямое произведение циклических групп, достаточно велик.

Теорема IV.1.47. Любая конечная абелева группа является прямым произведением циклических групп.

Доказательство этой теоремы превышает уровень первоначального знакомства с теорией групп и поэтому опускается. В то же время утверждение этой теоремы имеет большое значение при изучении строения групп обратимых элементов колец классов вычетов, которые используются в различных криптографических приложениях.

Лекция 2. РАЗЛИЧНЫЕ КЛАССЫ ГРУПП

Определение IV.2.1. Пусть $N_k = \{1, 2, \dots, k\}$ – отрезок натурального ряда чисел. Назовем подстановкой на множестве N_k произвольную биекцию на N_k .

Условимся записывать подстановку на N_k в виде таблицы из двух строк, где верхняя строка – упорядоченный набор чисел $1, 2, \dots, k$, а нижняя строка представляет собой некоторую перестановку элементов верхней строки, такую, что каждый элемент нижней строки является образом расположенного над ним элемента верхней строки. Например, подстановка

$$\phi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 5 & 4 & 2 \end{pmatrix}$$

является биекцией на N_5 , которая отображает 1 в 3, 2 в 1, 3 в 5, 4 в 4 и 5 в 2. Зададим алгебраическую операцию умножения на множестве подстановок на N_k как композицию биекций на N_k . Поскольку композиция отображений ассоциативна, то данная алгебраическая операция ассоциативна. Очевидно, что тождественная подстановка (т.е. подстановка на N_k , которая оставляет на месте каждый элемент N_k) является нейтральным элементом для умножения подстановок. Наконец, для каждой подстановки на N_k существует обратная подстановка относительно умножения подстановок, поскольку для каждой биекции существует обратная биекция. Очевидно, что обратная подстановка получается из данной подстановки перестановкой верхней и нижней ее строк с последующей (если необходимо) перестановкой столбцов так, чтобы новая верхняя строка была упорядочена в виде $1, 2, \dots, k$.

Например, для подстановки $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 5 & 4 & 2 \end{pmatrix}$ обратной является подстановка

$$\begin{pmatrix} 3 & 1 & 5 & 4 & 2 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 1 & 4 & 3 \end{pmatrix}.$$

Проверим это. Поскольку нейтральный элемент имеет вид $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}$, то имеем в результате умножения

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 5 & 4 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 1 & 4 & 3 \end{pmatrix} &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 1 & 4 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 5 & 4 & 2 \end{pmatrix}. \end{aligned}$$

Таким образом, множество подстановок на $\mathbb{N}_k$ относительно умножения образует группу, которая называется симметрической группой k -й степени S_k .

Пример IV.2.2. Рассмотрим группу S_3 .

Имеем

$$S_3 = \{g_0, g_1, g_2, g_3, g_4, g_5\},$$

$$g_0 = \begin{pmatrix} 123 \\ 123 \end{pmatrix}, g_1 = \begin{pmatrix} 123 \\ 132 \end{pmatrix}, g_2 = \begin{pmatrix} 123 \\ 321 \end{pmatrix}, g_3 = \begin{pmatrix} 123 \\ 213 \end{pmatrix},$$

где

$$g_4 = \begin{pmatrix} 123 \\ 231 \end{pmatrix}, g_5 = \begin{pmatrix} 123 \\ 312 \end{pmatrix}.$$

Сначала найдем порядки элементов S_3 . Имеем

$$O(g_1) = O(g_2) = O(g_3) = 2,$$

так как

$$g_1^2 = g_2^2 = g_3^2 = g_0 = 1_{S_3}.$$

Далее,

$$O(g_4) = O(g_5) = 3,$$

так как

$$g_4^3 = g_5^3 = g_0 = 1_{S_3}, \quad g_4^2 \neq g_0, \quad g_5^2 \neq g_0.$$

Тогда в группе S_3 есть три циклические подгруппы порядка два и одна циклическая подгруппа порядка три, а именно:

$G_1 = \langle g_1 \rangle, G_2 = \langle g_2 \rangle$ и $G_3 = \langle g_3 \rangle$ – циклические подгруппы порядка два;

$G_4 = \langle g_4 \rangle = \langle g_5 \rangle$ – циклическая подгруппа порядка три. Заметим, что $g_5 = g_4^2 \in G_4$, и поэтому $\langle g_4 \rangle = \langle g_5 \rangle$.

Других подгрупп, кроме вышеуказанных, в группе S_3 нет. Действительно, если H есть собственная подгруппа группы S_3 , то по теореме Лагранжа $|H|$ есть делитель $|S_3| = 6$, отличный от 1 или 6. Но тогда любая собственная подгруппа группы S_3 имеет порядок 2 или 3. Следовательно, любая собственная подгруппа группы S_3 является циклической подгруппой, поскольку любая группа простого порядка является циклической группой. Но совокупность циклических подгрупп группы S_3 исчерпывается подгруппами G_1, G_2, G_3 и G_4 . Следовательно, вышеуказанные 4 подгруппы составляют полный набор собственных подгрупп S_3 .

Найдем теперь центр $Z(S_3)$. Покажем, что $Z(S_3) = \{1\}$, где $1 = g_0$. Достаточно убедиться в том, что любой неединичный элемент S_3 не принадлежит центру S_3 , т.е. для каждого такого элемента можно указать хотя бы один перестановочный с ним элемент. Имеем

g_1 и g_2 не перестановочны, так как

$$g_1 g_2 = \begin{pmatrix} 123 \\ 132 \end{pmatrix} \begin{pmatrix} 123 \\ 321 \end{pmatrix} = \begin{pmatrix} 123 \\ 231 \end{pmatrix} = g_4,$$

$$g_2 g_1 = \begin{pmatrix} 123 \\ 321 \end{pmatrix} \begin{pmatrix} 123 \\ 132 \end{pmatrix} = \begin{pmatrix} 123 \\ 312 \end{pmatrix} = g_5$$

и $g_1 g_2 \neq g_2 g_1$.

Заметим, что умножение перестановок производится справа налево, так как определяется как композиция биекций, которая производится справа налево:

$$(f \circ g)(x) = f(g(x)).$$

Далее элементы g_3 и g_4 не перестановочны, так как

$$g_3 g_4 = \begin{pmatrix} 123 \\ 213 \end{pmatrix} \begin{pmatrix} 123 \\ 231 \end{pmatrix} = \begin{pmatrix} 123 \\ 132 \end{pmatrix} = g_1,$$

$$g_4 g_3 = \begin{pmatrix} 123 \\ 231 \end{pmatrix} \begin{pmatrix} 123 \\ 213 \end{pmatrix} = \begin{pmatrix} 123 \\ 321 \end{pmatrix} = g_2$$

и $g_3 g_4 \neq g_4 g_3$.

Наконец, элементы g_5 и g_1 не перестановочны, так как

$$g_5 g_1 = \begin{pmatrix} 123 \\ 312 \end{pmatrix} \begin{pmatrix} 123 \\ 132 \end{pmatrix} = \begin{pmatrix} 123 \\ 321 \end{pmatrix} = g_2,$$

$$g_1 g_5 = \begin{pmatrix} 123 \\ 132 \end{pmatrix} \begin{pmatrix} 123 \\ 312 \end{pmatrix} = \begin{pmatrix} 123 \\ 213 \end{pmatrix} = g_3$$

и $g_1 g_5 \neq g_5 g_1$.

Вернемся снова к изучению произвольной симметрической группы S_n . Имеет место следующее утверждение.

Теорема IV.2.3. Порядок симметрической группы S_n равен

$$|S_n| = n!$$

Доказательство. Пусть $\phi: \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$ есть произвольная подстановка. Для символа 1 имеется n возможностей для $\phi(1)$. Если зафиксировать $\phi(1)$, то для символа 2 имеется $n-1$ возможностей для $\phi(2)$ и т.д. Таким образом, число всех возможных выборов $\phi(1)$, $\phi(2), \dots, \phi(n)$, а следовательно, и число всех подстановок на n символах $1, 2, \dots, n$ равно

$$n \cdot (n-1) \dots 2 \cdot 1 = n!$$

Рассмотрим теперь возможность разложения подстановок в произведение более простых подстановок. Идею разложения поясним на примере.

Пример IV.2.4. В группе S_5 рассмотрим два элемента g и h , где

$$g = \begin{pmatrix} 12345 \\ 23451 \end{pmatrix}; \quad h = \begin{pmatrix} 12345 \\ 34152 \end{pmatrix}.$$

Эти элементы можно представить в виде следующей схемы (рис. 1):

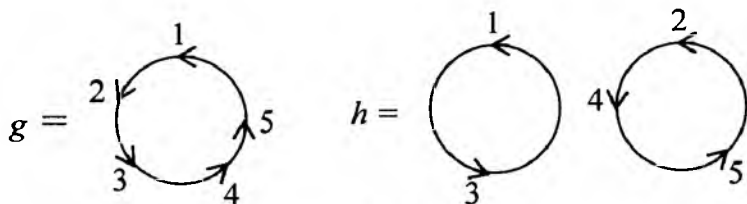


Рис. 1

В этом случае говорят, что подстановка g является циклом длины 5, а подстановка h является произведением двух независимых (т.е. непересекающихся) циклов длины 2 и 3, а именно

$$h = \begin{pmatrix} 12345 \\ 32145 \end{pmatrix} \begin{pmatrix} 12345 \\ 14352 \end{pmatrix}.$$

Условимся записывать подстановки-циклы в виде одной строки, где первый символ переходит при действии подстановки во второй символ, второй символ – в третий и т.д., наконец, последний символ цикла переходит в первый. Тогда имеем

$$\begin{pmatrix} 12345 \\ 23451 \end{pmatrix} = (12345),$$

$$\begin{pmatrix} 12345 \\ 32145 \end{pmatrix} = (13),$$

$$\begin{pmatrix} 12345 \\ 14352 \end{pmatrix} = (245).$$

Заметим, что в двух последних подстановках не указаны неподвижные символы (т.е. те символы, которые подстановка оставляет на месте), но их всегда можно восстановить, если группа подстановок фиксированная.

Возвращаясь к общему случаю, пусть ϕ – произвольная подстановка из S_n . На множестве символов $\{1, 2, \dots, n\}$ зададим бинарное отношение ρ_ϕ следующим образом: для любых $i, j \in \{1, 2, \dots, n\}$ пусть $i \rho_\phi j$ тогда и только тогда, когда существует целое число s , такое, что $j = \phi^s(i)$.

Непосредственно проверяется, что ρ_ϕ есть отношение эквивалентности. Действительно, имеем

а) ρ_ϕ рефлексивно, так как $i = \phi^0(i)$;

б) ρ_ϕ симметрично, так как если $j = \phi^s(i)$, то $i = \phi^{-s}(j)$;

в) ρ_ϕ транзитивно, так как если $j = \phi^s(i)$ и $k = \phi^t(j)$, то $k = \phi^{s+t}(i)$.

Но тогда множество $\{1, 2, \dots, n\}$ разбивается на классы эквивалентности относительно ρ_ϕ , которые называют ϕ -орбитами: если $\Phi = \{1, 2, \dots, n\}$, то

$$\Phi = \Phi_1 \cup \Phi_2 \cup \dots \cup \Phi_r,$$

где каждое множество $\Phi_i, i = 1, 2, \dots, r$, состоит из степеней ϕ , а именно

$$\Phi_i = \{v, \phi(v), \phi^2(v), \dots, \phi^{\ell_i-1}(v)\},$$

где ℓ_i — длина ϕ -орбиты Φ_i .

Полагая

$$\phi_i = (v \phi(v) \dots \phi^{\ell_i-1}(v)) = \begin{pmatrix} v & \phi(v) & \dots & \phi^{\ell_i-2}(v) \\ \phi(v) & \phi^2(v) & \dots & \phi^{\ell_i-1}(v) \end{pmatrix},$$

получаем подстановку, которая называется циклом длины ℓ_i . Действительно,

$$\ell_i \leq |\langle \phi \rangle| = O(\phi) \text{ и } \phi^{\ell_i}(v) = v,$$

причем ℓ_i есть наименьшее число с таким свойством.

Цикл ϕ_i оставляет на месте все символы из $\Phi \setminus \Phi_i$ и $\phi(w) = \phi_i(w)$ для любого $w \in \Phi_i$. Это позволяет называть ϕ_i, ϕ_j для $i \neq j$ независимыми или непересекающимися циклами. Поскольку $\phi^{\ell_i}(w) = w$ для любого $w \in \Phi_i$, то $\phi_i^{\ell_i} = e$, где e — тождественная подстановка (так как ℓ_i — наименьшее число с таким свойством, то $O(\Phi_i) = \ell_i$).

Таким образом, любая подстановка ϕ может быть представлена в виде

$$\phi = \phi_1 \cdot \phi_2 \dots \phi_r,$$

где все циклы $\phi_i, i = 1, 2, \dots, r$, перестановочны между собой.

Замечание. Циклы длины 1, соответствующие неподвижным точкам подстановки, в приведенном выше разложении принято опускать. Например,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 4 & 10 & 8 & 3 & 9 & 2 & 7 & 1 & 5 & 6 \end{pmatrix} = \\ = (1438)(2106)(59)(7) = (1438)(2106)(59).$$

Указанное выше разложение подстановки позволяет легко находить порядок подстановки.

Теорема IV.2.5. Порядок подстановки $\phi \in S_n$ (который равен порядку циклической подгруппы $\langle \phi \rangle$) равен наименьшему общему кратному длин независимых циклов, входящих в разложение ϕ .

Доказательство. Поскольку независимые циклы в разложении $\phi = \phi_1 \cdot \phi_2 \dots \phi_r$ перестановочны, то s -я степень подстановки ϕ записывается в виде

$$\phi^s = \phi_1^s \phi_2^s \dots \phi_r^s, \quad s = 0, 1, 2, \dots$$

Из независимости циклов $\phi_1, \phi_2, \dots, \phi_r$ следует, что $\phi' = e$ тогда и только тогда, когда $\phi_i' = e$ для любого $i = 1, 2, \dots, r$.

Но если $\phi_i' = e$ для любого $i = 1, 2, \dots, r$, то t делится на порядки $O(\phi_i)$ циклов ϕ_i , а следовательно, на длины ℓ_i циклов ϕ_i , поскольку $\ell_i = O(\phi_i)$ для $i = 1, 2, \dots, r$. Тогда t является общим кратным $\ell_i, i = 1, 2, \dots, r$. Пусть $t = O(\phi)$, т.е. t — наименьшее натуральное число, такое, что $\phi' = e$. Тогда t является наименьшим общим кратным чисел $\ell_i, i = 1, 2, \dots, r$, что и требовалось доказать.

Определение IV.2.6. Цикл длины 2 называется транспозицией.

Любая транспозиция имеет вид $\tau = (ij)$ и оставляет на месте все символы, отличные от i, j . Имеет место следующая

Теорема IV.2.7. Каждая подстановка $\phi \in S_n$ является произведением транспозиций.

Доказательство. Разложим подстановку ϕ в произведение независимых циклов

$$\phi = \phi_1 \cdot \phi_2 \dots \phi_r.$$

Каждый цикл можно разложить в произведение транспозиций, что можно сделать, например, так:

$$(1\ 2 \dots \ell - 1\ \ell) = (1\ \ell - 1) \dots (13)(12).$$

Но тогда ϕ есть произведение транспозиций.

Замечание. В отличие от разложения подстановок в произведение независимых циклов, которое однозначно с точностью до порядка записи сомножителей, разложение подстановок в произведение транспозиций неоднозначно.

Пример IV.2.8. Пусть $\phi \in S_n$ имеет вид

$$\phi = \begin{pmatrix} 12345 \\ 23154 \end{pmatrix} = (123)(45).$$

Для подстановки ϕ можно указать различные разложения в произведение транспозиций:

$$\phi = (13)(12)(45) = (23)(13)(45) = (13)(24)(12)(14)(45).$$

Хотя разложение подстановки в произведение транспозиций неоднозначно, один числовой инвариант для этих разложений можно указать. Как видно из примера IV.2.8, где приведены 3 и 5 транспозиций, т.е. нечетное число транспозиций, — это четность числа транспозиций в разложении.

Определение IV.2.9. Пусть $\phi \in S_n$ имеет вид

$$\phi = \psi_1 \cdot \psi_2 \dots \psi_p,$$

где $\psi_1, \psi_2, \dots, \psi_p$ — транспозиции.

Число $\varepsilon_\phi = (-1)^p$ называется четностью подстановки ϕ . Подстановка ϕ называется четной, если $\varepsilon_\phi = 1$, и нечетной, если $\varepsilon_\phi = -1$.

Другими словами, четные подстановки разложимы в четное число транспозиций, а нечетные подстановки — в нечетное число транспозиций. Заметим, что любая транспозиция является нечетной подстановкой. Приведем без доказательства формулировку следующей теоремы.

Теорема IV.2.10. Четность ε_ϕ подстановки $\phi \in S_n$ полностью определяется этой подстановкой и не зависит от способа разложения подстановки ϕ в произведение транспозиций.

Замечание. Все четные подстановки в группе S_n образуют подгруппу A_n , которая называется знакопеременной группой степени n . Действительно, если ϕ и ψ – четные подстановки, то, поскольку ψ^{-1} есть четная подстановка ($\psi \cdot \psi^{-1} = e$, где ψ и e – четные подстановки, следовательно, ψ^{-1} – четная подстановка), тогда $\phi\psi^{-1}$ есть четная подстановка. Таким образом, по критерию подгруппы имеем, что A_n есть подгруппа группы S_n .

Приведем также без доказательства формулу для вычисления четности подстановки по разложению в произведение независимых циклов.

Пусть подстановка $\phi \in S_n$ разложена в произведение независимых циклов ϕ_i длин соответственно $\ell_i, i = 1, 2, \dots, r$,

$$\phi = \phi_1 \cdot \phi_2 \dots \phi_r.$$

Тогда

$$\varepsilon_\phi = (-1)^{\sum_{i=1}^r (\ell_i - 1)}.$$

Определение IV.2.11. Множество $GL_n(\mathbb{R})$ действительных невырожденных квадратных матриц порядка n называется общей линейной группой порядка n над $\mathbb{R}$.

Корректность определения IV.2.11 вытекает из следующих утверждений:

а) произведение невырожденных действительных квадратных матриц порядка n есть невырожденная квадратная матрица порядка n (так как определитель произведения таких матриц равен произведению ненулевых определений матриц-сомножителей, то определитель произведения матриц отличен от нуля);

б) операция умножения произвольных (а следовательно, и невырожденных) действительных квадратных матриц порядка n ассоциативна;

в) единичная матрица порядка n является нейтральным элементом по умножению матриц в $GL_n(\mathbb{R})$;

г) для каждой матрицы $A \in GL_n(\mathbb{R})$ имеем

$$A^{-1} \in GL_n(\mathbb{R})$$

(так как определитель $|A^{-1}| = \frac{1}{|A|}$ и $|A| \neq 0$, то $|A^{-1}| \neq 0$ и, следовательно,

$$A^{-1} \in GL_n(\mathbb{R}).$$

Группа $GL_n(\mathbb{R})$ является бесконечной некоммутативной группой, и в ней есть как элементы конечного порядка, так и элементы бесконечного порядка.

Пример IV.2.12. В группе $GL_n(\mathbb{R})$ рассмотрим следующие элементы:

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}; \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}; \quad C = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Непосредственно проверяется, что матрица A коммутирует как с матрицей B , так и с матрицей C , но матрицы B и C не коммутируют:

$$AB = BA, \quad AC = CA, \quad BC \neq CB.$$

Матрицы A и B имеют конечный порядок:

$$O(A) = 2, \quad O(B) = 4,$$

поскольку $A^2 = E$ и $B^2 \neq E$, $B^3 \neq E$, $B^4 = E$, а матрица C имеет бесконечный порядок, так как $C^k \neq E$ для любого натурального k . Следовательно, мы имеем две конечные циклические подгруппы и одну бесконечную циклическую подгруппу:

$$\langle A \rangle = \{E, A\}, \quad |\langle A \rangle| = 2,$$

$$\langle B \rangle = \{E, B, B^2, B^3\}, \quad |\langle B \rangle| = 4,$$

$$\langle C \rangle = \{\dots, C^{-k}, \dots, C^{-2}, C^{-1}, E, C, C^2, \dots, C^k, \dots\}, \quad |\langle C \rangle| = \infty.$$

Группа $GL_n(\mathbb{R})$ содержит кроме циклических подгрупп также и нециклические подгруппы.

Теорема IV.2.13. Следующие подмножества группы $GL_n(\mathbb{R})$ являются в ней подгруппами:

а) $GL_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid |A| = 1\};$

б) $D_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid A = \begin{pmatrix} a_1 & a^2 & \dots & 0 \\ & & & a_n \end{pmatrix},$

$$0 \neq a_i \in \mathbb{R}, \quad i = 1, \dots, n\};$$

$$\text{в) } C_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid A = \begin{pmatrix} a & & & 0 \\ & a & \ddots & \\ & 0 & & a \end{pmatrix}, \quad 0 \neq a \in \mathbb{R}\};$$

$$\text{г) } T_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ & a_{22} & \ddots & a_{2n} \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}, \quad a_{ij} \in \mathbb{R}\};$$

$$\text{д) } UT_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid A = \begin{pmatrix} 1 & a_{12} & \dots & a_{1n} \\ & 1 & & a_{2n} \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix}, \quad a_{ij} \in \mathbb{R}\}.$$

Доказательство. Используем для доказательства критерий подгруппы (теорема IV.1.27).

а) Пусть $A, B \in SL_n(\mathbb{R})$, т.е. $|A| = |B| = 1$. Тогда

$$|AB^{-1}| = |A| \cdot |B^{-1}| = |A| \cdot \frac{1}{|B|} = 1,$$

и, следовательно, $AB^{-1} \in SL_n(\mathbb{R})$. Таким образом, по критерию подгруппы (теорема IV.1.27) имеем, что $SL_n(\mathbb{R})$ есть подгруппа группы $GL_n(\mathbb{R})$.

б) Пусть $A, B \in D_n(\mathbb{R})$, т.е. $A = \text{diag}(a_1, \dots, a_n)$, $a_i \neq 0$,

$$B = \text{diag}(b_1, \dots, b_n), \quad b_i \neq 0,$$

тогда

$$\begin{aligned} AB^{-1} &= \text{diag}(a_1, \dots, a_n) \cdot \text{diag}\left(\frac{1}{b_1}, \dots, \frac{1}{b_n}\right) = \\ &= \text{diag}\left(\frac{a_1}{b_1}, \dots, \frac{a_n}{b_n}\right) \in D_n(\mathbb{R}). \end{aligned}$$

Следовательно, по критерию подгруппы (теорема IV.1.27) имеем, что $D_n(\mathbb{R})$ есть подгруппа группы $GL_n(\mathbb{R})$.

в) Пусть $A, B \in C_n(\mathbb{R})$, т.е. $A = \text{diag}(a, \dots, a)$, $a \neq 0$,

$$B = \text{diag}(b, \dots, b), b \neq 0,$$

тогда

$$AB^{-1} = \text{diag}\left(\frac{a}{b}, \dots, \frac{a}{b}\right) \in C_n(\mathbb{R}).$$

Следовательно, $C_n(\mathbb{R})$ есть подгруппа группы $GL_n(\mathbb{R})$. Пункты «г» и «д» проверяются аналогично. Теорема доказана.

Подгруппы группы $GL_n(\mathbb{R})$, указанные в теореме IV.2.13, имеют следующие названия:

$GL_n(\mathbb{R})$ – специальная линейная группа порядка n над $\mathbb{R}$;

$D_n(\mathbb{R})$ – диагональная группа порядка n над $\mathbb{R}$;

$C_n(\mathbb{R})$ – скалярная группа порядка n над $\mathbb{R}$;

$T_n(\mathbb{R})$ – верхнетреугольная группа порядка n над $\mathbb{R}$;

$UT_n(\mathbb{R})$ – верхнеунитреугольная группа порядка n над $\mathbb{R}$.

Теперь укажем центр группы $GL_n(\mathbb{R})$.

Теорема IV.2.14. Подгруппа $C_n(\mathbb{R})$ является центром группы $GL_n(\mathbb{R})$.

Доказательство. Пусть Z есть центр группы $GL_n(\mathbb{R})$. Покажем сначала, что $C_n(\mathbb{R}) \subseteq Z$.

Пусть $A \in C_n(\mathbb{R})$, $A = \text{diag}(a, \dots, a)$, $a \neq 0$ и $B \in GL_n(\mathbb{R})$ – произвольный элемент. Покажем, что A и B коммутируют, т.е. $AB = BA$.

Пусть

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

тогда

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

$$BA = \begin{pmatrix} b_{11}a & b_{12}a & b_{1n}a \\ b_{21}a & b_{22}a & b_{2n}a \\ \dots & \dots & \dots \\ b_{n1}a & b_{n2}a & b_{nn}a \end{pmatrix}.$$

Поскольку умножение действительных чисел коммутативно, то $AB = BA$ и $A \in Z$. Таким образом, $C_n(\mathbb{R}) \subseteq Z$.

Докажем теперь обратное включение $Z \subseteq C_n(\mathbb{R})$. Пусть $A \in Z$ — произвольный элемент, покажем, что тогда A есть скалярная матрица, $A = \text{diag}(a, \dots, a)$, $a \neq 0$.

Рассмотрим матрицы вида

$$t_{ij}(\alpha) = \begin{pmatrix} 1 & 0 & 0 & 0 & \vdots & 0 \\ 0 & 1 & 0 & \alpha & \vdots & 0 \\ 0 & 0 & 1 & 0 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \vdots & 1 \end{pmatrix},$$

j

где $\alpha \neq 0$ находится на пересечении i -й строки и j -го столбца, все диагональные элементы равны 1, а все остальные элементы равны 0 — такие матрицы $t_{ij}(\alpha)$ называются трансвекциями. Заметим, что умножение произвольной матрицы X на трансвекцию $t_{ij}(1)$ справа равносильно прибавлению к j -му столбцу матрицы X i -го столбца этой матрицы, а умножение на трансвекцию $t_{ij}(1)$ слева равносильно прибавлению к j -й строке матрицы X i -й строки этой матрицы. Непосредственными вычислениями с матричными элементами можно показать, что если $A \in Z$, то из перестановочности матрицы A со всеми трансвекциями следует, что все внедиагональные элементы a_{ij} матрицы A , где $i \neq j$, равны нулю, а все диагональные элементы совпадают, т.е. $A = \text{diag}(a, \dots, a)$, $a \neq 0$.

Лекция 3. КОЛЬЦА

Определение IV.3.1. Будем говорить, что на множестве K задана структура кольца (или короче: K есть кольцо), если на множестве K заданы две алгебраические операции, которые принято обозначать «+» (сложение) и « $\cdot$ » (умножение), причём выполняются следующие свойства:

а) $(K; +)$ есть абелева группа;

б) $(K; \cdot)$ есть полугруппа;

в) имеет место двойная дистрибутивность умножения относительно сложения, т.е. для любых $a, b, c \in K$

$$a(b + c) = ab + ac,$$

$$(a + b)c = ac + bc.$$

Нейтральный элемент абелевой группы $(K; +)$, которую принято называть аддитивной группой кольца K , принято обозначать 0_K или просто 0 и называть нулём кольца K . Полугруппу $(K; \cdot)$ принято называть мультипликативной полугруппой кольца K . Если в мультипликативной полугруппе $(K; \cdot)$ кольца K есть нейтральный элемент, то его принято обозначать 1_K или просто 1 и называть единицей кольца K . В этом случае говорят, что K есть кольцо с единицей. Если умножение в кольце K коммутативно, то говорят, что K — коммутативное кольцо.

Пример IV.3.2. Множество $\mathbb{Z}$ всех целых чисел является кольцом с единицей относительно обычных операций сложения и умножения целых чисел. Аналогично $\mathbb{Z}$ можно убедиться, что $\mathbb{Q}$, $\mathbb{R}$ и $\mathbb{C}$ — кольца.

Рассмотрим теперь простейшие свойства колец.

Предложение IV.3.3. Пусть K — произвольное кольцо, тогда справедливы следующие утверждения:

а) $x \cdot 0 = 0 \cdot x = 0$ для любого $x \in K$;

б) $(-x) \cdot y = x \cdot (-y) = -(xy)$ для любых $x, y \in K$;

в) $(-x)(-y) = xy$ для любых $x, y \in K$;

г) если K — кольцо с единицей, то $(-1)x = -x$ для любого $x \in K$.

Доказательство.

а) Имеем

$$x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0. \quad (1)$$

К обеим частям равенства (1) прибавим $-(x \cdot 0)$, тогда имеем

$$-(x \cdot 0) + x \cdot 0 = -(x \cdot 0) + (x \cdot 0 + x \cdot 0). \quad (2)$$

Левая часть равенства (2) равна нулю, а правая часть

$$\begin{aligned} -(x \cdot 0) + (x \cdot 0 + x \cdot 0) &= (-x \cdot 0 + x \cdot 0) + x \cdot 0 = \\ &= 0 + x \cdot 0 = x \cdot 0. \end{aligned}$$

Таким образом, получаем

$$x \cdot 0 = 0.$$

Аналогично получается $0 \cdot x = 0$.

б) Рассмотрим следующую сумму элементов кольца K :

$$(-x) \cdot y + x \cdot y = (-x + x)y = 0 \cdot y.$$

По свойству «а» имеем $0 \cdot y = 0$, тогда $(-x) \cdot y + x \cdot y = 0$.

Это означает, что $(-x) \cdot y$ является противоположным к $x \cdot y$ (т.е. обратным к $x \cdot y$ относительно сложения). Таким образом,

$$(-x) \cdot y = -(x \cdot y).$$

Аналогично получаем, что $x \cdot (-y) = -(x \cdot y)$.

в) Дважды применяя свойство «б»), получаем

$$(-x) \cdot (-y) = -(x \cdot (-y)) = -(-(x \cdot y)) = x \cdot y.$$

г) Применяя свойство «б»), получаем

$$(-1) \cdot x = -(1 \cdot x) = -x.$$

Хотя понятие кольца возникло как обобщение свойств целых чисел, есть примеры, которые существенно различны по своим свойствам от кольца целых чисел.

Пример IV.3.4. Пусть G есть произвольная абелева группа (например, $G = (\mathbb{Z}; +)$ – группа всех целых чисел по сложению). Определим алгебраическую операцию умножения на G следующим образом:

$$\text{для любых } x, y \in G \text{ пусть } x \cdot y = 0_G,$$

где 0_G – нейтральный элемент относительно сложения в группе G .

Очевидно, что $(G; \cdot)$ – полугруппа и, кроме того, выполняется свойство двойкой дистрибутивности умножения относительно сложения в G . Таким образом, $(G; +, \cdot)$ есть кольцо, которое называется кольцом с

нулевым умножением. Заметим, что кольцо с нулевым умножением не имеет единицы.

Определение IV.3.5. Пусть K есть кольцо с единицей, рассмотрим суммы вида: $1, 1+1, \dots, \underbrace{1+1+\dots+1}_l, \dots$, обозначим их соответственно

$1, 2 \cdot 1, \dots, l \cdot 1, \dots$, где $l \in \mathbb{N}$, и назовём целыми элементами кольца K . Рассмотрим два случая.

Случай 1. Если все целые элементы $l \cdot 1$, где $l \neq 0$, являются ненулевыми элементами кольца K , то в этом случае говорят, что K есть кольцо нулевой характеристики и обозначают $\text{char } K = 0$. Заметим, что кольцо нулевой характеристики содержит бесконечное число целых элементов. Достаточно убедиться в том, что различным натуральным числам l_1 и l_2 соответствуют различные целые элементы $l_1 \cdot 1$ и $l_2 \cdot 1$ кольца K . Действительно, пусть $l_1 \neq l_2$ и $l_1 \cdot 1 = l_2 \cdot 1$, тогда имеем

$$\underbrace{1+1+\dots+1}_{l_1} = \underbrace{1+1+\dots+1}_{l_2}. \quad (1)$$

Пусть для определённости $l_1 > l_2$, тогда, прибавляя к обеим частям (1) элемент $-\underbrace{(1+1+\dots+1)}_{l_2} \in K$, получаем

$$\begin{aligned} & -\underbrace{(1+1+\dots+1)}_{l_2} + \underbrace{(1+1+\dots+1)}_{l_1} = \\ & = -\underbrace{(1+1+\dots+1)}_{l_2} + \underbrace{(1+1+\dots+1)}_{l_2}. \end{aligned} \quad (2)$$

Левая часть равенства (2) имеет вид

$$\underbrace{(-1) + (-1) + \dots + (-1)}_{l_2} + \underbrace{(1+1+\dots+1)}_{l_1} = (l_1 - l_2) \cdot 1,$$

а правая часть равенства (2) равна нулю. Таким образом, $(l_1 - l_2) \cdot 1 = 0$, где $(l_1 - l_2) \in \mathbb{N}$, противоречие с определением кольца нулевой характеристики.

Итак, в кольце K нулевой характеристики имеется бесконечное множество целых элементов:

$$1, 2 \cdot 1, \dots, l \cdot 1, \dots, l = 1, 2, \dots$$

Случай 2. Пусть для некоторого натурального числа l имеем $l \cdot 1 = 0$, тогда существует наименьшее натуральное l с этим свойством, которое называется характеристикой кольца $\mathbb{R}$, обозначается $\text{char } K = l$.

Пример IV.3.6. Кольца $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ и $\mathbb{C}$ являются кольцами нулевой характеристики. Примеры колец ненулевой характеристики будут приведены далее.

Определение IV.3.7. Пусть K – кольцо с единицей. Будем говорить, что элемент $u \in K$ является обратимым элементом, если существует элемент $u^{-1} \in K$ такой, что

$$u \cdot u^{-1} = u^{-1} \cdot u = 1.$$

Предложение IV.3.8. Для любого кольца K с единицей множество K^* всех обратимых элементов кольца K является группой относительно умножения.

Доказательство. Проверим, что умножение является алгебраической операцией во множестве K^* . Пусть $x, y \in K^*$, тогда существуют $x^{-1}, y^{-1} \in K^*$, такие, что

$$x \cdot x^{-1} = x^{-1} \cdot x = 1, \quad y \cdot y^{-1} = y^{-1} \cdot y = 1.$$

Отсюда следует, что $x, y \in K^*$, и тогда K^* замкнуто относительно умножения. Действительно, $(x \cdot y)^{-1}$ существует и $(x \cdot y)^{-1} = y^{-1} \cdot x^{-1}$, так как имеем

$$(y^{-1} \cdot x^{-1})(x \cdot y) = y^{-1}(x^{-1}x)y = y^{-1} \cdot 1 \cdot y = y^{-1}y = 1.$$

Ассоциативность умножения имеет место по условию. Единица кольца K является обратимым элементом и, следовательно, $1 \in K^*$. Наконец, поскольку любой $x \in K^*$ является обратимым элементом кольца K , то существует обратный к x элемент K , обозначаем x^{-1} , и тогда $(K^*; \cdot)$ является группой.

Пример IV.3.9. $\mathbb{Z}^* = \{-1, 1\}$, $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$, $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$, $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$.

Определение IV.3.10. Подмножество L кольца K называется подкольцом, если L является кольцом относительно операций сложения и умножения в кольце K (обозначение: $L \leq K$).

Теорема IV.3.11. Подмножество $L \subseteq K$ кольца K является подкольцом в K тогда и только тогда, когда выполняются следующие условия:

- а) для любых $x, y \in L$ имеем $x - y \in L$;
- б) для любых $x, y \in L$ имеем $x \cdot y \in L$.

Достаточность. Покажем, что $(L; +)$ есть подгруппа в $(K; +)$. По критерию подгруппы (теорема IV.1.27) достаточно проверить, что для любых $x, y \in L$ имеем $x - y \in L$. Но в нашем случае это верно по условию «а». Следовательно, $(L; +)$ есть абелева группа по сложению. По условию «б» получаем, что множество L замкнуто относительно умножения, и, следовательно, $(L; \cdot)$ есть полугруппа по умножению, ввиду ассоциативности умножения в K , а следовательно, и в L . Наконец, дистрибутивность умножения относительно сложения выполняется для всего кольца K , а следовательно, и для подмножества L кольца K . Таким образом, L является кольцом относительно операций сложения и умножения кольца K , т.е. L есть подкольцо кольца K .

Необходимость. Пусть L есть подкольцо кольца K , тогда L есть кольцо относительно операций сложения и умножения кольца K . Но тогда для любых $x, y \in L$ имеем, что $(-y) \in L$ и $x + (-y) = x - y \in L$, так как L – абелева группа по сложению. Далее, $x \cdot y \in L$, так как L есть полугруппа по умножению. Теорема доказана.

Пример IV.3.12. Каждое из колец $\mathbb{Z}$, $\mathbb{Q}$ и $\mathbb{R}$ является подкольцом соответственно в $\mathbb{Q}$, $\mathbb{R}$ и $\mathbb{C}$, т.е. $\mathbb{Z} < \mathbb{Q} < \mathbb{R} < \mathbb{C}$.

Приведём теперь примеры нечисловых колец.

Пример IV.3.13. Пусть K есть коммутативное кольцо с единицей, и рассмотрим множество $K[x]$ всевозможных многочленов от неизвестных x с коэффициентами из кольца K , т.е. любой $f \in K[x]$ имеет вид

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0,$$

где $a_n, \dots, a_1, a_0 \in K$.

Терминология: степень многочлена $f(x)$ есть $\deg(f) = n$, старший коэффициент многочлена $f(x)$ есть $a_n \neq 0$.

Если определить операции сложения и умножения многочленов по обычным «школьным» правилам выполнения этих операций, то непосредственно проверяется, что получается коммутативное кольцо с единицей.

Пример IV.3.14. Пусть K есть коммутативное кольцо с единицей, и рассмотрим множество $M_n(K)$ всех квадратных матриц порядка n с элементами из кольца K . Непосредственно проверяется, что относительно обычных операций сложения и умножения матриц $M_n(K)$ есть некоммутативное кольцо с единицей.

Пример IV.3.15. Пусть X – произвольное множество, K – произвольное кольцо с единицей. Обозначим через K^X множество всех отображений из X в K :

$$K^X = \{f | f: X \rightarrow K\}.$$

Если кольцевые операции в K обозначим, как обычно, через «+» и « $\cdot$ », то определим алгебраические операции в K^X следующим образом: для любых $f, g \in K^X$ пусть

$$(f \oplus g)(x) = f(x) + g(x), \quad x \in X,$$

$$(f \otimes g)(x) = f(x) \cdot g(x), \quad x \in X.$$

Непосредственно проверяется, что $(K^X; \oplus, \otimes)$ является кольцом с единицей (заметим, что нулём в этом кольце является отображение $0: x \mapsto 0_K$ для любого $x \in X$, а единицей – отображение $1: x \mapsto 1_K$ для любого $x \in X$), которое называется кольцом функций из X в K . В частном случае $X = K = \mathbb{R}$ получаем кольцо $\mathbb{R}^{\mathbb{R}}$ действительных функций действительного переменного.

Будем обозначать через $\text{Im } f$ – образ отображения f , т.е. для $f: X \rightarrow K$

$$\text{Im } f = \{y \in K | \exists x \in X \text{ такой, что } f(x) = y\},$$

а через $\text{Ker } f$ – множество тех элементов X , которые f отображает в нуль кольца K , а именно

$$\text{Ker } f = \{x \in X | f(x) = 0_K\}.$$

Другими словами, $\text{Ker } f$ есть множество нулей (корней) отображения (функции) f .

Пример IV.3.16. В кольце $\mathbb{Z}$ зададим бинарное отношение «сравнимости по модулю n » следующим образом: для любых $x, y \in \mathbb{Z}$ пусть x сравнимо с y по модулю n тогда и только тогда, когда $x - y$ нацело делится на n .

Легко видеть, что данное отношение является отношением эквивалентности и, следовательно, множество $\mathbb{Z}$ разбивается на классы эквивалентности по модулю n (обозначение: $\bar{x} = x \bmod n$). Обозначим через $\mathbb{Z}_n$ множество всех классов эквивалентности, которые принято называть

классами вычетов по модулю n . Определим в $\mathbb{Z}_n$ алгебраические операции сложения и умножения по следующему правилу:

$$x \bmod n + y \bmod n = (x + y) \bmod n,$$

$$(x \bmod n) \cdot (y \bmod n) = (xy) \bmod n.$$

Непосредственно проверяется, что $(\mathbb{Z}_n; +, \cdot)$ есть коммутативное кольцо с единицей.

Например, пусть $n = 4$, тогда

$$\mathbb{Z}_4 = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}\},$$

где $\bar{0} = \{4k \mid k \in \mathbb{Z}\},$

$$\bar{1} = \{4k + 1 \mid k \in \mathbb{Z}\},$$

$$\bar{2} = \{4k + 2 \mid k \in \mathbb{Z}\},$$

$$\bar{3} = \{4k + 3 \mid k \in \mathbb{Z}\}.$$

Операции в $\mathbb{Z}_4$ задаются табл. 1 и 2:

Таблица 1

+	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$

Таблица 2

$\times$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{0}$	$\bar{2}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

Рассмотрим теперь ещё одно свойство колец.

Определение IV.3.17. Будем говорить, что ненулевые элементы x и y кольца K являются делителями нуля, если $x \cdot y = 0$. Назовём кольцо K областью целостности, если K не содержит делителей нуля.

Пример IV.3.18. $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ и $\mathbb{C}$ являются областями целостности, а $\mathbb{Z}_4$ нет, так как $\bar{2} \cdot \bar{2} = 0$ в $\mathbb{Z}_4$.

Определение IV.3.19. Пусть $K_1, K_2, \dots, K_n$ – кольца, и рассмотрим декартово произведение $K = K_1 \times K_2 \times \dots \times K_n$. Определим в K структуру абелевой группы относительно алгебраической операции сложения сле-

дующим образом: для любых элементов $x = (x_1, x_2, \dots, x_n)$ и $y = (y_1, y_2, \dots, y_n)$ из множества K пусть

$$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n).$$

Тогда абелева группа $(K; +)$ является прямой суммой аддитивных групп колец $K_1, K_2, \dots, K_n$. Определим в K структуру полугруппы относительно алгебраической операции умножения следующим образом: для любых элементов $x = (x_1, x_2, \dots, x_n)$ и $y = (y_1, y_2, \dots, y_n)$ из множества K пусть

$$x \cdot y = (x_1 \cdot y_1, x_2 \cdot y_2, \dots, x_n \cdot y_n),$$

где $x_i \cdot y_i$ — произведение x_i и y_i как элементов кольца K_i для любого $i = 1, 2, \dots, n$. Заметим, что $(K; \cdot)$ является полугруппой, а именно прямым произведением полугрупп $(K_i; \cdot)$ для $i = \overline{1, n}$. Далее, умножение в K имеет свойство двоякой дистрибутивности относительно сложения. Действительно, для любых элементов $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n)$ и $z = (z_1, z_2, \dots, z_n)$ из множества K имеем

$$\begin{aligned} x \cdot (y + z) &= (x_1, \dots, x_n) \cdot (y_1 + z_1, \dots, y_n + z_n) = \\ &= (x_1(y_1 + z_1), x_2(y_2 + z_2), \dots, x_n(y_n + z_n)) = \\ &= (x_1 y_1 + x_1 z_1, x_2 y_2 + x_2 z_2, \dots, x_n y_n + x_n z_n) = \\ &= (x_1 y_1, x_2 y_2, \dots, x_n y_n) + (x_1 z_1, x_2 z_2, \dots, x_n z_n) = x \cdot y + x \cdot z. \end{aligned}$$

Аналогично проверяется, что

$$(x + y)z = xz + yz.$$

Таким образом, $(K; +, \cdot)$ является кольцом, которое называют прямым произведением колец $K_1, K_2, \dots, K_n$.

Пример IV.3.20. Пусть $K_1 = K_2 = \mathbb{Z}$ и $K = K_1 \times K_2$ — прямое произведение двух копий кольца $\mathbb{Z}$. Хотя $\mathbb{Z}$ — область целостности, но $K = \mathbb{Z} \times \mathbb{Z}$ имеет делители нуля. Действительно, $(1, 0)$ и $(0, 1)$ — ненулевые элементы $K = \mathbb{Z} \times \mathbb{Z}$ (так как нулём K является $(0, 0)$), но $(1, 0) \cdot (0, 1) = (0, 0)$.

Пример IV.3.21. В кольце $\mathbb{Z}_4$ есть делители нуля. Действительно, из табл. 3 имеем умножения в $\mathbb{Z}_4$, что $\bar{2} \cdot \bar{2} = \bar{0}$.

Таблица 3

$\cdot$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{0}$	$\bar{2}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

Пример IV.3.22. В кольце $\mathbb{Z}_3$ нет делителей нуля. Действительно, из рассмотрения табл. 4 умножения в кольце $\mathbb{Z}_3$ видно, что $\mathbb{Z}_3$ не имеет делителей нуля.

Таблица 4

$\cdot$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{1}$

Рассмотрим теперь группы $\mathbb{Z}_n^*$ обратимых элементов кольца $\mathbb{Z}_n$. Прежде всего нужно выяснить, когда ненулевой элемент кольца $\mathbb{Z}_n$ является обратимым элементом.

Теорема IV.3.23. Ненулевой элемент $\bar{0} \neq \bar{s} \in \mathbb{Z}_n$ является обратимым элементом тогда и только тогда, когда $\text{НОД}(s, n) = 1$.

Достаточность. Пусть $\bar{0} \neq \bar{s} \in \mathbb{Z}_n$ и $\text{НОД}(s, n) = 1$. Тогда существуют целые числа k и l , такие, что $ks + ln = 1$. Отсюда следует, что

$$\bar{1} = \overline{ks + ln} = \overline{ks} + \overline{ln} = \overline{k \cdot s} + \overline{l \cdot n} = \overline{k \cdot s},$$

так как $\bar{n} = \bar{0}$ и, следовательно, $\bar{l \cdot n} = \bar{0}$. Таким образом, $\bar{k} = \bar{s}^{-1}$ и $\bar{s}$ — обратимый элемент в $\mathbb{Z}_n$.

Необходимость. Пусть $\bar{0} \neq \bar{s}$ – обратимый элемент кольца $\mathbb{Z}_n$, т.е. существует обратный к $\bar{s}$ элемент $\bar{s}^{-1} \in \mathbb{Z}_n$, такой, что $\bar{s} \cdot \bar{s}^{-1} = \bar{1} = \bar{s}^{-1} \cdot \bar{s}$.

Случай 1. $s \mid n$.

В этом случае $\text{НОД}(s, n) = s$. Покажем, что отсюда следует $s = 1$.

Так как $s \mid n$, то $n = s \cdot t$ для некоторого $t \in \mathbb{Z}$. Имеем

$$\bar{t} = \bar{1} \cdot \bar{t} = \bar{1} \cdot \bar{t} = \left(\bar{s}^{-1} \cdot \bar{s} \right) \cdot \bar{t} = \bar{s}^{-1} \cdot (\bar{s} \cdot \bar{t}) = \bar{s}^{-1} \cdot \bar{n} = \bar{s}^{-1} \cdot \bar{0} = \bar{0}.$$

Но равенство $\bar{t} = \bar{0}$ означает, что t нацело делится на n , т.е. $n \mid t$. С другой стороны, из $n = s \cdot t$ следует, что $t \mid n$. Таким образом, $t = n$, и тогда из $n = s \cdot t$ вытекает, что $s = 1$. С учётом $\text{НОД}(s, n) = s$ в данном случае получаем, что $\text{НОД}(s, n) = 1$.

Случай 2. s не делит n .

В этом случае $\text{НОД}(s, n) = d < s$, в частности, $d \mid s$ и $d < s$, $d \mid n$ и $d < n$. Так как d делит s , то $s = d \cdot s'$ для некоторого $s' \in \mathbb{Z}$. Поскольку $\bar{s}$ обратим в $\mathbb{Z}_n$, то

$$\bar{1} = \bar{s} \cdot \bar{s}^{-1} = \overline{d \cdot s'} \cdot \bar{s}^{-1} = (\bar{d} \cdot \bar{s}') \cdot \bar{s}^{-1} = \bar{d} \cdot (\bar{s}' \cdot \bar{s}^{-1}).$$

Отсюда следует, что $\bar{d}$ является обратимым элементом $\mathbb{Z}_n$, причём $d \mid n$. Следовательно, $\bar{d}$ удовлетворяет случаю 1, и тогда $d = 1$. Таким образом, $\text{НОД}(s, n) = d = 1$.

Пример IV.3.24. Пусть $n = 6$, тогда обратимыми элементами в кольце $\mathbb{Z}_6$ являются $\bar{1}$, $\bar{5}$,

$\bar{1}$ – тривиальный обратимый элемент,
 $\bar{5}$ – обратим, так как $\text{НОД}(5, 6) = 1$.

Заметим, что $\bar{5}^{-1} = \bar{5}$, так как $\bar{5} \cdot \bar{5} = \overline{25} = \bar{1}$ в $\mathbb{Z}_6$. Таким образом, $\mathbb{Z}_6^* = \{\bar{1}, \bar{5}\}$, причём $\mathbb{Z}_6^* = \langle \bar{5} \rangle$ – циклическая группа второго порядка.

Всегда ли группа обратимых элементов кольца классов вычетов является циклической группой?

Пример IV.3.25. Пусть $n = 8$, тогда обратимыми элементами в кольце $\mathbb{Z}_8$ являются $\bar{1}$, $\bar{3}$, $\bar{5}$, $\bar{7}$,

$\bar{1}$ – единица группы $\mathbb{Z}_8^*$,

$\bar{3}$ – обратим, так как $\text{НОД}(3, 8) = 1$,

$\bar{5}$ – обратим, так как $\text{НОД}(5, 8) = 1$,

$\bar{7}$ – обратим, так как $\text{НОД}(7, 8) = 1$.

Таким образом $\mathbb{Z}_8^* = \{\bar{1}, \bar{3}, \bar{5}, \bar{7}\}$. Найдём порядки элементов группы $\mathbb{Z}_8^*$:

$$O(\bar{3}) = 2, \text{ так как } \bar{3}^2 = \bar{1},$$

$$O(\bar{5}) = 2, \text{ так как } \bar{5}^2 = \bar{1},$$

$$O(\bar{7}) = 2, \text{ так как } \bar{7}^2 = \bar{1}.$$

Порядок группы $\mathbb{Z}_8^*$ равен 4, поэтому если бы группа $\mathbb{Z}_8^*$ была циклической, то среди элементов этой группы должен быть элемент порядка 4. Поскольку таких элементов нет, то группа $\mathbb{Z}_8^*$ – нециклическая группа. Каково строение этой группы?

Поскольку кольца классов вычетов являются коммутативными, то группы обратимых элементов этих колец – конечные абелевы группы. Тогда группа $\mathbb{Z}_n^*$ для любого n является в общем случае прямым произведением циклических групп, в частности, для некоторых n – циклическими группами.

Найдём представление группы $\mathbb{Z}_8^*$ в виде прямого произведения циклических групп, а именно покажем, что

$$\mathbb{Z}_8^* = \langle \bar{3} \rangle \times \langle \bar{5} \rangle.$$

По определению прямого произведения подгрупп, поскольку группа $\mathbb{Z}_8^*$ – абелева, то пункт «б» этого определения выполняется (любые два элемента из подгрупп $\langle \bar{3} \rangle$ и $\langle \bar{5} \rangle$ перестановочны). Проверим справедливость пункта «а», т.е. любой элемент $\mathbb{Z}_8^*$ однозначно представим в виде произведения элементов из подгрупп $\langle \bar{3} \rangle$ и $\langle \bar{5} \rangle$. Однозначность этого представления равносильна условию $\langle \bar{3} \rangle \cap \langle \bar{5} \rangle = \{\bar{1}\}$. Поскольку $\langle \bar{3} \rangle = \{\bar{1}, \bar{3}\}$, $\langle \bar{5} \rangle = \{\bar{1}, \bar{5}\}$, то $\langle \bar{3} \rangle \cap \langle \bar{5} \rangle = \{\bar{1}\}$, следовательно, однозначность представления имеет место. Осталось показать существование

представления для любого элемента $\mathbb{Z}_8^*$. Имеем, где первый сомножитель из $\langle \bar{3} \rangle$, а второй сомножитель из $\langle \bar{5} \rangle$:

$$\begin{aligned}\bar{1} &= \bar{1} \cdot \bar{1}, \\ \bar{3} &= \bar{3} \cdot \bar{1}, \\ \bar{5} &= \bar{1} \cdot \bar{5}, \\ \bar{7} &= \bar{3} \cdot \bar{5}.\end{aligned}$$

Приведём ещё несколько примеров групп $\mathbb{Z}_n^*$.

Пример IV.3.26. Пусть $n = 7$, тогда $\mathbb{Z}_7^* = \{\bar{1}, \bar{2}, \bar{3}, \bar{4}, \bar{5}, \bar{6}\}$.

Группа $\mathbb{Z}_7^*$ – циклическая группа порядка 6, так как $\mathbb{Z}_7^* = \langle \bar{3} \rangle$. Действительно, имеем

$$\bar{3}^2 = \bar{2}, \bar{3}^3 = \bar{6}, \bar{3}^4 = \bar{4}, \bar{3}^5 = \bar{5}, \bar{3}^6 = \bar{1}.$$

Кроме образующего $\bar{3}$, группа $\mathbb{Z}_7^*$ имеет и другие образующие, а именно $\bar{5}$. Это можно проверить непосредственным вычислением степеней $\bar{5}$ или применением теоремы: если все элементы циклической группы $\langle \bar{3} \rangle$ порядка 6 записаны в виде $\bar{3}^k$, как это сделано выше, то образующими будут те элементы, для которых $\text{НОД}(k, 6) = 1$. Поскольку только для $k = 5$ имеем $\text{НОД}(5, 6) = 1$, то кроме $\bar{3}$ есть ещё один образующий, а именно $\bar{3}^5 = \bar{5}$.

Заметим, что группа $\mathbb{Z}_7^*$ имеет подгруппы $G_1 = \langle \bar{2} \rangle = \langle \bar{4} \rangle$ порядка 3 и $G_2 = \langle \bar{6} \rangle$ порядка 2, причём группа $\mathbb{Z}_7^*$ разложима в прямое произведение своих подгрупп G_1 и G_2 :

$$\mathbb{Z}_7^* = G_1 \times G_2.$$

Этот факт можно установить непосредственной проверкой аналогично предыдущему примеру. Частично упорядоченное по включению множество подгрупп группы $\mathbb{Z}_7^*$ изображается диаграммой рис. 1, где $A \rightarrow B$ означает: «подгруппа A группы содержится в подгруппе B группы G ».

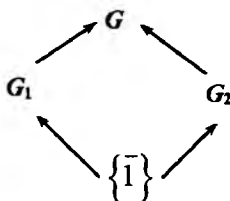


Рис. 1

Пример IV.3.27. Пусть $n = 10$, тогда $\mathbb{Z}_{10}^* = \{\bar{1}, \bar{3}, \bar{7}, \bar{9}\}$.

Имеем $\bar{3}^2 = \bar{9}$, $\bar{3}^3 = \bar{7}$, $\bar{3}^4 = \bar{1}$, следовательно, $\mathbb{Z}_{10}^*$ является циклической группой порядка 4 и порождается $\bar{3}$, т.е. $\mathbb{Z}_{10}^* = \langle \bar{3} \rangle$. Так как $\bar{7} = \bar{3}^3$ и $\text{НОД}(3, 4) = 1$, то $\bar{7}$ – ещё один образующий группы $\mathbb{Z}_{10}^*$:

$$G = \mathbb{Z}_{10}^* = \langle \bar{3} \rangle = \langle \bar{7} \rangle.$$

Единственная собственная подгруппа группы $\mathbb{Z}_{10}^*$ – это $H = \langle \bar{9} \rangle$ – циклическая подгруппа порядка 2, так как $\bar{9}^2 = \bar{1}$. Множество подгрупп в $\mathbb{Z}_{10}^*$ линейно упорядочено по включению, как видно из диаграммы на рис. 2:



Рис. 2

Часто в обозначениях элементов кольца $\mathbb{Z}_n$ черточки сверху элементов убирают и используют так называемую приведённую систему вычетов:

$$\mathbb{Z}_n = \{0, 1, 2, \dots, n-1\},$$

т.е. полный набор различных остатков от деления на n с операциями

$$a \oplus b = (a + b) \bmod n,$$

$$a \otimes b = (ab) \bmod n,$$

где $x \bmod n$ есть остаток от деления x на n . Поскольку каждый приведённый вычет x есть представитель класса вычетов $\bar{x}$, то из свойств классов эквивалентности следует, что если для приведённых вычетов x и y имеет место $x \equiv y \pmod{n}$, т.е. x и y принадлежат одному классу вычетов, то $x = y$.

Некоторый опыт работы с кольцом $\mathbb{Z}_n$ как приведённой системой вычетов позволяет отказаться от записи кольцевых операций в виде $\oplus$ и $\otimes$ в пользу использования обычных форм записи сложения и умножения в виде «+» и « $\cdot$ ». В дальнейшем, если не будет оговорено противное, кольцо $\mathbb{Z}_n$ будет записываться как приведённая система вычетов с соответствующими операциями.

Пример IV.3.28. Кольцо $\mathbb{Z}_6$ записывается в виде $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$.

Определение IV.3.29. Будем говорить, что кольцо K изоморфно кольцу L , если существует биекция φ из кольца K на кольцо L , которая сохраняет кольцевые операции, т.е. для любых $x, y \in K$ имеют место равенства

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

$$\varphi(xy) = \varphi(x)\varphi(y).$$

Если кольца K и L изоморфны, то будем обозначать это так:

$$K \cong L.$$

Лекция 4. ПОЛЯ И МНОГОЧЛЕНЫ

Определение IV.4.1. Кольцо K , в котором все ненулевые элементы образуют абелеву группу относительно умножения, называется полем.

Из определения следует, что группа обратимых элементов K^* поля K состоит из всех ненулевых элементов поля K и называется мультипликативной группой поля K . Заметим, что кольцо K является полем тогда и только тогда, когда K есть коммутативное кольцо с единицей, в котором каждый ненулевой элемент обратим.

Пример IV.4.2. Непосредственно проверяется, что $\mathbb{Q}$, $\mathbb{R}$ и $\mathbb{C}$ – поля, а $\mathbb{Z}$ полем не является.

Предложение IV.4.3. Пусть K – поле, тогда в K нет делителей нуля.

Доказательство. Пусть $0 \neq a$ и $0 \neq b$ – элементы поля K , такие, что $a \cdot b = 0$. Тогда

$$a^{-1}(ab) = a^{-1} \cdot 0 = 0$$

и, следовательно,

$$a^{-1}(ab) = (a^{-1}a)b = 1 \cdot b = b = 0,$$

что противоречит $b \neq 0$.

Определение IV.4.4. Подмножество L поля K называется подполем, если L само есть поле относительно операций поля K .

Теорема IV.4.5. Подмножество L поля K является подполем тогда и только тогда, когда для любых $a, b \in L$ имеем

$$a - b \in L, a \cdot b \in L$$

и для любого $0 \neq a \in L$

$$a^{-1} \in L.$$

Доказательство. Утверждение теоремы проверяется непосредственно.

Следующая теорема даёт ответ на вопрос, когда кольцо классов вычетов $\mathbb{Z}_n$ является полем.

Теорема IV.4.6. Кольцо $\mathbb{Z}_n$ является полем тогда и только тогда, когда $n = p$ есть простое число.

Необходимость. Пусть $\mathbb{Z}_n$ – поле и $0 \neq \bar{a} \in \mathbb{Z}_n$, тогда $\bar{a}$ обратим и по теореме IV.3.23 имеем $\text{НОД}(a, n) = 1$, где $a \in \{1, \dots, n-1\}$. Но тогда число n не имеет делителей, отличных от 1 и n , т.е. $n = p$ – простое число.

Достаточность. Пусть $n = p$ есть простое число, тогда любое из чисел $\{1, 2, \dots, p-1\}$ взаимно просто с p и по теореме IV.3.23 тогда любой $0 \neq \bar{a} \in \mathbb{Z}_p$ обратим. Поскольку $\mathbb{Z}_p$ – коммутативное кольцо с единицей, отсюда следует, что $\mathbb{Z}_p$ есть поле.

Следующий факт имеет большое значение в криптографии.

Теорема IV.4.7. Мультипликативная группа $\mathbb{Z}_p^*$ поля $\mathbb{Z}_p$, где p – простое число, является циклической группой порядка $p-1$.

Пример IV.4.8. Пусть $p = 11$, тогда $\mathbb{Z}_{11}^*$ есть циклическая группа порядка 10, причём $\mathbb{Z}_{11}^* = \langle \bar{2} \rangle$.

Образующий α циклической группы $\mathbb{Z}_p^*$ называется примитивным элементом поля $\mathbb{Z}_p$.

Для различных криптографических приложений часто приходится решать следующую задачу: имеется поле $\mathbb{Z}_p$, где p – большое случайное простое число (например, битовой длины 256 бит); найти циклическую подгруппу H группы $G = \mathbb{Z}_p^*$ большого порядка q , скажем, битовой длины 160 бит. Решение этой задачи существует, если $p-1$ имеет делитель q подходящего размера. Тогда находим примитивный элемент α поля $\mathbb{Z}_p$ и вычисляем $\beta = \alpha^{(p-1)/q}$. Легко видеть, что $H = \langle \beta \rangle$ имеет порядок q ,

причём H является единственной подгруппой группы $\mathbb{Z}_p^*$ порядка q .

Определение IV.4.9. Будем говорить, что поля F и K изоморфны, если существует биекция из одного из этих полей в другое, которая сохраняет операции поля.

Многочлены над полями имеют многочисленные приложения, но вначале дадим понятие многочлена над коммутативным кольцом.

Определение IV.4.10. Пусть здесь и далее K – коммутативное кольцо с единицей. Назовём многочленом над кольцом K произвольную конечную последовательность элементов $(a_0, a_1, \dots, a_n)$ кольца K . Если

$p = (a_0, a_1, \dots, a_n)$ – многочлен и $a_n \neq 0$, то n называется степенью многочлена p ($n = \deg(p)$), a_n называется старшим коэффициентом многочлена p , причём если $a_n = 1$, то многочлен называется нормированным. Многочлен нулевой степени $p = (a_0)$ называется константой. Условимся для многочлена $p = (a_0, a_1, \dots, a_n)$ степени n возможным дописывать произвольное число нулей после старшего коэффициента $a_0 \neq 0$. Например,

$$(1, 0, 2, 3) = (1, 0, 2, 3, 0, 0, 0) = (1, 0, 2, 3, 0).$$

Введём символ X следующим образом: X есть многочлен первой степени вида

$$X = (0, 1) = (0, 1, 0, \dots, 0).$$

Множество всех многочленов над кольцом K обозначается $K[X]$. Зададим в этом множестве операции сложения и умножения следующим образом: пусть $p = (a_0, a_1, \dots, a_n)$, $q = (b_0, b_1, \dots, b_m)$, тогда, если для определённости $n \leq m$, дописываем нужное число нулей для многочлена меньшей степени $p = (a_0, a_1, \dots, a_n, 0, \dots, 0)$ и зададим

$$p + q = (a_0 + b_0, a_1 + b_1, \dots, a_n + b_n, b_{n+1}, \dots, b_m),$$

$$p \cdot q = (c_0, c_1, \dots, c_k, \dots, c_{m+n}),$$

где $c_k = \sum_{i=0}^k a_i b_{k-i}$, $k = 0, 1, \dots, m+n$.

Относительно введённых выше операций $K[X]$ является коммутативным кольцом с единицей.

Покажем теперь, что многочлены можно представлять в привычном виде. Найдём вначале степени многочлена X :

$$X^2 = X \cdot X = (0, 1) \cdot (0, 1) = (0 \cdot 0 = 0, 0 \cdot 1 + 1 \cdot 0 = 0, 1 \cdot 1 = 1),$$

$$X^3 = X^2 \cdot X = (0, 0, 1) \cdot (0, 1) = (0, 0, 0, 1),$$

.....

$$X^i = \left(\underbrace{0, 0, \dots, 0}_i, 1 \right) \text{ для } i = 1, 2, \dots$$

Тогда, полагая $X^0 = (1, 0, \dots, 0)$ и $a_i = (a_i, 0, \dots, 0)$, можно представить произвольный многочлен $p = (a_0, a_1, \dots, a_n)$ в виде $p = a_0 + a_1 \cdot X + a_2 X^2 + \dots + a_n X^n$. Однако принципиальное отличие от

обычного понимания многочлена состоит в том, что здесь X не является переменной, а p соответственно не является функцией. Однако для каждого многочлена можно определить соответствующую функцию следующим образом.

Определение IV.4.11. Пусть

$$p = (a_0, \dots, a_n) = a_0 + a_1 \cdot X + a_2 X^2 + \dots + a_n X^n -$$

многочлен из $K[X]$, тогда определим функцию $\bar{p}: K \rightarrow K$ по правилу: $\bar{p}: r \mapsto a_0 + a_1 r + \dots + a_n r^n$ для $r \in K$, которую назовём полиномиальной функцией, индуцированной многочленом p .

Как будет видно ниже, в некоторых случаях, например для $\mathbb{Q}$, $\mathbb{R}$ и $\mathbb{C}$, возможно отождествление многочлена и соответствующей полиномиальной функции, однако существуют примеры, когда этого нельзя делать.

Пример IV.4.12. Пусть $K = \mathbb{Z}_n$ и рассмотрим в $\mathbb{Z}_n[X]$ многочлен

$$p = X \cdot (X + \bar{1}) \dots (X + (\overline{n-1})).$$

Многочлен p , очевидно, отличен от нуля и имеет степень n . Однако $\bar{p}: \mathbb{Z}_n \rightarrow \mathbb{Z}_n$ является нулевой функцией, поскольку для любого $r \in \{\bar{0}, \bar{1}, \dots, \overline{n-1}\}$ имеем

$$\bar{p}(r) = r \cdot (r + \bar{1}) \dots (r + (\overline{n-1})) = \bar{0}.$$

Понятие полиномиальной функции, индуцированной многочленом, полезно хотя бы потому, что даёт возможность ввести понятие корня многочлена.

Определение IV.4.13. Назовём $r \in K$ корнем многочлена $p \in K[X]$, если r является корнем соответствующей полиномиальной функции, т.е. $\bar{p}(r) = 0$.

Пример IV.4.14. Рассмотрим многочлен $p = \bar{1} + X + X^2$ в кольце $\mathbb{Z}_3[X]$, тогда $\bar{1}$ является корнем полиномиальной функции $\bar{p}$, так как

$$\bar{p}(\bar{1}) = \bar{1} + \bar{1} + \bar{1}^2 = \bar{0}.$$

Заметим, что $\bar{1}$ является единственным корнем многочлена p , так как

$$\overline{p}(\overline{0}) = \overline{1} + \overline{0} + \overline{0}^2 = \overline{1},$$

$$\overline{p}(\overline{2}) = \overline{1} + \overline{2} + \overline{2}^2 = \overline{1}.$$

Замечание. Из определения умножения в кольце $K[X]$ следует, что если в кольце K нет делителей нуля, то их нет и в $K[X]$, а обратимыми элементами в $K[X]$ являются многочлены вида

$$p = (a_0, 0, \dots, 0),$$

где $a_0 \in K$ – обратимый элемент.

Теперь и до конца лекции (если не оговорено противное) будем считать, что K – поле, и рассматривать кольцо $K[X]$ для этого случая.

Рассмотрим деление многочленов с остатком в кольце $K[X]$, где K – поле.

Алгоритм деления с остатком в $K[X]$, где K – поле.

Вход: $u, v \in K[X]$, где

$$u = u_0 + u_1X + \dots + u_mX^m,$$

$$v = v_0 + v_1X + \dots + v_nX^n,$$

причём $v_n \neq 0$, $m \geq n \geq 0$, $u_i, v_j \in K$.

Выход: частное $q = q_0 + q_1X + \dots + q_{m-n}X^{m-n}$,

остаток $r = r_0 + r_1X + \dots + r_{n-1}X^{n-1}$,

такие, что $u = q \cdot v + r$, $\deg(r) < \deg(v)$.

Д1 (Итерация по k). Выполнить Д2 для

$$k = m - n, m - n - 1, \dots, 0;$$

после этого работа алгоритма заканчивается с ответом

$$u_{n-1} = r_{n-1}, \dots, u_0 = r_0.$$

Д2 (Цикл деления). Вычислить $q_k \leftarrow u_{n+k} \cdot v_n^{-1}$, затем вычислить

$$u_j \leftarrow u_j - q_k \cdot v_{j-k} \text{ для } j = (n-1) + k, \dots, k.$$

Пример IV.4.15. Пусть $K = \mathbb{Q}$,

$$u = 1 - X + 3X^2 + 2X^3, \quad m = 3,$$

$$v = -2 - 5X + X^2, \quad n = 2,$$

тогда

$$u_0 = 1, u_1 = -1, u_2 = 3, u_3 = 2,$$

$$v_0 = -2, v_1 = -5, v_2 = 1.$$

Параметр k принимает следующие значения:

$$k: m - n = 1, m - n - 1 = 0;$$

$$k = 1 \begin{cases} q_1 = u_3 \cdot v_2^{-1} = 2, \\ j = 2: \dot{u}_2 = u_2 - q_1 \cdot v_1 = 3 - 2 \cdot (-5) = 13, \\ j = 1: \dot{u}_1 = u_1 - q_1 \cdot v_0 = -1 - 2 \cdot (-2) = 3; \end{cases}$$

$$k = 0 \begin{cases} q_0 = \dot{u}_2 \cdot v_2^{-1} = 13, \\ j = 1: \dot{u}_1 = \dot{u}_1 - q_0 v_1 = 3 - 13(-5) = 68, \\ j = 0: \dot{u}_0 = u_0 - q_0 v_0 = 1 - 13(-2) = 27. \end{cases}$$

Таким образом,

$$q = q_0 + q_1 X = 13 + 2X,$$

$$r = r_0 + r_1 X = \dot{u}_0 + \dot{u}_1 X = 27 + 68X.$$

Проверка: $u = q \cdot v + r$;

$$(13 + 2X)(-2 - 5X + X^2) + (27 + 68X) = 1 - X + 3X^2 + 2X^3.$$

Определение IV.4.16. Будем говорить, что многочлен $u \in K[X]$ делится нацело на ненулевой многочлен $0 \neq v \in K[X]$, если остаток $r \in K[X]$ от деления u на v равен нулю, т.е. $u = q \cdot v$ (обозначение $v \mid u$).

Одним из важных следствий алгоритма деления с остатком в $K[X]$, где K – поле, является следующая теорема.

Теорема IV.4.17 (теорема Безу). Пусть $u \in K[X]$, K – поле. Элемент $\alpha \in K$ является корнем многочлена u тогда и только тогда, когда

$$u = (x - \alpha) \cdot q$$

для некоторого $q \in K[X]$.

Доказательство. Вспомним, что корень многочлена – это корень соответствующей полиномиальной функции.

Необходимость. Пусть $\alpha \in K$ есть корень многочлена $u \in K[X]$. Разделим с остатком многочлен u на многочлен $(X - \alpha) \in K[X]$, тогда имеем

$$u = (X - \alpha) \cdot q + r.$$

Поскольку α – корень многочлена u , то для соответствующей полиномиальной функции $\bar{u}$ имеем

$$\bar{u}(\alpha) = 0.$$

С другой стороны, имеем

$$\bar{u}(\alpha) = (\alpha - \alpha) \cdot q + r = 0 \cdot q + r = r.$$

Следовательно, $r = 0$ и тогда

$$u = (X - \alpha) \cdot q.$$

Достаточность. Пусть

$$u = (X - \alpha) \cdot q,$$

тогда для соответствующей полиномиальной функции $\bar{u}$ имеем

$$\bar{u}(\alpha) = (\alpha - \alpha) \cdot q = 0 \cdot q = 0.$$

Это означает, что α есть корень многочлена u .

Следствие. Пусть $\alpha \in K$ – произвольный элемент и $f(x) \in K[X]$, где K – поле. Остаток от деления $f(x)$ на $x - \alpha$ равен $f(\alpha)$.

Доказательство. По алгоритму деления с остатком $f(x) = (x - \alpha) \cdot g(x) + r(x)$, где $\deg r(x) < \deg(x - \alpha)$. Но тогда $r(x)$ – многочлен нулевой степени, т.е. $r(x) = \beta \in K$. Найдём $r(x)$ так:

для $x = \alpha$ имеем $f(\alpha) = 0 \cdot g(\alpha) + r(\alpha)$, тогда $r(x) = f(\alpha)$.

Определение IV.4.18. Если $u, v \in K[X]$, K – поле, такие, что $u = q \cdot v$, причём $\deg(v) < \deg(u)$, то многочлен v называется собственным делителем многочлена u . Будем говорить, что многочлен u неприводим, если u не имеет собственных делителей, причём $\deg(u) \geq 1$.

Очевидно, что многочлены степени 1 неприводимы. Описание неприводимых многочленов над $\mathbb{R}$ и $\mathbb{C}$ приводится в следующей теореме.

Теорема IV.4.19.

а) В кольце $\mathbb{R}[X]$ неприводимые многочлены – это многочлены степени 1 и те многочлены степени 2 вида $a_0 + a_1X + a_2X^2$, где $a_1^2 - 4a_0a_2 < 0$.

б) В кольце $\mathbb{C}[X]$ неприводимые многочлены – это в точности многочлены степени 1.

Теорема IV.4.20. Пусть K – поле, тогда для любого ненулевого многочлена $f \in K[X]$ существует и единственно представление f в виде произведения неприводимых многочленов:

$$f = \alpha \cdot u_1 \cdot u_2 \cdot \dots \cdot u_k,$$

где $\alpha \in K$, $u_1, \dots, u_k$ – неприводимые многочлены из $K[X]$.

Имеет место следующая теорема.

Теорема IV.4.21. Пусть K – поле и $f, g \in K[X]$, причём $f \neq 0$ или $g \neq 0$. Тогда существует многочлен $d \in K[X]$ такой, что выполняются следующие условия:

а) d является общим делителем f и g , т.е. f и g нацело делятся на d (сокращённая запись $d | f$ и $d | g$);

б) для любого многочлена $d' \in K[X]$, если $d' | f$ и $d' | g$, то $d' | d$.

Доказательство. Применяя несколько раз алгоритм деления с остатком, получим следующую цепочку равенств:

$$f = q_1 g + r_1, \deg(r_1) < \deg(g),$$

$$g = q_2 r_1 + r_2, \deg(r_2) < \deg(r_1),$$

$$r_1 = q_3 r_2 + r_3, \deg(r_3) < \deg(r_2),$$

.....

$$r_{k-2} = q_k r_{k-1} + r_k, \deg(r_k) < \deg(r_{k-1}),$$

$$r_{k-1} = q_{k+1} r_k, r_{k+1} = 0.$$

Поскольку степени остатков образуют строго убывающую последовательность неотрицательных целых чисел, то через конечное число шагов обязательно получится нулевой остаток $r_{k+1} = 0$. Покажем, что последний ненулевой остаток r_k как раз и есть искомый многочлен d .

Проверим выполнение условия «а». Из последнего равенства имеем, что $r_k | r_{k-1}$, затем, поднимаясь вверх по цепочке равенств, получаем $r_k | r_{k-2}, \dots, r_k | g, r_k | f$. Итак, $r_k | f$ и $r_k | g$.

Проверим далее условие «б». Пусть $d' | f$ и $d' | g$, тогда, спускаясь вниз по цепочке равенств, получим:

$$d' | f, d' | g, d' | r_1, \dots, d' | r_{k-1}, d' | r_k.$$

Таким образом, многочлен $d = r_k$ удовлетворяет условиям «а» и «б» теоремы.

Определение IV.4.22. Нормированный многочлен $d \in K[X]$, K – поле, называется наибольшим общим делителем многочленов $f, g \in K[X]$, если d удовлетворяет условиям «а» и «б» теоремы IV.4.21. В этом случае d обозначается $d = \text{НОД}(f, g)$. Если $\text{НОД}(f, g) = 1$, то f и g называются взаимно простыми.

В доказательстве теоремы IV.4.21 содержится алгоритм нахождения наибольшего общего делителя, который называется алгоритмом Евклида. Следующая теорема получается как следствие теоремы Евклида.

Теорема IV.4.23 (расширенный алгоритм Евклида). В условиях теоремы IV.4.21 существуют многочлены $s, t \in K[X]$, такие, что выполняется равенство

$$s \cdot f + t \cdot g = d = \text{НОД}(f, g).$$

Доказательство. В цепочке равенств из доказательства теоремы IV.4.21 каждый из остатков r_i равен линейной комбинации двух предыдущих остатков (здесь коэффициенты линейных комбинаций – многочлены из $K[X]$). При этом r_1 линейно выражается через f и g :

$$r_1 = f - q_1 g,$$

r_2 линейно выражается через r_1 и g , а именно

$$r_2 = g - q_2 r_1.$$

Но тогда r_2 также линейно выражается через f и g :

$$\begin{aligned} r_2 &= g - q_2 r_1 = g - q_2 (f - q_1 g) = \\ &= g - q_2 f + q_2 q_1 g = (1 + q_2 q_1) g - q_2 f. \end{aligned}$$

Продолжая таким образом, получим, что и все остальные остатки $r_3, r_4, \dots, r_{k-1}, r_k$ линейно выражаются через f и g с коэффициентами из $K[X]$. Таким образом, $r_k = s \cdot f + t \cdot g$ для некоторых многочленов $s, t \in K[X]$. Поскольку $r_k = \text{НОД}(f, g)$, то теорема доказана.

Пример IV.4.24. Пусть $f, g \in \mathbb{Q}[X]$, где $f = 2 - 3X + X^2$, $g = 1 + X$. Найдём $\text{НОД}(f, g)$ по расширенному алгоритму Евклида.

Сначала делим f на g с остатком, получаем

$$f = (X - 4)g + 6.$$

Затем делим с остатком g на 6:

$$g = 6 \cdot \left(\frac{1}{6} + \frac{1}{6} X \right).$$

Поскольку получен нулевой остаток, то НОД(f, g) равен с точностью до множителя из K последнему из ненулевых остатков, т.е. 6. Нормируя этот многочлен, получим НОД(f, g) = 1. Выразим теперь НОД(f, g) как линейную комбинацию f и g . Имеем

$6 = f - (X - 4)g$ и, деля обе части равенства на 6, получим

$$\frac{1}{6} \cdot f - \left(\frac{1}{6} X - \frac{2}{3} \right) g = 1,$$

т.е. $s = \frac{1}{6}$, $t = \frac{2}{3} - \frac{1}{6} X$, таковы, что

$$s \cdot f + t \cdot g = d = \text{НОД}(f, g).$$

Кольцо многочленных вычетов определяется аналогично кольцу вычетов. Пусть K — поле, $K[X]$ — кольцо многочленов от неизвестного X над полем K и $f \in K[X]$ — произвольный фиксированный многочлен. Для любых многочленов g и h из $K[X]$ зададим бинарное отношение ρ_f следующим образом:

$$g \rho_f h \Leftrightarrow f | (g - h).$$

Аналогично отношению сравнимости по модулю n для целых чисел отношение ρ_f является отношением эквивалентности, и получаем разбиение кольца $K[X]$ на классы эквивалентности по модулю f . Легко видеть, что $g \rho_f h$ тогда и только тогда, когда g и h имеют одинаковые остатки от деления на f . Будем обозначать $g \rho_f h$ как $g \equiv h \pmod{f}$ и называть g и h сравнимыми по модулю f , классы эквивалентности по модулю f будем называть многочленными классами вычетов по модулю f или просто многочленными вычетами.

На множестве многочленных вычетов, которое обозначим $K[X]/(f)$, зададим алгебраические операции сложения и умножения следующим образом:

$$\overline{g} + \overline{h} = \overline{g + h},$$

$$\overline{g} \cdot \overline{h} = \overline{gh}$$

для любых $g, h \in K[X]/(f)$.

Аналогично классам вычетов по модулю n , многочленные классы вычетов по модулю f образуют кольцо относительно введенных выше операций сложения и умножения классов многочленных вычетов. Если $n = \deg f$, то множество всевозможных остатков от деления многочленов из $K[X]$ на f есть множество всех многочленов из $K[X]$, степень которых строго меньше n . Аналогично кольцу приведенных вычетов по модулю n , которое отождествлялось с кольцом вычетов по модулю n , определим кольцо приведенных многочленных вычетов по модулю f следующим образом.

Условимся обозначать через $K[X]/(f)$ множество нормированных (старший коэффициент равен 1) многочленов $K[X]$ степени $< n$. Зададим на этом множестве две алгебраические операции:

$$g \oplus h = g + h,$$

$$g \odot h = (gh) \bmod f,$$

где $g, h \in K[X]/(f)$ и $g \bmod f$ означает остаток от деления g на f .

Непосредственно проверяется, что относительно этих операций $K[X]/(f)$ является кольцом, которое будем называть кольцом многочленных вычетов по модулю f .

Замечание. Операции $\oplus$ и $\odot$ часто обозначают просто «+» и « $\cdot$ » (сложение и умножение).

Таким образом, в зависимости от контекста, обозначение $K[X]/(f)$ как кольца многочленных вычетов используется и для кольца классов многочленных вычетов по модулю f , и для кольца приведенных вычетов (остаток по модулю f).

Пусть F – поле, и предположим, что f есть неприводимый нормированный многочлен степени $k \geq 1$ из $F[x]$. Тогда кольцо многочленных вычетов $F[x]/(f)$ есть поле. Рассмотрим произвольный элемент $\overline{g} = g + (f)$ кольца $F[x]/(f)$. По алгоритму деления с остатком имеем $g = hf + r$, где $\deg(r) < \deg(f) = k$, и поскольку $hf \in (f)$, то

$g + (f) = r + (f)$ (поскольку $g - r = hf \in (f)$). Следовательно, любой элемент кольца $F[x]/(f)$ можно однозначно представить в виде

$$a_0 + a_1x + \dots + a_{k-1}x^{k-1} + (f), \quad (1)$$

где $a_i \in F$, $i = \overline{0, k-1}$.

Пример IV.4.25. Пусть $F = \mathbb{R}$ и $f = x^2 + 1$, тогда $\mathbb{R}[x]/(x^2 + 1)$ есть поле ($f = x^2 + 1$ неприводим над $\mathbb{R}$), и любой элемент $\mathbb{R}[x]/(x^2 + 1)$ однозначно представим в виде $r = r + (x^2 + 1) = \underbrace{a + bx}_{r} + (x^2 + 1)$, где $a, b \in \mathbb{R}$.

Заметим, что множество $\overline{F}$ смежных классов вида $\{a + (f) | a \in F\}$ образует подполе поля $F[x]/(f)$.

Теорема IV.4.26. Пусть F — поле, f — неприводимый нормированный многочлен, тогда поле $F[x]/(f)$ содержит подполе $\overline{F} = \{a + (f) | a \in F\}$, которое в дальнейшем будет отождествляться с полем F .

Имеет место следующая лемма (предполагается, что определение векторного пространства над произвольным полем F идентично определению векторного пространства над $\mathbb{R}$, если скаляры из F заменить на скаляры из $\mathbb{R}$).

Лемма IV.4.27. Пусть F — поле, F' — подполе. Тогда F может рассматриваться как векторное пространство над полем F' .

Доказательство получается непосредственной проверкой определения векторного пространства.

По лемме IV.4.27, если F — поле и f — неприводимый нормированный многочлен, то поле $F[x]/(f)$ является векторным пространством над полем $\overline{F} = F$. Заметим, что множество смежных классов

$$\{1 + (f) = [1], x + (f) = [x], \dots, x^{k-1} + (f) = [x^{k-1}]\}$$

линейно независимо. Действительно, из

$$\alpha_1 [1] + \alpha_2 [x] + \dots + \alpha_{k-1} [x^{k-1}] = [0]$$

следует $[\alpha_1 \cdot 1 + \alpha_2 \cdot x + \dots + \alpha_{k-1} x^{k-1}] = [0]$, и тогда

$$\alpha_1 \cdot 1 + \alpha_2 \cdot x + \dots + \alpha_{k-1} x^{k-1} \in [0] = [f],$$

что означает делимость многочлена $k-1$ -й степени $\alpha_1 \cdot 1 + \alpha_2 \cdot x + \dots + \alpha_{k-1} x^{k-1}$ на многочлен k -й степени f , что возможно, лишь если $\alpha_1 \cdot 1 + \alpha_2 \cdot x + \dots + \alpha_{k-1} x^{k-1} = 0$, что, в свою очередь, возможно лишь для $\alpha_1 = \alpha_2 = \dots = \alpha_{k-1} = 0$.

Более того, поскольку любой смежный класс поля $F[x]/(f)$ имеет вид $[g] = g + (f)$, где

$$g \bmod f = a_0 + a_1 x + \dots + a_{k-1} x^{k-1},$$

то $[g]$ однозначно представим в виде $[g] = a_0 [1] + a_1 [x] + \dots + a_{k-1} [x^{k-1}]$. Следовательно, $[1], [x], \dots, [x^{k-1}]$ есть базис F -векторного пространства $F[x]/(f)$, и тогда $\dim_F(F[x]/(f)) = k$. Таким образом, если обозначить $X = x + (f) = [x]$, то любой элемент $F[x]/(f)$ можно однозначно представить в виде

$$a_0 + a_1 X + a_2 X^2 + \dots + a_{k-1} X^{k-1}, \quad (2)$$

где $a_i \in F$, $i = \overline{0, k-1}$.

Вспоминая задание операций в кольце $F[x]/(f)$, мы получаем, что элементы поля $F[x]/(f)$ однозначно представимы как многочлены от X с коэффициентами из поля F , а операции сложения и умножения в этом поле являются операциями сложения и умножения многочленов от X с последующим взятием остатка от деления на многочлен $f = f(X)$.

Пусть $\overline{f}$ есть полиномиальная функция, индуцированная неприводимым нормированным многочленом f степени k над $F[x]/(f)$. Тогда имеем

$$\begin{aligned} \overline{f}(X) &= \overline{f}(x + (f)) = f(x + (f)) = \\ &= a_0 + a_1(x + (f)) + \dots + a_k(x + (f))^k = \\ &= a_0 + a_1 x + \dots + a_k x^k + (f) = f + (f) = (f) = [0], \end{aligned}$$

т.е. X есть корень функции $\bar{f}$ над полем $F[x]/(f)$.

Условимся в дальнейшем элементы поля $F[x]/(f)$ рассматривать как формальные многочлены от X вида (2) (см. с. 181), где X – новый символ со свойством $\bar{f}(X) = 0$.

Пример IV.4.28. Пусть $F = \mathbb{R}$, $f = x^2 + 1$ – неприводимый нормированный многочлен степени 2 над $\mathbb{R}$. Тогда $F[x]/(f)$ есть поле с элементами вида $a + bX$, $a, b \in \mathbb{R}$, где $X = x + (f)$, причём $X^2 + [1] = [0]$ в этом поле. Это означает, что в поле $F[x]/(f)$ разрешимо уравнение $z^2 + 1 = 0$ и решением его является X .

Определение IV.4.29. Пусть F – поле и F' – поле, такое, что $F \subset F'$. Тогда F' называется расширением поля F .

Теорема IV.4.30 (теорема Кронекера). Пусть F – поле, $g \in F[x]$ – произвольный многочлен степени ≥ 1 . Тогда существует такое расширение K поля F , что многочлен g имеет корень в K (точнее, полиномиальная функция $\bar{g}$, индуцированная многочленом g , имеет корень в K).

Доказательство. Если $\bar{g}$ имеет корень в K , то доказывать нечего. Пусть это не так, тогда g имеет делитель f , степень которого ≥ 2 и который является неприводимым нормированным многочленом над полем F (в частности, может быть $g = f$). Полагаем $K = F[x]/(f)$ и рассмотрим $\bar{g}$ и $\bar{f}$ как полиномиальные функции над K . Обозначим $x + (f) = \alpha$, тогда имеем

$$\bar{f}(\alpha) = f + (f) = (f) = [0]$$

в поле K , т.е. α есть корень $\bar{f}$ над K . Поскольку f – делитель g , то α является корнем и для полиномиальной функции $\bar{g}$ над K .

Если $F = \mathbb{R}$ и $K = F[x]/(x^2 + 1)$, то из примера IV.4.28 и теоремы IV.4.30 видно, что поле K является расширением поля F , а полиномиальная функция, индуцированная многочленом $x^2 + 1$, не имеет корней в F , однако имеет корень $X = x + (f)$ в K .

Определение IV.4.31. Многочлен $f \in F[x]$, F – поле, называется вполне разложимым над расширением K поля F , если f можно представить как произведение линейных множителей из $K[x]$. При этом K называется полем разложения многочлена f над полем F , если f есть вполне

разложимый многочлен над K , но не является таковым ни над каким собственным подполем в K , содержащим F .

Теорема IV.4.32. Пусть F – поле, $g \in F[x]$ – многочлен степени ≥ 1 . Тогда существует расширение K поля F , которое является полем разложения многочлена g .

Доказательство. По теореме IV.4.30 существует поле $K_1 \supseteq F$ такое, что многочлен g имеет корень α в K_1 . Тогда по теореме Безу многочлен g имеет линейный делитель $x - \alpha$ в $K_1[x]$. Если многочлен g не вполне разложим над K_1 , то повторяем эту процедуру и строим цепочку расширенных полей

$$K_1 \subseteq K_2 \subseteq \dots \subseteq K,$$

до тех пор, пока не получим требуемое расширение.

Определение IV.4.33. Пусть K есть расширение поля F и A есть произвольное множество элементов поля K . Обозначим через $F(A)$ пересечение всех подполей, которые содержат как F , так и A . Поле $F(A)$ называется расширением поля F , полученным присоединением к F элементов из A . Говорят также, что множество A порождает поле $F(A)$ над полем F . Если A есть одноэлементное множество, $A = \{a\}$, то в случае $a \notin F$ $F(\{a\})$ обозначается $F(a)$ и называется простым расширением поля F , а элемент a называется порождающим элементом $F(a)$ над F .

Пример IV.4.34. Пусть $F = \mathbb{Q}$, $a = \sqrt{2} \notin \mathbb{Q}$. Тогда $\mathbb{Q}(\sqrt{2})$ есть простое расширение поля $\mathbb{Q}$ посредством $\sqrt{2}$.

Определение IV.4.35. Пусть K есть расширение поля F . Элемент $\alpha \in K$ называется алгебраическим над F , если существует $0 \neq g \in F[x]$ такой, что $\bar{g}(\alpha) = 0$, в противном случае α называется трансцендентным элементом над F . Расширение K поля F называется алгебраическим, если любой элемент из K является алгебраическим над F . Если K содержит хотя бы один трансцендентный элемент над F , то K называется трансцендентным расширением. Размерность расширения K как векторного пространства над полем F называется степенью K над F и обозначается $[K:F]$.

Пример IV.4.36. Любой элемент α из поля F является корнем многочлена $x - \alpha \in F[x]$. Поэтому любой элемент F – алгебраический над F . Число $\sqrt{2} \in \mathbb{R}$, $\sqrt{2} \notin \mathbb{Q}$ является алгебраическим над $\mathbb{Q}$, поскольку $\sqrt{2}$

есть корень многочлена $x^2 - 2 \in \mathbb{Q}[x]$. Число π является трансцендентным над $\mathbb{Q}$, поэтому $\mathbb{R}$ есть трансцендентное расширение $\mathbb{Q}$.

Как устроены простые трансцендентные расширения, показывает следующая теорема.

Теорема IV.4.37. Пусть K есть расширение поля F и $\alpha \in K$ – трансцендентный элемент над F . Тогда простое расширение F посредством α , а именно $F(\alpha)$, изоморфно полю $F[x]$ алгебраических дробей вида $\frac{f(x)}{g(x)}$, где $0 \neq g(x), f(x) \in F[x]$ с обычными операциями сложения и умножения алгебраических дробей.

Определение IV.4.38. Пусть K есть расширение поля F , $F \subset K$ и $\alpha \in K \setminus F$. Обозначим

$$F[\alpha] = \left\{ a_0 + a_1\alpha + \dots + a_n\alpha^n \mid a_0, \dots, a_n \in F, n = 0, 1, \dots \right\}.$$

Непосредственно проверяется, что $F[\alpha]$ из определения IV.4.38 является подкольцом в кольце $K[x]$. Следующая теорема выявляет строение простых расширений посредством алгебраического элемента.

Теорема IV.4.39. Пусть K есть расширение поля F и $\alpha \in K \setminus F$ – алгебраический элемент над F . Тогда:

- а) $F(\alpha) = F[\alpha] \cong F[x]/(f)$, где f есть однозначно определённый нормированный многочлен над F , имеющий α своим корнем;
- б) α есть корень $g(x) \in F[x]$ тогда и только тогда, когда многочлен g делится на многочлен f из пункта «а»;
- в) если многочлен f из пункта «а» имеет степень n , то элементы $1, \alpha, \dots, \alpha^{n-1}$ образуют базис векторного пространства $F(\alpha)$ над F ; в частности, $[F(\alpha) : F] = n$ и любой элемент $F(\alpha)$ можно однозначно представить в виде

$$a_0 + a_1\alpha + \dots + a_{n-1}\alpha^{n-1}, a_i \in F.$$

Определение IV.4.40. Пусть K есть расширение поля F и $\alpha \in K$ – алгебраический элемент над F . Единственный неприводимый нормированный многочлен $f \in F[x]$, имеющий α своим корнем, называется минимальным многочленом элемента α . Степень минимального многочлена элемента α называется степенью α над полем F и обозначается $\deg_F(\alpha)$.

Пример IV.4.41. Минимальным многочленом числа $\sqrt{2}$ над $\mathbb{Q}$ есть многочлен $x^2 - 2$, причём $\mathbb{Q}(\sqrt{2}) = \mathbb{Q}[\sqrt{2}]$ и $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$, а числа $1, \sqrt{2}$ образуют базис двумерного векторного пространства $\mathbb{Q}(\sqrt{2})$ над $\mathbb{Q}$, так что любой элемент $\mathbb{Q}(\sqrt{2})$ однозначно записывается в виде $a + b\sqrt{2}$, $a, b \in \mathbb{Q}$.

Замечание. Элемент α из расширения K поля F является алгебраическим тогда и только тогда, когда α является корнем неприводимого многочлена $f \in F[x]$. Это утверждение следует из того, что α есть корень многочлена $g \in F[x]$ тогда и только тогда, когда α – корень неприводимого делителя f многочлена g .

Определение IV.4.42. Поле K называется конечным расширением поля F , если K есть расширение поля F и $[K : F] < \infty$. В противном случае K называется бесконечным расширением поля F . Степенью элемента $\alpha \in K$ над полем F называется $[F(\alpha) : F]$.

Следующие утверждения проверяются непосредственно.

1) Если $\alpha_1, \dots, \alpha_n$ есть базис в K над F , то

$$F(\alpha_1, \dots, \alpha_n) = \left\{ c_1 \alpha_1 + c_2 \alpha_2 + \dots + c_n \alpha_n \mid c_i \in F, i = 1, \dots, n \right\}.$$

2) Если K есть конечное расширение поля F степени n , то существует такое подмножество $\{\alpha_1, \dots, \alpha_n\}$ в K , что $K = F(\alpha_1, \dots, \alpha_n)$.

Имеет место следующая теорема.

Теорема IV.4.43. Любое конечное расширение поля F является алгебраическим.

Доказательство. Если $n = [K : F]$, то любые $n + 1$ элементов из K линейно зависимы. Пусть $\alpha \in K$, тогда $1, \alpha, \alpha^2, \dots, \alpha^n$ есть линейно зависимое множество над F , т.е. существуют $c_i \in F$, не все равные нулю одновременно, такие, что

$$c_0 + c_1 \alpha + \dots + c_n \alpha^n = 0.$$

Но тогда α есть корень многочлена $c_0 + c_1 x + \dots + c_n x^n \in F[x]$, т.е. α – алгебраический элемент над F .

Теорема IV.4.44. Пусть L есть конечное расширение поля K , а K есть конечное расширение поля F . Тогда

$$[L : F] = [L : K] \cdot [K : F].$$

Доказательство. Пусть $\{\alpha_i | i \in J\}$ – базис L над K , а $\{\beta_j | j \in I\}$ – базис K над F . Непосредственно проверяется, что множество $\{\alpha_i, \beta_j | i \in J, j \in I\}$ образует базис L над F из $|J| \cdot |I|$ элементов.

Следствие IV.4.45. Пусть K есть конечное расширение поля F . Тогда справедливы следующие утверждения:

- а) степень над F любого элемента из K делит $[K : F]$;
- б) $\alpha \in K$ порождает поле K над F (т.е. $K = F(\alpha)$) тогда и только тогда, когда $\deg_F(\alpha) = [K : F]$;
- в) если $[K : F] = 2^m$ и f есть неприводимый многочлен над F степени 3, то f есть неприводимый многочлен над K .

Заметим, что из IV.4.45 «в» вытекает неразрешимость классической древнегреческой задачи удвоения куба посредством циркуля и линейки.

Определение IV.4.46. Неприводимый многочлен $f \in F[x]$, F – поле, называется сепарабельным, если f не имеет кратных корней в своём поле разложения.

Произвольный многочлен $f \in F[x]$ называется сепарабельным, если каждый его неприводимый делитель является сепарабельным. Алгебраический элемент α называется сепарабельным над F , если его минимальный многочлен над F сепарабельный.

Теорема IV.4.47. Пусть F – поле, $\alpha_1, \dots, \alpha_n$ – алгебраические сепарабельные элементы над F . Тогда существует такой $\alpha \in F(x_1, \dots, x_n)$, что $F(\alpha_1, \dots, \alpha_n) = F(\alpha)$.

Зададим отображение дифференцирования

$$D: F[x] \rightarrow F[x]$$

по правилу

$$D: f = a_0 + a_1x + \dots + a_nx^n \mapsto f' = a_1 + a_2x + \dots + na_nx^{n-1}.$$

Легко видеть, что $D(af + bg) = aD(f) + bD(g)$, т.е. D является F -линейным отображением, причём $\text{Ker } D = F$. Заметим, что многочлен $f' = D(f)$ называется (формальной) производной многочлена f .

Следующие две теоремы проверяются непосредственно.

Теорема IV.4.48. Многочлен $f \in F[x]$, F – поле, является сепарабельным тогда и только тогда, когда $\text{НОД}(f, f') = 1$.

Теорема IV.4.49. Неприводимый многочлен f над полем F – сепарабельный тогда и только тогда, когда $f' \neq 0$.

Определение IV.4.50. Поле F называется алгебраически замкнутым, если любой многочлен $f \in F[x]$ такой, что $\deg f \geq 1$, является вполне разложимым над F . Поле $\overline{F}$ называется алгебраическим замыканием поля F , если F является алгебраически замкнутым и алгебраическим расширением поля F . (Это равносильно тому, что F не имеет алгебраических расширений, которые строго содержат $\overline{F}$.)

Из определения IV.4.50 следует, что если F есть алгебраически замкнутое поле, то любой многочлен $f \in F[x]$ степени $n \geq 1$ можно представить в виде произведения n линейных множителей.

Теорема IV.4.51. Существует поле F , которое является расширением поля $\mathbb{R}$.

Доказательство. Сначала построим поле F . Рассмотрим двумерное арифметическое векторное пространство $\mathbb{R}^2$ над полем $\mathbb{R}$ действительных чисел, т.е. множество всех упорядоченных пар вида (a, b) , где $a, b \in \mathbb{R}$ с операциями покомпонентного сложения векторов и умножения на скаляр. Зададим в $\mathbb{R}^2$ алгебраическую операцию умножения по следующему правилу:

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc).$$

Коммутативность умножения очевидна.

Ассоциативность умножения следует из равенств:

$$\begin{aligned} (a, b)[(c, d) \cdot (e, f)] &= (a, b) \cdot (ce - df, cf + de) = \\ &= (ace - adf - bcf - bde, acf + ade + bce - bdf), \end{aligned}$$

$$\begin{aligned} [(a, b)(c, d)](e, f) &= (ac - bd, ad + bc)(e, f) = \\ &= (ace - bde - adf - bcf, acf - bdf + ade + bce). \end{aligned}$$

Дистрибутивность умножения относительно сложения вытекает из равенств:

$$\begin{aligned} [(a, b) + (c, d)](e, f) &= (a + c, b + d)(e, f) = \\ &= (ae + ce - bf - df, af + cf + be + de), \end{aligned}$$

$$\begin{aligned} (a, b)(ef) + (c, d)(e, f) &= \\ &= (ae - bf, af + be) + (ce - df, cf + de) = \\ &= (ae - bf + ce - df, af + be + cf + de). \end{aligned}$$

Так как векторное пространство $\mathbb{R}^2$ есть абелева группа по сложению, то в $\mathbb{R}^2$ есть нуль — это $(0, 0)$ и для каждого $(a, b) \in \mathbb{R}^2$ есть противоположный элемент (т.е. обратный к (a, b) относительно сложения), а именно $(-a, -b)$.

Покажем теперь, что любой ненулевой элемент $\mathbb{R}^2$ имеет обратный относительно умножения. Пусть $(0, 0) \neq (a, b) \in \mathbb{R}^2$ (т.е. $a^2 + b^2 \neq 0$), и рассмотрим уравнение для нахождения обратного элемента:

$$(a, b)(x, y) = (1, 0).$$

Здесь $(1, 0) \in \mathbb{R}^2$ — единица $\mathbb{R}^2$, т.е. нейтральный элемент относительно умножения. Действительно, для любого $(c, d) \in \mathbb{R}^2$ имеем

$$(c, d) \cdot (1, 0) = (c, d).$$

Имеем

$$(a, b)(x, y) = (ax - by, ay + bx) = (1, 0),$$

тогда получаем систему уравнений:

$$\begin{cases} ax - by = 1, \\ ay + bx = 0. \end{cases}$$

Решая эту систему уравнений, получим

$$x = \frac{a}{a^2 + b^2}, \quad y = \frac{-b}{a^2 + b^2}.$$

Тогда

$$(a, b)^{-1} = \left(\frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2} \right).$$

Действительно,

$$\begin{aligned} (a, b) \left(\frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2} \right) &= \\ &= \left(\frac{a^2}{a^2 + b^2} - \frac{b^2}{a^2 + b^2}, -\frac{ab}{a^2 + b^2} + \frac{ba}{a^2 + b^2} \right) = (1, 0). \end{aligned}$$

Итак, $\mathbb{R}^2$ является коммутативным кольцом с единицей, в котором любой ненулевой элемент имеет обратный по умножению. Следовательно, $\mathbb{R}^2$ является полем, которое называется полем комплексных чи-

сел и обозначается $\mathbb{C}$. Заметим, что в поле $\mathbb{C}$ есть подполе $\mathbb{C}' = \{(a, 0) | a \in \mathbb{R}\}$. Действительно, по критерию подполя (теорема IV.4.5) имеем

для любых $(a, 0), (b, 0) \in \mathbb{C}'$

$$(a, 0) - (b, 0) = (a - b, 0) \in \mathbb{C}',$$

$$(a, 0) \cdot (b, 0) = (ab, 0) \in \mathbb{C}',$$

если $(0, 0) \neq (a, 0)$, т.е. $a \neq 0$, то

$$(a, 0)^{-1} = \left(\frac{1}{a}, 0\right) \in \mathbb{C}'.$$

Далее, поле $\mathbb{C}'$ изоморфно $\mathbb{R}$ (обозначается $\mathbb{C}' \cong \mathbb{R}$). Действительно, рассмотрим отображение $\phi: \mathbb{R} \rightarrow \mathbb{C}'$, заданное по правилу

$$\phi: a \mapsto (a, 0), \quad a \in \mathbb{R}.$$

Очевидно, что ϕ есть биекция, причём

$$\phi(a + b) = (a + b, 0) = (a, 0) + (b, 0) = \phi(a) + \phi(b),$$

$$\phi(ab) = (ab, 0) = (a, 0) \cdot (b, 0) = \phi(a) \cdot \phi(b),$$

$$\phi(1) = (1, 0).$$

Следовательно, ϕ есть изоморфизм полей.

Отождествим поле $\mathbb{C}'$ с полем $\mathbb{R}$ и будем считать, что поле $\mathbb{C}$ содержит подполе $\mathbb{R} (= \mathbb{C}')$. Тогда $\mathbb{C}$ является расширением поля $\mathbb{R}$. Покажем, что $\mathbb{C} = \mathbb{R}(\alpha)$ для некоторого $\alpha \in \mathbb{C}$.

Сначала найдём корень α многочлена $f = x^2 + 1$ в $\mathbb{C}$. Пусть $\alpha = (0, 1) \in \mathbb{C}$, тогда

$$\alpha^2 + (1, 0) = (-1, 0) + (1, 0) = (0, 0),$$

следовательно, $\alpha = (0, 1)$ является корнем многочлена $x^2 + 1$. Этот корень принято обозначать $(0, 1) = i$, причём i есть алгебраический элемент над $\mathbb{R}$. Так как $x^2 + 1$ является неприводимым нормированным многочленом из $\mathbb{R}[x]$, то $x^2 + 1$ есть минимальный многочлен элемента $i \in \mathbb{C}$ над $\mathbb{R}$. Тогда $(1, 0), i$ есть базис $\mathbb{C}$ как двумерного векторного пространства над $\mathbb{R}$. Следовательно, любой элемент из $\mathbb{C}$ может быть

однозначно представлен в виде $a + bi$, $a, b \in \mathbb{R}$. Но тогда по теореме IV.4.39 имеем

$$\mathbb{C} = \mathbb{R}[i] = \mathbb{R}(i) \cong \mathbb{R}[x]/(x^2 + 1).$$

Наконец, $\mathbb{C}$ есть поле разложения для $x^2 + 1 \in \mathbb{R}[x]$, так как

$$x^2 + 1 = (x - i)(x + i).$$

Имеет место следующее утверждение.

Теорема IV.4.52. Поле $\mathbb{C}$ есть алгебраически замкнутое поле.

Вопросы для самоконтроля

1. Верно ли, что для любой алгебраической структуры, если e_1 и e_2 являются ее нейтральными элементами, то $e_1 = e_2$?
2. Верно ли, что для любого моноида и любого элемента A , если b и c являются обратными элементами для a , то $b = c$?
3. Верно ли, что если квазигруппа Q является полугруппой, то Q является группой?
4. Верно ли, что если полугруппа P является квазигруппой, то P является группой?
5. Верно ли, что в любой группе $(G; \cdot)$ однозначно разрешимы уравнения $ax = b$ и $ya = b$?
6. Сформулируйте критерий подгруппы для аддитивной группы $(G; +)$.
7. Может ли группа порядка 4 быть подгруппой группы порядка 30?
8. Сформулируйте теорему Лагранжа и ее следствие.
9. Верно ли, что порядок циклической группы равен порядку ее образующего?
10. Верно ли, что в циклической группе простого порядка любой элемент, отличный от нейтрального, является ее образующим?
11. Верно ли, что четные подстановки чисел 1, 2, 3 (т.е. подстановки с четным суммарным числом инверсий в обеих строках) образуют подгруппу в симметрической группе S_3 ?
12. Верно ли, что в кольце с единицей обратимые элементы образуют группу относительно умножения?
13. Может ли обратимый элемент кольца с единицей быть делителем нуля?
14. Обязательно ли необратимый элемент кольца с единицей является делителем нуля?
15. Будет ли обязательно неделитель нуля в кольце с единицей обратимым элементом?
16. Может ли в кольце с единицей сумма делителей нуля быть обратимым элементом?

17. Может ли в кольце с единицей сумма обратимых элементов быть делителем нуля или нулем?
18. Для каких чисел n кольцо вычетов по модулю n является полем?
19. Обязательно ли будет полем коммутативное кольцо с единицей без делителей нуля?
20. Будет ли полем кольцо функции $\mathbb{R}^{\mathbb{R}}$?

ЛИТЕРАТУРА

1. *Ильин В.А., Позняк Э.Г.* Линейная алгебра. – М.: Физматлит, 2005.
2. *Кострикин А.И.* Введение в алгебру. – М.: Наука, 1987.
3. *Головина Л.И.* Линейная алгебра и некоторые ее приложения. – М.: Наука, 1984.

Оглавление

Предисловие.....	3
МОДУЛЬ I	4
Лекция 1. Множества, отображения и подстановки.....	4
Лекция 2. Матрицы и определители.....	21
Лекция 3. Ранг матрицы	35
Лекция 4. Обратная матрица и матричные уравнения.....	43
Вопросы для самоконтроля.....	51
МОДУЛЬ II	53
Лекция 1. Понятие векторного пространства и простейшие свойства	53
Лекция 2. Линейная зависимость векторов	59
Лекция 3. Базисы	66
Лекция 4. Подпространства.....	80
Вопросы для самоконтроля.....	88
МОДУЛЬ III	90
Лекция 1. Совместность систем линейных уравнений.....	90
Лекция 2. Однородные системы линейных уравнений	99
Лекция 3. Метод Гаусса решения систем линейных уравнений	107
Лекция 4. Треугольное разложение матриц (матричная интерпретация метода Гаусса)	115
Вопросы для самоконтроля.....	122
МОДУЛЬ IV	124
Лекция 1. Алгебраические операции и группы.....	124
Лекция 2. Различные классы групп	142
Лекция 3. Кольца.....	155
Лекция 4. Поля и многочлены	169
Вопросы для самоконтроля.....	191
ЛИТЕРАТУРА	193

Семен Константинович Росошек

Алгебра и геометрия

Часть 1

Учебное пособие

Редактор *Н. И. Шидловская*

Верстка *Л. В. Пермякова*

К-ОКП ОК-005-93, код продукции 954240

Изд. лиц. ИД № 04000 от 12.02.01. Подписано к печати 28.04.08.
Формат 60×84¹/₁₆. Бумага офсетная. Печать офсетная. Гарнитура «Times».
Усл. печ. л. 11,39. Уч.-изд. л. 12,76. Тираж 200 экз.

ООО «Издательство научно-технической литературы»
634050, г. Томск, пл. Ново-Соборная, 1, тел. (382-2) 53-33-35

Отпечатано в типографии ЗАО «М-Принт», г. Томск, ул. Пролетарская, 38/1