МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ

В. Г. Бардаков

ЛЕКЦИИ ПО АЛГЕБРЕ Ю. И. МЕРЗЛЯКОВА

Учебное пособие

Новосибирск
2012

УДК 512(075.8)
ББК В14я73-2
Б 247

Бардаков В. Г. Лекции по алгебре Ю. И. Мерзлякова: Учеб. пособие / Новосиб. гос. ун-т. Новосибирск, 2012. 311 с.

ISBN 978-5-4437-0061-8

Предлагаемый курс лекций читался автором на первом курсе механико-математического факультета Новосибирского государственного университета в 2008–2010 гг. Этот курс был разработан Ю. И. Мерзляковым, который читал лекции на первом курсе механико-математического факультета около 30 лет. В курсе изучаются основные алгебраические системы: группы, кольца, поля, векторные пространства, алгебры. Подробно изучаются группы подстановок, кольца матриц и многочленов, поле комплексных чисел, конечномерные векторные пространства. Также разбираются темы, входящие в стандартные курсы линейной алгебры: определители, системы линейных уравнений, линейные преобразования векторных пространств, жорданова форма матрицы, евклидовы и унитарные пространства, квадратичные формы. Кроме того, изучаются основы векторной алгебры. Пособие предназначено для студентов, изучающих курсы “Высшая алгебра” и “Линейная алгебра”.

Рецензенты:
д-р физ.-мат. наук, доц. А. П. Пожидаев,
канд. физ.-мат. наук, доц. М. В. Нещадим

ISBN 978-5-4437-0061-8

Оглавление

Предисловие

Ю. И. Мерзляков читал лекции на первом курсе механико-математического факультета с 1965 по 1995 годы. Студенты считали его одним из лучших лекторов механико-математического факультета. К сожалению, Юрий Иванович так и не издал свои лекции. На вопрос, почему он это не делает, Юрий Иванович отвечал: “Боюсь, что студенты перестанут посещать мои лекции”. Наверное, поэтому мне не удалось найти конспектов этих лекций или хотя бы черновиков, написанных самим Юрием Ивановичем. Пришлось восстанавливать весь курс по своим студенческим конспектам, а также по конспектам других слушателей.

Те, кому посчастливилось учиться у Юрия Ивановича, хорошо помнят его безупречные лекции, его аккуратный почерк и виртуозное владение доской. Отдавая себе отчет в том, что невозможно передать стиль Юрия Ивановича в учебном пособии, я, тем не менее, взял на себя смелость изложить этот курс, чтобы сохранить его для будущих поколений студентов.

Изложенный в настоящем учебном пособии курс Юрий Иванович успевал прочесть в течение двух семестров. Главы 1–3 излагались в первом семестре, а главы 4–9 – во втором. По объему это достаточно большой курс, и мне до сих пор не понятно, как Юрий Иванович при своей неторопливой манере, записывая все подробно на доске, успевал изложить весь материал. Во всяком случае, мне этого не удавалось и приходилось опускать некоторые разделы.

Благодарю всех, кто принимал участие в подготовке этих лекций: М. В. Нещадима и А. П. Пожидаева, внимательно прочитавших рукопись и внесших ряд полезных замечаний и предложений; О. В. Брюханова, Е. А. Захрямина, А. А. Коробова, М. В. Нещадима, Ю. В. Сосновского, беседы с которыми способствовали появлению настоящих лекций; О. В. Брюханова, разработавшего дизайн обложки и оказавшего помощь в подготовке оригинал-макета, Н. Л. Абашееву и Н. Б. Аюпову, помогавших освоить премудрости LaTeXа. Также благодарю студентов, прослушавших этот курс и внесших в него ряд полезных замечаний и предложений.

Обозначения

Введем обозначения, которые будем использовать на протяжении нашего курса. Символом

$$A = \{a_1, a_2, \ldots\}$$

обозначается множество A, состоящее из элементов $a_1, a_2, \ldots$ Запись $a \in A$ означает, что элемент a принадлежит множеству A; запись $a \notin A$ означает, что a не принадлежит множеству A. Если B является подмножеством множества A, то символически это обозначается так: $B \subseteq A$. Пустое множество будем обозначать символом $\emptyset$. Если $A = \{a_1, a_2, \ldots, a_n\}$ – конечное множество, то символом $|A| = n$ обозначается число элементов множества A.

Для числовых множеств будем использовать следующие обозначения:

$$\mathbb{N} = \{1, 2, \ldots\}$$

– *множество натуральных чисел* ;

$$\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$$

– *множество целых чисел* ;

$$\mathbb{Q} = \left\{ \frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N} \right\}$$

– *множество рациональных чисел* ; символом $\mathbb{R}$ будем обозначать множество вещественных чисел, которое можно представлять как множество точек на вещественной оси;

$$\mathbb{R}_+ = \{r \in \mathbb{R} \mid r > 0\}$$

– множество положительных вещественных чисел;

$$\mathbb{R}_{\geq 0} = \{r \in \mathbb{R} \mid r \geq 0\}$$

– множество неотрицательных вещественных чисел;

$$\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}, i^2 = -1\}$$

– *множество комплекснных чисел* .

Если $A_1, A_2, \ldots, A_n$ – непустые множества, то их *декартовым произведением* $A_1 \times A_2 \times \ldots \times A_n$ называется множество упорядоченных n-ок:

$$A_1 \times A_2 \times \ldots \times A_n = \{(a_1, a_2, \ldots, a_n) \mid a_i \in A_i\}.$$

В частности, если $A_1 = A_2 = \ldots = A_n = A$, то декартово произведение $A_1 \times A_2 \times \ldots \times A_n$ называется n-й *декартовой степенью* множества A и обозначается A^n.

Отображение множества A в множество B будем обозначать либо

$$\varphi : A \longrightarrow B, \text{ либо } A \xrightarrow{\varphi} B.$$

Если $a \in A$, то образ элемента a при отображении φ обозначается либо $\varphi(a)$, либо $a\varphi$.

Введение

Истоки алгебры зародились в цивилизациях Древнего Египта и Древней Греции. Именно там стали изучать действия над целыми и рациональными числами. В Древней Греции были сформулированы знаменитые задачи на построение при помощи циркуля и линейки: задача о трисекции угла, задача об удвоении куба и др. Напомним, что задача о трисекции угла состоит в том, чтобы разбить угол на три равных угла; задача об удвоении куба – в том, чтобы по заданному кубу построить куб, объем которого в два раза больше объема исходного куба.

Мы проследим развитие алгебры на примере решения уравнений. Рассмотрим следующее уравнение:

$$a\,x = b. \tag{1}$$

Здесь a и b – некоторые известные числа, а x – неизвестное. Это линейное алгебраическое уравнение от одной неизвестной. Что значит "решить уравнение"? Это значит, найти все его решения или доказать, что решений нет. При этом надо указывать множество, в котором ищется решение, так как может оказаться, что в одном множестве решений не существует, но если его вложить в некоторое большое множество, то в нем могут существовать решения. Под решением уравнения (1) мы понимаем такое число x^0, которое при подстановке в уравнение вместо неизвестной приводит к верному равенству.

При изучении любого уравнения (или системы уравнений) нас интересуют следующие два вопроса: имеет ли данное уравнение решение и, если ответ утвердительный, как найти все множество решений? Из школьного курса алгебры известно, что уравнение (1) разрешимо тогда и только тогда, когда либо a отлично от нуля, либо $a = b = 0$. В первом случае решение единственно и определяется равенством $x = a^{-1}\,b$, а во втором случае решением является любое число.

Более сложным, по сравнению с (1), является уравнение

$$a\,x^2 + b\,x + c = 0, \quad a \neq 0, \tag{2}$$

где опять a, b, c – заданные числа, а x – неизвестное. Это так называемое *квадратное уравнение* от одной неизвестной. Такие уравнения умели решать еще в IX в. на Востоке. В это же время возник и сам

термин “алгебра”. Решения уравнения (2) определяются по формуле

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4\,a\,c}}{2\,a}. \tag{3}$$

Теперь мы можем пойти дальше и определить алгебраическое уравнение степени n от одной неизвестной:

$$a_0\,x^n + a_1\,x^{n-1} + \ldots + a_{n-1}\,x + a_n = 0, \quad a_0 \neq 0. \tag{4}$$

Что известно про решения таких уравнений? При $n = 3$ известна *формула Кардано*, похожая на формулу (3), которая позволяет найти корни любого уравнения третьей степени. При $n = 4$ существует *метод Феррари*, позволяющий сводить решение уравнения 4-й степени к решению уравнения 3-й степени. Формулы для решения уравнений 3-й и 4-й степени были получены итальянскими математиками Кардано, Тарталья, Феррари в XVI в. После этого многие математики пытались найти аналогичные формулы для решения общего уравнения 5-й степени. Эти попытки продолжались до тех пор, пока в XIX в. А. Руффини (1765–1822) и Н. Абель (1802–1829) не доказали, что общее уравнение (4) при $n \geq 5$ неразрешимо в радикалах, т. е. не существует формул, выражающих решение через коэффициенты при помощи основных алгебраических операций: сложения, вычитания, умножения, деления, возведения в степень и извлечения корня. Подчеркнем, что речь идет именно об уравнении общего вида, так как легко указать конкретные уравнения сколь угодно высокой степени, разрешимые в радикалах. Например, уравнение

$$x^{100} - 3\,x^{50} + 2 = 0$$

сотой степени легко сводится к квадратному уравнению.

Французский математик Эварист Галуа (1811–1832), занимаясь условиями разрешимости уравнения в радикалах, создал теорию, которая в настоящее время называется *теорией Галуа*. Эту теорию можно считать началом современной алгебры. Э. Галуа впервые ввел такие понятия, как *группа*, *поле*, *автоморфизм*. Помимо критерия разрешимости уравнения в радикалах, теория Галуа позволяет доказать неразрешимость задачи о трисекции угла, задачи об удвоении куба и ряда других задач, сформулированных еще в Древней Греции.

Интересна судьба Э. Галуа. По нашим меркам у него не было даже высшего образования. При поступлении в Политехническую школу

Галуа провалился на экзамене по математике Он запустил тряпку для стирания с доски в голову экзаменатора, посчитав его вопросы слишком тривиальными. В 1829 г. Галуа поступил в Нормальную школу, из которой был отчислен за свои политические убеждения. Затем сидел в тюрьме и в 1832 г. убит на дуэли. За свою жизнь он написал несколько работ, которые представил во французскую Академию наук. К сожалению, некоторые из них были потеряны, другие же не получили признания современников. После этого имя Эвариста Галуа надолго было предано забвению. Все его математические работы попали к Огюсту Шевалье, но тот не смог найти никого, кто согласился бы их издать. Только в 1846 г. Ж. Лиувиль впервые опубликовал их в основанном им математическом журнале.

Математические работы Галуа составляют шестьдесят небольших страниц. Никогда еще труды столь малого объема не доставляли автору такой широкой известности. Через несколько десятков лет после смерти Галуа немецкий математик Давид Гильберт назвал теорию Галуа "установлением определенного остова понятий". Сам Галуа так писал о цели своих исследований: "Подчинить вычисления своей воле, сгруппировать математические операции, научиться их классифицировать по степени трудности, а не по внешним признакам – вот задачи математиков будущего так, как я их понимаю, вот путь, по которому я хочу пойти".

Мы разобрали одно из возможных обобщений уравнения (1), получив при этом уравнение n-й степени от одной неизвестной. Возможно и обобщение в другом направлении. Сохраним условие линейности, но будем рассматривать уравнения, зависящие от нескольких неизвестных. Придем к *системе линейных уравнений.* Систему m линейных уравнений от n неизвестных можно записать в виде

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

где $a_{i,j}, b_i$ – заданные числа, а x_j – неизвестные. При исследовании системы нас будут интересовать следующие вопросы: разрешима ли система? как описать все множество решений? *Решением* мы называем упорядоченную n-ку чисел $x^0 = (x_1^0, x_2^0, \ldots, x_n^0)$, которые при подстановке в систему дают верные равенства. Системы такого вида хорошо изучены, и мы научимся отвечать на оба поставленных

вопроса.

Если ввести следующие таблицы:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots\dots\dots\dots & & \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

которые называются *матрицей системы*, *столбцом неизвестных* и *столбцом свободных членов*, то нашу систему можно записать в виде

$$A\,X = B.$$

Сравним эту запись с уравнением (1). Видим, что по форме они очень похожи. Более того, иногда решение системы можно находить по формуле $X = A^{-1}\,B$.

Наконец, если, рассматривая уравнение (1), мы откажемся от условия линейности и допустим наличие нескольких неизвестных, то придем к полиномиальному уравнению или, более общо, к системе полиномиальных уравнений. Наиболее известным таким уравнением является уравнение

$$x^n + y^n = z^n.$$

где x, y, z – неизвестные целые числа. При $n = 2$ все такие решения описаны. Они образуют пифагоровы тройки. В частности,

$$3^2 + 4^2 = 5^2.$$

Оказывается, что при $n > 2$ это уравнение уже не имеет целочисленных решений. Эта теорема называется Великой, или Последней теоремой Ферма и была сформулирована Пьером Ферма в 1637 г. на полях книги “Арифметики” Диофанта. Ферма записал ее с припиской, что найденное им остроумное доказательство этой теоремы слишком длинно, чтобы его можно было поместить на полях книги: “Наоборот, невозможно разложить куб на два куба, биквадрат на два биквадрата и вообще никакую степень, большую квадрата, на две степени с тем же показателем. Я нашел этому поистине чудесное доказательство, но поля книги слишком узки для него”. С тех пор было предпринято много попыток найти элементарное доказательство этой теоремы, которые не увенчались успехом. Полное доказательство (совсем неэлементарное) было найдено в 1995 г. английским математиком Эндрю Уайлсом.

Решения систем полиномиальных уравнений изучаются в курсе алгебраической геометрии, и мы лишь слегка коснемся этой темы, доказав знаменитую теорему Гильберта, которая утверждает, что всякая система полиномиальных уравнений от конечного числа неизвестных равносильна некоторой конечной подсистеме.

Глава 1

ОСНОВНЫЕ ПОНЯТИЯ АЛГЕБРЫ

§ 1. Алгебраическая система. Изоморфизм

1.1. Алгебраическая операция, алгебраическая система. Множества и отображения на них — вот два основных объекта, к изучению которых сводится любая математическая теория. Пусть задано некоторое непустое множество A и функция $f : A^n \longrightarrow A$, аргументы которой пробегают множество A, и она принимает при этом значения из A. В этом случае мы говорим, что на множестве A определена n-арная *алгебраическая операция.* В частности, если $n = 1$, то операция называется *унарной,* если $n = 2$, то — *бинарной,* если $n = 3$, то — *тернарной* и т. д.

В школьном курсе вы уже встречались с такими операциями, как сложение и умножение. При этом для их обозначения используется не функциональная запись $f(x, y)$, а более привычная: $x + y$ и $x \cdot y$. В дальнейшем мы тоже будем использовать подобные обозначения для бинарных операций.

Как следует из определения, всякая алгебраическая операция определена на некотором множестве. Поэтому можно дать такое

О п р е д е л е н и е. *Алгебраической системой* называется непустое множество A с определенными на нем алгебраическими операциями:

$$\mathfrak{A} = \langle A;\ f_i\,(i \in I)\rangle,$$

где I – некоторое множество индексов, конечное или бесконечное.

Чтобы проиллюстрировать это понятие, приведем

П р и м е р ы алгебраических систем:

1) $\mathfrak{A}_1 = \langle$ {действительные числа}; взятие среднего арифметического, умножение на $10\rangle = \langle \mathbb{R}; \frac{a+b}{2}, 10a\rangle$;

2) $\mathfrak{A}_2 = \langle$ {положительные действительные числа}; взятие среднего геометрического, возведение в 10-ю степень$\rangle = \langle \mathbb{R}_+; \sqrt{ab}, a^{10}\rangle$;

3) $\mathfrak{A}_3 = \langle$ {точки на плоскости}; взятие центра тяжести треугольника с заданными вершинами$\rangle$.

Как видно уже из этих примеров, алгебраических систем существует достаточно много и изучать все системы довольно проблематично. Определяемое ниже понятие изоморфизма позволяет сократить число изучаемых систем.

1.2. Изоморфизм алгебраических систем. Заданы две алгебраические системы

$$\mathfrak{A} = \langle A;\ f_i\,(i \in I)\rangle, \quad \mathfrak{A}' = \langle A';\ f_i'\,(i \in I)\rangle$$

с одинаковыми наборами алгебраических операций (т. е. арность операции f_i равна арности операции f_i' для всех $i \in I$). Пусть установлено взаимно однозначное отображение

$$\varphi : A \longrightarrow A',$$

множества A *на* множество A', которое сохраняет операции, т. е. φ – такое отображение, для которого выполняются следующие свойства:

1) φ *однозначно*, т. е. одному элементу из A соответствует один элемент из A';

2) φ *унивалентно*, т. е. два разных элемента из A переходят в два разных элемента из A';

3) φ отображение *на*, т. е. для всякого $a' \in A'$ существует $a \in A$ такой, что $a\varphi = a'$;

4) φ *сохраняет операции*, т. е. для всех индексов $i \in I$ и всех наборов $a_1, \ldots, a_{n_i}$ элементов из A справедливо равенство

$$f_i(a_1, \ldots, a_{n_i})\varphi = f_i'(a_1\varphi, \ldots, a_{n_i}\varphi),$$

где n_i – арность операции f_i. В этом случае φ называется *изоморфным отображением*, или *изоморфизмом* системы $\mathfrak{A}$ на систему $\mathfrak{A}'$. При этом системы $\mathfrak{A}$ и $\mathfrak{A}'$ называются *изоморфными*, что символически записывается так: $\mathfrak{A} \simeq \mathfrak{A}'$.

Изоморфные системы с алгебраической точки зрения одинаковы, т. е. все свойства системы $\mathfrak{A}$ выполняются и в $\mathfrak{A}'$. Поэтому в алгебре

их не различают или рассматривают как точные копии друг друга – подобно тому, как мы не различаем экземпляры одного и того же романа, напечатанные разным шрифтом и на разной бумаге, если интересуемся только содержанием романа. Теперь мы можем дать определение нашего предмета. *Алгебра* – это наука, изучающая алгебраические системы с точностью до изоморфизма.

П р и м е р изоморфных систем. Покажем, что алгебраические системы $\mathfrak{A}_1$ и $\mathfrak{A}_2$ из приведенного выше примера изоморфны. Рассмотрим отображение

$$\varphi : \mathbb{R} \longrightarrow \mathbb{R}_+,$$

определенное правилом $a\varphi = 2^a$. Из школьного курса известно, что φ – взаимно однозначно и *на*. Чтобы проверить, что φ сохраняет операции, мы должны проверить следующие два равенства:

$$\left(\frac{a+b}{2}\right)\varphi = \sqrt{a\varphi \cdot b\varphi},$$

$$(10a)\varphi = (a\varphi)^{10},$$

справедливость которых следует из свойств показательной функции.

У п р а ж н е н и е. Для всякого натурального числа n на множестве целых чисел определим унарную операцию f_n правилом $f_n(x) = nx$. Докажите, что алгебраическая система $U_2 = \langle \mathbb{Z}; f_2 \rangle$ изоморфна алгебраической системе $U_3 = \langle \mathbb{Z}; f_3 \rangle$.

У п р а ж н е н и е. Для каких натуральных n и m имеет место изоморфизм $U_n \simeq U_m$?

Множество с определенной на нем одной унарной операцией называется *унаром*. Это простейшая (в смысле определения) алгебраическая система.

1.3. Подсистема. Если $\mathfrak{A} = \langle A;\ f_i\,(i \in I) \rangle$ – алгебраическая система, а B – непустое подмножество в A, то мы можем рассматривать операции f_i, $i \in I$, выбирая аргументы только из B. Такие операции называются *индуцированными* на B:

$$f_i|_B : B^{n_i} \longrightarrow A, \quad i \in I,$$

где n_i – арность операции f_i. Если при этом все операции $f_i|_B$, $i \in I$ являются алгебраическими на B, то получим алгебраическую систему $\mathfrak{B} = \langle B;\ f_i|_B\,(i \in I) \rangle$, которая называется *подсистемой* алгебраической системы $\mathfrak{A}$.

П р и м е р. Подсистемой алгебраической системы $\mathfrak{A} = \langle \mathbb{Z}; + \rangle$ является система $\mathfrak{B} = \langle 2\mathbb{Z}; + \rangle$, где $2\mathbb{Z} = \{2n \mid n \in \mathbb{Z}\}$ – множество четных чисел.

В этом параграфе мы ввели основные понятия: алгебраическая система, подсистема, изоморфизм. Далее будем изучать конкретные алгебраические системы – группы, кольца, поля, векторные пространства и в них рассматривать подгруппы, подкольца, подполя, подпространства, а также интерпретировать понятие изоморфизма.

§ 2. Группы

2.1. Аксиоматика. Некоторые алгебраические системы столь часто встречаются в различных областях математики, что их изучение стало предметом самостоятельных теорий. Именно таково понятие группы – предмет изучения теории групп. Группа – это множество с одной бинарной операцией, подчиняющейся некоторым аксиомам. В теории групп бинарную операцию называют обычно умножением и обозначают точкой (которую почти всегда опускают), реже используют $+$, $\odot$, $*$ и другие символы. Запись операции точкой называют еще *мультипликативной записью*, а запись плюсом – *аддитивной записью*.

О п р е д е л е н и е. *Группой* называется алгебраическая система $\langle G; \cdot \rangle$ с одной бинарной операцией $\cdot$, для которой выполнены следующие аксиомы.

1. Операция *ассоциативна*, т. е. $(ab)c = a(bc)$ для любых a, b, c из G.

2. Операция гарантирует единицу, т. е. в G существует такой элемент e – он называется *единицей*, что $ae = ea = a$ для любого a из G.

3. Операция гарантирует обратные элементы, т. е для любого a из G существует в G такой элемент x – он называется *обратным* к a, что $ax = xa = e$.

В дальнейшем, если понятно, о какой операции идет речь, будем обозначать группу $\langle G; \cdot \rangle$ символом G.

Установим некоторые следствия из определения.

С л е д с т в и е 1. *В группе существует единственный единичный элемент.*

Действительно, пусть e' и e'' – единичные элементы группы G.

Тогда $e'e'' = e''$, с другой стороны, $e'e'' = e'$. Следовательно, $e' = e''$.

С л е д с т в и е 2. *В каждой группе для каждого a существует единственный обратный элемент, который будем обозначать символом a^{-1}.*

Действительно, предположим, что x и y – два обратных элемента для элемента a. Рассмотрим равенство $(xa)y = x(ay)$, справедливое в силу аксиомы ассоциативности. По определению обратного элемента, левая часть этого равенства равна

$$(xa)y = ey = y,$$

аналогичным образом преобразуем правую часть:

$$x(ay) = xe = x.$$

Следовательно, $x = y$.

С л е д с т в и е 3. *Для любых элементов a, b группы G справедливо равенство $(ab)^{-1} = b^{-1}a^{-1}$.*

Действительно,

$$(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = aa^{-1} = e.$$

Благодаря ассоциативности в группах элемент $(ab)c = a(bc)$ можно записывать как abc, по той же причине однозначно определено произведение n элементов $a_1a_2 \dots a_n$ – без указания скобок, но в указанном порядке. Произведение n элементов, равных a, называется n-й *степенью* элемента a и обозначается a^n. Полагаем далее $a^0 = e$ и для $n < 0$ $a^n = (a^{-n})^{-1}$ или $a^n = (a^{-1})^{-n}$, что, как легко видеть, одно и то же.

У п р а ж н е н и е. Если a – произвольный элемент группы, m, n – целые числа, то $a^m a^n = a^{m+n}$, $(a^m)^n = a^{mn}$.

О п р е д е л е н и е. Группа $\langle G; \cdot \rangle$ называется *коммутативной* или *абелевой*, если операция $\cdot$ удовлетворяет следующей аксиоме коммутативности: для любых a, b из G справедливо равенство $ab = ba$.

Операцию в абелевой группе обычно обозначают символом $+$, единичный элемент называют *нулевым элементом* и обозначают символом 0, а обратный к элементу $a \in G$ называют *противоположным* и обозначают символом $-a$.

П р и м е р ы групп:

1) $\langle \mathbb{Z};\, + \rangle$ – множество целых чисел относительно операции сложения;

2) $\langle 2\mathbb{Z};\, + \rangle$ – множество четных чисел относительно операции сложения;

3) $\langle \mathbb{Q}^*;\, \cdot \rangle$ – множество ненулевых рациональных чисел относительно операции умножения;

4) $\langle$\{вращения квадрата\}; композиция вращений$\rangle$.

Нетрудно проверить, что все эти группы являются абелевыми.

2.2. Изоморфизм. В соответствии с общим определением изоморфизма алгебраических систем группы $\langle G;\, \cdot \rangle$ и $\langle G';\, \odot \rangle$ называются *изоморфными,* если существует взаимно однозначное отображение φ множества G *на* множество G', сохраняющее операцию умножения, т. е. выполнены следующие условия:

1) φ – однозначно;

2) φ – унивалентo;

3) φ – отображение *на*;

4) для любых элементов $x, y \in G$ справедливо равенство

$$(x \cdot y)\varphi = x\varphi \odot y\varphi.$$

Например, множество $\mathbb{R}_+$ положительных действительных чисел есть группа относительно обычного умножения чисел, множество $\mathbb{R}$ всех действительных чисел – группа относительно обычного сложения чисел, а отображение $\varphi : \mathbb{R}_+ \longrightarrow \mathbb{R}$, определяемое формулой $a\varphi = \log_2 a$, – изоморфизм $\mathbb{R}_+$ на $\mathbb{R}$.

2.3. Подгруппы. Подмножество группы G называется ее *подгруппой,* если оно замкнуто относительно операции, имеющейся в G, и само является группой относительно индуцированной операции. Если H – подгруппа группы G, то пишем $H \leq G$.

П р и м е р ы.

1. Группа $\langle \mathbb{Z};\, + \rangle$, ее подгруппа $\langle 2\mathbb{Z};\, + \rangle$.

2. Группа $\langle \mathbb{R}_+;\, \cdot \rangle$ не является подгруппой группы $\langle \mathbb{R};\, + \rangle$, так как имеет другую операцию.

Сформулируем необходимые и достаточные условия того, что некоторое подмножество является подгруппой.

Л е м м а 1. *Подмножество $H \subseteq G$ является подгруппой группы G в том и только том случае, когда выполнены следующие два условия*:

а) *из того, что $a, b \in H$, следует, что и $ab \in H$ (замкнутость относительно умножения)*;

б) *из того, что* $a \in H$, *следует, что и* $a^{-1} \in H$ (*замкнутость относительно взятия обратного*).

Д о к а з а т е л ь с т в о. Если H является подгруппой, то очевидно, что условия а и б выполняются.

Покажем теперь, что если выполнены условия а и б, то H является подгруппой. Для этого надо проверить аксиомы группы.

1. Аксиома ассоциативности $(ab)c = a(bc)$ выполнена для H, так как она выполняется для G.

3. Аксиома существования обратного элемента следует из условия б.

2. Пусть $a \in H$; по условию б $a^{-1} \in H$, найдем $aa^{-1} = e$, и по условию а $e \in H$.

Лемма доказана.

Условия леммы (замкнутость относительно умножения и взятия обратного) символически записывают так:

$$HH \subseteq H, \quad H^{-1} \subseteq H.$$

Установим следующее утверждение.

Л е м м а 2. *Пересечение любого семейства подгрупп некоторой группы является подгруппой.*

Д о к а з а т е л ь с т в о. Пусть в группе G задано семейство подгрупп H_α, $\alpha \in J$. Рассмотрим их пересечение $H = \bigcap_{\alpha \in J} H_\alpha$. Покажем, что H – подгруппа группы G. По предыдущей лемме достаточно доказать, что для H выполнены условия а и б. Выберем два элемента $a, b \in H$. Тогда $a, b \in H_\alpha$ для любого индекса $\alpha \in J$. Так как H_α – группа, то $ab \in H_\alpha$. Следовательно, $ab \in H$ и условие а установлено. Рассмотрим элемент $a \in H$. Тогда обратный элемент $a^{-1} \in H_\alpha$ для любого индекса $\alpha \in J$. Так как H_α является группой, то существует обратный элемент $a^{-1} \in H_\alpha$ для всех $\alpha \in J$. Следовательно, $a^{-1} \in H$ и условие б справедливо. Таким образом, пересечение H является подгруппой.

2.4. Порождающие множества. Если M – некоторое подмножество группы G, то пересечение (M) всех подгрупп, содержащих M, называется подгруппой, *порожденной множеством* M, а само M – *порождающим множеством* подгруппы (M):

$$(M) = \bigcap_{H \leq G, M \subseteq H} H.$$

В этом случае говорят, что элементы множества M являются *порождающими элементами* подгруппы (M). Подгруппу (M) иногда обозначают также через гр(M).

Т е о р е м а. *Если M – подмножество группы G, то*

$$(M) = \{m_1^{\varepsilon_1} m_2^{\varepsilon_2} \dots m_n^{\varepsilon_n} \mid m_i \in M, \varepsilon_i = \pm 1, n = 1, 2, \dots\}.$$

Д о к а з а т е л ь с т в о. Обозначим правую часть через H. Так как подгруппа (M) содержит все m_i из M, то справедливо включение $(M) \supseteq H$. С другой стороны, очевидно, что $HH \subseteq H$, $H^{-1} \subseteq H$, поэтому ввиду леммы 1 множество H – подгруппа, содержащая M. Отсюда $H \supseteq (M)$ и окончательно $H = (M)$.

Укажем порождающие множества некоторых встречавшихся ранее групп.

П р и м е р ы.

1) $\mathbb{Z} = (1)$, т. е. группа целых чисел по сложению порождается единицей;

2) $\mathbb{Q} = \left(\frac{1}{n} \mid n = 1, 2, \dots\right)$;

3) $\mathbb{Q}^* = (-1, 2, 3, 5, 7, 11, \dots)$.

Подгруппа, порожденная одним элементом a, называется *циклической.* По теореме она состоит из всевозможных степеней порождающего элемента:

$$(a) = \{a^n \mid n = 0, \pm 1, \pm 2, \dots\}.$$

Как видно из примера 1, группа $\mathbb{Z}$ является циклической. В следующем параграфе мы познакомимся с конечными циклическими группами. Все они устроены довольно просто. Если же мы рассмотрим группы, порожденные двумя элементами, то таких групп существует очень много. Более того, знаменитая теорема Хигмана утверждает, что всякая счетная группа является подгруппой некоторой двупорожденной группы.

§ 3. Кольца и поля

3.1. Определения и примеры. В школьном курсе вы уже встречались со множествами, на которых определены операции сложения и умножения. Таковыми, в частности, являются целые числа, рациональные, вещественные. Это и есть примеры колец и полей. В этом параграфе мы дадим формальные определения.

О п р е д е л е н и е. *Кольцом* называется алгебраическая система $\langle K; +, \cdot \rangle$ с двумя бинарными операциями ($+$ – сложение, $\cdot$ – умножение), для которых выполнены следующие аксиомы.

С1. Сложение *ассоциативно*, т. е. $(a+b)+c = a+(b+c)$ для любых a, b, c из K.

С2. Сложение *коммутативно*, т. е. $a + b = b + a$ для любых a, b из K.

С3. Существование нулевого элемента, т. е. в K существует такой элемент 0 – он называется *нулем*, что $a + 0 = a$ для любого a из K.

С4. Существование противоположного элемента, т. е. для любого a из K существует в K такой элемент x – он называется *противоположным* к a, что $a + x = 0$.

У1. Умножение *ассоциативно*, т. е. $a(bc) = (ab)c$ для любых a, b, c из K.

СУ1. Сложение и умножение удовлетворяют правой дистрибутивности, т. е. $(a + b)c = ac + bc$ для любых a, b, c из K.

СУ2. Сложение и умножение удовлетворяют левой дистрибутивности, т. е. $c(a + b) = ca + cb$ для любых a, b, c из K.

Иными словами, по сложению K является абелевой группой (выполнены аксиомы С1 – С4); по умножению K является полугруппой. *Полугруппой* называется алгебраическая система с одной бинарной ассоциативной операцией.

Так же, как и для групп, можно показать, что нулевой элемент единственный и для всякого элемента a из K существует единственный противоположный, который будем обозначать символом $-a$.

П р и м е р ы колец.

1. Множество целых (рациональных, вещественных) чисел с операциями сложения и умножения является кольцом.

2. Множество натуральных чисел с этими же операциями кольцом не является.

3. Рассмотрим множество функций

$$f : \mathbb{R} \longrightarrow \mathbb{R},$$

определенных для всех вещественных значений x и принимающих вещественные значения. Если определить операции сложения и умножения функций по правилу

$$(f + g)(x) = f(x) + g(x), \quad (f \cdot g)(x) = f(x) \cdot g(x),$$

то множество функций с этими операциями является кольцом.

Укажем некоторые следствия из аксиом кольца.

С л е д с т в и е. *Во всяком кольце произведение любого элемента на нулевой элемент есть нулевой элемент.*

Действительно,

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

Прибавляя к обеим частям этого равенства по элементу $-(a \cdot 0)$, получим, что $a \cdot 0 = 0$. Аналогично доказывается, что $0 \cdot a = 0$.

О п р е д е л е н и е. *Полем* называется алгебраическая система $\langle P; +, \cdot \rangle$ с двумя бинарными операциями ($+$ – сложение, $\cdot$ – умножение), для которых выполнены следующие аксиомы.

С1. Сложение *ассоциативно*, т. е. $(a+b)+c = a+(b+c)$ для любых a, b, c из P.

С2. Сложение *коммутативно*, т. е. $a + b = b + a$ для любых a, b из P.

С3. Существование нулевого элемента, т. е. в P существует такой элемент 0 – он называется *нулем*, что $a + 0 = a$ для любого a из P.

С4. Существование противоположного элемента, т. е. для любого a из P существует в P такой элемент x – он называется *противоположным* к a, что $a + x = 0$.

У1. Умножение *ассоциативно*, т. е. $a(bc) = (ab)c$ для любых a, b, c из P.

У2. Умножение *коммутативно*, т. е. $ab = ba$ для любых a, b из P.

У3. Существование единичного элемента, т. е. в P существует такой элемент $1 \neq 0$ – он называется *единицей*, что $1a = a$ для любого a из P.

У4. Существование обратного элемента, т. е. для любого a из P, отличного от нуля, существует в P такой элемент y – он называется *обратным* к a, что $ay = 1$.

СУ. Дистрибутивность, т. е. $(a + b)c = ac + bc$ для любых a, b, c из P.

Нетрудно показать, что в поле существует единственный нулевой элемент и единственный единичный элемент; противоположный и обратный к a определяются единственным образом и обозначаются соответственно $-a$ и a^{-1}.

Каждое поле является кольцом. С другой стороны, поле можно определить как кольцо, в котором выполнены аксиомы У2–У4. Если

положить $P^* = P \setminus \{0\}$, то $\langle P^*; \cdot\rangle$ – группа. Она называется *мультипликативной группой поля.* Из аксиомы У3 следует, что поле содержит, по крайней мере, два элемента. В силу коммутативности умножения в поле правая дистрибутивность равносильна левой дистрибутивности.

П р и м е р поля. Множество рациональных (вещественных) чисел с операциями сложения и умножения образует поле.

Существуют кольца, которые не являются полями. Например, кольцо целых чисел не является полем.

3.2. Кольца вычетов. Существуют кольца, состоящие из конечного числа элементов.

П р и м е р. Рассмотрим алгебраическую систему $\langle\{\overline{0}, \overline{1}, \overline{2}\}; +, \cdot\rangle$, где операции сложения и умножения определены правилами

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{0}$	$\overline{1}$

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{1}$

Нетрудно проверить, что полученная алгебраическая система является кольцом и полем.

П р и м е р. Алгебраическая система $\langle\{\overline{0}, \overline{1}, \overline{2}, \overline{3}\}; +, \cdot\rangle$, в которой операции сложения и умножения определены правилами

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{0}$	$\overline{2}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

является кольцом, но не полем (не выполняется аксиома У4).

Разобранные примеры являются частными случаями общего семейства *колец вычетов по модулю* n. Будем обозначать символом $\mathbb{Z}_n$ множество остатков от деления целых чисел на n. Для произвольного целого числа a символом $\overline{a}$ будем обозначать остаток от деления a на n. Определим на множестве $\mathbb{Z}_n$ операции сложения и умножения по правилу

$$\overline{a} + \overline{b} = \overline{a + b}, \quad \overline{a} \cdot \overline{b} = \overline{a \cdot b}.$$

У п р а ж н е н и е. Алгебраическая система

$$\mathbb{Z}_n = \langle\{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{n-1}\}; +, \cdot\rangle$$

всегда является кольцом, но полем является тогда и только тогда, когда n – простое число.

Нетрудно видеть, что если мы рассмотрим множество $\mathbb{Z}_n$ только относительно операции сложения, то оно является абелевой группой, которая порождается элементом $\overline{1}$, т. е. является циклической. Таким образом, мы знаем два примера циклических групп: $\mathbb{Z}$ и $\mathbb{Z}_n$. Оказывается, что с точностью до изоморфизма ими исчерпываются все циклические группы.

У п р а ж н е н и е. Любая бесконечная циклическая группа изоморфна группе $\mathbb{Z}$, любая циклическая группа конечного порядка n изоморфна группе $\mathbb{Z}_n$.

Порядком конечной группы мы называем число ее элементов.

У п р а ж н е н и е. Пусть $n \in \mathbb{N}$. Существует поле из n элементов тогда и только тогда, когда n – степень простого числа.

3.3. Делители нуля. Рассматривая кольцо вычетов $\mathbb{Z}_4$, видим, что $\overline{2}\cdot\overline{2} = \overline{0}$, т. е. произведение двух ненулевых элементов равно нулю.

О п р е д е л е н и е. Ненулевые элементы a и b кольца K такие, что $a \cdot b = 0$ называются *делителями нуля.*

Делители нуля существуют не только в конечных кольцах, но и в бесконечных.

П р и м е р. Покажем, что кольцо вещественных функций обладает делителями нуля. Действительно, полагая

$$f(x) = \begin{cases} 0 & \text{при} \quad x \le 0, \\ x & \text{при} \quad x > 0, \end{cases} \qquad g(x) = \begin{cases} x & \text{при} \quad x \le 0, \\ 0 & \text{при} \quad x > 0, \end{cases}$$

видим, что обе эти функции отличны от нуля, а их произведение равно нулю, т. е. функции, которая при любом x принимает значение 0.

У п р а ж н е н и е. Никакое поле не содержит делителей нуля.

Именно это свойство вещественных чисел и имелось в виду, когда в школе вас учили, что если произведение двух выражений равно нулю, то хотя бы одно из них равно нулю.

Из отсутствия делителей нуля в кольце K вытекает, что любое равенство можно сократить на ненулевой общий множитель. Действительно, если $c\,a = c\,b$, $a, b, c \in K$, и $c \neq 0$, то $c\,(a - b) = 0$, откуда заключаем, что $a - b = 0$, т. е. $a = b$.

3.4. Характеристика поля. Если P – поле, то в нем есть единица 1. Возьмем ее и будем складывать с собой. Возможно, что на

некотором шаге получим 0. Если впервые 0 получим на m-м шаге:

$$\underbrace{1+1+\ldots+1}_{m}=0,$$

то говорим, что *характеристика поля* P равна m.

Как мы знаем, в числовых полях такое невозможно. В этом случае говорим, что поле имеет характеристику 0. Примерами полей ненулевой характеристики служат все конечные поля; существуют, впрочем, и бесконечные поля, имеющие ненулевую характеристику.

Если поле имеет характеристику 2, то $1+1=0$, а потому и для любого элемента $a \in P$ сумма $a+a=0$, т. е. каждый элемент есть противоположный к себе.

У п р а ж н е н и е. Характеристика поля – либо 0, либо простое число.

3.5. Изоморфизм для колец и полей. Два кольца (поля) K и K' *изоморфны,* если существует взаимно однозначное отображение *на*

$$\varphi : K \longrightarrow K'$$

такое, что для всех $a, b \in K$ справедливы равенства:

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

$$(ab)\varphi = a\varphi \cdot b\varphi.$$

П р и м е р. Рассмотрим кольцо $\langle\{\overline{0},\overline{1}\};+,\cdot\rangle$, состоящее из двух элементов с операциями сложения и умножения, заданными таблицами

$+$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

$\cdot$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$

и другое кольцо $\langle\{$ н, ч $\};+,\cdot\rangle$, для которого

$+$	ч	н
ч	ч	н
н	н	ч

$\cdot$	ч	н
ч	ч	ч
н	ч	н

Нетрудно проверить, что эти два кольца изоморфны, если установить соответствие:

$$\overline{0} \longmapsto \text{ч}, \quad \overline{1} \longmapsto \text{н}.$$

3.6. Подкольцо, подполе. Подмножество кольца (поля) называется *подкольцом* (*подполем*), если оно замкнуто относительно сложения и умножения и само является кольцом (полем) относительно этих индуцированных операций.

Л е м м а 1. *Подмножество L кольца K тогда и только тогда является подкольцом, когда оно замкнуто относительно сложения, умножения и взятия противоположного элемента, т. е. когда выполняются следующие условия*:

а) *для любых $a, b \in L$ сумма $a + b \in L$*;

б) *для любых $a, b \in L$ произведение $ab \in L$*;

в) *для любого $a \in L$ противоположный $-a \in L$.*

Подмножество L поля P тогда и только тогда является подполем, когда оно замкнуто относительно сложения, умножения, взятия противоположного элемента и взятия обратного элемента. Последнее означает:

г) *для любого $a \in L, a \neq 0$ обратный элемент $a^{-1} \in L$.*

Д о к а з а т е л ь с т в о. Рассмотрим множество L с операцией сложения. Так как выполнены условия а и в, то ввиду леммы 1 из § 2 $\langle L; + \rangle$ является подгруппой группы $\langle K; + \rangle$, а так как выполнено условие б, то $\langle L; +, \cdot \rangle$ – подкольцо. Для доказательства второго утверждения достаточно заметить, что в силу условий б и г $\langle L \backslash \{0\}; \cdot \rangle$ является подгруппой группы $\langle P \backslash \{0\}; \cdot \rangle$, а потому $\langle L; +, \cdot \rangle$ является подполем.

Л е м м а 2. *Пересечение любого множества подколец* (*подполей*) *снова является подкольцом* (*подполем*).

Д о к а з а т е л ь с т в о. Пусть K – некоторое кольцо, L_α, $\alpha \in J$ – семейство подколец и $L = \bigcap_{\alpha \in J} L_\alpha$ – их пересечение. Чтобы доказать, что L – подкольцо, надо доказать, что L замкнуто относительно сложения, умножения и взятия противоположного, т. е.

а) если $a, b \in L$, то $a + b \in L$;

б) если $a, b \in L$, то $ab \in L$;

в) если $a \in L$, то $-a \in L$.

Для доказательства а заметим, что если $a, b \in L$, то $a, b \in L_\alpha$ для любого $\alpha \in J$. Следовательно, $a + b \in L_\alpha$ для любого $\alpha \in J$, но это и означает, что $a + b \in L$.

б) Если $a, b \in L$, то $a, b \in L_\alpha$ для любого $\alpha \in J$. Следовательно, $ab \in L_\alpha$ для любого $\alpha \in J$, т. е. $ab \in L$.

в) Пусть $a \in L$. Тогда $a \in L_\alpha$ для любого $\alpha \in J$, но тогда $-a \in L_\alpha$

для любого $\alpha \in J$ и, следовательно, $-a \in L$.

Если теперь K – поле, а L_α – семейство подполей, то мы должны установить следующее утверждение:

г) если $a \in L$, $a \neq 0$, то $a^{-1} \in L$.

Заметим, что если $a \in L$, то $a \in L_\alpha$ для любого $\alpha \in J$ и $a^{-1} \in L_\alpha$ для любого $\alpha \in J$, но это и означает, что $a^{-1} \in L$. Лемма доказана.

Подкольцом кольца K, порожденным множеством M, называется пересечение всех подколец, содержащих M, т. е.

$$(M) = \bigcap L_\alpha, \quad L_\alpha - \text{подкольцо в } K, \quad L_\alpha \supseteq M.$$

Подполем поля P, порожденным множеством M, называется пересечение всех подполей, содержащих M.

П р и м е р ы. 1) Кольцо $\langle \mathbb{Z}; +, \cdot \rangle$ порождается 1. Аналогично поле $\langle \mathbb{Q}; +, \cdot \rangle$ порождается 1.

2) В поле $\langle \mathbb{R}; +, \cdot \rangle$ рассмотрим подкольцо, порожденное элементами 1 и $\sqrt{2}$. Нетрудно заметить, что это подкольцо состоит из чисел

$$\{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}.$$

Более того, это подкольцо является подполем.

У п р а ж н е н и е. Поле L_2 неизоморфно полю L_3, где

$$L_n = \{a + b\sqrt{n} \mid a, b \in \mathbb{Q}\}, n \in \mathbb{N} \text{ и не является квадратом.}$$

У п р а ж н е н и е. При каких натуральных p и q имеет место изоморфизм $L_p \simeq L_q$?

§ 4. Группы подстановок

4.1. Определения. Пусть

$$M_n = \{1, 2, 3, \ldots, n\}$$

– конечное множество. *Подстановкой* множества M_n называется взаимно однозначное отображение этого множества на себя. Всякую подстановку можно записать в виде

$$a = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ i_1 & i_2 & i_3 & \ldots & i_n \end{pmatrix},$$

где внизу находятся элементы из M_n, в которые переходят верхние. Очевидно, что одна и та же подстановка может быть записана несколькими способами, например,

$$a = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 1 & 3 & 2 \\ 2 & 4 & 1 & 3 \end{pmatrix}.$$

На множестве подстановок определим умножение.

Произведением двух подстановок называется третья, равная последовательному выполнению первой, а затем второй.

П р и м е р. Если

$$a = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix},$$

то их произведения

$$ab = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}, \quad ba = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}.$$

Из этого примера видим, что операция умножения подстановок некоммутативна.

Обозначим

$$S_n = \{\text{подстановки множества } M_n\}$$

и будем называть *множеством подстановок степени* n. Нетрудно проверить, что $|S_n| = n!$, где символом $|A|$ обозначается число элементов множества A.

Справедлива

Т е о р е м а 1. *Множество S_n с операцией умножения образует группу. При $n \geq 3$ она некоммутативна.*

Д о к а з а т е л ь с т в о. То, что $\langle S_n; \cdot \rangle$ является алгебраической системой, следует из определения произведения подстановок. Проверим аксиомы группы.

1) Ассоциативность: проверим, что для любых подстановок a, b, c из S_n справедливо равенство

$$(ab)c = a(bc).$$

Для этого обозначим $ab = d$, $dc = f$ – левая часть равенства; $bc = g$, $ag = h$ – правая часть равенства. Выберем некоторый символ $i \in M_n = \{1, 2, \ldots, n\}$ и подействуем на него подстановкой f. Получим

$$if = (id)c = ((ia)b)c.$$

Действуя подстановкой h, получим

$$ih = (ia)g = ((ia)b)c.$$

Следовательно, на каждый символ $i \in M_n$ подстановки f и h действуют одинаково, но это означает, что они равны, т. е. ассоциативность умножения выполняется.

2) Легко проверить, что единичным элементом является подстановка

$$e = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix},$$

оставляющая все символы на месте.

3) Пусть

$$a = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$$

– произвольная подстановка. Нетрудно убедиться, что обратной является подстановка

$$a^{-1} = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ 1 & 2 & \dots & n \end{pmatrix}.$$

Покажем, что при $n \geq 3$ группа S_n некоммутативна. Положим

$$a = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 2 & 1 & 3 & 4 & \dots & n \end{pmatrix}, \qquad b = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 3 & 2 & 1 & 4 & \dots & n \end{pmatrix}.$$

Тогда $1(ab) = 2$, т. е. подстановка ab переводит символ 1 в символ 2, а $1(ba) = 3$, т. е. подстановка ba переводит символ 1 в символ 3. Следовательно, $ab \neq ba$. Теорема доказана.

Из этой теоремы, в частности, следует, что группа S_3 порядка 6 неабелева.

У п р а ж н е н и е. Существует ли неабелева группа порядка меньше 6?

4.2. Разложение подстановки в произведение независимых циклов. Рассмотрим подстановку

$$\begin{pmatrix} \dots & i_1 & \dots & i_2 & \dots & i_{s-1} & \dots & i_s & \dots \\ \dots & i_2 & \dots & i_3 & \dots & i_s & \dots & i_1 & \dots \end{pmatrix},$$

где все невыписанные символы остаются на месте. Тогда эту подстаноку будем записывать в виде $(i_1 i_2 \dots i_{s-1} i_s)$ и называть *циклической подстановкой*, или *циклом.*

П р и м е р. Подстановка

$$a = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 5 & 3 & 4 & 7 & 6 & 2 \end{pmatrix}$$

имеет следующее представление в виде цикла $a = (257)$. Заметим, что этот цикл можно записать несколькими способами:

$$(257) = (572) = (725).$$

Две циклические подстановки, или короче, два цикла называются *независимыми*, если их множества действительно перемещаемых символов не пересекаются. Легко заметить, что независимые циклы перестановочны.

П р и м е р. Пара зависимых циклов:

$$(123),\ (257)$$

(оба содержат символ 2); пара независимых циклов:

$$(134),\ (257).$$

Т е о р е м а 2. *Всякая нетождественная подстановка может быть разложена в произведение независимых циклов. Это разложение единственно с точностью до порядка множителей.*

Д о к а з а т е л ь с т в о. Пусть a – некоторая подстановка из S_n и t – число действительно перемещаемых символов. Проведем доказательство индукцией по t. Очевидно, что $t \geq 2$. При $t = 2$ имеем

$$a = \begin{pmatrix} \dots & i & \dots & j & \dots \\ \dots & j & \dots & i & \dots \end{pmatrix} = (ij)$$

– цикл, и утверждение теоремы справедливо.

При $t > 2$ представим нашу подстановку в таком виде:

$$a = \begin{pmatrix} \dots & i_1 & \dots & i_2 & \dots & i_{s-1} & \dots & i_s & \dots \\ \dots & i_2 & \dots & i_3 & \dots & i_s & \dots & i_1 & \dots \end{pmatrix}.$$

Здесь мы выбрали некоторый символ i_1 и следим за тем, куда он переходит. Начиная с некоторого момента символы будут повторяться, и первым повтором будет символ i_1 (символы, отмеченные точками, тоже могут перемещаться).

Определим подстановку

$$b = \begin{pmatrix} \dots & i_1 & \dots & i_2 & \dots & i_{s-1} & \dots & i_s & \dots \\ \dots & i_2 & \dots & i_3 & \dots & i_s & \dots & i_1 & \dots \end{pmatrix},$$

в которой невыписанные символы остаются на месте, и подстановку

$$c = \begin{pmatrix} \dots & i_1 & \dots & i_2 & \dots & i_{s-1} & \dots & i_s & \dots \\ \dots & i_1 & \dots & i_2 & \dots & i_{s-1} & \dots & i_s & \dots \end{pmatrix},$$

в которой невыписанные символы переходят в те, в которые переходят символы подстановки a, а выписанные остаются на месте.

Заметим, что $bc = a$. При этом b – это цикл, в b и c нет общих перемещаемых символов и в c число перемещаемых символов равно $t - s$, т. е. на s меньше, чем в a. По предположению индукции

$$c = c_1\, c_2\, \dots\, c_p$$

– произведение независимых циклов. Следовательно,

$$a = b\, c_1\, c_2\, \dots\, c_p$$

– произведение независимых циклов. Таким образом, мы представили всякую подстановку $a \in S_n$ в виде произведения независимых циклов.

Докажем единственность. Пусть

$$a = c_1\, c_2\, \dots\, c_k = d_1\, d_2\, \dots\, d_l$$

– два разложения подстановки a в произведения независимых циклов. Заметим, что если какой-то символ встречается в одной записи, то он встречается и в другой (если символ перемещается, это указывается в любой записи). Возьмем некоторый перемещаемый символ i и передвинем его в начало соответствующего цикла, а также поставим этот цикл на первое место. Будем иметь

$$c_1 = (i\, \alpha\, \beta\, \dots\, \delta), \qquad d_1 = (i\, \alpha'\, \beta'\, \dots\, \delta').$$

Это означает, что подстановка a содержит, с одной стороны, фрагмент

$$\begin{pmatrix} \dots & i & \dots \\ \dots & \alpha & \dots \end{pmatrix},$$

а с другой – фрагмент

$$\begin{pmatrix} \dots & i & \dots \\ \dots & \alpha' & \dots \end{pmatrix},$$

(учесть, что циклы независимы). Следовательно, $\alpha = \alpha'$. Теорема доказана.

4.3. Декремент. Четность подстановки. *Декрементом* подстановки называется разность между числом действительно перемещаемых символов и числом независимых циклов.

П р и м е р. Пусть

$$a = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 4 & 1 & 7 & 2 & 3 & 5 & 8 \end{pmatrix} = (163)(2475).$$

Тогда декремент d равен $7 - 2 = 5$.

Если декремент подстановки – четное число, то подстановка называется *четной*. Если декремент – нечетное число, то подстановка называется *нечетной*. Подстановка, переставляющая только два символа, называется *транспозицией*. Очевидно, транспозиция – нечетная подстановка.

Т е о р е м а 3. *При умножении произвольной подстановки на транспозицию ее четность меняется.*

Д о к а з а т е л ь с т в о. Пусть

$$a = (i_1\, i_2\, \dots\, i_r)\,(j_1\, j_2\, \dots\, j_s)\,\dots$$

– разложение подстановки a в произведение независимых циклов. Пусть при этом k – число действительно перемещаемых символов и l – число независимых циклов, декремент $d = k - l$. Рассмотрим транспозицию $b = (p\, q)$. При этом символы p и q могут либо входить, либо не входить в независимые циклы подстановки a. Разберем все эти случаи и результаты поместим в таблицу:

	Случаи	a
1	p и q не входят в a	$(i_1\, i_2\, \dots\, i_r)\,(j_1\, j_2\, \dots\, j_s)\,\dots$
2	p входит, q не входит в a	$(p\, i_2\, \dots\, i_r)\,(j_1\, j_2\, \dots\, j_s)\,\dots$
3	p не входит, q входит в a	$(q\, i_2\, \dots\, i_r)\,(j_1\, j_2\, \dots\, j_s)\,\dots$
4	p и q входят в один цикл	$(p\, \dots\, q\, \dots)\,(j_1\, j_2\, \dots\, j_s)\,\dots$
5	p и q входят в разные циклы	$(p\, \dots)\,(q\, \dots)\,(j_1\, \dots\, j_s)\,\dots$

ab	k'	l'	d'
$(i_1\, i_2\, \ldots\, i_r)\,(j_1\, j_2\, \ldots\, j_s)\,\ldots\,(p\, q)$	$k+2$	$l+1$	$d+1$
$(p\, i_2\, \ldots\, i_r\, q)\,(j_1\, j_2\, \ldots\, j_s)\,\ldots$	$k+1$	l	$d+1$
$(q\, i_2\, \ldots\, i_r\, p)\,(j_1\, j_2\, \ldots\, j_s)\,\ldots$	$k+1$	l	$d+1$
$(p\,\ldots)\,(q\,\ldots)\,(j_1\, j_2\, \ldots\, j_s)\,\ldots$	k	$l+1$	$d-1$
$(p\,\ldots\,q\,\ldots)\,(j_1\, j_2\, \ldots\, j_s)\,\ldots$	k	$l-1$	$d+1$

где k' – число действительно перемещаемых символов подстановки ab, l' – число ее независимых циклов, а $d' = k' - l'$ – декремент. Анализируя последний столбец полученной таблицы, получаем требуемое утверждение.

Т е о р е м а 4. *В группе S_n число четных и нечетных подстановок одно и то же и равно $\frac{1}{2}n!$.*

Д о к а з а т е л ь с т в о. Пусть

$$P = \{a_1,\, a_2,\, \ldots,\, a_s\}$$

– множество всех четных подстановок из S_n. Возьмем транспозицию $t = (1\,2)$ и рассмотрим множество подстановок:

$$Pt = \{a_1 t,\, a_2 t,\, \ldots,\, a_s t\}.$$

По теореме 3 все эти подстановки нечетные. Чтобы доказать теорему, надо установить, что

1) в Pt все подстановки различны;

2) всякая нечетная подстановка из S_n содержится в множестве Pt.

Докажем 1. Предположим, что $a_i t = a_j t$. Умножая обе части этого равенства справа на t, получим $a_i t^2 = a_j t^2$. Учитывая, что t^2 – тождественная подстановка, получим $a_i = a_j$, но так как в P все подстановки различны, то $a_i \neq a_j$.

Для доказательства 2 возьмем некоторую нечетную подстановку b и найдем $a = bt$. По теореме 3 a – четная подстановка, а все четные подстановки содержатся в множестве P. Следовательно, $a = a_i$ для некоторого i. Тогда в множестве Pt находим $a_i t = bt^2 = b$. Пункт 2 установлен.

Следовательно, число четных и нечетных подстановок в S_n одно и то же, а так как в S_n содержится $n!$ элементов, то это число равно $\frac{1}{2}n!$. Теорема доказана.

4.4. Порождающие множества. Транспозиции являются простейшими подстановками. Как показывает следующая теорема, множество транспозиций является порождающим множеством группы подстановок.

Т е о р е м а 5. *Всякая подстановка разлагается в произведение транспозиций. Это разложение неединственно, но четность числа транспозиций всегда одна и та же и совпадает с четностью самой подстановки.*

Следующее равенство показывает, что одна и та же подстановка может иметь различные разложения в произведение транспозиций:

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

Д о к а з а т е л ь с т в о т е о р е м ы. Рассмотрим некоторую подстановку a и представим ее в виде произведения независимых циклов:

$$a = c_1\,c_2\,\ldots c_s.$$

Легко проверить, что каждый цикл разлагается в произведение транспозиций:

$$(i_1\,i_2\,\ldots\,i_t) = (i_1\,i_2)\,(i_1\,i_3)\,\ldots\,(i_1\,i_t).$$

Следовательно, и сама подстановка a разлагается в произведение транспозиций:

$$a = t_1\,t_2\,\ldots\,t_k.$$

Так как транспозиция – нечетная подстановка, а при умножении ее на транспозицию получаем четную подстановку, то по теореме 3 четность k равна четности a. Теорема доказана.

Отметим, что найденное в теореме множество порождающих группы S_n при $n > 2$ не является минимальным ($S_2 = \text{гр}((1\,2))$ – циклическая группа порядка 2).

У п р а ж н е н и е. Группа S_n порождается множеством транспозиций $(1\,2)$, $(1\,3)$, $\ldots$, $(1\,n)$.

Оказывается, что и это множество не является минимальным.

У п р а ж н е н и е. Группа S_n, $n > 2$, порождается транспозицией $(1\,2)$ и циклом $(1\,2\,\ldots\,n)$.

Четные подстановки образуют подгруппу группы S_n, которая называется *знакопеременной группой* и обозначается символом A_n.

О п р е д е л е н и е. *Знаком подстановки* τ назовем следующее число:

$$\operatorname{sgn}\tau = \begin{cases} +1 & \text{если } \tau \text{ четная,} \\ -1 & \text{если } \tau \text{ нечетная.} \end{cases}$$

С л е д с т в и е. *Справедливы следующие равенства*:

$$\operatorname{sgn}(\sigma\cdot\tau) = \operatorname{sgn}\sigma\cdot\operatorname{sgn}\tau,$$

$$\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}\sigma.$$

Д о к а з а т е л ь с т в о. Представим подстановки σ и τ в виде произведения транспозиций:

$$\sigma = \sigma_1\,\sigma_2\,\ldots\,\sigma_s, \qquad \tau = \tau_1\,\tau_2\,\ldots\,\tau_t.$$

Тогда

$$\sigma\,\tau = \sigma_1\,\sigma_2\,\ldots\,\sigma_s\,\tau_1\,\tau_2\,\ldots\,\tau_t,$$
$$\sigma^{-1} = \sigma_s\,\ldots\,\sigma_2\,\sigma_1.$$

По теореме 5

$$\operatorname{sgn}\sigma = (-1)^s, \quad \operatorname{sgn}\tau = (-1)^t, \quad \operatorname{sgn}(\sigma\,\tau) = (-1)^{s+t}, \quad \operatorname{sgn}(\sigma^{-1}) = (-1)^s.$$

Следствие установлено.

§ 5. Кольца матриц

5.1. Суперпозиция линейных замен. Рассмотрим две линейные замены переменных с коэффициентами из кольца K:

$$\begin{cases} x_1 = a_{11}\,y_1 + \ldots + a_{1n}\,y_n, \\ x_2 = a_{21}\,y_1 + \ldots + a_{2n}\,y_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_r = a_{r1}\,y_1 + \ldots + a_{rn}\,y_n, \end{cases} \qquad a_{ij} \in K,$$

$$\begin{cases} y_1 = b_{11}\,z_1 + \ldots + b_{1s}\,z_s, \\ y_2 = b_{21}\,z_1 + \ldots + b_{2s}\,z_s, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = b_{n1}\,z_1 + \ldots + b_{ns}\,z_s, \end{cases} \qquad b_{kl} \in K.$$

Если подставить вторую замену в первую, то получим

$$x_i = a_{i1}\,y_1 + \ldots + a_{in}\,y_n = a_{i1}(b_{11}\,z_1 + \ldots + b_{1s}\,z_s) + \ldots + a_{in}(b_{n1}\,z_1 + \ldots + b_{ns}\,z_s) =$$
$$= (a_{i1}\,b_{11} + a_{i2}\,b_{21} + \ldots + a_{in}\,b_{n1})\,z_1 + \ldots + (a_{i1}\,b_{1s} + a_{i2}\,b_{2s} + \ldots + a_{in}\,b_{ns})\,z_s =$$
$$= c_{i1}\,z_1 + c_{i2}\,z_2 + \ldots + c_{is}\,z_s,$$

где

$$c_{ij} = a_{i1}\,b_{1j} + a_{i2}\,b_{2j} + \ldots + a_{in}\,b_{nj} = \sum_{k=1}^{n} a_{ik}\,b_{kj}.$$

Такая замена переменных, когда в одну подставляется другая, называется *суперпозицией линейных замен.* Получается опять линейная замена:

$$\begin{cases} x_1 = c_{11}\,z_1 \,+\, \ldots \,+ c_{1s}\,z_s, \\ x_2 = c_{21}\,z_1 \,+\, \ldots \,+ c_{2s}\,z_s, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_r = c_{r1}\,z_1 \,+\, \ldots \,+ c_{rs}\,z_s, \end{cases} \qquad c_{kl} \in K.$$

5.2. Матрицы. Сложение и умножение матриц. При работе с такими объектами удобно использовать матрицы. *Матрицей* размера $r \times n$ над кольцом K называется таблица

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots\ldots\ldots\ldots\ldots & & & \\ a_{r1} & a_{r2} & \ldots & a_{rn} \end{pmatrix}, \quad a_{ij} \in K.$$

В дальнейшем будем записывать такие матрицы как $A = (a_{ij})_{i,j=1}^{r,n}$ или, если размеры матрицы ясны, просто как $A = (a_{ij})$. Множество всех матриц размера $r \times n$ над кольцом K будем обозначать $M_{r\times n}(K)$.

Пусть заданы две матрицы, имеющие одинаковые размеры с элементами из одного кольца:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots\ldots\ldots\ldots\ldots & & & \\ a_{r1} & a_{r2} & \ldots & a_{rn} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \ldots\ldots\ldots\ldots\ldots & & & \\ b_{r1} & b_{r2} & \ldots & b_{rn} \end{pmatrix}.$$

Тогда их *суммой* называется матрица

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \ldots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \ldots & a_{2n} + b_{2n} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & & & \\ a_{r1} + b_{r1} & a_{r2} + b_{r2} & \ldots & a_{rn} + b_{rn} \end{pmatrix}.$$

Если матрицы

$$F = \begin{pmatrix} f_{11} & f_{12} & \ldots & f_{1n} \\ f_{21} & f_{22} & \ldots & f_{2n} \\ \ldots\ldots\ldots\ldots\ldots & & & \\ f_{r1} & f_{r2} & \ldots & f_{rn} \end{pmatrix}$$

и

$$G = \begin{pmatrix} g_{11} & g_{12} & \dots & g_{1s} \\ g_{21} & g_{22} & \dots & g_{2s} \\ \dots\dots\dots\dots\dots\dots \\ g_{n1} & g_{n2} & \dots & g_{ns} \end{pmatrix}$$

таковы, что число столбцов в F совпадает с числом строк в G, то мы можем определить *произведение*

$$F \cdot G = H = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1s} \\ h_{21} & h_{22} & \dots & h_{2s} \\ \dots\dots\dots\dots\dots\dots \\ h_{r1} & h_{r2} & \dots & h_{rs} \end{pmatrix},$$

где

$$h_{ij} = f_{i1}\, g_{1j} + f_{i2}\, g_{2j} + \dots + f_{in}\, g_{nj} = \sum_{k=1}^{n} f_{ik}\, g_{kj}.$$

Сравните полученную формулу с формулой для коэффициентов при суперпозиции линейных замен.

П р и м е р. Произведение двух матриц равно

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & 0 \end{pmatrix} \cdot \begin{pmatrix} 7 & 2 & 3 \\ -1 & 0 & 5 \\ 4 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 12 \\ 11 & 4 & 21 \end{pmatrix}.$$

5.3. Кольцо матриц. Как мы знаем, если матрицы квадратные, т. е. имеют размеры $n \times n$, то их можно складывать и умножать. Будем называть такие матрицы *матрицами степени* n. Определим множество

$$M_n(K) = \{\text{матрицы степени } n \text{ над } K\} = M_{n\times n}(K).$$

Т е о р е м а 1. *Если K – кольцо, то $\langle M_n(K);\, +,\, \cdot\rangle$ – кольцо. Если K – кольцо с единицей, то $\langle M_n(K);\, +,\, \cdot\rangle$ – кольцо с единицей.*

Д о к а з а т е л ь с т в о. То, что операции сложения и умножения матриц являются алгебраическими, следует из определения. Проверим аксиомы кольца.

С1. Рассмотрим матрицы $A, B, C \in M_n(K)$. Нам надо доказать, что

$$(A + B) + C = A + (B + C).$$

Рассмотрим элемент, стоящий на месте (i, j) в матрице $(A + B) + C$. Он равен $(a_{ij} + b_{ij}) + c_{ij}$. Так как K – кольцо, то

$$(a_{ij} + b_{ij}) + c_{ij} = a_{ij} + (b_{ij} + c_{ij}).$$

Следовательно, сложение матриц ассоциативно.

С2. Устанавливается аналогичным образом, при этом учитывается тот факт, что сложение в кольце K коммутативно.

С3. В качестве нулевого элемента надо взять матрицу, у которой на всех местах стоят нули.

С4. Легко проверить, что если матрица $A \in M_n(K)$ имеет вид

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

то противоположная

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

Для проверки ассоциативности умножения (аксиома У1) $(AB)C = A(BC)$, введем обозначения:

$$AB = D, \quad DC = F, \quad BC = G, \quad AG = H.$$

Нужно доказать, что $F = H$. Имеем:

$$f_{ij} = \sum_{k=1}^{n} d_{ik}\, c_{kj} = \sum_{k=1}^{n} \left(\sum_{l=1}^{n} a_{il}\, b_{lk} \right) c_{kj} = \sum_{k=1}^{n} \sum_{l=1}^{n} (a_{il}\, b_{lk})\, c_{kj}.$$

С другой стороны,

$$h_{ij} = \sum_{p=1}^{n} a_{ip}\, g_{pj} = \sum_{p=1}^{n} a_{ip} \left(\sum_{q=1}^{n} b_{pq}\, c_{qj} \right) =$$

$$= \sum_{p=1}^{n} \sum_{q=1}^{n} a_{ip}\, (b_{pq}\, c_{qj}) = \sum_{l=1}^{n} \sum_{k=1}^{n} a_{il}\, (b_{lk}\, c_{kj}),$$

где последнее равенство вытекает из следующего равенства:

$$\sum_{i=1}^{n} a_i = \sum_{p=1}^{n} a_p = a_1 + \ldots + a_n,$$

которое показывает, что индекс суммирования можно обозначить любой буквой.

Далее нам потребуется

Л е м м а 1. *Если* α_{ij}, $1 \le i \le r$, $1 \le j \le s$ – *элементы кольца* K, *то*

$$\sum_{i=1}^{r}\sum_{j=1}^{s} \alpha_{ij} = \sum_{j=1}^{s}\sum_{i=1}^{r} \alpha_{ij}.$$

Д о к а з а т е л ь с т в о. Построим матрицу

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1s} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2s} \\ \ldots\ldots\ldots\ldots\ldots & & & \\ \alpha_{r1} & \alpha_{r2} & \ldots & \alpha_{rs} \end{pmatrix}.$$

Сначала складываем элементы в каждой строке, а затем складываем полученные элементы:

$$(\alpha_{11} + \alpha_{12} + \ldots + \alpha_{1s}) + \ldots + (\alpha_{r1} + \alpha_{r2} + \ldots + \alpha_{rs}) = \sum_{i=1}^{r}\sum_{j=1}^{s} \alpha_{ij};$$

затем складываем элементы в каждом столбце и находим сумму полученных элементов:

$$(\alpha_{11} + \alpha_{21} + \ldots + \alpha_{r1}) + \ldots + (\alpha_{1s} + \alpha_{2s} + \ldots + \alpha_{rs}) = \sum_{j=1}^{s}\sum_{i=1}^{r} \alpha_{ij}.$$

Из коммутативности сложения в K следует, что эти две суммы совпадают. Лемма доказана.

Воспользовавшись этой леммой, получим:

$$\sum_{l=1}^{n}\sum_{k=1}^{n} a_{il}\,(b_{lk}\,c_{kj}) = \sum_{k=1}^{n}\sum_{l=1}^{n}(a_{il}\,b_{lk})\,c_{kj}.$$

Следовательно, аксиома У1 установлена.

Для доказательства СУ1, $(A+B)C = AC + BC$, введем обозначения

$$A + B = D, \quad DC = F, \quad AC = G, \quad BC = H, \quad G + H = U.$$

Нам надо доказать, что $F = U$. Для этого вычислим элемент, стоящий на месте (i, j) в матрице F. Имеем:

$$f_{ij} = \sum_{k=1}^{n} d_{ik}\, c_{kj} = \sum_{k=1}^{n} (a_{ik} + b_{ik})\, c_{kj}.$$

С другой стороны,

$$u_{ij} = g_{ij} + h_{ij} = \sum_{k=1}^{n} a_{ik}\, c_{kj} + \sum_{k=1}^{n} b_{ik}\, c_{kj} = \sum_{k=1}^{n} (a_{ik} + b_{ik}) c_{kj},$$

где в последнем равенстве мы воспользовались коммутативностью сложения и правой дистрибутивностью в кольце K.

Аксиома СУ2 проверяется аналогично.

Если K – кольцо с единицей 1, то единицей кольца $M_n(K)$ является единичная матрица E, у которой на главной диагонали стоят 1, а на всех остальных местах – нули. Теорема доказана.

5.4. Диагональные матрицы и трансвекции. В кольце $M_n(K)$ введем два класса матриц.

О п р е д е л е н и е. *Диагональной матрицей* называется матрица, у которой все элементы вне главной диагонали равны нулю:

$$\operatorname{diag}(\alpha_1, \ldots, \alpha_n) = \begin{pmatrix} \alpha_1 & 0 & \ldots & 0 \\ 0 & \alpha_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \alpha_n \end{pmatrix}.$$

О п р е д е л е н и е. *Трансвекцией* $T_{ij}(\beta)$, $1 \le i \ne j \le n$ называется матрица, у которой на главной диагонали стоят 1, на месте (i, j) – элемент $\beta \in K$, а на всех остальных местах – нули.

Посмотрим, что происходит при умножении произвольной матрицы на диагональную. Если матрицу умножить на диагональную справа, то получим

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 & 0 & \ldots & 0 \\ 0 & \alpha_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \alpha_n \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11}\alpha_1 & a_{12}\alpha_2 & \dots & a_{1n}\alpha_n \\ a_{21}\alpha_1 & a_{22}\alpha_2 & \dots & a_{2n}\alpha_n \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}\alpha_1 & a_{n2}\alpha_2 & \dots & a_{nn}\alpha_n \end{pmatrix}.$$

Если матрицу умножить на диагональную слева, то получим

$$\begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \dots\dots\dots\dots\dots \\ 0 & 0 & \dots & \alpha_n \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots\dots\dots\dots\dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} =$$

$$= \begin{pmatrix} \alpha_1 a_{11} & \alpha_1 a_{12} & \dots & \alpha_1 a_{1n} \\ \alpha_2 a_{21} & \alpha_2 a_{22} & \dots & \alpha_2 a_{2n} \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_n a_{n1} & \alpha_n a_{n2} & \dots & \alpha_n a_{nn} \end{pmatrix}.$$

Таким образом, справедлива

Л е м м а 2. *При умножении произвольной матрицы на диагональную матрицу* $\operatorname{diag}(\alpha_1, \dots, \alpha_n)$ *справа первый столбец умножается на* α_1, *второй – на* α_2 *и т. д. При умножении произвольной матрицы на диагональную матрицу* $\operatorname{diag}(\alpha_1, \dots, \alpha_n)$ *слева первая строка умножается на* α_1, *вторая – на* α_2 *и т. д.*

Пусть дана трансвекция $T_{ij}(\beta)$. При умножении ее на произвольную матрицу справа получим

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots\dots\dots\dots\dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \cdot T_{ij}(\beta) =$$

$$= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,j-1} & a_{1i}\beta + a_{1j} & a_{1,j+1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2,j-1} & a_{2i}\beta + a_{2j} & a_{2,j+1} & \dots & a_{2n} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1} & a_{n2} & \dots & a_{n,j-1} & a_{ni}\beta + a_{nj} & a_{n,j+1} & \dots & a_{nn} \end{pmatrix},$$

где сумма стоит в j-м столбце.

При умножении слева получим матрицу

$$T_{ij}(\beta)\cdot\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{i-1,1} & a_{i-1,2} & \dots & a_{i-1,n} \\ a_{i1}+\beta a_{j1} & a_{i2}+\beta a_{j2} & \dots & a_{in}+\beta a_{jn} \\ a_{i+1,1} & a_{i+1,2} & \dots & a_{i+1,n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

где сумма стоит в i-й строке.

Справедлива

Л е м м а 3. *При умножении произвольной матрицы на трансвекцию $T_{ij}(\beta)$ справа к ее j-му столбцу прибавляется i-й столбец, умноженный на β. При умножении произвольной матрицы на трансвекцию $T_{ij}(\beta)$ слева к ее i-й строке прибавляется j-я строка, умноженная на β.*

Отметим также следующие свойства трансвекций.

Л е м м а 4. 1) *Для произведения трансвекций справедливо равенство*

$$T_{ij}(\beta)\cdot T_{ij}(\gamma) = T_{ij}(\beta+\gamma).$$

2) *Обратной к трансвекции $T_{ij}(\beta)$ является трансвекция $T_{ij}(-\beta)$.*

Д о к а з а т е л ь с т в о. 1) По лемме 3 умножение $T_{ij}(\beta)$ справа на трансвекцию $T_{ij}(\gamma)$ соответствует тому, что мы к j-му столбцу матрицы $T_{ij}(\beta)$ прибавляем i-й столбец, умноженный на γ. В результате получим трансвекцию $T_{ij}(\beta+\gamma)$.

2) По пункту 1 имеем

$$T_{ij}(\beta)\cdot T_{ij}(-\beta) = T_{ij}(0),$$

а $T_{ij}(0)$ – единичная матрица.

5.5. Разложение матрицы в произведение диагональной и трансвекций. Основным результатом настоящего пункта является доказательство следующего утверждения.

Т е о р е м а 2. *Пусть P – поле. Всякая матрица $A \in M_n(P)$ разлагается в произведение*

$$A = T_1\, T_2\, \dots\, T_r\, D\, T_{r+1}\, T_{r+2}\, \dots\, T_s,$$

где D – диагональная матрица, T_i, $i = 1, 2, \dots, s$, – трансвекции.

Д о к а з а т е л ь с т в о. Назовем *элементарными преобразованиями матрицы* следующие преобразования:

1) прибавление к одной строке матрицы другой ее строки, умноженной на ненулевой элемент из P;

2) прибавление к одному столбцу матрицы другого ее столбца, умноженного на ненулевой элемент из P.

Используя индукцию по n, докажем, что при помощи этих элементарных преобразований всякую матрицу можно привести к диагональной матрице.

При $n = 1$ матрица A диагональная.

Предположим, что матрицы порядка $n-1$ мы умеем приводить к диагональному виду. Рассмотрим матрицу $A = (a_{ij}) \in M_n(P)$ степени n. В зависимости от вида этой матрицы рассмотрим несколько случаев.

Случай 1: $a_{11} \neq 0$. Так как P – поле, то существует элемент a_{11}^{-1}, обратный к элементу a_{11}. Умножим первый столбец матрицы A на $-a_{1j}a_{11}^{-1}$ и прибавим к j-му столбцу, где $j = 2, 3, \ldots, n$. Получим матрицу, у которой все элементы, стоящие в первой строке, за исключением первого элемента, равны нулю:

$$\begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ a_{21} & a'_{22} & \ldots & a'_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a'_{n2} & \ldots & a'_{nn} \end{pmatrix}.$$

Умножим первую строку на $-a_{i1}a_{11}^{-1}$ и прибавим к i-ой строке для всех $i = 2, 3, \ldots, n$. Получим матрицу

$$\left(\begin{array}{c|ccc} a_{11} & 0 & \ldots & 0 \\ \hline 0 & * & \ldots & * \\ \vdots & \multicolumn{3}{c}{\ldots\ldots\ldots} \\ 0 & * & \ldots & * \end{array}\right).$$

Применяя предположение индукции, приведем эту матрицу при помощи элементарных преобразований к диагональному виду:

$$\begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & b_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & b_n \end{pmatrix}.$$

Случай 2: $a_{11} = 0$, но какой-то элемент, a_{1j} или a_{i1}, $i, j \in \{2, 3, \ldots, n\}$, отличен от нуля. Если $a_{1j} \neq 0$, то умножим j-й столбец на единицу и прибавим к первому столбцу. Если $a_{i1} \neq 0$, то умножим i-ую строку на 1 и прибавим к первой строке. В обоих случаях приходим к разобранному выше случаю 1.

Случай 3: весь первый столбец и вся первая строка матрицы A состоят из нулей. В этом случае, воспользовавшись предположением индукции, приведем ее к диагональному виду.

Возвращаемся к доказательству теоремы. Как следует из леммы 3, умножение матрицы на трансвекцию справа соответствует элементарному преобразованию столбцов матрицы, а умножение матрицы на трансвекцию слева соответствует элементарному преобразованию строк матрицы. Таким образом, мы установили, что найдутся трансвекции $U_1, U_2, \ldots, U_r, U_{r+1}, \ldots, U_s$ такие, что

$$U_1 U_2 \ldots U_r A U_{r+1} \ldots U_s = D,$$

где D – некоторая диагональная матрица.

Воспользовавшись далее леммой 4, получим

$$A = U_r^{-1} U_{r-1}^{-1} \ldots U_1^{-1} D U_s^{-1} \ldots U_{r+1}^{-1},$$

где каждая U_i^{-1}, $i = 1, 2, \ldots, s$, как следует из леммы 4, является трансвекцией. Теорема доказана.

§ 6. Определители

6.1. Определитель обратимости матрицы. Если K – кольцо, то матрицы степени n над K образуют кольцо, но не поле. Нетрудно заметить, что не всякая ненулевая матрица обладает обратной.

П р и м е р. Рассмотрим ненулевую матрицу

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

и покажем, что у нее не существует обратной. Действительно, найдем произведение

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} z & t \\ 0 & 0 \end{pmatrix},$$

и очевидно, что ни для каких $x, y, z, t \in K$ последняя матрица не является единичной.

Как по произвольной матрице узнать, имеет ли она обратную?

Введем следующее

О п р е д е л е н и е. *Определитель обратимости матрицы*, или просто *определитель* матрицы $A \in M_n(P)$, есть следующий элемент из поля P:

$$\det A = |A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots\dots\dots\dots\dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \cdot a_{1,1\sigma}\, a_{2,2\sigma}\, \dots\, a_{n,n\sigma},$$

где S_n – множество подстановок степени n.

Если A имеет степень n, то $\det A$ называют *определителем порядка* n. Покажем, как, используя это определение, можно вычислять определители порядка 2 и 3.

П р и м е р 1. При $n = 2$ множество S_n содержит две подстановки:

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

Причем $\operatorname{sgn}\sigma_1 = +1$, $\operatorname{sgn}\sigma_2 = -1$. Тогда

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \operatorname{sgn}\sigma_1 \cdot a_{1,1\sigma_1}\, a_{2,2\sigma_1} + \operatorname{sgn}\sigma_2 \cdot a_{1,1\sigma_2}\, a_{2,2\sigma_2} = a_{11}\, a_{22} - a_{12}\, a_{21}.$$

П р и м е р 2. При $n = 3$ множество S_n содержит шесть подстановок:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$

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

При этом, три первых имеют знак $+1$, а три последних – знак -1. Следовательно,

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

$$- a_{13}\, a_{22}\, a_{31} - a_{12}\, a_{21}\, a_{33} - a_{11}\, a_{23}\, a_{32}.$$

6.2. Свойства определителей. Матрица $C \in M_n(P)$ называется *полураспавшейся,* если она имеет вид

$$C = \begin{pmatrix} A & * \\ \mathbf{0} & B \end{pmatrix},\ A \in M_r(P),\ B \in M_{n-r}(P),\ \mathbf{0} \in M_{(n-r)\times r}(P),\ 1 < r < n.$$

1. *Определитель полураспавшейся матрицы вычисляется по формуле*

$$\det \begin{pmatrix} A & * \\ \mathbf{0} & B \end{pmatrix} = \det A \cdot \det B.$$

Д о к а з а т е л ь с т в о. Рассмотрим полураспавшуюся матрицу

$$C = \left(\begin{array}{cccc|ccc} a_{11} & a_{12} & \dots & a_{1r} & a_{1,r+1} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rr} & a_{r,r+1} & \dots & a_{rn} \\ \hline 0 & 0 & \dots & 0 & b_{r+1,r+1} & \dots & b_{r+1,n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & b_{n,r+1} & \dots & b_{nn} \end{array}\right).$$

Найдем произведение

$$\det A \cdot \det B = \sum_{\sigma \in S_r} \operatorname{sgn}\sigma \cdot a_{1,1\sigma}\, a_{2,2\sigma}\, \dots\, a_{r,r\sigma} \cdot$$

$$\cdot \sum_{\tau \in S_{n-r}} \operatorname{sgn}\tau \cdot b_{r+1,(r+1)\tau}\, b_{r+2,(r+2)\tau}\, \dots\, b_{n,n\tau} =$$

$$= \sum_{\sigma \in S_r, \tau \in S_{n-r}} \operatorname{sgn}\sigma \operatorname{sgn}\tau \cdot a_{1,1\sigma}\, a_{2,2\sigma}\, \dots\, a_{r,r\sigma}\, b_{r+1,(r+1)\tau}\, b_{r+2,(r+2)\tau}\, \dots\, b_{n,n\tau},$$

где σ – подстановка множества $\{1, 2, \dots, r\}$, а τ – подстановка множества $\{r+1,\ r+2,\ \dots, n\}$.

Обозначим

$$\widetilde{\sigma} = \left(\begin{array}{cccc|ccc} 1 & 2 & \dots & r & r+1 & \dots & n \\ 1\sigma & 2\sigma & \dots & r\sigma & r+1 & \dots & n \end{array}\right),$$

$$\widetilde{\tau} = \left(\begin{array}{cccc|ccc} 1 & 2 & \dots & r & r+1 & \dots & n \\ 1 & 2 & \dots & r & (r+1)\tau & \dots & n\tau \end{array}\right)$$

– подстановки множества $\{1, 2, \dots, n\}$, т. е. обе эти подстановки действуют на одном множестве, лежат в группе S_n, и мы можем их перемножать. При этом

$$\operatorname{sgn}\sigma = \operatorname{sgn}\widetilde{\sigma}, \quad \operatorname{sgn}\tau = \operatorname{sgn}\widetilde{\tau}.$$

Следовательно,

$$\det A \cdot \det B = \sum_{\widetilde{\sigma},\widetilde{\tau}\in S_n} \operatorname{sgn}\widetilde{\sigma} \cdot \operatorname{sgn}\widetilde{\tau} \cdot a_{1,1\widetilde{\sigma}}\, a_{2,2\widetilde{\sigma}}\, \ldots\, a_{r,r\widetilde{\sigma}} \cdot$$

$$\cdot b_{r+1,(r+1)\widetilde{\tau}}\, b_{r+2,(r+2)\widetilde{\tau}}\, \ldots\, b_{n,n\widetilde{\tau}} =$$

$$= \sum_{\widetilde{\sigma},\widetilde{\tau}\in S_n} \operatorname{sgn}(\widetilde{\sigma}\cdot\widetilde{\tau}) \cdot a_{1,1\widetilde{\sigma}\widetilde{\tau}}\, a_{2,2\widetilde{\sigma}\widetilde{\tau}}\, \ldots\, a_{r,r\widetilde{\sigma}\widetilde{\tau}}\, b_{r+1,(r+1)\widetilde{\sigma}\widetilde{\tau}}\, b_{r+2,(r+2)\widetilde{\sigma}\widetilde{\tau}}\, \ldots\, b_{n,n\widetilde{\sigma}\widetilde{\tau}} =$$

$$= \sum_{\pi\in S_n} \operatorname{sgn}\pi \cdot a_{1,1\pi}\, a_{2,2\pi}\, \ldots\, a_{r,r\pi}\, b_{r+1,(r+1)\pi}\, b_{r+2,(r+2)\pi}\, \ldots\, b_{n,n\pi} = \det C.$$

Здесь мы воспользовались таким фактом: если некоторая подстановка $\pi \in S_n$ не представима в виде $\widetilde{\sigma}\widetilde{\tau}$, то произведение $c_{1,1\pi}\, c_{2,2\pi}\, \ldots\, c_{n,n\pi}$ обращается в нуль, так как содержит элемент $c_{kl} = 0$. Докажем его. Пусть подстановка имеет вид

$$\pi = \left(\begin{array}{ccc|ccc} \ldots & i & \ldots & \ldots & k & \ldots \\ \ldots & j & \ldots & \ldots & l & \ldots \end{array}\right), \quad i, l \le r, \quad k, j > r,$$

тогда в соответствующем произведении содержится множитель $c_{kl} = 0$.

О п р е д е л е н и е. Матрица $A^t \in M_{s\times r}(P)$ называется *транспонированной к матрице* $A \in M_{r\times s}(P)$, если ее i-й столбец совпадает с i-й строкой матрицы A.

П р и м е р. Для матрицы

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

транспонированной будет

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

Для квадратной матрицы транспонировать – это все равно что повернуть матрицу относительно главной диагонали. Для квадратных матриц справедливо следующее свойство.

2. *Определитель транспонированной матрицы равен определителю самой матрицы*:

$$\det A^t = \det A.$$

Д о к а з а т е л ь с т в о. Обозначим элементы матрицы A^t символами b_{ij}. Тогда $b_{ij} = a_{ji}$, $1 \le i, j \le n$. Находим определитель матрицы A^t:

$$\det A^t = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \cdot b_{1,1\sigma}\, b_{2,2\sigma} \,\ldots\, b_{n,n\sigma} = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \cdot a_{1\sigma,1}\, a_{2\sigma,2} \,\ldots\, a_{n\sigma,n} =$$

$$= \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \cdot a_{1,1\sigma}\, a_{2,2\sigma} \,\ldots\, a_{n,n\sigma} = \det A.$$

3. *При перестановке двух строк матрицы ее определитель меняет знак.*

Д о к а з а т е л ь с т в о. Пусть

$$B = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{j1} & a_{j2} & \ldots & a_{jn} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{i1} & a_{i2} & \ldots & a_{in} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

– матрица, полученная из A перестановкой i-й и j-й строк. Тогда

$$\det B = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \cdot a_{1,1\sigma}\, a_{2,2\sigma} \,\ldots\, a_{j,i\sigma} \,\ldots\, a_{i,j\sigma} \,\ldots\, a_{n,n\sigma} =$$

$$= \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \cdot a_{1,1\tau\sigma}\, a_{2,2\tau\sigma} \,\ldots\, a_{i,i\tau\sigma} \,\ldots\, a_{j,j\tau\sigma} \,\ldots\, a_{n,n\tau\sigma} =$$

$$= -\sum_{\tau\sigma \in S_n} \operatorname{sgn}\tau\sigma \cdot a_{1,1\tau\sigma}\, a_{2,2\tau\sigma} \,\ldots\, a_{i,i\tau\sigma} \,\ldots\, a_{j,j\tau\sigma} \,\ldots\, a_{n,n\tau\sigma},$$

где $\tau = (i, j)$ – транспозиция и действие подстановки $\tau\sigma$ определяется следующим образом:

$$i\tau\sigma = j\sigma, \quad j\tau\sigma = i\sigma, \quad k\tau\sigma = k\sigma \text{ если } k \ne i, j.$$

При этом мы использовали такой факт: если

$$\{\sigma_1, \sigma_2, \ldots, \sigma_{n!}\}$$

– множество всех подстановок из S_n, то

$$\{\tau\sigma_1, \tau\sigma_2, \ldots, \tau\sigma_{n!}\}$$

– тоже множество всех подстановок из S_n.

С л е д с т в и е. *Если две строки матрицы одинаковы, то ее определитель равен нулю.*

Действительно, если характеристика поля P отлична от 2, то, переставляя две одинаковые строки, получим $\det A = -\det A$.

У п р а ж н е н и е. Докажите следствие для полей характеристики 2.

4. *Если матрица в одной из строчек содержит сумму, то ее определитель записывается как сумма определителей:*

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{i-1,1} & a_{i-1,2} & \dots & a_{i-1,n} \\ b_1 + c_1 & b_2 + c_2 & \dots & b_n + c_n \\ a_{i+1,1} & a_{i+1,2} & \dots & a_{i+1,n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{i-1,1} & a_{i-1,2} & \dots & a_{i-1,n} \\ b_1 & b_2 & \dots & b_n \\ a_{i+1,1} & a_{i+1,2} & \dots & a_{i+1,n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} +$$

$$+ \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{i-1,1} & a_{i-1,2} & \dots & a_{i-1,n} \\ c_1 & c_2 & \dots & c_n \\ a_{i+1,1} & a_{i+1,2} & \dots & a_{i+1,n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix},$$

где сумма стоит в i-й строке.

Д о к а з а т е л ь с т в о следует из равенства

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \cdot a_{1,1\sigma}\, a_{2,2\sigma}\, \dots\, a_{i-1,(i-1)\sigma}\, (b_{i\sigma} + c_{i\sigma})\, a_{i+1,(i+1)\sigma}\, \dots\, a_{n,n\sigma} =$$

$$= \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \cdot a_{1,1\sigma}\, a_{2,2\sigma}\, \dots\, a_{i-1,(i-1)\sigma}\, b_{i\sigma}\, a_{i+1,(i+1)\sigma}\, \dots\, a_{n,n\sigma} +$$

$$+ \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \cdot a_{1,1\sigma}\, a_{2,2\sigma}\, \dots\, a_{i-1,(i-1)\sigma}\, c_{i\sigma}\, a_{i+1,(i+1)\sigma}\, \dots\, a_{n,n\sigma}.$$

5. *Если все элементы некоторой строки матрицы умножить на $\alpha \in P$, то ее определитель умножится на α.*

Действительно, пусть матрица B получается из матрицы A умножением i-й строки на α. Тогда

$$\det B = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \cdot a_{1,1\sigma}\, a_{2,2\sigma}\, \ldots\, (\alpha a_{i,i\sigma})\, \ldots\, a_{n,n\sigma} =$$

$$= \alpha \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \cdot a_{1,1\sigma}\, a_{2,2\sigma}\, \ldots\, a_{i,i\sigma}\, \ldots\, a_{n,n\sigma} = \alpha \det A.$$

С л е д с т в и е. *Если две строки матрицы пропорциональны, то ее определитель равен нулю.*

Для формулировки последнего свойства определителей введем

О п р е д е л е н и е. *Линейной комбинацией строк* матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ \ldots. \\ A_n \end{pmatrix}$$

с коэффициентами $\beta_1, \beta_2, \ldots, \beta_n$ из P называется строка

$$\beta_1 A_1 + \beta_2 A_2 + \ldots + \beta_n A_n = \left(\sum_{i=1}^{n} \beta_i a_{i1}, \sum_{i=1}^{n} \beta_i a_{i2}, \ldots, \sum_{i=1}^{n} \beta_i a_{in} \right).$$

Линейная комбинация называется *нетривиальной*, если не все коэффициенты β_i равны нулю.

П р и м е р. Линейной комбинацией строк матрицы

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

с коэффициентами $(-1, 0, 2)$ является строка

$$-1(2, -1, 3, 4) + 0(0, 5, 1, 2) + 2(3, -1, 0, 1) = (4, -1, -3, -2).$$

6. *Определитель матрицы не изменится, если к какой-нибудь ее строке прибавить линейную комбинацию остальных строк* (*кроме нее самой*).

Д о к а з а т е л ь с т в о. Рассмотрим матрицу

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ \ldots. \\ A_n \end{pmatrix},$$

где A_i – i-я строка матрицы A. Если к первой строке матрицы прибавить линейную комбинацию остальных строк, то ввиду свойства 4 определитель полученной матрицы равен

$$\begin{vmatrix} A_1 + \beta_2 A_2 + \ldots + \beta_n A_n \\ A_2 \\ \vdots \\ A_n \end{vmatrix} = \begin{vmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{vmatrix} + \begin{vmatrix} \beta_2 A_2 \\ A_2 \\ \vdots \\ A_n \end{vmatrix} + \ldots + \begin{vmatrix} \beta_n A_n \\ A_2 \\ \vdots \\ A_n \end{vmatrix}.$$

Заметим, что в правой части все определители, кроме первого, равны нулю. Действительно, во втором определителе первая строка пропорциональна второй, в третьем – первая строка пропорциональна третьей и т. д. и, наконец, в последнем определителе первая строка пропорциональна последней. По следствию свойства 5 заключаем, что все они равны нулю. Таким образом, сумма, стоящая в правой части, равна определителю матрицы A.

Говорят, что между строками матрицы A существует нетривиальная линейная зависимость, если некоторая нетривиальная линейная комбинация строк равна нулевой строке.

С л е д с т в и е. *Если между строками матрицы существует нетривиальная линейная зависимость, то ее определитель равен* 0.

Мы сформулировали свойства определителей 4–6 для строк матрицы, но ввиду свойства 2 они справедливы и для столбцов.

6.3. Существование обратной матрицы. Теперь мы можем сформулировать критерий существования обратной матрицы.

Т е о р е м а 1. *Матрица тогда и только тогда обратима, когда ее определитель не равен нулю.*

Эта теорема вытекает из следующего более общего утверждения

Т е о р е м а 2. *Пусть P – поле, $A \in M_n(P)$, тогда равносильны следующие утверждения*:

1) *существует $X \in M_n(P)$ такая, что $AX = XA = E$*;

2) *существует $Y \in M_n(P)$ такая, что $AY = E$*;

3) *существует $Z \in M_n(P)$ такая, что $ZA = E$*;

4) *определитель* $\det A$ *не равен нулю.*

Д о к а з а т е л ь с т в о. 1) $\Rightarrow$ 2), 1) $\Rightarrow$ 3). Очевидно.

2) $\Rightarrow$ 4) (импликация 3) $\Rightarrow$ 4) устанавливается аналогично). Представим матрицу A в виде произведения диагональной и трансвекций:

$$A = T_1\, T_2\, \ldots T_r\, D\, T_{r+1}\, \ldots\, T_s.$$

Покажем, что справедливо равенство

$$\det A = \det D.$$

Действительно, так как умножить матрицу справа на трансвекцию это все равно что к одной строке прибавить другую строку, умноженную на некоторый коэффициент, то по свойству 6 получаем требуемое равенство.

Следовательно,

$$\det A = \det D = \alpha_1\,\alpha_2\,\ldots\,\alpha_n,$$

где α_i – коэффициенты диагональной матрицы $D = \operatorname{diag}(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n)$.

Предположим, что $A\,Y = E$. Имеем

$$T_1\,T_2\,\ldots T_r\,D\,T_{r+1}\,\ldots\,T_s\,Y = E,$$

и надо доказать, что все $\alpha_i \neq 0$. Так как каждая трансвекция обратима, то

$$D\,T_{r+1}\,\ldots\,T_s\,Y = T_r^{-1}\,T_{r-1}^{-1}\,\ldots T_1^{-1},$$

откуда

$$D\,T_{r+1}\,\ldots\,T_s\,Y\,T_1\,T_2\,\ldots T_r = E.$$

Если бы некоторое $\alpha_i = 0$, то и вся строка матрицы D была бы нулевой, и при умножении ее на любую матрицу эта строка была бы нулевой. Приходим к противоречию, так как справа стоит единичная матрица E, у которой все строки ненулевые. Следовательно, все $\alpha_i \neq 0$, а потому и $\det A = \alpha_1\,\alpha_2\,\ldots\,\alpha_n \neq 0$.

4) $\Rightarrow$ 1). Так как

$$\alpha_1\,\alpha_2\,\ldots\,\alpha_n = \det D = \det A \neq 0,$$

то все $\alpha_i \neq 0$, а потому обладают обратными. Тогда и для матрицы D существует обратная:

$$D^{-1} = \operatorname{diag}(\alpha_1^{-1},\ \alpha_2^{-1},\ \ldots,\ \alpha_n^{-1}).$$

В качестве матрицы X возьмем матрицу

$$X = T_s^{-1}\,T_{s-1}^{-1}\,\ldots T_{r+1}^{-1}\,D^{-1}\,T_r^{-1}\,\ldots\,T_1^{-1}.$$

Теорема доказана.

6.4. Определитель произведения матриц. Справедлива

Т е о р е м а 3. *Определитель произведения равен произведению определителей*:

$$\det(A \cdot B) = \det A \cdot \det B.$$

Д о к а з а т е л ь с т в о разобьем на несколько случаев.

Случай 1: хотя бы одна из матриц, A или B, необратима. Тогда по теореме 2 либо $\det A = 0$, либо $\det B = 0$. Покажем, что в этом случае матрица AB также необратима. Предположим противное, т. е. найдется матрица X такая, что

$$(AB)X = X(AB) = E.$$

Тогда $A(BX) = E$ и по теореме 2 $\det A \neq 0$. С другой стороны, $(XA)B = E$ и опять по теореме 2 $\det B \neq 0$. Противоречие. Следовательно, обратной к матрице AB не существует, а потому $\det(AB) = 0$.

Случай 2: обе матрицы A и B обратимы. По теореме 2 $\det A \neq 0$, $\det B \neq 0$. Пусть

$$A = T_1\, T_2\, \ldots T_r\, D\, T_{r+1}\, \ldots\, T_s,$$

где $D = \operatorname{diag}(\alpha_1,\, \alpha_2,\, \ldots,\, \alpha_n)$ – диагональная матрица, а T_i – трансвекция. Аналогично

$$B = U_1\, U_2\, \ldots U_p\, F\, U_{p+1}\, \ldots\, U_q,$$

где $F = \operatorname{diag}(\beta_1,\, \beta_2,\, \ldots,\, \beta_n)$ – диагональная матрица, а U_j – трансвекция.

По свойству 6 имеем

$$\det A = \det D = \alpha_1\, \alpha_2\, \ldots\, \alpha_n,$$

$$\det B = \det F = \beta_1\, \beta_2\, \ldots\, \beta_n,$$

и произведение

$$DF = \operatorname{diag}(\alpha_1\beta_1,\, \alpha_2\beta_2,\, \ldots,\, \alpha_n\beta_n)$$

имеет определитель

$$\det(DF) = \alpha_1\, \alpha_2\, \ldots\, \alpha_n\, \beta_1\, \beta_2\, \ldots\, \beta_n = \det D \cdot \det F.$$

Рассмотрим произведение

$$AB = T_1\, T_2\, \ldots T_r\, D\, T_{r+1}\, \ldots\, T_s \cdot U_1\, U_2\, \ldots U_p\, F\, U_{p+1}\, \ldots\, U_q =$$

$$= T_1\,T_2\,\ldots T_r\,(D\,T_{r+1}D^{-1})\,(D\,T_{r+2}D^{-1})\ldots(D\,T_sD^{-1})\times$$
$$\times DF\,(F^{-1}U_1F)\;\ldots(F^{-1}U_pF)\,U_{p+1}\;\ldots\,U_q.$$

Заметим, что каждая матрица $D\,T_iD^{-1}$ и $F^{-1}\,U_jF$ является трансвекцией.

Л е м м а. *Если* $D = \operatorname{diag}(\alpha_1,\,\alpha_2,\,\ldots,\,\alpha_n)$ – *невырожденная диагональная матрица, то*

$$D^{-1}\,T_{ij}(\beta)\,D = T_{ij}\left(\beta\frac{\alpha_j}{\alpha_i}\right),$$

т. е. матрица, сопряженная с трансвекцией при помощи диагональной матрицы, опять является трансвекцией.

Из этой леммы следует, что матрицы DT_iD^{-1} и $F^{-1}U_jF$ являются трансвекциями, а потому

$$\det(AB) = \det(DF) = \det D\cdot\det F = \det A\cdot\det B.$$

Теорема доказана.

6.5. Разложение определителя по строке. Пусть $A = (a_{ij})$ – матрица из $M_n(P)$. Символом M_{ij} будем обозначать матрицу, полученную из A вычеркиванием i-й строки и j-го столбца. Очевидно, M_{ij} опять является квадратной матрицей. *Алгебраическим дополнением* в A к элементу a_{ij} называется элемент

$$A_{ij} = (-1)^{i+j}\det(M_{ij}).$$

При этом определитель $\det(M_{ij})$ называется *дополняющим минором.*

Т е о р е м а 4. *Для всякой строки матрицы* $A \in M_n(P)$ *справедливо равенство*

$$\sum_{k=1}^{n} a_{ik}\,A_{jk} = \begin{cases} \det A & \text{при } i = j,\\ 0 & \text{при } i \neq j.\end{cases}$$

Д о к а з а т е л ь с т в о разбивается на несколько случаев.

Случай 1. $i = j$. В этом случае

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n}\\ \ldots & \ldots & \ldots & \ldots\\ a_{i1} & a_{i2} & \ldots & a_{in}\\ \ldots & \ldots & \ldots & \ldots\\ a_{n1} & a_{n2} & \ldots & a_{nn}\end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n}\\ \ldots & \ldots & \ldots & \ldots\\ a_{i1} & 0 & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots\\ a_{n1} & a_{n2} & \ldots & a_{nn}\end{vmatrix} +$$

$$+\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ 0 & a_{i2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \dots + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

В каждом из определителей, стоящих в правой части, будем переставлять i-ю строку с предыдущей, пока не поставим ее на место первой строки. Затем переставляем столбцы так, чтобы в каждом определителе на месте $(1,1)$ стоял ненулевой элемент. Получим

$$\det A = (-1)^{i-1}\left|\begin{array}{c|ccc} a_{i1} & 0 & \dots & 0 \\ \hline * & & & \\ \vdots & & M_{i1} & \\ * & & & \end{array}\right| + (-1)^{i}\left|\begin{array}{c|ccc} a_{i2} & 0 & \dots & 0 \\ \hline * & & & \\ \vdots & & M_{i2} & \\ * & & & \end{array}\right| + \dots$$

$$+(-1)^{i+n-2}\left|\begin{array}{c|ccc} a_{in} & 0 & \dots & 0 \\ \hline * & & & \\ \vdots & & M_{in} & \\ * & & & \end{array}\right| =$$

$$= (-1)^{i-1}\, a_{i1}\det(M_{i1}) + (-1)^{i}\, a_{i2}\det(M_{i2}) + \dots + (-1)^{i+n-2}\, a_{in}\det(M_{in}) =$$

$$= a_{i1}\, A_{i1} + a_{i2}\, A_{i2} + \dots + a_{in}\, A_{in},$$

где в предпоследнем равенстве мы воспользовались первым свойством определителей (определитель полураспавшейся матрицы), а в последнем – определением алгебраического дополнения.

Случай 2. $i \neq j$. Рассмотрим матрицу B, которая получается из матрицы A, если вместо j-й строки поставить i-ю. Расписывая определитель матрицы B по j-й строке, ввиду предыдущего случая имеем

$$\det B = a_{i1}A_{j1} + a_{i2}A_{j2} + \dots + a_{in}A_{jn},$$

но, как мы знаем из свойств определителя, если в некоторой квадратной матрице две строки одинаковы, то ее определитель равен нулю. Следовательно, $\det B = 0$ и

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \dots + a_{in}A_{jn} = 0.$$

Теорема доказана.

6.6. Применение к системам линейных уравнений. *Системой линейных уравнений* называется следующее выражение:

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

Матрица

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

называется *матрицей системы*, а столбцы

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

– *столбцом неизвестных* и *столбцом свободных членов* соответственно. Кратко систему (1) можно записать в таком виде:

$$A\,X = B.$$

Решением системы (1) называется упорядоченный набор

$$(x_1^0,\, x_2^0,\, \ldots,\, x_n^0) \in P^n,$$

удовлетворяющий всем уравнениям системы.

Изучая системы, мы хотим ответить на следующие вопросы: имеет ли система решения, а если имеет, то как их найти?

Если в системе $m = n$, т. е. число уравнений равно числу неизвестных, то мы можем найти определитель матрицы A. Если он отличен от нуля, то говорим, что *система невырожденная*. Для невырожденных систем справедлива

Т е о р е м а 5. *Всякая невырожденная система* (1) *имеет единственное решение, которое дается формулой*

$$\left(\frac{d_1}{d},\, \frac{d_2}{d},\, \ldots,\, \frac{d_n}{d}\right),$$

где $d = \det A$ *– определитель матрицы системы, а* d_j *– определитель матрицы, получаемой из* A *заменой* j*-го столбца столбцом свободных членов.*

Доказательство. Вначале докажем существование решений. Так как по условию определитель матрицы A отличен от нуля, то она имеет обратную. Найдем $X = A^{-1} B$. Тогда

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

Следовательно, решение существует.

Докажем единственность решения. Пусть Y – решение, т. е. $A Y = B$. Умножим обе части этого равенства слева на A^{-1}. Получим

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

т. е. $Y = A^{-1} B$, а потому $X = Y$. Следовательно, решение единственно.

Покажем, что решение задается формулой

$$\left(\frac{d_1}{d}, \frac{d_2}{d}, \ldots, \frac{d_n}{d} \right).$$

По теореме 4 имеем

$$d_j = \begin{vmatrix} a_{11} & \ldots & a_{1,j-1} & b_1 & a_{1,j+1} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2,j-1} & b_2 & a_{2,j+1} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{n,j-1} & b_n & a_{n,j+1} & \ldots & a_{nn} \end{vmatrix} =$$

$$= b_1 A_{1j} + b_2 A_{2j} + \ldots + b_n A_{nj} = \sum_{k=1}^{n} b_k A_{kj}.$$

Подставим выражения $x_j = d_j/d$, $j = 1, 2, \ldots, n$ в i-е уравнение системы:

$$\sum_{j=1}^{n} a_{ij} x_j = \sum_{j=1}^{n} a_{ij} \frac{d_j}{d} = \frac{1}{d} \sum_{j=1}^{n} a_{ij} d_j = \frac{1}{d} \sum_{j=1}^{n} a_{ij} \left(\sum_{k=1}^{n} b_k A_{kj} \right) =$$

$$= \frac{1}{d} \sum_{j=1}^{n} \sum_{k=1}^{n} a_{ij} b_k A_{kj} = \frac{1}{d} \sum_{k=1}^{n} \sum_{j=1}^{n} a_{ij} b_k A_{kj} =$$

$$= \frac{1}{d} \sum_{k=1}^{n} b_k \left(\sum_{j=1}^{n} a_{ij} A_{kj} \right) = \frac{1}{d} b_i d = b_i.$$

Здесь в предпоследнем равенстве мы использовали теорему 4. Теорема доказана.

Формулы, полученные в этой теореме, называются *формулами Крамера.*

6.7. Применение к вычислению обратной матрицы. Как мы знаем, матрица $A \in M_n(P)$ обратима тогда и только тогда, когда ее определитель отличен от нуля. Справедлива

Т е о р е м а 6. *Если* $A = (a_{ij}) \in M_n(P)$ *и* $\det A \neq 0$, *то обратная матрица может быть найдена по формуле*

$$A^{-1} = \begin{pmatrix} \frac{A_{11}}{d} & \frac{A_{21}}{d} & \dots & \frac{A_{n1}}{d} \\ \dots\dots\dots\dots\dots\dots \\ \frac{A_{1r}}{d} & \frac{A_{2r}}{d} & \dots & \frac{A_{nr}}{d} \\ \dots\dots\dots\dots\dots\dots \\ \frac{A_{1n}}{d} & \frac{A_{2n}}{d} & \dots & \frac{A_{nn}}{d} \end{pmatrix},$$

где $d = \det A$, A_{ij} – *алгебраическое дополнение к элементу* a_{ij} *в матрице* A.

Д о к а з а т е л ь с т в о. Пусть $Y = (y_{ij}) = (\frac{A_{ji}}{d})$ – матрица из теоремы. Надо доказать, что $AY = YA = E$. Обозначим $AY = C$. Тогда

$$c_{pq} = \sum_{r=1}^{n} a_{pr} y_{rq} = \sum_{r=1}^{n} a_{pr} \frac{A_{qr}}{d} = \frac{1}{d} \sum_{r=1}^{n} a_{pr} A_{qr}.$$

По теореме 4 имеем

$$\frac{1}{d} \sum_{r=1}^{n} a_{pr} A_{qr} = \begin{cases} 1 & \text{если} \quad p = q, \\ 0 & \text{если} \quad p \neq q. \end{cases}$$

Следовательно, $C = E$ – единичная матрица. Аналогично проверяется, что $YA = E$. Теорема доказана.

Матрица

$$A^* = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ \dots\dots\dots\dots\dots\dots \\ A_{1r} & A_{2r} & \dots & A_{nr} \\ \dots\dots\dots\dots\dots\dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$$

называется *присоединенной матрицей* к матрице A. Очевидно, что

$$AA^* = A^*A = \operatorname{diag}(d, d, \ldots, d).$$

П р и м е р. Для матрицы

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

с отличным от нуля определителем $\det A = \alpha\,\delta - \beta\,\gamma \neq 0$ обратная находится по формуле

$$A^{-1} = \frac{1}{\alpha\,\delta - \beta\,\gamma} \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}.$$

§ 7. Поле комплексных чисел

7.1. Определение. Предположим, что мы знаем, что такое поле действительных чисел $\mathbb{R}$. Хотим построить поле комплексных чисел $\mathbb{C}$. Для этого построим поле P со следующими свойствами:

1) $\mathbb{R}$ является подполем поля P;

2) в P существует элемент ξ, для которого справедливо равенство $\xi^2 + 1 = 0$;

3) подполе поля P, порожденное $\mathbb{R}$ и ξ, совпадает с P.

Следующая теорема гарантирует существование такого поля P.

Т е о р е м а 1. *Поле P со свойствами* 1–3 *существует. Любые два поля со свойствами* 1–3 *изоморфны.* Любое из таких полей называется *полем комплексных чисел.*

Доказательство этой теоремы разобьем на две части. Вначале докажем существование, а потом единственность.

7.2. Существование поля комплексных чисел. Рассмотрим следующее подмножество матриц:

$$C' = \left\{ \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \;\middle|\; \alpha, \beta \in \mathbb{R} \right\} \subseteq M_2(\mathbb{R}).$$

Покажем, что это множество с операциями сложения и умножения матриц является алгебраической системой. Для этого надо убедиться, что операции сложения и умножения являются алгебраическими.

Возьмем две матрицы

$$A = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}, \quad B = \begin{pmatrix} \gamma & \delta \\ -\delta & \gamma \end{pmatrix}$$

из C'. Находя их сумму и произведение, получим

$$A+B = \begin{pmatrix} \alpha+\gamma & \beta+\delta \\ -(\beta+\delta) & \alpha+\gamma \end{pmatrix}, \quad A\cdot B = \begin{pmatrix} \alpha\gamma-\beta\delta & \alpha\delta+\beta\gamma \\ -(\beta\gamma+\alpha\delta) & -\beta\delta+\alpha\gamma \end{pmatrix}.$$

Следовательно, операции действительно являются алгебраическими, и мы имеем алгебраическую систему

$$\langle\, C';\, +,\, \cdot\,\rangle.$$

Покажем, что эта алгебраическая система является полем. Для этого надо проверить все аксиомы поля. Аксиомы С1, С2, У1, СУ1, СУ2 выполняются для всех квадратных матриц. Чтобы проверить аксиому С3, заметим, что нулевая матрица принадлежит множеству C'. Чтобы проверить С4, заметим, что если некоторая матрица лежит в C', то и противоположная лежит в C'.

Для проверки У2 находим

$$A\cdot B = \begin{pmatrix} \alpha\gamma-\beta\delta & \alpha\delta+\beta\gamma \\ -\beta\gamma-\alpha\delta & -\beta\delta+\alpha\gamma \end{pmatrix}, \quad B\cdot A = \begin{pmatrix} \gamma\alpha-\delta\beta & \gamma\beta+\delta\alpha \\ -\gamma\beta-\delta\alpha & -\delta\beta+\gamma\alpha \end{pmatrix}.$$

Следовательно, $A \cdot B = B \cdot A$.

У3. Единичная матрица

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

очевидно лежит в C'.

Покажем наконец, что для всякой ненулевой матрицы $A \in C'$ найдется обратная, лежащая в C'. Так как A – ненулевая, то либо $\alpha \neq 0$, либо $\beta \neq 0$, а потому $\det A = \alpha^2 + \beta^2 \neq 0$. Следовательно, обратная матрица существует. Используя теорему о вычислении обратной матрицы, находим

$$A^{-1} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = \begin{pmatrix} \frac{\alpha}{\alpha^2+\beta^2} & -\frac{\beta}{\alpha^2+\beta^2} \\ \frac{\beta}{\alpha^2+\beta^2} & \frac{\alpha}{\alpha^2+\beta^2} \end{pmatrix},$$

т. е. A^{-1} лежит в C'. Следовательно, C' является полем.

Рассмотрим отображение $\varphi : \mathbb{R} \longrightarrow C'$, определенное формулой

$$\alpha\varphi = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}, \quad \text{где} \quad \alpha \in \mathbb{R}.$$

Легко заметить, что это отображение однозначно и унивалентно. Покажем, что оно сохраняет операции. Для этого надо проверить, что для любых $\alpha, \beta \in \mathbb{R}$ справедливы равенства

$$(\alpha + \beta)\varphi = \alpha\varphi + \beta\varphi, \quad (\alpha \cdot \beta)\varphi = \alpha\varphi \cdot \beta\varphi.$$

Эти равенства следуют из правил сложения и умножения диагональных матриц

$$\begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix} + \begin{pmatrix} \beta & 0 \\ 0 & \beta \end{pmatrix} = \begin{pmatrix} \alpha + \beta & 0 \\ 0 & \alpha + \beta \end{pmatrix},$$

$$\begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix} \cdot \begin{pmatrix} \beta & 0 \\ 0 & \beta \end{pmatrix} = \begin{pmatrix} \alpha \cdot \beta & 0 \\ 0 & \alpha \cdot \beta \end{pmatrix}.$$

Отметим, что отображение φ не является отображением *на*, а является вложением.

Обозначим символом R' следующее подмножество диагональных матриц:

$$R' = \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix} \,\middle|\, \alpha \in \mathbb{R} \right\} \subseteq C'.$$

Фактически мы доказали, что $\varphi(\mathbb{R}) = R'$, т. е. поле $\mathbb{R}$ изоморфно полю R'.

Введем множество

$$C = (C' \setminus R') \cup \mathbb{R}.$$

Перенесем операции из множества матриц C' на множество C.

П р и м е р. Рассмотрим сумму следующих матриц из C':

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

В результате получим матрицу из C', но она не входит в C.

Определим на C операции $+$ и $\cdot$ следующим образом. Если z_1 и z_2 – два элемента из C, то положим

$$z_1+z_2=\begin{cases} z_1+z_2 & \text{если} \quad z_1\notin R',\, z_2\notin R',\, z_1+z_2\notin R',\\ \alpha & \text{если} \quad z_1\notin R',\, z_2\notin R',\\ & \qquad z_1+z_2=\begin{pmatrix}\alpha & 0\\ 0 & \alpha\end{pmatrix}\in R',\\ z_1+\begin{pmatrix}\beta & 0\\ 0 & \beta\end{pmatrix} & \text{если} \quad z_1\notin R',\, z_2=\beta\in\mathbb{R},\\ \begin{pmatrix}\alpha & 0\\ 0 & \alpha\end{pmatrix}+z_2 & \text{если} \quad z_1=\alpha\in\mathbb{R},\, z_2\notin R',\\ \alpha+\beta & \text{если} \quad z_1=\alpha\in\mathbb{R},\, z_2=\beta\in\mathbb{R}\end{cases}$$

– их сумма;

$$z_1\cdot z_2=\begin{cases} z_1\cdot z_2 & \text{если} \quad z_1\notin R',\, z_2\notin R',\, z_1\cdot z_2\notin R',\\ \alpha & \text{если} \quad z_1\notin R',\, z_2\notin R',\, z_1\cdot z_2=\begin{pmatrix}\alpha & 0\\ 0 & \alpha\end{pmatrix}\in R',\\ z_1\cdot\begin{pmatrix}\beta & 0\\ 0 & \beta\end{pmatrix} & \text{если} \quad z_1\notin R',\, z_2=\beta\in\mathbb{R},\\ \begin{pmatrix}\alpha & 0\\ 0 & \alpha\end{pmatrix}\cdot z_2 & \text{если} \quad z_1=\alpha\in\mathbb{R},\, z_2\notin R',\\ \alpha\cdot\beta & \text{если} \quad z_1=\alpha\in\mathbb{R},\, z_2=\beta\in\mathbb{R}\end{cases}$$

– их произведение.

По построению имеем изоморфизм:

$$\langle\, C;\,+,\,\cdot\,\rangle\simeq\langle\, C';\,+,\,\cdot\,\rangle.$$

Покажем, что C – то поле, которое мы хотели построить. Для этого надо проверить, что выполняются условия 1–3:

1) по построению поле $\mathbb{R}$ содержится в C;

2) покажем, что в качестве ξ можно взять матрицу

$$i=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\in C.$$

Действительно,

$$i^2=i\cdot i=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\cdot\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}=\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}.$$

Последняя матрица не лежит в C, но в C лежит число -1. Следовательно,

$$i^2 = -1, \quad \text{т. е.} \quad i^2 + 1 = 0;$$

3) покажем, что подполе поля C, порожденное $\mathbb{R}$ и $\imath$ совпадает с C. Для этого заметим, что

$$C = \{\alpha + i\,\beta \,|\, \alpha,\, \beta \in \mathbb{R}\}.$$

Действительно, рассмотрим произведение

$$i\beta = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \beta & 0 \\ 0 & \beta \end{pmatrix} = \begin{pmatrix} 0 & \beta \\ -\beta & 0 \end{pmatrix}$$

и сумму

$$\alpha + i\beta = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix} + \begin{pmatrix} 0 & \beta \\ -\beta & 0 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

Видим, что полученная матрица лежит в C, но из таких матриц и состоит множество C.

Пусть теперь L – подполе поля C, порожденное $\mathbb{R}$ и i. В этом поле лежат вещественные числа, лежит i, а так как поле замкнуто относительно операций сложения и умножения, то и всякий элемент $\alpha + i\beta$ лежит в L, т. е. $C \subseteq L$, по условию $L \subseteq C$. Следовательно, $L = C$, и существование поля со свойствами 1–3 доказано.

7.3. Единственность поля комплексных чисел. Надо доказать, что любые два поля со свойствами 1–3 изоморфны. Вначале поймем, как устроено поле P. Докажем, что

$$P = \{\alpha + \xi\,\beta \,|\, \alpha, \beta \in \mathbb{R}\}.$$

Обозначим правую часть этого равенства через

$$M = \{\alpha + \xi\,\beta \,|\, \alpha, \beta \in \mathbb{R}\}.$$

Надо доказать два включения:

$$P \supseteq M, \quad P \subseteq M.$$

Первое включение следует из того, что P содержит $\mathbb{R}$ и ξ.

Чтобы доказать второе включение, покажем, что M – подполе в P. Для этого надо доказать, что M замкнуто относительно сложения, умножения, взятия противоположного и обратного:

а) проверим, что M замкнуто относительно сложения, имеем

$$(\alpha + \xi\,\beta) + (\alpha' + \xi\,\beta') = (\alpha + \alpha') + \xi\,(\beta + \beta') \in M;$$

б) проверим, что M замкнуто относительно умножения, имеем

$$(\alpha + \xi\,\beta) \cdot (\alpha' + \xi\,\beta') = (\alpha\alpha' - \beta\beta') + \xi\,(\alpha\beta' + \alpha'\beta) \in M;$$

в) легко заметить, что противоположным к элементу $\alpha + \xi\,\beta$ является элемент $-(\alpha + \xi\,\beta) = (-\alpha) + \xi\,(-\beta) \in M$;

г) рассмотрим $\alpha + \xi\,\beta \neq 0$, тогда

$$(\alpha + \xi\,\beta)^{-1} = \frac{\alpha}{\alpha^2 + \beta^2} - \xi\frac{\beta}{\alpha^2 + \beta^2} \in M.$$

Следовательно, M является подполем поля P.

Покажем, что M содержит все действительные числа и элемент ξ. Первое следует из равенства $\alpha = \alpha + \xi \cdot 0 \in M$, а второе – из равенства $\xi = 0 + \xi \cdot 1 \in M$. Но тогда по свойству 3 определения поля P имеем включение $P \subseteq M$.

Докажем, что любое поле P со свойствами 1–3 изоморфно полю C'. Это следует из того, что если два поля изоморфны некоторому третьему полю, то они изоморфны между собой. Отсюда и последует нужное утверждение.

Т е о р е м а 2. *Отображение* $\omega : C' \longrightarrow P$, *действующее по правилу*

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \omega = \alpha + \xi\,\beta,$$

является изоморфизмом.

Д о к а з а т е л ь с т в о. По определению изоморфизма мы должны проверить следующие условия для ω:

1) однозначность;

2) унивалентность;

3) отображение *на*;

4) для любых $z_1, z_2 \in C'$ справедливо равенство $(z_1 + z_2)\omega = z_1\omega + z_2\omega$;

5) для любых $z_1, z_2 \in C'$ справедливо равенство $(z_1 \cdot z_2)\omega = z_1\omega \cdot z_2\omega$.

Условие 1 справедливо в силу определения.

2) Пусть

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \omega = \alpha + \xi\,\beta, \quad \begin{pmatrix} \alpha' & \beta' \\ -\beta' & \alpha' \end{pmatrix} \omega = \alpha' + \xi\,\beta',$$

и предположим, что $\alpha+\xi\,\beta=\alpha'+\xi\,\beta'$. Тогда $\alpha-\alpha'=\xi\,(\beta-\beta')$. Если $\beta-\beta'\neq 0$, то $\xi=\frac{\alpha-\alpha'}{\beta-\beta'}\in\mathbb{R}$, но в поле $\mathbb{R}$ нет элемента, удовлетворяющего равенству $\xi^2+1=0$. Противоречие. Следовательно, $\beta=\beta'$, $\alpha=\alpha'$.

3) Фактически уже доказано. Следует из того, что

$$P=\{\alpha+\xi\,\beta\,|\,\alpha,\beta\in\mathbb{R}\}.$$

4) Левая часть равенства имеет вид

$$\left[\begin{pmatrix}\alpha & \beta\\ -\beta & \alpha\end{pmatrix}+\begin{pmatrix}\alpha' & \beta'\\ -\beta' & \alpha'\end{pmatrix}\right]\omega=\begin{pmatrix}\alpha+\alpha' & \beta+\beta'\\ -(\beta+\beta') & \alpha+\alpha'\end{pmatrix}\omega=$$
$$=(\alpha+\alpha')+\xi\,(\beta+\beta'),$$

правая часть имеет вид

$$(\alpha+\xi\,\beta)+(\alpha'+\xi\,\beta')=(\alpha+\alpha')+\xi\,(\beta+\beta').$$

Эти части равны, а потому равенство справедливо. Свойство 5 устанавливается аналогично. Теорема доказана.

Множество комплексных чисел будем обозначать символом

$$\mathbb{C}=\{a+i\,b\,|\,a,b\in\mathbb{R}\}.$$

7.4. Геометрическая интерпретация поля комплексных чисел. Как мы установили выше, всякое комплексное число $z\in\mathbb{C}$ единственным образом представимо в виде $z=a+i\,b$. При этом a называется *действительной частью числа* z и обозначается $\operatorname{Re} z$; b называется *мнимой частью числа* z и обозначается $\operatorname{Im} z$; $r=\sqrt{a^2+b^2}$ называется *модулем числа* z и обозначается $|z|$. Сопоставим этому числу точку на плоскости с координатами (a,b).

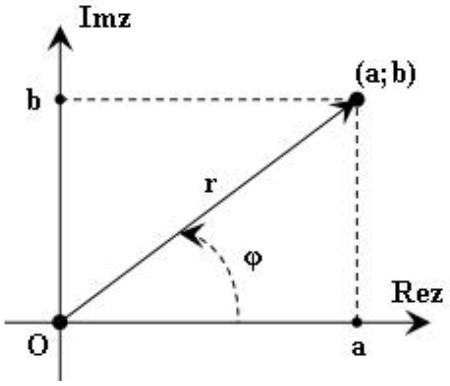

Рис.

С другой стороны, положение точки z на плоскости определяется заданием ее полярных координат: расстояния $r = |z|$ от начала координат до z и угла φ между положительным направлением оси абсцисс и направлением из начала координат на z (см. рис). Угол φ называется *аргументом числа* z и обозначается символом $\arg z = \varphi = \operatorname{arctg} y/x$. По определению $\arg z$ может принимать любые положительные и отрицательные значения, но при заданном r, углы, отличающиеся на целое кратное 2π, соответствуют одному и тому же числу. Аргумент не определен для числа 0 с модулем $|0| = 0$. Таким образом, мы приходим к *тригонометрической форме комплексного числа* $z = r(\cos\varphi + i\,\sin\varphi)$, $r \in \mathbb{R}$.

Комплексно-сопряженным к числу $z = a + i\,b$ называется число $\overline{z} = a - i\,b$. Очевидно, если $z \in \mathbb{R}$, то $\overline{z} = z$.

У п р а ж н е н и е. $\operatorname{Re}(z_1 + z_2) = \operatorname{Re} z_1 + \operatorname{Re} z_2$, $\operatorname{Im}(z_1 + z_2) = \operatorname{Im} z_1 + \operatorname{Im} z_2$.

У п р а ж н е н и е. $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$, $\quad \arg(z_1 \cdot z_2) = \arg z_1 + \arg z_2$.

§ 8. Общая линейная группа и ее важнейшие подгруппы

Пусть $\mathrm{M}_n(P)$ – множество всех матриц степени n над полем P. Определим следующие множества:

$$\mathrm{GL}_n(P) = \{A \in \mathrm{M}_n(P) \,|\, \det A \neq 0\},$$

$$\mathrm{SL}_n(P) = \{A \in \mathrm{M}_n(P) \,|\, \det A = 1\},$$

$$\mathrm{O}_n(P) = \left\{A \in \mathrm{M}_n(P) \,|\, A \cdot A^t = E\right\},$$

$$\mathrm{U}_n = \left\{A \in \mathrm{M}_n(\mathbb{C}) \,|\, A \cdot \overline{A^t} = E\right\},$$

где A^t – матрица, транспонированная к A, а $\overline{A^t}$ – матрица, полученная из A транспонированием и взятием комплексно-сопряженного к каждому элементу, т. е. если $A = (a_{ij})$, то в матрице A^t на месте (i, j) стоит элемент a_{ji}, а в матрице $\overline{A^t}$ на месте (i, j) стоит элемент $\overline{a_{ji}}$. Множество U_n мы определяем только в случае, когда $P = \mathbb{C}$ – поле комплексных чисел.

Заметим, что в определении $\mathrm{O}_n(P)$ мы требуем, чтобы выполнялось равенство $A \cdot A^t = E$, которое означает, что для A существует правая обратная. Но по теореме о существовании обратной матрицы отсюда следует, что A^t является и левой обратной, т. е. справедливо

равенство $A^t \cdot A = E$. Аналогичным образом получаем, что всякая матрица A из U_n удовлетворяет равенству $\overline{A^t} \cdot A = E$.

Покажем, что все эти множества являются группами относительно умножения матриц. Для этого нам потребуются два вспомогательных утверждения.

Л е м м а 1. *Для любых матриц* $A, B \in \mathrm{M}_n(P)$ *справедливо равенство*

$$(A \cdot B)^t = B^t \cdot A^t.$$

Д о к а з а т е л ь с т в о. Обозначим

$$A \cdot B = C, \quad C^t = D, \quad A^t = F, \quad B^t = G, \quad G \cdot F = H.$$

Надо доказать, что $D = H$. Имеем

$$d_{ij} = c_{ji} = \sum_{k=1}^{n} a_{jk}\, b_{ki},$$

$$h_{ij} = \sum_{k=1}^{n} g_{ik}\, f_{kj} = \sum_{k=1}^{n} b_{ki}\, a_{jk} = \sum_{k=1}^{n} a_{jk}\, b_{ki},$$

т. е. $d_{ij} = h_{ij}$. Лемма доказана.

Л е м м а 2. *Для любой матрицы* $A \in \mathrm{GL}_n(P)$ *справедливо равенство*

$$(A^t)^{-1} = (A^{-1})^t.$$

Д о к а з а т е л ь с т в о следует из равенства

$$(A^{-1})^t \cdot A^t = (A\, A^{-1})^t = E^t = E.$$

Т е о р е м а. *Относительно матричного умножения* $\mathrm{GL}_n(P)$ – *группа, а* $\mathrm{SL}_n(P)$, $\mathrm{O}_n(P)$ – *ее подгруппы.*

Д о к а з а т е л ь с т в о. Установим, что $\mathrm{GL}_n(P)$ – группа. Пусть A, B – две матрицы из $\mathrm{GL}_n(P)$. Тогда $\det A \neq 0$, $\det B \neq 0$ и по свойству определителей

$$\det(AB) = \det A \cdot \det B \neq 0.$$

Следовательно, операция умножения является алгебраической на множестве $\mathrm{GL}_n(P)$. Проверим аксиомы группы. Ассоциативность умножения выполнена в $\mathrm{M}_n(P)$, а потому и в $\mathrm{GL}_n(P)$. Единичная матрица существует в $\mathrm{M}_n(P)$ и, очевидно, она лежит в $\mathrm{GL}_n(P)$. Так

как для всякой матрицы $A \in \mathrm{GL}_n(P)$ ее определитель отличен от нуля, то существует A^{-1} такая, что $A \cdot A^{-1} = E$. Тогда из свойства определителей $\det A \cdot \det A^{-1} = 1$, а потому $\det A^{-1} \neq 0$. Следовательно, A^{-1} лежит в $\mathrm{GL}_n(P)$. Таким образом, мы доказали, что $\mathrm{GL}_n(P)$ является группой.

Очевидно, что множества $\mathrm{SL}_n(P)$, $\mathrm{O}_n(P)$ содержатся в $\mathrm{GL}_n(P)$. Чтобы доказать, что каждое из них является подгруппой, надо доказать замкнутость относительно умножения и взятия обратного.

Если A, B – две матрицы из $\mathrm{SL}_n(P)$, то

$$\det(AB) = \det A \cdot \det B = 1,$$

т. е. их произведение опять лежит в $\mathrm{SL}_n(P)$. Мы знаем, что A^{-1} существует, а из равенства $\det A \cdot \det A^{-1} = 1$ следует, что $\det A^{-1} = 1$. Таким образом, A^{-1} тоже лежит в $\mathrm{SL}_n(P)$, и мы установили, что $\mathrm{SL}_n(P)$ – подгруппа группы $\mathrm{GL}_n(P)$.

Рассмотрим множество $\mathrm{O}_n(P)$. Так как $A \cdot A^t = E$, то $\det A \neq 0$, а потому $\mathrm{O}_n(P)$ содержится в $\mathrm{GL}_n(P)$. Если A и B – две матрицы из $\mathrm{O}_n(P)$, то $A \cdot A^t = E$ и $B \cdot B^t = E$. Произведение $A \cdot B$ лежит в $\mathrm{O}_n(P)$, если

$$(AB) \cdot (AB)^t = E.$$

Обратная матрица A^{-1} лежит в $\mathrm{O}_n(P)$, если

$$(A^{-1}) \cdot (A^{-1})^t = E.$$

По лемме 1

$$(A\,B) \cdot (A\,B)^t = (A\,B) \cdot (B^t\,A^t) = A \cdot (B\,B^t)\,A^t = E.$$

Следовательно, множество $\mathrm{O}_n(P)$ замкнуто относительно умножения.

Для доказательства замкнутости относительно взятия обратного воспользуемся леммой 2. Таким образом, множество $\mathrm{O}_n(P)$ является подгруппой. Теорема доказана.

У п р а ж н е н и е. Докажите, что U_n является подгруппой группы $\mathrm{GL}_n(\mathbb{C})$.

Определенные нами группы носят специальные названия. Группа $\mathrm{GL}_n(P)$ называется *общей линейной группой*, группа $\mathrm{SL}_n(P)$ – *специальной линейной группой*, группа $\mathrm{O}_n(P)$ – *ортогональной линейной группой* и U_n – *унитарной линейной группой.*

Глава 2

ВЕКТОРНЫЕ ПРОСТРАНСТВА

§ 9. Векторное пространство над полем

9.1. Аксиоматика. Пусть задано поле P, элементы которого будем называть *скалярами,* и некоторое множество V, элементы которого будем называть *векторами.* На множестве векторов определена бинарная операция сложения $+$, сопоставляющая паре векторов некоторый вектор. Относительно операции сложения V является абелевой группой. Кроме того, определена функция

$$f : P \times V \longrightarrow V, \quad (\alpha, v) \longrightarrow \alpha \cdot v.$$

Будем обозначать значение $\alpha \cdot v$ символом $\alpha\, v$ и называть *произведением вектора на скаляр.* Предполагается, что выполнены следующие аксиомы:

$$\alpha\,(u + v) = \alpha\, u + \alpha\, v,$$

$$(\alpha + \beta)\, u = \alpha\, u + \beta\, u,$$

$$(\alpha\, \beta)\, u = \alpha\,(\beta\, u),$$

$$1 \cdot u = u$$

для любых $\alpha, \beta \in P$ и $u, v \in V$. В этом случае мы говорим, что задано *векторное пространство над полем* P.

Схематически векторное пространство можно представить следующим образом:

$$f : P \times V \longrightarrow V$$

$+, \cdot$	$+$
$P = \{\alpha, \beta, \gamma, \ldots\}$	$V = \{u, v, w, \ldots\}$

Из аксиом векторного пространства выведем некоторые следствия.

С л е д с т в и е 1. *Для всякого вектора v из V справедливо равенство*

$$0 \cdot v = 0,$$

где в правой части 0 означает нулевой вектор, а в левой – нулевой элемент поля скаляров.

Действительно,

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

Прибавив к обеим частям этого равенства элемент, противоположный к $0 \cdot v$, получим $0 \cdot v = 0$.

С л е д с т в и е 2. *Для всякого вектора v из V справедливо равенство*

$$-v = (-1) \cdot v.$$

Действительно,

$$v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1 - 1) \cdot v = 0 \cdot v = 0,$$

т. е. $(-1) \cdot v$ является противоположным к элементу v.

9.2. Векторное пространство как алгебраическая система. Чтобы представить векторное пространство как алгебраическую систему, выберем в качестве основного множества множество V и определим на нем, помимо операции сложения векторов, множество унарных операций f_α для каждого скаляра $\alpha \in P$, положив $f_\alpha(u) = \alpha u$ для любого вектора $u \in V$. Тогда векторное пространство является алгебраической системой:

$$\langle V; +, f_\alpha \ (\alpha \in P) \rangle.$$

Из аксиом векторного пространства следует, что операции f_α удовлетворяют следующим условиям:

$$1)\quad f_\alpha(u+v) = f_\alpha(u) + f_\alpha(v),$$

$$2)\quad f_{\alpha+\beta}(u) = f_\alpha(u) + f_\beta(u),$$

$$3)\quad f_{\alpha\beta}(u) = f_\alpha(f_\beta(u)),$$

$$4)\quad f_1(u) = u$$

для любых $\alpha, \beta \in P$ и $u, v \in V$.

9.3. Изоморфизм векторных пространств. Пусть заданы два векторных пространства

$$\langle V; +,\ f_\alpha\ (\alpha \in P)\rangle, \quad \langle V'; +,\ f_\alpha\ (\alpha \in P)\rangle$$

над одним и тем же полем P. Говорят, что они *изоморфны*, если существует взаимно однозначное отображение $\varphi : V \longrightarrow V'$ пространства V *на* пространство V', сохраняющее операции, т. е.

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

$$(f_\alpha(u))\varphi = f_\alpha(u\varphi).$$

Последнее равенство эквивалентно такому:

$$(\alpha u)\varphi = \alpha(u\varphi).$$

9.4. Примеры векторных пространств. 1) В качестве поля P возьмем поле вещественных чисел $\mathbb{R}$ с обычными операциями сложения и умножения. В качестве V – множество направленных отрезков на плоскости. Вектор αu является произведением вектора u на скаляр α.

2) В качестве множества векторов возьмем множество матриц $M_n(P)$ с операцией сложения, а в качестве поля – поле P. Множество $M_n(P)$ является векторным пространством над P, если для матрицы $A = (a_{ij})$ произведение $\alpha\, A$ есть матрица $(\alpha\, a_{ij})$.

3) Пусть P – произвольное поле. В качестве V возьмем множество многочленов $P[x]$ с коэффициентами из P от одной переменной x, а в качестве умножения на скаляр – умножение многочлена на элемент из P.

4) Следующий пример является в некотором смысле универсальным. Более точно, всякое конечномерное векторное пространство над полем P изоморфно этому векторному пространству. В качестве V возьмем множество P^n – множество всех упорядоченных последовательностей n элементов из P, т. е.

$$P^n = \{v = (\alpha_1, \alpha_2, \ldots, \alpha_n) \,|\, \alpha_i \in P\}\,.$$

На множестве таких последовательностей определим операцию сложения:

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

и операцию умножения на скаляр:

$$\gamma\, v = \gamma\,(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\gamma\,\alpha_1, \gamma\,\alpha_2, \ldots, \gamma\,\alpha_n), \quad \gamma \in P.$$

5) Поле как векторное пространство. Пусть L – некоторое поле, а l – его подполе. В качестве P возьмем l, а в качестве V – поле L. Операция умножения скаляра на вектор – обычное произведение в поле L. В частности, поле комплексных чисел $\mathbb{C}$ является векторным пространством над полем $\mathbb{R}$.

6) Множество функций

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

с операциями поточечного сложения и умножения на элементы из $\mathbb{R}$ образует векторное пространство над $\mathbb{R}$.

У п р а ж н е н и е. Проверьте, что в каждом из примеров выполнены все аксиомы векторного пространства.

У п р а ж н е н и е. Векторное пространство из пункта 1 изоморфно пространству $\mathbb{R}^2$. Векторное пространство из пункта 2 изоморфно пространству P^{n^2}. Множество комплексных чисел $\mathbb{C}$, рассматриваемых как векторное пространство, изоморфно пространству $\mathbb{R}^2$.

§ 10. Линейная зависимость векторов

10.1. Линейная комбинация. Пусть задано векторное пространство V над полем P и пусть $u_1, u_2, \ldots, u_s$ – система векторов (в системе, в отличие от множества, могут быть одинаковые векторы) из V, а $\alpha_1, \alpha_2, \ldots, \alpha_s$ – некоторые скаляры из поля P. Вектор

$$\alpha_1\, u_1 + \alpha_2\, u_2 + \ldots + \alpha_s\, u_s$$

называется *линейной комбинацией* векторов $u_1, u_2, \ldots, u_s$ с коэффициентами α_1, α_2, $\ldots$, α_s. Если $\alpha_1 = \alpha_2 = \ldots = \alpha_s = 0$, то линейная комбинация называется *тривиальной*. Очевидно, тривиальная линейная комбинация дает тривиальный вектор.

Говорят, что вектор u из V *линейно выражается через* векторы $u_1, u_2, \ldots, u_s$ из V, если

$$u = \beta_1 u_1 + \beta_2 u_2 + \ldots + \beta_s u_s$$

при подходящих $\beta_1, \beta_2, \ldots, \beta_s \in P$. Говорят, что вектор u *линейно выражается через бесконечную систему векторов*, если он линейно выражается через некоторую ее конечную подсистему.

Л е м м а 1. *Пусть V – векторное пространство над полем P, $u_1, u_2, \ldots, u_s \in V$, $2 \leq s < \infty$. Тогда следующие утверждения равносильны*:

1) *существует нетривиальная линейная комбинация векторов $u_1, u_2, \ldots, u_s$, равная нулю*;

2) *хотя бы один вектор u_{i_0}, $i_0 \in \{1, 2, \ldots, s\}$ линейно выражается через остальные.*

Д о к а з а т е л ь с т в о. 1) $\Rightarrow$ 2). Пусть

$$\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_s u_s = 0,$$

где не все α_i равны нулю. Возьмем $\alpha_{i_0} \neq 0$, тогда

$$\alpha_{i_0} u_{i_0} = -\alpha_1 u_1 - \alpha_2 u_2 - \ldots - \alpha_{i_0-1} u_{i_0-1} - \alpha_{i_0+1} u_{i_0+1} - \ldots - \alpha_s u_s.$$

Обе части этого равенства умножим на элемент, обратный к α_{i_0}, получим:

$$u_{i_0} = -\frac{\alpha_1}{\alpha_{i_0}} u_1 - \frac{\alpha_2}{\alpha_{i_0}} u_2 - \ldots - \frac{\alpha_{i_0-1}}{\alpha_{i_0}} u_{i_0-1} - \frac{\alpha_{i_0+1}}{\alpha_{i_0}} u_{i_0+1} - \ldots - \frac{\alpha_s}{\alpha_{i_0}} u_s.$$

2)$\Rightarrow$ 1). Предположим, что

$$u_{i_0} = \gamma_1 u_1 + \gamma_2 u_2 + \ldots + \gamma_{i_0-1} u_{i_0-1} + \gamma_{i_0+1} u_{i_0+1} + \ldots + \gamma_s u_s$$

для некоторого $u_{i_0} \neq 0$. Перенесем все члены в одну часть:

$$0 = \gamma_1 u_1 + \gamma_2 u_2 + \ldots + \gamma_{i_0-1} u_{i_0-1} - u_{i_0} + \gamma_{i_0+1} u_{i_0+1} + \ldots + \gamma_s u_s.$$

Ввиду следствия 2 из аксиом векторного пространства $-u_{i_0} = (-1)u_{i_0}$, а потому полученная линейная комбинация нетривиальна. Лемма доказана.

О п р е д е л е н и е. а) Система, состоящая из одного вектора, называется *линейно зависимой*, если этот вектор равен нулю. б) Конечная система $u_1, u_2, \ldots, u_s$, состоящая более чем из одного вектора, называется *линейно зависимой*, если выполнено одно из условий леммы 1. в) Бесконечная система векторов называется *линейно зависимой*, если линейно зависима некоторая ее конечная подсистема.

10.2. Линейно эквивалентные системы. Говорят, что одна система векторов *линейно выражается* через другую систему векторов, если каждый ее вектор линейно выражается через векторы другой системы. Две системы векторов называются *линейно эквивалентными*, если первая система линейно выражается через вторую и, в свою очередь, вторая система линейно выражается через первую.

Для трех систем векторов справедлива

Л е м м а 2. *Если имеются три системы векторов*

$$u_1, u_2, \ldots, u_r, \tag{1}$$

$$v_1, v_2, \ldots, v_s, \tag{2}$$

$$w_1, w_2, \ldots, w_t \tag{3}$$

и известно, что (1) *линейно выражается через* (2), *а* (2) *линейно выражается через* (3), *то* (1) *линейно выражается через* (3).

Д о к а з а т е л ь с т в о. По условию

$$u_i = \sum_{j=1}^{s} \alpha_{ij}\, v_j, \quad i = 1, 2, \ldots, r, \quad v_j = \sum_{k=1}^{t} \beta_{jk}\, w_k, \quad j = 1, 2, \ldots, s.$$

Тогда

$$u_i = \sum_{j=1}^{s} \alpha_{ij} \left(\sum_{k=1}^{t} \beta_{jk}\, w_k \right) = \sum_{j=1}^{s} \sum_{k=1}^{t} (\alpha_{ij}\, \beta_{jk})\; w_k =$$

$$= \sum_{k=1}^{t} \left(\sum_{j=1}^{s} \alpha_{ij}\, \beta_{jk} \right) w_k = \sum_{k=1}^{t} \gamma_{ik}\, w_k,$$

где $\gamma_{ik} = \sum_{j=1}^{s} \alpha_{ij}\, \beta_{jk}$.

10.3. Теорема о замене. Цель настоящего пункта – доказать следующее утверждение.

Т е о р е м а. *Пусть даны две системы векторов*

$$u_1, u_2, \ldots, u_r, \tag{1}$$

$$v_1, v_2, \ldots, v_s. \tag{2}$$

Причем (1) *линейно независима и линейно выражается через* (2). *Тогда* а) $r \leq s$, б) *найдутся такие* r *векторов в системе* (2), *которые можно заменить на векторы системы* (1), *и полученная система векторов* (2′) *будет линейно эквивалентна системе* (2).

Д о к а з а т е л ь с т в о проведем индукцией по r. Пусть вначале $r = 1$, т. е. система (1) состоит из одного вектора u_1. Так как эта система линейно независима, то $u_1 \neq 0$. Учитывая, что (1) линейно выражается через (2), имеем неравенство $s \geq 1$, т. е. пункт а установлен. Установим пункт б. Так как u_1 линейно выражается через (2), имеем равенство

$$u_1 = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_s v_s, \quad \alpha_i \in P.$$

При этом, ввиду того, что $u_1 \neq 0$, хотя бы один из коэффициентов α_i отличен от нуля. Не уменьшая общности, можно считать, что $\alpha_1 \neq 0$. Тогда

$$v_1 = \frac{1}{\alpha_1} u_1 - \frac{\alpha_2}{\alpha_1} v_2 - \ldots - \frac{\alpha_s}{\alpha_1} v_s.$$

Система векторов

$$u_1, v_2, \ldots, v_s$$

и будет искомой. Действительно, как мы заметили, v_1 линейно выражается через эту систему, а по условию теоремы u_1 линейно выражается через (2).

Пусть теперь $r > 1$. Рассмотрим систему векторов

$$u_1, u_2, \ldots, u_{r-1}. \tag{3}$$

Для систем (3) и (2) выполнены условия теоремы и из предположения индукции имеем:

а) $r - 1 \leq s$,

б) в системе (2) найдутся $r-1$ векторов (пусть они располагаются по порядку), которые можно заменить на векторы системы (3) так, что полученная система

$$u_1, u_2, \ldots, u_{r-1}, v_r, v_{r+1}, \ldots, v_s \tag{4}$$

будет линейно эквивалентна системе (2).

а) Докажем, что $r \leq s$. Предположим противное: $r > s$. Так как по предположению индукции $r-1 \leq s$, то $r = s+1$ и $r-1 = s$. Тогда в (4) векторов $v_r, v_{r+1}, \ldots, v_s$ нет и (4) принимает вид

$$u_1, u_2, \ldots, u_{r-1}.$$

Вектор u_r линейно выражается через (2), а так как (2) линейно эквивалентна (4), то u_r линейно выражается через (4), т. е. через u_1, u_2, $\ldots$, u_{r-1}, но это противоречит тому, что система (1) линейно независима. Следовательно, $r \leq s$.

Докажем утверждение б. По условию вектор u_r линейно выражается через (2), а так как (2) линейно эквивалентна (4), то u_r линейно выражается через (4):

$$u_r = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_{r-1} u_{r-1} + \beta_r v_r + \ldots + \beta_s v_s. \qquad (5)$$

Покажем, что хотя бы один из коэффициентов β_i отличен от нуля. Действительно, если предположить, что все они равны нулю, то

$$u_r = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_{r-1} u_{r-1},$$

т. е. u_r линейно выражается через $u_1, u_2, \ldots, u_{r-1}$, что противоречит условию линейной независимости системы (1).

Можно считать, что $\beta_r \neq 0$, тогда из (5):

$$v_r = \left(-\frac{\alpha_1}{\beta_r}\right) u_1 + \left(-\frac{\alpha_2}{\beta_r}\right) u_2 + \ldots + \left(-\frac{\alpha_{r-1}}{\beta_r}\right) u_{r-1} + \frac{1}{\beta_r} u_r + \\ + \left(-\frac{\beta_{r+1}}{\beta_r}\right) v_{r+1} + \ldots + \left(-\frac{\beta_s}{\beta_r}\right) v_s. \qquad (6)$$

Рассмотрим систему

$$u_1, \ldots, u_{r-1}, u_r, v_{r+1}, \ldots, v_s, \qquad (2')$$

которая получается из (2) удалением r векторов и заменой их векторами системы (1). Покажем, что (2′) линейно эквивалентна (2). Для этого заметим, что (2′) линейно эквивалентна (4). Вектор v_r из (4) линейно выражается через (2′) (см. выражение (6)). Векторы $u_1, \ldots, u_{r-1}, v_{r+1}, \ldots, v_s$ из (4) входят в (2′). Поэтому (4) линейно выражается через (2′). Аналогично с использованием (5) система (2′) линейно выражается через (4), т. е. (2′) линейно эквивалентна (4). По

предположению индукции (4) эквивалентна (2), а так как мы доказали, что (2′) эквивалентна (4), то по лемме 2 система (2′) линейно эквивалентна (2). Теорема доказана.

С л е д с т в и е. *Любые две линейно независимые, линейно эквивалентные системы векторов либо бесконечны, либо конечны и состоят из одного и того же числа векторов.*

Д о к а з а т е л ь с т в о. Пусть

$$u_1, u_2, \ldots, \tag{7}$$

$$v_1, v_2, \ldots \tag{8}$$

– две линейно независимые, линейно эквивалентные системы векторов. Предположим, что система (7) бесконечна. Покажем, что тогда и система (8) бесконечна. Пусть, напротив, (8) является конечной системой:

$$v_1, v_2, \ldots, v_s. \tag{8}$$

Рассмотрим систему векторов:

$$u_1, u_2, \ldots, u_{s+1}, \tag{9}$$

которая является подсистемой системы (7). Она состоит из линейно независимых векторов и каждый ее вектор линейно выражается через систему (8). По теореме о замене $s + 1 \leq s$ – противоречие. Следовательно, система (8) бесконечна.

Пусть теперь обе системы конечны:

$$u_1, u_2, \ldots, u_r,$$

$$v_1, v_2, \ldots, v_s.$$

Опять по теореме о замене $r \leq s$, а так как вторая система также линейно выражается через первую, то и $s \leq r$. Следовательно, $s = r$.

§ 11. База векторного пространства

11.1. Максимальная линейно независимая подсистема. Пусть

$$u_1, u_2, \ldots, u_r, \ldots \tag{1}$$

– система векторов некоторого векторного пространства. *Максимальной линейно независимой подсистемой* назовем подсистему, которая

сама линейно независима, а при добавлении любого вектора системы становится линейно зависимой. Такие подсистемы всегда существуют в любой системе. Если в (1) все векторы нулевые, то максимальная подсистема – это пустая система. Если $u_1 \neq 0$, то подсистема, состоящая из u_1, – линейно независима. Если есть линейно независимая подсистема векторов и она немаксимальна, то можно добавить еще один вектор. Если полученная подсистема немаксимальна, то можно добавить еще один вектор и т. д.

Л е м м а 1. *Любой вектор системы линейно выражается через максимальную линейно независимую подсистему.*

Д о к а з а т е л ь с т в о. Пусть $v_1, v_2, \ldots$ – максимальная линейно независимая подсистема системы (1) и u_i – некоторый вектор из (1). Рассмотрим подсистему

$$u_i, v_1, v_2, \ldots$$

системы (1). Она линейно зависима, т. е. найдется конечная подсистема $v_1, v_2, \ldots, v_s$ такая, что

$$\alpha\, u_i + \beta_1\, v_1 + \beta_2\, v_2 + \ldots + \beta_s\, v_s = 0$$

для некоторых коэффициентов $\alpha, \beta_1, \ldots, \beta_s$. Заметим, что $\alpha \neq 0$. Действительно, если $\alpha = 0$, то

$$\beta_1\, v_1 + \beta_2\, v_2 + \ldots + \beta_s\, v_s = 0.$$

Но так как векторы $v_1, v_2, \ldots, v_s$ линейно независимы, то $\beta_1 = \ldots = \beta_s = 0$, а это противоречит линейной зависимости. Следовательно, $\alpha \neq 0$, и мы можем выразить

$$u_i = \left(-\frac{\beta_1}{\alpha}\right) v_1 + \ldots + \left(-\frac{\beta_s}{\alpha}\right) v_s.$$

Таким образом, всякий вектор из (1) является линейной комбинацией векторов из максимальной линейно независимой подсистемы. Лемма доказана.

Из этой леммы, в частности, следует, что во всякой системе векторов любые две максимальные линейно независимые подсистемы линейно эквивалентны, а потому по следствию теоремы о замене либо обе бесконечны, либо имеют одинаковое число векторов. Поэтому мы можем дать

О п р е д е л е н и е. Максимальную линейно независимую подсистему называют *базисной подсистемой* системы (1). Число векторов в базисной подсистеме системы (1) называют *рангом системы*

(1). Если в качестве системы (1) рассматривать все векторы векторного пространства V, то ее базисную подсистему называют *базой пространства* V, а ее ранг – *размерностью пространства*. Размерность пространства V будем обозначать символом $\dim V$.

Как было замечено выше, все базы одного пространства имеют одинаковое число векторов.

Вернемся теперь к рассмотренным выше примерам векторных подпространств и укажем в каждом из них базу.

П р и м е р ы. 1) Базу пространства направленных отрезков на плоскости образуют два любых ненулевых непараллельных вектора.

2) Базу пространства $M_n(P)$ образуют так называемые *матричные единицы* $E_{ij} \in M_n(P)$, $1 \le i, j \le n$ – матрицы, у которых на месте (i, j) стоит единица, а на всех остальных местах нули.

3) Пространство многочленов $P[x]$ имеет бесконечную счетную размерность, и база состоит из многочленов: $1, x, x^2, \ldots$

4) База пространства P^n состоит из векторов e_i, $1 \le i \le n$, у которых на i-м месте стоит единица, а на всех остальных местах – нули.

5) Пространство $\mathbb{C}$ является двумерным векторным пространством над $\mathbb{R}$ с базой $1, \imath$.

В дальнейшем мы будем рассматривать только конечномерные пространства.

11.2. Координаты вектора. Пусть V – векторное пространство размерности $n = \dim V$ над полем P. Пусть $e_1, e_2, \ldots, e_n$ – база пространства V. Если $u \in V$, то по лемме 1

$$u = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n, \quad \alpha_i \in P,$$

и элементы $\alpha_1, \alpha_2, \ldots, \alpha_n$ называются *координатами вектора* u в базе $e_1, e_2, \ldots, e_n$. Покажем, что координаты определяются однозначно, т. е. если

$$u = \alpha'_1 e_1 + \alpha'_2 e_2 + \ldots + \alpha'_n e_n, \quad \alpha'_i \in P,$$

то $\alpha'_1 = \alpha_1$, $\alpha'_2 = \alpha_2$, $\ldots$, $\alpha'_n = \alpha_n$. Действительно, рассматривая разность двух представлений, получим

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

и из линейной независимости следуют нужные равенства, т. е. координаты вектора u определяются однозначно. Будем записывать базу

пространства в виде столбца

$$e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix},$$

а координаты вектора u – в виде строки

$$[u] = (\alpha_1, \alpha_2, \ldots, \alpha_n).$$

При этом имеет место равенство

$$[u]\, e = u.$$

Т е о р е м а. *Всякое векторное пространство V размерности n над полем P изоморфно пространству n-мерных строк (n-ок) P^n над P.*

Д о к а з а т е л ь с т в о. Покажем, что существует взаимно однозначное отображение пространства V на пространство n-мерных строк. Для этого зафиксируем в V некоторую базу e и каждому $u \in V$ поставим в соответствие координаты

$$u \longmapsto [u].$$

Это отображение (обозначим его φ) и есть искомый изоморфизм пространства V на пространство строк:

$$P^n = \{(\alpha_1, \alpha_2, \ldots, \alpha_n) \mid \alpha_i \in P\}.$$

Отображение

$$\varphi : V \longrightarrow P^n$$

взаимнооднозначно. Действительно, так как координаты в V определяются однозначно, то φ однозначно. Если $[u] = [v]$, т. е. векторы u и v имеют одни и те же координаты, то они равны. Отображение φ является отображением *на*, так как для любой строки $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ существует вектор $u = \alpha_1\, e_1 + \alpha_2\, e_2 + \ldots + \alpha_n\, e_n$ из V, являющийся прообразом. Чтобы показать, что φ сохраняет операции, надо доказать:

1) $(u + v)\varphi = u\varphi + v\varphi \Longleftrightarrow [u + v] = [u] + [v]$;
2) $(\alpha\, u)\varphi = \alpha\, u\varphi \Longleftrightarrow [\alpha\, u] = \alpha\, [u]$.

Для доказательства 1 положим

$$[u] = (\alpha_1, \alpha_2, \ldots, \alpha_n), \quad [v] = (\beta_1, \beta_2, \ldots, \beta_n).$$

Тогда сумма

$$u + v = (\alpha_1 + \beta_1)\, e_1 + (\alpha_2 + \beta_2)\, e_2 + \ldots + (\alpha_n + \beta_n)\, e_n$$

имеет координаты

$$[u + v] = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n) = [u] + [v].$$

Равенство 2 проверяется аналогично. Теорема доказана.

11.3. Преобразование координат вектора при смене базы. В каждом пространстве V существует много баз. Пусть имеются две базы:

$$e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}, \qquad e' = \begin{pmatrix} e'_1 \\ e'_2 \\ \vdots \\ e'_n \end{pmatrix}.$$

Как они связаны?

Любой вектор из e' линейно выражается через базу e:

$$\begin{cases} e'_1 = t_{11}\, e_1 + t_{12}\, e_2 + \ldots + t_{1n}\, e_n, \\ e'_2 = t_{21}\, e_1 + t_{22}\, e_2 + \ldots + t_{2n}\, e_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ e'_n = t_{n1}\, e_1 + t_{n2}\, e_2 + \ldots + t_{nn}\, e_n. \end{cases} \tag{2}$$

Соберем коэффициенты в матрицу

$$T = (t_{ij}) = \begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix}.$$

Она называется *матрицей перехода от* базы e к базе e'. При этом систему (2) можно записать в таком виде:

$$e' = T\, e. \tag{3}$$

Л е м м а 2. *Матрица перехода T – невырожденная матрица.*

Д о к а з а т е л ь с т в о. Найдем матрицу перехода от e' к e, которую обозначим символом S, т. е.

$$e = S\,e'.$$

Подставляя это выражение в равенство (3), получим

$$e' = E\,e' = T\,e = T\,(S\,e').$$

Для завершения доказательства надо показать, что из последнего равенства следует равенство $E = T\,S$. Из этого равенства ввиду теоремы об обратимости матрицы (см. теорему 2 из § 6) будет следовать, что матрица T невырождена.

Справедливость равенства $E = T\,S$ вытекает из следующего утверждения.

Л е м м а 3 (о сокращении). *Пусть $v_1, v_2, \ldots, v_k$ – линейно независимая система векторов, X и Y – матрицы размера $m \times k$. Если положить $v = (v_1, v_2, \ldots, v_k)^t$ – вектор-столбец, то из равенства $X\,v = Y\,v$ следует равенство матриц $X = Y$.*

Д о к а з а т е л ь с т в о. Пусть

$$X = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1k} \\ x_{21} & x_{22} & \ldots & x_{2k} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ x_{m1} & x_{m2} & \ldots & x_{mk} \end{pmatrix}, \quad Y = \begin{pmatrix} y_{11} & y_{12} & \ldots & y_{1k} \\ y_{21} & y_{22} & \ldots & y_{2k} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ y_{m1} & y_{m2} & \ldots & y_{mk} \end{pmatrix}.$$

Тогда равенство $X\,v = Y\,v$ запишется в виде:

$$\begin{pmatrix} x_{11}\,v_1 + x_{12}\,v_2 + \ldots + x_{1k}\,v_k \\ x_{21}\,v_1 + x_{22}\,v_2 + \ldots + x_{2k}\,v_k \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_{m1}\,v_1 + x_{m2}\,v_2 + \ldots + x_{mk}\,v_k \end{pmatrix} = \begin{pmatrix} y_{11}\,v_1 + y_{12}\,v_2 + \ldots + y_{1k}\,v_k \\ y_{21}\,v_1 + y_{22}\,v_2 + \ldots + y_{2k}\,v_k \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_{m1}\,v_1 + y_{m2}\,v_2 + \ldots + y_{mk}\,v_k \end{pmatrix},$$

или

$$\begin{cases} x_{11}\,v_1 + x_{12}\,v_2 + \ldots + x_{1k}\,v_k = y_{11}\,v_1 + y_{12}\,v_2 + \ldots + y_{1k}\,v_k, \\ x_{21}\,v_1 + x_{22}\,v_2 + \ldots + x_{2k}\,v_k = y_{21}\,v_1 + y_{22}\,v_2 + \ldots + y_{2k}\,v_k, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_{m1}\,v_1 + x_{m2}\,v_2 + \ldots + x_{mk}\,v_k = y_{m1}\,v_1 + y_{m2}\,v_2 + \ldots + y_{mk}\,v_k. \end{cases}$$

Перенося все слагаемые в левую часть, получим

$$\begin{cases} (x_{11} - y_{11})\,v_1 + (x_{12} - y_{12})\,v_2 + \ldots + (x_{1k} - y_{1k})\,v_k = 0, \\ (x_{21} - y_{21})\,v_1 + (x_{22} - y_{22})\,v_2 + \ldots + (x_{2k} - y_{2k})\,v_k = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ (x_{m1} - y_{m1})\,v_1 + (x_{m2} - y_{m2})\,v_2 + \ldots + (x_{mk} - y_{mk})\,v_k = 0. \end{cases}$$

Так как по условию система векторов $v_1, v_2, \ldots, v_k$ линейно независима, то $x_{ij} = y_{ij}$. Лемма установлена.

Если вектор u в базе e имеет координаты $[u] = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, а в базе e' – координаты $[u]' = (\alpha'_1, \alpha'_2, \ldots, \alpha'_n)$, то

$$u = [u]\, e = [u]'\, e' = [u]'\, T\, e.$$

Следовательно, по лемме о сокращении $[u]' = [u]\, T^{-1}$. Таким образом, координаты вектора в новой базе есть произведение координат в старой базе на обратную к матрице перехода.

§ 12. Подпространства

12.1. Определения и примеры. Подмножество U векторного пространства V называется *подпространством*, если оно замкнуто относительно сложения и умножения на скаляр и само является векторным пространством относительно этих индуцированных операций.

Следующая очевидная лемма позволяет проверить, будет ли данное множество подпространством.

Л е м м а 1. *Пусть V – векторное подпространство над полем P. Для подмножества $U \subseteq V$ равносильны следующие утверждения*:

1) *U – подпространство пространства V*;

2) *для любых $u, v \in U$ и $\alpha \in P$ сумма $u + v$ и произведение αu* лежат в U;

3) *для любых $u, v \in U$ и $\alpha, \beta \in P$ линейная комбинация $\alpha\, u + \beta\, v$* лежит в U.

Укажем в каждом из разобранных в примерах векторных пространств некоторые подпространства.

П р и м е р ы. 1) Множество направленных отрезков на плоскости, параллельных некоторой фиксированной прямой, является подпространством пространства направленных отрезков на плоскости.

2) Матрица $A \in M_n(P)$ называется *симметрической* (*кососимметрической*), если $A^t = A$ (соответственно $A^t = -A$). Множество симметрических (кососимметрических) матриц является подпространством пространства $M_n(P)$.

3) Если V – пространство многочленов над полем P, а

$$U = \{f(x) = a_1\, x + a_2\, x^2 + \ldots + a_n\, x^n \mid n \in \mathbb{N}, a_i \in P\}$$

– множество многочленов без свободного члена, то U – подпространство пространства V.

4) Множество n-ок из P^n, у которых первая координата нулевая, образует подпространство пространства P^n, которое изоморфно P^{n-1}.

5) Множество *гауссовых чисел*

$$\{a + i\,b \mid a, b \in \mathbb{Q}\}$$

образует подмножество, но не подпространство пространства комплексных чисел $\mathbb{C}$ над полем $\mathbb{R}$.

6) Множество непрерывных функций $g : \mathbb{R} \longrightarrow \mathbb{R}$ образует подпространство пространства всех функций $f : \mathbb{R} \longrightarrow \mathbb{R}$.

12.2. Сумма и пересечение подпространств. Пусть U_i, $i \in I$ – некоторое семейство подпространств пространства V. Их *суммой* называется следующее множество векторов:

$$\sum_{i\in I} U_i = \{u_{i_1} + u_{i_2} + \ldots + u_{i_s} \mid u_{i_j} \in U_{i_j}\}.$$

Если, в частности, $I = \{1, 2, \ldots, m\}$ – конечное множество, то

$$\sum_{i=1}^{m} U_i = \{u_1 + u_2 + \ldots + u_m \mid u_i \in U_i\}.$$

Л е м м а 2. *Сумма подпространств – подпространство векторного пространства.*

Д о к а з а т е л ь с т в о достаточно провести для случая конечного числа подпространств. Рассмотрим два вектора

$$u = u_1 + u_2 + \ldots + u_m, \quad v = v_1 + v_2 + \ldots + v_m, \quad u_i, v_i \in U_i,$$

из суммы $\sum_{i=1}^{m} U_i$ и возьмем $\alpha, \beta \in P$. Рассмотрим линейную комбинацию

$$\alpha\,u + \beta\,v = (\alpha\,u_1 + \beta\,v_1) + (\alpha\,u_2 + \beta\,v_2) + \ldots + (\alpha\,u_m + \beta\,v_m) =$$

$$= w_1 + w_2 + \ldots + w_m,$$

где символом w_i обозначена линейная комбинация $\alpha\,u_i + \beta\,v_i$, $i = 1, 2, \ldots, m$. Заметим, что $\alpha\,u_1 + \beta\,v_1$ лежит в подпространстве U_1, $\alpha\,u_2 + \beta\,v_2$ лежит в подпространстве U_2 и т. д. Следовательно, сумма

$w_1 + w_2 + \ldots + w_m$ лежит в $\sum_{i=1}^m U_i$, а потому $\sum_{i=1}^m U_i$ – подпространство. Лемма доказана.

Введем теперь операцию пересечения для подпространств. Пусть опять U_i, $i \in I$ – семейство подпространств пространства V. Их *пересечением* назовем следующее множество векторов:

$$\bigcap_{i\in I} U_i = \{v \in V \mid v \in U_i \text{ для всех } i \in I\}.$$

Заметим, что это множество непусто. Оно, в частности, содержит нулевой вектор, так как его содержит каждое подпространство U_i. Покажем, что пересечение $\bigcap_{i\in I} U_i$ является подпространством.

Л е м м а 3. *Пересечение подпространств – подпространство векторного пространства.*

Д о к а з а т е л ь с т в о. Пусть $u, v \in \bigcap_{i\in I} U_i$ и $\alpha, \beta \in P$. Рассмотрим линейную комбинацию $\alpha\, u + \beta\, v$, которая лежит в U_i для каждого $i \in I$. Следовательно, $\alpha\, u + \beta\, v$ лежит в пересечении подпространств. Таким образом, пересечение подпространств является подпространством.

12.3. Линейная оболочка. Пусть M – некоторое подмножество V. *Линейной оболочкой* $\mathcal{L}(M)$ множества M называется пересечение всех подпространств пространства V, содержащих M. Иными словами, линейная оболочка – это наименьшее подпространство, содержащее M.

Т е о р е м а 1. *Справедливо следующее равенство:*

$$\mathcal{L}(M) = \{\alpha_1\, m_1 + \alpha_2\, m_2 + \ldots + \alpha_k\, m_k \mid m_i \in M, \quad \alpha_i \in P\},$$

т. е. линейная оболочка – множество всевозможных линейных комбинаций из M.

Д о к а з а т е л ь с т в о. Обозначим

$$\overline{M} = \{\alpha_1\, m_1 + \alpha_2\, m_2 + \ldots + \alpha_k\, m_k \mid m_i \in M, \quad \alpha_i \in P\}.$$

Пусть U_i, $i \in I$ – все подпространства, содержащие множество M. По определению $\mathcal{L}(M) = \bigcap_{i\in I} U_i$. Очевидно, $U_i \supseteq \overline{M}$ для любого $i \in I$, а потому $\mathcal{L}(M) \supseteq \overline{M}$.

Для доказательства обратного включения покажем, что $\overline{M}$ – подпространство. Для этого возьмем два произвольных вектора

$$u = \alpha_1\, m_1 + \alpha_2\, m_2 + \ldots + \alpha_k\, m_k, \quad v = \alpha_{k+1}\, m_{k+1} + \alpha_{k+2}\, m_{k+2} + \ldots + \alpha_s\, m_s$$

из $\overline{M}$ и два произвольных скаляра $\alpha, \beta \in P$. Найдем их линейную комбинацию:

$$\alpha\, u + \beta\, v = \alpha\,(\alpha_1\, m_1 + \alpha_2\, m_2 + \ldots + \alpha_k\, m_k) +$$
$$+\beta\,(\alpha_{k+1}\, m_{k+1} + \alpha_{k+2}\, m_{k+2} + \ldots + \alpha_s\, m_s) =$$
$$= (\alpha\,\alpha_1)\, m_1 + (\alpha\,\alpha_2)\, m_2 + \ldots + (\alpha\,\alpha_k)\, m_k +$$
$$+(\beta\,\alpha_{k+1})\, m_{k+1} + (\beta\,\alpha_{k+2})\, m_{k+2} + \ldots + (\beta\,\alpha_s)\, m_s.$$

Видим, что этот вектор лежит в $\overline{M}$.

Очевидно, что $\overline{M} \supseteq M$, т. е. $\overline{M}$ совпадает с одним из подпространств U_i. Следовательно, $\overline{M} \supseteq \bigcap_{i\in I} U_i$. Теорема доказана.

12.4. Прямая сумма подпространств. Сумма подпространств $\sum_{i\in I} U_i$ называется *прямой суммой*, если для каждого ее вектора

$$u = u_{i_1} + u_{i_2} + \ldots + u_{i_s}, \quad u_{i_j} \in U_{i_j}$$

такая запись единственна. Прямая сумма подпространств U_i, $i \in I$ обозначается $\oplus_{i\in I} U_i$. Если $I = \{1, 2, \ldots, m\}$ – конечное множество, то пишут $U_1 \oplus U_2 \oplus \ldots \oplus U_m$.

Л е м м а 4. *Сумма подпространств $\sum_{i\in I} U_i$ является прямой суммой тогда и только тогда, когда*

$$U_i \cap \sum_{j\neq i} U_j = \{0\} \text{ для всех } i \in I.$$

Д о к а з а т е л ь с т в о. Предположим, что сумма прямая, но равенство не выполняется, т. е. существует индекс $i_0 \in I$ такой, что

$$U_{i_0} \cap \sum_{j\neq i_0} U_j \neq \{0\}.$$

Пусть $v \neq 0$ и $v \in U_{i_0} \cap \sum_{j\neq i_0} U_j$, тогда $v = v_{j_1} + v_{j_2} + \ldots + v_{j_s}$, где $j_1, j_2, \ldots, j_s \neq i_0$. Рассмотрим

$$-v + v_{j_1} + v_{j_2} + \ldots + v_{j_s} = 0 = 0 + 0 + \ldots + 0,$$

т. е. нулевой вектор записывается двумя разными способами. Противоречие.

Докажем обратное утверждение. Предположим, что сумма непрямая, т. е. найдется вектор, который записывается двумя способами:

$$w = u_{i_1} + u_{i_2} + \ldots + u_{i_s} = u'_{i_1} + u'_{i_2} + \ldots + u'_{i_s}, \quad u_{i_j}, u'_{i_j} \in U_{i_j}.$$

Будем считать, что $u_{i_1} \neq u'_{i_1}$ и оба вектора лежат в одном подпространстве U_{i_1}. Как этого добиться, можно заметить из следующего примера. Если

$$w = u_2 + u_5 = u'_3 + u'_7,$$

то это равенство можно представить в таком виде:

$$w = u_2 + 0 + u_5 + 0 = 0 + u'_3 + 0 + u'_7.$$

Далее имеем равенство

$$0 \neq u_{i_1} - u'_{i_1} = (u'_{i_2} - u_{i_2}) + (u'_{i_3} - u_{i_3}) + \ldots + (u'_{i_s} - u_{i_s}).$$

Заметим, что $u_{i_1} - u'_{i_1} \in U_{i_1}$, $u'_{i_2} - u_{i_2} \in U_{i_2}$, $\ldots$, $u'_{i_s} - u_{i_s} \in U_{i_s}$. Таким образом, вектор из одного пространства разложен в сумму векторов из других пространств, но по предположению пересечение этих подпространств равно нулевому вектору. Следовательно, вопреки предположению $u_{i_1} = u'_{i_1}$. Лемма доказана.

С л е д с т в и е. 1) *Сумма двух подпространств $U_1 + U_2$ пространства V прямая, если $U_1 \cap U_2 = \{0\}$.* 2) *Сумма трех подпространств $U_1 + U_2 + U_3$ пространства V прямая, если*

$$U_1 \cap (U_2 + U_3) = \{0\}, \quad U_2 \cap (U_1 + U_3) = \{0\}, \quad U_3 \cap (U_1 + U_2) = \{0\}.$$

Легко указать примеры, показывающие, что условия в пункте 2 нельзя заменить условием $U_i \cap U_j = \{0\}$ для всех $i \neq j$.

12.5. Размерность суммы и пересечения подпространств. В этом пункте мы установим формулу, связывающую размерность суммы и пересечения двух подпространств. Для этого нам потребуется

Л е м м а 5. *Любую линейно независимую систему векторов можно дополнить до базы пространства.*

Д о к а з а т е л ь с т в о. Пусть $u_1, u_2, \ldots, u_r$ – линейно независимая система векторов и $v_1, v_2, \ldots, v_n$ – база пространства. По теореме о замене можем заменить r векторов базы на векторы $u_1, u_2, \ldots, u_r$. Получим систему векторов

$$u_1, u_2, \ldots, u_r, v_{r+1}, v_{r+2}, \ldots, v_n.$$

Утверждается, что это база всего пространства. Действительно, она линейно эквивалентна базе пространства. Значит, любой вектор выражается через эту систему. Если бы полученная система была линейно зависимой, то, после удаления линейно зависимых векторов,

получили бы базу, содержащую меньше n векторов, но это невозможно, так как база пространства содержит n векторов. Лемма доказана.

Т е о р е м а 2. *Пусть U_1 и U_2 – подпространства пространства V, тогда справедливо равенство*

$$\dim(U_1 + U_2) = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2).$$

В частности, для прямой суммы справедливо равенство

$$\dim(U_1 \oplus U_2) = \dim(U_1) + \dim(U_2).$$

Д о к а з а т е л ь с т в о. Обозначим $D = U_1 \cap U_2$, $S = U_1 + U_2$, и пусть

$$e_1, e_2, \ldots, e_r \tag{1}$$

– база D.

По лемме 5 существует база

$$e_1, e_2, \ldots, e_r, f_{r+1}, f_{r+2}, \ldots, f_s \tag{2}$$

пространства U_1 и база

$$e_1, e_2, \ldots, e_r, g_{r+1}, g_{r+2}, \ldots, g_t \tag{3}$$

пространства U_2. Рассмотрим систему векторов:

$$e_1, e_2, \ldots, e_r, f_{r+1}, f_{r+2}, \ldots, f_s, g_{r+1}, g_{r+2}, \ldots, g_t. \tag{4}$$

Если мы докажем, что это база S, то отсюда и последует нужная формула.

Докажем, что система (4) линейно независима. Пусть

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_r e_r + \beta_{r+1} f_{r+1} + \beta_{r+2} f_{r+2} + \ldots + \beta_s f_s +$$
$$+\gamma_{r+1} g_{r+1} + \gamma_{r+2} g_{r+2} + \ldots + \gamma_t g_t = 0 \tag{5}$$

для некоторых α_i, β_j, γ_k из P. Надо доказать, что $\alpha_i = 0, \beta_j = 0, \gamma_k = 0$. Перенося часть слагаемых в правую часть, получим

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_r e_r + \beta_{r+1} f_{r+1} + \beta_{r+2} f_{r+2} + \ldots + \beta_s f_s =$$
$$= -\gamma_{r+1} g_{r+1} - \gamma_{r+2} g_{r+2} - \ldots - \gamma_t g_t.$$

Заметим, что вектор из левой части лежит в U_1, а вектор из правой части – в U_2. Следовательно, он лежит в D, т. е. в пересечении подпространств, а потому разлагается по базе (1), т. е. имеем равенство

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_r e_r + \beta_{r+1} f_{r+1} + \beta_{r+2} f_{r+2} + \ldots + \beta_s f_s =$$

$$= \delta_1 e_1 + \delta_2 e_2 + \ldots + \delta_r e_r$$

для некоторых δ_l из P. Отсюда

$$\delta_1 e_1 + \delta_2 e_2 + \ldots + \delta_r e_r + \gamma_{r+1} g_{r+1} + \gamma_{r+2} g_{r+2} + \ldots + \gamma_t g_t = 0,$$

но это линейная комбинация базисных векторов пространства U_2. Следовательно, все $\gamma_k = 0$, а так как система векторов (2) линейно независима, то и все $\alpha_i = 0$. Тогда из (5) следует, что и все $\beta_j = 0$.

Докажем, что система векторов (4) является максимальной. Рассмотрим произвольный вектор $w \in S$. По определению $w = u_1 + u_2$, где $u_1 \in U_1$, $u_2 \in U_2$. Так как $u_1 \in U_1$, то он разлагается по базе пространства U_1:

$$u_1 = \mu_1 e_1 + \mu_2 e_2 + \ldots + \mu_r e_r + \nu_{r+1} f_{r+1} + \nu_{r+2} f_{r+2} + \ldots + \nu_s f_s.$$

Аналогично u_2 разлагается по базе пространства U_2:

$$u_2 = \mu'_1 e_1 + \mu'_2 e_2 + \ldots + \mu'_r e_r + \lambda_{r+1} g_{r+1} + \lambda_{r+2} g_{r+2} + \ldots + \lambda_t g_t.$$

Следовательно,

$$w = (\mu_1+\mu'_1) e_1+(\mu_2+\mu'_2) e_2+\ldots+(\mu_r+\mu'_r) e_r+\nu_{r+1} f_{r+1}+\nu_{r+2} f_{r+2}+\ldots$$

$$+\nu_s f_s + \lambda_{r+1} g_{r+1} + \lambda_{r+2} g_{r+2} + \ldots + \lambda_t g_t.$$

Таким образом, w линейно выражается через систему (4). Теорема доказана.

§ 13. Ранг матрицы

13.1. Теорема о ранге. Всякую матрицу $A \in M_{m\times n}(P)$ можно рассматривать либо как систему строк, либо как систему столбцов:

$$A = (A^1, A^2, \ldots, A^n) = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{pmatrix},$$

где A^i, $i = 1, 2, \ldots, n$ – i-й столбец матрицы A, а A_j, $j = 1, 2, \ldots, m$ – j-я строка матрицы A.

О п р е д е л е н и е. *Рангом матрицы* называется ранг системы ее столбцов, т. е. в системе столбцов выбирается максимальная линейно независимая подсистема, число столбцов которой и называется рангом матрицы.

Если в некоторой матрице выбрать k столбцов и такое же число строк, то определитель полученной на пересечении матрицы, называется *минором порядка k*. *Базисный минор матрицы* – это такой отличный от нуля минор, что все миноры большего порядка равны нулю.

Т е о р е м а о р а н г е. *Ранг матрицы равен порядку ее базисного минора.*

Д о к а з а т е л ь с т в о. В матрице возьмем базисный минор. Предположим, что он находится в левом верхнем углу. Если это не так, то, переставляя столбцы, а затем строки (очевидно, что это преобразование не меняет ранг системы столбцов), приведем матрицу к нужному виду.

Пусть D – базисный минор, r – его порядок. Утверждается, что первые r столбцов составляют максимальную линейно независимую подсистему системы столбцов. Докажем линейную независимость. От противного: если бы первые r столбцов были линейно зависимы, то и столбцы матрицы D были бы линейно зависимы, и по свойству определителей $\det D = 0$, но это противоречит тому, что D – базисный минор.

Докажем максимальность. Для этого надо доказать, что всякий j-й столбец, $j > r$, линейно выражается через первые r столбцов. Рассмотрим i-ю строку. Построим минор на пересечении первых r столбцов и j-го столбца и первых r строк и i-й строки. Это минор порядка $r + 1$:

$$\left|\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1r} & a_{1j} \\ a_{21} & a_{22} & \dots & a_{2r} & a_{2j} \\ \dots\dots\dots\dots & & & \dots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rr} & a_{rj} \\ \hline a_{i1} & a_{i2} & \dots & a_{ir} & a_{ij} \end{array}\right| .$$

Он равен нулю. Действительно, если $i > r$, то это следует из того, что D – базисный минор, а если $i \le r$, то данный минор содержит две одинаковые строки.

Подсчитаем этот минор разлагая по последней строке

$$a_{i1}\, B_{1j} + a_{i2}\, B_{2j} + \ldots + a_{ir}\, B_{rj} + a_{ij}\, D = 0,$$

где $B_{1j}, B_{2j}, \ldots, B_{rj}$ – алгебраические дополнения к элементам a_{i1}, a_{i2}, $\ldots$, a_{ir} соответственно. Из этого равенства получим

$$a_{ij} = -\frac{B_{1j}}{D}\, a_{i1} - \frac{B_{2j}}{D}\, a_{i2} - \ldots - \frac{B_{rj}}{D}\, a_{ir}, \quad i = 1, 2, \ldots, m,$$

причем определители $B_{1j}, B_{2j}, \ldots, B_{rj}$ от i не зависят. Значит, для столбцов есть следующее равенство:

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = -\frac{B_{1j}}{D} \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} - \frac{B_{2j}}{D} \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} - \ldots - \frac{B_{rj}}{D} \begin{pmatrix} a_{1r} \\ a_{2r} \\ \vdots \\ a_{mr} \end{pmatrix}.$$

Следовательно, j-й столбец является линейной комбинацией первых r столбцов, а значит, первые r столбцов – максимальная линейно независимая подсистема. Теорема доказана.

С л е д с т в и е 1. *Ранг матрицы равен рангу системы ее строк.*

Д о к а з а т е л ь с т в о. Заметим, что ранг системы строк совпадает с рангом системы столбцов, так как при транспонировании базисный минор остается базисным минором. Отсюда получаем нужное утверждение.

С л е д с т в и е 2. *Определитель квадратной матрицы тогда и только тогда равен нулю, когда между ее столбцами существует нетривиальная линейная зависимость.*

Д о к а з а т е л ь с т в о. Разбирая свойства определителя, мы установили, что если столбцы некоторой квадратной матрицы линейно зависимы, то ее определитель равен нулю. Обратное утверждение следует из теоремы о ранге.

13.2. Ранг произведения матриц. Прежде чем формулировать основной результат, установим следующее утверждение.

Л е м м а 1. *Пусть в векторном пространстве V заданы две системы векторов:*

$$u_1, u_2, \ldots, u_r, \tag{1}$$

$$v_1, v_2, \ldots, v_s, \tag{2}$$

причем (1) *линейно выражается через* (2). *Тогда ранг системы* (1) *не превосходит ранга системы* (2).

Д о к а з а т е л ь с т в о. Пусть

$$u_{i_1}, u_{i_2}, \ldots, u_{i_p} \tag{1'}$$

– максимальная линейно независимая подсистема системы (1),

$$v_{j_1}, v_{j_2}, \ldots, v_{j_q}, \tag{2'}$$

– максимальная линейно независимая подсистема системы (2).

Заметим, что система (1′) линейно выражается через систему (1), а (1) линейно выражается через систему (2) по условию. Так как (2′) – максимальная линейно независимая подсистема, то (2) линейно выражается через (2′). Следовательно, (1′) линейно выражается через (2′). По теореме о замене $p \leq q$, где p – ранг системы (1), а q – ранг системы (2). Лемма доказана.

Т е о р е м а. *Ранг произведения матриц не превосходит рангов множителей.*

Д о к а з а т е л ь с т в о. Пусть $A \in M_{m\times n}(P), B \in M_{n\times s}(P)$, $A \cdot B = C \in M_{m\times s}(P)$ – их произведение. По определению

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj}, \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, s.$$

Зафиксируем индекс i и заставим j пробегать всевозможные значения. Тогда i-я строка матрицы C будет иметь вид

$$(c_{i1}, c_{i2}, \ldots, c_{is}) = \sum_{k=1}^{n} a_{ik}\, (b_{k1}, b_{k2}, \ldots, b_{ks}).$$

Следовательно, i-я строка матрицы C есть линейная комбинация строк матрицы B. Рассмотрим следующие системы векторов:

$$\{\text{строки матрицы } C\} = \{C_1, C_2, \ldots, C_m\}, \tag{3}$$

$$\{\text{строки матрицы } B\} = \{B_1, B_2, \ldots, B_n\}. \tag{4}$$

Как мы видели, система (3) линейно выражается через систему (4). Тогда по доказанной лемме 1 ранг системы (3) не превосходит ранга системы (4). Так как ранг матрицы равен рангу ее строк, то

$$\text{ранг } C \leq \text{ранг } B.$$

Если в равенстве

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj}, \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, s$$

заставить индекс i пробегать множество от 1 до m, то получим линейную комбинацию столбцов матрицы A. Так же, как и выше, устанавливается неравенство

$$\text{ранг } C \leq \text{ранг } A.$$

Теорема доказана.

Следующий пример показывает, что в общем случае утверждение теоремы улучшить нельзя.

П р и м е р. 1) Пусть

$$A = B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{ранг } A = \text{ранг } B = 2.$$

Тогда их произведение

$$A \cdot B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

имеет ранг 2.

2) Пусть

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{ранг } A = 2, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \text{ранг } B = 1.$$

Тогда их произведение

$$A \cdot B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

имеет ранг 1.

3) Пусть

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \text{ранг } A = 1, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{ранг } B = 2.$$

Тогда их произведение

$$A \cdot B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

имеет ранг 1.

4) Пусть

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \text{ранг } A = 1, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \text{ранг } B = 1.$$

Тогда их произведение

$$A \cdot B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

имеет ранг 0.

С л е д с т в и е. *Если матрицу умножить на невырожденную матрицу, то ее ранг не изменится.*

Д о к а з а т е л ь с т в о. Напомним, что невырожденной называется квадратная матрица, определитель которой отличен от нуля. Пусть A – некоторая матрица, а Q – невырожденная матрица из $M_n(P)$, и при этом определено произведение $C = A \cdot Q$. Тогда по теореме о ранге произведения матриц имеем неравенство

$$\text{ранг } C \leq \text{ранг } A.$$

С другой стороны, легко заметить, что определено и произведение $A = C \cdot Q^{-1}$, снова по теореме о ранге произведения матриц:

$$\text{ранг } A \leq \text{ранг } C.$$

Следовательно, ранг C = ранг A.

Если матрица умножается на невырожденную слева, то доказательство проводится аналогично.

§ 14. Системы линейных уравнений

14.1. Критерий совместности. Пусть P – поле. *Системой m линейных уравнений от n неизвестных* называется следующее выражение:

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

где коэффициенты a_{ij}, b_i – заданные элементы поля P. *Решением этой системы* называется n-ка $(x_1^0, x_2^0, \ldots, x_n^0) \in P^n$, которая при подстановке в систему дает верные равенства. При рассмотрении любого уравнения (системы уравнений) возникает следующий вопрос: когда решение существует?

Чтобы ответить на этот вопрос, свяжем с системой (1) следующие матрицы:

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

– *матрица системы,*

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

– *столбец свободных членов* и *столбец неизвестных* соответственно. Тогда систему (1) кратко можно записать в таком виде:

$$A \cdot X = B.$$

Если $A = (A^1, A^2, \dots, A^n)$, где A^i – i-й столбец матрицы A, то система (1) примет такой вид:

$$A^1 x_1 + A^2 x_2 + \dots + A^n x_n = B.$$

Введем так называемую *расширенную матрицу* $\overline{A}$, которая получается из матрицы A приписыванием справа столбца свободных членов:

$$\overline{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}.$$

Будем говорить, что система (1) *совместна,* если она имеет хотя бы одно решение. Теперь мы можем сформулировать критерий совместности системы.

Т е о р е м а 1. *Система* (1) *тогда и только тогда совместна, когда ранг ее матрицы равен рангу расширенной матрицы.*

Д о к а з а т е л ь с т в о. Докажем вначале импликацию слева направо. Предположим, что система совместна, и пусть $(x_1^0, x_2^0, \dots, x_n^0)$ – некоторое ее решение. Рассмотрим следующие системы векторов:

$$\{\text{столбцы матрицы } A\} = \{A^1, A^2, \dots, A^n\}, \tag{2}$$

$$\{\text{столбцы матрицы } \overline{A}\} = \{A^1, A^2, \ldots, A^n, B\}. \tag{3}$$

Заметим, что система (2) линейно выражается через систему (3), а система (3) линейно выражается через систему (2), что непосредственно следует из равенства

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} x_1^0 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} x_2^0 + \ldots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n^0 = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Следовательно, система (2) эквивалентна системе (3), а потому по лемме 1 из § 13 ранги этих двух систем совпадают.

Докажем импликацию справа налево. Пусть ранг матрицы A совпадает с рангом матрицы $\overline{A}$. Нужно доказать, что система совместна. В системе столбцов матрицы A выберем максимальную линейно независимую подсистему. Пусть это столбцы с номерами $j_1, j_2, \ldots, j_r$, т. е.

$$A^{j_1}, A^{j_2}, \ldots, A^{j_r}$$

– максимальная линейно независимая система столбцов матрицы A. Те же самые столбцы есть и в матрице $\overline{A}$, где они также образуют максимальную линейно независимую подсистему (так как ранг $A =$ ранг $\overline{A}$). Следовательно,

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = x_{j_1}^0 \begin{pmatrix} a_{1,j_1} \\ a_{2,j_1} \\ \vdots \\ a_{m,j_1} \end{pmatrix} + x_{j_2}^0 \begin{pmatrix} a_{1,j_2} \\ a_{2,j_2} \\ \vdots \\ a_{m,j_2} \end{pmatrix} + \ldots + x_{j_r}^0 \begin{pmatrix} a_{1,j_r} \\ a_{2,j_r} \\ \vdots \\ a_{m,j_r} \end{pmatrix},$$

при некоторых $x_{j_1}^0, x_{j_2}^0, \ldots, x_{j_r}^0$. Ясно, что тогда все числа $x_{j_1}^0$, $x_{j_2}^0$, …, $x_{j_r}^0$ входят в некоторую n-ку, которая является решением системы (1) (остальные элементы этой n-ки равны нулю). Теорема доказана.

Предположим теперь, что система имеет решения. Возникает следующий вопрос: найти все решения системы (1).

14.2. Общее решение. Начнем со следующего простого примера, который показывает, что система может иметь много решений.

П р и м е р. Система, состоящая из одного уравнения

$$x_1 + x_2 = 0,$$

имеет множество решений: $(t, -t)$, где $t \in P$.

Нам надо научиться описывать все решения системы.

О п р е д е л е н и е. *Общим решением системы* (1) называется n-ка функций от параметров $t_1, t_2, \ldots, t_d$:

$$(f_1(t_1, t_2, \ldots, t_d), f_2(t_1, t_2, \ldots, t_d), \ldots, f_n(t_1, t_2, \ldots, t_d)), \qquad (2)$$

где $f_i : P^d \longrightarrow P$, $i = 1, 2, \ldots, n$ такая, что

1) при любых $t_1^0, t_2^0, \ldots, t_d^0$ из P вектор

$$(f_1(t_1^0, t_2^0, \ldots, t_d^0), f_2(t_1^0, t_2^0, \ldots, t_d^0), \ldots, f_n(t_1^0, t_2^0, \ldots, t_d^0))$$

является решением системы (1);

2) любое решение системы (1) может быть получено из (2) при подходящих значениях параметров.

Опишем алгоритм, позволяющий находить общее решение системы (1).

I. Выяснить совместность системы (для этого надо проверить равенство ранг $A =$ ранг $\overline{A}$).

II. Предположим, что система совместна, и при отыскании ранга мы нашли базисный минор матрицы A. Пусть он имеет порядок r. Переставляя уравнения системы и меняя порядок неизвестных, будем считать, что он находится в левом верхнем углу, тогда первые строки образуют максимальную линейно независимую подсистему пространства строк матрицы A. Рассмотрим систему, составленную из первых r уравнений исходной системы:

$$\begin{cases} a_{11}\,x_1 + a_{12}\,x_2 + \ldots + a_{1n}\,x_n = b_1, \\ a_{21}\,x_1 + a_{22}\,x_2 + \ldots + a_{2n}\,x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}\,x_1 + a_{r2}\,x_2 + \ldots + a_{rn}\,x_n = b_r. \end{cases} \qquad (1')$$

Будем называть ее *базисной системой*. Докажем, что эта более короткая система равносильна системе (1), т. е. их множества решений совпадают.

Обозначим $A_{\text{баз}}$ и $\overline{A}_{\text{баз}}$ матрицу и расширенную матрицу системы $(1')$ соответственно.

Рассмотрим расширенную матрицу $\overline{A}$ системы (1) и расширенную матрицу $\overline{A}_{\text{баз}}$ базисной системы $(1')$. Используя тот факт, что всякая строка, входящая в $\overline{A}$, линейно выражается через строки матрицы

$\overline{A}_{\text{баз}}$, подберем такую матрицу C, что $C \cdot \overline{A}_{\text{баз}} = \overline{A}$. Для этого обозначим символами $\overline{A}_1$, $\overline{A}_2$, …, $\overline{A}_r$ первые r строк системы (1). Они являются строками матрицы $\overline{A}_{\text{баз}}$, т. е.

$$\overline{A}_{\text{баз}} = \begin{pmatrix} \overline{A}_1 \\ \overline{A}_2 \\ \vdots \\ \overline{A}_r \end{pmatrix}.$$

Так как строки $\overline{A}_1$, $\overline{A}_2$, …, $\overline{A}_r$ матрицы $\overline{A}$ образуют максимальную линейно независимую подсистему, то остальные строки выражаются через них, т. е. найдутся элементы $c_{ij} \in P$, для которых справедливы равенства

$$\begin{cases} \overline{A}_{r+1} = c_{r+1,1}\,\overline{A}_1 + c_{r+1,2}\,\overline{A}_2 + \ldots + c_{r+1,r}\,\overline{A}_r, \\ \overline{A}_{r+2} = c_{r+2,1}\,\overline{A}_1 + c_{r+2,2}\,\overline{A}_2 + \ldots + c_{r+2,r}\,\overline{A}_r, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \overline{A}_m = c_{m,1}\,\overline{A}_1 + c_{m,2}\,\overline{A}_2 + \ldots + c_{m,r}\,\overline{A}_r. \end{cases}$$

Рассмотрим матрицу

$$C = \left(\begin{array}{cccc} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \\ \hline c_{r+1,1} & c_{r+1,2} & \ldots & c_{r+1,r} \\ c_{r+2,1} & c_{r+2,2} & \ldots & c_{r+2,r} \\ \ldots & \ldots & \ldots & \ldots \\ c_{m,1} & c_{m,2} & \ldots & c_{m,r} \end{array}\right),$$

где вверху стоит единичная матрица степени r. Для нее справедливо равенство $C \cdot \overline{A}_{\text{баз}} = \overline{A}$.

Очевидно, что всякое решение системы (1) является решением системы (1′). Пусть $(x_1^0, x_2^0, \ldots, x_n^0)$ – решение базисной системы (1′). Докажем, что оно является решением всей системы. Подставим

$$\overline{A} \cdot \begin{pmatrix} x_1^0 \\ x_2^0 \\ \vdots \\ x_n^0 \\ -1 \end{pmatrix} = (C \cdot \overline{A}_{\text{баз}}) \begin{pmatrix} x_1^0 \\ x_2^0 \\ \vdots \\ x_n^0 \\ -1 \end{pmatrix} = C \cdot \left(\overline{A}_{\text{баз}} \begin{pmatrix} x_1^0 \\ x_2^0 \\ \vdots \\ x_n^0 \\ -1 \end{pmatrix} \right) = C \cdot \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} = \mathbf{0},$$

т. е. $(x_1^0, x_2^0, \ldots, x_n^0)$ – решение системы (1). Таким образом, мы должны научиться решать систему (1′), поскольку она равносильна системе (1).

III. Рассмотрим систему

$$\begin{cases} a_{11}\,x_1 + a_{12}\,x_2 + \ldots + a_{1r}\,x_r = b_1 - a_{1,r+1}\,x_{r+1} - \ldots - a_{1n}\,x_n, \\ a_{21}\,x_1 + a_{22}\,x_2 + \ldots + a_{2r}\,x_r = b_2 - a_{2,r+1}\,x_{r+1} - \ldots - a_{2n}\,x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}\,x_1 + a_{r2}\,x_2 + \ldots + a_{rr}\,x_r = b_r - a_{r,r+1}\,x_{r+1} - \ldots - a_{rn}\,x_n. \end{cases}$$

Переменные, входящие в правую часть, будем называть *свободными*, а переменные, входящие в левую часть, – *базисными*. Решаем систему относительно базисных переменных, считая, что ее правая часть известна. Так как определитель матрицы системы отличен от нуля, то такие системы мы уже умеем решать. Решая ее относительно $x_1, x_2, \ldots, x_r$, получим решение:

$$(L_1(x_{r+1}, x_{r+2}, \ldots, x_n), L_2(x_{r+1}, x_{r+2}, \ldots, x_n), \ldots, L_r(x_{r+1}, x_{r+2}, \ldots, x_n)).$$

Тогда

$$(L_1(x_{r+1}, x_{r+2}, \ldots, x_n), L_2(x_{r+1}, x_{r+2}, \ldots, x_n), \ldots, L_r(x_{r+1}, x_{r+2}, \ldots, x_n),$$
$$x_{r+1}, x_{r+2}, \ldots, x_n)$$

– решение системы (1). Чтобы доказать, что это общее решение, мы должны проверить выполнение условий 1 и 2.

1) Если придать некоторые значения свободным переменным x_{r+1}, x_{r+2}, …, x_n и найти по ним значения

$$L_1(x_{r+1}, x_{r+2}, \ldots, x_n), L_2(x_{r+1}, x_{r+2}, \ldots, x_n), \ldots, L_r(x_{r+1}, x_{r+2}, \ldots, x_n),$$

то построенная n-ка будет решение системы (1).

2) Покажем, что всякое решение системы (1) можно получить из нашей формулы.

Пусть $(x_1^0, x_2^0, \ldots, x_r^0, x_{r+1}^0, x_{r+2}^0, \ldots, x_n^0)$ – некоторое решение системы (1). Придадим значения $x_{r+1} = x_{r+1}^0$, $x_{r+2} = x_{r+2}^0$, …, $x_n = x_n^0$ и найдем первые r значений:

$$L_1(x_{r+1}^0, x_{r+2}^0, \ldots, x_n^0), L_2(x_{r+1}^0, x_{r+2}^0, \ldots, x_n^0), \ldots, L_r(x_{r+1}^0, x_{r+2}^0, \ldots, x_n^0).$$

Тогда набор

$$(L_1(x_{r+1}^0, x_{r+2}^0, \ldots, x_n^0), L_2(x_{r+1}^0, x_{r+2}^0, \ldots, x_n^0), \ldots, L_r(x_{r+1}^0, x_{r+2}^0, \ldots, x_n^0),$$

$$x^0_{r+1}, x^0_{r+2}, \ldots, x^0_n)$$

является решением системы (1). Так как невырожденная система имеет единственное решение, то

$$x^0_1 = L_1(x^0_{r+1}, x^0_{r+2}, \ldots, x^0_n), x^0_2 = L_2(x^0_{r+1}, x^0_{r+2}, \ldots, x^0_n), \ldots, x^0_r =$$
$$= L_r(x^0_{r+1}, x^0_{r+2}, \ldots, x^0_n).$$

Это значит, что произвольное решение можно задать формулой:

$$(L_1(x_{r+1}, x_{r+2}, \ldots, x_n), L_2(x_{r+1}, x_{r+2}, \ldots, x_n), \ldots, L_r(x_{r+1}, x_{r+2}, \ldots, x_n),$$
$$x_{r+1}, x_{r+2}, \ldots, x_n).$$

14.3. Связь неоднородных систем с однородными. Система линейных уравнений называется *однородной,* если все ее свободные члены равны нулю.

По системе линейных уравнений (1) мы можем построить соответствующую ей однородную систему:

$$\begin{cases} a_{11}\,x_1 + a_{12}\,x_2 + \ldots + a_{1n}\,x_n = 0, \\ a_{21}\,x_1 + a_{22}\,x_2 + \ldots + a_{2n}\,x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}\,x_1 + a_{m2}\,x_2 + \ldots + a_{mn}\,x_n = 0. \end{cases} \qquad (1_0)$$

Следующая теорема показывает, как связано множество решений неоднородной системы и множество решений однородной.

Т е о р е м а 2. *Если X^0 – некоторое решение системы* (1), *то множество решений системы* (1) *есть*

$$\{\text{решение системы (1)}\} = X^0 + \{\text{решение системы } (1_0)\}.$$

Д о к а з а т е л ь с т в о. Запишем системы (1) и (1_0) в матричной форме:

$$A \cdot X = B, \quad A \cdot X = \mathbf{0}.$$

Установим включение $\subseteq$. Пусть Y – некоторое решение системы (1). Надо доказать, что оно представимо в таком виде:

$$Y = X^0 + \text{решение системы } (1_0).$$

Рассмотрим

$$A(Y - X^0) = AY - AX^0 = B - B = \mathbf{0},$$

т. е. $Y - X^0 = Z$ – решение однородной системы. Следовательно, $Y = X^0 + Z$.

Обратное включение. Пусть теперь Z – решение системы (1_0). Покажем, что $X^0 + Z$ – решение системы (1). Подставляя в систему, получим:

$$A(X^0 + Z) = AX^0 + AZ = B + \mathbf{0} = B.$$

Теорема доказана.

Т е о р е м а 3. *Множество решений однородной системы* (1_0) *является подпространством пространства* P^n.

Д о к а з а т е л ь с т в о. Проверим, что если Z, W – два решения системы (1_0), а $\alpha, \beta \in P$, то $\alpha\, Z + \beta\, W$ – решение системы (1_0). Имеем:

$$A(\alpha\, Y + \beta\, W) = A\,\alpha\, Z + A\,\beta\, W = \alpha\,(A\, Z) + \beta\,(A\, W) = \alpha\,\mathbf{0} + \beta\,\mathbf{0} = \mathbf{0}.$$

Таким образом, линейная комбинация решений снова есть решение. Теорема доказана.

§ 15. Однородные системы

15.1. Фундаментальные системы решений. Так как все решения однородной системы (1_0) есть подпространство, то можно найти его базу и через нее выразить все множество решений.

О п р е д е л е н и е. База пространства решений однородной системы уравнений называется *фундаментальной системой решений.*

Решение однородной системы сводится к отысканию фундаментальной системы решений. Рассмотрим систему линейных однородных уравнений:

$$\begin{cases} a_{11}\, x_1 + a_{12}\, x_2 + \ldots + a_{1n}\, x_n = 0, \\ a_{21}\, x_1 + a_{22}\, x_2 + \ldots + a_{2n}\, x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}\, x_1 + a_{m2}\, x_2 + \ldots + a_{mn}\, x_n = 0. \end{cases} \qquad (1_0)$$

Ее *рангом* называется ранг матрицы $A = (a_{ij})$. Если $r = n$, то $X = \mathbf{0}$ – единственное решение.

Предположим, что $r < n$. Будем считать, что базисный минор матрицы системы находится в левом верхнем углу. Рассмотрим под-

систему системы (1_0), состоящую из первых r уравнений:

$$\begin{cases} a_{11}\,x_1 + a_{12}\,x_2 + \ldots + a_{1r}\,x_r = -a_{1,r+1}\,x_{r+1} - \ldots - a_{1n}\,x_n, \\ a_{21}\,x_1 + a_{22}\,x_2 + \ldots + a_{2r}\,x_r = -a_{2,r+1}\,x_{r+1} - \ldots - a_{2n}\,x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}\,x_1 + a_{r2}\,x_2 + \ldots + a_{rr}\,x_r = -a_{r,r+1}\,x_{r+1} - \ldots - a_{rn}\,x_n, \end{cases} \tag{2}$$

где

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1r} \\ a_{21} & a_{22} & \ldots & a_{2r} \\ \ldots\ldots\ldots\ldots\ldots \\ a_{r1} & a_{r2} & \ldots & a_{rr} \end{vmatrix}$$

– базисный минор.

В этих обозначениях справедлива

Т е о р е м а. *Пусть задана однородная система линейных уравнений* (1_0), r *– ее ранг,* n *– число неизвестных, причем* $r < n$ *и в системе выделены* $n - r$ *свободных переменных.*

а) *Пусть* D *– произвольный определитель порядка* $n - r$*, отличный от нуля. Используя его строки в качестве значений свободных неизвестных, найдем* $n - r$ *решений системы* (1_0)*. Это – фундаментальная система решений системы* (1_0)*.*

б) *Всякая фундаментальная система решений системы* (1_0) *может быть получена из некоторого ненулевого определителя порядка* $n - r$ *при помощи процедуры, описанной в пункте* а.

Д о к а з а т е л ь с т в о. Пусть

$$D = \begin{vmatrix} d_{1,r+1} & d_{1,r+2} & \ldots & d_{1n} \\ d_{2,r+1} & d_{2,r+2} & \ldots & d_{2n} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ d_{n-r,r+1} & d_{n-r,r+2} & \ldots & d_{n-r,n} \end{vmatrix} \neq 0.$$

Возьмем первую строку этого определителя и подставим вместо свободных переменных в общее решение:

$$(L_1(x_{r+1}, x_{r+2}, \ldots, x_n), L_2(x_{r+1}, x_{r+2}, \ldots, x_n), \ldots, L_r(x_{r+1}, x_{r+2}, \ldots, x_n),$$

$$x_{r+1}, x_{r+2}, \ldots, x_n),$$

получим вектор

$$u_1 = (d_{11}, d_{12}, \ldots, d_{1r}, d_{1,r+1}, \ldots, d_{1n}),$$

подставим вторую строку, получим вектор

$$u_2 = (d_{21}, d_{22}, \ldots, d_{2r}, d_{2,r+1}, \ldots, d_{2n}),$$

и т. д. Наконец, подставив последнюю строку, получим вектор

$$u_{n-r} = (d_{n-r,1}, d_{n-r,2}, \ldots, d_{n-r,r}, d_{n-r,r+1}, \ldots, d_{n-r,n}).$$

Построенные векторы образуют множество решений однородной системы (1_0). Чтобы доказать, что они образуют фундаментальную систему решений, надо проверить:

1) линейную независимость;

2) максимальность.

Для произвольного вектора $v \in P^n$ символом v^* будем обозначать его конечный отрезок длины $n - r$, т. е. если $v = (\alpha_1, \ldots, \alpha_r, \alpha_{r+1}, \ldots, \alpha_n) \in P^n$, то $v^* = (\alpha_{r+1}, \ldots, \alpha_n) \in P^{n-r}$.

Установим пункт 1 от противного. Предположим, что векторы линейно зависимы. Так как строки $u_1, u_2, \ldots, u_{n-r}$ линейно зависимы (т. е. их линейная комбинация обращается в нуль), то и векторы $u_1^*, u_2^*, \ldots, u_{n-r}^*$ линейно зависимы, т. е. определитель

$$D = \begin{vmatrix} u_1^* \\ u_2^* \\ \vdots \\ u_{n-r}^* \end{vmatrix}$$

обращается в нуль, а это не так.

2) Надо доказать, что произвольное решение системы (1_0) есть линейная комбинация системы векторов $u_1, u_2, \ldots, u_{n-r}$. Пусть

$$u = (d_1, d_2, \ldots, d_r, d_{r+1}, d_{r+2}, \ldots, d_n)$$

– решение нашей системы. Рассмотрим векторы $u_1^*, u_2^*, \ldots, u_{n-r}^*$. Они линейно независимы, так как это строки определителя D. В $n - r$-мерном пространстве P^{n-r} они образуют базу. Пусть вектор u^* линейно выражается через эту базу:

$$u^* = \alpha_1 u_1^* + \alpha_2 u_2^* + \ldots + \alpha_{n-r} u_{n-r}^*, \quad \alpha_i \in P.$$

Рассмотрим вектор

$$w = u - (\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_{n-r} u_{n-r}) \in P^n,$$

который по теореме 3 из § 14 является решением однородной системы. Очевидно, $w^* = (0, 0, \ldots, 0) \in P^{n-r}$, т. е. w – решение, полученное при нулевых значениях свободных неизвестных. Тогда $L_1, L_2, \ldots, L_r$ в формуле общего решения системы (2) тоже нули (однородная система, матрица которой не вырождена, имеет только нулевое решение) и $w = (0, 0, \ldots, 0) \in P^n$, т. е.

$$u = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_{n-r} u_{n-r}.$$

Следовательно, $u_1, u_2, \ldots, u_{n-r}$ – фундаментальная система решений.

Докажем утверждение б. Пусть задана некоторая фундаментальная система решений:

$$\begin{aligned}
&v_1 = (c_{11}, c_{12}, \ldots, c_{1r}, c_{1,r+1}, \ldots, c_{1n}),\\
&v_2 = (c_{21}, c_{22}, \ldots, c_{2r}, c_{2,r+1}, \ldots, c_{2n}),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&v_{n-r} = (c_{n-r,1}, c_{n-r,2}, \ldots, c_{n-r,r}, c_{n-r,r+1}, \ldots, c_{n-r,n}).
\end{aligned}$$

Надо показать, что эту систему можно получить при помощи процедуры, описанной в пункте а. Для этого возьмем в качестве D определитель

$$D = \begin{vmatrix} c_{1,r+1} & c_{1,r+2} & \ldots & c_{1n} \\ c_{2,r+1} & c_{2,r+2} & \ldots & c_{2n} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ c_{n-r,r+1} & c_{n-r,r+2} & \ldots & c_{n-r,n} \end{vmatrix}$$

и покажем, что $D \neq 0$. Предположим противное, т. е. $D = 0$, тогда между его строками существует линейная зависимость

$$\beta_1 v_1^* + \beta_2 v_2^* + \ldots + \beta_{n-r} v_{n-r}^* = 0, \quad \beta_i \in P,$$

где не все β_i обращаются в нуль. Рассмотрим вектор

$$v = \beta_1 v_1 + \beta_2 v_2 + \ldots + \beta_{n-r} v_{n-r},$$

который по теореме 3 из § 14 является решением системы 1_0. Очевидно, $v^* = 0$, а значит, $v = 0$, т. е.

$$\beta_1 v_1 + \beta_2 v_2 + \ldots + \beta_{n-r} v_{n-r} = 0,$$

но это невозможно, так как $v_1, v_2, \ldots, v_{n-r}$ образуют базис пространства решений, а потому линейно независимы. Следовательно, $D \neq 0$. Теорема доказана.

15.2. База суммы и пересечения двух подпространств. Задано векторное пространство V над полем P размерности n и в нем два подпространства U' и U''. Мы знаем, что сумма $U' + U''$ и пересечение $U' \cap U''$ являются подпространствами. Пусть

$$u_1, u_2, \ldots, u_r$$

– база U', а

$$u_{r+1}, u_{r+2}, \ldots, u_s$$

– база U''. Требуется найти базу суммы $U' + U''$ и базу пересечения $U' \cap U''$.

Вначале найдем базу суммы. По определению

$$U' + U'' = \{u' + u'' \mid u' \in U', u'' \in U''\}.$$

Следовательно, $U' + U''$ состоит из множества линейных комбинаций векторов $u_1,\ u_2,\ \ldots,\ u_r,\ u_{r+1}, u_{r+2}, \ldots, u_s$. Выберем из этой системы векторов максимальную линейно независимую подсистему. Это и будет искомая база суммы $U' + U''$.

Найдем теперь базу пересечения. Заметим, что $u \in U' \cap U''$ тогда и только тогда, когда u разлагается как по базе пространства U', так и по базе пространства U'', т. е.

$$u = \sum_{i=1}^{r} \alpha_i\, u_i = \sum_{i=r+1}^{s} \alpha_i\, u_i. \tag{3}$$

Зафиксируем в V какую-нибудь базу $e_1, e_2, \ldots, e_n$, и пусть в этой базе вектор u_i имеет следующие координаты:

$$[u_i] = (u_{i1}, u_{i2}, \ldots, u_{in}), \quad i = 1, 2, \ldots, s.$$

Тогда из (3) получим систему

$$\sum_{i=1}^{r} \alpha_i\, u_{ik} = \sum_{i=r+1}^{s} \alpha_i\, u_{ik}, \quad 1 \le k \le n$$

или, перенося все слагаемые в левую часть, приходим к системе линейных однородных уравнений

$$\sum_{i=1}^{r} \alpha_i\, u_{ik} - \sum_{i=r+1}^{s} \alpha_i\, u_{ik} = 0, \quad 1 \le k \le n \tag{4}$$

с неизвестными $\alpha_1, \alpha_2, \ldots, \alpha_s$.

Пусть

$$\alpha^{(l)} = (\alpha_1^{(l)}, \alpha_2^{(l)}, \ldots, \alpha_r^{(l)}, \alpha_{r+1}^{(l)}, \ldots, \alpha_s^{(l)}), \quad 1 \le l \le t,$$

– фундаментальная система решений системы (4). Утверждается, что векторы

$$u^{(l)} = \sum_{i=1}^{r} \alpha_i^{(l)}\, u_i = \sum_{i=r+1}^{s} \alpha_i^{(l)}\, u_i, \quad 1 \le l \le t$$

образуют базу пересечения $U' \cap U''$. Очевидно, все они лежат в пересечении. Проверим, что они линейно независимы. Пусть

$$\sum_{l=1}^{t} \gamma^{(l)}\, u^{(l)} = 0, \quad \gamma^{(l)} \in P.$$

Надо доказать, что все $\gamma^{(l)} = 0$. Имеем:

$$\sum_{l=1}^{t} \gamma^{(l)} \left(\sum_{i=1}^{r} \alpha_i^{(l)}\, u_i \right) = \sum_{l=1}^{t} \gamma^{(l)} \left(\sum_{i=r+1}^{s} \alpha_i^{(l)}\, u_i \right) = 0.$$

Переставляя знаки суммирования, получим

$$\sum_{i=1}^{r} \left(\sum_{l=1}^{t} \gamma^{(l)}\, \alpha_i^{(l)} \right) u_i = \sum_{i=r+1}^{s} \left(\sum_{l=1}^{t} \gamma^{(l)}\, \alpha_i^{(l)} \right) u_i = 0.$$

Отсюда, учитывая, что каждая из систем векторов $u_1, u_2, \ldots, u_r$ и $u_{r+1}, u_{r+2}, \ldots, u_n$ линейно независима, получим равенство

$$\sum_{l=1}^{t} \gamma^{(l)}\, \alpha_i^{(l)} = 0$$

при каждом $i \in \{1, 2, \ldots, s\}$, т. е.

$$\sum_{l=1}^{t} \gamma^{(l)}\, \alpha^{(l)} = 0.$$

А так как множество векторов $\{\alpha^{(l)}\}$ линейно независимо, то все $\gamma^{(l)} = 0$.

Докажем максимальность. Пусть $v \in U' \cap U''$ и

$$v = \sum_{i=1}^{r} \beta_i u_i = \sum_{i=r+1}^{s} \beta_i u_i$$

при некоторых β_i из P. Обозначим $\beta = (\beta_1, \beta_2, \ldots, \beta_s)$ и, учитывая, что $\{\alpha^{(l)}\}$ – фундаментальная система решений, имеем

$$\beta = \sum_{l=1}^{t} \delta^{(l)} \alpha^{(l)}$$

при подходящих $\delta^{(l)} \in P$. Покажем, что

$$v = \sum_{l=1}^{t} \delta^{(l)} u^{(l)}.$$

Действительно,

$$\sum_{l=1}^{t} \delta^{(l)} u^{(l)} = \sum_{l=1}^{t} \delta^{(l)} \left(\sum_{i=1}^{r} \alpha_i^{(l)} u_i \right) = \sum_{l=1}^{t} \delta^{(l)} \left(\sum_{i=r+1}^{s} \alpha_i^{(l)} u_i \right) =$$

$$= \sum_{i=1}^{r} \left(\sum_{l=1}^{t} \delta^{(l)} \alpha_i^{(l)} \right) u_i = \sum_{i=r+1}^{s} \left(\sum_{l=1}^{t} \delta^{(l)} \alpha_i^{(l)} \right) u_i = \sum_{i=1}^{r} \beta_i u_i = \sum_{i=r+1}^{s} \beta_i u_i = v,$$

что и требовалось доказать.

§ 16. Теорема Фредгольма

Рассмотрим систему линейных уравнений

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

Вводя матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

запишем систему в таком виде:

$$AX = B. \tag{1}$$

Также будем рассматривать систему линейных однородных уравнений

$$AX = \mathbf{0} \tag{1$_0$}$$

и транспонированную систему

$$A^t Y = \mathbf{0}, \tag{2}$$

однородных уравнений от неизвестных

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix},$$

где A^t – матрица, транспонированная к A.

Чтобы сформулировать теорему, дадим такое

О п р е д е л е н и е. Два вектор-столбца

$$u = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix}, \quad \alpha_i \in P, \quad v = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}, \quad \beta_i \in P$$

называются *ортогональными*, если

$$u^t\, v = \alpha_1\, \beta_1 + \alpha_2\, \beta_2 + \ldots + \alpha_m\, \beta_m = 0.$$

Справедлива

Т е о р е м а Ф р е д г о л ь м а. I. (Существование.) *Система* (1) *тогда и только тогда совместна, когда ее столбец свободных членов ортогонален каждому решению однородной транспонированной системы, т. е.* $B^t\, Y^0 = \mathbf{0}$ *для любого* Y^0*, удовлетворяющего системе* (2).

II. (Единственность.) *Система* (1) *имеет единственное решение тогда и только тогда, когда система* (1$_0$) *имеет только нулевое решение.*

III. (Число решений.) *Если матрица A – квадратная, то системы* (1_0) *и* (2) *имеют одно и то же число линейно независимых решений.*

Д о к а з а т е л ь с т в о. I. Пусть система совместна и $A\,X^0 = B$, т. е. X^0 – решение. Предположим, что Y^0 – решение системы (2), т. е. $A^t\,Y^0 = \mathbf{0}$. Надо доказать, что $B^t\,Y^0 = 0$. Имеем

$$B^t\,Y^0 = (A\,X^0)^t\,Y^0 = (X^0)^t\,(A^t\,Y^0) = (X^0)^t\,\mathbf{0} = \mathbf{0}.$$

Докажем обратное утверждение. Пусть столбец свободных членов ортогонален каждому решению однородной транспонированной системы, т. е. из того, что $A^tY^0 = \mathbf{0}$, следует, что $B^tY^0 = 0$. Надо доказать, что система (1) совместна. После транспонирования получаем, что из равенства $(Y^0)^tA = \mathbf{0}$ следует равенство $(Y^0)^tB = 0$. Следовательно, если в матрице A строки с номерами $i_1, i_2, \ldots, i_s$ образуют максимальную линейно независимую подсистему, то строки с этими же номерами образуют максимальную линейно независимую подсистему и расширенной матрицы $\overline{A} = (A|B)$, а потому

$$\text{ранг}\,A = \text{ранг}\,\overline{A},$$

т. е. система (1) совместна.

II. По теореме 2 из § 14 всякое решение системы (1) есть сумма некоторого частного решения X^0 и решения однородной системы. Отсюда следует, что если решение системы (1) единственно, то однородная система может иметь только нулевое решение.

III. Если r – ранг матрицы A, n – число ее столбцов, то фундаментальная система решений однородной системы (1_0) состоит из $n - r$ векторов. Так как транспонированная матрица A^t также имеет ранг r, то и фундаментальная система решений системы (2) также состоит из $n - r$ векторов. Теорема доказана.

§ 17. Фактор-пространства

17.1. Эквивалентности и фактор-множества. Бинарное отношение $\sim$ на множестве M называется *отношением эквивалентности,* если выполнены следующие три условия:

а) *рефлексивность*: для всякого $x \in M$ справедливо $x \sim x$;

б) *симметричность*: для любых x и y из M из того, что $x \sim y$, следует, что $y \sim x$;

в) *транзитивность*: для любых x, y, z из M из того, что $x \sim y$ и $y \sim z$, следует, что $x \sim z$.

П р и м е р ы. 1) Если на множестве $\mathbb{Q}$ рассмотреть отношение $<$, то очевидно, что оно транзитивно, но не рефлексивно и не симметрично. Если на этом же множестве рассмотреть отношение $\leq$, то оно рефлексивно и транзитивно, но не симметрично. Отношение же $=$ является отношением эквивалентности, так как оно рефлексивно, симметрично и транзитивно.

2) Рассмотрим множество всех треугольников на плоскости. Нетрудно проверить, что отношение подобия будет отношением эквивалентности.

Если на множестве M задано отношение эквивалентности $\sim$, то оно распадается на классы эквивалентности:

$$K_a = \{x \in M \mid x \sim a\},$$

которые называются *смежными классами.* Эти классы либо совпадают, либо не имеют общих элементов. Действительно, предположим, что два класса K_a и K_b пересекаются и $c \in K_a \cap K_b$. Если $x \in K_b$, то $x \sim b$, $b \sim c$ и $c \sim a$. Тогда $x \sim a$. Значит, $K_b \subseteq K_a$. Аналогично проверяется, что и $K_a \subseteq K_b$. Следовательно, $K_a = K_b$.

О п р е д е л е н и е. Множество смежных классов называется *фактор-множеством множества M по отношению эквивалентности* $\sim$ и обозначается $M/_\sim$.

Если множество разбито на непересекающиеся классы, то можно ввести отношение эквивалентности, полагая, что два элемента эквивалентны, если они лежат в одном классе. Нетрудно убедиться, что это отношение действительно является отношением эквивалентности. Следовательно, ввести на множестве отношение эквивалентности или разбить на смежные классы – это одно и то же.

17.2. Фактор-пространства. Пусть задано векторное пространство V над полем P и U – некоторое его подпространство. Будем говорить, что векторы x и y из V *эквивалентны по модулю U*, и писать $x \sim y \pmod U$, если $x - y \in U$.

Покажем, что так введенное отношение является отношением эквивалентности. Действительно,

а) так как $0 = x - x \in U$, то $x \sim x$ и рефлексивность выполняется;

б) пусть $x \sim y$, т. е. $x - y \in U$, но тогда $y - x = -(x - y) \in U$, т. е. если вектор лежит в подпространстве, то и противоположный лежит там же, следовательно, $y \sim x$, и симметричность выполняется;

в) пусть $x \sim y$ и $y \sim z$, надо доказать, что $x \sim z$. Иными словами, из того, что $x - y \in U$ и $y - z \in U$, следует, что $x - z \in U$. Действительно,

$$x - z = (x - y) + (y - z) \in U.$$

Таким образом, введенное отношение является отношением эквивалентности, и множество V распадается на классы эквивалентности: V/U. Для любого элемента a из V обозначим

$$a + U = \{a + u \mid u \in U\}.$$

Справедлива

Л е м м а. *Каждый смежный класс K_a, $a \in V$, из V/U имеет вид*

$$K_a = a + U.$$

Д о к а з а т е л ь с т в о. Установим включение $\subseteq$. Если $x \sim a$, то $x - a = u \in U$ и $x = a + u$. Обратное включение. Если $a + u = x$, где $u \in U$, то $x - a \in U$, а потому $x \sim a$. Лемма доказана.

Таким образом, пространство V распадается на смежные классы, и каждый класс имеет такой вид: $a + U$. Определим на фактормножестве V/U структуру векторного пространства над полем P (над тем же полем, над которым определялось векторное пространство V).

О п р е д е л е н и е. Определим *сумму двух смежных классов* равенством

$$(a + U) + (b + U) = a + b + U$$

и умножение класса на скаляр

$$\alpha(a + U) = \alpha a + U.$$

Т е о р е м а 1. *Данные выше определения операций сложения и умножения на скаляр не зависят от случайного выбора представителей смежных классов. Относительно этих операций множество V/U является векторным пространством.*

Д о к а з а т е л ь с т в о. Рассмотрим некоторый класс $a + U$. Выбирая другой представитель, зададим его в виде $a' + U$. Аналогично класс $b + U$ зададим в виде $b' + U$. Нам надо доказать следующие равенства:

$$a + b + U = a' + b' + U, \quad \alpha a + U = \alpha a' + U,$$

которые и означают, что определенные выше операции суммы и умножения на скаляр не зависят от выбора представителей. Из равенства классов

$$a + U = a' + U, \quad b + U = b' + U$$

имеем включения

$$a - a' \in U, \quad b - b' \in U,$$

но тогда

$$(a + b) - (a' + b') = (a - a') + (b - b') \in U, \quad \alpha a - \alpha a' = \alpha(a - a') \in U,$$

и первая часть теоремы установлена.

Чтобы проверить, что V/U является векторным пространством, надо проверить аксиомы абелевой группы и аксиомы, связывающие скаляры и векторы. Аксиома ассоциативности сложения

$$[(a + U) + (b + U)] + (c + U) = (a + U) + [(b + U) + (c + U)]$$

равносильна равенству

$$(a + b) + c + U = a + (b + c) + U,$$

справедливость которого следует из ассоциативности сложения в V.

Коммутативность сложения проверяется аналогично.

В качестве нулевого класса возьмем класс $0 + U = U$. Тогда

$$(a + U) + (0 + U) = a + U.$$

В качестве противоположного класса для $a+U$ возьмем класс $-a+U$. Тогда

$$(a + U) + (-a + U) = 0 + U = U.$$

Проверяем смешанные аксиомы:

$$\begin{aligned} &\alpha(A + B) = \alpha A + \alpha B, \\ &(\alpha + \beta)A = \alpha A + \beta A, \\ &(\alpha\beta)A = \alpha(\beta A), \\ &1 \cdot A = A, \end{aligned}$$

где $\alpha, \beta \in P$, $A, B \in V/U$.

Проверим первую аксиому:

$$\alpha\,[(a + U) + (b + U)] = \alpha(a + U) + \alpha(b + U).$$

Левую часть этого равенства представим в виде

$$\alpha\left[(a+U)+(b+U)\right]=\alpha(a+b)+U.$$

Преобразуем правую часть:

$$\alpha(a+U)+\alpha(b+U)=(\alpha a+U)+(\alpha b+U)=(\alpha a+\alpha b)+U.$$

Так как они равны, то аксиома установлена. Остальные аксиомы проверяются аналогично. Теорема доказана.

17.3. Размерность фактор-пространства. Связь между размерностями пространства, подпространства и соответствующего фактор-пространства дает

Т е о р е м а 2. *Размерность фактор-пространства V/U равна размерности V минус размерность U, т. е.*

$$\dim(V/U)=\dim V-\dim U.$$

Д о к а з а т е л ь с т в о. Выберем в подпространстве U произвольную базу $e_1,e_2,\ldots,e_r$ и дополним ее векторами $e_{r+1},e_{r+2},\ldots,e_n$ до базы всего пространства V. Для доказательства теоремы достаточно доказать, что классы

$$e_{r+1}+U,e_{r+2}+U,\ldots,e_n+U$$

образуют базу V/U.

Проверим линейную независимость. Пусть

$$\alpha_{r+1}(e_{r+1}+U)+\alpha_{r+2}(e_{r+2}+U)+\ldots+\alpha_n(e_n+U)=U.$$

Надо доказать, что $\alpha_{r+1}=\alpha_{r+2}=\ldots=\alpha_n=0$. Перепишем наше равенство в таком виде:

$$\alpha_{r+1}e_{r+1}+\alpha_{r+2}e_{r+2}+\ldots+\alpha_ne_n+U=U.$$

Это равносильно тому, что

$$\alpha_{r+1}e_{r+1}+\alpha_{r+2}e_{r+2}+\ldots+\alpha_ne_n\in U,$$

но если некоторый вектор лежит в U, то он раскладывается по базе пространства U. Следовательно, имеем равенство

$$\alpha_{r+1}e_{r+1}+\alpha_{r+2}e_{r+2}+\ldots+\alpha_ne_n=\beta_1e_1+\beta_2e_2+\ldots+\beta_re_r.$$

Учитывая линейную независимость векторов $e_1, e_2, \ldots, e_r$, e_{r+1}, e_{r+2}, $\ldots$, e_n, заключаем, что все $\alpha_i = 0$.

Докажем максимальность. Рассмотрим некоторый класс $v + U$. Надо представить его как линейную комбинацию базисных классов. Для вектора v справедливо равенство

$$v = \gamma_1 e_1 + \gamma_2 e_2 + \ldots + \gamma_r e_r + \gamma_{r+1} e_{r+1} + \ldots + \gamma_n e_n.$$

Тогда

$$v+U = \gamma_1 (e_1+U)+\gamma_2 (e_2+U)+\ldots+\gamma_r (e_r+U)+\gamma_{r+1} (e_{r+1}+U)+\ldots+$$

$$+\gamma_n (e_n + U) = \gamma_1 (e_1 + U) + \gamma_2 (e_2 + U) + \ldots + \gamma_r (e_r + U) + U.$$

Максимальность установлена. Теорема доказана.

Глава 3

КОЛЬЦА МНОГОЧЛЕНОВ

§ 18. Многочлены от одной переменной

18.1. Определения и основные свойства. *Многочленом от одной переменной над кольцом* K называется выражение

$$f = f(x) = a_0 + a_1 x + \ldots + a_n x^n = \sum_{i=0}^{n} a_i x^i, \quad a_i \in K,$$

где x – некоторая буква. Если $a_n \neq 0$, то a_n называется *старшим коэффициентом многочлена* f, а n *– степенью многочлена* f (обозначение: $n = \deg f$). Нулевому многочлену 0 степень не приписывается. Два многочлена *равны*, если равны коэффициенты при одинаковых степенях x. Множество всех многочленов от x над кольцом K будем обозначать символом $K[x]$. На множестве $K[x]$ определим операции сложения и умножения:

$$\sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{n} b_i x^i = \sum_{i=0}^{n} (a_i + b_i) x^i,$$

$$\sum_{i=0}^{n} a_i x^i \cdot \sum_{j=0}^{m} b_j x^j = \sum_{k=0}^{n+m} c_k x^k, \quad \text{где} \quad c_k = \sum_{i+j=k} a_i b_j.$$

При определении операции сложения мы добавляем нулевые слагаемые с тем, чтобы получить записи с одинаковыми степенями x.

Из этого определения видно, что $\langle K[x];\, +,\, \cdot\rangle$ – алгебраическая система, которую в дальнейшем будем обозначать просто $K[x]$. Более того, справедлива

Т е о р е м а 1. 1) *Если K – кольцо, то $K[x]$ – тоже кольцо.* 2) *Если K – коммутативное кольцо, то $K[x]$ – тоже коммутативное кольцо.* 3) *Если K содержит единицу, то $K[x]$ – тоже содержит единицу.* 4) *Если K не имеет делителей нуля, то $K[x]$ тоже не имеет делителей нуля.*

Д о к а з а т е л ь с т в о. 1) Проверим аксиомы кольца.

С1. Ассоциативность сложения следует из равенств

$$(a+b)+c=\sum_{i=0}^{n}\left[(a_i+b_i)+c_i\right]x^i=\sum_{i=0}^{n}\left[a_i+(b_i+c_i)\right]x^i=a+(b+c),$$

где $a, b, c \in K[x]$.

С2. Коммутативность сложения следует из равенств

$$a+b=\sum_{i=0}^{n}(a_i+b_i)x^i=\sum_{i=0}^{n}(b_i+a_i)x^i=b+a.$$

С3. Нулевым элементом является нулевой многочлен, т. е. $0 \in K$.

С4. Противоположным для многочлена $a=\sum_{i=0}^{n}a_ix^i$ будет многочлен

$$-a=\sum_{i=0}^{n}(-a_i)x^i.$$

У1. Надо доказать, что для любых многочленов $a, b, c \in K[x]$ справедливо равенство

$$(a\,b)\,c=a\,(b\,c).$$

Обозначим $ab=d$, $dc=f$, $bc=g$, $ag=h$. Надо проверить, что $f=h$. Пусть a_i – коэффициенты многочлена a, b_j – коэффициенты многочлена b, c_k – коэффициенты многочлена c и т. д. Тогда

$$f_i=\sum_{k+l=i}d_k\,c_l=\sum_{k+l=i}\left(\sum_{r+s=k}a_r\,b_s\right)c_l=\sum_{r+s+l=i}(a_r\,b_s)c_l.$$

С другой стороны,

$$h_i = \sum_{p+q=i} a_p\, g_q = \sum_{p+q=i} a_p \left(\sum_{u+v=q} b_u\, c_v \right) = \sum_{p+u+v=i} a_p\, (b_u c_v).$$

Учитывая, что K – кольцо, получаем $f_i = h_i$.

СУ1. Надо доказать, что $(a+b)c = ac+bc$ для любых $a, b, c \in K[x]$. Обозначим $a + b = d$, $dc = f$, $ac = g$, $bc = h$, $g + h = t$. Покажем, что $f = t$. Имеем:

$$f_i = \sum_{k+l=i} d_k\, c_l = \sum_{k+l=i} (a_k + b_k)\, c_l = \sum_{k+l=i} (a_k\, c_l + b_k\, c_l) =$$

$$= \sum_{k+l=i} a_k\, c_l + \sum_{k+l=i} b_k\, c_l.$$

С другой стороны,

$$t_i = g_i + h_i = \sum_{k+l=i} a_k\, c_l + \sum_{k+l=i} b_k\, c_l.$$

Аксиома СУ2 проверяется аналогично.

2) Коммутативность умножения следует из определения операции умножения в $K[x]$.

3) Так как $K \subset K[x]$, то единицей в $K[x]$ является 1 из K. Действительно,

$$1 \cdot \sum_{i=1}^{n} a_i\, x^i = \sum_{i=1}^{n} a_i\, x^i.$$

4) Докажем, что $K[x]$ не содержит делителей нуля. Пусть

$$a = a_0 + a_1 x + \ldots + a_n x^n, a_n \neq 0,$$

$$b = b_0 + b_1 x + \ldots + b_m x^m, b_m \neq 0$$

– два многочлена из $K[x]$. Рассмотрим их произведение:

$$a \cdot b = (a_0 + a_1 x + \ldots + a_n x^n)(b_0 + b_1 x + \ldots + b_m x^m) = a_0 b_0 + \ldots + a_n b_m x^{n+m}.$$

Старший коэффициент $a \cdot b$ равен $a_n b_m$. Так как кольцо K без делителей нуля и $a_n \neq 0$, $b_m \neq 0$, то $a_n b_m \neq 0$. Следовательно, $ab \neq 0$. Теорема доказана.

При доказательстве последней части теоремы мы установили, что степень произведения fg равна степени f плюс степень g, т. е.

$$\deg(fg) = \deg f + \deg g, \quad f \neq 0, \quad g \neq 0.$$

Отсюда, в частности, следует, что даже если K – поле, $K[x]$ не обязано быть полем. Действительно, если $fg = 1$, то

$$\deg(fg) = \deg f + \deg g = 0.$$

Следовательно, многочлены f и g должны быть ненулевыми элементами из K. Таким образом, нами установлено

С л е д с т в и е. *Группа обратимых элементов кольца $K[x]$ совпадает с группой обратимых элементов кольца K.*

18.2. Деление с остатком. На множестве целых чисел существует операция деления с остатком, т. е. если m и $n \neq 0$ – два целых числа, то мы можем разделить m на n с остатком, т. е. представить m в виде

$$m = nq + r, \text{ где } 0 \leq r < |n|$$

для некоторого целого q и неотрицательного целого r. Для кольца многочленов тоже существует операция деления с остатком.

Т е о р е м а 2. *Пусть P – поле. Для всяких $f, g \in P[x]$, где $g \neq 0$, существуют и единственные $q, r \in P[x]$, такие, что:*

а) $f = g \cdot q + r$;

б) $r = 0$ *или* $\deg r < \deg g$.

При этом q называется *частным,* а r – *остатком* от деления f на g.

Д о к а з а т е л ь с т в о. Докажем существование. Если $\deg f < \deg g$, то можно положить $q = 0$, $r = f$. Если $\deg f \geq \deg g$, то построим последовательность многочленов f_i, $i = 0, 1, \ldots$, положив $f_0 = f$ и для каждого $i \geq 0$ определив

$$f_{i+1} = f_i - \frac{\text{старший коэффициент } f_i}{\text{старший коэффициент } g} \cdot g\, x^{\deg(f_i) - \deg(g)}.$$

Полагая

$$h_i = \frac{\text{старший коэффициент } f_i}{\text{старший коэффициент } g} \cdot x^{\deg(f_i) - \deg(g)},$$

запишем многочлен f_{i+1} в виде $f_{i+1} = f_i - h_i g$. Видим, что степени этих многочленов убывают:

$$\deg f_0 > \deg f_1 > \deg f_1 > \ldots$$

Следовательно, не позже чем, через $n = \deg f$ шагов, мы получим многочлен f_k, который либо равен нулю, либо его степень будет меньше степени многочлена g.

Сложим все равенства

$$\begin{aligned} f_0 &= f, \\ f_1 &= f_0 - g\,h_0, \\ f_2 &= f_1 - g\,h_1, \\ &\ldots\ldots\ldots\ldots \\ f_k &= f_{k-1} - g\,h_{k-1}, \end{aligned}$$

получим

$$f_0 + f_1 + \ldots + f_k = f + f_0 + f_1 + \ldots + f_{k-1} - g(h_0 + h_1 + \ldots + h_{k-1}).$$

Отсюда

$$f = g(h_0 + h_1 + \ldots + h_{k-1}) + f_k.$$

Положим $(h_0 + h_1 + \ldots + h_{k-1}) = q$, $f_k = r$. Ясно, что r и q искомые и для них выполняются условия а и б.

Докажем единственность. Пусть имеются две пары многочленов (q_1, r_1) и (q_2, r_2), удовлетворяющие условию теоремы, т. е.

$$f = g \cdot q_1 + r_1, \quad \text{где } r_1 = 0 \text{ или } \deg r_1 < \deg g,$$

$$f = g \cdot q_2 + r_2, \quad \text{где } r_2 = 0 \text{ или } \deg r_2 < \deg g.$$

Вычитая одно равенство из другого, получим

$$g(q_1 - q_2) = r_2 - r_1.$$

Если $r_2 - r_1 \neq 0$, то $q_1 - q_2 \neq 0$. Так как эти многочлены отличны от нуля, рассмотрим их степени. Степень левой части равна $\deg g + \deg(q_1 - q_2) \geq \deg g$, а степень правой части меньше $\deg g$. Приходим к противоречию. Значит, $r_2 - r_1 = 0$, а тогда $q_1 - q_2 = 0$ (так как кольцо $P[x]$ без делителей нуля). Теорема доказана.

18.3. Наибольший общий делитель двух многочленов. Пусть f – некоторый многочлен из кольца $K[x]$. Говорим, что многочлен $d \in K[x]$ является *делителем* многочлена f (обозначаем $d|f$), если $f = d \cdot h$ для некоторого многочлена $h \in K[x]$. Говорим, что d – *наибольший общий делитель* многочленов f и g, если:

а) $d|f$ и $d|g$;

б) если некоторый многочлен $d' \in K[x]$ является делителем f и g, то $d'|d$.

Заметим, что если d – наибольший общий делитель, то kd – тоже наибольший общий делитель для любого $k \in K^*$, т. е. наибольший общий делитель определяется неоднозначно. Символом (f, g) будем обозначать *приведенный наибольший общий делитель* многочленов f и g, т. е. наибольший общий делитель со старшим коэффициентом 1.

Т е о р е м а 3. *Пусть P – поле. Для любых ненулевых многочленов $f, g \in P[x]$ справедливы следующие утверждения.* 1) *В $P[x]$ существует наибольший общий делитель многочленов f и g.* 2) *Приведенный наибольший общий делитель многочленов f и g единственный.* 3) *Наибольший общий делитель может быть найден при помощи алгоритма Евклида и поэтому не изменится, если мы будем рассматривать многочлены над большим полем.*

Д о к а з а т е л ь с т в о. 1) Применяя алгоритм Евклида к f и g, получим:

$$\begin{cases} f = g \cdot q_1 + r_1, & \deg r_1 < \deg g, \\ g = r_1 \cdot q_2 + r_2, & \deg r_2 < \deg r_1, \\ r_1 = r_2 \cdot q_3 + r_3, & \deg r_3 < \deg r_2, \\ \ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots \\ r_{k-3} = r_{k-2} \cdot q_{k-1} + r_{k-1}, & \deg r_{k-1} < \deg r_{k-2}, \\ r_{k-2} = r_{k-1} \cdot q_k + r_k, & \deg r_k < \deg r_{k-1}, \\ r_{k-1} = r_k \cdot q_{k+1}. \end{cases}$$

Тогда последний ненулевой остаток r_k и будет наибольшим общим делителем. Действительно, проверим выполнение условий из определения наибольшего общего делителя.

а) Просматривая систему равенств снизу вверх, из последнего равенства видим, что $r_k|r_{k-1}$, тогда из предпоследнего $r_k|r_{k-2}$, и т. д, наконец, из первых двух равенств заключаем, что $r_k|g$ и $r_k|f$. Следовательно, r_k делит f и g.

б) Пусть $d'|f$ и $d'|g$. Тогда из первого равенства следует, что d' делит r_1, из второго – что d' делит r_2 и т. д. Следовательно, d' делит r_k.

2) Пусть d_1 и d_2 – приведенные наибольшие общие делители f и g. Тогда по пункту б $d_1|d_2$ и $d_2|d_1$. Отсюда $\deg d_2 \leq \deg d_1$ и $\deg d_1 \leq \deg d_2$. Следовательно, $\deg d_1 = \deg d_2$. Предположим, что $d_2 = d_1 \cdot h$, где h – многочлен нулевой степени, т. е. элемент из P, а так как d_1 и d_2 приведены, то $h = 1$, а потому $d_1 = d_2$.

3) Рассмотрим некоторое поле L, содержащее поле P. Тогда $P[x] \subset L[x]$, но если мы рассматриваем многочлены $f, g \in P[x]$ и применяем к ним алгоритм Евклида, то наибольший общий делитель над P остается точно таким же и для поля L. Теорема доказана.

Покажем, что с изменением поля P делители многочленов меняются.

П р и м е р. Рассмотрим поля

$$\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Для колец многочленов имеем включения

$$\mathbb{Q}[x] \subset \mathbb{R}[x] \subset \mathbb{C}[x].$$

Следовательно, многочлен с рациональными коэффициентами можно рассматривать как многочлен с действительными и комплексными коэффициентами, но, как видно из следующей таблицы, делители могут быть разными.

Многочлен	$\mathbb{Q}[x]$	$\mathbb{R}[x]$	$\mathbb{C}[x]$
x^2-2	$1, x^2-2$	$1, x^2-2, x \pm \sqrt{2}$	$1, x^2-2, x \pm \sqrt{2}$
x^2+1	$1, x^2+1$	$1, x^2+1$	$1, x^2+1, x \pm i$

Видим, что с увеличением поля число делителей может увеличиваться.

§ 19. Линейное уравнение первой степени

19.1. Критерий разрешимости. В кольце многочленов $P[x]$ над полем P рассмотрим уравнение

$$f \cdot u + g \cdot v = h, \quad f, g, h \in P[x] \tag{1}$$

с неизвестными $u, v \in P[x]$. Следующая теорема дает необходимое и достаточное условие разрешимости этого уравнения.

Т е о р е м а 1. *Уравнение* (1) *имеет решение тогда и только тогда, когда наибольший общий делитель многочленов* f *и* g *делит* h.

Д о к а з а т е л ь с т в о. Предположим, что уравнение (1) имеет решение u_0, v_0. Тогда справедливо равенство

$$f \cdot u_0 + g \cdot v_0 = h.$$

Так как приведенный наибольший общий делитель (f, g) делит f и g, то (f, g) делит и h.

Обратно. Предположим, что $(f, g) \mid h$. Применяя алгоритм Евклида, найдем (f, g):

$$\begin{cases} f = g \cdot q_1 + r_1, & \deg r_1 < \deg g, \\ g = r_1 \cdot q_2 + r_2, & \deg r_2 < \deg r_1, \\ r_1 = r_2 \cdot q_3 + r_3, & \deg r_3 < \deg r_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots \\ r_{k-3} = r_{k-2} \cdot q_{k-1} + r_{k-1}, & \deg r_{k-1} < \deg r_{k-2}, \\ r_{k-2} = r_{k-1} \cdot q_k + r_k, & \deg r_k < \deg r_{k-1}, \\ r_{k-1} = r_k \cdot q_{k+1}. \end{cases}$$

Деля r_k на коэффициент при старшей степени получим приведенный наибольший общий делитель (f, g).

Обозначим

$$I = \{f\,a + g\,b \mid a, b \in P[x]\}.$$

Очевидно, множество I удовлетворяет следующим условиям:

1) если $h_1, h_2 \in I$, то разность $h_1 - h_2 \in I$;

2) если $h \in I$, $t \in P[x]$, то произведения $h \cdot t \in I$, $t \cdot h \in I$.

Подмножество кольца, удовлетворяющее условиям 1–2, называется *идеалом.* Так как $f = f \cdot 1 + g \cdot 0$, $g = f \cdot 0 + g \cdot 1$, то многочлены f и g лежат в I. Отсюда по алгоритму Евклида $r_1 \in I$, $r_2 \in I$, …, $r_k \in I$. Приведенный наибольший общий делитель (f, g) получается из r_k делением на его старший коэффициент, а потому и $(f, g) \in I$. Следовательно, $h \in I$, так как h делится на (f, g), а потому найдутся $u, v \in P[x]$, при которых $f \cdot u + g \cdot v = h$. Таким образом, уравнение (1) имеет решение. Теорема доказана.

19.2. Взаимно простые многочлены. Если $(f, g) = 1$, т. е. приведенный наибольший общий делитель многочленов f и g равен 1, то говорим, что f и g *взаимно просты.* Следующая теорема описывает свойства взаимно простых многочленов.

Т е о р е м а 2. *Справедливы следующие утверждения*:

а) $(f, g) = 1$ *тогда и только тогда, когда существуют* u *и* v *такие, что* $f \cdot u + g \cdot v = 1$;

б) *если* $(f, \varphi) = 1$ *и* $(f, \psi) = 1$, *то* $(f, \varphi \cdot \psi) = 1$;
в) *если* $d \mid (fg)$ *и* $(d, f) = 1$, *то* $d \mid g$;
г) *если* $d_1 \mid f$, $d_2 \mid f$ *и при этом* $(d_1, d_2) = 1$, *то* $d_1 d_2 \mid f$.
Д о к а з а т е л ь с т в о. а) Рассмотрим уравнение

$$f \cdot u + g \cdot v = 1.$$

По теореме 1 это уравнение имеет решение тогда и только тогда, когда $(f, g) = 1$.

б) По пункту а существуют u_1, v_1 такие, что

$$f \cdot u_1 + \varphi \cdot v_1 = 1,$$

а также такие u_2, v_2, что

$$f \cdot u_2 + \psi \cdot v_2 = 1.$$

Перемножая эти два равенства, получим

$$f \cdot (f u_1 u_2 + \psi u_1 v_2 + \varphi v_1 u_2) + \varphi\psi \cdot v_1 v_2 = 1,$$

т. е. нашлись многочлены u, v, удовлетворяющие равенству

$$f \cdot u + \varphi\psi \cdot v = 1.$$

Тогда по пункту а $(f, \varphi\psi) = 1$.

в) Опять по пункту а существуют u, v такие, что

$$d \cdot u + f \cdot v = 1.$$

Умножим обе части этого равенства на g, получим

$$d \cdot ug + fg \cdot v = g.$$

Видим, что первое и второе слагаемое делятся на d. Следовательно, и сумма делится на d, а потому $d \mid g$.

г) Опять ввиду а существуют u, v такие, что

$$d_1 \cdot u + d_2 \cdot v = 1.$$

Умножая обе части на f, получим

$$d_1 \cdot uf + d_2 \cdot vf = f.$$

Так как $d_1|f$, то $f = d_1 f_1$ для некоторого многочлена f_1. Аналогично из того, что $d_2|f$, заключаем, что $f = d_2 f_2$ для некоторого многочлена f_2. Подставив эти выражения для f в левую часть предыдущего равенства, получим

$$d_1 d_2 \cdot u f_2 + d_1 d_2 \cdot v f_1 = f.$$

Видим, что первое и второе слагаемое в левой части делятся на $d_1 d_2$, следовательно, $d_1 d_2 \mid f$. Теорема доказана.

19.3. Общее решение уравнения $f \cdot u + g \cdot v = 1$. Рассмотрим уравнение

$$f \cdot u + g \cdot v = 1, \quad (f, g) = 1. \tag{2}$$

Очевидно, что, научившись решать такие уравнения, мы сможем решать и уравнения с произвольной правой частью. Справедлива

Т е о р е м а 3. 1) *Общее решение уравнения* (2) *имеет вид* $(u_0 + gt, v_0 - ft)$, *где* (u_0, v_0) – *некоторое частное решение, а* t – произвольный многочлен из $P[x]$. 2) *Если степени* f *и* g *больше нуля, то существует единственное решение* (u, v) *с условием: степень* u *меньше степени* g, *а степень* v *меньше степени* f.

Д о к а з а т е л ь с т в о. 1) То, что пара $(u_0 + gt, v_0 - ft)$ является решением уравнения (2), проверяется прямой подстановкой. Проверим, что любое наперед заданное решение представимо в таком виде. Пусть (u_*, v_*) – некоторое решение уравнения (2). Имеем два равенства:

$$f \cdot u_* + g \cdot v_* = 1,$$
$$f \cdot u_0 + g \cdot v_0 = 1.$$

Вычитая из первого второе, получим

$$f \cdot (u_* - u_0) = g \cdot (v_0 - v_*). \tag{3}$$

Видим, что этот многочлен делится на f и на g, т. е.

$$f \mid g(v_0 - v_*), \quad g \mid f(u_* - u_0).$$

Учитывая, что $(f, g) = 1$ по пункту в теоремы 2 имеем $f \mid (v_0 - v_*)$, т. е. $v_0 - v_* = ft$ для некоторого многочлена $t \in P[x]$. Аналогично $u_* - u_0 = gt_1$ для некоторого многочлена t_1, но, учитывая (3), заключаем, что $t = t_1$. Следовательно,

$$(u_*, v_*) = (u_0 + gt, v_0 - ft).$$

2) Пусть $\deg f > 0$, $\deg g > 0$ и (u_0, v_0) – какое-нибудь решение уравнения (2). Поделим u_0 с остатком на g, а v_0 – на f, получим:

$$u_0 = g \cdot u_1 + u_2, \quad \text{где } u_2 = 0, \text{ или } \deg u_2 < \deg g;$$

$$v_0 = f \cdot v_1 + v_2, \quad \text{где } v_2 = 0, \text{ или } \deg v_2 < \deg f.$$

Заметим, что остатки u_2 и v_2 ненулевые. Действительно, если $u_2 = 0$, то из равенства

$$f \cdot u_0 + g \cdot v_0 = 1$$

заключаем, что

$$fg \cdot u_1 + g \cdot v_0 = 1,$$

но это равенство невозможно, так как левая часть делится на многочлен g ненулевой степени, а правая – нет. Аналогично проверяется, что $v_2 \neq 0$.

Докажем, что пара (u_2, v_2) является решением уравнения (2). Имеем:

$$1 = f \cdot u_0 + g \cdot v_0 = f(gu_1 + u_2) + g(fv_1 + v_2) = fg(u_1 + v_1) + fu_2 + gv_2,$$

т. е.

$$fg(u_1 + v_1) = 1 - fu_2 - gv_2. \tag{4}$$

Предположим, что обе части равенства (4) ненулевые. Тогда степень левой части равна

$$\deg f + \deg g + \deg(u_1 + v_1) \geq \deg f + \deg g,$$

а степень правой – меньше чем

$$\deg f + \deg g.$$

Противоречие. Следовательно, обе части равенства (4) равны нулю, т. е.

$$1 - fu_2 - gv_2 = 0,$$

а потому

$$fu_2 + gv_2 = 1.$$

Заметим, что любое другое решение не годится, что следует из пункта 1. Теорема доказана.

У п р а ж н е н и е. Что можно сказать про аналог уравнения Пелля

$$u^2 - f \cdot v^2 = 1, \quad f \in P[x],$$

в кольце $P[x]$?

§ 20. Корни и значения многочлена

20.1. Теорема Безу. До сих пор мы смотрели на многочлены как на формальные выражения, которые можно складывать и умножать. Существует и другая точка зрения, рассматривающая многочлен $f(x) \in P[x]$ как функцию $f : P \longrightarrow P$. Если

$$f(x) = a_0 + a_1x + \ldots + a_nx^n \in P[x],$$

то для всякого $c \in P$ *значением многочлена в точке* c назовем элемент

$$f(c) = a_0 + a_1c + \ldots + a_nc^n \in P.$$

Если $f(c) = 0$, то c называется *корнем многочлена* $f(x)$.

Следующая теорема показывает, что задача нахождения корней многочлена равносильна задаче нахождения его линейных делителей.

Т е о р е м а (Безу). *Элемент $c \in P$ является корнем многочлена $f(x) \in P[x]$ тогда и только тогда, когда $(x - c) \mid f(x)$.*

Д о к а з а т е л ь с т в о. Разделим $f(x)$ с остатком на $x - c$, получим

$$f(x) = (x - c)q(x) + r, \quad q(x) \in P[x], \quad r \in P.$$

Отсюда при $x = c$ получим $f(c) = r$. Из этого равенства и следует нужное утверждение.

20.2. Формула Тейлора. Дадим вначале

О п р е д е л е н и е. Если

$$f(x) = a_0 + a_1x + \ldots + a_nx^n \in P[x]$$

– некоторый многочлен, то его *производной* (*первой производной*) называется многочлен

$$f'(x) = a_1 + 2a_2x + \ldots + na_nx^{n-1}.$$

Производная от производной называется *второй производной* и обозначается $f''(x)$. Вообще для произвольного натурального $i > 1$ i-я производная определяется правилом

$$f^{(i)}(x) = (f^{(i-1)}(x))'.$$

Мы приводим формальное определение производной многочлена, не привлекая понятие предела и других прелестей математического анализа. Тем не менее, можно показать, что привычные формулы для производных справедливы и в нашем случае.

У п р а ж н е н и е. Проверьте следующие равенства:

1) $(f+g)' = f' + g'$;

2) $(f \cdot g)' = f'g + g'f$;

3) $(cf)' = cf'$, $c \in P$.

Т е о р е м а 1. *Пусть P – поле нулевой характеристики. Если f – многочлен степени n из $P[x]$, то для всякого элемента $c \in P$ справедливо равенство*

$$f(x) = \sum_{i=0}^{n} \frac{f^{(i)}(c)}{i!}(x-c)^i.$$

Эта формула называется *формулой Тейлора.*

Д о к а з а т е л ь с т в о. Положим

$$f(x) = b_0 + b_1(x-c) + \ldots + b_n(x-c)^n.$$

Тогда

$$\begin{aligned}
&f'(x) = b_1 + 2b_2(x-c) + \ldots + nb_n(x-c)^{n-1},\\
&f''(x) = 2b_2 + 3\cdot 2b_3(x-c) + \ldots + n(n-1)b_n(x-c)^{n-2},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&f^{(i)}(x) = i!b_i + \ldots + n(n-1)\ldots(n-i+1)b_n(x-c)^{n-i},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&f^{(n)}(x) = n!b_n.
\end{aligned}$$

При $x = c$ в этих формулах остаются только свободные члены:

$$f^{(i)}(c) = i!\, b_i, \quad i = 0, 1, \ldots, n.$$

Значит,

$$b_i = \frac{f^{(i)}(c)}{i!}, \quad i = 0, 1, \ldots, n.$$

Теорема доказана.

20.3. Интерполяционная формула Лагранжа.

З а д а ч а и н т е р п о л я ц и и. Пусть задано $n+1$ различных элементов $x_0, x_1, \ldots, x_n$ поля P и $n+1$ элементов $y_0, y_1, \ldots, y_n$ из P.

Требуется найти такую функцию $f(x) : P \longrightarrow P$, для которой выполняются равенства $y_i = f(x_i)$, $i = 0, 1, \ldots, n$. Наглядно это можно представить в виде следующей таблицы.

x	x_0	x_1	$\ldots$	x_i	$\ldots$	x_n
$f(x)$	y_0	y_1	$\ldots$	y_i	$\ldots$	y_n

Понятно, что при такой постановке задача имеет множество решений. Если же искать функцию в виде многочлена, то можно показать, что существует единственный многочлен степени не выше n, удовлетворяющий условиям этой задачи. Действительно, будем искать $f(x)$ в виде многочлена

$$f(x) = a_0 + a_1 x + \ldots + a_n x^n \in P[x]$$

с неизвестными коэффициентами a_i. Подставляя в него x_j, $j = 0, 1, \ldots, n$, получим систему $n + 1$ уравнения с $n + 1$ неизвестным:

$$\begin{cases} a_0 + a_1 x_0 + \ldots + a_n x_0^n = y_0, \\ a_0 + a_1 x_1 + \ldots + a_n x_1^n = y_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0 + a_1 x_n + \ldots + a_n x_n^n = y_n. \end{cases}$$

Решая эту систему, найдем искомый многочлен. Ввиду того, что определитель системы отличен от нуля, она имеет единственное решение.

Существует готовая формула (*формула Лагранжа*), позволяющая сразу написать этот многочлен:

$$f(x) = \sum_{i=0}^{n} y_i \cdot \frac{(x - x_0)(x - x_1)\ldots(x - x_{i-1})(x - x_{i+1})\ldots(x - x_n)}{(x_i - x_0)(x_i - x_1)\ldots(x_i - x_{i-1})(x_i - x_{i+1})\ldots(x_i - x_n)}.$$

Легко проверить, что $y_i = f(x_i)$, $i = 0, 1, \ldots, n$.

20.4. Кратные корни. Как следует из теоремы Безу, если c является корнем многочлена, то он делится на $x - c$. Может случиться, что многочлен делится не только на $x - c$, но и на некоторую его степень.

О п р е д е л е н и е. Если

$$f(x) = (x - c)^k g(x)$$

и $g(c) \neq 0$, то c называется *k-кратным корнем многочлена* $f(x)$. Если $k = 1$, то говорят, что c – *простой корень.*

Т е о р е м а 2. *Над полем характеристики нуль k-кратный корень многочлена является $k-1$-кратным корнем его производной.*

Д о к а з а т е л ь с т в о. Пусть

$$f(x) = (x-c)^k g(x), \quad g(c) \neq 0.$$

Тогда

$$f'(x) = k(x-c)^{k-1}g(x) + (x-c)^k g'(x) = (x-c)^{k-1}[kg(x) + (x-c)g'(x)].$$

Нетрудно проверить, что выражение в квадратных скобках не обращается в нуль при $x = c$.

У п р а ж н е н и е. Для полей ненулевой характеристики теорема неверна.

§ 21. Кольца с однозначным разложением

21.1. Определения и примеры. Из основной теоремы арифметики следует, что всякое целое число a единственным способом представимо в виде произведения простых чисел:

$$a = \varepsilon p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}, \quad \varepsilon = \pm 1, \alpha_i \in \mathbb{N},$$

где p_i – простые числа. Возникает естественный вопрос: верно ли аналогичное утверждение для кольца многочленов $P[x]$? В настоящем параграфе дается утвердительный ответ на этот вопрос.

Пусть K – *целостное кольцо*, т. е. коммутативное кольцо без делителей нуля. Предположим также, что K содержит единицу. Очевидно, что этим условиям удовлетворяет, в частности, кольцо целых чисел и кольцо многочленов над полем. Далее в этом параграфе будем считать, что K – целостное кольцо с единицей.

О п р е д е л е н и е. Элементы a и b из K называются *ассоциированными,* если $a = b \cdot \varepsilon$, где ε – обратимый элемент кольца K.

О п р е д е л е н и е. Ненулевой необратимый элемент a из K называется *неразложимым,* если из равенства $a = b \cdot c$ следует, что либо b обратимый, либо c обратимый.

О п р е д е л е н и е. Целостное кольцо K с единицей называется *кольцом с однозначным разложением,* если:

а) всякий ненулевой необратимый элемент из K разлагается в произведение неразложимых множителей;

б) если $a = p_1 p_2 \dots p_r = q_1 q_2 \dots q_s$ – два разложения на неразложимые множители, то $r = s$ и, возможно после перенумерации сомножителей, p_1 ассоциирован q_1, p_2 ассоциирован q_2 и т. д., наконец, p_r ассоциирован q_r.

П р и м е р ы. 1) В кольце $\mathbb{Z}$ обратимыми являются элементы $-1, 1$, ассоциированные элементы: a и $-a$; неразложимые элементы: $\pm p$, где p – простое число.

2) В кольце $P[x]$ обратимыми являются элементы $\alpha \in P$, $\alpha \neq 0$, ассоциированные элементы: $f, \alpha f$, $\alpha \neq 0$; неразложимые элементы: неразложимые элементы кольца многочленов $P[x]$. Как мы видели ранее, множество неразложимых элементов зависит от поля P.

Оба кольца $\mathbb{Z}$ и $P[x]$ являются кольцами с однозначным разложением. Для кольца $\mathbb{Z}$ это следует из основной теоремы арифметики, а для кольца $P[x]$ мы докажем это ниже.

21.2. Кольцо многочленов как кольцо с однозначным разложением. Для доказательства основного утверждения настоящего пункта нам потребуется

Л е м м а 1. *Если $f \in P[x]$, p – неразложим в $P[x]$, то либо $p \mid f$, либо $(p, f) = 1$.*

Д о к а з а т е л ь с т в о. Пусть $d = (p, f)$, т. е. $p = d \cdot h$ для некоторого многочлена $h \in P[x]$. Так как p неразложим, то либо $d \in P^*$, либо $h \in P^*$. Если $d \in P^*$, то $d = 1$, так как (p, f) – приведенный наибольший общий делитель. Если $h \in P^*$, то $d = \frac{1}{h} p$, т. е. $p \mid f$. Лемма доказана.

Теперь мы готовы доказать следующее утверждение.

Т е о р е м а. *Если P – поле, то $P[x]$ – кольцо с однозначным разложением.*

Д о к а з а т е л ь с т в о. Для доказательства надо проверить условия а и б из определения кольца с однозначным разложением. Пусть f – ненулевой необратимый элемент из $P[x]$.

а) Если f неразложим, то доказывать нечего, в противном случае разлагаем f в произведение двух многочленов: $f = f_1 \cdot f_2$, где $\deg f_1 < \deg f$, $\deg f_2 < \deg f$. Если какой-то f_i разложим, то разлагаем его: $f_i = f_{i1} \cdot f_{i2}$, где $\deg f_{i1} < \deg f_i$, $\deg f_{i2} < \deg f_i$ и т. д. Видим, что на каждом шаге мы имеем многочлены меньших степеней, а потому процесс оборвется, т. е. на некотором шаге получим разложение f в произведение неразложимых множителей.

б) Пусть

$$f = p_1 p_2 \dots p_r = q_1 q_2 \dots q_s$$

– два разложения в произведение неразложимых множителей. Надо доказать, что $r = s$ и после перенумерации $p_i = q_i \cdot \varepsilon_i$, $i = 1, 2, \dots, r$, где ε_i – обратимый элемент. Доказательство проведем индукцией по r. При $r = 1$ нужное утверждение следует из определения неразложимого элемента.

Предположим, что утверждение доказано для $r-1$ и надо установить его для r. Имеем $p_1 \mid q_1 q_2 \dots q_s$. Покажем, что p_1 делит некоторый q_i. Действительно, по лемме 1 либо $p_1 \mid q_1$, либо $(p_1, q_1) = 1$. Если p_1 делит q_1, то это нас устраивает. Если $(p_1, q_1) = 1$, то $p_1 \mid q_2 \dots q_s$ (по теореме о взаимной простоте). Опять по лемме 1 либо $p_1 \mid q_2$, либо $p_1 \mid q_3 \dots q_s$ и т. д. Через несколько шагов получим, что $p_1 \mid q_i$ для некоторого i. Выполняя перенумерацию сомножителей, будем считать, что $p_1 \mid q_1$, т. е. $q_1 = p_1 \varepsilon_1$, где $\varepsilon_1 \in P^*$, а это значит, что p_1 и q_1 ассоциированы. Так как $P[x]$ без делителей нуля, то, сокращая обе части нашего равенства на p_1, приходим к равенству

$$p_2 \dots p_r = \varepsilon_1 q_2 \dots q_s.$$

По индукционному предположению $r = s$ и p_2 ассоциирован с $\varepsilon_1 q_2$, который ассоциирован с q_2, p_3 ассоциирован с q_3 и т. д., наконец, p_r ассоциирован с q_r. Теорема доказана.

П р и м е р. Если рассматривать многочлен $x^2 - 2$ как многочлен из $\mathbb{Q}[x]$, то он неразложим. Если рассматривать его как многочлен из $\mathbb{R}[x]$, то он разложим и является произведением двух неразложимых многочленов:

$$x^2 - 2 = (x + \sqrt{2})(x - \sqrt{2}).$$

При этом по-другому его можно представить в виде

$$x^2 - 2 = [r(x + \sqrt{2})][r^{-1}(x - \sqrt{2})], \quad r \in \mathbb{R}$$

и легко заметить, что $x + \sqrt{2}$ ассоциирован $r(x + \sqrt{2})$, а $x - \sqrt{2}$ ассоциирован $r^{-1}(x - \sqrt{2})$.

21.3. Примеры целостных колец, не являющихся кольцами с однозначным разложением.

П р и м е р 1. В этом примере разложение на неразложимые сомножители не обрывается (не выполняется условие а определения кольца с однозначным разложением).

Пусть $K = \mathbb{Z}[2^{\frac{1}{2}}, 2^{\frac{1}{4}}, 2^{\frac{1}{8}}, \ldots, 2^{\frac{1}{2^i}}, \ldots]$ – подкольцо поля $\mathbb{R}$, порожденное множеством целых чисел $\mathbb{Z}$ и числами $2^{\frac{1}{2}}, 2^{\frac{1}{4}}, 2^{\frac{1}{8}}, \ldots, 2^{\frac{1}{2^i}}, \ldots$ То, что K является целостным кольцом, т. е. коммутативным кольцом без делителей нуля, следует из включения $K \subseteq \mathbb{R}$. Также очевидно, что K содержит единицу.

Чтобы понять, как устроены элементы из K, определим кольца

$$K_r = \mathbb{Z}[2^{\frac{1}{2}}, 2^{\frac{1}{4}}, 2^{\frac{1}{8}}, \ldots, 2^{\frac{1}{2^r}}], \quad r = 1, 2, \ldots$$

Очевидно,

$$K = \bigcup_{r=1}^{\infty} K_r.$$

Кроме того, легко заметить, что

$$K_r = \mathbb{Z}[2^{\frac{1}{2^r}}].$$

Если рассмотреть произвольный элемент a из K, то он лежит в некотором K_r, а потому найдется многочлен $g \in \mathbb{Z}[x]$ такой, что $a = g(\theta)$, где $\theta = 2^{\frac{1}{2^r}}$.

Покажем, что процесс разложения на неразложимые множители в K не обрывается:

$$2 = 2^{\frac{1}{2}} \cdot 2^{\frac{1}{2}} = 2^{\frac{1}{2}} \cdot 2^{\frac{1}{4}} \cdot 2^{\frac{1}{4}} = 2^{\frac{1}{2}} \cdot 2^{\frac{1}{4}} \cdot 2^{\frac{1}{8}} \cdot 2^{\frac{1}{8}} = \ldots,$$

как видно, этот процесс не оборвется. Но надо показать, что в этих разложениях нет обратимых элементов.

Л е м м а 2. *Все элементы* $2^{\frac{1}{2^m}}$, $m = 0, 1, 2, \ldots$ *необратимы в кольце* K.

Д о к а з а т е л ь с т в о. Пусть некоторый элемент $2^{\frac{1}{2^m}}$ обратим, т. е. найдется многочлен $g(x) \in \mathbb{Z}[x]$ и элемент $\theta = 2^{\frac{1}{2^r}}$ такие, что $2^{\frac{1}{2^m}} \cdot g(\theta) = 1$. Можно считать, что $r \geq m$. Тогда $f(\theta) = 1$ для многочлена $f(x) = x^{2^{r-m}} g(x)$ из $\mathbb{Z}[x]$. При этом $f(x)$ – многочлен без свободного члена. С другой стороны, $\theta^{2^r} = 2$.

Рассмотрим многочлены $f(x) - 1$ и $x^{2^r} - 2$. Для них θ является корнем. Поэтому по теореме Безу

$$(x - \theta) \mid (f(x) - 1) \ \text{ и } \ (x - \theta) \mid (x^{2^r} - 2),$$

а потому они не взаимно просты, т. е. $(f(x) - 1, x^{2^r} - 2) \neq 1$. Покажем, что $x^{2^r} - 2$ неразложим в $\mathbb{Z}[x]$. Пусть, напротив,

$$x^{2^r} - 2 = g_1 \cdot g_2, \quad g_i \in \mathbb{Z}[x], \quad \deg g_i > 0.$$

Определим отображение $\mathbb{Z}[x] \longrightarrow \mathbb{Z}_2[x]$, где $\mathbb{Z}_2 = \{\overline{0}, \overline{1}\}$ – поле, состоящее из двух элементов, как отображение, переводящее многочлен

$$a = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \in \mathbb{Z}[x]$$

в многочлен

$$\overline{a} = \overline{a}_0 + \overline{a}_1 x + \overline{a}_2 x^2 + \ldots + \overline{a}_n x^n \in \mathbb{Z}_2[x],$$

где $\overline{a}_i$ – остаток от деления a_i на 2. Тогда равенство

$$x^{2^r} - 2 = g_1 \cdot g_2$$

перейдет в равенство

$$x^{2^r} = \overline{g_1} \cdot \overline{g_2}.$$

П р и м е р. Равенство

$$(x^2 - 3\,x + 2)(x^4 - 1) = x^6 - 3\,x^5 + 2\,x^4 - x^2 + 3\,x - 2$$

переходит в равенство

$$(x^2 + x)(x^4 + \overline{1}) = x^6 + x^5 + x^2 + x.$$

Нетрудно показать, что $\overline{g_1} = x^s$ и $\overline{g_2} = x^{2^r - s}$ для некоторого натурального s. Значит, в g_1 и g_2 все коэффициенты, кроме старших, четные. В частности, свободные члены многочленов g_i четные. При перемножении многочленов свободные члены перемножаются и свободный член произведения $g_1 \cdot g_2$ делится на 4, а в левой части нашего равенства

$$x^{2^r} - 2 = g_1 \cdot g_2$$

свободный член не делится на 4. Противоречие. Следовательно, $x^{2^r} - 2$ неразложим в $\mathbb{Z}[x]$.

Теперь мы хотим воспользоваться леммой 1, но в ней требуется, чтобы многочлен был неразложим в $P[x]$, где P – поле.

Следующее утверждение будет доказано позже (см. следствие леммы 5 из § 29).

Л е м м а 3. *Многочлен из $\mathbb{Z}[x]$ неразложим в $\mathbb{Q}[x]$ тогда и только тогда, когда он неразложим в $\mathbb{Z}[x]$.*

По этой лемме $x^{2^r} - 2$ неразложим в $\mathbb{Q}[x]$, а потому ввиду леммы 1

$$(x^{2^r} - 2) \mid (f(x) - 1),$$

т. е.

$$f(x) - 1 = (x^{2^r} - 2) \cdot h(x),$$

где $h(x) \in \mathbb{Z}[x]$. Следовательно, свободный член многочлена $f(x) - 1$ равен свободному члену многочлена $(x^{2^r} - 2) \cdot h(x)$, но это невозможно, так как свободный член первого равен -1, а свободный член второго делится на 2. Следовательно, все элементы в нашем разложении необратимы и процесс разложения на неразложимые множители не обрывается.

П р и м е р 2. В этом примере существует разложение на неразложимые множители, но оно неоднозначно. Рассмотрим кольцо $K = \mathbb{Z}[\sqrt{-3}]$. Это подкольцо поля $\mathbb{C}$, порожденное $\mathbb{Z}$ и $\sqrt{-3}$. Нетрудно проверить, что

$$K = \{a + b\sqrt{-3} \mid a, b \in \mathbb{Z}\}.$$

Так как $K \subseteq \mathbb{C}$, то K – целостное кольцо и, очевидно, содержит единицу.

Покажем, что каждый элемент из K разлагается на неразложимые множители. Для всякого элемента $\alpha = a + b\sqrt{-3}$ из K определим норму: $N : K \longrightarrow \mathbb{N} \cup \{0\}$, полагая $N(\alpha) = a^2 + 3\,b^2$. Нетрудно проверить, что норма удовлетворяет следующему равенству:

$$N(\alpha \cdot \beta) = N(\alpha) \cdot N(\beta).$$

Действительно, если $\beta = c + d\sqrt{-3}$ – другой элемент из K, то

$$\alpha \cdot \beta = (a\,c - 3\,b\,d) + (a\,d + b\,c)\sqrt{-3},$$

и для нормы произведения справедливо равенство

$$N(\alpha \cdot \beta) = (a\,c - 3\,b\,d)^2 + 3\,(a\,d + b\,c)^2 = a^2\,c^2 + 9b^2\,d^2 + 3a^2\,d^2 + 3b^2\,c^2.$$

С другой стороны,

$$N(\alpha) \cdot N(\beta) = (a^2 + 3\,b^2)(c^2 + 3\,d^2) = a^2\,c^2 + 3a^2\,d^2 + 3b^2\,c^2 + 9b^2\,d^2,$$

т. е. $N(\alpha \cdot \beta) = N(\alpha) \cdot N(\beta)$.

Используя установленное равенство, дадим описание обратимых элементов кольца K.

Л е м м а 4. *В кольце K обратимыми являются только элементы* 1 *и* -1.

Д о к а з а т е л ь с т в о. Предположим, что $\alpha = a + b\sqrt{-3} \in K$ обратим. Тогда для него найдется $\beta \in K$ такой, что $\alpha\beta = 1$. По свойству нормы из этого равенства получим $N(\alpha) \cdot N(\beta) = 1$. Следовательно, $N(\alpha) = N(\beta) = 1$, но это означает, что $a^2 + 3\,b^2 = 1$, а это равенство выполняется лишь при условии $a = \pm 1$, $b = 0$. Лемма доказана.

Рассмотрим некоторый ненулевой необратимый элемент α из K и будем разлагать его в произведение:

$$\alpha = \alpha_1 \cdot \alpha_2 = \alpha_{11} \cdot \alpha_{12} \cdot \alpha_{21} \cdot \alpha_{22} = \dots$$

Этому разложению соответствует разложение для норм:

$$N(\alpha) = N(\alpha_1) \cdot N(\alpha_2) = N(\alpha_{11}) \cdot N(\alpha_{12}) \cdot N(\alpha_{21}) \cdot N(\alpha_{22}) = \dots$$

Как мы знаем, $N(\alpha) = 1$ тогда и только тогда, когда $\alpha = \pm 1$, т. е. является обратимым элементом в K. Учитывая, что норма принимает целые неотрицательные значения, видим, что для норм этот процесс оборвется. Таким образом, всякий элемент $\alpha \in K$ разлагается на неразложимые сомножители.

Покажем, что это разложение неединственно. Действительно,

$$4 = 2 \cdot 2 = (1 + \sqrt{-3}) \cdot (1 - \sqrt{-3}),$$

т. е. 4 имеет два разложения на множители. Ясно, что 2 не ассоциировано с $1 \pm \sqrt{-3}$. Покажем, что 2 неразложим. Предположим, что $2 = \alpha \cdot \beta$, $\alpha, \beta \in K$. Тогда $N(2) = N(\alpha) \cdot N(\beta)$. Так как $N(2) = 4$ имеет единственное нетривиальное разложение $4 = 2 \cdot 2$ в $\mathbb{Z}$, то достаточно заметить, что 2 не является нормой никакого числа из K. Если бы $N(\alpha) = a^2 + 3\,b^2 = 2$, то отсюда следовало бы, что $b = 0$, а так как a – целое число, то это равенство невозможно.

Аналогично устанавливается, что и $1 \pm \sqrt{-3}$ неразложим (заметим, что $N(1 \pm \sqrt{-3}) = 4$).

§ 22. Идеалы. Фактор-кольца

22.1. Определения и примеры. С идеалом мы фактически уже встречались, когда доказывали критерий разрешимости линейного уравнения с двумя неизвестными. Введенное там множество I являлось идеалом. Дадим формальное определение.

О п р е д е л е н и е. Подмножество I кольца K называется *идеалом* (обозначение: $I \lhd K$), если выполнены следующие два условия:

а) если $a, b \in I$, то $a - b \in I$;

б) если $a \in I$, $c \in K$, то $a \cdot c \in I$, $c \cdot a \in I$.

Из этого определения, в частности, следует, что идеал является подкольцом. С другой стороны, кольцо целых чисел является подкольцом поля рациональных чисел, но не является идеалом.

П р и м е р 1. Пусть $\mathbb{Z}$ – кольцо целых чисел. Тогда множество

$$n\mathbb{Z} = \{\text{целые числа, делящиеся на } n\}$$

является идеалом для любого целого неотрицательного n.

У п р а ж н е н и е. Докажите, что идеалы $n\mathbb{Z}$ исчерпывают все идеалы в $\mathbb{Z}$.

П р и м е р 2. Пусть $K = P[x]$, а f – некоторый многочлен из K. Тогда идеалом является множество всех многочленов, которые делятся на f, т. е.

$$I = \{f \cdot g \mid g \in K\}.$$

Ниже мы покажем, что такими идеалами исчерпываются все идеалы кольца K.

У п р а ж н е н и е. *Во всяком поле P только два идеала: нулевой и само P.*

Р е ш е н и е. Действительно, пусть I – идеал в P и $I \neq 0$. Возьмем элемент $a \in I$, $a \neq 0$. Тогда $1 = a \cdot a^{-1} \in I$, но по определению идеала отсюда следует, что $1 \cdot b$ для любого элемента $b \in P$. Следовательно, $I = P$.

22.2. Порождающее множество идеала. Так же, как и в случае колец (см. лемму 2 из § 3), доказывается

Л е м м а 1. *Пересечение любого семейства идеалов является идеалом.*

Если M – некоторое подмножество кольца K, то символом ид(M) или просто (M) обозначим пересечение всех идеалов в K, содержащих M, иными словами, (M) – наименьший идеал, содержащий множество M. Если $I = (M)$, то говорим, что I *порождается множеством* M или что M является *базой* идеала I. Если множество M конечно, то говорят, что идеал I *конечно порожден.*

Более конструктивное описание идеала (M) дает

Л е м м а 2. *Пусть K – коммутативное кольцо с единицей и $M \subseteq K$. Тогда*

$$(M) = \left\{ \sum a_i\, b_i \mid a_i \in M, b_i \in K \right\}.$$

В частности, если $M = \{a_1, a_2, \ldots, a_n\}$, то

$$(a_1, a_2, \ldots, a_n) = \left\{ \sum_{i=1}^{n} a_i\, b_i \;\middle|\; b_i \in K \right\}.$$

Д о к а з а т е л ь с т в о. Включение справа налево очевидно. Проверим, что множество сумм, стоящих в правой части, действительно образуют идеал:

а) так как $\sum a_i\, b_i - \sum a_i\, b'_i = \sum a_i\, (b_i - b'_i)$, то разность является такой же суммой;

б) так как $(\sum a_i\, b_i) \cdot c = \sum a_i\, (b_i c)$, то умножение на c переводит сумму в аналогичную сумму.

Таким образом, мы установили, что множество

$$\left\{ \sum a_i\, b_i \;\middle|\; a_i \in M, b_i \in K \right\}$$

является наименьшим идеалом, содержащим M.

О п р е д е л е н и е. Идеал, порожденный одним элементом, называется *главным.*

П р и м е р. Очевидно, $n\mathbb{Z} = (n)$, n – порождающий идеала $n\mathbb{Z}$. При этом

$$n\mathbb{Z} = \{0, \pm n, \pm 2n, \ldots\}.$$

22.3. Фактор-кольца. Пусть K – кольцо, I – идеал в K и a, b – два элемента из K. Говорим, что *a сравним с b по модулю идеала I*, и пишем: $a \equiv b \pmod{I}$, или просто $a \equiv b$, если $a - b \in I$. Заметим, что отношение $\equiv$ является отношением эквивалентности. Действительно,

1) так как I – подкольцо, то $0 = a - a \in I$, т. е $a \equiv a$;

2) если $a - b \in I$, то и $-(a - b) = b - a \in I$, т. е. из того, что $a \equiv b$, следует, что $b \equiv a$;

3) если $a - b \in I$ и $b - c \in I$, то $(a - b) + (b - c) = a - c \in I$, т. е. из того, что $a \equiv b$ и $b \equiv c$, следует, что $a \equiv c$.

Заметим, что при доказательстве этих пунктов мы использовали лишь то, что I подкольцо.

Как мы знаем, множество с определенным на нем отношением эквивалентности разбивается на классы эквивалентности. Множество K распадается на смежные классы:

$$K/I = \{\text{смежные классы кольца } K \text{ по идеалу } I\}.$$

Легко видеть, что всякий смежный класс

$$K_a = \{x \in K \mid x \equiv a (\mathrm{mod}\ I)\}$$

имеет вид

$$K_a = a + I,$$

где мы обозначаем

$$a + I = \{a + i \mid i \in I\}.$$

На множестве всех смежных классов можно ввести операции сложения и умножения.

О п р е д е л е н и е. Для смежных классов $a + I$ и $b + I$ из K/I положим

$$(a + I) + (b + I) = a + b + I,$$
$$(a + I) \cdot (b + I) = a \cdot b + I.$$

Справедлива

Т е о р е м а 1. 1) *Сложение и умножение смежных классов не зависят от случайного выбора представителей в смежных классах.* 2) *Множество K/I с операциями сложения и умножения является кольцом.* Оно называется *фактор-кольцом кольца K по идеалу I.*

Д о к а з а т е л ь с т в о. 1) Пусть $a + I = a' + I$ и $b + I = b' + I$. Надо доказать, что

$$a + b + I = a' + b' + I,$$
$$a \cdot b + I = a' \cdot b' + I.$$

Иными словами, из того, что $a - a' \in I$ и $b - b' \in I$, надо доказать, что

$$(a + b) - (a' + b') \in I,$$
$$a \cdot b - a' \cdot b' \in I.$$

Так как сумма элементов из идеала лежит в идеале, то

$$(a - a') + (b - b') = (a + b) - (a' + b') \in I,$$

и первое включение установлено.

Рассмотрим далее

$$a \cdot b - a' \cdot b' = a \cdot b - a' \cdot b + a' \cdot b - a' \cdot b' = (a - a') \cdot b + a' \cdot (b - b'),$$

ввиду того, что $(a - a') \cdot b \in I$ и $a' \cdot (b - b') \in I$, заключаем, что и вся сумма в правой части лежит в идеале I. Заметим, что здесь мы использовали полное определение идеала.

2) Надо проверить аксиомы кольца. Они следуют из соответствующих аксиом для K. Например, аксиома ассоциативности сложения

$$[(a + I) + (b + I)] + (c + I) = (a + I) + [(b + I) + (c + I)]$$

следует из равенства

$$(a + b) + c + I = a + (b + c) + I.$$

Аналогично проверяется коммутативность сложения, ассоциативность умножения, дистрибутивность.

Нулевым классом является класс $0 + I = I$.

Противоположным к классу $a + I$ будет класс $-(a + I) = -a + I$. Теорема доказана.

П р и м е р. Пусть $\mathbb{Z}$ – кольцо целых чисел, а $I = (5)$ – множество всех чисел, кратных 5. Тогда $\mathbb{Z}/I \simeq \mathbb{Z}_5$ – кольцо вычетов по модулю 5.

§ 23. Идеалы в кольце многочленов

23.1. Кольцо многочленов как кольцо главных идеалов. Как мы знаем, в кольце $\mathbb{Z}$ всякий идеал является главным. Аналогичным свойством обладает и кольцо $P[x]$.

Т е о р е м а 1. *В кольце $P[x]$ каждый идеал является главным, т. е. имеет вид*

$$(f) = \{f \cdot g \mid g \in P[x]\}$$

для некоторого f из $P[x]$.

Д о к а з а т е л ь с т в о. Пусть I – некоторый идеал в $P[x]$. Если $I = \{0\}$, то $I = (0)$. Предположим, что $I \neq (0)$, и выберем многочлен f наименьшей степени, лежащий в I. Покажем, что $I = (f)$. Действительно, включение $I \supseteq (f)$ очевидно. Пусть g – произвольный многочлен из I. Разделим g с остатком на f, получим:

$$g = fq + r, \quad \text{где } r = 0 \text{ или } \deg r < \deg f.$$

Если при этом $r \neq 0$, то $r = g - fq \in I$, что противоречит тому, что f – многочлен наименьшей степени в I. Следовательно, $g = fq$, т. е. $g \in (f)$, а потому $I = (f)$.

У п р а ж н е н и е. Укажите в кольце $P[x, y]$ идеал, который не является главным.

У к а з а н и е: рассмотрите идеал (x, y), состоящий из многочленов без свободного члена.

23.2. Кольца с условием максимальности.

Т е о р е м а 2. *Для всякого кольца K следующие условия равносильны*:

а) *всякий идеал кольца K порождается конечным множеством элементов*;

б) *всякая возрастающая цепочка идеалов*

$$I_1 \leq I_2 \leq \dots$$

стабилизируется на некотором номере n, т. е. $I_n = I_{n+1} = \dots$

При любом из этих условий K называется *кольцом с условием максимальности.* Класс всех таких колец обозначается Max.

Д о к а з а т е л ь с т в о. а) $\Rightarrow$ б). Предположим противное: существует цепочка идеалов, которая неограниченно растет:

$$I_1 < I_2 < \dots < I_n < I_{n+1} < \dots$$

Возьмем множество

$$I = \bigcup_{k=1}^{\infty} I_k.$$

Покажем, что I – идеал в K. Действительно, если $a, b \in I$, то $a \in I_n$, $b \in I_m$ для некоторых натуральных n и m. Следовательно, $a, b \in I_s$ при $s = \max\{n, m\}$. Так как I_s – идеал, то $a - b \in I_s$, а потому $a - b \in I$. Пусть теперь $a \in I$, $c \in K$. Следовательно, $a \in I_n$ для некоторого n. Следовательно, $ac \in I_n$, а потому $ac \in I$. Таким образом, I действительно является идеалом.

Допустим, что $I = (a_1, a_2, \dots, a_m)$, т. е. I порождается конечным множеством элементов. Пусть $a_1 \in I_{n_1}$, $a_2 \in I_{n_2}$, $\dots$, $a_m \in I_{n_m}$. Положим $n = \max\{n_1, n_2, \dots, n_m\}$. Тогда идеал I_n содержит элементы $a_1, a_2, \dots, a_m$, но тогда I_n содержит и I. Значит, $I_n = I$, и, следовательно,

$$I_n = I_{n+1} = \dots$$

Противоречие.

б) $\Rightarrow$ а). От противного. Пусть идеал I не конечно порожденный. Выберем элемент $a_1 \in I$, тогда $(a_1) < I$; выберем $a_2 \in I \setminus (a_1)$, тогда $(a_1, a_2) < I$; выберем $a_3 \in I \setminus (a_1, a_2)$, тогда $(a_1, a_2, a_3) < I$, и так как идеал I не является конечно порожденным, то этот процесс можно продолжать до бесконечности. Получаем цепочку идеалов

$$(a_1) < (a_1, a_2) < (a_1, a_2, a_3) < \dots,$$

которая не стабилизируется. Противоречие. Теорема доказана.

23.3. Теорема Гильберта о базах. Если рассмотреть бесконечную систему линейных уравнений от переменных $x_1, x_2, \dots, x_n$, то она эквивалентна некоторой конечной подсистеме. Если мы рассмотрим произвольную систему полиномиальных уравнений

$$f_\alpha(x_1, \dots, x_n) = 0, \quad \alpha \in A,$$

то возникает естественный вопрос: будет ли она равносильна некоторой своей конечной подсистеме

$$g_1(x_1, \dots, x_n) = 0, \quad g_2(x_1, \dots, x_n) = 0, \dots, g_s(x_1, \dots, x_n) = 0?$$

Положительный ответ на этот вопрос следует из теоремы Гильберта.

Т е о р е м а (Д. Гильберт, 1890). *Пусть K – коммутативное кольцо с единицей. Если $K \in$ Max, то $K[x] \in$ Max.*

Д о к а з а т е л ь с т в о теоремы проведем от противного. Допустим, что K – коммутативное кольцо с единицей, удовлетворяющее условию максимальности ($K \in$ Max), но $K[x] \notin$ Max. Следовательно, найдется цепочка идеалов, которая не стабилизируется, или, что то же самое, I – идеал в $K[x]$, который не конечно порожден. Выберем в I множество многочленов, полагая $f_0 = 0$, и для каждого $i = 0, 1, \dots$ выберем многочлен f_{i+1} наименьшей степени, который не лежит в идеале $(f_0, f_1, \dots, f_i)$, т. е. $f_{i+1} \in I \setminus (f_0, f_1, \dots, f_i)$. Пусть n_i – степень многочлена f_i, а a_i – его старший коэффициент, $i = 1, 2, \dots$ Очевидно,

$$n_1 \le n_2 \le \dots$$

Достаточно доказать, что цепочка идеалов

$$(a_1) < (a_1, a_2) < \dots < (a_1, a_2, \dots, a_i) < (a_1, a_2, \dots, a_i, a_{i+1}) < \dots$$

не стабилизируется. Пусть, напротив,

$$(a_1, a_2, \ldots, a_i) = (a_1, a_2, \ldots, a_i, a_{i+1})$$

при некотором i, тогда

$$a_{i+1} = \sum_{k=1}^{i} a_k\, b_k \text{ при подходящих } b_k \in K.$$

Рассмотрим многочлен

$$g(x) = f_{i+1}(x) - \sum_{k=1}^{i} f_k(x) b_k\, x^{n_{i+1}-n_k}.$$

Ясно, что

$$g \in I \setminus (f_0, f_1, \ldots, f_i).$$

Если бы $g \in (f_0, f_1, \ldots, f_i)$, то и $f_{i+1} \in (f_0, f_1, \ldots, f_i)$ и степень g была бы меньше степени f_{i+1}. Противоречие. Теорема доказана.

У п р а ж н е н и е. Где используется наличие единицы в кольце K?

Как мы знаем, кольцо целых чисел является кольцом с условием максимальности, так как каждый идеал порождается одним элементом. Заметим также, что поле является кольцом с условием максимальности. Действительно, мы знаем, что каждый идеал поля либо нулевой, либо совпадает со всем полем. В первом случае он порождается нулем, а во втором – единицей. Замечая, что $K[x_1, \ldots, x_n] = K[x_1, \ldots, x_{n-1}][x_n]$, индукцией по n из теоремы Гильберта получаем

С л е д с т в и е 1. *Если K – поле или кольцо $\mathbb{Z}$, то $K[x_1, \ldots, x_n] \in$* Max.

С л е д с т в и е 2. *Всякая система полиномиальных уравнений от n неизвестных над любым полем или кольцом $\mathbb{Z}$ равносильна некоторой конечной подсистеме.*

Д о к а з а т е л ь с т в о. Пусть K – поле или кольцо $\mathbb{Z}$. Пусть $M \subseteq K^n$ – это множество всех решений нашей системы:

$$f_\alpha(x_1, \ldots, x_n) = 0, \quad \alpha \in A. \tag{1}$$

Символом I обозначим множество

$$\{f \in K[x_1, \ldots, x_n] \mid f(x_1^0, x_2^0, \ldots, x_n^0) = 0 \text{ при всех } (x_1^0, x_2^0, \ldots, x_n^0) \in M\}$$

– множество всех многочленов, корни которых лежат в M. Легко проверить, что I – идеал в кольце $K[x_1, \ldots, x_n]$. По следствию 1 всякий идеал конечно порожден. Следовательно, найдутся многочлены g_1, g_2, $\ldots$, g_s из $K[x_1, \ldots, x_n]$ такие, что $(f_\alpha \mid \alpha \in A) = (g_1, g_2, \ldots, g_s)$. Покажем, что система (1) равносильна системе

$$g_1(x_1, \ldots, x_n) = 0, \quad g_2(x_1, \ldots, x_n) = 0, \ldots, g_s(x_1, \ldots, x_n) = 0, \qquad (2)$$

т. е.

$$\{\text{решение системы (1)}\} = \{\text{решение системы(2)}\}.$$

Действительно, для всякого многочлена f_α найдутся многочлены h_{α_1}, h_{α_2}, $\ldots$, h_{α_s} из $K[x_1, \ldots, x_n]$ такие, что

$$f_\alpha = g_1 h_{\alpha_1} + g_2 h_{\alpha_2} + \ldots + g_s h_{\alpha_s},$$

а потому если какая-то n-ка является решением системы (2), то она является решением системы (1). Обратное включение очевидно.

§ 24. Теорема о существовании корня

24.1. Постановка задачи. Сформулируем следующую задачу. Пусть P – поле, f – многочлен степени ≥ 1 из $P[x]$. Построить поле L, удовлетворяющее следующим условиям:

1) L содержит P как подполе;

2) в L существует элемент α такой, что $f(\alpha) = 0$, т. е. α – корень многочлена f.

Учитывая, что многочлен f можно представить в виде произведения $f = f_1 f_2 \ldots f_m$ неразложимых многочленов, можно считать, что f неразложим, так как если α – корень многочлена f_i, то α – корень и многочлена f.

Т е о р е м а о с у щ е с т в о в а н и и к о р н я. *Для всякого поля P и всякого неразложимого многочлена $f \in P[x]$ степени $n > 0$ существует поле L со свойствами* 1, 2. *Если, кроме того, выполнено условие*

3) *подполе в L, порожденное P и α, совпадает с L,*

то L единственное с точностью до изоморфизма, т. е. любые два поля со свойствами 1–3 *изоморфны.*

Доказательство теоремы разобьем на две части. Вначале покажем, что такое поле L действительно существует, а затем докажем единственность.

24.2. Существование. Рассмотрим фактор-кольцо $L' = P[x]/(f)$, где (f) – главный идеал, порожденный многочленом f. Проверим, что L' поле. Так как операции сложения и умножения определены на представителях, то все аксиомы кольца выполняются. Проверяем оставшиеся аксиомы.

У2) $(g+(f))\cdot(h+(f)) = (h+(f))\cdot(g+(f))$.

Эта аксиома выполнена в силу определения умножения смежных классов и коммутативности умножения в $P[x]$.

У3) Легко заметить, что смежный класс $1+(f)$ является единицей в L'.

У4) Пусть $g+(f) \neq (f)$, т. е. $g \notin (f)$, а потому g не делится на f. Так как f неразложим, то по лемме 1 из § 21 $(f,g) = 1$, а значит существуют $u, v \in P[x]$ такие, что

$$fu + gv = 1.$$

Тогда обратным к классу $g+(f)$ будет $(g+(f))^{-1} = v+(f)$. Действительно,

$$(g+(f))\cdot(v+(f)) = gv+(f) = 1 - fu + (f) = 1 + (f).$$

Таким образом, L' – поле.

В L' укажем подполе, изоморфное полю P. Положим

$$P' = \{a+(f) \mid a \in P\}.$$

Пусть $\omega : P \longrightarrow P'$ определяется следующим образом: $a \longmapsto a+(f)$, $a \in P$. Покажем, что ω – изоморфизм P на P', т. е.

1) ω – однозначно;
2) ω – унивалентно;
3) ω – отображение *на*;
4) $(a+b)\omega = a\omega + b\omega$;
5) $(a\cdot b)\omega = a\omega \cdot b\omega$.

1) То, что отображение ω однозначно – очевидно. 2) Если $a \neq b$, то надо показать, что $a+(f) \neq b+(f)$. Пусть $a+(f) = b+(f)$, тогда $a-b \in (f)$ и $f \mid (a-b)$, тогда $a-b = f\cdot g$ и, следовательно, $a-b = 0$ (учесть, что a и b – элементы из поля, а потому имеют нулевую степень), т. е. $a = b$. 3) Очевидно. 4) Левая часть: $(a+b)\omega = a+b+(f)$; правая часть: $a\omega + b\omega = (a+(f))+(b+(f)) = a+b+(f)$. 5) Проверяется аналогично. Таким образом, ω – изоморфизм.

Теперь возьмем $L = (L'\backslash P')\cup P$ с перенесенными из L' операциями:

$$A+B=\begin{cases} a+b & \text{при } A=a\in P,\quad B=b\in P;\\ g+b+(f) & \text{при } A=g+(f)\notin P',\quad B=b\in P;\\ a+h+(f) & \text{при } A=a\in P,\quad B=h+(f)\notin P';\\ g+h+(f) & \text{при } A=g+(f)\notin P',\quad B=h+(f)\notin P',\\ & \quad A+B\notin P';\\ c & \text{при } A=g+(f)\notin P',\quad B=h+(f)\notin P',\\ & \quad A+B=c+(f)\in P';\end{cases}$$

$$A\cdot B=\begin{cases} ab & \text{при } A=a\in P,\quad B=b\in P;\\ gb+(f) & \text{при } A=g+(f)\notin P',\quad B=b\in P;\\ ah+(f) & \text{при } A=a\in P,\quad B=h+(f)\notin P';\\ gh+(f) & \text{при } A=g+(f)\notin P',\quad B=h+(f)\notin P',\\ & \quad A\cdot B\notin P';\\ c & \text{при } A=g+(f)\notin P',\quad B=h+(f)\notin P',\\ & \quad A\cdot B=c+(f)\in P'.\end{cases}$$

Проверим, что L – искомое. То, что для него выполняется свойство 1, очевидно. Покажем, что выполняется свойство 2. Возьмем $\alpha = x+(f)$ и покажем, что $f(\alpha)=0$. Пусть

$$f(x)=a_0+a_1\,x+\ldots+a_n\,x^n,\quad a_i\in P,$$

тогда

$$f(\alpha)=a_0+a_1\,\alpha+\ldots+a_n\,\alpha^n=a_0+a_1\,(x+(f))+\ldots+a_n\,(x+(f))^n=$$

$$=a_0+a_1\,x+\ldots+a_n\,x^n+(f)=f+(f)=0.$$

У п р а ж н е н и е. Что будет, если многочлен f разложим?

24.3. Единственность. Докажем, что построенное поле единственно. Воспользуемся тем, что если два поля изоморфны одному и тому же полю, то они изоморфны между собой. Пусть M – поле, удовлетворяющее условиям 1–3, т. е.

1) $M\supseteq P$;

2) найдется элемент $\beta\in M$ такой, что $f(\beta)=0$;

3) подполе поля M, порожденное P и β, совпадает с M.

Надо доказать, что $M\simeq L'$. Предварительно заметим, что

$$M=\{g(\beta)\mid g\in P[x]\}.$$

Понятно, что

$$M \supseteq \{g(\beta) \mid g \in P[x]\},$$

т. е. всякий элемент $b_0 + b_1\,\beta + \ldots + a_n\,\beta^n$ лежит в M. Покажем, что множество

$$\{g(\beta) \mid g \in P[x]\}$$

является подполем, содержащим P и β.

Заметим, что если $g_1, g_2 \in P[x]$, то и элементы $g_1(\beta) + g_2(\beta)$, $g_1(\beta) \cdot g_2(\beta)$, $-g_1(\beta)$ лежат в нашем множестве. Пусть теперь $g(\beta) \neq 0$. Тогда $f \nmid g$ (иначе $g = f \cdot h$ и $g(\beta) = f(\beta) \cdot h(\beta) = 0$). Так как f неразложим, то опять по лемме 1 из § 21 $(f, g) = 1$. Значит, существуют u, v такие, что $fu + gv = 1$. При $x = \beta$ имеем

$$f(\beta) \cdot u(\beta) + g(\beta) \cdot v(\beta) = 1.$$

Так как $f(\beta) \cdot u(\beta) = 0$, то $v(\beta) = g(\beta)^{-1}$. Следовательно, мы установили, что множество

$$\{g(\beta) \mid g \in P[x]\}$$

является полем, а так как оно содержит P и β, то оно содержит и M.

Докажем, что M и L' изоморфны. Рассмотрим отображение

$$\varphi : L' \longrightarrow M,$$

определенное правилом

$$g + (f) \longmapsto g(\beta),$$

т. е. смежному классу $g + (f)$ сопоставим элемент $g(\beta)$ из M. Проверим, что φ – изоморфизм L' на M. Для этого надо проверить следующие условия:

1) φ – однозначно;
2) φ – унивалентно;
3) φ – отображение *на*;
4) $(A + B)\varphi = A\varphi + B\varphi$;
5) $(A \cdot B)\varphi = A\varphi \cdot B\varphi$.

1) Покажем, что от выбора представителя наше определение не зависит. Пусть $g + (f) = g_1 + (f)$. Тогда $f \mid (g - g_1)$, т. е. $g - g_1 = f \cdot f_1$, отсюда $g(\beta) - g_1(\beta) = 0$ и $g(\beta) = g_1(\beta)$.

2) Пусть $g+(f) \neq h+(f)$. Надо доказать, что $g(\beta) \neq h(\beta)$. Пусть, напротив, $g(\beta) = h(\beta)$. Тогда β – корень $g-h$ и f. По теореме Безу

$$(x-\beta) \mid (g(x)-h(x)) \text{ и } (x-\beta) \mid f(x).$$

Значит, $(g-h, f) \neq 1$, а так как f неразложим, то опять по лемме 1 из § 21 имеем: $f \mid (g-h)$, откуда $g+(f) = h+(f)$. Противоречие.

3) Следует из установленного равенства:

$$M = \{g(\beta) \mid g \in P[x]\}.$$

4) Пусть $A = g+(f)$, $B = h+(f)$, тогда левая часть

$$(A+B)\varphi = (g+h+(f))\varphi = g(\beta)+h(\beta),$$

правая часть

$$A\varphi + B\varphi = g(\beta)+h(\beta).$$

5) Устанавливается аналогично.

Теорема доказана.

Как установлено в доказательстве теоремы, наименьшее поле, содержащее P и некоторый элемент β, единственно. Оно называется *расширением поля P при помощи элемента β* и обозначается $P(\beta)$. Также из доказательства следует, что

$$P(\beta) = \{g(\beta) \mid g \in P[x]\}.$$

У п р а ж н е н и е. Возьмите в качестве поля P поле вещественных чисел $\mathbb{R}$, в качестве многочлена f многочлен x^2+1 и постройте наименьшее поле, в котором этот многочлен имеет корень.

Глава 4

КОЛЬЦА МНОГОЧЛЕНОВ (продолжение)

§ 25. Результант. Исключение неизвестного. Дискриминант

Даны два многочлена:

$$f(x) = a_0x^k + a_1x^{k-1} + \ldots + a_k, \quad a_i \in P;$$

$$g(x) = b_0x^l + b_1x^{l-1} + \ldots + b_l, \quad b_j \in P.$$

При этом мы не предполагаем, что $a_0 \neq 0$ и $b_0 \neq 0$. Можно сформулировать следующий вопрос: существуют ли у них общие корни?

Мы уже знаем, что многочлены f и g тогда и только тогда обладают общим корнем в некотором расширении поля P, если они не являются взаимно простыми. Таким образом, вопрос о существовании общих корней у данных многочленов может быть решен применением к ним алгоритма Евклида.

25.1. Результант двух многочленов от одного неизвестного. Укажем другой метод, позволяющий ответить на поставленный вопрос.

О п р е д е л е н и е. *Результантом* многочленов f и g называется

определитель

$$\mathrm{Res}(f,g)=\begin{vmatrix} a_0 & a_1 & \dots & a_k & 0 & 0 & \dots & 0 \\ 0 & a_0 & \dots & a_{k-1} & a_k & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_0 & a_1 & \dots & a_k \\ b_0 & b_1 & \dots & b_l & 0 & 0 & \dots & 0 \\ 0 & b_0 & \dots & b_{l-1} & b_l & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & b_0 & b_1 & \dots & b_l \end{vmatrix}$$

порядка $k+l$.

Из свойств определителей следует равенство

$$\mathrm{Res}(g,f)=(-1)^{kl}\mathrm{Res}(f,g).$$

Целью настоящего параграфа является доказательство следующего утверждения

Т е о р е м а 1. *Пусть*

$$f(x)=a_0x^k+a_1x^{k-1}+\ldots+a_k=a_0(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_k),\quad a_0\neq 0;$$

$$g(x)=b_0x^l+b_1x^{l-1}+\ldots+b_l=b_0(x-\beta_1)(x-\beta_2)\ldots(x-\beta_l),\quad b_0\neq 0$$

– два многочлена из $P[x]$. *Тогда*

$$\mathrm{Res}(f,g)=a_0^l\,b_0^k\prod_{\substack{1\le i\le k\\ 1\le j\le l}}(\alpha_i-\beta_j).$$

Для доказательства нам потребуется

Л е м м а 1 (определитель Вандермонда). *Для любых элементов* $z_1,z_2,\ldots,z_n$, $n\ge 2$ *из поля* P *справедливо равенство*

$$\begin{vmatrix} z_1^{n-1} & z_2^{n-1} & \dots & z_n^{n-1} \\ z_1^{n-2} & z_2^{n-2} & \dots & z_n^{n-2} \\ \dots & \dots & \dots & \dots \\ z_1 & z_2 & \dots & z_n \\ 1 & 1 & \dots & 1 \end{vmatrix}=\prod_{1\le i<j\le n}(z_i-z_j).$$

Д о к а з а т е л ь с т в о проведем индукцией по n.

При $n = 2$ имеем

$$\begin{vmatrix} z_1 & z_2 \\ 1 & 1 \end{vmatrix} = z_1 - z_2.$$

Предположим, что формула справедлива при $n-1$, и рассмотрим определитель порядка n. Вычитаем из первой строки вторую, умноженную на z_n, затем из второй – третью, умноженную на z_n, и т. д., и наконец, из $(n-1)$-й строки вычитаем n-ю, умноженную на z_n. Получим:

$$\begin{vmatrix} z_1^{n-1} - z_n \cdot z_1^{n-2} & \dots & z_{n-1}^{n-1} - z_n \cdot z_{n-1}^{n-2} & 0 \\ z_1^{n-2} - z_n \cdot z_1^{n-3} & \dots & z_{n-1}^{n-2} - z_n \cdot z_{n-1}^{n-3} & 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & & \\ z_1 - z_n & \dots & z_{n-1} - z_n & 0 \\ 1 & \dots & 1 & 1 \end{vmatrix}.$$

Разлагая этот определитель по последнему столбцу, приходим к определителю порядка $n-1$:

$$\begin{vmatrix} z_1^{n-1} - z_n \cdot z_1^{n-2} & \dots & z_{n-1}^{n-1} - z_n \cdot z_{n-1}^{n-2} \\ z_1^{n-2} - z_n \cdot z_1^{n-3} & \dots & z_{n-1}^{n-2} - z_n \cdot z_{n-1}^{n-3} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & \\ z_1 - z_n & \dots & z_{n-1} - z_n \end{vmatrix}.$$

Вынося из первого столбца $(z_1 - z_n)$, из второго $(z_2 - z_n)$ и т. д., и наконец, из $(n-1)$-го $(z_{n-1} - z_n)$, получим:

$$(z_1 - z_n)(z_2 - z_n)\dots(z_{n-1} - z_n) \begin{vmatrix} z_1^{n-2} & z_2^{n-2} & \dots & z_{n-1}^{n-2} \\ z_1^{n-3} & z_2^{n-3} & \dots & z_{n-1}^{n-3} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & & \\ z_1 & z_2 & \dots & z_{n-1} \\ 1 & 1 & \dots & 1 \end{vmatrix}.$$

Воспользовавшись предположением индукции, получим нужное равенство. Лемма доказана.

Следующая лемма легко устанавливается непосредственной проверкой.

Л е м м а 2 (формулы Виета). *Если*

$$x^n + a_1x^{n-1} + \ldots + a_n = (x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n),$$

то

$a_1 = -(\alpha_1 + \alpha_2 + \ldots + \alpha_n)$,

$a_2 = \alpha_1\cdot\alpha_2 + \alpha_1\cdot\alpha_3 + \ldots + \alpha_1\cdot\alpha_n + \alpha_2\cdot\alpha_3 + \ldots + \alpha_{n-1}\cdot\alpha_n$,

.....

$a_i = (-1)^i \sum\limits_{1\le k_1<k_2<\ldots<k_i\le n} \alpha_{k_1}\alpha_{k_2}\ldots\alpha_{k_i}$,

.....

$a_n = (-1)^n\alpha_1\alpha_2\ldots\alpha_n$.

Д о к а з а т е л ь с т в о теоремы 1. Рассмотрим матрицу

$$P = \begin{pmatrix} a_0 & a_1 & \ldots & a_k & 0 & 0 & \ldots & 0 \\ 0 & a_0 & \ldots & a_{k-1} & a_k & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & a_0 & a_1 & \ldots & a_k \\ b_0 & b_1 & \ldots & b_l & 0 & 0 & \ldots & 0 \\ 0 & b_0 & \ldots & b_{l-1} & b_l & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & b_0 & b_1 & \ldots & b_l \end{pmatrix}$$

и матрицу

$$Q = \left(\begin{array}{cccc|cccc} \beta_1^{k+l-1} & \beta_2^{k+l-1} & \ldots & \beta_l^{k+l-1} & \alpha_1^{k+l-1} & \alpha_2^{k+l-1} & \ldots & \alpha_k^{k+l-1} \\ \beta_1^{k+l-2} & \beta_2^{k+l-2} & \ldots & \beta_l^{k+l-2} & \alpha_1^{k+l-2} & \alpha_2^{k+l-2} & \ldots & \alpha_k^{k+l-2} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \beta_1 & \beta_2 & \ldots & \beta_l & \alpha_1 & \alpha_2 & \ldots & \alpha_k \\ 1 & 1 & \ldots & 1 & 1 & 1 & \ldots & 1 \end{array}\right).$$

Найдем их произведение $P\cdot Q$. Заметим, что в матрице $P\cdot Q$ на месте $(1,1)$ будет стоять элемент

$a_0\beta_1^{k+l-1} + a_1\beta_1^{k+l-2} + \ldots + a_k\beta_1^{l-1} = \beta_1^{l-1}(a_0\beta_1^k + a_1\beta_1^{k-1} + \ldots + a_k) = \beta_1^{l-1}f(\beta_1)$;
на месте $(2,1)$ – элемент

$a_0\beta_1^{k+l-2} + a_1\beta_1^{k+l-3} + \ldots + a_k\beta_1^{l-1} = \beta_1^{l-2}f(\beta_1)$
и т. д. Наконец, на месте $(l,1)$ будет стоять элемент

$a_0\beta_1^k + a_1\beta_1^{k-1} + \ldots + a_k = f(\beta_1)$.
На месте $(l+1,1)$ будет стоять элемент

$b_0\beta_1^{k+l-1} + b_1\beta_1^{k+l-2} + \ldots + b_l\beta_1^{k-1} = \beta_1^{k-1}g(\beta_1) = 0$
и т. д.

Рассмотрим элементы, которые получаются при умножении $l+1$-го столбца матрицы Q на матрицу P. На месте $(1,l+1)$ будет стоять элемент

$a_0\alpha_1^{k+l-1} + a_1\alpha_1^{k+l-2} + \ldots + a_k\alpha_1^{l-1} = \alpha_1^{l-1}(a_0\alpha_1^k + a_1\alpha_1^{k-1} + \ldots + a_k) = \alpha_1^{l-1}f(\alpha_1) = 0$;
на месте $(l+1,l+1)$ – элемент

$b_0\alpha_1^{k+l-1} + b_1\alpha_1^{k+l-2} + \ldots + b_l\alpha_1^{k-1} = \alpha_1^{k-1}g(\alpha_1)$;
на месте $(l+k,l+1)$ будет стоять элемент

$b_0\alpha_1^l + b_1\alpha_1^{l-1} + \ldots + b_l = g(\alpha_1)$.

Вычисляя аналогичным образом другие элементы, получим:

$$P\cdot Q = \left(\begin{array}{ccc|ccc} \beta_1^{l-1}f(\beta_1) & \ldots & \beta_l^{l-1}f(\beta_l) & 0 & \ldots & 0 \\ \beta_1^{l-2}f(\beta_1) & \ldots & \beta_l^{l-2}f(\beta_l) & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ f(\beta_1) & \ldots & f(\beta_l) & 0 & \ldots & 0 \\ \hline 0 & \ldots & 0 & \alpha_1^{k-1}g(\alpha_1) & \ldots & \alpha_k^{k-1}g(\alpha_k) \\ 0 & \ldots & 0 & \alpha_1^{k-2}g(\alpha_1) & \ldots & \alpha_k^{k-2}g(\alpha_k) \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & g(\alpha_1) & \ldots & g(\alpha_k) \end{array}\right).$$

Вычислим определители этих матриц:

$$\det P = \operatorname{Res}(f,g), \quad \det Q = \prod_{1\le i<j\le k}(\alpha_i-\alpha_j)\cdot\prod_{1\le r<s\le l}(\beta_r-\beta_s)\cdot\prod_{\substack{1\le r\le l\\ 1\le i\le k}}(\beta_r-\alpha_i),$$

$$\det(P\cdot Q)=\prod_{1\le r\le l} f(\beta_r)\cdot\prod_{1\le r<s\le l}(\beta_r-\beta_s)\cdot\prod_{1\le i\le k} g(\alpha_i)\cdot\prod_{1\le i<j\le k}(\alpha_i-\alpha_j)=$$

$$=\prod_{1\le i<j\le k}(\alpha_i-\alpha_j)\cdot\prod_{1\le r<s\le l}(\beta_r-\beta_s)\cdot a_0^l\prod_{\substack{1\le r\le l\\1\le i\le k}}(\beta_r-\alpha_i)\cdot b_0^k\prod_{\substack{1\le i\le k\\1\le r\le l}}(\alpha_i-\beta_r).$$

Учитывая, что

$$\det P\cdot\det Q=\det(P\cdot Q),$$

после сокращения, получим

$$\det P=a_0^l\,b_0^k\prod_{\substack{1\le i\le k\\1\le j\le l}}(\alpha_i-\beta_j).$$

Теорема доказана.

С л е д с т в и е. *Справедливо равенство*

$$\operatorname{Res}(f,g)=a_0^l\prod_{i=1}^{k} g(\alpha_i).$$

У п р а ж н е н и я. 1) Где нужно, чтобы a_0 и b_0 были отличны от нуля? 2) Почему выполненное сокращение законно?

25.2. Критерий совместности двух уравнений с одним неизвестным. Теперь мы готовы дать ответ на вопрос, сформулированный в начале параграфа.

Т е о р е м а 2. *Для многочленов $f,g\in P[x]$ равносильны следующие утверждения*:

1) $\operatorname{Res}(f,g)=0$;

2) *f и g имеют общий корень или $a_0=b_0=0$.*

Д о к а з а т е л ь с т в о разбивается на несколько случаев.

Случай 1. $a_0\neq 0$, $b_0\neq 0$. Сразу следует из теоремы 1.

Случай 2. $a_0=b_0=0$. Тогда целый столбик в определителе $\operatorname{Res}(f,g)$ равен нулю, а потому $\operatorname{Res}(f,g)=0$.

Случай 3. $a_0\neq 0$, $b_0=0$. Если $g(x)\equiv 0$, то $\operatorname{Res}(f,g)=0$, f и g имеют общий корень. Пусть $g(x)\not\equiv 0$, но $b_0=\ldots=b_{i-1}=0$, $b_i\neq 0$.

Тогда

$$\operatorname{Res}(f,g)=\begin{vmatrix} a_0 & a_1 & \dots & a_k & 0 & 0 & \dots & 0 \\ 0 & a_0 & \dots & a_{k-1} & a_k & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_0 & a_1 & \dots & a_k \\ 0 & \dots & 0 & b_i & \dots & b_l & \dots & 0 \\ 0 & \dots & 0 & 0 & b_i & \dots & b_l \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & b_i & \dots & b_l \end{vmatrix} = a_0^i \operatorname{Res}(f,\overline{g}),$$

где $\overline{g}(x)=b_ix^{l-i}+\ldots+b_l$. Тогда

$$\operatorname{Res}(f,g)=0 \Leftrightarrow \operatorname{Res}(f,\overline{g})=0.$$

По теореме 1 равенство $\operatorname{Res}(f,\overline{g})=0$ равносильно тому, что f и $\overline{g}$ имеют общий корень, а это равносильно тому, что f и g имеют общий корень.

Случай 4. $a_0=0$, $b_0\neq 0$. Разбирается аналогично случаю 3. Теорема доказана.

25.3. Исключение неизвестных. Рассмотрим следующую систему алгебраических уравнений:

$$\begin{cases} f_1(x_1,x_2,\ldots,x_n)=0, \\ f_2(x_1,x_2,\ldots,x_n)=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f_m(x_1,x_2,\ldots,x_n)=0, \end{cases}$$

где $f_i\in P[x_1,x_2,\ldots,x_n]$. Для ее решения можно попробовать применить метод исключения неизвестных, используемый для решения систем линейных уравнений, т. е. исключить вначале одну неизвестную, получив систему от меньшего числа неизвестных, затем вторую и т. д. Проделаем это для случая системы двух уравнений с двумя неизвестными.

Пусть мы имеем систему двух уравнений от двух неизвестных:

$$\begin{cases} f(x,y)=a_0(y)x^k+a_1(y)x^{k-1}+\ldots+a_k(y)=0, \\ g(x,y)=b_0(y)x^l+b_1(y)x^{l-1}+\ldots+b_l(y)=0, \end{cases} \tag{1}$$

где $a_i(y),b_j(y)\in P[y]$. Допустим, что мы умеем решать уравнения с одним неизвестным произвольной степени. Надо свести нашу систему

к уравнению от одной неизвестной. Рассмотрим

$$\mathrm{Res}_x(f,g) = F(y),$$

т. е., рассматривая $f(x,y)$ и $g(x,y)$ как многочлены от x и вычисляя результант, получим многочлен $F(y)$ от переменной y. Если (α,β) – решение системы (1), т. е. общий корень $f(x,y)$ и $g(x,y)$, то $F(\beta)=0$. Подставив это значение в систему (1), получим два многочлена от одной неизвестных. Мы уже умеем определять, имеют ли они общие корни. Заметим, что среди пар (α,β) могут существовать «лишние». Это такие β, которые обращают $a_0(y)$ или $b_0(y)$ в нуль.

25.4. Дискриминант. Для многочлена $f \in P[x]$ естественно сформулировать такой вопрос: когда f имеет кратные корни? Алгоритмический ответ ясен. Надо найти наибольший общий делитель многочлена f и его производной f'. Если он является многочленом не нулевой степени, то f имеет кратные корни. Как найти явный ответ?

О п р е д е л е н и е. *Дискриминантом* многочлена

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a^n = a_0(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n), \quad a_0 \neq 0$$

называется элемент

$$\mathrm{Dis}(f) = a_0^{2n-2} \prod_{1\le i<j\le n} (\alpha_i - \alpha_j)^2.$$

Из этого определения видим, что если $a_0 \neq 0$, то дискриминант равен нулю тогда и только тогда, когда многочлен имеет кратные корни. Но, чтобы вычислить дискриминант, используя это определение, мы должны знать корни. Следующая теорема позволяет вычислить дискриминант по коэффициентам многочлена.

Т е о р е м а 3. *Справедливо равенство*

$$\mathrm{Dis}(f) = (-1)^{C_n^2}\frac{1}{a_0}\mathrm{Res}(f,f'), \quad C_n^m = \frac{n!}{m!(n-m)!}.$$

Д о к а з а т е л ь с т в о. Пусть $\beta_1,\beta_2,\ldots,\beta_{n-1}$ – корни производной f', т. е.

$$f'(x) = na_0(x-\beta_1)(x-\beta_2)\ldots(x-\beta_{n-1}).$$

Тогда по теореме 1 имеем:

$$\mathrm{Res}(f,f') = a_0^{n-1}(na_0)^n \prod_{\substack{1\le i\le n\\ 1\le j\le n-1}} (\alpha_i - \beta_j) = a_0^{n-1}\prod_{i=1}^{n} f'(\alpha_i).$$

С другой стороны, для многочлена

$$f(x) = a_0(x-\alpha_1)(x-\alpha_2)\dots(x-\alpha_n)$$

найдем производную по формуле дифференцирования произведения:

$$f'(x) = a_0 \sum_{i=1}^{n} \prod_{\substack{j\neq i\\ 1\leq j\leq n}} (x-\alpha_j).$$

Подставив значение α_i, получим

$$f'(\alpha_i) = a_0 \prod_{\substack{j\neq i\\ 1\leq j\leq n}} (\alpha_i-\alpha_j).$$

Тогда

$$\operatorname{Res}(f,f') = a_0^{2n-1} \prod_{i=1}^{n} \prod_{\substack{j\neq i\\ 1\leq j\leq n}} (\alpha_i-\alpha_j) = a_0^{2n-2}(-1)^{C_n^2} a_0 \prod_{1\leq i<j\leq n} (\alpha_i-\alpha_j)^2 =$$

$$= (-1)^{C_n^2} a_0 \operatorname{Dis}(f).$$

Теорема доказана.

П р и м е р. Пусть $f(x) = ax^2 + bx + c$. Тогда $f' = 2ax + b$, и, вычисляя дискриминант, получим

$$\operatorname{Dis}(f) = (-1)^{C_2^2} \frac{1}{a} \begin{vmatrix} a & b & c \\ 2a & b & 0 \\ 0 & 2a & b \end{vmatrix} = b^2 - 4ac.$$

§ 26. Многочлены от нескольких переменных

26.1. Кольцо многочленов от нескольких переменных. *Многочленом над кольцом K от n переменных* $x_1, x_2, \dots, x_n$ называется выражение

$$f = f(x_1, x_2, \dots, x_n) = \sum_{k_1,k_2,\dots,k_n} a_{k_1k_2\dots k_n} x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}, \quad a_{k_1k_2\dots k_n} \in K,$$

в котором лишь конечное число коэффициентов $a_{k_1k_2\ldots k_n}$ отлично от нуля. *Степенью многочлена* f называется число

$$\max_{k_1,k_2,\ldots,k_n}(k_1+k_2+\ldots+k_n),$$

где максимум берется по всем наборам $k_1, k_2, \ldots, k_n$, для которых $a_{k_1k_2\ldots k_n} \neq 0$. На множестве многочленов естественным образом определяется сумма

$$\sum_{k_1,k_2,\ldots,k_n} a_{k_1k_2\ldots k_n}x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} + \sum_{k_1,k_2,\ldots,k_n} b_{k_1k_2\ldots k_n}x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} =$$
$$= \sum_{k_1,k_2,\ldots,k_n} (a_{k_1k_2\ldots k_n}+b_{k_1k_2\ldots k_n})x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}$$

и произведение

$$\sum_{k_1,k_2,\ldots,k_n} a_{k_1k_2\ldots k_n}x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \cdot \sum_{l_1,l_2,\ldots,l_n} b_{l_1l_2\ldots l_n}x_1^{l_1}x_2^{l_2}\ldots x_n^{l_n} =$$
$$= \sum_{m_1,m_2,\ldots,m_n} c_{m_1m_2\ldots m_n}x_1^{m_1}x_2^{m_2}\ldots x_n^{m_n},$$

где

$$c_{m_1m_2\ldots m_n} = \sum_{k_1+l_1=m_1,\ldots,k_n+l_n=m_n} a_{k_1k_2\ldots k_n}b_{l_1l_2\ldots l_n}.$$

Множество многочленов над кольцом K от переменных $x_1, x_2, \ldots, x_n$ будем обозначать $K[x_1, x_2, \ldots, x_n]$.

Т е о р е м а 1. *Если K – коммутативное кольцо без делителей нуля с единицей, то $K[x_1, x_2, \ldots, x_n]$ – коммутативное кольцо без делителей нуля с единицей.*

Д о к а з а т е л ь с т в о проведем индукцией по n. При $n = 1$ теорема доказана ранее.

Пусть $n > 1$. Рассмотрим отображение

$$K[x_1, x_2, \ldots, x_n] \longrightarrow K[x_1, x_2, \ldots, x_{n-1}][x_n],$$

которое сопоставляет многочлену $f(x_1, x_2, \ldots, x_n)$ его представление в виде многочлена от одной переменной x_n с коэффициентами из $K[x_1, x_2, \ldots, x_{n-1}]$. Это отображение является изоморфизмом, т. е.

$K[x_1, x_2, \ldots, x_n] \simeq K[x_1, x_2, \ldots, x_{n-1}][x_n]$. По предположению индукции $K[x_1, x_2, \ldots, x_{n-1}]$ – коммутативное кольцо без делителей нуля с единицей. По доказанной теореме о многочленах от одной переменной заключаем, что $K[x_1, x_2, \ldots, x_n]$ – коммутативное кольцо без делителей нуля с единицей. Теорема доказана.

П р и м е р. Пусть K – коммутативное кольцо с единицей, P – его подполе, $\alpha_1, \alpha_2, \ldots, \alpha_n$ – некоторые элементы из K. Пусть L – подкольцо, порожденное P и $\alpha_1, \alpha_2, \ldots, \alpha_n$. Тогда

$$L = \{f(\alpha_1, \alpha_2, \ldots, \alpha_n) \mid f \in P[x_1, x_2, \ldots, x_n]\}.$$

Действительно, включение $\supseteq$ очевидно. Чтобы проверить включение $\subseteq$, достаточно убедиться, что множество элементов вида

$$\sum_{k_1,k_2,\ldots,k_n} a_{k_1k_2\ldots k_n}\alpha_1^{k_1}\alpha_2^{k_2}\ldots\alpha_n^{k_n}, \quad a_{k_1k_2\ldots k_n} \in P$$

образует подкольцо, которое содержит P и $\alpha_1, \alpha_2, \ldots, \alpha_n$, а потому содержится в L.

В связи с этим примером дадим такое

О п р е д е л е н и е. Элементы $\alpha_1, \alpha_2, \ldots, \alpha_n$ из K называются *алгебраически независимыми над* P, если каждый элемент из L имеет единственную запись $f(\alpha_1, \alpha_2, \ldots, \alpha_n)$, где $f(x_1, x_2, \ldots, x_n) \in P[x]$.

Таким образом, элементы $\alpha_1, \alpha_2, \ldots, \alpha_n$ алгебраически независимы, если любой элемент

$$\sum_{k_1,k_2,\ldots,k_n} a_{k_1k_2\ldots k_n}\alpha_1^{k_1}\alpha_2^{k_2}\ldots\alpha_n^{k_n}$$

имеет единственную запись или, что равносильно, 0 записывается единственным образом. Действительно, если $\beta = f_1(\alpha_1, \alpha_2, \ldots, \alpha_n) = f_2(\alpha_1, \alpha_2, \ldots, \alpha_n)$, и $f_1 \neq f_2$, то $0 = (f_1 - f_2)(\alpha_1, \alpha_2, \ldots, \alpha_n)$, т. е. 0 записывается двумя способами. Иными словами, элементы $\alpha_1, \alpha_2, \ldots, \alpha_n$ алгебраически независимы над P, если n–ка $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ не является корнем никакого многочлена из $P[x_1, x_2, \ldots, x_n]$.

У п р а ж н е н и е. Если $\alpha_1, \alpha_2, \ldots, \alpha_n$ алгебраически независимы над P, то отображение

$$P[x_1, x_2, \ldots, x_n] \longrightarrow L,$$

определенное правилом $f(x_1, x_2, \ldots, x_n) \longmapsto f(\alpha_1, \alpha_2, \ldots, \alpha_n)$ является изоморфизмом. Таким образом, $L \simeq P[x_1, x_2, \ldots, x_n]$.

При $n = 1$ элемент α, алгебраически зависимый над P, называется *алгебраическим над* P; алгебраически независимый элемент называется *трансцендентным над* P.

26.2. Словарное упорядочивание многочленов. Если мы рассматриваем многочлен

$$f(x_1, x_2, x_3) = x_1^3 x_2 + x_1 x_2 x_3^2 + x_3^4,$$

то не известно какой коэффициент назвать старшим коэффициентом. Введем линейный порядок на одночленах.

О п р е д е л е н и е. Говорим, что одночлен $ax_1^{k_1}x_2^{k_2}\dots x_n^{k_n}$ *выше* одночлена $a'x_1^{k'_1}x_2^{k'_2}\dots x_n^{k'_n}$, если $k_1 = k'_1$, $k_2 = k'_2$, $\dots$, $k_{i-1} = k'_{i-1}$, но $k_i > k'_i$.

Высший одночлен в теории многочленов от нескольких переменных играет ту же роль, что и старший член в теории многочленов от одной переменной.

Л е м м а. *Для любых многочленов f и g из $K[x_1, x_2, \dots, x_n]$ высший член произведения $f \cdot g$ равен произведению высшего члена f и высшего члена g.*

Д о к а з а т е л ь с т в о. Пусть $ax_1^{k_1}x_2^{k_2}\dots x_n^{k_n}$ – высший член многочлена f, а $a'x_1^{k'_1}x_2^{k'_2}\dots x_n^{k'_n}$ – любой другой член f; $bx_1^{l_1}x_2^{l_2}\dots x_n^{l_n}$ – высший член многочлена g, а $b'x_1^{l'_1}x_2^{l'_2}\dots x_n^{l'_n}$ – любой другой член g. Пусть

$$k_1 = k'_1, k_2 = k'_2, \dots, k_{i-1} = k'_{i-1}, k_i > k'_i;$$
$$l_1 = l'_1, l_2 = l'_2, \dots, l_{j-1} = l'_{j-1}, l_j > l'_j.$$

Надо доказать, что произведение $abx_1^{k_1+l_1}x_2^{k_2+l_2}\dots x_n^{k_n+l_n}$ выше любого из членов

$$ab'x_1^{k_1+l'_1}x_2^{k_2+l'_2}\dots x_n^{k_n+l'_n}, \quad a'bx_1^{k'_1+l_1}x_2^{k'_2+l_2}\dots x_n^{k'_n+l_n},$$
$$a'b'x_1^{k'_1+l'_1}x_2^{k'_2+l'_2}\dots x_n^{k'_n+l'_n}.$$

Для первых двух членов это очевидно. Докажем для третьего.

Пусть $i \le j$, тогда

$$k_1+l_1 = k'_1+l'_1, k_2+l_2 = k'_2+l'_2, \dots, k_{i-1}+l_{i-1} = k'_{i-1}+l'_{i-1}, k_i+l_i > k'_i+l'_i.$$

Если $i \ge j$, то

$$k_1+l_1 = k'_1+l'_1, k_2+l_2 = k'_2+l'_2, \dots, k_{i-1}+l_{i-1} = k'_{j-1}+l'_{j-1}, k_j+l_j > k'_j+l'_j,$$

т. е. мы доказали, что член $abx_1^{k_1+l_1}x_2^{k_2+l_2}\ldots x_n^{k_n+l_n}$ выше члена $a'b'x_1^{k'_1+l'_1}\cdot x_2^{k'_2+l'_2}\ldots x_n^{k'_n+l'_n}$. Лемма доказана.

26.3. Симметрические многочлены. В кольце $K[x_1, x_2, \ldots, x_n]$ можно выделить подкольцо симметрических многочленов.

О п р е д е л е н и е. Многочлен называется *симметрическим*, если он не изменяется ни при какой перестановке переменных.

П р и м е р ы. 1) Нетрудно проверить, что многочлен

$$x_1^3x_2 + x_2^3x_3 + x_1x_2^3 + x_1x_3^3 + x_2x_3^3 + x_1^3x_3$$

является симметрическим.

2) Следующие многочлены называются *элементарными симметрическими многочленами*:

$$\begin{aligned}
&\sigma_1 = x_1 + x_2 + \ldots + x_n,\\
&\sigma_2 = x_1x_2 + x_1x_3 + \ldots x_1x_n + x_2x_3 + \ldots + x_{n-1}x_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\sigma_k = \sum_{1\le i_1<i_2<\ldots<i_k\le n} x_{i_1}x_{i_2}\ldots x_{i_k},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\sigma_n = x_1x_2\ldots x_n.
\end{aligned}$$

Они возникают в формуле Виета.

Элементарные симметрические многочлены порождают все симметрические многочлены. Более точно, справедлива

Т е о р е м а 2. *Пусть K – коммутативное кольцо с единицей, $K_s[x_1, x_2, \ldots, x_n]$ – множество всех симметрических многочленов над K от $x_1, x_2, \ldots, x_n$. Тогда*

1) *$K_s[x_1, x_2, \ldots, x_n]$ – подкольцо кольца $K[x_1, x_2, \ldots, x_n]$;*

2) *$K_s[x_1, x_2, \ldots, x_n]$ порождается кольцом K и многочленами σ_1, σ_2, …, σ_n.*

Д о к а з а т е л ь с т в о. 1) Чтобы проверить, что $K_s[x_1, x_2, \ldots, x_n]$ является подкольцом, достаточно проверить, что если f и g – симметрические многочлены, то многочлены $f+g$, $-f$ и $f\cdot g$ также являются симметрическими. Это непосредственно следует из определения.

2) Проверим, что $K_s[x_1, x_2, \ldots, x_n]$ порождается K и σ_1, σ_2, …, σ_n, т. е. каждый симметрический многочлен можно представить как

многочлен над K от σ_1, σ_2, ..., σ_n. Пусть $f \in K_s[x_1, x_2, \ldots, x_n]$ и его высший член равен $ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}$. Так как f симметрический, то $k_1 \geq k_2 \geq \ldots \geq k_n$. Действительно, если бы оказалось, что $k_i < k_{i+1}$, то, переставляя x_i и x_{i+1}, получили бы член, который содержится в f и выше $ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}$.

Рассмотрим

$$f_1 = a\sigma_1^{k_1-k_2}\sigma_2^{k_2-k_3}\ldots\sigma_{n-1}^{k_{n-1}-k_n}\sigma_n^{k_n}.$$

По лемме о высшем члене произведения видим, что высший член f_1 равен

$$ax_1^{k_1-k_2}(x_1x_2)^{k_2-k_3}\ldots(x_1x_2\ldots x_n)^{k_n} = ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n},$$

т. е. совпадает с высшим членом многочлена f. Рассмотрим многочлен $f - f_1$. Его высший член ниже высшего члена многочлена f. Пусть $bx_1^{l_1}x_2^{l_2}\ldots x_n^{l_n}$ – высший член многочлена $f - f_1$. Берем многочлен

$$f_2 = b\sigma_1^{l_1-l_2}\sigma_2^{l_2-l_3}\ldots\sigma_{n-1}^{l_{n-1}-l_n}\sigma_n^{l_n}.$$

Так же как и выше, замечаем, что высший член f_2 равен высшему члену многочлена $f - f_1$. Поэтому высший член $f - f_1 - f_2$ ниже высшего члена $f - f_1$. Продолжая эту процедуру, видим, что процесс оборвется на нулевом многочлене, и в результате мы получим искомое разложение

$$f = f_1 + f_2 + \ldots + f_s.$$

Теорема доказана.

26.4. Симметрические многочлены от корней многочлена от одной переменной. Рассмотрим многочлен

$$f(x) = x^2 + 1$$

над полем рациональных чисел $\mathbb{Q}$. Как мы знаем, он имеет два комплексных корня: $\alpha_1 = -\mathrm{i}$, $\alpha_2 = \mathrm{i}$. Если мы подставим эти корни в многочлен от двух переменных

$$h(x_1, x_2) = x_1^2 + x_1x_2 + x_2$$

над $\mathbb{Q}$, то получим элемент $h(\alpha_1, \alpha_2) = \mathrm{i}$, который не лежит в $\mathbb{Q}$. Если же подставить корни в симметрический многочлен $g(x_1, x_2) =$

$x_1^2 + x_1x_2 + x_2^2$, то получим элемент $g(\alpha_1, \alpha_2) = -1$ из $\mathbb{Q}$. Оказывается, что справедливо

П р е д л о ж е н и е. *Пусть $f(x)$ многочлен из $P[x]$ степени n, имеющий корни $\alpha_1, \alpha_2, \ldots, \alpha_n$ в некотором расширении L поля P, а $g = g(x_1, x_2, \ldots, x_n)$ – симметрический многочлен над P. Тогда его значение $g(\alpha_1, \alpha_2, \ldots, \alpha_n)$ от корней $\alpha_1, \alpha_2, \ldots, \alpha_n$ лежит в поле P.*

Пусть многочлен

$$f(x) = a_0x_n + a_1x^{n-1} + \ldots + a_n, \quad a_i \in P$$

имеет корни $\alpha_1, \alpha_2, \ldots, \alpha_n$ в некотором поле L, которое является расширением поля P, т. е.

$$f(x) = a_0(x - \alpha_1)(x - \alpha_2)\ldots(x - \alpha_n).$$

Если g – симметрический, то по теореме 1 его можно записать как многочлен от $\sigma_1, \sigma_2, \ldots, \sigma_n$ над P, т. е. $g = F(\sigma_1, \sigma_2, \ldots, \sigma_n)$. Подставив корни, по теореме Виета получим

$$g(\alpha_1, \alpha_2, \ldots, \alpha_n) = F(\sigma_1(\alpha_1, \alpha_2, \ldots, \alpha_n), \sigma_2(\alpha_1, \alpha_2, \ldots, \alpha_n), \ldots,$$

$$\sigma_n(\alpha_1, \alpha_2, \ldots, \alpha_n)) = F\left(-\frac{a_1}{a_0}, \frac{a_2}{a_0}, \ldots, (-1)^n\frac{a_n}{a_0}\right) \in P.$$

Предложение доказано.

26.5. Алгебраическая независимость элементарных симметрических многочленов. Как мы знаем, кольцо симметрических многочленов $P_s[x_1, x_2, \ldots, x_n]$ порождается полем P и элементарными симметрическими многочленами σ_1, σ_2, …, σ_n. Справедлива

Т е о р е м а 3. *Элементарные симметрические многочлены σ_1, σ_2, …, σ_n алгебраически независимы над основным полем P.*

Д о к а з а т е л ь с т в о. Предположим, что для некоторого многочлена $\varphi \in P[y_1, y_2, \ldots, y_n]$ имеем $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n) = 0$, т. е. 0 представим двумя разными способами. Надо доказать, что $\varphi \equiv 0$.

Пусть $\varphi = \sum\limits_i \varphi_i$, где $\varphi_i = a_i y_1^{k_{i_1}} y_2^{k_{i_2}} \ldots y_n^{k_{i_n}}$ и все наборы $(k_{i_1}, k_{i_2}, \ldots, k_{i_n})$ различны при различных i. Отсюда

$$\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n) = \sum_i \varphi_i(\sigma_1, \sigma_2, \ldots, \sigma_n).$$

Пусть

$$\varphi_i(\sigma_1, \sigma_2, \ldots, \sigma_n) = f_i(x_1, x_2, \ldots, x_n) \in P[x_1, x_2, \ldots, x_n].$$

По лемме о высшем члене произведения имеем

$$(\text{высший член } f_i) =$$

$$= a_i \cdot (\text{высший член } \sigma_1)^{k_{i_1}} \cdot \ldots \cdot (\text{высший член } \sigma_n)^{k_{i_n}} =$$

$$= a_i x_1^{k_{i_1}+k_{i_2}+\ldots+k_{i_n}} x_2^{k_{i_2}+k_{i_3}+\ldots+k_{i_n}} \ldots x_n^{k_{i_n}}.$$

Если высший член многочлена f_i известен, то φ_i можно однозначно восстановить. Тогда по формуле

$$\varphi = \sum_i \varphi_i$$

можем восстановить φ. Если бы высшие члены у f_i и f_j совпадали, то $\varphi_i = \varphi_j$, а этого быть не может. Следовательно, все высшие члены у f_i различны. Пусть $bx_1^{l_1} x_2^{l_2} \ldots x_n^{l_n}$ – высший из всех высших членов f_i. Тогда он ни с чем не сократится, а потому $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n) \neq 0$, противоречие. Теорема доказана.

Из доказанной теоремы следует, что кольцо симметрических многочленов изоморфно кольцу многочленов:

$$P[\sigma_1, \sigma_2, \ldots, \sigma_n] = P_s[x_1, x_2, \ldots, x_n] \simeq P[x_1, x_2, \ldots, x_n].$$

§ 27. Комплексные многочлены от одной переменной

Особый интерес представляют поля, которые не надо расширять – алгебраически замкнутые.

О п р е д е л е н и е. Поле P называется *алгебраически замкнутым,* если любой многочлен ненулевой степени из $P[x]$ имеет в P хотя бы один корень.

У п р а ж н е н и е. Поле алгебраически замкнуто тогда и только тогда, когда любой многочлен над этим полем разлагается над ним на линейные множители. (У к а з а н и е: использовать теорему Безу.)

27.1. Алгебраическая замкнутость поля комплексных чисел. В этом пункте мы докажем утверждение о том, что поле комплексных чисел алгебраически замкнуто. Это утверждение иногда

называют основной теоремой алгебры. Более правильно было бы назвать ее основной теоремой алгебры многочленов.

Предварительно докажем следующее утверждение, которое представляет и самостоятельный интерес.

Л е м м а 1 (о модуле старшего члена). *Для многочлена*

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n, \quad a_0 \neq 0, \quad a_i \in \mathbb{C}$$

найдется такое вещественное неотрицательное число N, что для всех $x \in \mathbb{C}$, таких, что $|x| \geq N$, справедливо неравенство

$$|a_0x^n| > |a_1x^{n-1} + \ldots + a_n|.$$

Д о к а з а т е л ь с т в о. Имеем

$$|a_1x^{n-1} + \ldots + a_n| \leq |a_1| \cdot |x|^{n-1} + |a_2| \cdot |x|^{n-2} + \ldots + |a_n|.$$

Положим

$$A = \max\{|a_1|, |a_2|, \ldots, |a_n|\}.$$

Тогда при $|x| \neq 1$

$$|a_1| \cdot |x|^{n-1} + |a_2| \cdot |x|^{n-2} + \ldots + |a_n| \leq A(|x|^{n-1} + |x|^{n-2} + \ldots + 1) =$$

$$= A\frac{|x|^n - 1}{|x| - 1} < A\frac{|x|^n}{|x| - 1}.$$

Найдем те значения x, для которых выполняется неравенство

$$A\frac{|x|^n}{|x| - 1} \leq |a_0x^n|.$$

Нетрудно проверить, что оно выполняется тогда и только тогда, когда

$$|x| > 1 + \frac{A}{|a_0|} \text{ и } |x| < 1.$$

Следовательно, полагая

$$N = 1 + \frac{A}{|a_0|},$$

получим нужное утверждение. Лемма доказана.

Из этой леммы, в частности, следует, что все комплексные корни многочлена $f(x)$ лежат внутри круга с центром в начале координат и радиуса N.

Т е о р е м а. *Поле комплексных чисел алгебраически замкнуто.*

Д о к а з а т е л ь с т в о. Пусть

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n, \quad a_0 \neq 0, \quad a_i \in \mathbb{C}.$$

Надо доказать, что f имеет комплексный корень.

Случай 1. Предположим вначале, что $f \in \mathbb{R}[x]$, т. е. все коэффициенты a_i лежат в $\mathbb{R}$. Представим степень многочлена n в таком виде: $n = 2^k q$, где q – нечетное число. Воспользуемся индукцией по k. При $k = 0$ число n нечетно. Ввиду леммы о модуле старшего члена найдется такое неотрицательное вещественное число N, что $f(-N)$ и $f(N)$ имеют разные знаки. Справедлива следующая лемма, доказываемая в курсе математического анализа.

Л е м м а 2. *Всякая вещественная функция, непрерывная на отрезке $[a, b]$ и принимающая на его концах значения разных знаков, обращается в 0 в некоторой точке $c \in [a, b]$.*

Следовательно, найдется вещественное число c из отрезка $[-N, N]$, которое является корнем многочлена f. Основание индукции установлено.

Пусть теперь $k > 0$. Предположим, что утверждение справедливо для $k - 1$, и докажем его для k. По теореме о существовании корня найдется поле P, являющееся расширением поля $\mathbb{R}$, над которым многочлен f разлагается на линейные множители:

$$f(x) = a_0(x - \alpha_1)(x - \alpha_2)\ldots(x - \alpha_n), \quad \alpha_i \in P.$$

Пусть $c \in \mathbb{R}$. Определим числа $\beta_{ij} = \alpha_i\alpha_j + (\alpha_i + \alpha_j)c$, $1 \leq i < j \leq n$ и рассмотрим многочлен

$$g(x) = \prod_{1 \leq i < j \leq n} (x - \beta_{ij}).$$

Очевидно, что его степень

$$\deg g = \frac{n(n-1)}{2} = 2^{k-1}q',$$

где $q' = q(n - 1)$ – нечетно, так как $k \geq 1$. Надо убедиться, что коэффициенты многочлена g лежат в $\mathbb{R}$. Заметим, что всякий коэффициент многочлена g является элементарным симметрическим многочленом от β_{ij} над $\mathbb{R}$, т. е. если переставить пару индексов (i, j) и (r, s), то многочлен $g(x)$ не изменится. Отсюда нетрудно вывести, что

если переставить местами любые корни α_i и α_j, то многочлен также не изменится. Следовательно, всякий коэффициент g является симметрическим многочленом над $\mathbb{R}$ от $\alpha_1, \alpha_2, \ldots, \alpha_n$. По доказанному выше предложению о симметрическом многочлене от корней уравнения заключаем, что всякий коэффициент многочлена g лежит в $\mathbb{R}$, а потому и $g \in \mathbb{R}[x]$. По предположению индукции g имеет хотя бы один корень в $\mathbb{C}$, т. е. хотя бы одно $\beta_{ij} \in \mathbb{C}$.

Вспоминаем, что β_{ij} зависит от вещественного параметра c. Выбирая другое значение c, получим, что какой-то другой β_{rs} является комплексным корнем, но в качестве c мы можем брать любое вещественное число, а число перестановок корней β_{ij} конечно. Следовательно, найдется пара вещественных чисел c_1 и c_2, $c_1 \neq c_2$ таких, что

$$\alpha_i\alpha_j + (\alpha_i + \alpha_j)c_1 = b_1 \in \mathbb{C},$$
$$\alpha_i\alpha_j + (\alpha_i + \alpha_j)c_2 = b_2 \in \mathbb{C}.$$

Тогда

$$\alpha_i + \alpha_j = \frac{b_1 - b_2}{c_1 - c_2} \in \mathbb{C}, \quad \alpha_i\alpha_j \in \mathbb{C}.$$

Пара α_i, α_j является корнями уравнения

$$z^2 - (\alpha_i + \alpha_j)z + \alpha_i\alpha_j = 0$$

с комплексными коэффициентами. По формуле нахождения корней квадратного многочлена видим, что α_i и α_j являются комплексными числами. Следовательно, мы нашли два комплексных корня многочлена f.

Случай 2: $f \in \mathbb{C}[x]$, т. е. все a_i лежат в $\mathbb{C}$. Рассмотрим многочлен

$$\overline{f}(x) = \overline{a_0}x^n + \overline{a_1}x^{n-1} + \ldots + \overline{a_n}$$

и определим

$$F(x) = f(x) \cdot \overline{f}(x) = \sum_{k=0}^{2n} c_k x^k, \text{ где } c_k = \sum_{r+s=k} a_r\overline{a_s}.$$

Ясно, что

$$\overline{c_k} = \overline{\sum_{r+s=k} a_r\overline{a_s}} = \sum_{r+s=k} \overline{a_r\overline{a_s}} = \sum_{r+s=k} \overline{a_r}a_s = c_k,$$

т. е. $c_k \in \mathbb{R}$, а потому $F(x) \in \mathbb{R}[x]$. По установленному случаю 1 у него существует корень, т. е. найдется некоторое $\gamma \in \mathbb{C}$ такое, что

$$F(\gamma) = f(\gamma) \cdot \overline{f}(\gamma) = 0.$$

Если $f(\gamma) = 0$, то γ – корень f. Если $\overline{f}(\gamma) = 0$, то

$$\overline{a_0}\gamma^n + \overline{a_1}\gamma^{n-1} + \ldots + \overline{a_n} = 0,$$

и, переходя к комплексно-сопряженному, получим

$$a_0\overline{\gamma}^n + a_1\overline{\gamma}^{n-1} + \ldots + a_n = \overline{0},$$

т. е. $\overline{\gamma}$ – корень многочлена f и $\gamma \in \mathbb{C}$. Теорема доказана.

27.2. Некоторые приложения. Пусть $z = a + i\,b$, $a, b \in \mathbb{R}$ – некоторое комплексное число. Как мы знаем, в тригонометрической форме его можно записать так: $z = r(\cos\varphi + i\,\sin\varphi)$. Обозначим $\cos\varphi + i\,\sin\varphi = \mathrm{e}^{i\alpha}$, $\alpha \in \mathbb{R}$ и рассмотрим N-угольник. Тогда его вершинами будут являться точки $a_k = \mathrm{e}^{\frac{2k\pi i}{N}}$, $k = 0, 1, \ldots, N-1$ на комплексной плоскости. Справедлива

Т е о р е м а (о значении многочлена в центре правильного многоугольника). *Для любого многочлена $f \in \mathbb{C}[x]$ степени n и всякого правильного N-угольника, $N > n$ с центром в точке a и вершинами a_k, $k = 0, 1, \ldots, N-1$, справедливо равенство*

$$f(a) = \frac{1}{N}\sum_{k=0}^{N-1} f(a_k).$$

Д о к а з а т е л ь с т в о. Можно считать, что $a_k = a + be^{\frac{2\pi ki}{N}}$ при подходящем $b \in \mathbb{C}$. Рассмотрим многочлен $g(x) = f(a + bx)$. Очевидно, $g(0) = f(a)$, и надо доказать, что

$$g(0) = \frac{1}{N}\sum_{k=0}^{N-1} g(e^{\frac{2\pi ki}{N}}).$$

Пусть $g(x) = \sum_{s=0}^{n} a_s x^s$. Тогда при подстановке получим:

$$\frac{1}{N}\sum_{k=0}^{N-1} g(e^{\frac{2\pi ki}{N}}) = \frac{1}{N}\sum_{k=0}^{N-1}\sum_{s=0}^{n} a_s\left(e^{\frac{2\pi ki}{N}}\right)^s = \frac{1}{N}\sum_{s=0}^{n} a_s\left(\sum_{k=0}^{N-1} e^{\frac{2\pi ksi}{N}}\right) =$$

$$= \frac{1}{N}a_0 N = a_0 = g(0).$$

Так как

$$\sum_{k=0}^{N-1} e^{\frac{2\pi ksi}{N}} = 1 + e^{\frac{2\pi si}{N}} + e^{\frac{2\pi s2i}{N}} + \dots$$

– геометрическая прогрессия, то

$$\sum_{k=0}^{N-1} e^{\frac{2\pi ksi}{N}} = \begin{cases} N & \text{при } s = 0, \\ \frac{q^N-1}{q-1} = 0 & \text{при } s \neq 0, \end{cases}$$

где $q = e^{\frac{2\pi si}{N}}$. Теорема доказана.

У п р а ж н е н и е. Где использовалось, что $N > n$?

Всякий многочлен $f(x) \in \mathbb{C}[x]$ мы можем рассматривать как функцию $f : \mathbb{C} \longrightarrow \mathbb{C}$. При этом $|f(x)|$ является функцией, определенной на $\mathbb{C}$ со значениями в $\mathbb{R}$, т. е. $|f| : \mathbb{C} \longrightarrow \mathbb{R}$. Справедлива

Т е о р е м а (отсутствие локальных максимумов модуля). *Ни для какого многочлена $f(x) \in \mathbb{C}[x] \setminus \mathbb{C}$ не существует локальных максимумов функции $|f|$.*

Д о к а з а т е л ь с т в о. Предположим, что точка a является точкой локального максимума функции $|f|$, т. е. $|f(x)| < |f(a)|$ при условии, что $x \neq a$ удовлетворяет неравенству $|x - a| \leq \delta$ для некоторого $\delta > 0$. Пусть $N > n$, где n – степень многочлена f. Обозначим a_k, $k = 0, 1, \dots, N-1$ – вершины правильного N-угольника с центром в точке a. Тогда по теореме 1 справедливо неравенство

$$|f(a)| \leq \frac{1}{N} \sum_{k=0}^{N-1} |f(a_k)|,$$

из которого следует, что $|f(a)| < |f(a)|$. Противоречие.

Отметим, что теоремы 1 и 2 справедливы не только для многочленов, но и для произвольных аналитических функций

$$f(x) = \sum_{k=0}^{\infty} a_k x^k, \quad a_k \in \mathbb{C}.$$

§ 28. Поле частных

Мы знаем, что кольцо целых чисел вкладывается в поле рациональных чисел, кольцо многочленов над полем вкладывается в поле рациональных дробей. Возникает естественный вопрос: всякое ли

кольцо вкладывается в поле? Ответ, очевидно, отрицательный, так как некоммутативное кольцо, или кольцо, обладающее делителями нуля, не может быть вложено в поле. Оказывается, что некоммутативность и наличие делителей нуля являются единственными препятствиями к требуемому вложению.

28.1. Вложение целостного кольца в поле. Напомним, что кольцо называется *целостным,* если оно коммутативно и не имеет делителей нуля. В частности, любое поле является целостным кольцом. Действительно, если $ab = 0$ и $a \neq 0$, то существует элемент a^{-1}, а потому $a^{-1}ab = b = 0$, т. е. $b = 0$. Кроме того, легко заметить, что любое подкольцо поля является целостным кольцом.

Т е о р е м а 1. 1) *Любое целостное кольцо K изоморфно вкладывается в некоторое поле L.* 2) *Подполе, порожденное образом кольца K в поле L, определено однозначно с точностью до изоморфизма.*

Д о к а з а т е л ь с т в о. Вначале докажем существование такого поля. Для этого рассмотрим множество пар

$$\{(a, b) \mid a, b \in K, b \neq 0\}$$

и определим на нем отношение $\sim$, полагая $(a, b) \sim (c, d)$, если $ad = bc$. Проверим, что это отношение является отношением эквивалентности. Для этого надо проверить:

1) $(a, b) \sim (a, b)$;

2) если $(a, b) \sim (c, d)$, то $(c, d) \sim (a, b)$;

3) если $(a, b) \sim (c, d)$ и $(c, d) \sim (e, f)$, то $(a, b) \sim (e, f)$.

Первые два свойства очевидны. Установим свойство 3. Пусть

$$(a, b) \sim (c, d) \text{ и } (c, d) \sim (e, f).$$

Это равносильно тому, что $ad = bc$ и $cf = de$. Надо доказать, что $af = be$. Умножая первое равенство на f, а второе на b, получим

$$adf = bcf, \quad bcf = bde.$$

Из этих равенств следует, что $adf = bde$, и, сокращая на d, получим $af = be$, но это и означает, что $(a, b) \sim (e, f)$. Следовательно, отношение $\sim$ действительно является отношением эквивалентности. Относительно этого отношения множество

$$\{(a, b) \mid a, b \in K, b \neq 0\}$$

распадается на классы эквивалентности. Обозначим $\overline{(a,b)}$ множество пар, эквивалентных (a,b), т. е.

$$\overline{(a,b)} = \{(c,d) \mid (c,d) \sim (a,b)\},$$

и рассмотрим фактор-множество

$$L = \{(a,b) \mid a,b \in K, b \neq 0\}/_{\sim}.$$

Определим на L операции сложения и умножения, полагая

$$\overline{(a,b)} + \overline{(c,d)} = \overline{(ad+bc,bd)},$$

$$\overline{(a,b)} \cdot \overline{(c,d)} = \overline{(ac,bd)}.$$

Проверим, что эти операции определены корректно, т. е. не зависят от выбора представителей. Пусть

$$\overline{(a,b)} = \overline{(a',b')}, \quad \overline{(c,d)} = \overline{(c',d')},$$

т. е.

$$(a,b) \sim (a',b'), \quad (c,d) \sim (c',d').$$

Надо проверить, что справедливы равенства

$$\overline{(ad+bc,bd)} = \overline{(a'd'+b'c',b'd')}, \quad \overline{(ac,bd)} = \overline{(a'c',b'd')},$$

т. е.

$$(ad+bc,bd) \sim (a'd'+b'c',b'd'), \quad (ac,bd) \sim (a'c',b'd').$$

Для проверки первого равенства умножим равенство $ab' = ba'$ на dd', а равенство $cd' = dc'$ на bb', складывая эти равенства, получим

$$(ad+bc)b'd' = bd(a'd'+b'c').$$

Следовательно, операция сложения определена корректно.

Проверим справедливость второго равенства. Имеем:

$$ab' = ba', \quad cd' = dc'.$$

Перемножая эти равенства почленно, получим $acb'd' = bda'c'$. Следовательно, операция умножения определена корректно.

Докажем, что алгебраическая система $\langle L; +, \cdot \rangle$ является полем. Для этого надо проверить все аксиомы поля.

С1) Ассоциативность сложения: $(A+B)+C = A+(B+C)$.

Пусть $A = \overline{(a,b)}$, $B = \overline{(c,d)}$, $C = \overline{(e,f)}$. Тогда

$$A + B = \overline{(ad+bc, bd)}, \quad (A+B)+C = \overline{((ad+bc)f + bde, (bd)f)}.$$

С другой стороны,

$$B + C = \overline{(cf+de, df)}, \quad A + (B+C) = \overline{(adf + b(cf+de), b(df))} =$$

$$= \overline{((ad+bc)f + bde, (bd)f)},$$

т. е. ассоциативность сложения выполняется.

С2) Коммутативность сложения: $A + B = B + A$. Очевидно.

С3) В качестве нулевого элемента возьмем класс $\overline{(0,u)}$, где $u \neq 0$.

С4) Противоположным к классу $\overline{(a,b)}$ является класс $-\overline{(a,b)} = \overline{(-a,b)}$.

Аксиомы У1 и У2 выполняются в силу соответствующих аксиом кольца K.

У3) Единичным классом является класс $\overline{(u,u)}$, $u \neq 0$.

У4) Для любого класса $\overline{(a,b)} \neq \overline{(0,u)}$ обратным будет $\overline{(a,b)}^{-1} = \overline{(b,a)}$.

Таким образом, L действительно является полем.

Рассмотрим отображение

$$\varphi : K \longrightarrow L,$$

сопоставляющее каждому элементу из K элемент из L по правилу $a\varphi = \overline{(au,u)}$, $u \in K \setminus \{0\}$. Покажем, что это изоморфизм K в L. Действительно, если $a \neq b$, но $\overline{(au,u)} = \overline{(bu,u)}$, т. е. $(au,u) \sim (bu,u)$, то $au^2 = bu^2$. Учитывая, что K целостное, вопреки предположению, получим $a = b$. Следовательно, отображение φ унивалентно.

Проверим, что φ сохраняет операцию сложения, т. е.

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

Левая часть этого равенства:

$$(a+b)\varphi = \overline{((a+b)u, u)};$$

правая часть:

$$a\varphi + b\varphi = \overline{(au,u)} + \overline{(bu,u)} = \overline{(au^2+bu^2, u^2)} = \overline{((a+b)u, u)}.$$

Следовательно, операция сложения сохраняется.

Аналогично проверяется, что сохраняется и операция умножения.

Таким образом, пункт 1 теоремы установлен.

2) Предположим, что существуют два изоморфных вложения φ' и φ'' кольца K в поля L' и L'' соответственно. Обозначим M – подполе поля L', порожденное $\varphi'(K)$, а M' – подполе поля L'', порожденное $\varphi''(K)$. Докажем, что каждое из этих подполей изоморфно полю L. Заметим вначале, что

$$M = \{ab^{-1} \mid a, b \in K, b \neq 0\},$$

где мы отождествляем элементы из K с их образами в $\varphi'(K)$. Действительно, если $a, b \in K$, то в M лежат обратные элементы и произведения. Следовательно, включение $\supseteq$ выполняется. Чтобы установить обратное включение покажем, что элементы ab^{-1} образуют подполе. Для этого надо проверить замкнутость относительно сложения, умножения, взятия противоположного и взятия обратного. Так как

$$ab^{-1} + cd^{-1} = (ad + bc)(bd)^{-1}$$

и

$$ab^{-1} \cdot cd^{-1} = (ac)(bd)^{-1},$$

то наше множество замкнуто относительно сложения и умножения. Противоположным к элементу ab^{-1} является элемент $-ab^{-1} = (-a)b^{-1}$. Если $ab^{-1} \neq 0$, то $a \neq 0$ и $(ab^{-1})^{-1} = ba^{-1}$. Заметим, что M содержит K. Действительно, если $a \in K$, то его можно представить в виде $a = (ab) \cdot b^{-1}$. Таким образом, множество элементов ab^{-1} образует подполе, которое содержит K, а так как M наименьшее с этим свойством, то отсюда следует включение $\subseteq$.

Рассмотрим теперь отображение

$$\omega : M \longrightarrow L,$$

действующее по правилу

$$ab^{-1} \longmapsto \overline{(a, b)}.$$

Покажем, что это отображение является изоморфизмом. То, что ω однозначно, следует из определения. Проверим унивалентность: предположим, что $(ab^{-1})\omega = (cd^{-1})\omega$, т. е. $\overline{(a, b)} = \overline{(c, d)}$, но последнее означает, что $(a, b) \sim (c, d)$, т. е. $ad = bc$, а потому $ab^{-1} = cd^{-1}$.

То, что ω сохраняет операции, следует из определения операций в L. Теорема доказана.

О п р е д е л е н и е. Наименьшее поле, содержащее данное целостное кольцо в качестве подкольца, называется *полем частных* этого кольца.

Поле частных полностью определяется кольцом K. Будем обозначать его поле частных символом $\ddot{K}$. Очевидно, что для кольца целых чисел $\mathbb{Z}$ его поле частных изоморфно полю рациональных чисел $\mathbb{Q}$.

При работе с полем частных пару (a, b), $b \neq 0$ записывают в виде дроби $\frac{a}{b}$. Очевидно, что класс $\overline{(a, b)}$ состоит из дробей $\frac{au}{bu}$, $u \neq 0$. Нетрудно проверить, что операции сложения и умножения классов в поле частных согласуются с обычными операциями сложения и умножения дробей.

28.2. Поле рациональных дробей. Рассмотрим кольцо многочленов $P[x_1, x_2, \ldots, x_n]$ над полем P. Как мы знаем, это кольцо является целостным. По доказанной теореме оно вкладывается в поле. Полем частных кольца многочленов $P[x_1, x_2, \ldots, x_n]$ является поле рациональных дробей:

$$P(x_1, x_2, \ldots, x_n) = \left\{ \frac{f}{g} \;\middle|\; f, g \in P[x_1, x_2, \ldots, x_n], g \neq 0 \right\}.$$

При этом $\frac{f}{g} = \frac{f'}{g'}$, если $fg' = gf'$. Сложение и умножение рациональных дробей определяются следующими правилами:

$$\frac{f}{g} + \frac{f'}{g'} = \frac{fg' + gf'}{gg'}, \quad \frac{f}{g} \cdot \frac{f'}{g'} = \frac{ff'}{gg'}.$$

Каждый многочлен можно рассматривать как функцию $P^n \longrightarrow P$. Для рациональной дроби может оказаться, что ее знаменатель обращается в нуль при некоторых значениях переменных.

28.3. База векторного пространства $P(x)$ над полем P. Множество $P[x_1, x_2, \ldots, x_n]$ образует векторное пространство над полем P. Его базу образуют одночлены вида

$$x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}, \quad k_i \in \mathbb{N} \cup \{0\}.$$

Также легко проверить, что и множество рациональных дробей $P(x_1, x_2, \ldots, x_n)$ образует векторное пространство над полем P. Возникает вопрос об описании базы этого пространства. Здесь мы рассмотрим этот вопрос для случая $n = 1$.

Векторное пространство $P[x]$ обладает базой

$$1, x, x^2, \ldots,$$

и любой многочлен является линейной комбинацией этих многочленов. Но, как легко заметить, эти многочлены уже не образуют базу векторного пространства $P(x)$.

О п р е д е л е н и е. Несократимая дробь $\frac{x^n}{p^m}$, $n \in \mathbb{N} \cup \{0\}$, $m \in \mathbb{N}$ называется *простейшей,* если p – неразложимый многочлен над P со старшим коэффициентом 1 и $\deg p > n$.

П р и м е р. Простейшими дробями в $\mathbb{R}(x)$ будут дроби

$$\frac{1}{(x-\alpha)^m}, \quad \frac{1}{(x^2+px+q)^m}, \quad \frac{x}{(x^2+px+q)^m},$$

где $\alpha \in \mathbb{R}$, а x^2+px+q – многочлен, неразложимый над $\mathbb{R}$.

Т е о р е м а 2. *Многочлены* $1, x, x^2, \ldots$ *и простейшие дроби образуют базу векторного пространства* $P(x)$ *над полем* P.

Д о к а з а т е л ь с т в о. Докажем вначале линейную независимость. Предположим, что некоторая линейная комбинация равна нулю, т. е.

$$f + \sum_{m=1}^{m_1} \frac{f_{1,m}}{p_1^m} + \sum_{m=1}^{m_2} \frac{f_{2,m}}{p_2^m} + \ldots + \sum_{m=1}^{m_s} \frac{f_{s,m}}{p_s^m} = 0,$$

где $f \in P[x]$, $f_{i,m} \in P[x]$ и $\deg f_{i,m} < \deg p_i$. Здесь мы объединили линейную комбинацию базисных многочленов в f и считаем, что все p_i различны. Так как если у двух дробей один и тот же множитель p_i в знаменателе, то мы можем найти их сумму и записать в виде одной дроби. Надо доказать, что $f = 0$ и все $f_{i,m} = 0$. Умножим обе части нашего равенства на $p_1^{m_1} p_2^{m_2} \ldots p_s^{m_s}$ и если $f \neq 0$, то мы получим равенство

$$(\text{старший коэффициент } f)x^{\deg f + m_1 \deg p_1 + \ldots + m_s \deg p_s} +$$

$$+\{\text{остальные слагаемые}\} = 0.$$

Покажем, что это равенство невозможно, так как степень первого слагаемого выше степени любого другого слагаемого. Действительно, если рассмотреть первое слагаемое первой суммы

$$f_{1,1} p_1^{m_1 - 1} p_2^{m_2} \ldots p_s^{m_s},$$

то оно имеет степень

$$\deg f_{1,1} + (m_1 - 1)\deg p_1 + m_2 \deg p_2 + \ldots + m_s \deg p_s <$$

$$< m_1 \deg p_1 + m_2 \deg p_2 + \ldots + m_s \deg p_s.$$

Аналогично проверяется, что и любое другое слагаемое имеет меньшую степень. Следовательно, $f = 0$.

Предположим теперь, что $f_{1,m_1} \neq 0$. Умножая на то же произведение $p_1^{m_1}\ p_2^{m_2}\ \ldots p_s^{m_s}$ слагаемое $\frac{f_{1,m_1}}{p_1^{m_1}}$, получим $f_{1,m_1} p_2^{m_2} \ldots p_s^{m_s}$, а все остальные слагаемые будут содержать множитель p_1. Следовательно,

$$f_{1,m_1}\, p_2(x)^{m_2} \ldots p_s(x)^{m_s} + p_1[\ldots] = 0.$$

Так как f_{1,m_1} и ни один из сомножителей $p_2, \ldots, p_s$ не делятся на p_1 (все p_i неразложимы), то $f_{1,m_1} = 0$. Аналогично проверяется, что и все $f_{i,m_1} = 0$.

Докажем максимальность. Возьмем произвольную дробь $\frac{f}{g} \in P(x)$. Пусть $g = p_1^{m_1} p_2^{m_2} \ldots p_s^{m_s}$, где $p_1, p_2, \ldots, p_m$ – различные неразложимые многочлены и старший коэффициент у каждого p_i равен 1. Проведем индукцию по s.

Если $s = 0$, то $g = 1$ и $\frac{f}{g}$ – просто многочлен, а потому разлагается в линейную комбинацию многочленов $1, x, x^2, \ldots$

Если $s = 1$, то $\frac{f}{g} = \frac{f}{p_1^{m_1}}$. Разделим f с остатком на p_1, получим

$$f = p_1 q_1 + r_1,$$

где $r_1 = 0$ или $\deg r_1 < \deg p_1$. Тогда

$$\frac{f}{p_1^{m_1}} = \frac{q_1}{p_1^{m_1-1}} + \frac{r_1}{p_1^{m_1}},$$

и дробь $\frac{r_1}{p_1^{m_1}}$ либо равна нулю, либо является простейшей. Если дробь, $\frac{q_1}{p_1^{m_1-1}}$ не является простейшей, то представим q_1 в виде $q_1 = p_1 q_2 + r_2$ и т. д. Получим разложение $\frac{f}{p_1^{m_1}}$ в виде линейной комбинации простейших дробей.

Предположим, что для $s-1$ утверждение справедливо, и рассмотрим два многочлена $p_1^{m_1}$ и $p_2^{m_2} p_3^{m_3} \ldots p_s^{m_s}$. Они взаимно просты, а потому существуют многочлены $u, v \in P[x]$ такие, что

$$p_1^{m_1} u + p_2^{m_2} p_3^{m_3} \ldots p_s^{m_s} v = 1.$$

Тогда

$$\frac{f}{g} = \frac{fu}{p_2^{m_2} p_3^{m_3} \dots p_s^{m_s}} + \frac{fv}{p_1^{m_1}}.$$

По предположению индукции и разобранному выше случаю $s = 1$ каждая из дробей в правой части является линейной комбинацией простейших дробей. Теорема доказана.

В о п р о с. Из каких элементов состоит база векторного пространства $P(x_1, x_2, \dots, x_n)$ при $n > 1$?

§ 29. Кольца с однозначным разложением

29.1. Равносильные определения кольца с однозначным разложением. Пусть K – целостное кольцо с единицей. Напомним некоторые определения, введенные ранее. Элементы a и b кольца K называются *ассоциированными*, если $a = b \cdot \varepsilon$ для некоторого $\varepsilon \in K^*$, где K^* – множество обратимых элементов кольца K. Ненулевой, необратимый элемент $a \in K$ называется *неразложимым,* если из того, что $a = bc$, следует, что $b \in K^*$ или $c \in K^*$. Целостное кольцо K с единицей называется *кольцом с однозначным разложением*, если:

а) всякий ненулевой необратимый элемент из K разлагается в произведение неразложимых;

б) если $a = p_1 p_2 \dots p_r = q_1 q_2 \dots q_s$ – два таких разложения, то $r = s$ и p_i ассоциирован с q_i (возможно после перенумерации) для всех $i = 1, 2, \dots, r$.

П р и м е р ы колец с однозначным разложением. 1) Кольцо целых чисел $\mathbb{Z}$. Для него $\mathbb{Z}^* = \{\pm 1\}$, ассоциированными являются элементы $\pm a$, а неразложимыми $\pm p$, где p – простое натуральное число.

2) Кольцо многочленов $P[x]$ от одной переменной над полем P. В этом случае множество обратимых элементов совпадает с $P^* = P \setminus \{0\}$, ассоциированными являются элементы αf, где $0 \neq \alpha \in P$, $f \in P[x]$.

Многочлен $f(x)$, неразложимый в кольце $K[x]$, часто называют *неразложимым над* K, или *неприводимым над* K.

Следующая теорема дает другое определение кольца с однозначным разложением.

Т е о р е м а 1. *Условия* а *и* б *равносильны условиям* а *и* б′, *где*

б′) *если* $p \mid ab$ *и* p *– неразложим, то* $p \mid a$ *или* $p \mid b$.

Д о к а з а т е л ь с т в о. Докажем вначале, что из условий а и б следуют условия а и б′. Пусть $p \mid ab$, т. е. $ab = pc$ и

$$a = a_1 a_2 \dots a_k, \quad b = b_1 b_2 \dots b_l, \quad c = c_1 c_2 \dots c_n$$

– разложения в произведение неразложимых. Тогда

$$a_1 a_2 \dots a_k b_1 b_2 \dots b_l = p c_1 c_2 \dots c_n$$

– две записи одного и того же элемента в произведение неразложимых. Тогда по б элемент p ассоциирован с некоторым a_i или с некоторым b_j, но это означает, что $p \mid a$ или $p \mid b$.

Докажем теперь, что из условий а и б′ следуют условия а и б. Пусть

$$p_1 p_2 \dots p_r = q_1 q_2 \dots q_s$$

– два разложения одного элемента в произведения неразложимых. Если $p_i \mid q_1 q_2 \dots q_s$, то по б′ $p_i \mid q_1$ или $p_i \mid q_2 \dots q_s$. Если $p_i \mid q_1$, то $q_1 = p_i c_i$ для некоторого $c_i \in K$, а так как q_1 неразложим, то $c_i \in K^*$, а потому p_i ассоциирован с q_1. Если $p_i \mid q_2 \dots q_s$, то опять либо $p_i \mid q_2$ либо $p_i \mid q_3 \dots q_s$ и т. д. Следовательно, p_i ассоциирован с некоторым q_j. Теорема доказана.

Л е м м а 1. *Если* K *– кольцо с однозначным разложением, то для любых ненулевых* $a_1, a_2, \dots, a_s$ *из* K *существует элемент* $d \in K$, *для которого выполняются следующие условия*:

а) $d \mid a_1, \quad d \mid a_2, \dots, d \mid a_s$;

б) *если некоторый* d' *делит* $a_1, a_2, \dots, a_s$, *то* $d' \mid d$.

Это d единственно с точностью до множителя из K^* и называется *наибольшим общим делителем* элементов $a_1, a_2, \dots, a_s$.

Д о к а з а т е л ь с т в о. Пусть $a_i = \varepsilon_i p_1^{\alpha_{i1}} p_2^{\alpha_{i2}} \dots p_r^{\alpha_{ir}}$, где $\varepsilon_i \in K^*$, $p_1, p_2, \dots, p_r$ – различные неразложимые и α_{ij} – неотрицательные целые. Возьмем $d = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$, где $\alpha_j = \min\{\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{sj}\}$. Это и есть наибольший общий делитель. Действительно, так как $d \mid a_i$, $i = 1, 2, \dots, s$, то условие а выполняется. Если $d' \mid a_i$, то $d' = p_1^{\beta_1} p_2^{\beta_2} \dots p_s^{\beta_s}$, где $\beta_i \le \alpha_i$. Следовательно, $d' \mid d$ и условие б также выполняется. Лемма доказана.

29.2. Примитивные многочлены. Пусть K – кольцо с однозначным разложением. Многочлен $f \in K[x]$ называется *примитивным,* если наибольший общий делитель его коэффициентов обратим в K, т. е. лежит в K^*.

П р и м е р. Пусть $K = \mathbb{Z}$ и $f(x) = -30x^2 + 12x - 3$. Легко проверить, что наибольший общий делитель его коэффициентов равен 3, а потому $f(x)$ не является примитивным (напомним, что $\mathbb{Z}^* = \{\pm 1\}$).

Л е м м а 2. *Произведение двух примитивных многочленов является примитивным многочленом.*

Д о к а з а т е л ь с т в о. Предположим, что утверждение леммы неверно, т. е. существуют примитивные многочлены

$$f(x) = a_0 + a_1 x + \ldots + a_n x^n$$

и

$$g(x) = b_0 + b_1 x + \ldots + b_m x^m$$

такие, что их произведение

$$f(x)g(x) = a_0 b_0 + (a_0 b_1 + a_1 b_0)x + \ldots + a_n b_m x^{n+m}$$

не является примитивным многочленом, т. е. найдется общий делитель коэффициентов, который неразложим в K. Пусть p – неразложимый элемент из K, делящий все коэффициенты многочлена fg. Заметим, что p не может делить все коэффициенты f и все коэффициенты g. Пусть

$$p \mid a_0, \quad p \mid a_1, \ldots, p \mid a_{i-1}, \quad \text{но} \quad p \nmid a_i;$$

$$p \mid b_0, \quad p \mid b_1, \ldots, p \mid b_{j-1}, \quad \text{но} \quad p \nmid b_j.$$

В произведении fg коэффициент при x^{i+j} имеет вид

$$a_0 b_{i+j} + a_1 b_{i+j-1} + \ldots + a_{i-1} b_{j+1} + a_i b_j + a_{i+1} b_{j-1} + \ldots + a_{i+j} b_0.$$

Заметим, что все слагаемые, стоящие перед $a_i b_j$, делятся на p и все слагаемые, стоящие после $a_i b_j$, также делятся на p. Так как по предположению все коэффициенты у fg делятся на p, то произведение $a_i b_j$ делится на p. Учитывая, что K – кольцо с однозначным разложением, заключаем, что $p \mid a_i$ или $p \mid b_j$. Полученное противоречивое и доказывает лемму.

29.3. Кольцо многочленов над кольцом с однозначным разложением – само кольцо с однозначным разложением. Докажем предварительно некоторые утверждения, которые потребуются нам в дальнейшем.

Л е м м а 3. *Множество обратимых элементов кольца $K[x]$ совпадает с множеством обратимых элементов кольца K.*

Д о к а з а т е л ь с т в о. Пусть $f \in K[x]$ обратим. Тогда для некоторого многочлена $g \in K[x]$ имеем $fg = 1$, но так как при умножении многочленов степени складываются, то многочлены f и g имеют нулевую степень, т. е. лежат в K, а так как их произведение равно 1, то f является обратимым элементом кольца K.

Обратно, если $\alpha \in K^*$, то $\alpha \in K[x]$, а потому является обратимым элементом кольца $K[x]$. Лемма доказана.

Л е м м а 4. *Элемент $a \in K$ разложим в K тогда и только тогда, когда a разложим в $K[x]$.*

Д о к а з а т е л ь с т в о. Пусть $a = bc$ – разложение в K, где $b, c \notin K^*$, но по лемме 3 это значит, что $b, c \notin K[x]^*$, а потому a разложим в $K[x]$.

Обратно. Пусть $a = bc$ – разложение в $K[x]$, т. е. $b, c \in K[x] \setminus K[x]^*$, но, учитывая, что $\deg a = 0$, заключаем, что и $\deg b = \deg c = 0$, т. е. $b, c \in K$, а потому $b, c \in K \setminus K^*$. Лемма доказана.

Л е м м а 5. *Пусть K – кольцо с однозначным разложением, $\ddot{K}$ – его поле частных. Тогда*: 1) *для всякого $f \in \ddot{K}[x]$ найдутся $\alpha \in \ddot{K}$ и примитивный многочлен g из $K[x]$ такие, что $f = \alpha g$*; 2) *эти α и g определяются единственным образом с точностью до множителя из K^**; 3) *многочлен f разложим в $\ddot{K}[x]$ тогда и только тогда, когда g разложим в $K[x]$.*

Д о к а з а т е л ь с т в о. 1) Докажем существование. Напомним, что

$$\ddot{K} = \left\{ \frac{a}{b} \mid a, b \in K, b \neq 0 \right\}.$$

Пусть

$$f(x) = \frac{a_0}{b_0} + \frac{a_1}{b_1}x + \ldots + \frac{a_n}{b_n}x^n.$$

Вынося знаменатели всех коэффициентов, получим

$$f(x) = \frac{1}{b_0 b_1 \ldots b_n} F(x),$$

где

$$F(x) = a_0 b_1 b_2 \ldots b_n + a_1 b_0 b_2 \ldots b_n x + \ldots + a_n b_0 b_1 \ldots b_{n-1} x^n$$

– многочлен из $K[x]$. Пусть d – наибольший общий делитель его коэффициентов, т. е. $F(x) = dg(x)$, где $g(x)$ – примитивный многочлен из $K[x]$. Полагая $\alpha = \frac{d}{b_0 b_1 \ldots b_n}$, получим искомое разложение.

П р и м е р. Для многочлена

$$f(x) = \frac{15}{7}x^3 - \frac{30}{35}x + \frac{5}{7} = \frac{5}{35}\left(15x^3 - 6x + 5\right)$$

имеем $\alpha = \frac{5}{35}$ и $g(x) = 15x^3 - 6x + 5$.

2) Докажем единственность. Пусть $f = \alpha g$, где $\alpha \in \ddot{K}$, а g – примитивный многочлен из $K[x]$. С другой стороны, f имеет разложение $f = \alpha' g'$, где $\alpha' \in \ddot{K}$, а g' – примитивный многочлен из $K[x]$. Пусть $\alpha = \frac{c}{d}$, $\alpha' = \frac{c'}{d'}$ для некоторых c, d, c', d' из K. Тогда из равенства $\alpha g = \alpha' g'$ получим $cd'g = c'dg'$. Заметим, что наибольший общий делитель коэффициентов многочлена из левой части равен cd', а наибольший общий делитель коэффициентов многочлена из правой части равен $c'd$. По лемме 1 $cd' = \varepsilon c'd$, где $\varepsilon \in K^*$. Отсюда

$$\frac{c}{d} = \varepsilon\frac{c'}{d'},$$

т. е. $\alpha = \varepsilon\alpha'$. Из равенства $\alpha g = \alpha' g'$ заключаем, что $g' = \varepsilon g$.

3) Докажем импликацию $\Rightarrow$. Пусть f разложим в $\ddot{K}[x]$, т. е. $f = f_1 f_2$, где f_1 и f_2 – необратимые элементы в $\ddot{K}[x]$, т. е. f_1 и f_2 – многочлены ненулевой степени. Имеем

$$f = \alpha g, \quad f_1 = \alpha_1 g_1, \quad f_2 = \alpha_2 g_2,$$

где $\alpha, \alpha_1, \alpha_2 \in \ddot{K}$, g, g_1, g_2 – примитивные многочлены из $K[x]$. Отсюда $\alpha g = \alpha_1\alpha_2 g_1 g_2$ и по лемме 2 многочлен $g_1 g_2$ является примитивным. По предположению леммы о единственности разложения $g = \varepsilon g_1 g_2$, $\varepsilon \in K^*$ и g_1, g_2 – многочлены ненулевой степени. Значит, g разложим в $K[x]$.

Докажем импликацию $\Leftarrow$. Пусть g разложим в $K[x]$, т. е. $g = g_1 g_2$, где g_1 и g_2 необратимы в $K[x]$. Так как g_1 и g_2 не являются обратимыми элементами кольца $K[x]$, то они не лежат в $K[x]^* = K^*$. Если $g_1 \in K$, то $g_2 \in K \setminus K^*$, но g_1 не может лежать в K, так как это противоречит примитивности g, т. е. $g_1, g_2 \notin K$, а потому g_1 и g_2 являются нетривиальными многочленами (имеют степень больше нуля). Ввиду установленного пункта 1 имеем

$$f = \alpha g = \alpha g_1 g_2,$$

где αg_1 и g_2 – многочлены ненулевой степени, а потому мы построили нетривиальное разложение в $\ddot{K}[x]$. Следовательно, f разложим в $\ddot{K}[x]$. Лемма доказана.

Из этой леммы вытекает такое следствие, которое мы использовали при построении примера целостного кольца с неоднозначным разложением.

С л е д с т в и е. *Многочлен $f(x) \in \mathbb{Z}[x]$ разложим в $\mathbb{Z}[x]$ тогда и только тогда, когда он разложим в $\mathbb{Q}[x]$.*

Теперь мы готовы доказать основное утверждение настоящего параграфа.

Т е о р е м а 2. *Если K – кольцо с однозначным разложением, то $K[x]$ – кольцо с однозначным разложением.*

Д о к а з а т е л ь с т в о. Мы знаем, что $\ddot{K}[x]$ и K – кольца с однозначным разложением. Надо проверить, что кольцо $K[x]$ удовлетворяет определению кольца с однозначным разложением.

а) Пусть f – ненулевой, необратимый элемент из $K[x]$. Как элемент из $\ddot{K}[x]$ многочлен f разлагается на неразложимые сомножители:

$$f = f_1 f_2 \dots f_s,$$

где f_i – неразложимы и лежат в $\ddot{K}[x]$. Тогда по лемме 5

$$f_1 f_2 \dots f_s = (\alpha_1 g_1)(\alpha_2 g_2) \dots (\alpha_s g_s) = \alpha g_1 g_2 \dots g_s, \quad \alpha = \alpha_1 \alpha_2 \dots \alpha_s \in K,$$

где α_i лежат в $\ddot{K}$, а g_i – примитивные из $K[x]$. Учитывая, что K – кольцо с однозначным разложением, получим разложение в $K[x]$:

$$f = \alpha g_1 g_2 \dots g_s = \overline{\alpha}_1 \overline{\alpha}_2 \dots \overline{\alpha}_r g_1 g_2 \dots g_s,$$

где $\alpha = \overline{\alpha}_1 \overline{\alpha}_2 \dots \overline{\alpha}_r$, все $\overline{\alpha}_j$ неразложимы и лежат в K, что следует из леммы 4, а все элементы g_i неразложимы в $K[x]$, что следует из леммы 5. Таким образом, мы нашли разложение на неразложимые множители.

б) Пусть мы имеем два разложения на неразложимые множители:

$$f = \alpha_1 \alpha_2 \dots \alpha_r g_1 g_2 \dots g_s = \alpha'_1 \alpha'_2 \dots \alpha'_{r'} g'_1 g'_2 \dots g'_{s'},$$

где $\alpha_i, \alpha'_i \in K$, а g_j, g'_j неразложимы в $K[x]$ и не лежат в K. Надо доказать, что число множителей в этих разложениях одинаково и сами множители ассоциированы.

Можно считать, что все g_j, g'_j примитивны. По лемме 5

$$\alpha_1 \alpha_2 \dots \alpha_r = \varepsilon \alpha'_1 \alpha'_2 \dots \alpha'_{r'} \tag{1}$$

и

$$\varepsilon g_1 g_2 \dots g_s = g'_1 g'_2 \dots g'_{s'}. \tag{2}$$

По лемме 4 все α_i, α_i' неразложимы в K. Так как K – кольцо с однозначным разложением, то $r = r'$ и возможно после перенумерации:

$$\alpha_1 \text{ ассоциирован с } \varepsilon\alpha_1', \text{ т. е. ассоциирован с } \alpha_1',$$

$$\alpha_2 \text{ ассоциирован с } \alpha_2',$$

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

$$\alpha_r \text{ ассоциирован с } \alpha_r'.$$

Рассмотрим теперь разложение (2). В кольце $\ddot{K}[x]$ имеем $s = s'$ и возможно после перенумерации:

$$g_1 \text{ ассоциирован с } \varepsilon g_1, \text{ т. е. ассоциирован с } g_1',$$

$$g_2 \text{ ассоциирован с } g_2',$$

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

$$g_s \text{ ассоциирован с } g_s'.$$

Если $1g_j = \delta_j g_j'$, $\delta_j \in \left(\ddot{K}[x]\right)^* = \ddot{K}^*$, а так как $\delta_j = 1\varepsilon_j$, $\varepsilon_j \in K^*$, то $\delta_j \in K^* = K[x]^*$. Теорема доказана.

Из доказанной теоремы легко получается такое

С л е д с т в и е. *Всякое кольцо $K[x_1, x_2, \ldots, x_n]$, где K – кольцо с однозначным разложением, тоже является кольцом с однозначным разложением. В частности, $\mathbb{Z}[x_1, x_2, \ldots, x_n]$ и $P[x_1, x_2, \ldots, x_n]$ являются кольцами с однозначным разложением.*

29.4. Неразложимые многочлены над $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$ и $\mathbb{Z}$. Неразложимые многочлены над $\mathbb{C}$ – это линейные многочлены вида $x - \alpha$, $\alpha \in \mathbb{C}$, что следует из алгебраической замкнутости поля $\mathbb{C}$ и теоремы Безу. Действительно, так как всякий многочлен $f(x) \in \mathbb{C}[x]$ ненулевой степени имеет корень в $\mathbb{C}$, то

$$f(x) = (x - \alpha)g(x), \quad g(x) \in \mathbb{C}[x].$$

Многочлены первой степени уже неразложимы.

Многочлены, неразложимые над $\mathbb{R}$, могут быть первой и второй степени. Пусть $f(x) \in \mathbb{R}[x]$ и степень $\deg f \geq 3$. Если α – корень многочлена $f(x)$, то $\overline{\alpha}$ – также корень $f(x)$. Если $\alpha \in \mathbb{R}$, то по теореме Безу

$$f(x) = (x - \alpha)g(x), \quad g(x) \in \mathbb{R}[x].$$

Если $\alpha \notin \mathbb{R}$, то $\overline{\alpha} \neq \alpha$, и опять по теореме Безу

$$f(x) = (x-\alpha)(x-\overline{\alpha})h(x), \quad h(x) \in \mathbb{C}[x].$$

Рассмотрим многочлен

$$(x-\alpha)(x-\overline{\alpha}) = x^2 - (\alpha+\overline{\alpha}) + \alpha\overline{\alpha}.$$

Так как $\alpha+\overline{\alpha} \in \mathbb{R}$ и $\alpha\overline{\alpha} = |\alpha|^2 \in \mathbb{R}$, то $(x-\alpha)(x-\overline{\alpha}) \in \mathbb{R}[x]$, а потому $f(x)$ делится на квадратичный многочлен с действительными коэффициентами, а потому и $h(x) \in \mathbb{R}[x]$, т. е. $f(x)$ разложим над $\mathbb{R}$.

Над полем $\mathbb{Q}$ существуют неразложимые многочлены любой степени.

У п р а ж н е н и е. Доказать, что для любого простого числа p и всякого натурального n многочлен $x^n - p$ неразложим над $\mathbb{Q}$.

Заметим, что если многочлен с целыми коэффициентами неразложим над $\mathbb{Q}$, то по следствию леммы 5 он неразложим и над $\mathbb{Z}$. Укажем алгоритм распознавания разложимости или неразложимости многочлена над $\mathbb{Z}$.

Пусть $f \in \mathbb{Z}[x]$ и n – его степень. Хотим проверить, представим ли он в виде $f = dh$, где $d, h \in \mathbb{Z}[x]$ и $\deg d > 0$, $\deg h > 0$. Можно считать, что $\deg d = N \leq \left[\frac{n}{2}\right]$, где $[a]$ – целая часть числа a. Выберем различные целые числа $x_0, x_1, \ldots, x_N$. Если $d(x)|f(x)$, то $d(x_i)|f(x_i)$, $i = 0, 1, \ldots, N$. Найдем все возможные делители $f(x_i)$. Затем по интерполяционной формуле Лагранжа найдем некоторый многочлен $p(x)$, пользуясь таблицей

x	x_0	$\ldots$	x_i	$\ldots$	x_N
$f(x)$	$f(x_0)$	$\ldots$	$f(x_i)$	$\ldots$	$f(x_N)$
$d(x)$	$d(x_0)$	$\ldots$	$d(x_i)$	$\ldots$	$d(x_N)$

т. е. $p(x_i) = d(x_i)$. Если $p(x) \mid f(x)$, то мы найдем разложение многочлена $f(x)$. Если, перебрав все такие делители, не найдем многочлена $p(x)$, который делит $f(x)$, то последний неразложим над $\mathbb{Z}$.

Глава 5

ЛИНЕЙНЫЕ ПРЕОБРАЗОВАНИЯ ВЕКТОРНЫХ ПРОСТРАНСТВ

§ 30. Линейные преобразования

30.1. Определения и примеры. Пусть V – векторное пространство над полем P. Отображение $\varphi : V \longrightarrow V$ называется *линейным преобразованием* пространства V, если

$$(\alpha u + \beta v)\varphi = \alpha(u\varphi) + \beta(v\varphi)$$

для всех $\alpha, \beta \in P$ и $u, v \in V$.

П р и м е р ы линейных преобразований.

1) Нулевое преобразование 0, переводящее всякий вектор в нулевой вектор.

2) Тождественное преобразование ε, переводящее всякий вектор $u \in V$ в себя.

3) Пусть $P[x]$ – пространство многочленов. Дифференцирование является линейным преобразованием, что следует из свойства производных: $(\alpha f + \beta g)' = \alpha f' + \beta g'$ для любых $\alpha, \beta \in P$ и $f, g \in P[x]$.

4) Для всякой подстановки π из S_n можно определить преобразование φ_π на пространстве n–ок P^n, полагая

$$(\alpha_1, \alpha_2, \ldots, \alpha_n)\varphi_\pi = (\alpha_{1\pi}, \alpha_{2\pi}, \ldots, \alpha_{n\pi}).$$

В частности, если $\pi = (12)$, то

$$(\alpha_1, \alpha_2, \ldots, \alpha_n)\varphi_{(12)} = (\alpha_2, \alpha_1, \alpha_3, \ldots, \alpha_n).$$

5) Функция $f : V \longrightarrow P$ называется *линейной*, если

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)$$

для всех $\alpha, \beta \in P$ и $u, v \in V$. Пусть $e_1, e_2, \ldots, e_n$ – база пространства V, $f_1, f_2, \ldots, f_n$ – линейные функции из V в P. Определим преобразование $\varphi : V \longrightarrow V$, положив

$$u\varphi = f_1(u)e_1 + f_2(u)e_2 + \ldots + f_n(u)e_n$$

для всякого вектора u из V. Это преобразование линейно. В некотором смысле это общий пример линейного преобразования.

30.2. Матрица линейного преобразования. Пусть φ – линейное преобразование векторного пространства V,

$$e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

– база V. Действуя на каждый базисный вектор e_i пространства V преобразованием φ, получим

$$\begin{cases} e_1\varphi = \alpha_{11}e_1 + \alpha_{12}e_2 + \ldots + \alpha_{1n}e_n, \\ e_2\varphi = \alpha_{21}e_1 + \alpha_{22}e_2 + \ldots + \alpha_{2n}e_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ e_n\varphi = \alpha_{n1}e_1 + \alpha_{n2}e_2 + \ldots + \alpha_{nn}e_n, \end{cases} \tag{1}$$

где $\alpha_{ij} \in P$.

Матрицу

$$[\varphi] = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{pmatrix}$$

будем называть *матрицей линейного преобразования φ в базе e*. Если необходимо подчеркнуть, что эта матрица в базе e, то будем писать $[\varphi]_e$. Кратко система (1) записывается в виде

$$e\varphi = [\varphi]e.$$

П р и м е р. В пространстве многочленов $P[x]$ выберем базу 1, x, $x^2, \ldots$ Чтобы найти матрицу преобразования дифференцирования φ, действуем на каждый вектор преобразованием φ. Получим

$$\begin{cases} 1\varphi = 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + \ldots, \\ x\varphi = 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + \ldots, \\ x^2\varphi = 2x = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 + \ldots, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

Тогда

$$[\varphi] = \begin{pmatrix} 0 & 0 & 0 & \ldots \\ 1 & 0 & 0 & \ldots \\ 0 & 2 & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{pmatrix}.$$

Если рассмотреть пространство $P_{\leq n}[x]$ – многочлены степени, не превосходящей n, для некоторого фиксированного n, то это конечномерное пространство с базой $1, x, x^2, \ldots, x^n$, в которой матрица дифференцирования имеет вид

$$[\varphi] = \begin{pmatrix} 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & n & 0 \end{pmatrix}.$$

30.3. Координаты образа вектора. Пусть u – произвольный вектор из V, разлагая его по базе e, получим

$$u = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = [u]e, \tag{2}$$

где $[u] = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ – координаты вектора u в базе e. Подействуем теперь на вектор u преобразованием φ, получим

$$u\varphi = [u\varphi]e, \tag{3}$$

где $[u\varphi]$ – координаты вектора $u\varphi$. Хотим найти связь между координатами $[u]$ и $[u\varphi]$.

Вспоминая определение матрицы преобразования φ, имеем

$$e\varphi = [\varphi]e.$$

Ввиду линейности φ

$$u\varphi = (\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n)\varphi = \alpha_1(e_1\varphi) + \alpha_2(e_2\varphi) + \ldots + \alpha_n(e_n\varphi) =$$

$$= [u](e\varphi) = [u]([\varphi]e) = ([u][\varphi])e.$$

Напомним следующее утверждение (см. лемму 3 из § 11).

Л е м м а (о сокращении). *Пусть $v_1, v_2, \ldots, v_n$ – линейно независимая система векторов, X и Y – матрицы размера $m \times n$. Если положить $v = (v_1, v_2, \ldots, v_n)^t$, где t означает транспонирование, то из равенства $Xv = Yv$ следует равенство матриц: $X = Y$.*

Используя эту лемму, получим искомое равенство

$$[u][\varphi] = [u\varphi],$$

т. е. координаты вектора $u\varphi$ есть произведение координат вектора u на матрицу преобразования φ.

30.4. Связь между матрицами линейного преобразования в разных базах. Пусть

$$e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}, \quad e' = \begin{pmatrix} e'_1 \\ e'_2 \\ \vdots \\ e'_n \end{pmatrix}$$

– две базы векторного пространства V. Мы можем найти матрицы линейного преобразования φ в этих базах:

$$e\varphi = [\varphi]e, \tag{4}$$

$$e'\varphi = [\varphi]'e', \tag{5}$$

получим соответственно матрицы $[\varphi]$ и $[\varphi]'$. Возникает естественный вопрос: как связаны эти матрицы?

Мы знаем, что базы e и e' связаны матрицей перехода, т. е.

$$e' = Te \tag{6}$$

для некоторой невырожденной матрицы $T = (t_{ij}) \in M_n(P)$. Тогда, используя равенства (4)–(5) и (6), получим

$$(T[\varphi])e = T([\varphi]e) = T(e\varphi) = (Te)\varphi = e'\varphi = [\varphi]'e' = [\varphi]'(Te) = ([\varphi]'T)e.$$

Воспользовавшись сформулированной выше леммой, получим

$$T[\varphi] = [\varphi]'T.$$

Домножив обе части справа на T^{-1}, получим окончательную формулу

$$[\varphi]' = T[\varphi]T^{-1}.$$

30.5. Алгебра линейных преобразований. Символом L обозначим множество линейных преобразований векторного пространства V. Определим на L операции сложения, умножения и умножения на скаляр из P.

Пусть φ и ψ – два линейных преобразования из L. Их *суммой* $\varphi + \psi$ назовем преобразование, которое действует по правилу

$$u(\varphi + \psi) = u\varphi + u\psi, \quad u \in V.$$

Их *произведением* $\varphi \cdot \psi$ назовем преобразование, которое действует по правилу

$$u(\varphi \cdot \psi) = (u\varphi)\psi,$$

и, наконец, для произвольного α из P определим преобразование $f_\alpha(\varphi) = \alpha\varphi$, полученное умножением на скаляр α формулой

$$u(\alpha\varphi) = (\alpha u)\varphi.$$

Легко убедиться, что L относительно этих операций является алгебраической системой:

$$\langle\, L;\, +,\, \cdot,\, f_\alpha\,(\alpha \in P)\,\rangle.$$

О п р е д е л е н и е. Алгебраическая система

$$\langle\, A;\, +,\, \cdot,\, f_\alpha\,(\alpha \in P)\,\rangle$$

называется *линейной алгеброй* над полем P, если алгебраическая система $\langle\, A;\, +,\, f_\alpha\,(\alpha \in P)\,\rangle$ является векторным пространством над полем P, а $\langle\, A;\, +,\, \cdot\,\rangle$ является кольцом и выполняется следующая аксиома:

$$\alpha(\varphi\psi) = (\alpha\varphi)\psi = \varphi(\alpha\psi), \text{ для любых } \varphi, \psi \in A, \quad \alpha \in P.$$

П р и м е р ы линейных алгебр. 1) Множество квадратных матриц $M_n(P)$ над полем P с операциями сложения, умножения матриц и умножения на скаляр образует линейную алгебру над полем P.

2) Множество многочленов $P[x_1, x_2, \ldots, x_n]$ над полем P с операциями сложения, умножения многочленов и умножения на скаляр образует линейную алгебру над полем P.

Покажем, что множество линейных преобразований векторного пространства является линейной алгеброй, которая изоморфна алгебре матриц. Более точно, справедлива

Т е о р е м а. *Множество линейных преобразований векторного пространства V размерности n над полем P является линейной алгеброй, изоморфной алгебре матриц $M_n(P)$.*

Д о к а з а т е л ь с т в о. Установим изоморфизм. Пусть e – некоторая база пространства V. Сопоставим каждому линейному преобразованию φ векторного пространства V его матрицу $[\varphi]$ в базе e. Покажем, что это отображение

$$\omega : L \longrightarrow M_n(P) \text{ по правилу } \varphi\omega = [\varphi]$$

и есть искомый изоморфизм.

Так как каждому линейному преобразованию соответствует его матрица, которая определяется однозначно при фиксированной базе, то наше отображение однозначно.

Предположим, что для двух линейных преобразований $\varphi, \psi \in L$ их матрицы равны: $[\varphi] = [\psi]$. Покажем, что тогда и $\varphi = \psi$. Для этого достаточно показать, что для произвольного вектора $v \in V$ справедливо равенство $v\varphi = v\psi$. Так как $v = [v]e$, то $v\varphi = [v][\varphi]e$. С другой стороны, $v\psi = [v][\psi]e$, а так как матрицы равны, то отсюда следует, что $v\varphi = v\psi$.

Проверим, что отображение $\omega : L \longrightarrow M_n(P)$ является отображением *на*. Пусть $A \in M_n(P)$, определим отображение φ по правилу $v\varphi = ([v]A)e \in V$. Проверим, что φ линейно. Выберем $v_1, v_2 \in V$, $\alpha_1, \alpha_2 \in P$, тогда

$$(\alpha_1 v_1 + \alpha_2 v_2)\varphi = [\alpha_1 v_1 + \alpha_2 v_2]Ae = (\alpha_1[v_1] + \alpha_2[v_2])Ae.$$

Воспользовавшись дистрибутивностью, получим

$$(\alpha_1[v_1] + \alpha_2[v_2])Ae = (\alpha_1[v_1]A + \alpha_2[v_2]A)e = \alpha_1[v_1]Ae + \alpha_2[v_2]Ae =$$

$$= \alpha_1(v_1\varphi) + \alpha_2(v_2\varphi).$$

Проверим, что матрица $[\varphi]$ в базе e равна A. Возьмем вектор e_i из базы e и подействуем на него преобразованием φ:

$$e_i\varphi = [0, \ldots, 0, 1, 0, \ldots, 0]Ae = a_{i1}e_1 + a_{i2}e_2 + \ldots + a_{in}e_n.$$

Следовательно, $[\varphi] = A$, т. е. для каждой матрицы A построили линейное преобразование, матрица которого равна A.

Проверим, что построенное отображение ω сохраняет операции. Возьмем преобразования $\varphi, \psi \in L$ и подействуем на базу их произведением:

$$e(\varphi \cdot \psi) = (e\varphi)\psi = ([\varphi]e)\psi = [\varphi](e\psi) = [\varphi]([\psi]e) = ([\varphi][\psi])e.$$

С другой стороны,

$$e(\varphi\psi) = [\varphi\psi]e.$$

Отсюда ввиду леммы $[\varphi\psi] = [\varphi][\psi]$. Таким образом, операция умножения сохраняется.

Рассмотрим теперь сумму двух линейных преобразований. По определению матрицы линейного преобразования $e(\varphi+\psi) = [\varphi+\psi]e$, а по определению суммы линейных преобразований получим:

$$e(\varphi + \psi) = e\varphi + e\psi = [\varphi]e + [\psi]e = ([\varphi] + [\psi])e.$$

Опять ввиду леммы приходим к равенству $[\varphi + \psi] = [\varphi] + [\psi]$. Следовательно, операция сложения сохраняется.

Пусть теперь $\alpha \in P$. Тогда

$$e(\alpha\varphi) = (\alpha e)\varphi = (\alpha[\varphi])e.$$

С другой стороны, $e(\alpha\varphi) = [\alpha\varphi]e$, а потому $[\alpha\varphi] = \alpha[\varphi]$. Изоморфизм установлен. Теорема доказана.

Из этой теоремы, в частности, следует, что линейные преобразования удовлетворяют аксиоме ассоциативности умножения, дистрибутивности и т. д.

§ 31. Образ и ядро линейного преобразования

31.1. Образ и ядро – подпространства векторного пространства. С каждым линейным преобразованием φ можно связать его *образ*

$$\operatorname{Im}\varphi = \{v\varphi \mid v \in V\} = V\varphi$$

и *ядро*

$$\operatorname{Ker}\varphi = \{v \in V \mid v\varphi = 0\}.$$

Справедлива

Т е о р е м а 1. *Для всякого линейного преобразования φ его образ $\operatorname{Im}\varphi$ и ядро $\operatorname{Ker}\varphi$ являются подпространствами пространства V.*

Д о к а з а т е л ь с т в о. Чтобы проверить, что $\operatorname{Im}\varphi$ является подпространством, возьмем $u_1, u_2 \in \operatorname{Im}\varphi$ и $\alpha_1,\ \alpha_2 \in P$ и покажем, что их линейная комбинация $\alpha_1 u_1 + \alpha_2 u_2$ лежит в образе $\operatorname{Im}\varphi$. Так как u_1 и u_2 лежат в образе, то для некоторых векторов $v_1,\ v_2$ имеем равенства $v_1\varphi = u_1,\ v_2\varphi = u_2$. Следовательно,

$$\alpha_1 u_1 + \alpha_2 u_2 = \alpha_1(v_1\varphi) + \alpha_2(v_2\varphi) = (\alpha_1 v_1 + \alpha_2 v_2)\varphi \in \operatorname{Im}\varphi.$$

Выберем теперь векторы u_1 и u_2 из $\operatorname{Ker}\varphi$. Действуя на линейную комбинацию $\alpha_1 u_1 + \alpha_2 u_2$ преобразованием φ, получим

$$(\alpha_1 u_1 + \alpha_2 u_2)\varphi = \alpha_1(u_1\varphi) + \alpha_2(u_2\varphi) = \alpha_1 0 + \alpha_2 0 = 0,$$

т. е. линейная комбинация $\alpha_1 u_1 + \alpha_2 u_2$ лежит в ядре $\operatorname{Ker}\varphi$. Теорема доказана.

Следующая теорема устанавливает связь между размерностью ядра и образа линейного преобразования.

Т е о р е м а 2. *Справедлива следующая формула*:

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

Д о к а з а т е л ь с т в о. Пусть $u_1, u_2, \ldots, u_s$ – база ядра $\operatorname{Ker}\varphi$, а $v_1, v_2, \ldots, v_r$ – база образа $\operatorname{Im}\varphi$. Для каждого v_i символом w_i обозначим его прообраз при преобразовании φ, т. е. $w_i\varphi = v_i$, $i = 1, 2, \ldots, r$.

Покажем, что система

$$u_1, u_2, \ldots, u_s, w_1, w_2, \ldots, w_r$$

образует базу пространства V. Проверим вначале, что эти векторы линейно независимы. Рассмотрим их линейную комбинацию:

$$\sum_{i=1}^{s} \alpha_i u_i + \sum_{j=1}^{r} \beta_j w_j = 0,$$

и, применяя преобразование φ, получим:

$$0 = \sum_{i=1}^{s} \alpha_i(u_i\varphi) + \sum_{j=1}^{r} \beta_j(w_j\varphi) = \sum_{j=1}^{r} \beta_j v_j.$$

Так как векторы $v_1, v_2, \ldots, v_r$ линейно независимы, то все $\beta_j = 0$. Из равенства

$$\sum_{i=1}^{s} \alpha_i u_i = 0$$

ввиду линейной независимости системы $u_1, u_2, \ldots, u_s$ заключаем, что и $\alpha_i = 0$. Следовательно, система векторов линейно независима.

Покажем, что система векторов является максимальной. Рассмотрим произвольный вектор $v \in V$. Тогда вектор $v\varphi$ лежит в $\operatorname{Im}\varphi$, а потому разлагается по базе, т. е.

$$v\varphi = \sum_{j=1}^{r} \gamma_j v_j$$

для некоторых $\gamma_j \in P$. Рассмотрим вектор

$$v - (\sum_{j=1}^{r} \gamma_j w_j)$$

и покажем, что он лежит в ядре $\operatorname{Ker}\varphi$. Действительно,

$$\left(v - \sum_{j=1}^{r} \gamma_j w_j\right)\varphi = v\varphi - \sum_{j=1}^{r} \gamma_j v_j = 0.$$

Следовательно,

$$v - \sum_{j=1}^{r} \gamma_j w_j = \sum_{i=1}^{s} \delta_i u_i$$

для некоторых δ_i из P, а потому мы имеем искомое разложение:

$$v = \sum_{j=1}^{r} \gamma_j w_j + \sum_{i=1}^{s} \delta_i u_i.$$

Теорема доказана.

31.2. Невырожденные линейные преобразования.

Т е о р е м а 3. *Для линейного преобразования φ векторного пространства V следующие утверждения эквивалентны*:

1) $\operatorname{Im}\varphi = V$;

2) $\operatorname{Ker}\varphi = 0$;

3) φ – *взаимно однозначно и на*;

4) *обратное преобразование φ^{-1} существует и линейно*;

5) *матрица $[\varphi]$ невырождена в любой базе*;

6) *матрица $[\varphi]$ невырождена в некоторой базе.*

Если выполнено любое из этих условий, то преобразование φ называется *невырожденным.*

Д о к а з а т е л ь с т в о. 1) $\Leftrightarrow$ 2). Равенство $\operatorname{Im}\varphi = V$, по предыдущей теореме, равносильно тому, что $\dim(\operatorname{Im}\varphi) = \dim V$ и $\dim(\operatorname{Ker}\varphi) = 0$, а это равносильно тому, что, $\operatorname{Ker}\varphi = 0$.

2) $\Rightarrow$ 3). Предположим, что $u\varphi = v\varphi$, тогда $(u - v)\varphi = 0$, т. е. $u - v \in \operatorname{Ker}\varphi$, а так как $\operatorname{Ker}\varphi = 0$, то $u - v = 0$ и $u = v$. Следовательно, φ унивалентно. Так как из 2) следует 1), то $\operatorname{Im}\varphi = V$, а это и означает, что φ – отображение *на*.

3) $\Rightarrow$ 4). Так как для любого вектора $v \in V$ существует единственный вектор $u \in V$ такой, что $u\varphi = v$, то в качестве φ^{-1} возьмем отображение, которое переводит вектор v в вектор u, т. е. $v\varphi^{-1} = u$. Проверим, что справедливы равенства $\varphi\varphi^{-1} = \varphi^{-1}\varphi = \varepsilon$, где ε – тождественное отображение. Действительно,

$$u(\varphi\varphi^{-1}) = (u\varphi)\varphi^{-1} = v\varphi^{-1} = u$$

и аналогично

$$v\varphi^{-1}\varphi = u\varphi = v.$$

Следовательно, φ^{-1} действительно является обратным отображением. Чтобы показать, что оно линейно, возьмем v_1 и v_2 из V такие, что $u_1\varphi = v_1$ и $u_2\varphi = v_2$. Так как

$$(\alpha_1 u_1 + \alpha_2 u_2)\varphi = \alpha_1 v_1 + \alpha_2 v_2,$$

то это и показывает, что φ^{-1} линейно.

4) $\Rightarrow$ 5). Из равенства $\varphi\varphi^{-1} = \varepsilon$ следует равенство матриц $[\varphi\varphi^{-1}] = [\varepsilon] = E$. В свою очередь, $[\varphi\varphi^{-1}] = [\varphi][\varphi^{-1}]$, а потому $[\varphi]$ – невырождено и $\det[\varphi] \cdot \det[\varphi^{-1}] \neq 0$.

5) $\Rightarrow$ 6). Очевидно.

6) $\Rightarrow$ 1). Пусть $e_1, e_2, \ldots, e_n$ – такая база пространства V, в которой матрица $[\varphi]$ невырождена. Возьмем вектор $v \in V$ и найдем его координаты $[v]$ в базе e. Рассмотрим систему $X[\varphi] = [v]$ относительно неизвестных $X = (x_1, x_2, \ldots, x_n)$. Эта система имеет единственное решение X^0. Положим $u = X^0 e$, тогда

$$u\varphi = X^0(e\varphi) = X^0([\varphi]e) = [v]e = v.$$

Таким образом, мы показали, что для любого вектора v найдется вектор $u \in V$ такой, что $u\varphi = v$, но это и означает, что $\operatorname{Im}\varphi = V$. Теорема доказана.

§ 32. Инвариантные подпространства

32.1. Инвариантные подпространства и неприводимость. Пусть U – подпространство векторного пространства V, а φ – линейное преобразование пространства V.

О п р е д е л е н и е. Говорим, что подпространство U *инвариантно относительно* φ, если $U\varphi \subseteq U$.

П р и м е р. Пусть $V = P[x]$ – пространство всех многочленов над полем P, $U = P_{\leq 3}[x]$ – подпространство многочленов степени не выше 3-й, φ – дифференцирование. Тогда $U\varphi \subseteq U$, т. е. U инвариантно относительно φ.

О п р е д е л е н и е. Пространство V называется *неприводимым относительно* φ, если V не содержит собственных (отличных от нулевого и всего V) инвариантных относительно φ подпространств.

О п р е д е л е н и е. Матрица вида

$$\left(\begin{array}{c|c} A & * \\ \hline \mathbf{0} & B \end{array}\right), \quad A \in M_r(P), \quad B \in M_s(P)$$

называется *полураспавшейся*.

Т е о р е м а 1. *Для линейного преобразования φ векторного пространства V равносильны следующие утверждения*:

1) *V содержит собственное, инвариантное относительно φ подпространство*;

2) *в подходящей базе матрица $[\varphi]$ полураспавшаяся.*

Д о к а з а т е л ь с т в о. 1) $\Rightarrow$ 2). Пусть U – собственное инвариантное относительно φ подпространство, т. е. $0 < U < V$ и $U\varphi \subseteq U$. Тогда найдется такая база

$$e_1, e_2, \ldots, e_r, e_{r+1}, e_{r+2}, \ldots, e_n$$

пространства V, что векторы $e_{r+1}, e_{r+2}, \ldots, e_n$ образуют базу подпространства U. Легко видеть, что в этой базе

$$[\varphi] = \left(\begin{array}{c|c} * & * \\ \hline \mathbf{0} & * \end{array}\right),$$

т. е. является полураспавшейся.

2) $\Rightarrow$ 1). Пусть в некоторой базе $e_1, e_2, \ldots, e_n$ матрица $[\varphi]$ является полураспавшейся:

$$[\varphi] = \left(\begin{array}{c|c} * & * \\ \hline \mathbf{0} & * \end{array}\right).$$

Возьмем в качестве U подпространство, натянутое на векторы e_{r+1}, e_{r+2}, $\ldots$, e_n. Ясно, что U является собственным подпространством и $U\varphi \subseteq U$. Теорема доказана.

32.2. Индуцированные преобразования. Пусть U – инвариантное подпространство относительно преобразования φ векторного пространства V. Определим преобразование

$$\varphi_U : U \longrightarrow U$$

по правилу

$$\varphi_U : u \longmapsto u\varphi$$

и будем называть *индуцированным преобразованием подпространства* U.

Рассмотрим фактор-пространство V/U пространства V по U и определим преобразование

$$\overline{\varphi}_U : V/U \longrightarrow V/U$$

по правилу

$$\overline{\varphi}_U : v + U \longmapsto v\varphi + U, \quad v \in V.$$

Покажем, что это определение не зависит от случайного выбора представителей смежных классов. Пусть $v_1 + U = v_2 + U$, т. е. $v_1 - v_2 \in U$. Тогда $v_1\varphi - v_2\varphi \in U$, т. е. $v_1\varphi + U = v_2\varphi + U$, а потому определение $\overline{\varphi}_U$ определено корректно.

Преобразование $\overline{\varphi}_U$ называется преобразованием, *индуцированным* φ в фактор-пространстве. Легко проверить, что преобразования φ_U и $\overline{\varphi}_U$ являются линейными. Связь трех преобразований φ, φ_U и $\overline{\varphi}_U$ дает

Т е о р е м а 2. *В согласованных базах матрица преобразования* φ *имеет такой вид*:

$$[\varphi] = \left(\begin{array}{c|c} [\overline{\varphi}_U] & * \\ \hline \mathbf{0} & [\varphi_U] \end{array}\right).$$

Д о к а з а т е л ь с т в о. Мы знаем, что справедлива формула

$$\dim V/U = \dim V - \dim U.$$

Укажем согласованные базы, в которых матрица имеет нужный вид. Пусть

$$e_1, e_2, \ldots, e_r, e_{r+1}, e_{r+2}, \ldots, e_n$$

– такая база векторного пространства V, что векторы $e_{r+1}, e_{r+2}, \ldots,$ e_n образуют базу подпространства U. Как мы знаем, смежные классы

$$e_1 + U, e_2 + U, \ldots, e_r + U$$

образуют базу фактор-пространства V/U. Действуя, на базисные векторы пространства V преобразованием φ, получим:

$$\begin{cases} e_1\varphi &= \alpha_{11}e_1 + \ldots + \alpha_{1r}e_r + \alpha_{1,r+1}e_{r+1} + \ldots + \alpha_{1n}e_n, \\ e_2\varphi &= \alpha_{21}e_1 + \ldots + \alpha_{2r}e_r + \alpha_{2,r+1}e_{r+1} + \ldots + \alpha_{2n}e_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ e_r\varphi &= \alpha_{r1}e_1 + \ldots + \alpha_{rr}e_r + \alpha_{r,r+1}e_{r+1} + \ldots + \alpha_{rn}e_n, \\ e_{r+1}\varphi &= 0 + \ldots + 0 + \alpha_{r+1,r+1}e_{r+1} + \ldots + \alpha_{r+1,n}e_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ e_n\varphi &= 0 + \ldots + 0 + \alpha_{n,r+1}e_{r+1} + \ldots + \alpha_{nn}e_n. \end{cases}$$

По определению

$$(e_i + U)\overline{\varphi}_U = e_i\varphi + U, \quad i = 1, 2, \ldots, r,$$

а потому

$$(e_i + U)\overline{\varphi}_U = \alpha_{i1}(e_1 + U) + \alpha_{i2}(e_2 + U) + \ldots \alpha_{ir}(e_r + U).$$

Таким образом,

$$\begin{array}{l} (e_1 + U)\overline{\varphi}_U = e_1\varphi + U = \alpha_{11}(e_1 + U) + \alpha_{12}(e_2 + U) + \ldots \alpha_{1r}(e_r + U), \\ (e_2 + U)\overline{\varphi}_U = e_2\varphi + U = \alpha_{21}(e_1 + U) + \alpha_{22}(e_2 + U) + \ldots \alpha_{2r}(e_r + U), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ (e_r + U)\overline{\varphi}_U = e_r\varphi + U = \alpha_{r1}(e_1 + U) + \alpha_{r2}(e_2 + U) + \ldots \alpha_{rr}(e_r + U). \end{array}$$

Следовательно, матрица $[\varphi]$ действительно имеет указанный вид. Теорема доказана.

32.3. Одномерные инвариантные подпространства, собственные векторы и собственные значения. Предположим, что подпространство U, инвариантное относительно φ, имеет размерность 1. Тогда U порождено любым ненулевым вектором $u \in U$ и

$$u\varphi = \alpha u, \quad \alpha \in P.$$

Поиск одномерных инвариантных подпространств сводится к отысканию таких u и α. При этом u называется *собственным вектором преобразования* φ, а α – его *собственным значением.*

Пусть $e_1, e_2, \ldots, e_n$ – база V, $[u]$ – координаты вектора u, а $[\varphi]$ – матрица преобразования φ в этой базе. Тогда

$$[u\varphi] = [\alpha u],$$

или

$$[u][\varphi] = \alpha[u].$$

Вынося координатную строку $[u]$, получим

$$[u]([\varphi] - \alpha E) = 0.$$

Чтобы существовал ненулевой вектор u, удовлетворяющий этому равенству, необходимо и достаточно, чтобы определитель

$$\det([\varphi] - \alpha E)$$

был равен нулю. Рассмотрим определитель $\chi(\lambda) = \det([\varphi] - \lambda E)$. Он является многочленом степени n относительно λ и собственные значения должны быть его корнями. Этот многочлен называется *характеристическим многочленом преобразования* φ. Справедлива

Т е о р е м а 3. а) *Характеристический многочлен не зависит от выбора базы, по которой он построен.* б) *Всякое собственное значение преобразования φ является корнем характеристического многочлена.* в) *Всякий корень характеристического многочлена, лежащий в основном поле, является собственным значением.*

Д о к а з а т е л ь с т в о. а) Пусть $e'_1, e'_2, \ldots, e'_n$ – другая база V. Тогда $e' = Te$, где T – матрица перехода от базы e к базе e'. Обозначим $[\varphi]'$ матрицу преобразования φ в базе e'. Как мы знаем,

$$[\varphi]' = T[\varphi]T^{-1}.$$

Поэтому

$$\det([\varphi]' - \lambda E) = \det(T[\varphi]T^{-1} - \lambda E) = \det(T\{[\varphi] - \lambda E\}T^{-1}) =$$

$$= \det T \cdot \det([\varphi] - \lambda E) \cdot \det T^{-1} = \det([\varphi] - \lambda E).$$

Следовательно, характеристический многочлен зависит от преобразования φ, а не от вида матрицы, соответствующей φ.

б) Это доказательство проведено выше. Каждое собственное значение обязано быть корнем характеристического многочлена.

в) Пусть $\chi(\alpha) = 0$, т. е. α – корень характеристического многочлена, лежащий в P. Надо доказать, что α – собственное значение.

Возьмем какую-нибудь базу $e_1, e_2, \ldots, e_n$ пространства V. Рассмотрим систему линейных уравнений

$$X([\varphi] - \alpha E) = 0, \quad X = (x_1, x_2, \ldots, x_n).$$

Так как ранг этой системы меньше числа неизвестных, существует ненулевое решение $X^0 = (x_1^0, x_2^0, \ldots, x_n^0)$. Рассмотрим вектор

$$u = x_1^0 e_1 + x_2^0 e_2 + \ldots + x_n^0 e_n.$$

Для него

$$[u]([\varphi] - \alpha E) = 0,$$

а потому $u(\varphi - \alpha\varepsilon) = 0$ или $u\varphi = \alpha u$. Значит, $u \neq 0$ – собственный вектор, а α – собственное значение. Теорема доказана.

У п р а ж н е н и е. Где нужно, чтобы α лежало в основном поле?

32.4. Теорема Гамильтона – Кэли.

Т е о р е м а 4. *Всякое линейное преобразование является корнем своего характеристического многочлена, т. е. если $\chi(\lambda)$ – характеристический многочлен преобразования φ, то $\chi(\varphi) = 0$.*

Д о к а з а т е л ь с т в о. Достаточно доказать эту теорему для матриц, так как алгебра линейных преобразований изоморфна алгебре матриц. Пусть $A \in M_n(P)$ и

$$\chi(\lambda) = \det(A - \lambda E)$$

– ее характеристический многочлен. Надо доказать, что $\chi(A) = 0$. Пусть

$$\chi(\lambda) = a_0 + a_1\lambda + \ldots + a_n\lambda^n.$$

Символом $B(\lambda)$ обозначим матрицу, присоединенную к матрице $A - \lambda E$ (напомним, что присоединенной к матрице $C = (c_{ij})$ называется матрица $(C_{ij})^t$, составленная из алгебраических дополнений C_{ij} к элементам c_{ij} и транспонированная). Тогда

$$(A - \lambda E)B(\lambda) = \chi(\lambda)E.$$

П р и м е р. Если дана матрица

$$\begin{pmatrix} 7 - 3\lambda^2 + \lambda^3 & 1 - \lambda \\ -8 + \lambda & -3 + \lambda^2 \end{pmatrix},$$

то ее можно представить в таком виде:

$$\begin{pmatrix} 7 - 3\lambda^2 + \lambda^3 & 1 - \lambda \\ -8 + \lambda & \lambda^2 - 3 \end{pmatrix} = \begin{pmatrix} 7 & 1 \\ -8 & -3 \end{pmatrix} + \lambda \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} +$$

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

Раскладывая матрицу $B(\lambda)$ по степеням λ, получим

$$B(\lambda) = B_0 + B_1\lambda + B_2\lambda^2 + \ldots + B_{n-1}\lambda^{n-1}, \quad B_k \in M_n(K).$$

Из равенства $(A-\lambda E)B(\lambda) = \chi(\lambda)E$, приравнивая матрицы при одинаковых степенях λ, получаем следующую систему равенств:

$$\begin{array}{ll} AB_0 & = a_0E, \\ AB_1 - B_0 & = a_1E, \\ AB_2 - B_1 & = a_2E, \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ AB_{n-1} - B_{n-2} & = a_{n-1}E, \\ -B_{n-1} & = a_nE. \end{array}$$

Умножим первое из этих равенств на E, второе – на A, третье – на A^2 и т. д. Наконец, последнее равенство умножим на A^n. Складывая полученные равенства, получим в левой части 0, а в правой

$$a_0E + a_1A + a_2A^2 + \ldots + a_nA^n = \chi(A).$$

Следовательно,

$$0 = \chi(A).$$

Теорема доказана.

У п р а ж н е н и е. Можно предложить такое доказательство теоремы Гамильтона – Кэли. В равенство

$$\chi(\lambda) = \det(A - \lambda E)$$

вместо λ подставим матрицу A, получим

$$\chi(A) = \det(A - AE) = 0.$$

Найдите ошибку в этом рассуждении.

§ 33. Нильпотентные и полупростые преобразования

33.1. Нильпотентные преобразования. Следующая теорема является определением и устанавливает некоторые свойства нильпотентных преобразований.

Т е о р е м а 1. *Для линейного преобразования φ следующие условия равносильны*:

1) $\varphi^N = 0$ *при подходящем* $N = N(\varphi) \in \mathbb{N}$;

2) *в подходящей базе*

$$[\varphi] = \begin{pmatrix} 0 & & * \\ & \ddots & \\ \mathbf{0} & & 0 \end{pmatrix}.$$

При любом из этих условий преобразование φ называется *нильпотентным.*

Д о к а з а т е л ь с т в о. 1) $\Rightarrow$ 2). Индукция по $n = \dim V$. При $n = 1$ очевидно, что $\varphi = 0$. Пусть далее $n > 1$. Если $\varphi = 0$, то утверждение очевидно. Пусть $\varphi \neq 0$, выберем N такое, что $\varphi^N = 0$, но $\varphi^{N-1} \neq 0$. Значит, существует вектор $u \in V$ такой, что $u\varphi^{N-1} \neq 0$. Обозначим $v = u\varphi^{N-1}$. Понятно, что $v \neq 0$ и $v\varphi = 0$. Пусть e_1, e_2, $\ldots$, e_{n-1}, v – база V и W – подпространство, порожденное вектором v. Тогда в этой базе

$$[\varphi] = \left(\begin{array}{c|c} & * \\ [\overline{\varphi}_W] & \vdots \\ & * \\ \hline 0 \ldots 0 & 0 \end{array}\right).$$

По определению индуцированного преобразования

$$\overline{\varphi}_W : a + W \longmapsto a\varphi + W.$$

Применить это преобразование N раз это все равно, что к a применить преобразование φ, т. е.

$$(a + W)\overline{\varphi}_W^N = a\varphi^N + W = W.$$

Следовательно, $\overline{\varphi}_W$ удовлетворяет условию 1, т. е. любой класс переходит под действием $\overline{\varphi}_W^N$ в 0.

Учитывая, что $\dim V/W = n - 1$, из предположения индукции заключаем, что в подходящей базе индуцированное преобразование имеет матрицу

$$[\overline{\varphi}_W]' = \begin{pmatrix} 0 & & * \\ & \ddots & \\ \mathbf{0} & & 0 \end{pmatrix}.$$

Возьмем в смежных классах

$$e'_1 + W, e'_2 + W, \ldots, e'_{n-1} + W$$

этой базы по представителю: e'_1, e'_2, $\ldots$, e'_{n-1}. Тогда в базе

$$e'_1, e'_2, \ldots, e'_{n-1}, v$$

имеем матрицу

$$[\varphi]' = \left(\begin{array}{ccc|c} 0 & & * & * \\ & \ddots & & \vdots \\ \mathbf{0} & & 0 & * \\ \hline 0 & \ldots & 0 & 0 \end{array}\right).$$

2) $\Rightarrow$ 1). Пусть в некоторой базе матрица преобразования φ имеет вид

$$A = \begin{pmatrix} 0 & & * \\ & \ddots & \\ \mathbf{0} & & 0 \end{pmatrix}.$$

Надо доказать, что в некоторой степени эта матрица равна нулевой матрице. Индукцией по i покажем, что

$$A^i = \begin{pmatrix} 0 & \ldots & 0 & & * \\ & \ddots & & \ddots & \\ & & \ddots & & 0 \\ & & & \ddots & \vdots \\ \mathbf{0} & & & & 0 \end{pmatrix},$$

т. е. матрица A^i имеет i нулевых диагоналей, начиная с главной. Для $i = 1$ утверждение справедливо, так как по условию матрица A имеет нулевую главную диагональ.

Предположим, что равенство справедливо для $i - 1$, и установим его для i. Обозначим символом E_{rs} *матричную единицу*, т. е. матрицу из $M_n(P)$, у которой на месте (r, s) стоит 1, а на всех остальных местах – нули. Тогда матрица A имеет вид

$$A = \sum_{q-p\geq 1} a_{pq} E_{pq}.$$

Заметим, что для матричных единиц справедливо следующее правило умножения:

$$E_{kl} \cdot E_{pq} = \begin{cases} E_{kq} & \text{при} \quad l = p, \\ 0 & \text{при} \quad l \neq p. \end{cases}$$

Воспользовавшись предположением индукции, получим

$$A^i = A^{i-1} \cdot A = \sum_{l-k \geq i-1} b_{kl} E_{kl} \cdot \sum_{q-p \geq 1} a_{pq} E_{pq} = \sum_{q-k \geq i} b_{kp}\, a_{pq} E_{kq}.$$

Эта матрица имеет i нулевых диагоналей.

Из доказанного равенства следует, что если положить $N \geq n$, то $A^N = 0$. Теорема доказана.

Непосредственно из доказательства получается такое

С л е д с т в и е. *В качестве* $N(\varphi)$ *можно взять* $n = \dim V$.

33.2. Полупростые преобразования.

Т е о р е м а 2. *Для линейного преобразования* φ *следующие условия равносильны*:

1) *все пространство* V *распадается в прямую сумму* $V = V_1 \oplus V_2 \oplus \ldots \oplus V_s$ *инвариантных и неприводимых относительно* φ *подпространств*;

2) *в подходящей базе*

$$[\varphi] = \begin{pmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_s \end{pmatrix},$$

где никакая матрица A_i *не подобна полураспавшейся матрице, т. е. ни в какой базе не станет полураспавшейся.* При любом из этих условий преобразование φ *называется полупростым.*

Д о к а з а т е л ь с т в о. 1) $\Rightarrow$ 2). Пусть $V = V_1 \oplus V_2 \oplus \ldots \oplus V_s$ – прямая сумма инвариантных и неприводимых относительно φ подпространств. Берем в каждом V_i по базе и объединяем их. В полученной базе

$$e_1, e_2, \ldots, e_n$$

матрица преобразования φ имеет вид

$$[\varphi] = \begin{pmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_s \end{pmatrix}.$$

При этом $V_i\varphi \subseteq V_i$. Нужно понять, что никакая клетка A_i не подобна полураспавшейся матрице, а это равносильно неприводимости подпространства V_i относительно φ.

2) $\Rightarrow$ 1). Пусть в базе

$$e_1, e_2, \ldots, e_n$$

матрица преобразования

$$[\varphi] = \begin{pmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_s \end{pmatrix},$$

где каждая клетка A_i не подобна полураспавшейся матрице. Разобьем базу пространства V на отрезки, стоящие против соответствующих клеток. Пусть V_i – подпространство, натянутое на i-й отрезок. Ясно, что $V = V_1 \oplus V_2 \oplus \ldots \oplus V_s$, $V_i\varphi \subseteq V_i$, и остается понять, что каждое подпространство V_i неприводимо. Если бы оно было приводимым, то нашлось бы собственное подпространство U такое, что

$$0 < U < V_i, \quad U\varphi \subseteq U,$$

и в этом случае матрица A_i была бы подобна полураспавшейся. Противоречие. Теорема доказана.

С л е д с т в и е. *Если φ – нильпотентно и полупросто, то $\varphi = 0$.*

Д о к а з а т е л ь с т в о. Если φ полупросто, то по теореме 2 в некоторой базе

$$[\varphi] = \begin{pmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_s \end{pmatrix},$$

а так как φ – нильпотентно, то каждая матрица A_i нильпотентна, а потому по теореме 1 подобна матрице такого вида:

$$\begin{pmatrix} 0 & & * \\ & \ddots & \\ \mathbf{0} & & 0 \end{pmatrix},$$

но из полупростоты следует, что A_i не подобна полураспавшейся. Следовательно, матрица A_i является квадратной степени 1 нулевой матрицей, но это и означает, что φ – нулевое преобразование.

33.3. Полупростота над полем, содержащим характеристические корни преобразования. В дальнейшем будем называть

корни характеристического многочлена *характеристическими корнями.* Справедлива

Т е о р е м а 3. *Пусть φ – линейное преобразование векторного пространства V над полем P, причем P содержит все характеристические корни преобразования φ. Тогда равносильны следующие утверждения*:

1) *φ полупросто*;

2) *в подходящей базе $[\varphi]$ диагональна*;

3) *существует многочлен $f \in P[\lambda]$, без кратных корней и такой, что $f(\varphi) = 0$.*

Д о к а з а т е л ь с т в о. 1) $\Rightarrow$ 2). Пусть φ полупросто, т. е. в подходящей базе

$$[\varphi] = \begin{pmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_s \end{pmatrix},$$

где никакая матрица A_i не подобна полураспавшейся матрице. Докажем, что все эти клетки одномерные. Допустим, что какая-то клетка имеет размерность ≥ 2, т. е. $\dim V_i \geq 2$ для соответствующего подпространства V_i. Выберем в V_i вектор v, являющийся собственным относительно φ, т. е. $v\varphi = \alpha v$. Очевидно, характеристический корень преобразования $\varphi|_{V_i}$ является характеристическим корнем преобразования φ, а потому α лежит в P. Пусть

$$e_1, e_2, \ldots, e_m, v$$

– база подпространства V_i. В этой базе

$$[\varphi|_{V_i}] = \left(\begin{array}{ccc|c} & & & * \\ & * & & \vdots \\ & & & * \\ \hline 0 & \ldots & 0 & \alpha \end{array}\right),$$

т. е. матрица A_i подобна полураспавшейся. Противоречие.

2) $\Rightarrow$ 3). Пусть в некоторой базе матрица $[\varphi]$ диагональна и многочлен $F(\lambda) \in P[\lambda]$ аннулирует преобразование φ, т. е. $F(\varphi) = 0$. В частности, в качестве $F(\lambda)$ можно взять характеристический многочлен преобразования φ. Разложим его на линейные множители:

$$F(\lambda) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \ldots (\lambda - \lambda_s)^{r_s},$$

где все корни λ_j различны. Возьмем многочлен

$$f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\dots(\lambda - \lambda_s)$$

и число

$$r = \max\{r_1, r_2, \dots, r_s\}.$$

Тогда $f(\lambda)^r = F(\lambda)d(\lambda)$ для некоторого многочлена $d(\lambda)$ и матрица $f([\varphi])$ диагональна. Рассмотрим матрицу $[f(\varphi)] = f([\varphi])$. Если ее возвести в степень r, то получим нулевую матрицу, но это значит, что на диагонали стояли нули, т. е. $f(\varphi) = 0$.

3) $\Rightarrow$ 1). Пусть $f(\varphi) = 0$ и

$$f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\dots(\lambda - \lambda_s),$$

где все λ_j различны. Надо доказать, что φ полупросто. Рассмотрим многочлены

$$f_i(\lambda) = \frac{f(\lambda)}{\lambda - \lambda_i}, \quad i = 1, 2, \dots, s.$$

Они в совокупности взаимно просты, а потому существуют многочлены $g_i \in P[\lambda]$ такие, что

$$g_1 f_1 + g_2 f_2 + \dots + g_s f_s = 1.$$

Обозначим

$$V_i = V f_i(\varphi), \quad i = 1, 2, \dots, s$$

и докажем, что V является прямой суммой подпространств V_i.

а) Докажем вначале, что $V = V_1 + V_2 + \dots + V_s$. Действительно, пусть $v \in V$, тогда из равенства

$$g_1(\varphi)f_1(\varphi) + g_2(\varphi)f_2(\varphi) + \dots + g_s(\varphi)f_s(\varphi) = \varepsilon,$$

где ε – тождественное преобразование, получим

$$v = v\varepsilon = vg_1(\varphi)f_1(\varphi) + vg_2(\varphi)f_2(\varphi) + \dots + vg_s(\varphi)f_s(\varphi).$$

Заметим, что первое слагаемое лежит в V_1, второе – в V_2 и т. д. и, наконец, последнее лежит в V_s. Следовательно, $V = V_1 + V_2 + \dots + V_s$.

б) Докажем, что сумма $V = V_1 + V_2 + \dots + V_s$ является прямой. Для этого достаточно доказать, что нулевой вектор имеет единственное разложение. Пусть

$$0 = u_1 f_1(\varphi) + u_2 f_2(\varphi) + \dots + u_s f_s(\varphi).$$

Покажем, что все $u_i f_i(\varphi) = 0$. Для этого к обеим частям равенства применим преобразование $f_i(\varphi)$. Заметим, что если $k \neq i$, то

$$u_k f_k(\varphi) f_i(\varphi) = u_k f(\varphi) h_{ki}(\varphi) = 0 \quad \text{для некоторых} \quad h_{ki}(\lambda) \in P[\lambda],$$

и мы имеем равенство

$$0 = u_i f_i^2(\varphi).$$

Так как $f_i(\lambda)$ и $\lambda - \lambda_i$ взаимно просты, то существуют многочлены p_i и q_i из $P[\lambda]$ такие, что

$$f_i(\lambda) p_i(\lambda) + (\lambda - \lambda_i) q_i(\lambda) = 1.$$

Отсюда

$$f_i(\varphi) p_i(\varphi) + (\varphi - \lambda_i \varepsilon) q_i(\varphi) = \varepsilon$$

и

$$u_i f_i(\varphi) = u_i f_i(\varphi)\varepsilon = u_i f_i^2(\varphi) p_i(\varphi) + u_i f(\varphi) q_i(\varphi) = 0,$$

так как $u_i f_i(\varphi)^2 = 0$ и $u_i f(\varphi) = 0$, т. е. эта сумма $V = V_1 + V_2 + \ldots + V_s$ является прямой.

в) Чтобы показать, что каждое подпространство V_i инвариантно относительно φ, заметим, что $V_i(\varphi - \lambda_i \varepsilon) = 0$. Действительно, произвольный вектор из V_i имеет вид $u f_i(\varphi)$ для некоторого $u \in V$. Действуя преобразованием $\varphi - \lambda_i \varepsilon$, получим

$$(u f_i(\varphi))(\varphi - \lambda_i \varepsilon) = u f(\varphi) = u0 = 0.$$

Следовательно, если $u_i \in V_i$, то $u_i \varphi = \lambda_i u_i$, а потому V_i инвариантно относительно φ.

Ввиду того, что преобразование φ_{V_i} является умножением на λ_i раз, то, выбирая базу в каждом V_i и объединяя их, получим базу пространства V, в которой матрица преобразования φ имеет вид

$$[\varphi] = \left(\begin{array}{ccc|c|cc} \lambda_1 & & \mathbf{0} & & & \\ & \ddots & & & \mathbf{0} & \\ \mathbf{0} & & \lambda_1 & & & \\ \hline & \mathbf{0} & & \ddots & \mathbf{0} & \\ \hline & & & & \lambda_s & \mathbf{0} \\ & \mathbf{0} & & & \ddots & \\ & & & & \mathbf{0} & \lambda_s \end{array}\right),$$

т. е. является диагональной матрицей, а потому φ полупросто. Теорема доказана.

Заметим, что диагональная матрица всегда полупроста, но обратное верно не всегда.

У п р а ж н е н и е. Привести пример полупростой матрицы, не представимой в диагональном виде.

33.4. Разложение преобразования на полупростую и нильпотентную компоненты.

Т е о р е м а 4. *Пусть φ – линейное преобразование векторного пространства V над полем P и P содержит все характеристические корни φ. Тогда*

а) *существует разложение*

$$\varphi = \psi + \theta,$$

где ψ – полупростое преобразование, а θ – нильпотентное, при этом $\psi\theta = \theta\psi$;

б) *такое разложение единственно*;

в) *ψ и θ являются многочленами от φ над полем P.*

При этом ψ называется *полупростой компонентой*, а θ – *нильпотентной*.

Д о к а з а т е л ь с т в о. Докажем существование. Покажем, что найдется многочлен $f(\lambda)$ из $P[\lambda]$ без кратных корней и $m \in \mathbb{N}$ такие, что $f^m(\varphi) = 0$. Пусть $F(\varphi) = 0$ для некоторого многочлена $F(\lambda) \in P[\lambda]$ (по теореме Гамильтона – Кэли такой многочлен существует). Разложим его на линейные множители:

$$F(\lambda) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \dots (\lambda - \lambda_s)^{r_s}, \quad \lambda_i \neq \lambda_j \text{ при } i \neq j.$$

Возьмем

$$f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_s).$$

Тогда $f^m(\varphi) = 0$ при $m = \max\{r_1, r_2, \dots, r_s\}$. Так как f без кратных корней, то $(f, f') = 1$, а потому существуют $u, v \in P[\lambda]$ такие, что

$$f\,u + f'\,v = 1.$$

Рассмотрим рекуррентную схему, позволяющую по каждому многочлену $g(\lambda)$ из $P[\lambda]$ построить последовательность многочленов $g^{(0)}(\lambda)$, $g^{(1)}(\lambda)$, ... из $P[\lambda]$:

$$\begin{cases} g^{(0)}(\lambda) = g(\lambda), \\ g^{(i+1)}(\lambda) = g^{(i)}(\lambda - f(\lambda)\,v(\lambda)), \quad i = 0, 1, \dots, \end{cases}$$

и применим ее к многочлену $e(\lambda) = \lambda$. Получим последовательность

$$e^{(0)}(\lambda), e^{(1)}(\lambda), \ldots, e^{(k)}(\lambda),$$

где k – наименьшее натуральное число, удовлетворяющее неравенству $2^k \geq m$. Положим $\psi = e^{(k)}(\varphi)$ и $\theta = \varphi - \psi$. Покажем, что эти преобразования и будут искомыми.

По теореме 3 преобразование полупросто, если оно является корнем многочлена без кратных корней. Рассмотрим

$$f(\psi) = f(e^{(k)}(\varphi)).$$

Если мы докажем, что $f(e^{(k)}(\lambda)) = f^{(k)}(\lambda)$ и $f^{(k)}(\lambda)$ делится на $f^{2^k}(\lambda)$, то, учитывая, что $f^{2^k}(\varphi) = 0$, получим, что $f^{(k)}(\varphi) = 0$. Таким образом, полупростота преобразования ψ будет доказана.

Чтобы доказать, что θ нильпотентно, надо доказать, что $\lambda - e^{(k)}(\lambda)$ делится на $f(\lambda)$.

Л е м м а 1. *Справедливо равенство*

$$f(e^{(i)}(\lambda)) = f^{(i)}(\lambda).$$

Д о к а з а т е л ь с т в о проведем индукцией по i. При $i = 0$ имеем

$$f(e^{(0)}(\lambda)) = f(\lambda) = f^{(0)}(\lambda).$$

Предположим, что утверждение справедливо для i, и докажем его для $i + 1$. Имеем

$$f(e^{(i+1)}(\lambda)) = f(e^{(i)}(\lambda - f(\lambda)\, v(\lambda))) = f^{(i)}(\lambda - f(\lambda)\, v(\lambda)) = f^{(i+1)}(\lambda).$$

Лемма доказана.

Л е м м а 2. $f^{(i)}(\lambda)$ *кратно* $f^{2^i}(\lambda)$.

Д о к а з а т е л ь с т в о проведем индукцией по i. При $i = 0$ видим, что $f^{(0)}(\lambda)$ кратно $f^{2^0}(\lambda) = f(\lambda)$. При $i = 1$ имеем

$$f^{(1)}(\lambda) = f(\lambda - f(\lambda)\, v(\lambda)).$$

Воспользовавшись формулой Тейлора

$$f(\lambda - a) = f(\lambda) - \frac{f'(\lambda)}{1!}a + \frac{f''(\lambda)}{2!}a^2 - \ldots,$$

получим

$$f(\lambda - f(\lambda)\, v(\lambda)) = f(\lambda) - \frac{f'(\lambda)}{1!}f(\lambda)\, v(\lambda) + \frac{f''(\lambda)}{2!}f^2(\lambda)v^2(\lambda) - \ldots =$$

$$= f(\lambda)\,[1 - f'(\lambda)\,v(\lambda) + f(\lambda)\,[\ldots]] = f(\lambda)\,[f(\lambda)u(\lambda) + f(\lambda)\,[\ldots]] =$$
$$= f^2(\lambda)\,[\ldots].$$

Предположим, что формула верна для i, и докажем ее для $i+1$. Имеем

$$f^{(i+1)}(\lambda) = f^{(i)}(\lambda - f(\lambda)\,v(\lambda)) = f^{2^i}(\lambda - f(\lambda)\,v(\lambda))\cdot[\ldots] = [f^{(1)}(\lambda)]^{2^i}[\ldots] =$$
$$= [f(\lambda)]^{2^{i+1}}[\ldots],$$

где мы использовали установленный выше факт, что $f^{(1)}(\lambda)$ делится на $f^2(\lambda)$. Лемма доказана.

Л е м м а 3. *Многочлен $\lambda - e^{(i)}(\lambda)$ делится на многочлен $f(\lambda)$.*

Д о к а з а т е л ь с т в о проведем индукцией по i. При $i = 0$ видим, что $\lambda - e^{(0)}(\lambda) = 0$, а 0 кратен $f(\lambda)$, так как $0 = f(\lambda)\cdot 0$.

Предположим, что лемма справедлива для i. Рассмотрим

$$\lambda - e^{(i+1)}(\lambda) = \lambda - e^{(i)}(\lambda - f(\lambda)\,v(\lambda)) - f(\lambda)\,v(\lambda) + f(\lambda)\,v(\lambda).$$

Так как $\mu - e^{(i)}(\mu)$ кратно $f(\mu)$, где $\mu = \lambda - f(\lambda)\,v(\lambda)$, то

$$\lambda - e^{(i+1)}(\lambda) = f(\lambda - f(\lambda)\,v(\lambda))\cdot[\ldots] + f(\lambda)\,v(\lambda).$$

В предыдущей лемме было установлено, что $f(\lambda - f(\lambda)\,v(\lambda))$ делится на $f(\lambda)$. Следовательно, $\lambda - e^{(i+1)}(\lambda)$ делится на $f(\lambda)$. Лемма доказана.

Мы установили, что преобразование ψ полупросто, а θ – нильпотентно. Как следует из построения, ψ и θ являются многочленами от φ, а потому $\psi\theta = \theta\psi$. Следовательно, утверждения а и в теоремы установлены.

Установим пункт б. Допустим, что найдется еще одна пара ψ', θ' преобразований, таких, что $\varphi = \psi' + \theta'$, и при этом ψ' полупросто, а θ' нильпотентно. Рассматривая разность двух равенств

$$\varphi = \psi + \theta, \quad \varphi = \psi' + \theta',$$

получим

$$\psi - \psi' = \theta' - \theta.$$

У п р а ж н е н и е. Привести пример двух полупростых преобразований, разность которых не является полупростым. Привести пример двух нильпотентных преобразований, разность которых не является нильпотентным.

Л е м м а 4. *Преобразования ψ и ψ' перестановочны, преобразования θ и θ' перестановочны.*

Д о к а з а т е л ь с т в о. Очевидно, что ψ' перестановочно с ψ'. По условию ψ' перестановочно с θ'. Следовательно, ψ' перестановочно с $\psi' + \theta' = \varphi$, т. е. ψ' перестановочно с φ, а потому перестановочно с любым многочленом от φ, но среди таких многочленов есть многочлен, равный ψ. Следовательно, ψ' перестановочно с ψ. Аналогично проверяется, что θ' перестановочно с θ.

Л е м м а 5. *Если преобразования ψ_1 и ψ_2 полупросты и перестановочны, то $\psi_1 \pm \psi_2$ полупросто.*

Д о к а з а т е л ь с т в о. По теореме 3 в некоторых базах матрицы преобразований ψ_1 и ψ_2 диагональны, но эти базы, вообще говоря, различны. Возьмем базу, в которой $[\psi_1]$ диагональна:

$$[\psi_1] = \left(\begin{array}{ccc|ccc|c|ccc} \alpha & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & \alpha & & & & & & & \\ \hline & & & \beta & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & \beta & & & & \\ \hline & & & & & & \ddots & & & \\ \hline & & & & & & & \omega & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \omega \end{array}\right),$$

и разобьем матрицу на клетки так, что в каждой клетке стоят одинаковые диагональные элементы. Пусть в этой же базе

$$[\psi_2] = \left(\begin{array}{c|c|c|c} X_{11} & X_{12} & \cdots & X_{1s} \\ \hline X_{21} & X_{22} & \cdots & X_{2s} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline X_{s1} & X_{s2} & \cdots & X_{ss} \end{array}\right),$$

где матрица разбита на клетки тех же размеров, что и $[\psi_1]$. Так как преобразования перестановочны, то

$$[\psi_1][\psi_2] = \left(\begin{array}{c|c|c|c} \alpha X_{11} & \alpha X_{12} & \cdots & \alpha X_{1s} \\ \hline \beta X_{21} & \beta X_{22} & \cdots & \beta X_{2s} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline \omega X_{s1} & \omega X_{s2} & \cdots & \omega X_{ss} \end{array}\right),$$

$$[\psi_2][\psi_1] = \left(\begin{array}{c|c|c|c} X_{11}\alpha & X_{12}\beta & \cdots & X_{1s}\omega \\ \hline X_{21}\alpha & X_{22}\beta & \cdots & X_{2s}\omega \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline X_{s1}\alpha & X_{s2}\beta & \cdots & X_{ss}\omega \end{array}\right).$$

Следовательно,

$$\left\{\begin{array}{llll} \alpha X_{11} = X_{11}\alpha, & \alpha X_{12} = X_{12}\beta, & \cdots & \alpha X_{1s} = X_{1s}\omega, \\ \beta X_{21} = X_{21}\alpha, & \beta X_{22} = X_{22}\beta, & \cdots & \beta X_{2s} = X_{2s}\omega, \\ \cdots & \cdots & \cdots & \cdots \\ \omega X_{s1} = X_{s1}\alpha, & \omega X_{s2} = X_{s2}\beta, & \cdots & \omega X_{ss} = X_{ss}\omega. \end{array}\right.$$

Из этой системы видим, что все клетки вне главной диагонали нулевые, т. е.

$$[\psi_2] = \left(\begin{array}{c|c|c|c} X_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & X_{22} & \cdots & \mathbf{0} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & X_{ss} \end{array}\right),$$

и каждая из диагональных клеток полупроста. Если клеточно-диагональная матрица является корнем многочлена без кратных корней, то и каждая диагональная клетка является корнем этого многочлена. Поэтому найдется база, в которой каждая клетка матрицы $[\psi_2]'$ имеет диагональный вид:

$$[\psi_2]' = \left(\begin{array}{c|c|c|c} * & & & \\ & \ddots & & \\ & & * & \\ \hline & * & & \\ & & \ddots & \\ & & & * \\ \hline & & \ddots & \\ \hline & & & * \\ & & & & \ddots \\ & & & & & * \end{array}\right),$$

но матрица $[\psi_1]$ при этом не изменится, так как каждой клетке соответствует подпространство, состоящее из собственных векторов пре-

образования $[\psi_1]$, а потому

$$[\psi_1]' = \left(\begin{array}{c|c|c|c} \alpha & & & \\ \ddots & & & \\ \alpha & & & \\ \hline & \beta & & \\ & \ddots & & \\ & \beta & & \\ \hline & & \ddots & \\ \hline & & & \omega \\ & & & \ddots \\ & & & \omega \end{array}\right).$$

Следовательно, $[\psi_1 \pm \psi_2]'$ диагональна. Лемма доказана.

Л е м м а 6. *Если преобразования θ_1 и θ_2 нильпотентны и перестановочны, то $\theta_1 \pm \theta_2$ нильпотентно.*

Д о к а з а т е л ь с т в о. Пусть $\theta_1^{n_1} = 0$, $\theta_2^{n_2} = 0$, тогда, воспользовавшись перестановочностью θ_1 и θ_2, получим

$$(\theta_1 + \theta_2)^N = \sum_{k=0}^{N} C_N^k \theta_1^k \theta_2^{N-k} = 0 \text{ при } N = n_1 + n_2.$$

Теорема доказана.

33.5. Полупростота над полем действительных чисел. Над полем комплексных чисел полупростота равносильна диагонализируемости матрицы. Над полем действительных чисел это уже неверно.

Т е о р е м а 5. *Над полем действительных чисел линейное преобразование φ полупросто тогда и только тогда, когда в некоторой базе его матрица клеточно-диагональна и каждая диагональная клетка либо одномерна, либо двумерна и не подобна полураспавшейся.*

Д о к а з а т е л ь с т в о. По теореме о полупростом преобразовании в некоторой базе его матрица клеточно-диагональна. Покажем, что каждая клетка имеет порядок ≤ 2. Для этого достаточно доказать, что всякая $A \in M_n(\mathbb{R})$ при $n \geq 3$ подобна полураспавшейся. Для этого покажем, что при $n \geq$ есть инвариантные подпространства.

Возьмем характеристический многочлен матрицы A и найдем его характеристический корень. Если корень вещественный, то собствен-

ный вектор порождает инвариантное подпространство, а потому матрица A подобна полураспавшейся. Пусть характеристический корень комплексный: $\alpha + i\beta$, $\alpha, \beta \in \mathbb{R}$. Рассмотрим $\mathbb{C}^n$ – пространство строк над полем комплексных чисел. Тогда характеристический корень есть собственное значение над $\mathbb{C}$. Следовательно, существует w – собственный вектор, отвечающий собственному значению $\alpha + i\beta$ т. е.

$$wA = (\alpha + i\beta)w, \quad w \neq 0.$$

У вектора w можно отделить действительную и мнимую части:

$$w = u + iv, \quad u, v \in \mathbb{R}^n,$$

тогда

$$(u + iv)A = (\alpha + i\beta)(u + iv).$$

Рассмотрим подпространство U, натянутое на векторы u и v. Учитывая равенства

$$\begin{cases} uA = \alpha u - \beta v, \\ vA = \beta u + \alpha v, \end{cases}$$

заключаем, что $UA \subseteq U$, т. е. подпространство U инвариантно относительно A и U – не более чем двумерно. Следовательно, исходная матрица подобна полураспавшейся. Теорема доказана.

Глава 6

ЖОРДАНОВА ФОРМА МАТРИЦЫ

§ 34. Теорема Жордана

34.1. Задача о подобии матриц. Как мы знаем, матрицы одного преобразования в разных базах связаны равенством

$$[\varphi]' = T[\varphi]T^{-1},$$

где T – невырожденная матрица. Возникает задача: распознать, когда две матрицы являются матрицами одного преобразования, записанного в разных базах.

О п р е д е л е н и е. Говорим, что матрица $A \in M_n(P)$ *подобна* матрице $B \in M_n(P)$, если существует матрица $X \in GL_n(P)$ такая, что $A = X^{-1}BX$.

Заметим вначале, что отношение подобия является отношением эквивалентности.

1) A подобна A, что следует из равенства $A = E^{-1}AE$.

2) Если A подобна B, то B подобна A. Действительно, если $A = X^{-1}BX$, то $B = XAX^{-1}$.

3) Если A подобна B, а B подобна C, то A подобна C. Действительно, если $A = X^{-1}BX$ и $B = Y^{-1}CY$, то $A = (YX)^{-1}C(YX)$.

Следовательно, все множество матриц $M_n(P)$ распадается на смежные классы относительно отношения подобия. Мы хотим научиться выбирать в каждом смежном классе по представителю, а

все другие матрицы из этого класса приводить к выбранному представителю.

О п р е д е л е н и е. 1) *Жордановой клеткой* степени r называется матрица следующего вида:

$$\begin{pmatrix} \alpha & 1 & & & \mathbf{0} \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha \end{pmatrix} \in M_r(P).$$

2) *Жорданова матрица* – матрица, которая распадается на жордановы клетки:

$$\left(\begin{array}{c|c|c|c} J_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & J_2 & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & J_s \end{array}\right),$$

где $J_1, J_2, \ldots, J_s$ – жордановы клетки.

П р и м е р. При $n = 4$ имеем следующие жордановы матрицы:

$$\begin{pmatrix} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 1 & 0 \\ 0 & 0 & \alpha & 1 \\ 0 & 0 & 0 & \alpha \end{pmatrix}, \quad \left(\begin{array}{ccc|c} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 1 & 0 \\ 0 & 0 & \alpha & 0 \\ \hline 0 & 0 & 0 & \beta \end{array}\right), \quad \left(\begin{array}{cc|cc} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ \hline 0 & 0 & \beta & 1 \\ 0 & 0 & 0 & \beta \end{array}\right),$$

$$\left(\begin{array}{cc|c|c} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ \hline 0 & 0 & \beta & 0 \\ \hline 0 & 0 & 0 & \gamma \end{array}\right), \quad \left(\begin{array}{c|c|c|c} \alpha & 0 & 0 & 0 \\ \hline 0 & \beta & 0 & 0 \\ \hline 0 & 0 & \gamma & 0 \\ \hline 0 & 0 & 0 & \delta \end{array}\right).$$

В каждом смежном классе существует несколько жордановых матриц, но они отличаются лишь порядком клеток.

Целью настоящего параграфа является доказательство следующей теоремы.

Т е о р е м а Ж о р д а н а. *Всякая квадратная матрица над полем, содержащим все ее характеристические корни, подобна некоторой жордановой матрице. Эта жорданова матрица определяется однозначно с точностью до порядка жордановых клеток.*

Для доказательства теоремы нам надо найти базу, в которой матрица имеет жорданову форму. Построение такой базы разбивается на несколько этапов. Вначале найдем базу, в которой матрица распадается на клетки с одним и тем же характеристическим корнем α, а потом уже каждую из этих клеток разобьем на жордановы клетки.

34.2. Корневое разложение.

О п р е д е л е н и е. Пусть V – векторное пространство над полем P, φ – линейное преобразование V, α – его собственное значение. Вектор $v \in V$ называется *корневым относительно* φ, если $v(\varphi - \alpha\varepsilon)^\rho = 0$ при подходящем $\rho \in \{0, 1, 2, \ldots\}$. Наименьшее такое ρ называется *высотой* v относительно φ.

Если $\rho = 1$, то $v\varphi = \alpha v$, т. е. v – собственный вектор. Следовательно, собственный вектор – корневой вектор высоты 1.

Т е о р е м а 1. 1) *Множество всех корневых векторов относительно преобразования φ, отвечающих одному и тому же собственному значению α преобразования φ, является подпространством пространства V.* 2) *Это подпространство инвариантно относительно φ.* Оно называется *корневым подпространством.*

Д о к а з а т е л ь с т в о. 1) Пусть V_α – множество корневых векторов относительно φ, отвечающих собственному значению α. Пусть

$$v_1(\varphi - \alpha\varepsilon)^{\rho_1} = 0, \quad v_2(\varphi - \alpha\varepsilon)^{\rho_2} = 0$$

и $\beta_1, \beta_2 \in P$. Тогда

$$(\beta_1 v_1 + \beta_2 v_2)(\varphi - \alpha\varepsilon)^\rho = 0, \quad \text{где } \rho = \max\{\rho_1, \rho_2\}.$$

Следовательно, V_α является подпространством.

2) Докажем, что $V_\alpha\varphi \subseteq V_\alpha$. Надо доказать, что если $v_1 \in V_\alpha$, то вектор $v_1\varphi$ является корневым. Рассмотрим

$$(v_1\varphi)(\varphi - \alpha\varepsilon)^{\rho_1}$$

и, учитывая, что всякий многочлен от φ перестановочен с самим φ, получим

$$(v_1\varphi)(\varphi - \alpha\varepsilon)^{\rho_1} = v_1(\varphi - \alpha\varepsilon)^{\rho_1}\varphi = 0\varphi = 0.$$

Теорема доказана.

У п р а ж н е н и е. Высоты векторов из V_α не превосходят кратности корня α в характеристическом многочлене преобразования φ, т. е. если

$$\chi(\lambda) = (\lambda - \alpha)^k \chi_1(\lambda), \quad \chi_1(\alpha) \neq 0,$$

то для любого v из V_α имеем

$$v(\varphi - \alpha\varepsilon)^k = 0.$$

Т е о р е м а 2. *Пусть поле P содержит все характеристические корни: α_1, α_2, ..., α_s преобразования φ. Пусть V_1, V_2, ..., V_s – соответствующие корневые подпространства, отвечающие α_1, α_2, ..., α_s. Тогда $V = V_1 \oplus V_2 \oplus \ldots \oplus V_s$.* Это разложение называется *корневым разложением.*

Д о к а з а т е л ь с т в о. Пусть характеристический многочлен преобразования φ имеет вид

$$\chi(\lambda) = (\lambda - \alpha_1)^{r_1}(\lambda - \alpha_2)^{r_2} \ldots (\lambda - \alpha_s)^{r_s}, \quad \alpha_i \neq \alpha_j \ \text{ при } i \neq j.$$

Рассмотрим многочлены

$$\chi_i(\lambda) = \frac{\chi(\lambda)}{(\lambda - \alpha_i)^{r_i}}, \quad i = 1, 2, \ldots, s.$$

а) Покажем, что $V_i = V\chi_i(\varphi)$. Установим включение $\supseteq$. Рассмотрим

$$v\chi_i(\varphi)(\varphi - \alpha_i\varepsilon)^{r_i} = v\chi(\varphi).$$

По теореме Гамильтона – Кэли $\chi(\varphi) = 0$, а потому $v\chi_i(\varphi)$ – корневой вектор. Следовательно, включение $\supseteq$ установлено.

Обратно. Пусть $w \in V_i$ и $w(\varphi - \alpha_i\varepsilon)^\rho = 0$. Надо доказать, что $w \in V\chi_i(\varphi)$. Так как $(\lambda - \alpha_i)^\rho$ и $\chi_i(\lambda)$ взаимно просты, то существуют многочлены $p_i(\lambda)$, $q_i(\lambda)$ из $P[\lambda]$ такие, что

$$p_i(\lambda)(\lambda - \alpha_i)^\rho + q_i(\lambda)\chi_i(\lambda) = 1.$$

Тогда

$$p_i(\varphi)(\varphi - \alpha_i\varepsilon)^\rho + q_i(\varphi)\chi_i(\varphi) = \varepsilon,$$

и отсюда

$$w = w\varepsilon = w[p_i(\varphi)(\varphi - \alpha_i\varepsilon)^\rho + q_i(\varphi)\chi_i(\varphi)] = [wq_i(\varphi)]\chi_i(\varphi),$$

т. е. равенство $V_i = V\chi_i(\varphi)$ доказано.

б) Покажем, что $V = V_1 + V_2 + \ldots + V_s$. Пусть $v \in V$. Так как многочлены $\chi_1(\lambda)$, $\chi_2(\lambda)$, ..., $\chi_s(\lambda)$ взаимно просты, то существуют многочлены $h_1(\lambda)$, $h_2(\lambda)$, ..., $h_s(\lambda)$ такие, что

$$h_1(\lambda)\chi_1(\lambda) + h_2(\lambda)\chi_2(\lambda) + \ldots + h_s(\lambda)\chi_s(\lambda) = 1.$$

Тогда

$$h_1(\varphi)\chi_1(\varphi) + h_2(\varphi)\chi_2(\varphi) + \ldots + h_s(\varphi)\chi_s(\varphi) = \varepsilon.$$

Отсюда

$$v = v\varepsilon = (vh_1(\varphi))\chi_1(\varphi) + (vh_2(\varphi))\chi_2(\varphi) + \ldots + (vh_s(\varphi))\chi_s(\varphi).$$

Заметим, что первое слагаемое лежит в V_1, второе в V_2 и т. д. Следовательно, мы получили нужное разложение.

в) Покажем, что $V = V_1 \oplus V_2 \oplus \ldots \oplus V_s$. Для этого достаточно показать, что нулевой вектор имеет единственную запись. Пусть

$$0 = u_1\chi_1(\varphi) + u_2\chi_2(\varphi) + \ldots + u_s\chi_s(\varphi).$$

Надо доказать, что все $u_i\chi_i(\varphi) = 0$. Применяя к нашему равенству преобразование $\chi_i(\varphi)$, получим

$$0 = u_i\chi_i(\varphi)^2,$$

так как слагаемые $u_k\chi_k(\varphi)\chi_i(\varphi)$ при $k \neq i$ обращаются в нуль. Далее, учитывая, что многочлены $\chi_i(\lambda)$ и $(\lambda - \alpha_i)^{r_i}$ взаимно просты, найдем многочлены $f_i(\lambda)$, $g_i(\lambda)$ такие, что

$$\chi_i(\lambda)f_i(\lambda) + (\lambda - \alpha_i)^{r_i}g_i(\lambda) = 1.$$

Тогда

$$\chi_i(\varphi)f_i(\varphi) + (\varphi - \alpha_i\varepsilon)^{r_i}g_i(\varphi) = \varepsilon$$

и

$$u_i\chi_i(\varphi) = u_i\chi_i(\varphi)\varepsilon = u_i\chi_i(\varphi)[\chi_i(\varphi)f_i(\varphi) + (\varphi - \alpha_i\varepsilon)^{r_i}g_i(\varphi)] =$$

$$= (u_i\chi_i^2(\varphi))f_i(\varphi) + u_i(\chi_i(\varphi)(\varphi - \alpha_i\varepsilon)^{r_i})g_i(\varphi) = u_i\chi(\varphi)g_i(\varphi) = 0,$$

т. е. сумма действительно прямая. Теорема доказана.

Таким образом, мы можем разложить V в сумму инвариантных подпространств. Выбрав в каждом из подпространств базу и объединив их, найдем базу V. В этой базе матрица преобразования φ имеет вид

$$[\varphi] = \left(\begin{array}{c|c|c|c} * & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & * & \cdots & \mathbf{0} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & * \end{array}\right),$$

где каждой диагональной клетке соответствует единственное собственное значение.

34.3. Канонический вид нильпотентного преобразования. Ограничимся отдельным корневым подпространством. Всякий вектор из этого подпространства удовлетворяет равенству

$$v(\varphi - \alpha_i\varepsilon)^{r_i} = 0.$$

Следовательно, преобразование $\psi = \varphi - \alpha_i\varepsilon$ является нильпотентным преобразованием корневого подпространства V_i. Для этого преобразования мы должны научиться строить базу, в которой матрица распадается на жордановы клетки.

О п р е д е л е н и е. Пусть ψ – нильпотентное преобразование векторного пространства V. Назовем подпространство U *циклическим относительно* ψ, если оно обладает базой e_1, e_2, …, e_k такой, что

$$e_1\psi = e_2, e_2\psi = e_3, \ldots, e_k\psi = 0.$$

Сама эта база называется *циклической.* Заметим, что циклическое подпространство U инвариантно относительно ψ и в циклической базе матрица преобразования имеет вид

$$[\psi_U] = \begin{pmatrix} 0 & 1 & & & \mathbf{0} \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & 0 \end{pmatrix} \in M_k(P). \qquad (1)$$

Покажем, что для всякого нильпотентного преобразования существует база, в которой матрица преобразования является клеточно-диагональной, а каждая диагональная клетка имеет вид (1).

Т е о р е м а 3. *Пусть ψ – нильпотентное преобразование пространства V. Тогда* 1) *существует разложение V в прямую сумму циклических относительно ψ подпространств*; 2) *если z_i – количество i-мерных циклических подпространств в этом разложении, то оно вычисляется по формуле*

$$z_i = r_{i-1} - 2r_i + r_{i+1},$$

где $r_i = \operatorname{rank}(\psi^i) = \dim(V\psi^i)$.

Заметим, что из второй части этой теоремы вытекает, в частности, что какое бы разложение на циклические слагаемые не взяли, количество слагаемых будет одно и то же.

Д о к а з а т е л ь с т в о. 1) Пусть $\psi^p = 0$, $\psi^{p-1} \neq 0$. Надо построить базу, которая состоит из циклических отрезков. Рассмотрим

$$W_i = \operatorname{Ker} \psi^i, \quad i = 0, 1, 2, \ldots, p.$$

Понятно, что все ядра вложены друг в друга:

$$0 = W_0 \subseteq W_1 \subseteq \ldots \subseteq W_{p-1} \subseteq W_p = V.$$

Действительно, если $v\psi^{i-1} = 0$, то $v\psi^i = 0$. Будем двигаться по цепочке справа налево. Рассмотрим фактор-пространство W_i/W_{i-1}. Векторами в нем являются смежные классы. При этом классы

$$w_1 + W_{i-1}, \quad w_2 + W_{i-1}, \quad \ldots, \quad w_l + W_{i-1}$$

линейно независимы, если $w_1, w_2, \ldots, w_l$ линейно независимы по модулю W_{i-1}.

В дальнейшем всякое утверждение о смежных классах будем формулировать как такое же утверждение об их представителях, добавляя «по модулю. . . ».

Строим базу для цепочки

$$0 = W_0 \subseteq W_1 \subseteq \ldots \subseteq W_{p-1} \subseteq W_p = V.$$

Чтобы найти базу $e_1, e_2, \ldots, e_q$ по $\operatorname{mod} W_{p-1}$, надо взять базу W_{p-1}, дополнить ее до базы V и взять дополняющие векторы в качестве $e_1, e_2, \ldots, e_q$. Иными словами,

$$e_1, e_2, \ldots, e_q - \text{ база } W_p \text{ по } \operatorname{mod} W_{p-1}.$$

Это значит, что

$$e_1 + W_{p-1}, e_2 + W_{p-1}, \ldots, e_q + W_{p-1}$$

– база W_p/W_{p-1}. Применим преобразование ψ, получим векторы

$$e_1\psi, e_2\psi, \ldots, e_q\psi,$$

которые лежат в W_{p-1}.

Так как $\psi^p = 0$, а $\psi^{p-1} \neq 0$, то построенные векторы линейно независимы по $\bmod W_{p-2}$. Действительно, если

$$\alpha_1(e_1\psi) + \alpha_2(e_2\psi) + \ldots + \alpha_q(e_q\psi) \equiv 0 \ (\bmod W_{p-2}), \quad \alpha_i \in P,$$

то это равносильно тому, что

$$\alpha_1(e_1\psi + W_{p-2}) + \alpha_2(e_2\psi + W_{p-2}) + \ldots + \alpha_q(e_q\psi + W_{p-2}) = W_{p-2},$$

т. е.

$$\alpha_1(e_1\psi) + \alpha_2(e_2\psi) + \ldots + \alpha_q(e_q\psi) \in W_{p-2}$$

или

$$(\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_q e_q)\psi \in W_{p-2}.$$

Следовательно,

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_q e_q \in W_{p-1},$$

но мы выбирали векторы $e_1, e_2, \ldots, e_q$ как дополняющие базу W_{p-1} до базы W_p, а потому $\alpha_1 = \alpha_2 = \ldots = \alpha_q = 0$, т. е. $e_1\psi$, $e_2\psi$, …, $e_q\psi$ линейно независимы по $\bmod W_{p-2}$.

Всякую линейно независимую систему можно дополнить до базы

$$e_1\psi, e_2\psi, \ldots, e_q\psi, f_1, f_2, \ldots, f_s - \text{ база } W_{p-1} \text{ по } \bmod W_{p-2}.$$

Подействовав ψ, получим векторы

$$e_1\psi^2, e_2\psi^2, \ldots, e_q\psi^2, f_1\psi, f_2\psi, \ldots, f_s\psi,$$

которые лежат в W_{p-2}.

Покажем что все они линейно независимы по $\bmod W_{p-3}$.

Действительно, пусть

$$\alpha'_1(e_1\psi^2) + \alpha'_2(e_2\psi^2) + \ldots + \alpha'_q(e_q\psi^2) + \beta_1(f_1\psi) + \beta_2(f_2\psi) + \ldots +$$

$$+\beta_s(f_s\psi) \equiv 0 \,(\bmod W_{p-3})$$

для некоторых $\alpha'_i, \beta_j \in P$. Тогда

$$[\alpha'_1(e_1\psi) + \alpha'_2(e_2\psi) + \ldots + \alpha'_q(e_q\psi) + \beta_1 f_1 + \beta_2 f_2 + \ldots + \beta_s f_s]\psi \in W_{p-3}$$

и

$$[\alpha'_1(e_1\psi) + \alpha'_2(e_2\psi) + \ldots + \alpha'_q(e_q\psi) + \beta_1 f_1 + \beta_2 f_2 + \ldots + \beta_s f_s] \in W_{p-2},$$

но векторы $e_1\psi$, $e_2\psi$, …, $e_q\psi$, f_1, f_2, …, f_s были линейно независимы по $\bmod W_{p-2}$. Следовательно, все $\alpha_i' = \beta_j = 0$.

Дополним до базы

$$e_1\psi^2, e_2\psi^2, \ldots, e_q\psi^2, f_1\psi, f_2\psi, \ldots, f_s\psi, g_1, g_2, \ldots, g_t$$

$$- \text{ база } W_{p-2} \text{ по } \bmod W_{p-3}.$$

Продолжая этот процесс, получим

$$e_1\psi^{p-1}, e_2\psi^{p-1}, \ldots, e_q\psi^{p-1}, f_1\psi^{p-2}, f_2\psi^{p-2}, \ldots, f_s\psi^{p-2},$$

$$g_1\psi^{p-3}, g_2\psi^{p-3}, \ldots, g_t\psi^{p-3}, \ldots, h_1, h_2, \ldots, h_l - \text{ база } W_1 \text{ по } \bmod W_0.$$

Покажем, что если объединить все построенные системы векторов, то получим базу V. Докажем вначале, что эта система линейно независима. Пусть

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_q e_q + \alpha_1'(e_1\psi) + \alpha_2'(e_2\psi) + \ldots + \alpha_q'(e_q\psi) +$$
$$+\beta_1(f_1\psi) + \beta_2(f_2\psi) + \ldots + \beta_s(f_s\psi) + \ldots +$$
$$+\gamma_1 h_1 + \gamma_2 h_2 + \ldots + \gamma_t h_t = 0.$$

Заметим, что вектор из левой части лежит в W_{p-1}. Следовательно, все $\alpha_1 = \alpha_2 = \ldots = \alpha_q = 0$. Оставшийся вектор лежит в W_{p-1}. Следовательно, все $\alpha_i' = \beta_j = 0$ и т. д. Таким образом, система векторов линейно независима.

Докажем, что построенная система векторов максимальна. Пусть $v \in V$. Тогда

$$v \equiv \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_q e_q \,(\bmod W_{p-1})$$

при подходящих α_i. Рассмотрим вектор

$$v - (\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_q e_q) \in W_{p-1}.$$

Для него

$$v - (\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_q e_q) \equiv \alpha_1'(e_1\psi) + \alpha_2'(e_2\psi) + \ldots + \alpha_q'(e_q\psi) +$$
$$+\beta_1 f_1 + \beta_2 f_2 + \ldots + \beta_s f_s \,(\bmod W_{p-2}).$$

Далее рассмотрим

$$v - (\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_q e_q) - (\alpha_1'(e_1\psi) + \alpha_2'(e_2\psi) + \ldots + \alpha_q'(e_q\psi) +$$

$$+\beta_1 f_1 + \beta_2 f_2 + \ldots + \beta_s f_s) \in W_{p-2}$$

и т. д. Продолжая этот процесс, получим, что вектор v есть линейная комбинация векторов построенной системы.

Весь процесс построения удобно заносить в следующую таблицу:

$$\begin{array}{rll}
V = W_p & & \\
\cup & e_1, \ldots, e_q & - \text{ база } W_p \,(\mathrm{mod}\, W_{p-1}); \\
W_{p-1} & & \\
\cup & e_1\psi, \ldots, e_q\psi, & f_1, \ldots, f_s - \text{ база } W_{p-1} \,(\mathrm{mod}\, W_{p-2}); \\
\vdots & & \\
\cup & & \\
W_1 & & \\
\cup & e_1\psi^{p-1}, \ldots, e_q\psi^{p-1}, & f_1\psi^{p-2}, \ldots, f_s\psi^{p-2}, \\
 & & \ldots, h_1, h_2, \ldots, h_l - \text{ база } W_1 \,(\mathrm{mod}\, W_0) \\
0 = W_0. & &
\end{array}$$

Видим, что каждый столбец этой системы образует циклическую базу.

2) Пусть z_i – число i-мерных циклических подпространств. Обозначим $d_i = \dim W_i$. Из первой строки таблицы видим, что число z_p циклических подпространств размерности p равно q и равно $\dim W_p - \dim W_{p-1}$. Из второй строки таблицы видим, что число z_{p-1} циклических подпространств размерности $p-1$ равно s и сумма $z_p + z_{p-1}$ равна $\dim W_{p-1} - \dim W_{p-2}$ и т. д.

В общем случае, из i-й строки построенной базы получаем равенство

$$z_{i+1} + z_{i+2} + \ldots + z_p = d_{i+1} - d_i,$$

из $(i+1)$-й – равенство

$$z_i + z_{i+1} + \ldots + z_p = d_i - d_{i-1}.$$

Вычитая из второго равенства первое, получим

$$z_i = -d_{i-1} + 2d_i - d_{i+1}.$$

Учитывая, что $r_i + d_i = n$ (размерность образа плюс размерность ядра равна размерности пространства), приходим к равенству

$$z_i = -d_{i-1} + 2d_i - d_{i+1} = r_{i-1} - 2r_i + r_{i+1}.$$

Теорема доказана.

34.4. Доказательство теоремы Жордана. Докажем вначале существование.

Пусть дано линейное преобразование φ векторного пространства V. Надо построить такую базу пространства V, в которой матрица преобразования φ имеет жорданову форму. Представим V в виде прямой суммы корневых подпространств:

$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_s,$$

где каждое подпространство V_i соответствует собственному значению α_i. Ограничение φ на V_i индуцирует преобразование φ_{V_i}. Это линейное преобразование подпространства V_i. Если рассмотреть ограничение $(\varphi - \alpha_i\varepsilon)|_{V_i}$, то это тоже линейное преобразование подпространства V_i. Как мы знаем, справедливы равенства

$$\begin{gathered} V_1(\varphi - \alpha_1\varepsilon)^{\rho_1} = 0, \\ V_2(\varphi - \alpha_2\varepsilon)^{\rho_2} = 0, \\ \vdots \\ V_s(\varphi - \alpha_s\varepsilon)^{\rho_s} = 0. \end{gathered}$$

Выбирая максимальный показатель $\rho = \max\{\rho_1, \rho_2, \ldots, \rho_s\}$, видим, что индуцированное преобразование $(\varphi - \alpha_i\varepsilon)|_{V_i} = (\varphi - \alpha_i\varepsilon)_{V_i}$ – нильпотентно ступени не выше ρ. Допустим, что

$$V_i = V_{i_1} \oplus V_{i_2} \oplus \ldots \oplus V_{i_k}$$

– прямая сумма циклических подпространств. Выбирая в каждом V_{i_j} циклическую базу и объединяя их, получим, что в этой базе

$$[(\varphi - \alpha_i\varepsilon)_{V_i}] = \left(\begin{array}{c|c|c|c} J_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & J_2 & \cdots & \mathbf{0} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & J_k \end{array}\right),$$

где каждая клетка имеет такой вид:

$$J_l = \begin{pmatrix} 0 & 1 & & & \mathbf{0} \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & 0 \end{pmatrix}.$$

Если рассмотреть в этой базе матрицу $[\varphi_{V_i}]$, то она будет иметь вид

$$[\varphi_{V_i}] = \left(\begin{array}{c|c|c|c} J_1(\alpha_i) & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & J_2(\alpha_i) & \cdots & \mathbf{0} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & J_k(\alpha_i) \end{array}\right),$$

где каждая клетка является клеткой жордана

$$J_l(\alpha_i) = \begin{pmatrix} \alpha_i & 1 & & & \mathbf{0} \\ & \alpha_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha_i \end{pmatrix}.$$

Объединяя все такие базы подпространств V_i, получим базу, в которой матрица преобразования φ имеет жорданову форму.

Докажем единственность. Пусть некоторое линейное преобразование имеет две жордановы матрицы. Как мы знаем, z_i – количество i-мерных клеток вычисляется по формуле

$$z_i = r_{i-1} - 2r_i + r_{i+1}.$$

Следовательно, количество клеток в обеих жордановых матрицах одинаково. Может меняться только порядок клеток. Теорема доказана.

§ 35. Полиномиальные матрицы

35.1. Задача об эквивалентности λ-матриц. *Полиномиальной* называют матрицу, элементами которой являются многочлены. Рассмотрим кольцо многочленов $P[\lambda]$ над полем P от переменной λ. Тогда полиномиальная матрица имеет вид

$$F = \begin{pmatrix} f_{11} & f_{12} & \dots & f_{1n} \\ f_{21} & f_{22} & \dots & f_{2n} \\ \dots & \dots & \dots & \dots \\ f_{n1} & f_{n2} & \dots & f_{nn} \end{pmatrix}, \quad f_{ij} \in P[\lambda].$$

Такие полиномиальные матрицы также называют «λ-матрицами». Частным случаем λ-матриц являются характеристические матрицы.

О п р е д е л е н и е. *Элементарными преобразованиями* λ-матриц называются следующие преобразования: 1) умножение строки на ненулевой скаляр; 2) прибавление к одной строке другой, умноженной на некоторый многочлен; 1)$'$ умножение столбца на ненулевой скаляр; 2)$'$ прибавление к одному столбцу другого, умноженного на некоторый многочлен.

О п р е д е л е н и е. Две λ-матрицы F_1 и F_2 называются *эквивалентными*, если от F_1 можно перейти к F_2 цепочкой элементарных преобразований:

$$F_1 = G_1 \longrightarrow G_2 \longrightarrow \ldots \longrightarrow G_s = F_2.$$

П р и м е р. Если в λ-матрице переставить две строки, то полученная матрица ей эквивалентна. Действительно, пусть F – некоторая λ-матрица. Выполним следующие элементарные преобразования:

1) умножим j-ю строку на 1 и прибавим к i-й;

2) умножим i-ю строку на -1 и прибавим к j-й;

3) умножим j-ю строку на 1 и прибавим к i-й;

4) j-ю строку умножим на -1.

Видим, что полученная матрица получается из F перестановкой i-й и j-й строк.

Нетрудно убедиться, что введенное отношение является отношением эквивалентности. Относительно этого отношения эквивалентности множество λ-матриц распадается на смежные классы.

О п р е д е л е н и е. *Канонической* называется диагональная λ-матрица:

$$\operatorname{diag}(d_1(\lambda), d_2(\lambda), \ldots, d_n(\lambda)),$$

такая, что: 1) $d_i(\lambda) | d_{i+1}(\lambda)$, $i = 1, 2, \ldots, n-1$; 2) если многочлен $d_i(\lambda)$ отличен от нуля, то его старший коэффициент равен единице.

Из этого определения следует, что каноническая матрица имеет вид

$$\operatorname{diag}(1, 1, \ldots, 1, f_1, f_2, \ldots, f_s, 0, 0, \ldots, 0),$$

т. е. единицы стоят вверху, а нули – внизу.

О п р е д е л е н и е. Элементы $d_i(\lambda)$ называются *инвариантными множителями* канонической матрицы.

Докажем, что в каждом смежном классе лежит по одной канонической матрице.

35.2. Приведение λ-матрицы к каноническому виду.

Т е о р е м а 1. *Всякая λ-матрица с помощью элементарных преобразований может быть приведена к каноническому виду.*

Д о к а з а т е л ь с т в о проведем индукцией по n – размерности матрицы. При $n = 1$ матрица является многочленом. Если его старший коэффициент равен единице, то мы имеем каноническую матрицу. В противном случае, поделив на старший коэффициент, приходим к канонической матрице.

Предположим, что теорема справедлива для матриц размера $n-1$, и рассмотрим λ-матрицу F размера n. Если $F = 0$, то доказывать нечего. Пусть $F \neq 0$. Среди всех матриц, эквивалентных F, выберем матрицу, у которой на месте $(1,1)$ стоит многочлен наименьшей степени со старшим коэффициентом 1, т. е.

$$F \text{ эквивалентна матрице } \begin{pmatrix} f_{11} & f_{12} & \dots & f_{1n} \\ f_{21} & f_{22} & \dots & f_{2n} \\ \dots & \dots & \dots & \dots \\ f_{n1} & f_{n2} & \dots & f_{nn} \end{pmatrix}.$$

Утверждается, что $f_{11}|f_{1i}$ и $f_{11}|f_{i1}$ при $i = 2, 3, \dots, n$. Действительно, если $f_{1i} = f_{11}q_{1i} + r_{1i}$ и $r_{1i} \neq 0$, то $\deg r_{1i} < \deg f_{1i}$, но тогда, используя элементарные преобразования, получим матрицу, эквивалентную F, у которой на месте $(1,1)$ стоит r_{1i}. Противоречие. Аналогично проверяется, что и $f_{11}|f_{i1}$. Используя элементарные преобразования, перейдем от матрицы F к матрице

$$\begin{pmatrix} f_{11} & 0 & \dots & 0 \\ 0 & g_{22} & \dots & g_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & g_{n2} & \dots & g_{nn} \end{pmatrix}, \quad g_{ij} \in P[\lambda].$$

Заметим, что в ней $f_{11}|g_{ij}$. Действительно, прибавляя i-ю строку к первой, так же как и выше, устанавливаем, что $f_{11}|g_{ij}$. По предположению индукции существует цепочка элементарных преобразований:

$$\begin{pmatrix} g_{22} & \dots & g_{2n} \\ \dots & \dots & \dots \\ g_{n2} & \dots & g_{nn} \end{pmatrix} \longrightarrow \dots \longrightarrow \begin{pmatrix} h_2 & \dots & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & \dots & h_n \end{pmatrix},$$

где последняя матрица диагональна, $h_i(\lambda)|h_{i+j}(\lambda)$, и если $h_i(\lambda) \neq 0$, то старший коэффициент равен единице. Распространяя эту цепочку

элементарных преобразований с матрицы (g_{ij}) до нашей матрицы, получим

$$\left(\begin{array}{c|ccc} f_{11} & 0 & \dots & 0 \\ \hline 0 & g_{22} & \dots & g_{2n} \\ \vdots & \multicolumn{3}{c}{\dots\dots\dots\dots} \\ 0 & g_{n2} & \dots & g_{nn} \end{array}\right) \longrightarrow \dots \longrightarrow \left(\begin{array}{c|ccc} f_{11} & 0 & \dots & 0 \\ \hline 0 & h_2 & \dots & \mathbf{0} \\ \vdots & & \ddots & \\ 0 & \mathbf{0} & \dots & h_n \end{array}\right).$$

Как легко заметить, последняя матрица является канонической. Теорема доказана.

35.3. Единственность канонического вида λ-матриц. Пусть F – λ-матрица. Для каждого $k = 1, 2, \dots, n$ обозначим $D_{k,F}(\lambda)$ (или просто $D_k(\lambda)$) – приведенный наибольший общий делитель миноров k-го порядка матрицы F, если они не все равны нулю, и положим $D_{k,F}(\lambda) = 0$ в противном случае. Таким образом, каждой λ-матрице мы сопоставим n многочленов $D_1(\lambda)$, $D_2(\lambda)$, …, $D_n(\lambda)$. Покажем, что они остаются инвариантными при элементарных преобразованиях.

Т е о р е м а 2. *Если матрицы F и G эквивалентны, то $D_{k,F}(\lambda) = D_{k,G}(\lambda)$, $k = 1, 2, \dots, n$.*

Д о к а з а т е л ь с т в о. Так как F эквивалентна G, то существует цепочка элементарных преобразований, переводящая F в G. Если мы покажем, что $D_{k,F}(\lambda)$ не меняется при выполнении одного элементарного преобразования, то и любая цепочка не меняет $D_{k,F}(\lambda)$.

1) Предположим, что G получается из F при помощи умножения некоторой строки на ненулевой скаляр. Если эта строка входит в минор, то после применения элементарного преобразования он умножается на скаляр, но на вычисление приведенного общего делителя эти скаляры никак не повлияют. Аналогично проверяется, что и при умножении столбца на ненулевой скаляр $D_{k,F}(\lambda)$ не меняются.

Случай, когда G получается из F при помощи элементарного преобразования $1'$, разбирается аналогично.

2) Пусть G получается из F после умножения i-й строки на $f \in P[x]$, $f \neq 0$, и прибавления к j-й строке. В зависимости от того, как расположен рассматриваемый минор, возможны четыре случая:

а) i-я и j-я строки проходят через минор;

б) i-я строка проходит через минор, а j-я – нет;

в) ни i-я, ни j-я строки не проходят через минор;

г) j-я строка проходит через минор, а i-я – нет.

В случаях а–в очевидно, что после элементарного преобразования 2 минор не меняется.

Докажем, что и в случае г утверждение теоремы справедливо. Пусть соответствующий минор матрицы G имеет вид

$$M_G = \begin{vmatrix} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ f_{j,\nu_1} + f f_{i,\nu_1} \quad f_{j,\nu_2} + f f_{i,\nu_2} \quad \dots \quad f_{j,\nu_k} + f f_{i,\nu_k} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \end{vmatrix},$$

где выделена j-я строка. Представим его в виде суммы двух определителей:

$$M_G = \begin{vmatrix} \dots\dots\dots\dots\dots\dots \\ f_{j,\nu_1} \quad f_{j,\nu_2} \quad \dots \quad f_{j,\nu_k} \\ \dots\dots\dots\dots\dots\dots \end{vmatrix} + f \begin{vmatrix} \dots\dots\dots\dots\dots\dots \\ f_{i,\nu_1} \quad f_{i,\nu_2} \quad \dots \quad f_{i,\nu_k} \\ \dots\dots\dots\dots\dots\dots \end{vmatrix} = M_F + f N_F,$$

где символом N_F обозначили второй минор, который отличается от соответствующего минора матрицы F тем, что у него вместо j-й строки стоит соответствующая часть i-й строки матрицы F. Рассмотрим $D_{k,F}$. По условию

$$D_{k,F} | M_F, \quad D_{k,F} | N_F.$$

Следовательно, $D_{k,F} | M_G$, а потому $D_{k,F} | D_{k,G}$. Аналогично проверяется, что $D_{k,G} | D_{k,F}$. Следовательно, $D_{k,F} = D_{k,G}$.

Случай, когда G получается из F при помощи элементарного преобразования $2'$, разбирается аналогично. Теорема доказана.

Т е о р е м а 3. *Для всякой λ-матрицы F: 1) $d_1 d_2 \dots d_k = D_{k,F}$; 2) канонический вид λ-матрицы определяется ею однозначно.*

Д о к а з а т е л ь с т в о. 1) Как мы знаем, λ-матрица F эквивалентна диагональной матрице

$$\operatorname{diag}(d_1, d_2, \dots, d_n).$$

Рассмотрим миноры порядка k полученной диагональной матрицы. Очевидно, что если некоторая строка проходит через минор, а одноименный столбец – нет, то этот минор равен нулю. Следовательно, всякий ненулевой минор порядка k равен $d_{i_1} d_{i_2} \dots d_{i_k}$. Среди таких миноров есть и минор $d_1 d_2 \dots d_k$, это и будет $D_{k,F}$. Действительно, так как $d_1 | d_2$, $d_2 | d_3$ и т. д., то

$$d_1 d_2 \dots d_k | d_{i_1} d_{i_2} \dots d_{i_k},$$

т. е. $d_1 d_2 \dots d_k$ делит любой минор порядка k, а так как увеличить его нельзя, то он и будет наибольшим общим делителем миноров порядка k.

2) По теореме 2 набор $D_{1,F}, D_{2,F}, \dots, D_{n,F}$ определяется единственным образом. Так как $d_k = D_{k,F}/D_{k-1,F}$, то и d_k определяется единственным образом. Следовательно, каноническая матрица определяется единственным образом. Теорема доказана.

Таким образом, мы решили исходную задачу, показав, что всякая λ-матрица эквивалентна единственной канонической матрице.

35.4. Эквивалентность λ-матриц унимодулярным матрицам.

О п р е д е л е н и е. *Унимодулярная матрица* – это λ-матрица, определитель которой лежит в самом поле P и не равен нулю.

Т е о р е м а 4. *Матрицы F и G эквивалентны тогда и только тогда, когда существуют унимодулярные матрицы X и Y такие, что $F = XGY$.*

Д о к а з а т е л ь с т в о. Установим импликацию $\Rightarrow$. Пусть F эквивалентна G, т. е. существует цепочка элементарных преобразований

$$F \longrightarrow \dots \longrightarrow G. \tag{1}$$

Надо построить унимодулярные матрицы X и Y, удовлетворяющие равенству $F = XGY$.

Заметим, что каждому элементарному преобразованию соответствует умножение на некоторую матрицу. Например, элементарному преобразованию 1 (умножение i-й строки матрицы на ненулевой скаляр α) соответствует умножение слева на диагональную матрицу

$$\operatorname{diag}(1, \dots, 1, \alpha, 1, \dots, 1),$$

где α стоит на i-м месте. Элементарному преобразованию 2 (умножение i-й строки на ненулевой элемент f и прибавление к j-й строке) соответствует умножение слева на трансвекцию $T_{ji}(f)$. Элементарным преобразованиям $1'$ и $2'$ соответствуют умножение справа на диагональную матрицу и соответственно на трансвекцию.

Таким образом, цепочке элементарных преобразований (1) соответствует произведение матриц

$$X_1 X_2 \dots X_s F X_{s+1} X_{s+2} \dots X_t = G,$$

где каждая X_i либо диагональная матрица, либо трансвекция. Учитывая, что

$$\det \operatorname{diag}(1, \ldots, 1, \alpha, 1, \ldots, 1) = \alpha, \quad \det T_{ij}(f) = 1,$$

заключаем, что матрицы

$$X_1 X_2 \ldots X_s = X^{-1}, \quad X_{s+1} X_{s+2} \ldots X_t = Y^{-1}$$

являются унимодулярными, а так как обратные к ним также унимодулярные, то построенные матрицы и будут искомыми.

$\Leftarrow$. Пусть $F = XGY$, где X и Y – унимодулярные матрицы. Сама матрица X зависит от λ, а ее определитель – не зависит. Очевидно, что $D_{n,X}(\lambda) = 1$, так как является приведенным многочленом и с точностью до умножения на скаляр совпадает с $\det X$. Все остальные $D_{k,X}(\lambda)$ при $1 \leq k < n$ равны 1, что следует из того, что $d_1 d_2 \ldots d_n = 1$ и ни один из d_i не содержит λ. Следовательно, все $D_{k,X}(\lambda) = 1$, а потому по теореме 3 и все $d_i = 1$. Таким образом, матрица X эквивалентна единичной матрице. Аналогично проверяется, что любая унимодулярная матрица эквивалентна единичной матрице. Следовательно, существуют цепочки элементарных преобразований:

$$X \longrightarrow \ldots \longrightarrow E,$$

$$Y \longrightarrow \ldots \longrightarrow E,$$

или на матричном языке:

$$X = X_1 X_2 \ldots X_s E X_{s+1} X_{s+2} \ldots X_t,$$

$$Y = Y_1 Y_2 \ldots Y_p E Y_{p+1} Y_{p+2} \ldots Y_q,$$

где X_i, Y_j – матрицы, соответствующие элементарным преобразованиям. Отсюда

$$F = XGY = X_1 X_2 \ldots X_t G Y_1 Y_2 \ldots Y_q,$$

а значит, существует цепочка элементарных преобразований:

$$G \longrightarrow \ldots \longrightarrow F.$$

Следовательно, G эквивалентна F. Теорема доказана.

35.5. Деление λ-матриц. Если

$$A(\lambda) = \left(\sum_k a_{ij}^k \lambda^k \right)_{i,j=1}^n$$

– некоторая λ-матрица, то ее можно представить в виде

$$A(\lambda) = A_0 + A_1\lambda + \ldots + A_s\lambda^s,$$

где каждая матрица A_i является матрицей из $M_n(P)$. Если в этом разложении $A_s \neq 0$, то назовем s *степенью* матрицы $A(\lambda)$ и будем обозначать $s = \deg A(\lambda)$.

П р и м е р.

$$\begin{pmatrix} \lambda^2+1 & \lambda \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \lambda + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \lambda^2.$$

О п р е д е л е н и е. λ-матрица $A(\lambda)$ называется *регулярной*, если $\det A_s \neq 0$.

Л е м м а 1. *Для любых λ-матриц $A(\lambda), B(\lambda)$ справедливы неравенства:*

$$\deg\left(A(\lambda) + B(\lambda)\right) \leq \max\left\{\deg A(\lambda), \deg B(\lambda)\right\},$$

$$\deg\left(A(\lambda) \cdot B(\lambda)\right) \leq \deg A(\lambda) + \deg B(\lambda).$$

Если хотя бы одна из матриц, $A(\lambda)$ или $B(\lambda)$, регулярна, то последнее неравенство превращается в равенство.

Л е м м а 2. *Пусть $A(\lambda)$ и $B(\lambda)$ – две λ-матрицы одной размерности, причем $A(\lambda)$ регулярна. Тогда существуют однозначно определенные матрицы $Q_l(\lambda)$, $R_l(\lambda)$, $Q_r(\lambda)$, $R_r(\lambda)$, такие, что*

$$B(\lambda) = A(\lambda)Q_l(\lambda) + R_l(\lambda), \quad B(\lambda) = Q_r(\lambda)A(\lambda) + R_r(\lambda),$$

где $R_l(\lambda)$ либо равна нулю, либо ее степень меньше $\deg(A(\lambda))$, *тоже для $R_r(\lambda)$.*

Доказательство лемм 1 и 2 такое же, как для обычных многочленов.

35.6. Скалярная эквивалентность λ-матриц. Матрицы $A(\lambda)$ и $B(\lambda)$ называются *скалярно эквивалентными*, если существуют невырожденные скалярные матрицы P и Q (их элементы не зависят от λ) такие, что

$$A(\lambda) = PB(\lambda)Q.$$

Т е о р е м а 5 (критерий эквивалентности). *Регулярные λ-матрицы $A\lambda + B$ и $C\lambda + D$ первой степени, в которых A, B, C, $D \in M_n(P)$, эквивалентны тогда и только тогда, когда они скалярно эквивалентны.*

Д о к а з а т е л ь с т в о. $\Leftarrow$. Следует из того, что условие скалярной эквивалентности более сильное, чем условие унимодулярной эквивалентности.

$\Rightarrow$. Пусть

$$A\lambda + B = X(C\lambda + D)Y, \tag{2}$$

где X, Y – унимодулярные матрицы. Разделим X слева на $A\lambda + B$, а Y – справа:

$$X = (A\lambda + B)X_1 + X_2, \quad Y = Y_1(A\lambda + B) + Y_2. \tag{3}$$

Степени X_2, Y_2 нулевые, или эти матрицы равны нулю. В любом случае, они не зависят от λ. Следовательно, из равенства (2) имеем

$$X^{-1}(A\lambda + B) = (C\lambda + D)Y_1(A\lambda + B) + (C\lambda + D)Y_2,$$

или

$$[X^{-1} - (C\lambda + D)Y_1](A\lambda + B) = (C\lambda + D)Y_2.$$

Как мы знаем, если X – унимодулярна, то X^{-1} тоже унимодулярна. Положим

$$T = X^{-1} - (C\lambda + D)Y_1. \tag{4}$$

Тогда

$$T(A\lambda + B) = (C\lambda + D)Y_2. \tag{5}$$

Сравнивая степени левой и правой частей, видим, что T и Y_2 – скалярные матрицы. Покажем, что T – невырожденная матрица. Домножив обе части равенства (4) слева на X, получим

$$XT = E - X(C\lambda + D)Y_1$$

или

$$E = XT + X(C\lambda + D)Y_1.$$

Первое вхождение X заменим на выражение из (3), а произведение $X(C\lambda + D)$ – на выражение из (2), получим

$$E = (A\lambda + B)X_1T + X_2T + (A\lambda + B)Y^{-1}Y_1,$$

т. е.

$$E - X_2T = (A\lambda + B)(X_1T + Y^{-1}Y_1).$$

В левой части стоит скалярная матрица, а в правой – матрица $A\lambda+B$, которая регулярна и содержит λ. Следовательно, $X_1T + Y^{-1}Y_1 = 0$. Значит

$$E - X_2T = 0 \text{ или } X_2T = E,$$

а потому $\det T \neq 0$. Из равенства (5)

$$A\lambda + B = T^{-1}(C\lambda + D)Y_2.$$

Приравнивая коэффициенты при λ, получим

$$A = T^{-1}CY_2.$$

Отсюда

$$\det A = \det T^{-1} \cdot \det C \cdot \det Y_2.$$

Так как $\det A \neq 0$, то и $\det Y_2 \neq 0$. Теорема доказана.

35.7. Критерий подобия скалярных матриц.

Т е о р е м а 6. *Две скалярные матрицы A и B подобны тогда и только тогда, когда их характеристические матрицы эквивалентны.*

Д о к а з а т е л ь с т в о. Пусть A и B подобны, т. е. $A = T^{-1}BT$ для некоторой невырожденной матрицы T. Тогда

$$T^{-1}(B - \lambda E)T = T^{-1}BT - \lambda T^{-1}ET = A - \lambda E,$$

т. е.

$$A - \lambda E \sim B - \lambda E.$$

Обратно. Пусть

$$A - \lambda E \sim B - \lambda E.$$

Так как это регулярные матрицы первой степени, то по доказанной теореме 5 $A - \lambda E$ и $B - \lambda E$ скалярно эквивалентны, т. е. существуют невырожденные скалярные матрицы P и Q такие, что

$$A - \lambda E = P(B - \lambda E)Q.$$

Приравнивая матрицы при одинаковых степенях λ, получим

$$A = PBQ, \quad E = PEQ.$$

Из последнего равенства

$$E = PQ,$$

т. е.

$$P = Q^{-1}.$$

Следовательно,

$$A = Q^{-1}BQ.$$

Теорема доказана.

35.8. Элементарные делители λ-матриц. Любая λ-матрица, имеющая нормальную диагональную форму, определяется набором линейных множителей.

О п р е д е л е н и е. Пусть F – нормальная диагональная форма некоторой λ-матрицы A:

$$F = \mathrm{Diag}(f_1, f_2, \ldots, f_n), \quad \text{старший коэффициент } f_1 \text{ равен } 1.$$

Выбросим из набора $f_1, f_2, \ldots, f_n$ те, которые равны 0 или 1. Остальные разложим над основным полем P на неразложимые множители:

$$f_i = \varepsilon_{i_1}^{\alpha_{i_1}} \varepsilon_{i_2}^{\alpha_{i_2}} \ldots \varepsilon_{i_{s_i}}^{\alpha_{i_{s_i}}}, \quad \varepsilon_{i_j} \neq \varepsilon_{i_t} \text{ при } j \neq t,$$

где ε_{i_j} – неразложим над P. Тогда множество

$$\{\varepsilon_{i_j}^{\alpha_{i_j}}\}$$

называется *системой элементарных делителей* матрицы A.

П р и м е р. Пусть

$$A \sim F = \mathrm{Diag}(1, 1, \lambda, \lambda, \lambda(1+\lambda), \lambda^2(1+\lambda), 0, 0).$$

Тогда

$$\lambda, \lambda, \lambda, (\lambda+1), \lambda^2, \lambda+1$$

– система элементарных делителей матрицы A.

В настоящем пункте будет доказана

Т е о р е м а 7. *Система элементарных делителей матрицы A, ее ранг и размерность определяют A с точностью до эквивалентности.*

П р и м е р. Пусть размерность матрицы A равна 7, ранг равен 5, а серия элементарных делителей

$$\lambda, \lambda, \lambda^2, \lambda^2, (\lambda+1)^2, \lambda+1.$$

Найдем нормальную диагональную форму матрицы A. Так как $7 - 5 = 2$, то в конце стоит два нуля и

$$A \sim \mathrm{Diag}(1, \lambda, \lambda, \lambda^2(1+\lambda), f_5, 0, 0).$$

При этом $f_5 = \lambda^i(1+\lambda)^j$ для некоторых i и j. Учитывая, что все f_i делят f_5, заключаем, что $f_5 = \lambda^2(1+\lambda)^2$.

Т е о р е м а 8. *Система элементарных делителей клеточно-диагональной λ-матрицы*

$$G = \mathrm{Diag}(G_1, G_2, \ldots, G_t)$$

совпадает с объединением систем элементарных делителей ее диагональных клеток G_1, G_2, …, G_t.

Д о к а з а т е л ь с т в о. 1) Пусть G – диагональная матрица

$$G = \mathrm{diag}(g_1, g_2, \ldots, g_n)$$

и ее нормальная диагональная форма имеет вид

$$F = \mathrm{diag}(f_1, f_2, \ldots, f_n).$$

Покажем, что элементарные делители для $f_1, f_2, \ldots, f_n$ совпадают с элементарными делителями для $g_1, g_2, \ldots, g_n$.

Можно считать, что $g_i \neq 0$, $i = 1, 2, \ldots, n$ и

$$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_s$$

– все элементарные делители $g_1, g_2, \ldots, g_n$, т. е.

$$g_i = a_i \varepsilon_1^{m_{i,1}} \varepsilon_2^{m_{i,2}} \ldots \varepsilon_s^{m_{i,s}}.$$

Покажем, что ε_i входят в f_i с теми же степенями. Пусть

$$\det G = g_1 g_2 \ldots g_n = a f_1 f_2 \ldots f_n = a \prod_{i=1}^{n} f_i = a \prod_{i,j} \varepsilon_{ij}^{\alpha_{ij}}$$

– разложение по серии элементарных делителей F. Учитывая, что разложение на неразложимые множители однозначно, видим, что каждый элементарный делитель матрицы F входит в множество элементарных делителей матрицы G.

Как выглядит элементарный делитель матрицы G, соответствующий ε_1? Вычислим миноры k-го порядка матрицы G. Не нарушая общности, можно считать, что

$$m_{11} \leq m_{21} \leq \ldots \leq m_{n1}.$$

Вычислим наибольший общий делитель миноров k-го порядка матрицы G и обозначим его $D_k(G)$:

$$\{g_{i_1} g_{i_2} \dots g_{i_k} \mid i_1 \le i_2 \le \dots \le i_k\}.$$

Видим, что

$$g_{i_1} g_{i_2} \dots g_{i_k} = \varepsilon_1^{m_{i_1,1}+m_{i_2,1}+\dots+m_{i_k,1}} h(\lambda).$$

Наименьшая степень

$$\varepsilon_1^{m_{1,1}+m_{2,1}+\dots+m_{k,1}} F_k(\lambda),$$

$$f_k = d_k(G) = \frac{D_k(G)}{D_{k-1}(G)} = \varepsilon_1^{m_{k1}} F_k^*(\lambda).$$

Так как $F_k^*(\lambda)$ не делится на $F_k(\varepsilon_1)$. Набор элементарных делителей ε_1 совпадает.

2) Пусть

$$F = \operatorname{diag}(F_1, F_2, \dots, F_s)$$

– клеточно-диагональная матрица. Цепочкой элементарных преобразований приведем каждую клетку к диагональному виду:

$$F_1 \text{ эквивалентна } G_1,$$

$$F_2 \text{ эквивалентна } G_2,$$

$$\dots\dots\dots\dots\dots\dots\dots\dots$$

$$F_s \text{ эквивалентна } G_s,$$

где каждая G_i является канонической диагональной. Тогда F эквивалентна диагональной матрице

$$G = \operatorname{diag}(G_1, G_2, \dots, G_s).$$

По рассмотренному выше случаю

$$\{\text{элементарные делителей } F\} =$$

$$= \{\text{элементарные делителей } G\} =$$

$$= \bigcup_{i=1}^{s} \{\text{ элементарные делители } G_i\} = \bigcup_{i=1}^{s} \{\text{ элементарные делители } F_i\}.$$

Теорема доказана.

35.9. Другое доказательство теоремы Жордана. Докажем предварительно такое утверждение.

Л е м м а 3. *Если*

$$A = \begin{pmatrix} \alpha & 1 & & & \mathbf{0} \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha \end{pmatrix}$$

– жорданова клетка степени m, то характеристическая матрица $\lambda E - A$ имеет единственный элементарный делитель $(\lambda - \alpha)^m$.

Д о к а з а т е л ь с т в о. Рассмотрим характеристическую матрицу

$$\lambda E - A = \begin{pmatrix} \lambda - \alpha & -1 & & & \mathbf{0} \\ & \lambda - \alpha & -1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & -1 \\ \mathbf{0} & & & & \lambda - \alpha \end{pmatrix}.$$

Чтобы найти инвариантные множители, найдем наибольшие общие делители миноров:

$$D_m(\lambda) = (\lambda - \alpha)^m, \quad D_{m-1}(\lambda) = 1.$$

Поэтому и все $D_k(\lambda) = 1$, $k = m - 2, \ldots, 1$. Учитывая, что $D_k = d_1 d_2 \ldots d_k$, заключаем:

$$d_1(\lambda) = d_2(\lambda) = \ldots = d_{m-1}(\lambda) = 1, \quad d_m(\lambda) = (\lambda - \alpha)^m,$$

и наша матрица эквивалентна матрице

$$\begin{pmatrix} 1 & & & & \mathbf{0} \\ & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ \mathbf{0} & & & & (\lambda - \alpha)^m \end{pmatrix}.$$

По определению элементарного делителя $(\lambda - \alpha)^m$ – единственный элементарный делитель матрицы $\lambda E - A$. Лемма доказана.

Переходим непосредственно к доказательству теоремы Жордана. Докажем вначале существование.

Пусть $A \in M_n(P)$ и P содержит все характеристические корни матрицы A. Рассмотрим характеристическую матрицу $\lambda E - A$ и найдем все ее элементарные делители:

$$(\lambda - \alpha_i)^{n_{ij}}.$$

По каждому из этих многочленов напишем клетку жордана:

$$J_{ij} = \begin{pmatrix} \alpha_i & 1 & & & \mathbf{0} \\ & \alpha_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha_i \end{pmatrix}$$

– клетка жордана степени n_{ij}. Из этих клеток составим матрицу жордана J. Утверждается, что эта матрица искомая. Действительно, по предыдущей теореме и лемме множество элементарных делителей характеристической матрицы $\lambda E - A$ совпадает с множеством элементарных делителей матрицы $\lambda E - J$. Матрица

$$\begin{pmatrix} \lambda E - J_{11} & & & & \mathbf{0} \\ & \lambda E - J_{12} & & & \\ & & \ddots & & \\ & & & \ddots & \\ \mathbf{0} & & & & \lambda E - J_{rs} \end{pmatrix}$$

является клеточно-диагональной, и каждая клетка дает по одному элементарному делителю $(\lambda - \alpha_i)^{n_{ij}}$. Поэтому канонический вид матрицы $\lambda E - A$ совпадает с каноническим видом матрицы $\lambda E - J$, т. е. $\lambda E - A$ эквивалентна $\lambda E - J$. Как мы знаем, если матрицы унимодулярны и эквивалентны, то они скалярно эквивалентны, т. е.

$$\lambda E - A = U(\lambda E - J)V,$$

где U и V не зависят от λ. Сравнивая матрицы при одинаковых степенях λ в левой и правой частях, получим

$$E = UV, \quad A = UJV.$$

Следовательно,

$$A = V^{-1}JV$$

и мы доказали, что матрица A подобна жордановой матрице.

Докажем единственность. Надо доказать, что если жорданова матрица J_1 подобна жордановой матрице J_2, то они совпадают с точностью до порядка жордановых клеток. Из подобия жордановых матриц следует, что

$$\lambda E - J_1 \text{ эквивалентна } \lambda E - J_2.$$

Отсюда

$$\{\text{элементарные делители } \lambda E - J_1\} = \{\text{элементарные делители } \lambda E - J_2\},$$

а потому

$$\{\text{клетки жордана в } J_1\} = \{\text{клетки жордана в } J_2\}.$$

Теорема доказана.

С л е д с т в и е. *Матрица, у которой все характеристические корни лежат в основном поле, подобна диагональной тогда и только тогда, когда все элементарные делители ее характеристической матрицы линейны, т. е. имеют вид* $\lambda - \alpha_i$.

§ 36. Функции от матриц

36.1. Многочлены от матриц. В этом параграфе будем рассматривать случай, когда поле является полем комплексных чисел. Как мы знаем, в этом случае всякая матрица приводится к жордановой форме.

Т е о р е м а 1. *Если* $A \in M_n(\mathbb{C})$, $f \in \mathbb{C}[\lambda]$ *и* A *подобна матрице жордана:* $A = T^{-1}JT$, *где*

$$J = J_1 \oplus J_2 \oplus \ldots \oplus J_s = \begin{pmatrix} J_1 & & & \mathbf{0} \\ & J_2 & & \\ & & \ddots & \\ & & & \ddots \\ \mathbf{0} & & & J_s \end{pmatrix},$$

то

1) $f(A) = T^{-1}f(J)T$;
2) $f(J) = f(J_1) \oplus f(J_2) \oplus \ldots \oplus f(J_s)$;
3) значение f от жордановой клетки степени m определяется равенством

$$f\left(\begin{pmatrix} \alpha & 1 & & & \mathbf{0} \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha \end{pmatrix}\right) = \begin{pmatrix} f(\alpha) & \frac{f'(\alpha)}{1!} & \ldots & \frac{f^{(m-1)}(\alpha)}{(m-1)!} \\ 0 & f(\alpha) & \ldots & \frac{f^{(m-2)}(\alpha)}{(m-2)!} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & f(\alpha) \end{pmatrix}.$$

Д о к а з а т е л ь с т в о. 1) Если

$$f(\lambda) = a_0\lambda^m + a_1\lambda^{m-1} + \ldots + a_m,$$

то

$$f(A) = f(T^{-1}JT) = a_0(T^{-1}JT)^m + a_1(T^{-1}JT)^{m-1} + \ldots + a_mE =$$
$$= a_0(T^{-1}J^mT) + a_1(T^{-1}J^{m-1}T) + \ldots + a_mE = T^{-1}f(J)T.$$

2) Если $B = B_1 \oplus B_2 \oplus \ldots \oplus B_s$ и $C = C_1 \oplus C_2 \oplus \ldots \oplus C_s$ – две клеточно-диагональные матрицы и размеры клетки B_i совпадают с размерами клетки C_i для всех $i = 1, 2, \ldots, s$, то

$$B + C = (B_1 + C_1) \oplus (B_2 + C_2) \oplus \ldots \oplus (B_s + C_s),$$
$$BC = B_1C_1 \oplus B_2C_2 \oplus \ldots \oplus B_sC_s,$$
$$\alpha B = \alpha B_1 \oplus \alpha B_2 \oplus \ldots \oplus \alpha B_s.$$

Отсюда

$$f(J) = f(J_1) \oplus f(J_2) \oplus \ldots \oplus f(J_s).$$

3) По формуле Тейлора

$$f(\lambda) = \sum_{k=0}^{m} \frac{f^{(k)}(\alpha)}{k!}(\lambda - \alpha)^k, \text{ для любого } \alpha \in \mathbb{C}.$$

Рассмотрим клетку жордана:

$$B = \begin{pmatrix} \alpha & 1 & & & \mathbf{0} \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha \end{pmatrix}.$$

Для нее

$$f(B) = \sum_{k=0}^{m} \frac{f^{(k)}(\alpha)}{k!}(B - \alpha E)^k.$$

Обозначим

$$C = B - \alpha E = \begin{pmatrix} 0 & 1 & & & \mathbf{0} \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & 0 \end{pmatrix}.$$

Остается доказать, что

$$C^k = \begin{pmatrix} 0 & \dots & 1 & & & \mathbf{0} \\ & 0 & \dots & 1 & & \\ & & \ddots & & \ddots & \\ & & & \ddots & & 1 \\ & & & & \ddots & \vdots \\ \mathbf{0} & & & & & 0 \end{pmatrix},$$

где единицы стоят на k-й диагонали от главной.

Индукция по k. Для $k = 1$ утверждение тривиально. Предположим, что формула справедлива для k, и рассмотрим $k + 1$. Имеем

$$C^k = \sum_{i=1}^{r-k} E_{i,i+k},$$

где E_{ij} – матрица, у которой на месте (i, j) стоит единица, а на всех остальных местах – нули. Для таких матриц справедлива следующая формула умножения:

$$E_{kl}E_{rs} = \begin{cases} E_{ks} & \text{при} \quad l = r, \\ 0 & \text{при} \quad l \neq r. \end{cases}$$

Тогда

$$C^{k+1} = C^k C = \sum_{i=1}^{r-k} E_{i,i+k} \sum_{j=1}^{r-1} E_{j,j+1} = \sum_{i=1}^{r-k-1} E_{i,i+k+1}.$$

Теорема доказана.

36.2. Функции от матриц. Многочлен Лагранжа – Сильвестера. Пусть $f : \mathbb{C} \longrightarrow \mathbb{C}$ – некоторая функция, $A \in M_n(\mathbb{C})$ и $(\lambda-\alpha_i)^{n_{ij}}$ – все элементарные делители характеристической матрицы $\lambda E-A$. Хотим определить $f(A)$, т. е. значение функции f от матрицы A.

Предположим, что выполняется следующее условие:

$$\textit{существуют все производные } f^{(k)}(\alpha_i),$$

$$k = 0, 1, \ldots, n_{ij} - 1, \quad i = 1, 2, \ldots, s. \tag{1}$$

Найдем жорданову форму матрицы A:

$$J = J_1 \oplus J_2 \oplus \ldots \oplus J_l,$$

где J_i – жордановы клетки, $A = T^{-1}JT$.

О п р е д е л е н и е. Положим:

1) $f(A) = T^{-1}f(J)T$;

2) $f(J) = f(J_1) \oplus f(J_2) \oplus \ldots \oplus f(J_l)$;

3) $$f\left(\begin{pmatrix} \alpha_i & 1 & & & \mathbf{0} \\ & \alpha_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha_i \end{pmatrix}\right) = \begin{pmatrix} f(\alpha_i) & \frac{f'(\alpha_i)}{1!} & \ldots & \frac{f^{(m-1)}(\alpha_i)}{(m-1)!} \\ 0 & f(\alpha_i) & \ldots & \frac{f^{(m-2)}(\alpha_i)}{(m-2)!} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & f(\alpha_i) \end{pmatrix},$$

где m – размер рассматриваемой жордановой клетки.

Надо доказать корректность этого определения, т. е. показать, что оно не зависит от T и J.

Т е о р е м а 2. *Пусть для функции $f : \mathbb{C} \longrightarrow \mathbb{C}$ и матрицы $A \in M_n(\mathbb{C})$ выполнено условие* (1). *Тогда* 1) *существует многочлен $p(\lambda) \in \mathbb{C}[\lambda]$, такой, что*

$$p^{(k)}(\alpha_i) = f^{(k)}(\alpha_i), \quad k = 0, 1, \ldots, n_{ij} - 1, \quad i = 1, 2, \ldots, s; \tag{2}$$

2) $p(A) = f(A)$, *и поэтому $f(A)$ от случайного выбора J и T не зависит.*

Д о к а з а т е л ь с т в о. Если доказана первая часть теоремы, то из нее следует и вторая часть. Докажем первую часть.

По значениям многочлена мы можем восстановить многочлен, используя формулу Лагранжа, но у нас есть и значения производных. Будем искать многочлен $p(\lambda)$ в виде

$$p(\lambda) = \sum_{i=1}^{s} \left\{ \left[\beta_{i0} + \beta_{i1}(\lambda - \alpha_i) + \ldots + \beta_{i,n_i-1}(\lambda - \alpha_i)^{n_i-1} \right] \cdot \right.$$

$$\left. \cdot (\lambda - \alpha_1)^{n_1} \ldots (\lambda - \alpha_{i-1})^{n_{i-1}} (\lambda - \alpha_{i+1})^{n_{i+1}} \ldots (\lambda - \alpha_s)^{n_s} \right\},$$

где $n_i = \max_j n_{ij}$. Покажем, что коэффициенты β_{ij} можно подобрать так, чтобы выполнялось условие (2). Подставив в нашу формулу значение α_i, получим

$$p(\alpha_i) = \beta_{i0} A_{i0}^{(0)},$$

где

$$A_{i0}^{(0)} = (\alpha_i - \alpha_1)^{n_1} \ldots (\alpha_i - \alpha_{i-1})^{n_{i-1}} (\alpha_i - \alpha_{i+1})^{n_{i+1}} \ldots (\alpha_i - \alpha_s)^{n_s} \neq 0,$$

т. е. от всей суммы остается только слагаемое с номером i, от выражения в квадратных скобках остается только коэффициент β_{i0}.

Найдем значение производной в точке α_i. Для этого обозначим

$$u(\lambda) = \left[\beta_{i0} + \beta_{i1}(\lambda - \alpha_i) + \ldots + \beta_{i,n_i-1}(\lambda - \alpha_i)^{n_i-1} \right],$$

$$v(\lambda) = (\lambda - \alpha_1)^{n_1} \ldots (\lambda - \alpha_{i-1})^{n_{i-1}} (\lambda - \alpha_{i+1})^{n_{i+1}} \ldots (\lambda - \alpha_s)^{n_s}.$$

Нетрудно убедиться, что

$$p'(\alpha_i) = u(\alpha_i)\, v'(\alpha_i) + u'(\alpha_i)\, v(\alpha_i) = \beta_{i0} A_{i0}^{(1)} + \beta_{i1} A_{i1}^{(1)}, \text{ где } A_{i1}^{(1)} = A_{i0}^{(0)}.$$

Далее находим вторую производную:

$$(u'v + uv')' = uv'' + 2u'v' + u''v,$$

и, подставляя значение α_i, получим

$$p''(\alpha_i) = u(\alpha_i)v''(\alpha_i) + 2u'(\alpha_i)v'(\alpha_i) + u''(\alpha_i)v(\alpha_i).$$

Аналогичным образом приходим к формулам

$$p''(\alpha_i) = \beta_{i0} A_{i0}^{(2)} + \beta_{i1} A_{i1}^{(2)} + 2\beta_{i2} A_{i2}^{(2)}, \text{ где } A_{i2}^{(2)} = A_{i0}^{(0)},$$

.....

$$p^{(k)}(\alpha_i) = \beta_{i0} A_{i0}^{(k)} + \beta_{i1} A_{i1}^{(k)} + \ldots + k!\beta_{ik} A_{ik}^{(k)}, \text{ где } A_{ik}^{(k)} = A_{i0}^{(0)},$$

.....

$$p^{(n_i-1)}(\alpha_i) = \beta_{i0} A_{i0}^{(n_i-1)} + \beta_{i1} A_{i1}^{(n_i-1)} + \ldots + (n_i-1)!\beta_{i,n_i-1} A_{i,n_i-1}^{(n_i-1)},$$

$$\text{где } A_{i,n_i-1}^{(n_i-1)} = A_{i0}^{(0)}.$$

Это треугольная система линейных уравнений относительно β_{ij}. Решая ее, найдем β_{ij}. Теорема доказана.

В этой теореме мы интерполировали функцию по значениям многочлена и значениям производной. Построенный многочлен $p(\lambda)$ называется *многочленом Лагранжа – Сильвестера.*

36.3. Ряды от матриц. Рассмотрим ряд

$$a_0 + a_1\lambda + a_2\lambda^2 + \ldots, \quad a_i \in \mathbb{R}.$$

Пусть $A \in M_n(\mathbb{R})$. Можем определить ряд

$$a_0E + a_1A + a_2A^2 + \ldots$$

Его частичной суммой называется матрица

$$f_k(A) = a_0E + a_1A + a_2A^2 + \ldots + a_kA^k.$$

Если существуют пределы

$$\lim_{k\to\infty} (f_k(A))_{ij} = b_{ij}, \quad 1 \le i, j \le n,$$

то матрица $B = (b_{ij})$ называется *суммой ряда.*

Если A – жорданова клетка

$$A = \begin{pmatrix} \alpha & 1 & & & \mathbf{0} \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha \end{pmatrix},$$

то

$$f_k(A) = \begin{pmatrix} f_k(\alpha) & \frac{f_k'(\alpha)}{1!} & \ldots & \frac{f_k^{(m-1)}(\alpha)}{(m-1)!} \\ 0 & f_k(\alpha) & \ldots & \frac{f_k^{(m-2)}(\alpha)}{(m-2)!} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & f_k(\alpha) \end{pmatrix}.$$

Тогда из теоремы 2 получается

Т е о р е м а 3. *Ряд*

$$\sum_{k=0}^{\infty} a_kA^k$$

сходится тогда и только тогда, когда сходится ряд

$$\sum_{k=0}^{\infty} a_k \lambda^k$$

и все его производные до $(n_i - 1)$*-го порядка в точке* α_i*, где* $(\lambda - \alpha_i)^{n_{ij}}$ *– элементарные делители матрицы* A*,* $n_i = \max_j n_{ij}$.

П р и м е р. Мы знаем, что

$$\mathrm{e}^{\lambda} = 1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots$$

Для матрицы A положим

$$\mathrm{e}^{A} = E + \frac{1}{1!}A + \frac{1}{2!}A^2 + \dots$$

По теореме 3 этот ряд сходится.

У п р а ж н е н и е. Если $AB = BA$, то $\mathrm{e}^A \cdot \mathrm{e}^B = \mathrm{e}^{A+B}$.

Рассмотрим множество матриц

$$\{At \mid t \in \mathbb{R}\}.$$

По упражнению

$$\mathrm{e}^{At} \cdot \mathrm{e}^{As} = \mathrm{e}^{A(t+s)},$$

т. е. отображение $t \longmapsto \mathrm{e}^{At}$ переводит сумму в произведение. В частности, множество

$$\{\mathrm{e}^{At} \mid t \in \mathbb{R}\}$$

образует подгруппу по умножению. Она называется *однопараметрической подгруппой.* Группу $GL_n(\mathbb{R})$ можно получить как конечное произведение однопараметрических подгрупп (группа Ли).

36.4. Приложение теории функций от матриц к решению систем линейных дифференциальных уравнений. Рассмотрим простейшее дифференциальное уравнение

$$y' = ay, \quad a \in \mathbb{R},$$

где $y = y(x)$ – неизвестная функция. Решим его методом разделения переменных. Из уравнения

$$\frac{\mathrm{d}y}{\mathrm{d}x} = ay$$

получим

$$\frac{\mathrm{d}y}{y} = a\mathrm{d}x.$$

Интегрируя это равенство, найдем общее решение:

$$y = c\mathrm{e}^{ax},$$

где c – произвольная постоянная.

Рассмотрим теперь систему обыкновенных дифференциальных уравнений с постоянными коэффициентами:

$$Y' = AY,$$

где

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad Y' = \begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{pmatrix}, \quad A \in M_n(\mathbb{R}).$$

Тогда легко проверить, что ее решение имеет вид

$$Y = \mathrm{e}^{Ax}C,$$

где e^{Ax} – функция от матрицы Ax, а C – вектор-столбец постоянных:

$$C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}.$$

Глава 7

ЕВКЛИДОВЫ И УНИТАРНЫЕ ПРОСТРАНСТВА

§ 37. Свойства евклидовых и унитарных пространств

37.1. Аксиоматика и примеры. Будем рассматривать векторное пространство V либо над полем вещественных чисел, либо над полем комплексных чисел. Вводя новую операцию скалярного произведения

$$(,) : V \times V \longrightarrow P,$$

получим в первом случае евклидово пространство, а во втором – унитарное. Определения этих пространств очень похожи, и мы будем давать их параллельно, рассматривая в левой части страницы случай поля вещественных чисел, а в правой – случай поля комплексных чисел.

Евклидово пространство	*Унитарное пространство*
− векторное пространство над $\mathbb{R}$	− векторное пространство над $\mathbb{C}$
$(\ ,\ ) : V \times V \longrightarrow \mathbb{R}$.	$(\ ,\ ) : V \times V \longrightarrow \mathbb{C}$.
Аксиомы:	Аксиомы:
1. $(a, b) = (b, a)$,	1. $(a, b) = \overline{(b, a)}$,
2. $(a' + a'', b) = (a', b) + (a'', b)$,	2. $(a' + a'', b) = (a', b) + (a'', b)$,
3. $(\alpha a, b) = \alpha(a, b)$, $\alpha \in \mathbb{R}$	3. $(\alpha a, b) = \alpha(a, b)$, $\alpha \in \mathbb{C}$
4. $(a, a) > 0$, при $a \neq 0$,	4. $(a, a) > 0$, при $a \neq 0$.
Пример:	Пример:
$(a, b) = \sum_{i=1}^{n} \alpha_i \beta_i$, где $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, $b = (\beta_1, \beta_2, \ldots, \beta_n)$.	$(a, b) = \sum_{i=1}^{n} \alpha_i \overline{\beta_i}$, где $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, $b = (\beta_1, \beta_2, \ldots, \beta_n)$.
Следствия из аксиом:	Следствия из аксиом:
$2'$. $(a, b' + b'') = (a, b') + (a, b'')$,	$2'$. $(a, b' + b'') = (a, b') + (a, b'')$,
$3'$. $(a, \beta b) = \beta(a, b)$,	$3'$. $(a, \beta b) = \overline{\beta}(a, b)$.

Докажем $3'$ для унитарного пространства. Для евклидова пространства годится это же рассуждение. Нужное равенство следует из цепочки, где над знаком равенства указан номер соответствующей аксиомы:

$$(a, \beta b) \overset{1}{=} \overline{(\beta b, a)} \overset{3}{=} \overline{\beta(b, a)} = \overline{\beta}\,\overline{(b, a)} \overset{1}{=} \overline{\beta}(a, b).$$

П р и м е р скалярного произведения на пространстве функций. Пусть V – множество веществознозначных (комплекснознач-

ных) функций, непрерывных на отрезке $[0,1]$. Тогда

$$(f,g)=\int_0^1 f(t)g(t)\mathrm{d}t,\quad \left(\text{соответственно } (f,g)=\int_0^1 f(t)\overline{g(t)}\mathrm{d}t\right).$$

При изучении евклидовых и унитарных пространств будем рассматривать, в основном, евклидовы пространства. Соответствующие определения и теоремы без особого труда переносятся и на унитарные пространства.

37.2. Ортонормированные системы векторов.

О п р е д е л е н и е. Пусть V – евклидово пространство. Векторы a и b из V называются *ортогональными*, если $(a,b)=0$.

О п р е д е л е н и е. Вектор $a\in V$ называется *нормированным*, если $(a,a)=1$.

О п р е д е л е н и е. Система векторов $a_1,a_2,\ldots,a_s$ называется *ортонормированной*, если

$$(a_i,a_j)=\begin{cases}1 & \text{при } i=j,\\ 0 & \text{при } i\neq j.\end{cases}$$

Т е о р е м а 1. 1) *Всякая ортонормированная система векторов линейно независима.* 2) *Всякую линейно независимую систему векторов можно переделать в ортонормированную с той же самой линейной оболочкой.* 3) *В ортонормированной базе для векторов* $a=(\alpha_1,\alpha_2,\ldots,\alpha_n)$ *и* $b=(\beta_1,\beta_2,\ldots,\beta_n)$ *скалярное произведение определяется равенством*

$$(a,b)=\sum_{i=1}^{n}\alpha_i\beta_i. \tag{1}$$

4) *Если в какой-то базе скалярное произведение задается формулой* (1), *то эта база ортонормированная.*

Д о к а з а т е л ь с т в о. 1) Пусть $a_1,a_2,\ldots,a_s$ – ортонормированная система векторов. Рассмотрим их линейную комбинацию и приравняем к нулю:

$$\alpha_1a_1+\alpha_2a_2+\ldots+\alpha_sa_s=0.$$

Умножим обе части скалярно на a_i:

$$\alpha_1(a_1,a_i)+\alpha_2(a_2,a_i)+\ldots+\alpha_s(a_s,a_i)=(0,a_i).$$

Учитывая, что

$$(a_i, a_j) = \begin{cases} 1 & \text{при } i = j, \\ 0 & \text{при } i \neq j, \end{cases}$$

получим равенство

$$\alpha_i = (0, a_i).$$

Из следующей цепочки равенств

$$(0, a) = (b - b, a) \overset{2}{=} (b, a) + (-b, a) \overset{3}{=} (b, a) - (b, a) = 0$$

заключаем, что $(0, a_i) = 0$, а потому все α_i равны нулю, т. е. ортонормированная система линейно независима.

2) Пусть $a_1, a_2, \ldots, a_s$ – линейно независимая система векторов. Приведем эту систему к ортогональной с той же линейной оболочкой. В качестве первого вектора новой системы возьмем $b_1 = a_1$. Допустим, что попарно ортогональные векторы b_1, b_2, $\ldots$, b_k уже построены и при этом справедлива система равенств

$$\begin{cases} b_1 & = a_1, \\ \beta_{21} b_1 + b_2 & = a_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ \beta_{k1} b_1 + \beta_{k2} b_2 + \ldots + \beta_{k,k-1} b_{k-1} + b_k & = a_k. \end{cases}$$

Ищем вектор b_{k+1} из соотношения

$$\beta_{k+1,1} b_1 + \beta_{k+1,2} b_2 + \ldots + \beta_{k+1,k} b_k + b_{k+1} = a_{k+1}.$$

Требуется, чтобы $(b_{k+1}, b_j) = 0$ при $j = 1, 2, \ldots, k$. Умножим обе части нашего равенства скалярно на b_j, получим

$$\beta_{k+1,1}(b_1, b_j) + \ldots + \beta_{k+1,j}(b_j, b_j) + \ldots + (b_{k+1}, b_j) = (a_{k+1}, b_j).$$

Отсюда находим:

$$\beta_{k+1,j} = \frac{(a_{k+1}, b_j)}{(b_j, b_j)}.$$

Надо заметить, что все b_j отличны от нуля. Если $b_j = 0$, то в j-й строке нашей системы стоит равенство

$$\beta_{j1} b_1 + \beta_{j2} b_2 + \ldots + \beta_{j,j-1} b_{j-1} = a_j,$$

а так как векторы $b_1, b_2, \ldots, b_{j-1}$ выражаются через векторы a_1, a_2, $\ldots$, a_{j-1}, то приходим к равенству

$$a_j = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_{j-1} a_{j-1},$$

которое противоречит тому, что векторы $a_1, a_2, \ldots, a_s$ линейно независимы.

Нормируя векторы $b_1, b_2, \ldots, b_s$, получим искомую систему векторов:

$$c_i = \frac{1}{\sqrt{(b_i, b_i)}} b_i, \quad i = 1, 2, \ldots, s.$$

3) Пусть $e_1, e_2, \ldots, e_n$ – ортонормированная база пространства V. Рассмотрим два вектора:

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

$$b = \beta_1 e_1 + \beta_2 e_2 + \ldots + \beta_n e_n.$$

Тогда

$$(a, b) = \left(\sum_{i=1}^{n} \alpha_i e_i, \sum_{j=1}^{n} \beta_j e_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j (e_i, e_j) = \sum_{i=1}^{n} \alpha_i \beta_i.$$

4) Пусть $e_1, e_2, \ldots, e_n$ – база пространства V, в которой скалярное произведение определяется равенством (1). Рассмотрим координаты двух базисных векторов:

$$[e_i] = (0, \ldots, 0, 1, 0, \ldots, 0),$$

где 1 стоит на i-м месте,

$$[e_j] = (0, \ldots, 0, 1, 0, \ldots, 0),$$

где 1 стоит на j-м месте. Тогда

$$(e_i, e_j) = \delta_{ij} = \begin{cases} 1 \text{ при } i = j, \\ 0 \text{ при } i \neq j, \end{cases}$$

но это и означает, что $e_1, e_2, \ldots, e_n$ образуют ортонормированную базу. Теорема доказана.

Процесс, описанный в пункте 2, называется *процессом ортогонализации Грама – Шмидта.*

37.3. Изоморфизм евклидовых пространств.

О п р е д е л е н и е. Евклидовы пространства V и V' *изоморфны*, если существует взаимно однозначное отображение $\varphi : V \xrightarrow{\text{на}} V'$, удовлетворяющее условиям:

1) $(a+b)\varphi = a\varphi + b\varphi$;
2) $(\alpha a)\varphi = \alpha(a\varphi)$;
3) $(a\varphi, b\varphi) = (a, b)$;
для всех $\alpha \in \mathbb{R}$, $a, b \in V$.

Т е о р е м а 2. *Евклидовы пространства изоморфны тогда и только тогда, когда они имеют одинаковую размерность.*

Д о к а з а т е л ь с т в о. $\Rightarrow$. Если евклидовы пространства изоморфны, то они изоморфны как векторные пространства. Следовательно, их размерности совпадают.

$\Leftarrow$. Пусть $\dim V = \dim V'$ и $e_1, e_2, \ldots, e_n$ – ортонормированная база V; $e'_1, e'_2, \ldots, e'_n$ – ортонормированная база V'. Рассмотрим отображение

$$\varphi : \sum_{i=1}^{n} \alpha_i e_i \longmapsto \sum_{i=1}^{n} \alpha_i e'_i.$$

Это отображение однозначно и *на*, кроме того, если

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n, \quad b = \beta_1 e_1 + \beta_2 e_2 + \ldots + \beta_n e_n$$

– два вектора из V, то

$$(a, b) = \left(\sum_{i=1}^{n} \alpha_i e_i, \sum_{j=1}^{n} \beta_j e_j \right) = \sum_{i=1}^{n} \alpha_i \beta_i,$$

$$(a\varphi, b\varphi) = \left(\sum_{i=1}^{n} \alpha_i e'_i, \sum_{j=1}^{n} \beta_j e'_j \right) = \sum_{i=1}^{n} \alpha_i \beta_i,$$

т. е. $(a, b) = (a\varphi, b\varphi)$. Теорема доказана.

С л е д с т в и е. *Любое евклидово пространство изоморфно* $\mathbb{R}^n$ *для некоторого натурального* n.

37.4. Норма вектора. Пусть V – евклидово пространство.

О п р е д е л е н и е. *Нормой вектора* u из V называется число $||u|| = \sqrt{(u, u)}$.

Л е м м а 1. *Норма вектора обладает следующими свойствами*:
а) $|(u, v)| \leq ||u|| \cdot ||v||$ – *неравенство Коши – Буняковского*;
б) $||u + v|| \leq ||u|| + ||v||$ – *неравенство треугольника*;
в) $||\alpha u|| = |\alpha| \cdot ||u||$;
для любых векторов u, v *и скаляра* $\alpha \in \mathbb{R}$.

Д о к а з а т е л ь с т в о. а) Для любого $\lambda \in \mathbb{R}$ имеем

$$0 \leq (u + \lambda v, u + \lambda v) = (u, u) + 2\lambda(u, v) + \lambda^2 (v, v).$$

Это квадратный трехчлен относительно λ. Так как он принимает неотрицательные значения, то его дискриминант меньше либо равен нулю, т. е.

$$4(u,v)^2 - 4(u,u)(v,v) \le 0,$$

откуда

$$|(u,v)| \le ||u|| \cdot ||v||.$$

Это и есть неравенство Коши – Буняковского.

б) Для любых u, v из V имеем

$$||u+v||^2 = (u+v, u+v) = (u,u) + 2(u,v) + (v,v) \le$$

$$\le ||u||^2 + 2||u|| \cdot ||v|| + ||v||^2 = (||u|| + ||v||)^2,$$

откуда

$$||u+v|| \le ||u|| + ||v||.$$

в) Из цепочки равенств

$$||\alpha u||^2 = (\alpha u, \alpha u) = \alpha^2 (u,u) = \alpha^2 ||u||^2$$

имеем

$$||\alpha u|| = |\alpha| \cdot ||u||.$$

Лемма доказана.

Неравенство Коши – Буняковского в координатной форме:

$$\left| \sum_{i=1}^{n} \alpha_i \beta_i \right| \le \sqrt{\sum_{i=1}^{n} \alpha_i^2} \cdot \sqrt{\sum_{i=1}^{n} \beta_i^2}.$$

Л е м м а 2. *Любое линейное преобразование φ евклидова пространства V ограничено, т. е. существует $M \in \mathbb{R}$ такое, что*

$$||u\varphi|| \le M||u||$$

для всех u из V.

Д о к а з а т е л ь с т в о. Пусть $e_1, e_2, \ldots, e_n$ – ортонормированная база V. Действуя преобразованием φ, получим

$$e_i\varphi = \sum_{j=1}^{n} \alpha_{ij} e_j, \quad i = 1, 2, \ldots, n.$$

Выберем вектор

$$u = \sum_{i=1}^{n} \xi_i e_i$$

из V и подействуем на него φ:

$$u\varphi = \sum_{i=1}^{n} \xi_i (e_i\varphi) = \sum_{i=1}^{n} \xi_i \left(\sum_{j=1}^{n} \alpha_{ij} e_j \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \xi_i \alpha_{ij} \right) e_j.$$

Отсюда

$$||u\varphi||^2 = \sum_{j=1}^{n} \left| \sum_{i=1}^{n} \xi_i \alpha_{ij} \right|^2,$$

и ввиду неравенства Коши – Буняковского

$$\sum_{j=1}^{n} \left| \sum_{i=1}^{n} \xi_i \alpha_{ij} \right|^2 \leq \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \xi_i^2 \cdot \sum_{i=1}^{n} \alpha_{ij}^2 \right) = ||u||^2 \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij}^2 =$$

$$= ||u||^2 \sum_{i,j=1}^{n} \alpha_{ij}^2.$$

Поэтому

$$||u\varphi|| \leq M \cdot ||u||,$$

где

$$M = \sqrt{\sum_{i,j=1}^{n} \alpha_{ij}^2}.$$

Лемма доказана.

37.5. Норма линейного преобразования. Из доказанной леммы следует, что отношение

$$\frac{||u\varphi||}{||u||}$$

ограничено. Поэтому мы можем дать такое

О п р е д е л е н и е. *Нормой линейного преобразования* φ евклидова пространства V называется число

$$||\varphi|| = \sup_{u \neq 0} \frac{||u\varphi||}{||u||} = \sup_{||u||=1} ||u\varphi||.$$

Л е м м а 3. *Если φ и ψ – линейные преобразования евклидова пространства V, то*

а) $||\varphi + \psi|| \leq ||\varphi|| + ||\psi||$;

б) $||\varphi \cdot \psi|| \leq ||\varphi|| \cdot ||\psi||$;

в) $||\alpha\varphi|| = |\alpha| \cdot ||\varphi||$;

г) $||\varphi|| > 0$ если $\varphi \neq 0$ и $||0|| = 0$;

д) *если α – собственное значение φ, то* $||\varphi|| \geq |\alpha|$.

Д о к а з а т е л ь с т в о. Пункты а–в следуют из ниже приведенных цепочек равенств и неравенств:

а) $||\varphi + \psi|| = \sup\limits_{||u||=1} ||u(\varphi + \psi)|| = \sup\limits_{||u||=1} ||u\varphi + u\psi|| \leq \sup\limits_{||u||=1} ||u\varphi|| +$
$+ \sup\limits_{||u||=1} ||u\psi|| = ||\varphi|| + ||\psi||$;

б) $||\varphi \cdot \psi|| = \sup\limits_{||u||=1} ||(u\varphi)\psi|| \leq ||\psi|| \sup\limits_{||u||=1} ||u\varphi|| = ||\varphi|| \cdot ||\psi||$;

в) $||\alpha\varphi|| = \sup\limits_{||u||=1} ||u(\alpha\varphi)|| = |\alpha| \sup\limits_{||u||=1} ||u\varphi|| = |\alpha|||\varphi||$.

г) Пусть $w\varphi \neq 0$. Тогда

$$\frac{||w\varphi||}{||w||} > 0,$$

и следовательно $||\varphi|| > 0$.

д) Пусть $w\varphi = \alpha w$. Тогда $||w\varphi|| = ||\alpha w|| = |\alpha|||w||$ и, деля на $||w||$, получим

$$\frac{||w\varphi||}{||w||} = \frac{|\alpha|||w||}{||w||} = |\alpha|.$$

Лемма доказана.

37.6. Сопряженные отображения. Пусть V, W – два евклидовых пространства и

$$\varphi : V \longrightarrow W$$

– линейное отображение одного на другое.

О п р е д е л е н и е. Линейное отображение $\varphi^* : W \longrightarrow V$ называется *сопряженным* к φ, если для любых $x \in V$ и $y \in W$ справедливо равенство

$$(x\varphi, y) = (x, y\varphi^*).$$

Т е о р е м а 3. *Для всякого линейного отображения $\varphi : V \longrightarrow W$ существует единственное сопряженное отображение $\varphi^* : W \longrightarrow V$. Если в V и W фиксированы ортонормированные базы, то в них $[\varphi^*] = [\varphi]^t$.*

Д о к а з а т е л ь с т в о. Докажем вначале, что если φ^* существует, то оно единственно. Пусть $e_1, e_2, \ldots, e_n$ – ортонормированная база V. Если $x \in V$, то

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

Тогда $x_j = (x, e_j)$, т. е.

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

Далее имеем

$$(x, y\varphi^*) = (x\varphi, y) = \left(\sum_{i=1}^{n} (x, e_i)(e_i\varphi), y \right) = \sum_{i=1}^{n} (x, e_i)(e_i\varphi, y) =$$

$$= \sum_{i=1}^{n} (x, (e_i\varphi, y)e_i) = \left(x, \sum_{i=1}^{n} (e_i\varphi, y)e_i \right).$$

Заметим, что если $(x, z) = (x, t)$ для любого вектора x из V, то $z = t$. Действительно, из равенства $(x, z) = (x, t)$ получим $(x, z - t) = 0$. Взяв в качестве x вектор $x = z - t$, из аксиом векторного пространства заключаем, что $z = t$.

Используя это свойство, видим, что

$$y\varphi^* = \sum_{i=1}^{n} (e_i\varphi, y)e_i,$$

и единственность доказана.

Докажем существование. Положим

$$y\varphi^* = \sum_{i=1}^{n} (e_i\varphi, y)e_i \text{ для любого } y \in W.$$

Проверим, что определенное таким образом отображение φ^* линейно. Имеем

$$(\alpha y_1 + \beta y_2)\varphi^* = \sum_{i=1}^{n}(e_i\varphi, \alpha y_1 + \beta y_2)e_i = \sum_{i=1}^{n}\left[\alpha(e_i\varphi, y_1) + \beta(e_i\varphi, y_2)\right]e_i =$$

$$= \alpha(y_1\varphi^*) + \beta(y_2\varphi^*).$$

Далее

$$(x, y\varphi^*) = \left(x, \sum_{i=1}^{n}(e_i\varphi, y)e_i\right) = \sum_{i=1}^{n}(x, e_i)(e_i\varphi, y) =$$

$$= \sum_{i=1}^{n}((x, e_i)(e_i\varphi), y) = (x\varphi, y)$$

и существование установлено.

Докажем вторую часть теоремы. Пусть

$$e_1, e_2, \ldots, e_n - \text{ ортонормированная база } V;$$

$$f_1, f_2, \ldots, f_m - \text{ ортонормированная база } W.$$

Действуя на эти базы преобразованиями φ и φ^* соответственно, получим

$$e_i\varphi = \sum_{j=1}^{m}\alpha_{ij}f_j, \quad i = 1, 2, \ldots, n,$$

т. е. $[\varphi] = (\alpha_{kl})$;

$$f_j\varphi^* = \sum_{i=1}^{n}\beta_{ji}e_i, \quad j = 1, 2, \ldots, m,$$

т. е. $[\varphi^*] = (\beta_{kl})$. Из этих равенств

$$(e_k\varphi, f_l) = \left(\sum_{j=1}^{m}\alpha_{kj}f_j, f_l\right) = \alpha_{kl},$$

$$(e_k, f_l\varphi^*) = \left(e_k, \sum_{i=1}^{n}\beta_{li}e_i\right) = \beta_{lk}.$$

Учитывая равенство $(e_k\varphi, f_l) = (e_k, f_l\varphi^*)$, заключаем, что $\alpha_{kl} = \beta_{lk}$. Следовательно, $[\varphi^*] = [\varphi]'$. Теорема доказана.

У п р а ж н е н и е. Для сопряженных отображений справедливы равенства:

1) $(\varphi^*)^* = \varphi$;

2) $(\varphi + \psi)^* = \varphi^* + \psi^*$;

3) $(\varphi \cdot \psi)^* = \psi^* \cdot \varphi^*$;

4) $(\alpha\varphi)^* = \alpha\varphi^*$;

5) $(\varphi^{-1})^* = (\varphi^*)^{-1}$.

§ 38. Ортогональные преобразования

38.1. Определения и свойства ортогональных преобразований.

Т е о р е м а 1. *Пусть V – евклидово пространство, φ – его линейное преобразование. Тогда следующие условия равносильны*:

а) $||a\varphi|| = ||a||$ *для любого* a *из* V;

б) φ *сохраняет скалярное произведение*: $(a\varphi, b\varphi) = (a, b)$ *для любых* a *и* b *из* V;

в) φ *переводит всякую ортонормированную базу в ортонормированную*;

г) φ *переводит некоторую ортонормированную базу в ортонормированную*;

д) *в некоторой ортонормированной базе матрица* $[\varphi]$ *ортогональна* (т. е. $[\varphi][\varphi]^t = E$);

е) *во всякой ортонормированной базе матрица* $[\varphi]$ *ортогональна.*

При любом из этих условий преобразование φ называется *ортогональным.*

Д о к а з а т е л ь с т в о. а) $\Rightarrow$ б). Имеем

$$||(a + b)\varphi|| = ||a + b||.$$

По определению нормы отсюда следует равенство

$$((a + b)\varphi, (a + b)\varphi) = (a + b, a + b).$$

Воспользовавшись линейностью преобразования φ и свойством скалярного произведения, получим

$$(a\varphi, a\varphi) + 2(a\varphi, b\varphi) + (b\varphi, b\varphi) = (a, a) + 2(a, b) + (b, b).$$

Учитывая, что норма вектора сохраняется, имеем равенства

$$(a\varphi, a\varphi) = (a, a), \quad (b\varphi, b\varphi) = (b, b).$$

Следовательно,

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

б) $\Rightarrow$ в). Пусть $e_1, e_2, \ldots, e_n$ – ортонормированная база, т. е. $(e_i, e_j) = \delta_{ij}$. Тогда

$$(e_i\varphi, e_j\varphi) = (e_i, e_j) = \delta_{ij},$$

т. е. $e_1\varphi, e_2\varphi, \ldots, e_n\varphi$ – ортонормированная база.

в) $\Rightarrow$ г). Тривиально.

г) $\Rightarrow$ д). Пусть $e_1, e_2, \ldots, e_n$ – ортонормированная база и $e_1\varphi$, $e_2\varphi$, $\ldots$, $e_n\varphi$ – ортонормированная база. Тогда

$$e_i\varphi = \sum_{k=1}^{n} \alpha_{ik} e_k$$

и надо показать, что матрица $A = (\alpha_{ik})$ ортогональна. Имеем

$$\delta_{ij} = (e_i\varphi, e_j\varphi) = \left(\sum_{k=1}^{n} \alpha_{ik} e_k, \sum_{l=1}^{n} \alpha_{jl} e_l \right) = \sum_{k,l=1}^{n} \alpha_{ik}\alpha_{jl}(e_k, e_l) = \sum_{k=1}^{n} \alpha_{ik}\alpha_{jk},$$

но это и означает, что $A \cdot A^t = E$, т. е. матрица $A = [\varphi]$ ортогональна.

д) $\Rightarrow$ е). Пусть $e_1, e_2, \ldots, e_n$ – ортонормированная база

$$e_i\varphi = \sum_{k=1}^{n} \alpha_{ik} e_k, \quad i = 1, 2, \ldots, n,$$

и матрица $[\varphi] = (\alpha_{ik})$ ортогональна. Пусть $e'_1, e'_2, \ldots, e'_n$ – другая ортонормированная база. Тогда $e' = Te$, где T – матрица перехода. Мы знаем, что $[\varphi]$ – матрица преобразования φ в базе e ортогональна. Надо доказать, что $[\varphi]'$ – матрица преобразования φ в базе e' тоже ортогональна. Вспоминая связь между матрицами преобразования в разных базах, имеем равенство

$$[\varphi]' = T[\varphi]T^{-1}.$$

Если при этом T – ортогональная матрица, то и $[\varphi]'$ будет ортогональной, так как ортогональные матрицы образуют группу. Имеем систему равенств

$$e'_i = \sum_{k=1}^{n} t_{ik} e_k, \quad i = 1, 2, \ldots, n, \quad T = (t_{ik}).$$

Учитывая, что e' – ортонормированная база,

$$(e'_i, e'_j) = \delta_{ij},$$

получим

$$\delta_{ij} = (e'_i, e'_j) = \left(\sum_{k=1}^{n} t_{ik} e_k, \sum_{l=1}^{n} t_{jl} e_k, \right) = \sum_{k=1}^{n} \sum_{l=1}^{n} t_{ik} t_{jl} (e_k, e_l) = \sum_{k=1}^{n} t_{ik} t_{jk},$$

т. е. $E = TT^t$, но это и означает, что матрица T ортогональна.

е) $\Rightarrow$ а). Пусть $e_1, e_2, \ldots, e_n$ – ортонормированная база и в ней матрица $[\varphi]$ ортогональна. Возьмем два вектора $a = [a]e$, $b = [b]e$, где

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

и рассмотрим их скалярное произведение:

$$(a, b) = \sum_{i=1}^{n} \alpha_i \beta_i = [a][b]^t.$$

С другой стороны,

$$||a\varphi||^2 = (a\varphi, a\varphi) = [a\varphi]\,[a\varphi]^t = ([a][\varphi])\,([a][\varphi])^t = [a]\,[\varphi]\,[\varphi]^t\,[a]^t =$$

$$= [a]\,[a]^t = (a, a) = ||a||^2.$$

Следовательно, $||a\varphi|| = ||a||$. Теорема доказана.

Как следует из этой теоремы, геометрический смысл ортогонального преобразования в том, что оно сохраняет длины векторов.

38.2. Канонический вид матрицы ортогонального преобразования.

Т е о р е м а 2. *Для всякого ортогонального преобразования φ евклидова пространства V существует такая ортонормированная база, в которой матрица $[\varphi]$ является клеточно-диагональной*

и каждая клетка либо одномерна (и имеет вид ± 1*), либо двумерна и имеет вид*

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \quad \alpha \in \mathbb{R},$$

т. е.

$$[\varphi] = E_k \oplus (-E_l) \oplus \begin{pmatrix} \cos\alpha_1 & -\sin\alpha_1 \\ \sin\alpha_1 & \cos\alpha_1 \end{pmatrix} \oplus \ldots \oplus \begin{pmatrix} \cos\alpha_m & -\sin\alpha_m \\ \sin\alpha_m & \cos\alpha_m \end{pmatrix},$$

где $k + l + 2m = n$.

Иными словами, всякое ортогональное преобразование сводится к нескольким поворотам и отражениям.

Д о к а з а т е л ь с т в о. В теореме о полупростоте над полем действительных чисел мы доказывали, что если некоторое подпространство векторного пространства над полем $\mathbb{R}$ неприводимо относительно линейного преобразования, то его размерность ≤ 2. Покажем, что каждая матрица над полем действительных чисел в подходящей базе имеет вид

$$\left(\begin{array}{c|c|c|c} * & * & * & * \\ \hline 0 & * & * & * \\ \hline 0 & 0 & \ddots & * \\ \hline 0 & 0 & 0 & * \end{array}\right),$$

где диагональные клетки имеют размерность 1 или 2.

Доказательство проведем индукцией по размерности $\dim V = n$. Если $\dim V = 1$, то доказывать нечего. Пусть далее $\dim V > 1$. Мы знаем, что нулевое подпространство и все пространство V инвариантны относительно φ. Если между ними есть собственное инвариантное подпространство U:

$$0 < U < V,$$

то, как мы знаем, в некоторой базе матрица $[\varphi]$ имеет вид

$$[\varphi] = \begin{pmatrix} [\overline{\varphi}_U] & * \\ 0 & [\varphi_U] \end{pmatrix}.$$

Будем строить цепочку инвариантных подпространств пространства V до тех пор, пока не получим неуплотняемую цепочку

$$0 < U_1 < U_2 < \ldots < U_m = V$$

инвариантных подпространств. Фактор-пространство U_{i+1}/U_i неприводимо, а как мы знаем, неприводимое подпространство в нашем случае не более чем двумерно. Получим матрицу нужного вида.

Выберем базу $a_1, a_2, \ldots, a_n$, согласованную с этой цепочкой. Построим по ней ортонормированную базу:

$$\begin{aligned} b_n &= \alpha_{nn}a_n, \\ b_{n-1} &= \alpha_{n-1,n-1}a_{n-1} + \alpha_{n-1,n}a_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ b_1 &= \alpha_{11}a_1 + \alpha_{12}a_2 + \ldots + \alpha_{1n}a_n. \end{aligned}$$

Очевидно, что матрица перехода является треугольной матрицей. Найдем матрицу преобразования φ в базе b:

$$[\varphi]_b = T[\varphi]T^{-1}.$$

Процесс ортогонализации имеет треугольную матрицу. Матрица T треугольная, и ее можно считать клеточно-треугольной. Обратная к треугольной – опять треугольная.

Так как преобразование φ ортогонально, то матрица $[\varphi]_b$ ортогональная, т. е. $[\varphi]_b[\varphi]_b^t = E$, а потому обратная к ортогональной это просто транспонированная. Учитывая, что $[\varphi]_b$ клеточно-треугольная, заключаем, что она просто клеточно-диагональная. Следовательно, $[\varphi]_b$ – клеточно-диагональная, и каждая клетка является ортогональной. Учитывая, что они либо одномерны, либо двумерны, заключаем, что если клетка одномерна, т. е. равна элементу α, то $\alpha^2 = 1$, а потому $\alpha = \pm 1$. Предположим, что клетка двумерна и имеет вид

$$\begin{pmatrix} x & y \\ z & t \end{pmatrix}.$$

Из условия ортогональности следует равенство:

$$\begin{pmatrix} x & y \\ z & t \end{pmatrix}\begin{pmatrix} x & z \\ y & t \end{pmatrix} = \begin{pmatrix} x^2+y^2 & xz+yt \\ xz+yt & z^2+t^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

т. е.

$$\begin{cases} x^2+y^2=1, \\ z^2+t^2=1, \\ xz+yt=0, \end{cases}$$

откуда заключаем, что клетка имеет вид

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

Так как

$$xz + yt = \cos\alpha \sin\beta - \sin\alpha \cos\beta = \sin(\beta - \alpha) = 0,$$

то

$$\beta - \alpha = \pi k, \quad k \in \mathbb{Z},$$

а потому $\beta = \alpha + \pi k$. Если k – четно, то можно считать, что $\beta = \alpha$, и мы имеем клетку

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

Если k нечетно, то можно считать, что $\beta = \alpha + \pi$, и получим клетку

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

У п р а ж н е н и е. Почему в формулировке теоремы ничего не говорится о клетках второго типа?

§ 39. Симметрические преобразования

39.1. Определения и свойства симметрических преобразований.

Т е о р е м а 1. *Для линейного преобразования φ евклидова пространства V следующие условия равносильны*:

а) $(a\varphi, b) = (a, b\varphi)$ *для любых векторов a и b из V*;

б) *во всякой ортонормированной базе матрица $[\varphi]$ симметрическая* (т. е. $[\varphi] = [\varphi]^t$);

в) *в некоторой ортонормированной базе матрица $[\varphi]$ симметрическая.*

При любом из этих условий преобразование φ называется *симметрическим.*

Д о к а з а т е л ь с т в о. а) $\Rightarrow$ б). Пусть $e_1, e_2, \ldots, e_n$ – ортонормированная база пространства V. Действуя на нее преобразованием φ, получим

$$e_i\varphi = \sum_{k=1}^{n} \alpha_{ik} e_k, \quad i = 1, 2, \ldots, n.$$

Рассмотрим скалярное произведение:

$$(e_i\varphi, e_j) = \left(\sum_{k=1}^{n} \alpha_{ik} e_k, e_j \right) = \sum_{k=1}^{n} \alpha_{ik} (e_k, e_j) = \alpha_{ij}.$$

С другой стороны,

$$(e_i, e_j\varphi) = \left(e_i, \sum_{l=1}^{n} \alpha_{jl} e_l\right) = \sum_{l=1}^{n} \alpha_{jl}(e_i, e_l) = \alpha_{ji}.$$

Учитывая, что $(e_i\varphi, e_j) = (e_i, e_j\varphi)$, получаем равенство $\alpha_{ij} = \alpha_{ji}$, т. е. матрица $[\varphi]$ симметрическая.

б) $\Rightarrow$ в). Тривиально.

в) $\Rightarrow$ а). Пусть $e_1, e_2, \ldots, e_n$ – ортонормированная база, в которой матрица $[\varphi]$ симметрична. Тогда

$$(a\varphi, b) = [a\varphi]\,[b]^t = [a]\,[\varphi]\,[b]^t$$

и

$$(a, b\varphi) = [a]\,[b\varphi]^t = [a]\,([b]\,[\varphi])^t = [a]\,[\varphi]^t\,[b]^t.$$

Так как $[\varphi] = [\varphi]^t$, то $(a\varphi, b) = (a, b\varphi)$. Теорема доказана.

39.2. Характеристические корни симметрического преобразования.

Т е о р е м а 2. *Все характеристические корни симметрического преобразования являются действительными числами.*

Д о к а з а т е л ь с т в о. Пусть φ – симметрическое преобразование, а A – соответствующая ему матрица в ортонормированной базе. Тогда A – симметрическая матрица из $M_n(\mathbb{R})$. Рассмотрим унитарное пространство $\mathbb{C}^n$ с базой

$$\begin{gathered}(1, 0, \ldots, 0),\\ (0, 1, \ldots, 0),\\ \ldots\ldots\ldots\ldots\\ (0, 0, \ldots, 1)\end{gathered}$$

и скалярным произведением

$$(a, b) = \sum_{i=1}^{n} \alpha_i \overline{\beta}_i.$$

Определим линейное преобразование ψ пространства $\mathbb{C}^n$ по правилу:

$$\psi : (\alpha_1, \alpha_2, \ldots, \alpha_n) \longmapsto (\alpha_1, \alpha_2, \ldots, \alpha_n)A.$$

Покажем, что $(a\psi, b) = (a, b\psi)$ для любых a, b из $\mathbb{C}^n$. Действительно, имеем

$$(a\psi, b) = [a\psi]\,\overline{[b]}^t = ([a]\,[\psi])\,\overline{[b]}^t.$$

С другой стороны,

$$(a, b\psi) = [a]\,\overline{[b\psi]}^t = [a]\,(\overline{[b][\psi]}^t) = [a]\,\overline{[\psi]}^t\,\overline{[b]}^t,$$

но так как $[\psi] = A$ симметрическая и лежит в $\mathrm{M}_n(\mathbb{R})$, то $(a\psi, b) = (a, b\psi)$, т. е. ψ – симметрическое преобразование пространства $\mathbb{C}^n$.

Пусть α – характеристический корень матрицы A. Для ψ он является собственным значением, т. е. существует вектор $u \in \mathbb{C}^n$ такой, что $u\psi = \alpha u$. По предыдущему имеем

$$(u\psi, u) = (u, u\psi).$$

Но

$$(u\psi, u) = (\alpha u, u) = \alpha(u, u),$$
$$(u, u\psi) = (u, \alpha u) = \overline{\alpha}(u, u),$$

откуда

$$\alpha(u, u) = \overline{\alpha}(u, u).$$

Так как по определению $u \neq 0$, то $(u, u) > 0$ и $\alpha = \overline{\alpha}$, т. е. $\alpha \in \mathbb{R}$. Теорема доказана.

39.3. Канонический вид матрицы симметрического преобразования.

Т е о р е м а 3. *Для всякого симметрического преобразования φ евклидова пространства V существует ортонормированная база из собственных векторов.*

Д о к а з а т е л ь с т в о. 1) Докажем вначале, что существует некоторая база из собственных векторов (не заботясь об ортогональности). Докажем, что φ – полупросто. По теореме о разложении преобразования в сумму полупростого и нильпотентного имеем

$$\varphi = \psi + \theta,$$

где ψ – полупростое, θ – нильпотентное и $\psi\theta = \theta\psi$. Хотим доказать, что $\theta = 0$. Пусть $\theta \neq 0$, тогда $\theta^N = 0$, а $\theta^{N-1} \neq 0$ для некоторого натурального N. Очевидно, что $N \geq 2$. Выберем вектор $v \in V$ такой, что $v\theta^{N-1} \neq 0$.

Симметричность преобразования означает, что в некоторой ортонормированной базе матрица преобразования симметрическая. Имеем: θ – многочлен от φ. В ортонормированной базе матрица $[\varphi]$ симметрическая. В этой же базе матрица $[\theta]$ является многочленом от

матрицы $[\varphi]$. Значит матрица $[\theta]$ симметрическая, а потому преобразование θ симметрическое. Рассмотрим

$$(v\theta^{N-1}, v\theta^{N-1}) > 0.$$

Используя определение симметрического преобразования, получим

$$(v\theta^{N-1}, v\theta^{N-1}) = (v, v\theta^{2N-2}).$$

Так как $2N-2 \geq N$, то

$$(v, v\theta^{2N-2}) = (v, 0) = 0.$$

Противоречие. Следовательно, $\theta = 0$, а потому φ – полупросто. Учитывая, что все характеристические корни преобразования φ лежат в $\mathbb{R}$, заключаем, что в подходящей базе матрица $[\varphi]$ диагональна. Эта база и будет искомой. Она состоит из собственных векторов.

2) Теперь мы должны ортонормировать найденную базу. Пусть

$$e_1, e_2, \ldots, e_k \text{ отвечают собственному значению } \alpha_1,$$

$$e_{k+1}, e_{k+2}, \ldots, e_m \text{ отвечают собственному значению } \alpha_2 \text{ и т. д.}$$

Будем ортонормировать каждый кусок этой базы. Ясно, что при этом новые векторы также будут собственными и им соответствуют те же самые собственные значения. Необходимо проверить, что векторы из разных кусков будут попарно ортогональны. Действительно, если $a\varphi = \alpha a$ и $b\varphi = \beta b$, где $\alpha \neq \beta$, то

$$(a\varphi, b) = (\alpha a, b) = \alpha(a, b).$$

С другой стороны,

$$(a, b\varphi) = (a, \beta b) = \beta(a, b).$$

Учитывая равенство $(a\varphi, b) = (a, b\varphi)$, заключаем, что $(a, b) = 0$. Теорема доказана.

С л е д с т в и е. *Для всякой действительной симметрической матрицы $A \in \mathrm{M}_n(\mathbb{R})$ существует ортогональная матрица $T \in \mathrm{O}_n(\mathbb{R})$ такая, что TAT^{-1} – диагональная матрица.*

Д о к а з а т е л ь с т в о. Пусть φ – симметрическое преобразование евклидова пространства V размерности n, матрица которого в ортонормированной базе $e_1, e_2, \ldots, e_n$ равна $[\varphi] = A$. По теореме 3

существует ортонормированная база $e'_1, e'_2, \ldots, e'_n$ из собственных векторов преобразования φ. Матрица перехода T от ортонормированной базы к ортонормированной является ортогональной. Следовательно, $[\varphi]' = T[\varphi]T^{-1} = TAT^{-1}$ и матрица $[\varphi]'$ диагональна. Следствие доказано.

§ 40. Полярное разложение

Целью настоящего параграфа является доказательство следующего утверждения.

Т е о р е м а. *Для любой матрицы $A \in M_n(\mathbb{R})$ существует разложение $A = SU$, $S, U \in M_n(\mathbb{R})$, где S – симметрическая матрица с неотрицательными собственными значениями, а U ортогональна. При этом S единственна, а если A не вырождена, то и U единственна.* Такое разложение матрицы A называется *полярным.*

Д о к а з а т е л ь с т в о. Докажем единственность разложения. Пусть $A = SU$, где S – симметрическая с неотрицательными собственными значениями, U ортогональна. Возьмем $A^t = U^t S^t$ и, перемножая, получим

$$AA^t = SUU^t S^t = S^2.$$

Так как $(AA^t)^t = AA^t$, то матрица AA^t симметрическая, а потому ее характеристические корни действительны. По следствию из теоремы 1 предыдущего параграфа существует такая ортогональная матрица Q, что

$$Q^{-1}(AA^t)Q = D,$$

где D диагональна. Пусть

$$D = \operatorname{diag}(\underbrace{\alpha, \ldots, \alpha}_{r}, \underbrace{\beta, \ldots, \beta}_{s}, \ldots),$$

где все числа, стоящие на диагонали, действительные. Покажем, что они неотрицательны. Действительно, матрица S подобна диагональной матрице

$$\operatorname{diag}(s_1, s_2, \ldots, s_n),$$

т. е.

$$S = X^{-1}\operatorname{diag}(s_1, s_2, \ldots, s_n)X.$$

Возводя обе части этого равенства в квадрат, получим

$$S^2 = X^{-1}\mathrm{diag}(s_1^2, s_2^2, \ldots, s_n^2)X.$$

Следовательно, у D все характеристические корни неотрицательны. Положим

$$\sqrt{D} = \mathrm{diag}(\sqrt{\alpha}, \ldots, \sqrt{\alpha}, \sqrt{\beta}, \ldots, \sqrt{\beta}, \ldots),$$

где $\sqrt{\ }$ – арифметический квадратный корень. Тогда

$$S = \sqrt{AA^t} = Q\sqrt{D}Q^{-1}.$$

Но так как корень квадратный из матрицы определяется единственным образом, то матрица S единственна.

Предположим теперь, что A не вырождена. Тогда S и U не вырождены. Следовательно, у S существует обратная матрица, поэтому $U = S^{-1}A$, т. е. U единственна.

Докажем существование.

1) Предположим, что $\det A \neq 0$. Так как AA^t симметрическая, то для некоторой ортогональной матрицы Q матрица $Q^{-1}(AA^t)Q$ диагональная. Обозначим ее через D и положим

$$S = \sqrt{AA^t} = Q\sqrt{D}Q^{-1}, \quad U = S^{-1}A.$$

Покажем, что построенные матрицы удовлетворяют всем необходимым требованиям. Покажем, что матрица S является симметрической с неотрицательными собственными значениями. Вычисляя транспонированную и учитывая, что $Q^t = Q^{-1}$, получим

$$S^t = (Q\sqrt{D}Q^t)^t = Q\sqrt{D}Q^t = S,$$

т. е. S – симметрическая. Так как у подобных матриц характеристические корни одни и те же, то

$$\{\text{характеристические корни матрицы } S\} =$$

$$= \{\text{характеристические корни матрицы } \sqrt{D}\}.$$

Покажем, что все эти корни неотрицательны. Пусть α – характеристический корень матрицы D, он же является и характеристическим корнем матрицы AA^t. Так как матрица AA^t симметрическая, то все ее

характеристические корни действительные. Покажем, что они неотрицательные. Пусть $u \in \mathbb{R}^n$ такой, что $u(AA^t) = \alpha u$. Тогда

$$0 \leq (uA, uA) = (uA)(uA)^t = u(AA^t)u^t = \alpha u u^t = \alpha(u, u).$$

Отсюда $\alpha \geq 0$.

Проверим ортогональность матрицы U:

$$UU^t = (S^{-1}A)(S^{-1}A)^t = S^{-1}AA^t(S^{-1})^t = E,$$

где мы воспользовались тем, что $AA^t = S^2$, т. е. U ортогональна.

2) Пусть теперь $\det A = 0$. Представим ее как предел

$$A = \lim_{k\to\infty} A_k$$

невырожденных матриц (например, можно положить $A_k = A - \frac{1}{k}E$). Справедливость теоремы для каждой матрицы A_k доказана. Следовательно, $A_k = S_k U_k$, где S_k – симметрическая с положительными собственными значениями, а U_k – ортогональная. По следствию из теоремы 1 предыдущего параграфа

$$S_k = V_k^{-1} D_k V_k,$$

где V_k – ортогональная матрица, а D_k – диагональная. Тогда

$$A_k = V_k^{-1} D_k V_k U_k$$

и из этого равенства

$$V_k A_k U_k^{-1} V_k^{-1} = D_k.$$

При $k \to \infty$ матрицы D_k стремятся к диагональной. Воспользуемся следующим утверждением.

Л е м м а 1. *Если множество замкнуто и ограничено, то из всякой бесконечной последовательности его элементов можно выбрать сходящуюся подпоследовательность.*

Л е м м а 2. *Множество ортогональных матриц* $\mathrm{O}_n(\mathbb{R})$ *замкнуто и ограничено.*

Д о к а з а т е л ь с т в о. 1) Проверим замкнутость. Замкнутость множества M означает, что если некоторая последовательность из M сходится к некоторому элементу, то этот элемент лежит в M. Пусть

$$X_0 = \lim_{k\to\infty} X_k,$$

где все матрицы X_k ортогональны. Покажем, что тогда и X_0 ортогональна. Действительно, ввиду ортогональности матриц X_k имеем

$$X_k X_k^t = E.$$

Переходя к пределу, получим

$$\lim_{k\to\infty}(X_k X_k^t) = \lim_{k\to\infty} X_k \cdot \lim_{k\to\infty} X_k^t = X_0 X_0^t = E.$$

Следовательно, X_0 ортогональна.

2) Проверим ограниченность. Надо указать константу, которая ограничивает каждую матрицу. Мы знаем, что если $X = (x_{ij})$ ортогональна, то

$$\sum_{j=1}^{n} x_{ij}^2 = 1,$$

а потому для каждого ее элемента справедливо неравенство $|x_{ij}| \le 1$. Лемма доказана.

Таким образом, мы можем воспользоваться леммой 1. У нас есть последовательность матриц. Выбираем такую подпоследовательность, чтобы элементы, стоящие на месте (1, 1), сходились к некоторому элементу. Далее, из выбранной подпоследовательности выбираем такую подпоследовательность, чтобы все элементы, стоящие на месте (1, 2), сходились к некоторому элементу и т. д. В результате найдем подпоследовательность матриц, которая сходится к некоторой матрице.

Выберем сходящиеся подпоследовательности $\{U_{k_l}\}$, $\{V_{k_l}\}$. Пусть

$$U_0 = \lim_{l\to\infty} U_{k_l}, \quad V_0 = \lim_{l\to\infty} V_{k_l}.$$

Пределом диагональных матриц может быть только диагональная матрица. Если при этом на диагонали стоят положительные числа, то в пределе там не могут возникнуть отрицательные. Предел последовательности

$$V_{k_l} A_{k_l} U_{k_l}^{-1} V_{k_l}^{-1} = D_{k_l}$$

равен

$$V_0 A U_0^{-1} V_0^{-1} = D_0,$$

где D_0 – диагональная с неотрицательными элементами. Отсюда

$$A = V_0^{-1} D_0 V_0 U_0,$$

и, полагая $S = V_0^{-1} D_0 V_0$, $U = U_0$, получим полярное разложение. Теорема доказана.

Глава 8

КВАДРАТИЧНЫЕ ФОРМЫ

§ 41. Квадратичные формы

Пусть P – поле, $x_1, x_2, \ldots, x_n$ – набор переменных. *Формой степени* m называется однородный многочлен степени m из кольца $P[x_1, x_2, \ldots, x_n]$. При $m = 1$ форма называется *линейной*, при $m = 2$ – *квадратичной*, при $m = 3$ – *кубической* и т. д.

41.1. Поведение матрицы квадратичной формы при линейной замене переменных. *Квадратичной формой* над полем P называется следующее выражение:

$$f = f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j, \quad a_{ij} \in P.$$

Учитывая, что переменные перестановочны, т. е. $x_i x_j = x_j x_i$, будем считать, что $a_{ij} = a_{ji}$. В этом случае нашей квадратичной форме соответствует симметрическая матрица коэффициентов

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad a_{ij} = a_{ji}.$$

Обозначим

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

– столбец неизвестных. Тогда нашу форму можно представить в таком виде:

$$f = X^t A X.$$

В дальнейшем мы будем часто преобразовывать переменные x_1, x_2, ..., x_n. Как при этом будет меняться форма?

Л е м м а 1. *Если переменные $x_1, x_2, \ldots, x_n$ подвергнуть линейной замене с матрицей $Q \in \mathrm{M}_n(P)$, т. е. положить $X = QY$, где*

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} - \text{ новые переменные,}$$

то квадратичная форма $f = X^t A X$ перейдет в квадратичную форму с матрицей $Q^t A Q$.

Д о к а з а т е л ь с т в о. Подставляя выражения старых переменных через новые, получим

$$f = X^t A X = (QY)^t A (QY) = Y^t (Q^t A Q) Y,$$

и в новых переменных наша форма имеет матрицу $Q^t A Q$.

41.2. Приведение квадратичной формы к каноническому виду. Будем искать линейные замены, приводящие квадратичные формы к наиболее простому виду. При этом будем использовать невырожденные замены переменных, т. е. замены, матрицы которых невырождены.

О п р е д е л е н и е. Квадратичная форма называется *канонической*, если она не содержит смешанных произведений, т. е. имеет вид

$$f = b_1 y_1^2 + b_2 y_2^2 + \ldots + b_n y_n^2, \quad b_i \in P.$$

Т е о р е м а 1. а) *Всякую квадратичную форму над полем P характеристики, не равной 2, можно привести линейной невырожденной заменой переменных с коэффициентами из P к каноническому виду.* б) *Для полей характеристики 2 это неверно.* в) *Число ненулевых коэффициентов в каноническом виде квадратичной формы не*

зависит от выбора линейной невырожденной замены, приводящей к этому виду, и равно рангу матрицы исходной формы.

Д о к а з а т е л ь с т в о. а) Пусть

$$f = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j, \quad a_{ij} \in P$$

– квадратичная форма. Проведем доказательство индукцией по n.

При $n = 1$ имеем форму $f = a_{11} x_1^2$, которая уже каноническая.

Предположим, что утверждение справедливо для $n - 1$, и установим его для n.

Случай 1. Разберем вначале случай, когда некоторый коэффициент a_{ii} отличен от нуля. Не уменьшая общности, можно считать, что $a_{11} \neq 0$ (иначе можно перенумеровать переменные – это линейное, невырожденное преобразование). Рассмотрим форму

$$f - \frac{1}{a_{11}} (a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n)^2.$$

После приведения подобных пропадут слагаемые, содержащие x_1, и мы получим квадратичную форму $g(x_2, \ldots, x_n)$, которая уже не зависит от x_1. Это наблюдение подсказывает следующую линейную замену:

$$\begin{cases} y_1 = a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n, \\ y_2 = x_2, \\ \ldots\ldots\ldots\ldots \\ y_n = x_n. \end{cases}$$

Нетрудно проверить, что определитель матрицы этой замены равен a_{11}, а потому она не вырождена. После применения этой замены, получим

$$f = \frac{1}{a_{11}} y_1^2 + g(y_2, \ldots, y_n).$$

По индукционному предположению существует линейная невырожденная замена переменных

$$\begin{cases} y_2 = c_{22} z_2 + \cdots + c_{2n} z_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = c_{n2} z_2 + \cdots + c_{nn} z_n, \end{cases} \qquad c_{ij} \in P,$$

после выполнения которой форма примет вид

$$g = b_2 z_2^2 + \ldots + b_n z_n^2.$$

Рассмотрим замену

$$\begin{cases} y_1 = z_1, \\ y_2 = c_{22}z_2 + \cdots + c_{2n}z_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = c_{n2}z_2 + \cdots + c_{nn}z_n. \end{cases}$$

После нее

$$f = \frac{1}{a_{11}}z_1^2 + b_2z_2^2 + \ldots + b_nz_n^2,$$

т. е. мы привели нашу форму к каноническому виду.

Случай 2. Все $a_{ii} = 0$. Если при этом и все другие коэффициенты a_{ij} равны нулю, то $f = 0$ – каноническая. Поэтому пусть $a_{12} \neq 0$ (если это не так, то используем перенумерацию переменных), т. е.

$$f = 2a_{12}x_1x_2 + \ldots$$

Выполним замену

$$\begin{cases} x_1 = \ \ t_1 - t_2, \\ x_2 = \ \ t_1 + t_2, \\ x_3 = \ \ \ \ t_3, \\ \ldots\ldots\ldots\ldots \\ x_n = \ \ \ \ t_n. \end{cases}$$

После нее форма примет вид

$$f = 2a_{12}t_1^2 - 2a_{12}t_2^2 + \ldots$$

Заметим, что выписанные слагаемые не могут уничтожиться с невыписанными. Следовательно, мы пришли к разобранному случаю 1).

У п р а ж н е н и е. Где используется, что характеристика поля не равна 2?

б) Укажем пример, когда теорема неверна. Рассмотрим поле P характеристики 2. В этом поле $1 + 1 = 0$, а потому $a = -a$ для любого a из P, т. е. всякий элемент является противоположным к себе. Как мы знаем, простейшим таким полем является поле вычетов $\mathbb{Z}_2$. Покажем, что квадратичную форму $f = x_1x_2$ от двух переменных нельзя привести к каноническому виду. Действительно, если

$$\begin{cases} x_1 = b_{11}y_1 + b_{12}y_2, \\ x_2 = b_{21}y_1 + b_{22}y_2 \end{cases}$$

– некоторая невырожденная линейная замена переменных, то в новых переменных форма примет вид

$$f = b_{11}b_{21}y_1^2 + (b_{11}b_{22} + b_{12}b_{21})y_1y_2 + b_{12}b_{22}y_2^2.$$

Чтобы форма приняла канонический вид, необходимо, чтобы коэффициент при y_1y_2 обращался в нуль, т. е.

$$b_{11}b_{22} + b_{12}b_{21} = b_{11}b_{22} - b_{12}b_{21} = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = 0,$$

но в этом случае матрица линейной замены будет вырожденной.

в) Пусть

$$f = b_1z_1^2 + b_2z_2^2 + \ldots + b_nz_n^2$$

– канонический вид квадратичной формы. Рассмотрим ее матрицу:

$$D = \begin{pmatrix} b_1 & & & \mathbf{0} \\ & b_2 & & \\ & & \ddots & \\ \mathbf{0} & & & b_n \end{pmatrix}.$$

Ранг этой матрицы равен числу ненулевых диагональных элементов. Если некоторая матрица A имеет ранг r, то матрица Q^tAQ, где Q – невырожденная матрица, также имеет ранг r, так как при умножении матрицы на невырожденную ее ранг не меняется. Учитывая, что $Q^tAQ = D$, заключаем, что r равно числу ненулевых слагаемых в каноническом виде. Теорема доказана.

О п р е д е л е н и е. *Ранг формы* – число ненулевых коэффициентов при квадратах в каноническом виде.

41.3. Закон инерции действительной квадратичной формы. Квадратичная форма называется *действительной* (*комплексной*), если ее коэффициенты лежат в поле $\mathbb{R}$ (соответственно в $\mathbb{C}$).

О п р е д е л е н и е. *Сигнатурой* действительной квадратичной формы называется перечень знаков коэффициентов любого ее канонического вида:

$$(\underbrace{+, \ldots, +, -, \ldots, -, 0, \ldots, 0}_{n}).$$

Покажем, что это определение корректно, т. е. не зависит от канонического вида.

Т е о р е м а (закон инерции). *Сигнатура действительной квадратичной формы не зависит от выбора невырожденной действительной замены переменных, приводящей форму к каноническому виду.*

Д о к а з а т е л ь с т в о проведем от противного. Допустим, что нашлись две невырожденные действительные замены переменных, приводящие форму к каноническим видам с разными сигнатурами:

$$f = y_1^2 + \ldots + y_k^2 - y_{k+1}^2 - \ldots - y_r^2 = z_1^2 + \ldots + z_l^2 - z_{l+1}^2 - \ldots - z_r^2,$$

$$0 \le k < l \le n,$$

где

$$Y = QX, \quad \det Q \neq 0, \quad Z = SX, \quad \det S \neq 0. \tag{1}$$

Рассмотрим систему линейных уравнений

$$\begin{cases} q_{11}x_1 + q_{12}x_2 + \cdots + q_{1n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ q_{k1}x_1 + q_{k2}x_2 + \cdots + q_{kn}x_n = 0, \\ s_{l+1,1}x_1 + s_{l+1,2}x_2 + \cdots + s_{l+1,n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ s_{n1}x_1 + s_{n2}x_2 + \cdots + s_{nn}x_n = 0, \end{cases} \tag{2}$$

где первые k равенств берутся из замены $Y = QX$, а последние $n - l$ – из замены $Z = SX$. В этой системе $k + n - l$ уравнений и n неизвестных. При этом

$$k + n - l = n - (l - k) < n,$$

т. е. ранг системы меньше n. Следовательно, система имеет ненулевые решения. Пусть $x^0 = (x_1^0, x_2^0, \ldots, x_n^0)$ – одно из них. По формулам (1) найдем $y^0 = Qx^0$ и $z^0 = Sx^0$. Тогда

$$f(x^0) = -(y_{k+1}^0)^2 - (y_{k+2}^0)^2 - \ldots - (y_r^0)^2 = (z_1^0)^2 + (z_2^0)^2 + \ldots + (z_l^0)^2.$$

Перенося все слагаемые в одну часть, получим сумму квадратов, равную нулю. В частности, $z_1^0 = z_2^0 = \ldots = z_l^0$. Значит,

$$\begin{cases} z_1^0 = s_{11}x_1^0 + s_{12}x_2^0 + \cdots + s_{1n}x_n^0 = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ z_l^0 = s_{l1}x_1^0 + s_{l2}x_2^0 + \cdots + s_{ln}x_n^0 = 0. \end{cases}$$

Отсюда, учитывая, что x^0 удовлетворяет системе (2), заключаем: $Sx^0 = 0$. Но это однородная система с невырожденной матрицей, а потому она имеет только нулевое решение. Это противоречит тому, что $x^0 \neq 0$. Следовательно, $k = l$, и теорема доказана.

§ 42. Эквивалентность квадратичных форм

Пусть P – поле, G – подгруппа общей линейной группы $\mathrm{GL}_n(P)$.

О п р е д е л е н и е. Квадратичные формы f и g от x_1, x_2, $\ldots$, x_n над P называются *эквивалетными* относительно G, если от f к g можно перейти при помощи линейной замены переменных с матрицей из G.

Будем записывать это так $f \sim g$. Покажем, что отношение $\sim$ является отношением эквивалентности.

а) $f \sim f$, так как $f \overset{E}{\to} f$, т. е. от формы f можно перейти к f при помощи единичной матрицы E, которая лежит в любой подгруппе $G \leq \mathrm{GL}_n(P)$;

б) если $f \sim g$, то $g \sim f$. Действительно, если от f можно перейти к g при помощи матрицы A из G, то от g можно перейти к f при помощи обратной матрицы A^{-1}, которая также лежит в G;

в) если $f \sim g$ и $g \sim h$, то $f \sim h$. Действительно, если от f можно перейти к g при помощи матрицы A из G, а от f можно перейти к h при помощи матрицы B из G, то от g можно перейти к h при помощи матрицы AB, которая очевидно лежит в G.

Относительно этого отношения эквивалентности все квадратичные формы распадаются на классы эквивалентности, которые зависят от P и G. Рассмотрим далее случаи, когда G одна из следующих групп: $GL_n(\mathbb{C})$, $GL_n(\mathbb{R})$, $O_n(\mathbb{R})$.

42.1. Эквивалентность комплексных квадратичных форм относительно группы $\mathrm{GL}_n(\mathbb{C})$.

Т е о р е м а 1. *Две комплексные квадратичные формы эквивалентны относительно группы $\mathrm{GL}_n(\mathbb{C})$ тогда и только тогда, когда они имеют одинаковый ранг. В частности, в качестве представителей смежных классов можно взять следующие формы*:

$$z_1^2 + z_2^2 + \ldots + z_r^2, \quad 0 \leq r \leq n.$$

Каждая комплексная форма эквивалентна одной из этих форм.

Д о к а з а т е л ь с т в о. 1) $\Rightarrow$. Пусть $f \sim g$, т. е. $f \overset{A}{\to} g$, $A \in \mathrm{GL}_n(\mathbb{C})$. Приведем g к каноническому виду:

$$g \to b_1 y_1^2 + b_2 y_2^2 + \ldots + b_r y_r^2, \quad b_i \in \mathbb{C}, \quad b_i \neq 0.$$

Выполняя замену

$$\begin{cases} z_1 = \sqrt{b_1}\, y_1, \\ \ldots\ldots\ldots\ldots \\ z_r = \sqrt{b_r}\, y_r, \\ z_{r+1} = y_{r+1}, \\ \ldots\ldots\ldots\ldots \\ z_n = y_n, \end{cases}$$

придем к форме

$$g \to z_1^2 + z_2^2 + \ldots + z_r^2.$$

Это доказывает последнюю часть теоремы.

Имеем диаграмму

$$\begin{array}{ccc} g & \overset{Q}{\to} & z_1^2 + z_2^2 + \ldots + z_r^2 \\ A\uparrow & \nearrow & \\ f & & \end{array}$$

для некоторой матрицы $Q \in \mathrm{GL}_n(\mathbb{C})$, т. е.

$$f \overset{AQ}{\to} z_1^2 + z_2^2 + \ldots + z_r^2.$$

Значит ранг f равен рангу g и равен r.

2) $\Leftarrow$. Пусть ранг формы f равен рангу формы g и этот ранг равен r. Тогда

$$f \overset{Q}{\to} z_1^2 + z_2^2 + \ldots + z_r^2$$

и

$$g \overset{S}{\to} z_1^2 + z_2^2 + \ldots + z_r^2.$$

Следовательно, $f \overset{QS^{-1}}{\longrightarrow} g$, но это и значит, что $f \sim g$. Теорема доказана.

42.2. Эквивалентность действительных квадратичных форм относительно группы $\mathrm{GL}_n(\mathbb{R})$.

Т е о р е м а 2. *Две действительные квадратичные формы эквивалентны относительно $\mathrm{GL}_n(\mathbb{R})$ тогда и только тогда, когда их*

сигнатуры равны. В частности, в качестве представителей смежных классов можно взять формы

$$z_1^2 + z_2^2 + \ldots + z_k^2 - z_{k+1}^2 - z_{k+2}^2 - \ldots - z_{k+l}^2, \quad 0 \le k + l \le n.$$

Каждая действительная квадратичная форма эквивалентна одной из этих форм.

Д о к а з а т е л ь с т в о. 1) $\Rightarrow$. Пусть $f \sim g$, т. е. $f \xrightarrow{A} g$, $A \in \mathrm{GL}_n(\mathbb{R})$. Приведем g к каноническому виду:

$$g \xrightarrow{Q} z_1^2 + z_2^2 + \ldots + z_k^2 - z_{k+1}^2 - z_{k+2}^2 - \ldots - z_{k+l}^2.$$

При этом квадратичная форма f приводится к тому же самому каноническому виду при помощи матрицы AQ:

$$\begin{array}{ccc} g & \xrightarrow{Q} & z_1^2 + z_2^2 + \ldots + z_k^2 - z_{k+1}^2 - z_{k+2}^2 - \ldots - z_{k+l}^2. \\ A\uparrow & \nearrow AQ & \\ f & & \end{array}$$

По закону инерции, сигнатуры f и g равны.

2) $\Leftarrow$. Пусть сигнатуры f и g равны. Надо доказать, что f и g эквивалентны. Приведем обе формы к каноническому виду:

$$f \xrightarrow{Q} z_1^2 + z_2^2 + \ldots + z_k^2 - z_{k+1}^2 - z_{k+2}^2 - \ldots - z_{k+l}^2,$$

$$g \xrightarrow{S} z_1^2 + z_2^2 + \ldots + z_k^2 - z_{k+1}^2 - z_{k+2}^2 - \ldots - z_{k+l}^2.$$

Понятно, что $f \xrightarrow{QS^{-1}} g$. Следовательно, по определению $f \sim g$. Теорема доказана.

42.3. Эквивалентность действительных квадратичных форм относительно ортогональной группы $\mathrm{O}_n(\mathbb{R})$.

О п р е д е л е н и е. Набор характеристических корней матрицы, взятых с их кратностью, называется *спектром матрицы. Спектр квадратичной формы* – спектр ее матрицы.

Т е о р е м а (приведение к главным осям). *Две действительные квадратичные формы эквивалентны относительно $\mathrm{O}_n(\mathbb{R})$ тогда и только тогда, когда их спектры равны. В частности, в качестве представителей смежных классов можно взять такие формы:*

$$c_1 z_1^2 + c_2 z_2^2 + \ldots + c_n z_n^2, \quad c_i \in \mathbb{R}, \quad c_1 \le c_2 \le \ldots \le c_n.$$

Д о к а з а т е л ь с т в о. 1) $\Rightarrow$. Пусть $f \sim g$, т. е. $f \xrightarrow{Q} g$, $Q \in \mathrm{O}_n(\mathbb{R})$. Пусть A – матрица формы f, а B – матрица формы g. Мы знаем, что $B = Q^t A Q$, где Q – ортогональна. Отсюда $B = Q^{-1} A Q$, а потому матрицы A и B подобны. Следовательно, у них характеристические многочлены одинаковы, поэтому спектр A равен спектру B.

2) $\Leftarrow$. Пусть спектр матрицы A равен спектру матрицы B. У нас была теорема о том, что всякая симметрическая матрица сопряжена при помощи ортогональной матрицы с диагональной. Следовательно,

$$UAU^{-1} = D, \quad U - \text{ ортогональная}, \quad D - \text{ диагональная};$$

$$VBV^{-1} = F, \quad V - \text{ ортогональная}, \quad F - \text{ диагональная}.$$

Очевидно,

$$\text{спектр}\, D = \text{ спектр}\, A = \text{ спектр}\, B = \text{ спектр}\, F.$$

Здесь первое и последнее равенства следуют из подобия матриц, а второе равенство – из условия теоремы.

Пусть

$$D = \mathrm{diag}(d_1, d_2, \ldots, d_n).$$

После перестановки коэффициентов D совпадает с F. Заметим, что этой перестановке соответствует мономиальная матрица (в каждой строке и каждом столбце лишь один элемент отличен от нуля и равен единице), а всякая мономиальная матрица ортогональна. Пусть для мономиальной матрицы W выполняется равенство

$$F = W^{-1} D W.$$

Тогда

$$A = U^{-1} D U = U^{-1}(W F W^{-1})U = U^{-1} W V B V^{-1} W^{-1} U =$$
$$= (V^{-1} W^{-1} U)^{-1} B (V^{-1} W^{-1} U) = (V^{-1} W^{-1} U)^t B (V^{-1} W^{-1} U),$$

но это и означает, что f эквивалентна g. Теорема доказана.

§ 43. Положительно определенные квадратичные формы

43.1. Свойства положительно определенных квадратичных форм. Мы знаем, что скалярное произведение на пространстве $V = \mathbb{R}^n$ задается формулой

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

$x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$, и отображение

$$(,) : V \times V \longrightarrow \mathbb{R}$$

является билинейной (т. е. линейной по каждому аргументу) формой.

Предположим, что

$$f(x, y) = \sum_{i,j=1}^{n} a_{ij} x_i y_j, \quad a_{ij} \in \mathbb{R}$$

– некоторая вещественная билинейная форма. Когда она задает скалярное произведение в $\mathbb{R}^n$?

Из определения скалярного произведения следует, что она должна удовлетворять следующим аксиомам:

1) $f(x, y) = f(y, x)$;
2) $f(x' + x'', y) = f(x', y) + f(x'', y)$;
3) $f(\alpha x, y) = \alpha f(x, y)$, $\alpha \in \mathbb{R}$;
4) $f(x, x) > 0$ при $x \neq 0$.

Из аксиомы 1 следует, что коэффициенты билинейной формы должны образовывать симметрическую матрицу. Аксиомы 2 и 3 никаких ограничений не накладывают, так как вытекают из билинейности формы. Заметим, что, рассматривая $f(x, x)$, $x \in \mathbb{R}$, получим квадратичную форму. Остается понять, каковы коэффициенты и сигнатура этой квадратичной формы, если она удовлетворяет аксиоме 4.

О п р е д е л е н и е. *Угловым минором порядка* k, $1 \leq k \leq n$ матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots\ldots\ldots\ldots\ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \in M_n(\mathbb{R})$$

называется минор

$$\Delta_k = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ \ldots\ldots\ldots\ldots\ldots \\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{vmatrix}.$$

Т е о р е м а 1. *Для действительной квадратичной формы f следующие утверждения равносильны*:

1) $f(x) > 0$ *при всех* $x \in \mathbb{R}^n$, $x \neq 0$;

2) *сигнатура формы f равна $(+,+,\ldots,+)$, т. е. f имеет канонический вид*

$$z_1^2 + z_2^2 + \ldots + z_n^2;$$

3) *все угловые миноры матрицы формы f положительны.*

При любом из этих условий форма f называется *положительно определенной.*

Д о к а з а т е л ь с т в о. 1) $\Rightarrow$ 2) От противного. Пусть f имеет канонический вид

$$f = b_1 y_1^2 + b_2 y_2^2 + \ldots + b_n y_n^2, \quad b_n \le 0,$$

который получается после замены

$$\begin{cases} y_1 = q_{11}x_1 + q_{12}x_2 + \ldots + q_{1n}x_n, \\ y_2 = q_{21}x_1 + q_{22}x_2 + \ldots + q_{2n}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = q_{n1}x_1 + q_{n2}x_2 + \ldots + q_{nn}x_n, \end{cases}$$

т. е. $Y = QX$, где матрица $Q = (q_{ij})$ невырождена. Рассмотрим систему линейных уравнений

$$\begin{cases} q_{11}x_1 + q_{12}x_2 + \ldots + q_{1n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ q_{n-1,1}x_1 + q_{n-1,2}x_2 + \ldots + q_{n-1,n}x_n = 0, \\ q_{n1}x_1 + q_{n2}x_2 + \ldots + q_{nn}x_n = 1. \end{cases}$$

Пусть $x^0 = (x_1^0, x_2^0, \ldots, x_n^0)$ – ее решение. Ясно, что оно отлично от нуля. Имеем $f(x^0) = b_n \le 0$. Противоречие.

2) $\Rightarrow$ 1) Имеем:

$$f(z) = z_1^2 + z_2^2 + \ldots + z_n^2.$$

Выберем некоторую n-ку $x^0 \in \mathbb{R}^n$, $x^0 \ne 0$. По ней найдем соответствующую n-ку $z^0 = z^0(x^0) \in \mathbb{R}^n$. Тогда $f(z^0) = f(x^0) > 0$.

Покажем, что из 1 и 2 следует 3. Индукция по n. При $n = 1$ имеем $f = a_{11}x_1^2$, $a_{11} > 0$, а потому все угловые миноры положительны. Предположим, что утверждение справедливо для $n-1$, и установим его для n. Имеем

$$f(x_1, x_2, \ldots, x_n) = g(x_1, x_2, \ldots, x_{n-1}) + 2\sum_{i=1}^{n-1} a_{in}x_i x_n + a_{nn}x_n^2.$$

Проверим, что для g выполняется условие 1. Возьмем $(x_1^0, x_2^0, \ldots, x_{n-1}^0) \neq 0$. Тогда

$$g(x_1^0, x_2^0, \ldots, x_{n-1}^0) = f(x_1^0, x_2^0, \ldots, x_{n-1}^0, 0) > 0$$

по условию. По индуктивному предположению

$$\Delta_1 > 0, \quad \Delta_2 > 0, \ldots, \Delta_{n-1} > 0,$$

так как они являются угловыми минорами для g. Надо доказать, что $\Delta_n > 0$. Ввиду 2 форма f приводится к сумме квадратов:

$$f \xrightarrow{Q} z_1^2 + z_2^2 + \ldots + z_n^2.$$

Если A – матрица формы f, то мы имеем равенство $Q^t A Q = E$, из которого $|A||Q|^2 = 1$ и $\Delta_n = |A| > 0$.

3) $\Rightarrow$ 2) Индукция по n. При $n = 1$ утверждение очевидно. Пусть $n > 1$. Так же как и выше, представим f в виде

$$f(x_1, x_2, \ldots, x_n) = g(x_1, x_2, \ldots, x_{n-1}) + 2\sum_{i=1}^{n-1} a_{in} x_i x_n + a_{nn} x_n^2.$$

По индуктивному предположению форма g приводится к виду

$$g = y_1^2 + y_2^2 + \ldots + y_{n-1}^2$$

при помощи некоторой замены:

$$\begin{cases} x_1 = q_{11} y_1 + q_{12} y_2 + \ldots + q_{1,n-1} y_{n-1}, \\ x_2 = q_{21} y_1 + q_{22} y_2 + \ldots + q_{2,n-1} y_{n-1}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_{n-1} = q_{n-1,1} y_1 + q_{n-1,2} y_2 + \ldots + q_{n-1,n-1} y_{n-1}. \end{cases}$$

Рассмотрим далее замену

$$\begin{cases} x_1 = q_{11} y_1 + q_{12} y_2 + \ldots + q_{1,n-1} y_{n-1}, \\ x_2 = q_{21} y_1 + q_{22} y_2 + \ldots + q_{2,n-1} y_{n-1}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_{n-1} = q_{n-1,1} y_1 + q_{n-1,2} y_2 + \ldots + q_{n-1,n-1} y_{n-1}, \\ x_n = \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad y_n. \end{cases}$$

После нее

$$f = y_1^2 + y_2^2 + \ldots + y_{n-1}^2 + 2\sum_{i=1}^{n-1} b_{in} y_i y_n + b_{nn} y_n^2 = \sum_{i=1}^{n-1} (y_i + b_{in} y_n)^2 + c y_n^2$$

для некоторого коэффициента c. Применим замену

$$\begin{cases} z_1 = y_1 + b_{1n}y_n, \\ z_2 = y_2 + b_{2n}y_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ z_{n-1} = y_{n-1} + b_{n-1,n}y_n, \\ z_n = y_n, \end{cases}$$

получим

$$f = z_1^2 + z_2^2 + \ldots + z_{n-1}^2 + cz_n^2.$$

Если матрица формы f была равна A, то мы получили равенство

$$\operatorname{diag}(1, 1, \ldots, 1, c) = S^t AS,$$

где S – матрица замены. Вычисляя определители от обеих частей, имеем

$$c = |A||S|^2 = \Delta_n |S|^2.$$

Учитывая, что $\Delta_n > 0$, заключаем, что и $c > 0$. Произведя замену

$$\begin{cases} t_1 = z_1, \\ t_2 = z_2, \\ \ldots\ldots\ldots \\ t_n = \sqrt{c}z_n, \end{cases}$$

получим утверждение 2. Теорема доказана.

43.2. Пары форм. В этом пункте мы рассмотрим следующую задачу: даны две действительные квадратичные формы, можно ли их одним линейным невырожденным преобразованием привести к каноническому виду?

П р и м е р. Пусть

$$f = x_1^2, \quad g = x_1x_2$$

– две действительные квадратичные формы. Покажем, что эти формы нельзя одновременно привести к каноническому виду невырожденной заменой. Действительно, если

$$\begin{cases} x_1 = q_{11}y_1 + q_{12}y_2, \\ x_2 = q_{21}y_1 + q_{22}y_2 \end{cases}$$

– линейная замена переменных, то

$$f = q_{11}^2y_1^2 + 2q_{11}q_{12}y_1y_2 + q_{12}^2y_2^2, \quad g = q_{11}q_{21}y_1^2 + (q_{11}q_{22} + q_{12}q_{21})y_1y_2 + q_{12}q_{22}y_2^2.$$

Пусть

$$\begin{cases} q_{11}q_{12} = 0, \\ q_{11}q_{22} + q_{12}q_{21} = 0. \end{cases}$$

Тогда нетрудно показать, что

$$\begin{vmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{vmatrix} = q_{11}q_{22} - q_{12}q_{21} = 0,$$

т. е. замена обязательно вырожденная.

Т е о р е м а 2. *Две квадратичные формы, из которых одна положительно определенная, можно одновременно привести к каноническому виду при помощи невырожденной вещественной линейной замены переменных.*

Д о к а з а т е л ь с т в о следует из цепочки преобразований:

$$\begin{cases} f \overset{\text{невыр}}{\longrightarrow} & y_1^2 + y_2^2 + \ldots + y_n^2 \quad - - - > \overset{\text{ортог}}{\longrightarrow} \quad z_1^2 + z_2^2 + \ldots + z_n^2, \\ g - - - > & \widetilde{g} \qquad\qquad\qquad\qquad\quad c_1z_1^2 + c_2z_2^2 + \ldots + c_nz_n^2. \end{cases}$$

Вначале невырожденным линейным преобразованием приводим положительно определенную форму f к каноническому виду. В силу теоремы о положительно определенных формах получим форму, у которой матрица единичная. При этом, выполняя ту же замену, приводим форму g к некоторой форме $\widetilde{g}$. Теперь, используя ортогональное преобразование, приведем форму $\widetilde{g}$ к каноническому виду. При этом, как легко заметить, матрица формы f по-прежнему останется единичной. Следовательно, композиция этих двух линейных замен приведет формы f и g к каноническому виду. Теорема доказана.

§ 44. Значения действительной квадратичной формы на единичной сфере

В настоящем параграфе докажем следующее утверждение.

Т е о р е м а. *Пусть $\mathbb{R}^n$ – евклидово пространство с обычным скалярным произведением*

$$(x, y) = \sum_{i=1}^{n} x_i y_i,$$

φ – симметрическое преобразование пространства $\mathbb{R}^n$. Определим квадратичную форму

$$f(x) = (x\varphi, x) = [x][\varphi][x]^t$$

на единичной сфере

$$S = \{v \in V \mid ||v|| = 1\}.$$

Тогда

1) *$f(x)$ достигает своего минимума. Если в точке $u \in S$ достигается минимум и $f(u) = \alpha$, то $u\varphi = \alpha u$, т. е. α – собственное значение, а u – собственный вектор преобразования φ;*

2) *α – наименьшее из всех собственных значений.*

Д о к а з а т е л ь с т в о. Установим вначале пункт 2. Если $\beta < \alpha$ и $w\varphi = \beta w$, где $||w|| = 1$, то

$$f(w) = (w\varphi, w) = \beta(w, w) = \beta < \alpha = \min_{v \in S} f(v).$$

Противоречие.

Для доказательства пункта 1 нам потребуется

Л е м м а. *Пусть S – замкнутое и ограниченное подмножество из $\mathbb{R}^n$, $f : S \longrightarrow \mathbb{R}$ – непрерывная функция. Тогда*

а) *f ограничена на S;*

б) *f достигает на S своего наибольшего и наименьшего значений.*

Д о к а з а т е л ь с т в о. а) Достаточно доказать, что f ограничена сверху, так как при переходе к функции $-f$ получим и ограниченность снизу. Предположим, что f не ограничена, т. е. для любого натурального n существует x_n из S такой, что $f(x_n) > n$. Возникает последовательность $x_1, x_2, \ldots$. Так как S замкнуто и ограничено, то из этой последовательности можно выбрать сходящуюся подпоследовательность $\{x_{n_k}\}$. Пусть эта подпоследовательность сходится к x_0. Тогда $x_0 \in S$. Учитывая, что f непрерывна, получим

$$\lim_{k \to \infty} f(x_{n_k}) = f(x_0).$$

Значит, для любого $\varepsilon > 0$ существует k_0 такой, что

$$|f(x_{n_k}) - f(x_0)| < \varepsilon \text{ при } k \geq k_0.$$

Тогда

$$|f(x_{n_k})| \leq |f(x_0)| + \varepsilon \text{ при } k \geq k_0.$$

Приходим к противоречию.

б) Пусть

$$M = \sup_{x \in S} f(x), \quad m = \inf_{x \in S} f(x).$$

Достаточно доказать утверждение для верхней границы, т. е., что существует $u \in S$ такое, что $f(u) = M$. По определению точной верхней грани:

1) $f(x) \leq M$ при любых $x \in S$;

2) для любого натурального n существует $y_n \in S$ такое, что

$$f(y_n) > M - \frac{1}{n}.$$

Так как S замкнуто и ограничено, то из каждой его последовательности можно выбрать сходящуюся подпоследовательность. Выберем такую подпоследовательность y_{n_k} из последовательности y_n и пусть

$$\lim_{k \to \infty} y_{n_k} = y_0.$$

Тогда $f(y_{n_k}) \leq M$. С другой стороны, $f(y_{n_k}) > M - \frac{1}{n_k}$. Пусть теперь k стремится к бесконечности. Тогда

$$M - \frac{1}{n_k} \longrightarrow M.$$

Следовательно,

$$f(y_{n_k}) \longrightarrow f(y_0)$$

и по лемме о двух милиционерах $f(y_0) = M$. Так как S замкнуто, то по определению $y_0 \in S$. Лемма доказана.

Проверим, что единичная сфера S замкнута и ограничена. Так как для любой точки сферы выполняется равенство

$$\sum_{i=1}^{n} x_i^2 = 1,$$

то ограниченность очевидна.

Докажем замкнутость. Пусть $x_n \longrightarrow x_0$ при $n \longrightarrow \infty$ и все x_n лежат в S. Надо доказать, что тогда и x_0 лежит в S. Рассмотрим

$$(x_n, x_n) - (x_0, x_0) = (x_n, x_n) + (x_n, x_0) - (x_n, x_0) - (x_0, x_0) = (x_n, x_n - x_0) +$$

$$+ (x_n - x_0, x_0) = (x_n - x_0, x_n + x_0) = (x_n - x_0, x_n - x_0) + 2(x_n - x_0, x_0).$$

Отсюда

$$|(x_n, x_n) - (x_0, x_0)| \leq ||x_n - x_0||^2 + 2||x_n - x_0|| \cdot ||x_0|| \underset{n\to\infty}{\longrightarrow} 0.$$

Таким образом, мы установили, что

$$|(x_n, x_n) - (x_0, x_0)| = |1 - (x_0, x_0)| \underset{n\to\infty}{\longrightarrow} 0,$$

но $|1 - (x_0, x_0)|$ является константой, а потому $|1 - (x_0, x_0)| = 0$, т. е. $||x_0|| = 1$. Замкнутость установлена.

Проверим, что выполняются и все другие условия леммы.

Проверим, что f непрерывна. Имеем

$$f(x) - f(y) = (x\varphi, x) - (y\varphi, y) = (x\varphi, x) - (x\varphi, y) + (x\varphi, y) - (y\varphi, y) =$$

$$= (x\varphi, x - y) + ((x - y)\varphi, y).$$

Отсюда

$$|f(x) - f(y)| \leq |(x\varphi, x - y)| + |((x - y)\varphi, y)|.$$

Используя неравенство Коши – Буняковского, получим

$$|(x\varphi, x - y)| + |((x - y)\varphi, y)| \leq ||x\varphi|| \cdot ||x - y|| + ||(x - y)\varphi|| \cdot ||y|| \leq$$

$$\leq ||x|| \cdot ||\varphi|| \cdot ||x - y|| + ||x - y|| \cdot ||\varphi|| \cdot ||y|| = (||x|| + ||y||) \cdot ||\varphi|| \cdot ||x - y||.$$

Из этого неравенства при $x, y \in S$ имеем неравенство

$$|f(x) - f(y)| \leq 2||\varphi|| \cdot ||x - y||.$$

Следовательно, f непрерывна.

Таким образом, мы можем применять лемму. Существует $u \in S$ такое, что

$$\alpha = f(u) = \min_{x\in S} f(x).$$

Надо доказать, что $u\varphi = \alpha u$, или, обозначая $\psi = \varphi - \alpha\varepsilon$, приходим к равенству $u\psi = 0$. Имеем

$$(x\varphi, x) = f(x) \geq \alpha \text{ при любом } x \in S;$$

$$(x\varphi, x) \geq \alpha(x, x) \text{ при любом } x \in \mathbb{R}^n;$$

$$(x\psi, x) \geq 0 \text{ при любом } x \in \mathbb{R}^n.$$

Пусть $x = u + \lambda v$, где $\lambda \in \mathbb{R}$, $v \in \mathbb{R}^n$. Тогда

$$0 \leq (x\psi, x) = (u\psi + \lambda(v\psi), u + \lambda v) = (u\psi, u) + \lambda(v\psi, u) + \lambda(u\psi, v) + \lambda^2(v\psi, v).$$

Так как ψ – симметрическое преобразование, то $(v\psi, u) = (v, u\psi) = (u\psi, v)$, а потому

$$(u\psi, u)+\lambda(v\psi, u)+\lambda(u\psi, v)+\lambda^2(v\psi, v) = (u\psi, u)+2\lambda(v\psi, u)+\lambda^2(v\psi, v).$$

Рассматривая это выражение как квадратный трехчлен относительно λ и учитывая, что он неотрицательный, имеем неравенство

$$4(v\psi, u)^2 - 4(u\psi, u)(v\psi, v) \leq 0.$$

Так как $f(u) = (u\varphi, u) = \alpha(u, u)$, т. е. $(u\psi, u) = 0$, то $(v\psi, u) = 0$ или $(u\psi, v) = 0$. Из того, что вектор v произвольный, получим $u\psi = 0$. Теорема доказана.

Глава 9

ПОЛИЛИНЕЙНАЯ АЛГЕБРА

§ 45. Тензоры

45.1. Пространство полилинейных форм. Пусть P – поле, $V_1, V_2, \ldots, V_m, W$ – векторные пространства над P. Отображение

$$\varphi : V_1 \times V_2 \times \ldots \times V_m \longrightarrow W$$

называется *полилинейным или m-линейным,* если для любых $\alpha, \beta \in P$ и $v_i', v_i'' \in V_i$ справедливо равенство

$$(\ldots, \alpha v_i' + \beta v_i'', \ldots)\varphi = \alpha(\ldots, v_i', \ldots)\varphi + \beta(\ldots, v_i'', \ldots)\varphi.$$

Если $W = P$, то полилинейное отображение называется *полилинейной формой,* или *функцией.*

Т е о р е м а 1. 1) *Множество $L = L(V_1, V_2, \ldots, V_m)$ всех полилинейных форм на $V_1 \times V_2 \times \ldots \times V_m$ является векторным пространством над P относительно сложения и умножения функций на скаляр.*

2) *Если $\{e_{i_n}^{(k)}\}$ – база V_k, то в L существует единственная база $\{e^*_{i_1 \ldots i_m}\}$ такая, что*

$$e^*_{i_1 \ldots i_m}(e_{j_1}^{(1)}, e_{j_2}^{(2)}, \ldots, e_{j_m}^{(m)}) = \delta_{i_1 j_1}\delta_{i_2 j_2} \ldots \delta_{i_m j_m}.$$

Она называется базой, *сопряженной* к базам $\{e_{i_n}^{(k)}\}$.

Д о к а з а т е л ь с т в о. 1) Заметим, что если две формы линейны по каждому аргументу, то после сложения сумма также будет линейной. Аналогично для умножения на скаляр. Далее мы должны проверить четыре аксиомы абелевой группы и четыре смешанные аксиомы:

1) $\alpha(f+g)=\alpha f+\alpha g$;
2) $(\alpha+\beta)f=\alpha f+\beta f$;
3) $(\alpha\beta)f=\alpha(\beta f)$;
4) $1\cdot f=f$;

для любых α,β из P и f,g из L. Проверка этих аксиом очевидна.

2) Пусть в каждом V_k фиксировано по базе $\{e_{i_n}^{(k)}\}$. Надо понять, каким должно быть значение $e^*_{i_1\dots i_m}$ на произвольной m-ке $(x_1,x_2,\dots,x_m)$. Очевидно, что

$$e^*_{i_1\dots i_m}(x_1,x_2,\dots,x_m)=$$

$$=e^*_{i_1\dots i_m}\left(\sum_{j_1}x_{1,j_1}e^{(1)}_{j_1},\sum_{j_2}x_{2,j_2}e^{(2)}_{j_2},\dots,\sum_{j_m}x_{m,j_m}e^{(m)}_{j_m}\right)=$$

$$=\sum_{j_1}\sum_{j_2}\dots\sum_{j_m}x_{1,j_1}x_{2,j_2}\dots x_{m,j_m}e^*_{i_1\dots i_m}(e^{(1)}_{j_1},e^{(2)}_{j_2},\dots,e^{(m))}_{j_m})=$$

$$=x_{1,i_1}x_{2,i_2}\dots x_{m,i_m}.$$

Отсюда вытекает единственность e^*.

Докажем существование. Положим

$$e^*_{i_1\dots i_m}(x_1,x_2,\dots,x_m)=x_{1,i_1}x_{2,i_2}\dots x_{m,i_m}.$$

Надо доказать, что это база. Проверим условия сопряженности, т. е. как e^* действует на базисные векторы. Имеем

$$e^*_{i_1\dots i_m}(e^{(1)}_{j_1},e^{(2)}_{j_2},\dots,e^{(m)}_{j_m})=\delta_{i_1j_1}\delta_{i_2j_2}\dots\delta_{i_mj_m}.$$

Проверим линейную независимость системы $\{e^*_{i_1\dots i_m}\}$. Пусть

$$\sum\alpha_{i_1\dots i_m}e^*_{i_1\dots i_m}=0.$$

Надо доказать, что все $\alpha_{i_1\dots i_m}=0$. Вычислив обе части этого равенства в точке $(e^{(1)}_{j_1},e^{(2)}_{j_2},\dots,e^{(m)}_{j_m})$, получим $\alpha_{i_1\dots i_m}=0$.

Проверим максимальность. Пусть $f \in L$, надо показать, что ее можно представить в виде линейной комбинации. Имеем

$$f(x_1, x_2, \ldots, x_m) = f\left(\sum_{j_1} x_{1,j_1} e^{(1)}_{j_1}, \sum_{j_2} x_{2,j_2} e^{(2)}_{j_2}, \ldots, \sum_{j_m} x_{m,j_m} e^{(m)}_{j_m}\right) =$$

$$= \sum_{j_1}\sum_{j_2}\cdots\sum_{j_m} x_{1,j_1} x_{2,j_2} \ldots x_{m,j_m} f\left(e^{(1)}_{j_1}, e^{(2)}_{j_2}, \ldots, e^{(m))}_{j_m}\right),$$

где $f\left(e^{(1)}_{j_1}, e^{(2)}_{j_2}, \ldots, e^{(m))}_{j_m}\right)$ – координаты f в базе e^*. Теорема доказана.

45.2. Тензоры. Пусть V – векторное пространство над P, $V^* = L(V)$ – сопряженное пространство (множество линейных функций на V). Рассмотрим

$$L = L(\underbrace{V, \ldots, V}_{p}, \underbrace{V^*, \ldots, V^*}_{q}).$$

Пусть $e_1, e_2, \ldots, e_n$ – база V, $e^*_1, e^*_2, \ldots, e^*_n$ – сопряженная база в V^*. Можно в пространстве L выбрать сопряженную по всем базам, т. е. определить $\{e^*_{i_1\ldots i_p, j_1\ldots j_q}\}$ базу L сопряженную к $\{e_i\}$ и $\{e^*_j\}$.

Каждая полилинейная функция записывается через базу L. Пусть $f \in L$ имеет в базе $\{e^*_{i_1\ldots i_p, j_1\ldots j_q}\}$ координаты $f_{i_1\ldots i_p, j_1\ldots j_q}$. В V выберем другую базу. При этом изменится сопряженная база в V^* и база L. Как при этом изменятся координаты?

Пусть $e'_1, e'_2, \ldots, e'_n$ – другая база пространства V, $e'^*_1, e'^*_2, \ldots, e'^*_n$ – сопряженная к ней база сопряженного пространства V^*, $\{e'^*_{i_1\ldots i_p, j_1\ldots j_q}\}$ – сопряженная к ним база L, $f'_{i_1\ldots i_p, j_1\ldots j_q}$ – координаты f в этой базе. Как связаны $f'_{i_1\ldots i_p, j_1\ldots j_q}$ с $f_{i_1\ldots i_p, j_1\ldots j_q}$?

Пусть $e' = Te$, где T – матрица перехода. Как связаны сопряженные базы?

Л е м м а. *Справедливо равенство* $e'^* = \check{T} e^*$, *где* $\check{T} = (T^{-1})^t$.

Д о к а з а т е л ь с т в о. Если

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

и

$$e'^*_r = \sum_{s=1}^{n} \check{t}_{rs} e^*_s, \text{ где } \check{T} = (\check{t}_{rs}),$$

то

$$e'^*_r(e'_k) = \sum_{s=1}^{n} \check{t}_{rs} e^*_s \left(\sum_{j=1}^{n} t_{kj} e_j \right) = \sum_{s=1}^{n} \sum_{j=1}^{n} \check{t}_{rs} t_{kj} e^*_s(e_j) = \sum_{s=1}^{n} \check{t}_{rs} t_{ks} = \delta_{rk},$$

так как $e^*_s(e_j) = \delta_{sj}$. Лемма доказана.

Вычислим новые координаты, учитывая линейность f:

$$f'_{i_1 \dots i_p, j_1 \dots j_q} = f(e'_{i_1}, e'_{i_2}, \dots, e'_{i_p}, e'^*_{j_1}, e'^*_{j_2}, \dots, e'^*_{j_q}) =$$

$$= f \left(\sum_{k_1=1}^{n} t_{i_1,k_1} e_{k_1}, \sum_{k_2=1}^{n} t_{i_2,k_2} e_{k_2}, \dots, \sum_{k_p=1}^{n} t_{i_p,k_p} e_{k_p}, \sum_{l_1=1}^{n} \check{t}_{j_1,l_1} e^*_{l_1}, \right.$$

$$\left. \sum_{l_2=1}^{n} \check{t}_{j_2,l_2} e^*_{l_2}, \dots, \sum_{l_q=1}^{n} \check{t}_{j_q,l_q} e^*_{l_q} \right) =$$

$$= \sum_{k_1=1}^{n} \dots \sum_{k_p=1}^{n} \sum_{l_1=1}^{n} \dots \sum_{l_q=1}^{n} t_{i_1,k_1} t_{i_2,k_2} \dots t_{i_p,k_p} \check{t}_{j_1,l_1} \check{t}_{j_2,l_2} \dots \check{t}_{j_q,l_q} f_{k_1 \dots k_p, l_1 \dots l_q},$$

т. е. новые координаты вычисляются через старые по формуле

$$f'_{i_1 \dots i_p, j_1 \dots j_q} =$$

$$= \sum_{k_1=1}^{n} \dots \sum_{k_p=1}^{n} \sum_{l_1=1}^{n} \dots \sum_{l_q=1}^{n} t_{i_1,k_1} t_{i_2,k_2} \dots t_{i_p,k_p} \check{t}_{j_1,l_1} \check{t}_{j_2,l_2} \dots \check{t}_{j_q,l_q} f_{k_1 \dots k_p, l_1 \dots l_q}. \tag{1}$$

О п р е д е л е н и е. Говорят, что задан *p раз ковариантный и q раз контравариантный тензор* над векторным пространством V, если каждой базе векторного пространства сопоставлен набор из n^{p+q} элементов $f_{i_1 \dots i_p, j_1 \dots j_q} \in P$ так, что при изменении базы эти наборы преобразуются по формуле (1). Число $p + q$ называется *валентностью* этого тензора.

Иными словами, тензор – полилинейная функция некоторого специального вида.

45.3. Тензорное произведение векторных пространств.

З а д а ч а. Заданы несколько векторных пространств $V_1, V_2, \dots, V_m$ над полем P. Надо построить векторное пространство T и полилинейное отображение

$$\tau : V_1 \times V_2 \times \dots \times V_m \longrightarrow T,$$

такие, что:

а) $\operatorname{Im}\tau$ порождает T как векторное пространство;

б) для всякого полилинейного отображения φ из $V_1\times V_2\times\ldots\times V_m$ в произвольное векторное пространство W существует линейное отображение $\varphi_0 : T \longrightarrow W$, такое, что $\varphi = \tau\varphi_0$, т. е. следующая диаграмма коммутативна

$$\begin{array}{ccc} V_1\times V_2\times\ldots\times V_m & \xrightarrow{\varphi} & W \\ \tau\downarrow & \nearrow\varphi_0 & \\ T & & \end{array}$$

Заметим, что если такие T и τ существуют, то полилинейная алгебра сводится к линейной.

Т е о р е м а 2. 1) *Для всяких $V_1, V_2, \ldots, V_m$ существуют T и τ, решающие поставленную задачу.* 2) *Эти T и τ единственны с точностью до изоморфизма.*

Пространство T называется *тензорным произведением* пространств $V_1, V_2, \ldots, V_m$ и обозначается $V_1\otimes V_2\otimes\ldots\otimes V_m$.

Д о к а з а т е л ь с т в о. 1) Укажем T и τ. В качестве T возьмем $T = L^*$, где $L = L(V_1, V_2, \ldots, V_m)$ – пространство полилинейных форм. Для всяких $v_1\in V_1$, $v_2\in V_2$, …, $v_m\in V_m$ обозначим $v_1\otimes v_2\otimes\ldots\otimes v_m$ функцию $L\longrightarrow P$, определенную правилом

$$f \longrightarrow f(v_1, v_2, \ldots, v_m).$$

Это отображение линейно, т. е. для $v_1\otimes v_2\otimes\ldots\otimes v_m\in T$ имеем

$$(v_1\otimes v_2\otimes\ldots\otimes v_m)(\alpha f+\beta g) = \alpha f(v_1, v_2, \ldots, v_m) + \beta g(v_1, v_2, \ldots, v_m).$$

Положим

$$\tau : (v_1, v_2, \ldots, v_m) \longrightarrow v_1\otimes v_2\otimes\ldots\otimes v_m.$$

Очевидно, что τ полилинейно. Действительно,

$$(\ldots, \alpha v_i' + \beta v_i'', \ldots)\tau = \ldots\otimes(\alpha v_i' + \beta v_i'')\otimes\ldots =$$
$$= \alpha(\ldots\otimes v_i'\otimes\ldots) + \beta(\ldots\otimes v_i''\otimes\ldots).$$

Проверим свойство а из нашей задачи. Пусть $\{e_{i_n}^{(k)}\}$ – база V_k. Достаточно доказать, что $\{e_{i_1}^{(1)}\otimes e_{i_2}^{(2)}\otimes\ldots\otimes e_{i_m}^{(m)}\}$ – база T, т. е. они порождают все T. Возьмем в L базу $\{e^*_{i_1 i_2\ldots i_m}\}$, сопряженную к базам $\{e_{i_n}^{(k)}\}$. Тогда по определению

$$\left(e_{i_1}^{(1)}\otimes e_{i_2}^{(2)}\otimes\ldots\otimes e_{i_m}^{(m)}\right)(e^*_{j_1 j_2\ldots j_m}) = e^*_{j_1 j_2\ldots j_m}(e_{i_1}^{(1)}, e_{i_2}^{(2)}, \ldots, e_{i_m}^{(m)}) =$$

$$= \delta_{i_1,j_1}\delta_{i_2,j_2}\dots\delta_{i_m,j_m}.$$

Рассмотрим $e^*_{j_1j_2\dots j_m}$ как вектор из L, а $\left(e^{(1)}_{i_1}\otimes e^{(2)}_{i_2}\otimes\dots\otimes e^{(m)}_{i_m}\right)$ как линейную функцию. До этого $e^*_{j_1j_2\dots j_m}$ выступала как функция, т. е. $\{e^{(1)}_{i_1}\otimes e^{(2)}_{i_2}\otimes\dots\otimes e^{(m)}_{i_m}\}$ – база в T, сопряженная к базе $\{e^*_{j_1j_2\dots j_m}\}$ в L.

Докажем б. Пусть задано φ. Возьмем в T базу

$$\{e^{(1)}_{j_1}\otimes e^{(2)}_{j_2}\otimes\dots\otimes e^{(m)}_{j_m}\}.$$

Мы должны положить

$$\left(e^{(1)}_{j_1}\otimes e^{(2)}_{j_2}\otimes\dots\otimes e^{(m)}_{j_m}\right)\varphi_0 = \left(e^{(1)}_{j_1}, e^{(2)}_{j_2}, \dots, e^{(m)}_{j_m}\right)\varphi.$$

Продолжим φ_0 на все T по линейности. Тогда для всяких $v_1\in V_1$, $v_2\in V_2$, …, $v_m\in V_m$ имеем

$$(x_1,x_2,\dots,x_m)\varphi = \left(\sum_{j_1}x_{1,j_1}e^{(1)}_{j_1}, \sum_{j_2}x_{2,j_2}e^{(2)}_{j_2}, \dots, \sum_{j_m}x_{m,j_m}e^{(m)}_{j_m}\right)\varphi =$$

$$= \sum_{j_1}\sum_{j_2}\dots\sum_{j_m}x_{1,j_1}x_{2,j_2}\dots x_{m,j_m}(e^{(1)}_{j_1}, e^{(2)}_{j_2}, \dots, e^{(m)}_{j_m})\varphi =$$

$$= \sum_{j_1}\sum_{j_2}\dots\sum_{j_m}x_{1,j_1}x_{2,j_2}\dots x_{m,j_m}(e^{(1)}_{j_1}\otimes e^{(2)}_{j_2}\otimes\dots\otimes e^{(m)}_{j_m})\varphi_0 =$$

$$= \sum_{j_1}\sum_{j_2}\dots\sum_{j_m}x_{1,j_1}x_{2,j_2}\dots x_{m,j_m}(e^{(1)}_{j_1}, e^{(2)}_{j_2}, \dots, e^{(m)}_{j_m})\tau\varphi_0 =$$

$$= (x_1,x_2,\dots,x_m)\tau\varphi_0.$$

Следовательно, $\varphi = \tau\varphi_0$.

2) Докажем единственность. Пусть T', τ' – другая такая пара со свойствами а и б. Тогда вместо φ возьмем τ'. Ввиду полилинейности существует линейное отображение

$$\tau'_0 : T \longrightarrow T',$$

такое, что

$$\tau\tau'_0 = \tau', \tag{1}$$

$$\begin{array}{ccc} V_1\times V_2\times\dots\times V_m & \xrightarrow{\tau'} & T, \\ \tau\downarrow & \nearrow\tau'_0 & \\ T' & & \end{array}$$

и существует

$$\tau_0 : T' \longrightarrow T,$$

которое линейно, и

$$\tau'\tau_0 = \tau. \tag{2}$$

Из равенств (1) и (2) имеем

$$\tau'\tau_0\tau_0' = \tau',$$

т. е.

$$(z\tau')\tau_0\tau_0' = z\tau' \in \operatorname{Im}\tau'$$

для всякого

$$z \in V_1 \times V_2 \times \ldots \times V_m,$$

т. е. $\tau_0\tau_0' = 1$. Аналогично $\tau_0'\tau_0 = 1$, т. е. τ_0 и τ_0' обратны друг другу, а потому $T \simeq T'$. Теорема доказана.

45.4. Тензорное произведение линейных отображений. Пусть

$$V_1 \xrightarrow{\varphi_1} W_1, \quad V_2 \xrightarrow{\varphi_2} W_2$$

– два линейных отображения. Тогда по предыдущей теореме существует единственное линейное отображение

$$V_1 \otimes V_2 \xrightarrow{\varphi_1\otimes\varphi_2} W_1 \otimes W_2$$

со свойством

$$(v_1 \otimes v_2)(\varphi_1 \otimes \varphi_2) = (v_1\varphi_1) \otimes (v_2\varphi_2).$$

У п р а ж н е н и е. Доказать, что

$$(\varphi\psi) \otimes (\eta\theta) = (\varphi \otimes \eta)(\psi \otimes \theta).$$

Список литературы

1. *Ван дер Варден Б. Л.* Алгебра. М.: Наука, 1979.
2. *Винберг Э. Б.* Курс алгебры. М.: Факториал Пресс, 2002.
3. *Воеводин В. В.* Линейная алгебра. СПб.: Лань, 2006.
4. *Каргаполов М. И., Мерзляков Ю. И.* Основы теории групп. М.: Наука, 1996.
5. *Кострикин А. И.* Введение в алгебру. М.: Физматлит, 2004.
6. *Курош А. Г.* Курс высшей алгебры. СПб.: Лань, 2008.
7. *Мальцев А. И.* Основы линейной алгебры. М.: Наука, 2005.
8. *Мальцев А. И.* Алгебраические системы. М.: Наука, 1970.
9. *Мерзляков Ю. И.* Рациональные группы. Новосибирск, 1987.
10. *Халмош П.* Конечномерные векторные пространства. М.: Физматгиз, 1963.

Предметный указатель

M

H

Р

С

Учебное издание

Бардаков Валерий Георгиевич

ЛЕКЦИИ ПО АЛГЕБРЕ Ю. И. МЕРЗЛЯКОВА

Учебное пособие

Редактор Е. П. Войтенко

Дизайн обложки О. В. Брюханов

Подписано в печать 23.04.2012 г.
Формат 70×100 1/16.
Уч.-изд. л. 19,4. Усл. печ. л. 25. Тираж 100 экз.
Заказ №

Редакционно-издательский центр НГУ.
630090, Новосибирск-90, ул. Пирогова, 2.