С.Н.Тронин

ЛЕКЦИИ ПО АЛГЕБРЕ

Семестр 2

Выпуск II

Жорданова нормальная форма матрицы

Казань — 2012

УДК 512.64

Представляется на сайте университета
по решению Редакционно-издательского совета
ФГАОУВПО «Казанский (Приволжский) федеральный университет»

методической комиссии Института математики и механики
им. Н.И.Лобачевского
Протокол №7 от 19 апреля 2012 г.

заседания кафедры алгебры и математической логики
Протокол №10 от 3 апреля 2012 г.

Автор-составитель
доктор физ.-мат. наук, доц. С.Н. Тронин

Научный редактор
доктор физ.-мат. наук, профессор С.М. Скрябин

Рецензент
кандидат физ.-мат. наук, доц. А.Н. Абызов

Лекции по алгебре. Семестр 2. Выпуск II. **Жорданова нормальная форма матрицы:** Учебно-методическое пособие / С.Н. Тронин. — Казань: Казанский (Приволжский) федеральный университет, 2012. — 78 с.

Данное учебно-методическое пособие пособие предназначено для студентов-математиков первого курса университета, изучающих алгебру. Оно представляет собой обработанные записи лекций, неоднократно читавшихся автором во втором семестре, и издается в виде нескольких выпусков. Во втором выпуске излагается теория жордановой нормальной формы матрицы линейного оператора. Кроме того, в этот же выпуск помещена справочная информация о векторных пространствах, группах, кольцах, полях и многочленах. Это те сведения из лекций первого семестра, которые используются во втором семестре. Содержание данного пособия полностью соответствует программе курса “Алгебра” для студентов-математиков, действующей в Казанском (Приволжском) федеральном университете.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Данное учебно-методическое пособие пособие предназначено для студентов-математиков первого курса университета, изучающих алгебру. Оно представляет собой обработанные записи лекций, неоднократно читавшихся автором во втором семестре, и издается в виде нескольких выпусков. В первом выпуске излагаются основы теории линейных отображений и линейных операторов.

Данный второй выпуск посвящен жордановой нормальной форме линейного оператора, и, кроме того, содержит справочную информацию по материалу первого семестра. Теорема о жордановой нормальной форме — это одна из важнейших теорем всего курса алгебры, и, безусловно, технически самая трудная теорема курса. В полном объеме, со всеми подготовительными результатами, она излагается в данном пособии на протяжении трех параграфов второй главы. В процессе доказательства используются многие факты, содержащиеся в первом выпуске пособия.

Отметим также, что в третьем выпуске содержится теория евклидовых и унитарных пространств, а также теория линейных операторов, действующих на таких пространствах.

Содержание данного учебного пособия — записи лекционного курса. Это значит, что излагается минимум того, что известно, только самые основные понятия, примеры и теоремы, без которых совершенно невозможно обойтись. Студент, желающий узнать больше, должен обратиться к книгам более солидного объема. Прежде всего это учебник А.И.Кострикина [9], [10], [11]. Второй том этого учебника [10] — это та книга, которую можно рекомендовать в первую очередь для основательного изучения материала, относящегося ко всем выпускам нашего пособия. Необходимо также отметить предшествовавшую [10] книгу высокого уровня [8], из которой можно узнать, например, много интерес-

ного о приложениях линейной алгебры в геометрии и физике. В списке литературы мы приводим также несколько других учебников (далеко не все, имеющиеся в наличии). Нельзя не упомянуть классические книги И.М. Гельфанда [3], В.А.Ильина и Э.Г.Позняка [5], А.Г. Куроша [13], и А.И. Мальцева [14]. Из более современных учебников отметим книги Э.Б. Винберга [2], В.А. Артамонова [1] и Г.С. Шевцова [15]. Наконец, в список литературы включено несколько учебных пособий сотрудников кафедры алгебры и математической логики Казанского федерального университета [4], [6], [7], [17].

Материал первого выпуска данного учебного пособия [16] предполагается известным, и будет использоваться, как правило, без особых оговорок.

Определения, примеры, теоремы, леммы, следствия и формулы в данном учебном пособии нумеруются с помощью трех цифр (чисел), из которых первое означает номер главы, второе — номер параграфа, и третье — номер определения, теоремы и т.п. внутри данного параграфа. При этом каждый из перечисленных видов нумеруемых объектов автоматически нумеруется отдельно и независимо от остальных (особенность издательской системы LaTeX). Например, в одном и том же параграфе 2.5 второй главы может присутствовать определение 2.5.2, теорема 2.5.2 и формула (2.5.2). Опыт показывает, что к этому можно легко привыкнуть. Доказательства заканчиваются символом $\square$.

Содержание данного пособия полностью соответствует программе курса “Алгебра” для студентов-математиков, действующей в Казанском (Приволжском) федеральном университете.

ГЛАВА II. ЖОРДАНОВА НОРМАЛЬНАЯ ФОРМА МАТРИЦЫ

В этой главе будет показано, к какому простейшему виду можно привести квадратную матрицу A с помощью преобразования подобия: $A \mapsto B^{-1}AB$. В случае, если рассматриваются матрицы над полем комплексных чисел, то такой простейший вид всегда существует, и называется *жордановой нормальной формой* матрицы A. Название происходит от фамилии автора, обнаружившего этот факт, известного французского математика К. Жордана (1838-1922).

Доказательство основного результата будет проводиться поэтапно, каждый параграф данной главы (кроме последнего) можно рассматривать как очередной этап доказательства.

2.1. Корневые подпространства

Пусть V — векторное пространство над полем K, $\mathcal{A} : V \to V$ — линейный оператор, и $\lambda \in K$. Рассмотрим множество $V(\lambda)$, состоящее из всех $v \in V$, для которых найдется целое $k \geq 0$, такое, что $(\mathcal{A}-\lambda\mathcal{E})^k v = 0$.

Лемма 2.1.1. *$V(\lambda)$ является подпространством векторного пространства V, инвариантным относительно $\mathcal{A}$ (а следовательно, и относительно любого $\mathcal{A} - \gamma\mathcal{E}$). $V(\lambda) \neq \{0\}$ тогда и только тогда, если λ есть собственное значение оператора $\mathcal{A}$, и в этом случае $V^\lambda \subseteq V(\lambda)$.*

Доказательство. Пусть $v_1, v_2 \in V(\lambda)$, $\alpha_1, \alpha_2 \in K$. Тогда найдутся такие целые $k_1, k_2 \geq 0$, что $(\mathcal{A} - \lambda\mathcal{E})^{k_1} v_1 = 0$ и $(\mathcal{A} - \lambda\mathcal{E})^{k_2} v_2 = 0$. Если $k = \max(k_1, k_2)$, то $(\mathcal{A}-\lambda\mathcal{E})^k v_1 = 0$ и $(\mathcal{A}-\lambda\mathcal{E})^k v_2 = 0$, откуда следует, что

$(\mathcal{A}-\lambda\mathcal{E})^k(\alpha_1 v_1+\alpha_2 v_2)=0$. Это означает, что $V(\lambda)$ есть подпространство векторного пространства V.

Проверим инвариантность этого подпространства относительно оператора A. Начнем с очевидного равенства: $\mathcal{A}(\mathcal{A}-\lambda\mathcal{E})=\mathcal{A}^2-\lambda\mathcal{A}=(\mathcal{A}-\lambda\mathcal{E})\mathcal{A}$. Отсюда легко следует, что для любого целого неотрицательного k справедливо равенство $\mathcal{A}(\mathcal{A}-\lambda\mathcal{E})^k=(\mathcal{A}-\lambda\mathcal{E})^k\mathcal{A}$. Теперь, если $v=\in V(\lambda)$, то для некоторого k справедливо равенство $(\mathcal{A}-\lambda\mathcal{E})^k v=0$. Но тогда $(\mathcal{A}-\lambda\mathcal{E})^k\mathcal{A}v=\mathcal{A}(\mathcal{A}-\lambda\mathcal{E})^k v=0$. Это означает, что $\mathcal{A}v\in V(\lambda)$.

Если λ — собственное значение $\mathcal{A}$, то ясно, что $\{0\}\neq V^\lambda\subseteq V(\lambda)$. Обратно, пусть $V(\lambda)\neq\{0\}$. Это значит, что существуют вектор $v\neq 0$, и число $k\geq 1$ такие, что $(\mathcal{A}-\lambda\mathcal{E})^k v=0$, но $(\mathcal{A}-\lambda\mathcal{E})^{k-1}v\neq 0$. Тогда вектор $(\mathcal{A}-\lambda\mathcal{E})^{k-1}v$ является собственным вектором $\mathcal{A}$, отвечающим собственному значению λ. □

Определение 2.1.1. Если λ — собственное значение оператора $\mathcal{A}$, то пространство $V(\lambda)$ называется *корневым подпространством* оператора $\mathcal{A}$.

Определение 2.1.2. Оператор $\mathcal{B}: V\to V$ называется *нильпотентным*, если существует такое $k\geq 1$, что $\mathcal{B}^k=0$. (Иными словами, для каждого $v\in V$ должно быть $\mathcal{B}^k v=0$.) Если $\mathcal{B}^k=0$, но $\mathcal{B}^{k-1}\neq 0$, то число k называется *степенью нильпотентности* оператора $\mathcal{B}$.

Пример 2.1.1. Рассмотрим линейный оператор, матрица которого в некотором базисе имеет слкдующий вид:

$$A=\begin{pmatrix} 0 & a_{1,2} & a_{1,3} & \dots & a_{1,n} \\ 0 & 0 & a_{2,3} & \dots & a_{2,n} \\ 0 & 0 & 0 & \dots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Тогда $A^n = 0$ (убедитесь в этом!), и данный оператор является нильпотентным.

Лемма 2.1.2. *Пусть λ — собственное значение оператора $\mathcal{A}$. Тогда ограничение $\mathcal{A} - \lambda\mathcal{E}$ на $V(\lambda)$ является нильпотентным оператором.*

Доказательство. Пусть $v_1, \dots, v_m$ — некоторый базис подпространства $V(\lambda)$. Тогда найдутся целые положительные числа $k_1, \dots, k_m$ такие, что $(\mathcal{A} - \lambda\mathcal{E})^{k_j} v_j = 0$ для всех $j = 1, \dots, m$. Если взять $k = \max(k_1, \dots, k_m)$, то $(\mathcal{A} - \lambda\mathcal{E})^k v_j = 0$ для всех $j = 1, \dots, m$. Произвольный вектор $v \in V(\lambda)$ можно представить в виде $v = \sum\limits_{j=1}^{m} \alpha_j v_j$. Поэтому

$$(\mathcal{A} - \lambda\mathcal{E})^k v = (\mathcal{A} - \lambda\mathcal{E})^k (\sum_{j=1}^{m} \alpha_j v_j) = \sum_{j=1}^{m} \alpha_j ((\mathcal{A} - \lambda\mathcal{E})^k v_j) = 0.$$

Но это именно то, что и утверждается в формулировке леммы. В частности, мы видим, что степень нильпотентности ограничения $\mathcal{A} - \lambda\mathcal{E}$ на $V(\lambda)$ не превосходит k. □

Лемма 2.1.3. *Пусть $\mathcal{B} : V \to V$ — нильпотентный оператор, причем $\dim(V) = n$. Тогда степень нильпотентности $\mathcal{B}$ не превосходит n. В частности, $\mathcal{B}^n = 0$.*

Заметим, что это утверждение справедливо для произвольного поля K.

Доказательство. Пусть $\mathcal{B}^k = 0$, но $\mathcal{B}^{k-1} \neq 0$. Это значит, что для каждого $x \in V$ выполнено равенство $\mathcal{B}^k x = 0$, но существует вектор $y \in V$ такой, что $\mathcal{B}^{k-1} y \neq 0$. Обозначим через $\mathcal{B}^m V$ образ оператора $\mathcal{B}^m$, т.е. множество всех векторов вида $\mathcal{B}^m v$, где v пробегает все пространство V. Тогда имеет место цепочка включений:

$$V \supseteq \mathcal{B}V \supseteq \mathcal{B}^2 V \supseteq \dots \mathcal{B}^m V \supseteq \mathcal{B}^{m+1} V \supseteq \dots \supseteq \mathcal{B}^{k-1} V \supset \mathcal{B}^k V = \{0\}.$$

Допустим, что все включения $\mathcal{B}^{i-1}V \supseteq \mathcal{B}^iV$ при $0 \leq i \leq m$ строгие, а $\mathcal{B}^mV = \mathcal{B}^{m+1}V$. Если $\mathcal{B}^mV = \{0\}$, т.е. $m = k$, то получим строго убывающую цепочку неравенств:

$$\dim(V) = n > \dim(\mathcal{B}V) > \dim(\mathcal{B}^2V) > \ldots > \dim(\mathcal{B}^{k-1}) > \dim(\mathcal{B}^kV) = 0.$$

Отсюда следует, что $k \leq n$. Допустим теперь, что $\mathcal{B}^mV = \mathcal{B}^{m+1}V$, и это подпространство отлично от нуля, так что $m+1 < k$. Из того, что $\mathcal{B}^mV \subseteq \mathcal{B}^{m+1}V$ следует, что для любого $v \in V$ существует $w \in V$ такой, что $\mathcal{B}^mv = \mathcal{B}^{m+1}w$. Можно выбрать вектор v таким, что $\mathcal{B}^{k-1}v \neq 0$. Так как $m < k-1$, то тем более $\mathcal{B}^mv \neq 0$. Применим теперь к левой и правой частям равенства $\mathcal{B}^mv = \mathcal{B}^{m+1}w$ оператор $\mathcal{B}^{k-m-1}$. Получим равенство $\mathcal{B}^{k-1}v = \mathcal{B}^kw$. Слева по предположению стоит ненулевой вектор, но $\mathcal{B}^kw = 0$, поскольку k есть степень нильпотентности $\mathcal{B}$. Полученное противоречие доказывает лемму. □

Теорема 2.1.1. *Пусть V — векторное пространство над полем действительных или комплексных чисел, $\mathcal{A} : V \to V$ — линейный оператор, и пусть $\chi_\mathcal{A} = (-1)^n(x-\lambda_1)^{n_1}\ldots(x-\lambda_m)^{n_m}$, где $\lambda_1, \ldots, \lambda_m \in K$ — все собственные значения $\mathcal{A}$. Тогда имеет место разложение в прямую сумму:*

$$V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_m).$$

Заметим, что предположение относительно $\chi_\mathcal{A}$ автоматически выполняется в случае, когда $K = \mathbb{C}$ — поле комплексных чисел.

Доказательство. Пусть $\chi_i(x) = \prod\limits_{j=1, j\neq i}^{m}(x-\lambda_j)^{n_j}$. Тогда НОД$(\chi_1, \ldots, \chi_m) = 1$, откуда следует, что существуют многочлены $h_1, \ldots, h_m$ такие, что

$$\sum_{i=1}^{m} \chi_i(x)h_i(x) = 1.$$

После подстановки вместо x оператора $\mathcal{A}$ получим:

$$\sum_{i=1}^{m} \chi_i(\mathcal{A}) h_i(\mathcal{A}) = \mathcal{E}. \tag{2.1.1}$$

Положим $W_i = \chi_i(\mathcal{A}) h_i(\mathcal{A}) V$. Так как $(x - \lambda_i)^{n_i} \chi_i(x) = (-1)^n \chi_{\mathcal{A}}(x)$, и $\chi_{\mathcal{A}}(\mathcal{A}) = 0$ (теорема Гамильтона-Кэли, именно здесь и требуется предполагать, что все происходит над полем действительных или комплексных чисел), то
$(\mathcal{A} - \lambda_i \mathcal{E})^{n_i} W_i = \{0\}$. Отсюда следует, что $W_i \subseteq V(\lambda_i)$ для всех i, $1 \leq i \leq m$. Применяя (2.1.1) к произвольному вектору $v \in V$, получим

$$v = v_1 + \cdots + v_m,$$

где $v_i = \chi_i(\mathcal{A}) h_i(\mathcal{A}) v \in W_i \subseteq V(\lambda_i)$ для каждого i. Отсюда следует, что

$$V = \sum_{i=1}^{m} W_i = \sum_{i=1}^{m} V(\lambda_i).$$

Остается показать, что эта сумма прямая.

Пусть $v_i \in V(\lambda_i)$, $1 \leq i \leq m$, $v_1 + \cdots + v_m = 0$, и некоторый $v_i \neq 0$. Тогда $v_i = \sum\limits_{j, j \neq i} (-v_j) = \sum\limits_{j, j \neq i} v'_j$, где $v'_j = -v_j \in V(\lambda_j)$.

Так как ограничение $\mathcal{A} - \lambda_j \mathcal{E}$ на $V(\lambda_j)$ нильпотентно для всех j, $1 \leq j \leq m$, и степень нильпотентности этого оператора не превышает размерности $V(\lambda_j)$, которая, в свою очередь, не больше размерности пространства V, равной n, то $(\mathcal{A} - \lambda_i \mathcal{E})^n v_i = 0$ и $(\mathcal{A} - \lambda_j \mathcal{E})^n v'_j = 0$ при $j \neq i$. Положим $h(x) = \prod\limits_{j \neq i} (x - \lambda_j)^n$. Тогда многочлены $(x - \lambda_i)^n$ и $h(x)$ взаимно просты, откуда следует, что найдутся такие многочлены $a(x)$ и $b(x)$, что

$$a(x)(x - \lambda_i)^n + b(x) h(x) = 1.$$

Отсюда получаем

$$a(\mathcal{A})(\mathcal{A} - \lambda_i \mathcal{E})^n + b(\mathcal{A}) h(\mathcal{A}) = \mathcal{E}.$$

Применим левую и правую части этого равенства к вектору $v_i \neq 0$. Уже известно, что $(\mathcal{A}-\lambda_i\mathcal{E})^n v_i = 0$. Рассмотрим вектор $b(\mathcal{A})h(\mathcal{A})v_i$, а точнее, даже вектор $h(\mathcal{A})v_i$. Заменим v_i на $\sum\limits_{j,j\neq i} v'_j$, и получим $\sum\limits_{j,j\neq i} h(\mathcal{A})v'_j$. Но так как для каждого $j \neq i$ оператор $h(\mathcal{A})$ содержит множитель $(\mathcal{A} - \lambda_j\mathcal{E})^n$, эти множители можно переставлять, и так как по леммам 2.1.2 и 2.1.3 выполняется равенство $(\mathcal{A} - \lambda_j\mathcal{E})^n v'_j = 0$, то $h(\mathcal{A})v'_j = 0$. Тем более, $b(\mathcal{A})h(\mathcal{A})v'_j = 0$, для всех $j \neq i$, откуда следует $b(\mathcal{A})h(\mathcal{A})v_i = 0$. Теперь из равенства

$$a(\mathcal{A})(\mathcal{A} - \lambda_i\mathcal{E})^n v_i + b(\mathcal{A})h(\mathcal{A})v_i = \mathcal{E}v_i = v_i$$

получаем $0 = v_i$, что противоречит сделанному ранее предположению $v_i \neq 0$. Таким образом, сумма

$$V = \sum_{i=1}^{m} V(\lambda_i)$$

действительно является прямой суммой.

Покажем еще, что $V(\lambda_i) = W_i$ для всех i. С одной стороны, $W_i \subseteq V(\lambda_i)$. С другой стороны, так как $V = \sum\limits_{i=1}^{m} W_i$, то произвольный $v \in V(\lambda_i)$ представляется в виде суммы $v = w_1 + \cdots + w_i + \cdots + w_m$, где $w_j \in W_j \subseteq V(\lambda_j)$. Отсюда $w_1 + \cdots + (w_i - v) + \cdots + w_m = 0$ и, по определению прямой суммы, $w_1 = 0, \ldots, w_i - v = 0, \ldots, w_m = 0$. Таким образом, $v = w_i \in W_i$. Теорема доказана. □

Далее мы сформулируем и докажем три дополнения к теореме 2.1.1. Конечно, можно было бы сформулировать саму эту теорему таким образом, чтобы то, что мы назыываем ниже “дополнниями”, стало частью теоремы (как это и сделано в учебнике А.И. Кострикина [10]). Но тогда эта теорема стала бы довольно громоздкой, что вызвало бы определенные проблемы как при составлении вопросов для экзаменационных билетов, так и на экзамене. Чтобы несколько уменьшить эти проблемы,

мы разделили теорему на части, с тем, чтобы вместо одного вопроса в билете появилось бы по крайней мере два.

Итак, пусть сохраняются все обозначения и соглашения из доазательства теоремы. Рассмотрим $V(\lambda_i) = W_i = \chi_i(\mathcal{A})h_i(\mathcal{A})V$. Это означает, что

$$(\mathcal{A} - \lambda_i\mathcal{E})^{n_i}V(\lambda_i) = \{0\}.$$

Отсюда выводится следующее утверждение:

Дополнение 1. *При тех же предположениях и обозначениях единственным собственным значением ограничения оператора $\mathcal{A}$ на $V(\lambda_i)$ является λ_i, $1 \leq i \leq m$.*

Доказательство. Так как у оператора $\mathcal{A}$ существует собственный вектор $w \neq 0$, отвечающий собственному значению λ_i, и $w \in V(\lambda_i)$, то этот вектор будет также собственным вектором ограничения $\mathcal{A}$ на $V(\lambda_i)$, отвечающим собственному значению λ_i.

Допустим, что существует какое-то другое собственное значение $\gamma \neq \lambda_i$ ограничения $\mathcal{A}$ на $V(\lambda_i)$. Это значит, что найдется ненулевой $v \in V(\lambda_i)$ такой, что $(\mathcal{A} - \gamma\mathcal{E})v = 0$. Многочлены $x - \gamma$ и $(x - \lambda_i)^{n_i}$ взаимно просты. Отсюда следует, что найдутся многочлены f и g такие, что $f(x)(x - \gamma) + g(x)(x - \lambda_i)^{n_i} = 1$. Подставляя опратор $\mathcal{A}$, получим равенство:

$$f(\mathcal{A})(\mathcal{A} - \gamma\mathcal{E}) + g(\mathcal{A})(\mathcal{A} - \lambda_i\mathcal{E})^{n_i} = \mathcal{E}.$$

Применим оператор, стоящий в левой части, к ненулевому вектору v. Но $(\mathcal{A} - \gamma\mathcal{E})v = 0$ по выбору v, и $(\mathcal{A} - \lambda_i\mathcal{E})^{n_i}v = 0$, так как, по доказанному выше, $(\mathcal{A} - \lambda_i\mathcal{E})^{n_i}V(\lambda_i) = \{0\}$. Отсюда получаем, что $v = 0$, в противоречии с выбором v. Утверждение доказано. □

Дополнение 2. *При тех же предположениях и обозначениях* $\dim(V(\lambda_i)) = n_i$, $1 \leq i \leq m$.

Доказательство. Пусть $k_i = \dim(V(\lambda_i)), 1 \leq k \leq m$. Тогда $k_1 + \cdots + k_m = n = n_1 + \cdots + n_m$. Выберем в каждом подпространстве $V(\lambda_i)$ какой-нибудь базис, и пусть A_i есть матрица оператора $\mathcal{A}|_{V(\lambda_i)}$ в этом базисе. Это матрица размера $k_i \times k_i$. Объединение всех выбранных базисов подпространств $V(\lambda_i)$ будет базисом всего пространства V (так как уже доказано, что $V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_m)$), и матрицей оператора $\mathcal{A}$ в этом базисе будет матрица

$$\begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_m \end{pmatrix}$$

Очевидно, что характеристически многочлен этой матрицы (а следовательно, и оператора $\mathcal{A}$) равен произведению характеристических многочленов матриц A_i, $1 \leq i \leq m$. Но характеристический многочлен матрицы A_i как матрицы оператора $\mathcal{A}|_{V(\lambda_i)}$ должен иметь своими корнями в точности собственный значения этого оператора. В Дополнении 1 уже было показано, что такое собственное значение только одно, и это само число λ_i. Так как степень характеристического многочлена матрицы равна порядку матрицы, то это будет многочлен $(-1)^{k_i}(x - \lambda_i)^{k_i}$. Таким образом, мы вычислили характеристический многочлен оператора $\mathcal{A}$, и он оказался равным

$$\begin{aligned}(-1)^{k_1}(x - \lambda_1)^{k_1} \dots (-1)^{k_m}(x - \lambda_i)^{k_m} = \\ (-1)^{k_1 + \cdots + k_m}(x - \lambda_1)^{k_1} \dots (x - \lambda_i)^{k_m} = \\ (-1)^n(x - \lambda_1)^{k_1} \dots (x - \lambda_i)^{k_m}.\end{aligned}$$

Но тот же самый многочлен по условию равен $(-1)^n(x - \lambda_1)^{n_1} \dots (x - \lambda_i)^{n_m}$. Так как разложение многочленов на неприводимые множители однозначно, то отсюда следует, что для всех i, $1 \leq i \leq m$, выполняются равенства $n_i = k_i = \dim(V(\lambda_i))$. □

Дополнение 3. *Ограничение линейного оператора $\mathcal{A}-\lambda_i\mathcal{E}$ на инвариантное подпространство $\bigoplus\limits_{j,j\neq i} V(\lambda_j)$ есть невырожденный линейный оператор.*

Доказательство. Достаточно убедиться, что ядро ограничения $\mathcal{A}-\lambda_i\mathcal{E}$ на $\bigoplus\limits_{j,j\neq i} V(\lambda_j)$ равно нулю (см. теорему 1.3.7). Пусть $v_j \in V(\lambda_j)$, $j \neq i$, и $(\mathcal{A}-\lambda_i\mathcal{E})\sum\limits_{j,j\neq i} v_j = 0$. Так как каждое подпространство $V(\lambda_j)$ инвариантно относительно $\mathcal{A}-\lambda_i\mathcal{E}$, то получаем равенство

$$\sum_{j,j\neq i}(\mathcal{A}-\lambda_i\mathcal{E})v_j = 0,$$

в котором каждый вектор $(\mathcal{A}-\lambda_i\mathcal{E})v_j$ принадлежит слагаемому $V(\lambda_j)$ прямой суммы $\bigoplus\limits_{j,j\neq i} V(\lambda_j)$. По определению прямой суммы это означает, что $(\mathcal{A}-\lambda_i\mathcal{E})v_j = 0$ для каждого $j \neq i$. Таким образом, $\mathcal{A}v_j = \lambda_i v_j$. Если бы $v_j \neq 0$, то это означало бы, что у ограничения $\mathcal{A}$ на $V(\lambda_j)$ имеются два различных собственных значения, λ_i и λ_j, а это противоречит Дополнению 1. □

2.2. Существование жордановой нормальной формы

Определение 2.2.1. *Жордановой клеткой* порядка k, соответствующей собственному значению λ, называется $k \times k$-матрица

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

Заметим, что $J_k(\lambda) + \alpha E_k = J_k(\lambda+\alpha)$.

Пример 2.2.1. В случае $k = 1, 2, 3$ соответствующие жордановы клетки таковы:

$$J_1(\lambda) = (\lambda), \qquad J_2(\lambda) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \qquad J_3(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Определение 2.2.2. Говорят, что матрица A является *жордановой* (или имеет *жорданову нормальную форму*), если A есть блочно-диагональная матрица, на диагонали которой стоят жордановы клетки, т.е.

$$A = \begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{k_2}(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ddots & J_{k_l}(\lambda_l) \end{pmatrix}$$

Здесь порядки жордановых клеток $k_1, \dots, k_l$ не обязательно различны, и элементы поля $\lambda_1, \dots, \lambda_l$ также не обязательно различны.

Определение 2.2.3. Говорят, что квадратная матрица A *приводится к жордановой нормальной форме*, если существует невырожденная матрица B такая, что матрица $B^{-1}AB$ является жордановой (т.е. имеет жорданову номальную форму).

Приведение матрицы A к жордановой нормальной форме — это процесс нахождения жордановой матрицы вида $B^{-1}AB$. Иногда кроме вычисления жордановой матрицы требуется и нахождение матрицы B. На языке линейных операторов это означает нахождение базиса, в котором матрица данного линейного оператора имеет жорданову нормальную форму.

Определение 2.2.4. Пусть $\mathcal{A} : V \to V$ — линейный оператор. Базис пространства V называется *жордановым базисом* для $\mathcal{A}$, если матрица $\mathcal{A}$ в этом базисе является жордановой.

Доказательство существования жордановой нормальной формы для любой квадратной матрицы над полем комплексных чисел, которое дается ниже, состоит, по сути, именно в обосновании факта существования жорданова базиса для произвольного линейного оператора. Это делается в несколько этапов, путем сведения задачи к частному случаю.

Лемма 2.2.1. *Пусть $V = U_1 \oplus \cdots \oplus U_l$, где подпространства U_i инвариантны относительно оператора $\mathcal{A}$, и A_i есть матрица ограничения $\mathcal{A}$ на U_i, $1 \le i \le l$. Если каждая из матриц A_i приводится к жордановой нормальной форме, то и матрица $\mathcal{A}$ также приводится к жордановой нормальной форме. Если J_i — жорданова нормальная форма матрицы A_i, то жорданова нормальная форма матрицы A имеет вид:*

$$\begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_l \end{pmatrix}$$

Доказательство. Выберем базис V, состоящий из объединения базисов инвариантных подпространств $U_1, \dots, U_l$. Тогда матрицей $\mathcal{A}$ в этом базисе будет матрица

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_l \end{pmatrix}$$

Допустим, что $B_i^{-1} A_i B_i = J_i$ для $1 \le i \le l$. Это равносильно существованию в каждом из подпространств U_i жорданова базиса для оператора

$\mathcal{A}|_{U_i}$. Объединение этих базисов снова будет базисом пространства V, в котором матрица оператора $\mathcal{A}$ будет иметь требуемый вид. В частности, если положить

$$B = \begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_l \end{pmatrix}$$

то

$$B^{-1} = \begin{pmatrix} B_1^{-1} & 0 & \dots & 0 \\ 0 & B_2^{-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_l^{-1} \end{pmatrix}$$

и

$$J = B^{-1}AB = \begin{pmatrix} B_1^{-1}AB_1 & 0 & \dots & 0 \\ 0 & B_2^{-1}AB_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_l^{-1}AB_l \end{pmatrix} = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_l \end{pmatrix}$$

Блоки $J_1, \dots, J_l$ — это блочно-диагональные матрицы, на диагоналях которых расположены жордановы клетки. Очевидно, что и матрица J является блочно-диагональной матрицей, на блочной диагонали которой располагаются все те жордановы клетки, которые были в матрицах $J_1, \dots, J_l$. По определению, это означает, что матрица J является жордановой. □

Из доказательства леммы 2.2.1 следует также, что если для каждого оператора $\mathcal{A}|_{U_i} : U_i \to U_i$ в подпространстве U_i существует жорданов базис, то объединение всех этих базисов подпространств $U_1, \dots, U_l$ будет жордановым базисом всего пространства V.

Лемма 2.2.2. *Пусть A — некоторая матрица порядка n над произвольным полем K, и $\lambda \in K$. Если матрица $A - \lambda E_n$ приводится к*

жордановой нормальной форме, то A также можно привести к жордановой нормальной форме.

Доказательство. Допустим, что существует такая обратимая матрица B, что

$$B^{-1}(A-\lambda E)B = J = \begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{k_2}(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{k_l}(\lambda_l) \end{pmatrix}$$

Тогда

$$J = B^{-1}(A-\lambda E)B = B^{-1}AB - B^{-1}(\lambda E)B = B^{-1}AB - \lambda B^{-1}B = B^{-1}AB - \lambda E.$$

Отсюда получаем $B^{-1}AB = J + \lambda E$. Но матрицу λE можно представить в виде

$$\begin{pmatrix} \lambda E_{k_1} & 0 & \dots & 0 \\ 0 & \lambda E_{k_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda E_{k_l} \end{pmatrix}$$

Вспоминая, что $J_{k_i} + \lambda E_{k_i} = J_{k_i}(\lambda_i + \lambda)$, получаем равенство:

$$J + \lambda E = \begin{pmatrix} J_{k_1}(\lambda_1 + \lambda) & 0 & \dots & 0 \\ 0 & J_{k_2}(\lambda_2 + \lambda) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{k_l}(\lambda_l + \lambda) \end{pmatrix}$$

Справа в этом равенстве стоит матрица, находящаяся в жордановой нормальной форме. □

Из доказательства леммы 2.2.2 следует, что матрица B, приводящая к жордановой нормальной форме матрицу $A - \lambda E$, приводит к жордановой нормальной форме и матрицу A. Это также означает, что если у

некоторого оператора $\mathcal{A} - \mathcal{E}$ имеется жорданов базис, то этот же базис будет жордановым базисом и для оператора $\mathcal{A}$.

Рассмотрим случай $K = \mathbb{C}$, и пусть дан линейный оператор $\mathcal{A} : V \to V$, и $V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_m)$ — разложение в прямую сумму корневых подпространств этого оператора. Применяя к этому разложению лемму 2.2.1, получаем, что задача о существовании жордановой нормальной формы для матрицы оператора $\mathcal{A}$ сводится к случаю, когда $\mathcal{A}$ имеет единственное собственное значение λ и $V = V(\lambda)$.

Из леммы 2.2.2 теперь следует, что при этих предположениях достаточно научиться приводить к жордановой нормальной форме матрицу оператора $\mathcal{A} - \lambda\mathcal{E}$. Но этот оператор при сделанных предположениях (т.е. при $V = V(\lambda)$) является нильпотентным. Таким образом, существование жордановой нормальной формы для матрицы произвольного оператора $\mathcal{A}$ над полем $\mathbb{C}$ будет доказано, если удастся доказать, что такая форма существует для матриц нильпотентных операторов.

Пусть $\mathcal{B} : V \to V$ — некоторый нильпотентный оператор, причем поле K можно считать произвольным.

Лемма 2.2.3. *Единственное собственное значение нильпотентного оператора — элемент $0 \in K$. Над полем комплексных чисел характеристический многочлен нильпотентного оператора равен $(-1)^n x^n$ (где n — размерность пространства). Обратно, если у какого-то оператора $\mathcal{B}$ над полем действительных или комплексных чисел характеристический многочлен равен $(-1)^n x^n$, то этот оператор является нильпотентным.*

Доказательство. В самом деле, если $\mathcal{B}v = \lambda v$, $v \neq 0$, то, ввиду нильпотентности $\mathcal{B}$, найдется такое $r > 0$, что $\mathcal{B}^r v = 0$, но $\mathcal{B}^{r-1} v \neq 0$. Применим оператор $\mathcal{B}^{r-1}$ к обеим частям равенства $\mathcal{B}v = \lambda v$, и получим

$0 = \mathcal{B}^r v = \lambda \mathcal{B}^{r-1} v$. Отсюда следует, что $\lambda = 0$.

Предположим, что мы имеем дело с оператором, действующим на векторном пространстве над полем комплексных чисел. Тогда характеристический многочлен этого оператора записывается в виде $(-1)^n(x - \lambda_1)^{n_1} \dots (x - \lambda_m)^{n_m}$, где $\lambda_1, \dots, \lambda_m$ — все различные собственные значения оператора. Однако, как только что показано, собственное значение только одно, и это нуль. Поэтому характеристический многочлен равен $(-1)^n x^n$.

Обратно, если характеристический многочлен оператора $\mathcal{B}$ равен $(-1)^n x^n$, и оператор действует на векторном пространстве над полем действительных или комплексных чисел, то можно применить теорему Гамильтона-Кэли, из которой следует, что $(-1)^n \mathcal{B}^n = 0$, т.е. $\mathcal{B}^n = 0$. $\square$

Заметим, что на самом деле утверждения этой леммы справедливы для операторов над каким угодно полем. Однако у нас пока нет возможности доказать теорему Гамильтона-Кэли для произвольного поля.

Пример 2.2.2. Рассмотрим оператор, в некотором базисе имеющий верхнетреугольную матрицу с нулями на главной диагонали, т.е. матрицу вида

$$\begin{pmatrix} 0 & a_{1,2} & a_{1,3} & a_{1,4} & \dots & a_{1,n-1} & a_{1,n} \\ 0 & 0 & a_{2,3} & a_{2,4} & \dots & a_{2,n-1} & a_{2,n} \\ 0 & 0 & 0 & a_{3,4} & \dots & a_{3,n-1} & a_{3,n} \\ 0 & 0 & 0 & 0 & \dots & a_{4,n-1} & a_{4,n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & a_{n-1,n} \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

Прямыми вычислениями можно убедиться (см. пример 2.1.1), что эта матрица нильпотентна при любом выборе поля K. Если же поле, в котором содержатся компоненты матрицы, является полем действительных

или комплексных чисел, то вычисления не нужны, так как характеристический многочлен матрицы равен $(-1)^n x^n$ и можно применить доказанную выше лемму.

То же самое справедливо и для нижнетреугольных матриц.

Итак, если $\mathcal{B}$ есть нильпотентный оператор, то для каждого $v \in V$ либо $v = 0$, либо существует целое $k > 0$ такое, что $\mathcal{B}^k v = 0$.

Лемма 2.2.4. *Пусть $v \neq 0$, $\mathcal{B}^{k-1} v \neq 0$, и $\mathcal{B}^k v = 0$ для некоторого целого $k \geq 1$. Тогда векторы $v, \mathcal{B}v, \mathcal{B}^2 v, \ldots, \mathcal{B}^{k-1} v$ линейно независимы.*

Доказательство. Вспомним, что $\mathcal{B}^0 = \mathcal{E}$, так что $v = \mathcal{B}^0 v$, и рассмотрим некоторую нетривиальную линейную зависимость:

$$\alpha_0 \mathcal{B}^0 v + \alpha_1 \mathcal{B} v + \cdots + \alpha_{k-1} \mathcal{B}^{k-1} v = 0.$$

Пусть $\alpha_0 = \cdots = \alpha_{r-1} = 0$, $\alpha_r \neq 0$. Разделив на α_r, получим зависимость вида:

$$\mathcal{B}^r v + \gamma_1 \mathcal{B}^{r+1} v + \cdots + \gamma_{k-1-r} \mathcal{B}^{k-1} v = 0.$$

Применим к обеим частям этого равенства оператор $\mathcal{B}^{k-1-r}$. При $j > 0$ получим $\mathcal{B}^{k-1-r} \mathcal{B}^{r+j} v = \mathcal{B}^{j-1} \mathcal{B}^k v = 0$, так как $\mathcal{B}^k v = 0$ по предположению. Поэтому от всей последней линейной зависимости остается только равенство: $\mathcal{B}^{k-1} v = 0$, которое противоречит предположению о том, что $\mathcal{B}^{k-1} v \neq 0$. Лемма доказана. □

Определение 2.2.5. Пусть, как и выше, $\mathcal{B}$ — нильпотентный оператор, $v \in V$, $v \neq 0$, $\mathcal{B}^{k-1} v \neq 0$, и $\mathcal{B}^k v = 0$ для некоторого целого $k \geq 1$. Обозначим подпространство $\langle v, \mathcal{B}v, \mathcal{B}^2 v, \ldots, \mathcal{B}^{k-1} v \rangle$ через $C(v, k)$, и будем называть его *циклическим подпространством* нильпотентного оператора $\mathcal{B}$.

Следующая лемма является очень важной для дальнейшего.

Лемма 2.2.5. *Циклическое подпространство $C(v,k)$ является инвариантным относительно оператора $\mathcal{B}$. Матрица ограничения $\mathcal{B}$ на $C(v,k)$ в базисе $\mathcal{B}^{k-1}v, \mathcal{B}^{k-2}v, \dots, \mathcal{B}v, v$ есть жорданова клетка $J_k(0)$ порядка k*

Доказательство. Для того, чтобы установить инвариантность $C(v,k)$ относительно оператора $\mathcal{B}$, достаточно убедиться, что при действии $\mathcal{B}$ на элементы базиса получаются элементы того же подпространства $C(v,k)$. Но это очевидно: при действии $\mathcal{B}$ на элементы $\mathcal{B}^{k-2}v, \dots, \mathcal{B}v, v$ получаются элементы того же базиса $\mathcal{B}^{k-1}v, \mathcal{B}^{k-2}v, \dots, \mathcal{B}v$, а при действии $\mathcal{B}$ на $\mathcal{B}^{k-1}v$ получается $\mathcal{B}^k v = 0$.

Вычислим матрицу ограничения $\mathcal{B}$ на $C(v,k)$ в указанном базисе. Положим $e_1 = \mathcal{B}^{k-1}v$, $e_2 = \mathcal{B}^{k-2}v$, …, $e_j = \mathcal{B}^{k-j}v$, …, $e_k = v$. Тогда

$$\mathcal{B}e_1 = 0, \mathcal{B}e_2 = e_1, \dots, \mathcal{B}e_j = e_{j-1}, \dots, \mathcal{B}e_k = e_{k-1}.$$

Соответствующая матрица такова:

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

Это жорданова клетка порядка k, отвечающая собственному значению 0 оператора $\mathcal{B}$. □

Теорема 2.2.1. *Пусть $\mathcal{B} : V \to V$ — нильпотентный оператор (поле K произвольно). Тогда существует представление в виде прямой суммы циклических подпространств:*

$$V = \bigoplus_{i=1}^{s} C(v_j, m_j) \tag{2.2.1}$$

Доказательство. Проведем индукцию по $n = \dim(V)$. В случае $n = 1$ у пространства V есть базис из одного вектора $v \neq 0$, причем по лемме 2.1.3 будем иметь $\mathcal{B}v = 0$. Таким образом, само пространство $V = \langle v \rangle$ — циклическое подпространство.

Допустим, что утверждение лммы выполняется для всех векторных пространств, размерности которых строго меньше n, и всех действующих на таких пространствах нильпотентных операторов. Пусть $\dim(V) = n > 1$. Из доказательства леммы 2.1.3 следует, что включение $\mathcal{B}V \subset V$ строгое, так что $\dim(\mathcal{B}V) < n$. Выберем любое подпространство U пространства V, содержащее $\mathcal{B}V$ и имеющее размерность $\dim(U) = n - 1$. Тогда U будет инвариантным относительно $\mathcal{B}$. В самом деле, если $u \in U \subset V$, то $\mathcal{B}u \in \mathcal{B}V \subseteq U$. Ограничение оператора $\mathcal{B}$ на инвариантное подпространство U есть линейный нильпотентный оператор на $(n-1)$-мерном пространстве U. Следовательно, к нему применимо предположение индукции, и

$$U = \bigoplus_{j=1}^{s} C(v_j, m_j),$$

где $C(v_j, m_j) = \langle v_j, \mathcal{B}v_j, \ldots, \mathcal{B}^{m_j-1}v_j \rangle$, $v_j \in U$, $\mathcal{B}^{m_j-1}v_j \neq 0$, $\mathcal{B}^{m_j}v_j = 0$. Подпространства $C(v_j, m_j)$ инвариантны относительно ограничения $\mathcal{B}$ на U, следовательно, они инвариантны и относительного самого оператора $\mathcal{B}$. Ясно, что можно выбрать индексы так, чтобы $m_1 \geq m_2 \geq \ldots \geq m_s$.

Если $v \in V$ — любой вектор, не содержащийся в U, то подпространство, порожденное вектором v и подпространством U, должно совпадать с V (так как $v \notin U$, то оно строго больше U, и потому его размерность строго больше $n - 1 = \dim(U)$, а значит, она равна $n = \dim(V)$). Таким образом, у нас есть базис всего пространства V, состоящий из вектора v, и из всех базисов всех подпространств $C(v_j, m_j)$. Вектор $\mathcal{B}v \in \mathcal{B}V \subseteq U$

можно выразить через базис подпространства U:

$$\mathcal{B}v = \sum_{j=1}^{s} \sum_{k=0}^{m_j-1} \alpha_{j,k} \mathcal{B}^k v_j.$$

После перегруппировки слагаемых получим:

$$\mathcal{B}v = \sum_{j=1}^{s} \alpha_{j,0} v_j + \sum_{j=1}^{s} \sum_{k=1}^{m_j-1} \alpha_{j,k} \mathcal{B}^k v_j = \sum_{j=1}^{s} \alpha_{j,0} v_j + \mathcal{B}(\sum_{j=1}^{s} \sum_{k=1}^{m_j-1} \alpha_{j,k} \mathcal{B}^{k-1} v_j).$$

Иными словами,

$$\mathcal{B}v = \sum_{j=1}^{s} \alpha_{j,0} v_j + \mathcal{B}u,$$

где $u = \sum\limits_{j=1}^{s} \sum\limits_{k=1}^{m_j-1} \alpha_{j,k} \mathcal{B}^{k-1} v_j \in U$.

Положим $v' = v - u$. Вектор v' не принадлежит подпространству U, так как из $v' \in U$ ввиду $u \in U$ следовало бы $v \in U$, а это не так. Следовательно, как и в случае с v, вектор v' вместе с подпространством U порождает все пространство V, и вместе с базисом U образует базис V. При этом

$$\mathcal{B}v' = \mathcal{B}(v - u) = \mathcal{B}v - \mathcal{B}u = \sum_{j=1}^{s} \alpha_{j,0} v_j + \mathcal{B}u - \mathcal{B}u = \sum_{j=1}^{s} \alpha_{j,0} v_j.$$

Если $\alpha_{j,0} = 0$ для всех j, $1 \leq j \leq s$, то $\mathcal{B}v' = 0$, и подпространство $\langle v' \rangle$ является циклическим подпространством $C(v', 1)$. Поэтому

$$V = C(v', 1) \oplus \bigoplus_{j=1}^{s} C(v_j, m_j),$$

и теорема доказана.

Предположим, что $\alpha_{1,0} = \cdots = \alpha_{r-1,0} = 0$, $\alpha_{r,0} \neq 0$, так что

$$\mathcal{B}v' = \sum_{j=r}^{s} \alpha_{j,0} v_j.$$

Разделив на ненулевой элемент поля $\alpha_{r,0}$, получим:

$$\mathcal{B}(\frac{1}{\alpha_{r,0}}v') = v_r + \sum_{j=r+1}^{s} \frac{\alpha_{j,0}}{\alpha_{r,0}} v_j.$$

Введем новые обозначения: $v'' = \frac{1}{\alpha_{r,0}}v'$, $\beta_j = \frac{\alpha_{j,0}}{\alpha_{r,0}}v_j$ при $j > r$. Легко убедиться, что $v'' \notin U$. Полученное выше равенство записывается в виде:

$$\mathcal{B}v'' = v_r + \sum_{j=r+1}^{s} \beta_j v_j.$$

Так как $\mathcal{B}^{m_j}v_j = 0$, и $m_r \geq m_{r+1} \geq \cdots \geq m_s$, то $\mathcal{B}^{m_r+1}v'' = 0$. Так как $\mathcal{B}^{m_r-1}v_r \neq 0$, то $\mathcal{B}^{m_r}v'' \neq 0$. В самом деле,

$$\mathcal{B}^{m_r}v'' = \mathcal{B}^{m_r-1}v_r + \sum_{j=r+1}^{s} \beta_j \mathcal{B}^{m_r-1}v_j, \tag{2.2.2}$$

$\mathcal{B}^{m_r-1}v_r \in C(v_r, m_r)$, $\beta_j\mathcal{B}^{m_r-1}v_j \in C(v_j, m_j)$, и сумма подпространств вида $C(v_k, m_k)$ прямая. Поэтому, если выражение (2.2.2) равно нулю, то нулю равны и все его слагаемые. Но одно из них, а именно $\mathcal{B}^{m_r-1}v_r$, по предположению не является нулем.

Таким образом, существует циклическое подпространство $C(v'', m_r+1)$. Покажем, что

$$V = (\overset{s}{\underset{j=1,j\neq r}{\oplus}} C(v_j, m_j)) \oplus C(v'', m_r + 1).$$

Положим $W = \sum\limits_{j+1,j\neq r}^{s} C(v_j, m_j) + C(v'', m_r + 1)$. Начнем с того, что покажем равенство $V = W$. Ясно, что $W \subseteq V$. Покажем, что $V \subseteq W$. Для этого достаточно убедиться, что в W содержатся все элементы какого-либо базиса V. В нашем распоряжении имеется базис V, состоящий из базисов циклических пространств $C(v_j, m_j)$, $1 \leq j \leq s$, и вектора v''. Ясно, что $v'' \in C(v'', m_r+1) \subseteq W$. Далее, $C(v_j, m_j) \subseteq W$ при $j \neq r$. Остается разобраться только с базисными элементами $C(v_r, m_r)$. Поскольку

W есть сумма инвариантных (относительно $\mathcal{B}$) подпространств, то достаточно только убедиться, что $v_r \in W$. Используем полученное выше равенство: $\mathcal{B}v'' = v_r + \sum\limits_{j=r+1}^{s} \beta_j v_j$. Из него следует

$$v_r = \mathcal{B}v'' - \sum_{j=r+1}^{s} \beta_j v_j.$$

Все слагаемые в правой части принадлежат W. В самом деле, так как подпространство W инвариантно, то из $v'' \in W$ следует $\mathcal{B}v'' \in W$. Наконец, при $j \neq r$ элементы v_j принадлежат W по самому определению W.

Итак, $V = W$. Отсюда следует, что V есть линейная оболочка базисных элементов всех $C(v_j, m_j)$ при $j \neq r$, и базисных элементов $C(v'', m_r+1)$. Общее количество этих элементов равно $\sum\limits_{j=1, j\neq r}^{s} m_j + m_r + 1$, то есть равно числу элементов базиса V. Если бы данное множество образующих пространства V было линейно зависимым, то из него можно было бы выбрать базис V с меньшим числом элементов, что противоречит независимости числа элементов базиса от выбора базиса. Таким образом, базисные элементы всех $C(v_j, m_j)$ при $j \neq r$ и базисные элементы $C(v'', m_r + 1)$ в совокупности линейно независимы, и образуют базис V. Из теоремы 3.1.5 (это известное свойство прямых сумм) теперь следует, что в определении W сумма подпространств на самом деле является прямой. Итак, пространство V и в этом случае представляется в виде прямой суммы циклических подпространств. □

Следствие 2.2.1. *Матрица любого нильпотентного оператора (над произвольным полем) приводится к жордановой нормальной форме.*

Доказательство. Это следует из предыдущей теоремы и из предшествующей ей леммы, причем каждому прямому слагаемому $C(v_j, m_j)$

из (2.2.1) сответствует жорданова клетка $J_{m_j}(0)$. Таким образом, разложению (2.2.1) соответствует жорданова нормальная форма матрицы оператора $\mathcal{B}$:

$$\begin{pmatrix} J_{m_1}(0) & 0 & \dots & 0 \\ 0 & J_{m_2}(0) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ddots & J_{m_s}(0) \end{pmatrix}$$

□

Теорема 2.2.2. *Матрица произвольного линейного оператора над полем комплексных чисел может быть приведена к жордановой нормальной форме.*

Доказательство теоремы — это фактически те рассуждения, которые показывают, как задача о нахождении жордановой нормальной формы произвольного оператора сводится к вопросу о существовании жордановой нормальной формы для нильпотентного оператора. □

Отметим, что если проанализировать всю цепочку рассуждений, составляющих доказательство этой теоремы (начиная с теоремы Гамильтона-Кэли), то будет ясно, что над полем действительных чисел матрицу можно привести к жордановой нормальной форме тогда и только тогда, если ее характеристический многочлен можно разложить над полем $\mathbb{R}$ на множители первой степени.

2.3. Единственность жордановой нормальной формы и способ ее вычисления

Итак, с точки зрения теории приведение матрицы линейного оператора $\mathcal{A}$ к жордановой нормальной форме состоит из следующих шагов. Сначала находятся все собственные значения $\lambda_1, \dots, \lambda_m$ операто-

ра $\mathcal{A}$, и разложение в прямую сумму инвариантных подпространств $V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_m)$. Затем рассматривается ограничение $\mathcal{A}$ на каждое из подпространств $V(\lambda_i)$, и задача решается для этого случая. Выберем произвольное λ_i, обозначим его через λ, и пусть $V' = \bigoplus\limits_{j,j\neq i} V(\lambda_j)$, так что $V = V(\lambda) \oplus V'$, где ограничение $\mathcal{A} - \lambda\mathcal{E}$ на $V(\lambda)$ нильпотентно (а значит, нильпотентны и ограничения на $V(\lambda)$ всех $(\mathcal{A} - \lambda\mathcal{E})^k$ при $k \geq 1$), а ограничение $\mathcal{A} - \lambda\mathcal{E}$ на V' невырожденно (а значит, невырожденны и ограничения на $V(\lambda)$ всех $(\mathcal{A} - \lambda\mathcal{E})^k$ при $k \geq 1$). Поэтому, согласно теореме 2.2.1,

$$V(\lambda) = \bigoplus_{j=1}^{s} C(v_j, m_j),$$

где $C(v_j, m_j) = \langle v_j, (\mathcal{A} - \lambda\mathcal{E})v_j, \ldots, (\mathcal{A} - \lambda\mathcal{E})^{m_j-1}v_j \rangle$, причем $(\mathcal{A} - \lambda\mathcal{E})^{m_j-1}v_j \neq 0$, $(\mathcal{A} - \lambda\mathcal{E})^{m_j}v_j = 0$. При этом существует взаимнооднозначное соответствие между подпространствами $C(v_j, m_j)$ и жордановыми клетками вида $J_{m_j}(\lambda)$ в жордановой нормальной форме матрицы A линейного оператора $\mathcal{A}$. Таким образом, можно выразить количество $N(\lambda, k)$ жордановых клеток вида $J_k(\lambda)$, содержащихся в жордановой нормальной форме матрицы A, в виде суммы $N(\lambda, k) = \sum\limits_{j,m_j=k} 1$.

Теорема 2.3.1. *Жорданова нормальная форма матрицы A линейного оператора $\mathcal{A}$ над полем комплексных чисел определена однозначно. Точнее, для каждого собственного значения λ матрицы A и для каждого натурального числа k однозначно определено количество $N(\lambda, k)$ жордановых клеток вида $J_k(\lambda)$, содержащихся в жордановой нормальной форме матрицы A. Это количество вычисляется по формуле:*

$$N(\lambda, k) = r_{k-1} - 2r_k + r_{k+1},$$

где $r_0 = n$ (порядок матрицы), и $r_k = \mathrm{rk}((A - \lambda E)^k)$ для $k \geq 1$.

Доказательство. Пусть λ — некоторое собственное значение матрицы A. Положим $V' = \bigoplus\limits_{\lambda' \neq \lambda} V(\lambda')$, где сумма берется по всем собственным значениям A, отличным от λ. Тогда

$$V = V(\lambda) \oplus V'.$$

Ограничение $\mathcal{A} - \lambda\mathcal{E}$ на $V(\lambda)$ нильпотентно, поэтому

$$V(\lambda) = \bigoplus_{j=1}^{s} C(v_j, m_j),$$

где циклические подпространства

$$C_j = C(v_j, m_j) = \langle v_j, (\mathcal{A} - \lambda\mathcal{E})v_j, \dots, (\mathcal{A} - \lambda\mathcal{E})^{m_j - 1} v_j \rangle$$

инвариантны относительно $\mathcal{A} - \lambda\mathcal{E}$, а потому и относительно $\mathcal{A}$. Так как выбор обозначений зависит от нас, то можно предполагать, что $m_1 \leq \dots \leq m_s$. Положим

$$r_k = \mathrm{rk}((A - \lambda E)^k) = \dim((\mathcal{A} - \lambda\mathcal{E})^k V).$$

Заметим, что $(\mathcal{A} - \lambda\mathcal{E})^k C_j = \{0\}$ при $k \geq m_j$. Это следует из того, что для всех базисных элементов $(\mathcal{A} - \lambda\mathcal{E})^i v_j$ пространства C_j после применения к ним оператора $(\mathcal{A} - \lambda\mathcal{E})^k$ получается нулевой вектор.

Теперь отметим два общих свойства векторных пространств и линейных операторв.

а) Пусть $W = W_1 \oplus \dots \oplus W_s$ — разложение в прямую сумму подпространств, каждое из которых инвариантно относительно оператора $\mathcal{B}$. Тогда

$$\mathcal{B}W = \mathcal{B}W_1 \oplus \dots \oplus \mathcal{B}W_s.$$

б) Если при этом $W_i = \langle u_1, \dots, u_l \rangle$, то $\mathcal{B}W_i = \langle \mathcal{B}u_1, \dots, \mathcal{B}u_l \rangle$.

Пункт а) доказывается так. Любой вектор $w \in W$ представляетя в виде $w = w_1 + \dots + w_s$, где $w_i \in W_i$ для каждого i. Применяя оператор

$\mathcal{B}$, получим $\mathcal{B}w = \mathcal{B}w_1 + \cdots + \mathcal{B}w_s$. Отсюда следует, что $\mathcal{B}W \subseteq \sum\limits_{i=1}^{s} \mathcal{B}W_i$. Обратно, если дан вектор вида $\mathcal{B}w_1 + \cdots + \mathcal{B}w_s$, то он равен вектору $\mathcal{B}w$, где $w = w_1 + \cdots + w_s$. Отсюда следует обратное включение. Ввиду инвариантности относительно $\mathcal{B}$ все слагаемые $\mathcal{B}w_i$ принадлежат подпространствам W_i. Так как сумма подпространств W_i прямая, то из равенства $\mathcal{B}w'_1 + \cdots + \mathcal{B}w'_s = \mathcal{B}w''_1 + \cdots + \mathcal{B}w''_s$ будет следовать, что $\mathcal{B}w'_i = \mathcal{B}w''_i$ для всех i. Но это означает, что сумма подпространств $\mathcal{B}W_i$ также будет прямой.

Пункт б) доказывается еще проще. Произвольный вектор $w \in W_i$ имеет вид $w = \alpha_1 u_1 + \cdots + \alpha_l u_l$. Следовательно, $\mathcal{B}w = \alpha_1 \mathcal{B}u_1 + \cdots + \alpha_l \mathcal{B}u_l \in \langle \mathcal{B}u_1, \ldots, \mathcal{B}u_l \rangle$. Это значит, что $\mathcal{B}W_i \subseteq \langle \mathcal{B}u_1, \ldots, \mathcal{B}u_l \rangle$. Обратно, если дан вектор $\alpha_1 \mathcal{B}u_1 + \cdots + \alpha_l \mathcal{B}u_l \in \langle \mathcal{B}u_1, \ldots, \mathcal{B}u_l \rangle$, то его можно представить в виде $\mathcal{B}(\alpha_1 u_1 + \cdots + \alpha_l u_l) \in \mathcal{B}W_i$.

Применим эти свойства к разложению в прямую сумму

$$V = \bigoplus_{j=1}^{s} C_j \oplus V'$$

где каждое прямое слагаемое инвариантно относительно оператора $(\mathcal{A} - \lambda\mathcal{E})^k$. Получим следующее:

$$(\mathcal{A} - \lambda\mathcal{E})^k V = \bigoplus_{j=1}^{s} (\mathcal{A} - \lambda\mathcal{E})^k C_j \oplus (\mathcal{A} - \lambda\mathcal{E})^k V'. \qquad (2.3.1)$$

Ограничение $(\mathcal{A} - \lambda\mathcal{E})^k$ на V' — невырожденный (т.е. биективный) оператор, и поэтому $(\mathcal{A} - \lambda\mathcal{E})^k V' = V'$. Как уже было замечено, $(\mathcal{A} - \lambda\mathcal{E})^k C_j = \{0\}$ при $k \geq m_j$. При $k < m_j$ (с учетом пункта б)) получим:

$$(\mathcal{A} - \lambda\mathcal{E})^k C_j = \langle (\mathcal{A} - \lambda\mathcal{E})^k v_j, (\mathcal{A} - \lambda\mathcal{E})^{k+1} v_j, (\mathcal{A} - \lambda\mathcal{E})^{m_j - 1} v_j \rangle.$$

Так как $(\mathcal{A} - \lambda\mathcal{E})^{m_j} v_j = 0$ по определению, то остальные базисные элементы C_j при действии на них $(\mathcal{A} - \lambda\mathcal{E})^k$ обращаются в нуль.

Итак, при $k \leq m_j$ имеется равенство

$$\dim((\mathcal{A} - \lambda\mathcal{E})^k C_j) = m_j - k.$$

Вычисляя размерности левой и правой частей (2.3.1), получим:

$$r_k = \sum_{m_j > k} (m_j - k) + \dim(V')$$

Заменяя k на $k+1$, будем иметь

$$r_{k+1} = \sum_{m_j > k+1} (m_j - k - 1) + \dim(V').$$

Теперь вычислим $r_k - r_{k+1}$.

$$\begin{aligned} r_k - r_{k+1} &= \sum_{m_j > k} (m_j - k) - \sum_{m_j > k+1} (m_j - k - 1) = \\ &= \sum_{m_j = k+1} (k + 1 - k) + \sum_{m_j > k+1} (m_j - k) - \sum_{m_j > k+1} (m_j - k) + \sum_{m_j > k+1} 1 = \\ &= \sum_{m_j = k+1} 1 + \sum_{m_j > k+1} 1 = \\ &= N(\lambda, k+1) + N(\lambda, k+2) + \dots . \end{aligned}$$

Итак,

$$\begin{aligned} r_k - r_{k+1} &= N(\lambda, k+1) + N(\lambda, k+2) + \dots , \\ r_{k-1} - r_k &= N(\lambda, k) + N(\lambda, k+1) + N(\lambda, k+2) + \dots . \end{aligned}$$

Вычитая из второго равенства первое, получаем искомое равенство:

$$N(\lambda, k) = r_{k-1} - 2r_k + r_{k+1},$$

справедливое для всех $k \geq 1$. Так как $r_k = \dim((\mathcal{A} - \lambda\mathcal{E})^k V)$, и нулевая степень любого оператора по определению есть тождественный оператор $\mathcal{E}$, то $r_0 = \dim(V) = n$.

Таким образом, для каждого собственного значения оператора $\mathcal{A}$ количество жордановых клеток данного порядка k, отвечающих собственному значению λ, определяется однозначно, и не зависит от способа

приведения матрицы оператора к жордановой нормальной форме. Следовательно, и сама жорданова нормальная форма определена однозначно с точностью до расположения жордановых клеток. □

2.4. Минимальный многочлен линейного оператора

Всюду в этом параграфе предполагается, что поле K — это либо поле компплексных чисел $\mathbb{C}$, либо поле действительных чисел $\mathbb{R}$. Впрочем, некоторые утверждения (и их доказательства) справедливы для любого поля.

Теорема 2.4.1. *Пусть A — квадратная матрица с элементами из поля K. Существуют ненулевые многочлены $f(x) \in K[x]$, такие, что $f(A) = 0$. Выберем среди всех таких многочленов ненулевой многочлен наименьшей положительной степени, и обозначим его через $h(x)$. Тогда выполняются следующие свойства:*

1) $h(A) = 0$;

2) $f(A) = 0$ *тогда и только тогда, если* $f(x) = h(x)g(x)$ *для некоторого многочлена* $g(x) \in K[x]$;

3) *Матрица A обратима тогда и только тогда, если свободный член многочлена $h(x)$ отличен от нуля;*

4) *Если $\lambda \in K$ — корень $h(x)$, то λ — собственное значение матрицы A. Обратно, если λ — собственное значение A, то λ есть корень $h(x)$.*

Доказательство. Случай, когда $A = 0$, тривиален, и в дальнейшем мы будем предполагать, что $A \neq 0$. Векторное пространство квадратных матриц фиксированного размера является конечномерным

(например, если это $n \times n$-матрицы, то $\dim(M_n(K)) = n^2$). Поэтому счетная последовательность $E, A, A^2, \ldots, A^k, \ldots$ не может состоять из линейно независимых элементов (напомним, что через E обозначается единичная матрица). Должна найтись нетривиальная линейная комбинация вида $a_0E + a_1A + a_2A^2 + \cdots + a_mA^m = 0$, где по крайней мере один коэффициент a_i отличен от нуля. Можно даже считать, что $a_m \neq 0$. Тогда многочлен $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m$ отличен от нуля, и $f(A) = a_0E + a_1A + a_2A^2 + \cdots + a_mA^m = 0$. Таким образом, ненулевые многочлены со свойством $f(A) = 0$ существуют для любой квадратной матрицы A. (Для нулевой матрицы можно взять многочлен $f(x) = x$.) Выберем среди всех таких (ненулевых) многочленов многочлен $h(x)$, имеющий наименьшую степень. Из самого способа выбора h следует, что $h(A) = 0$.

Если $f(x) = h(x)g(x)$, то $f(A) = h(A)g(A) = 0 \cdot g(A) = 0$. (Необходимо помнить, что символ "0" в этом равенстве означает нулевую матрицу!) С другой стороны, пусть $f(A) = 0$. Разделим многочлен $f(x)$ на многочлен $h(x)$ с остатком:

$$f(x) = h(x)g(x) + r(x), \qquad \deg(r(x)) < \deg(h(x)).$$

Подставляя матрицу A вместо x, и вспоминая, что $f(A) = 0$ и $h(A) = 0$, получаем отсюда, что $r(A) = 0$. Но если $r(x) \neq 0$, то неравенство $\deg(r(x)) < \deg(h(x))$ приводит к противоречию с выбором $h(x)$ как многочлена с *наименьшей* степенью среди всех тех многочленов, которые при подстановке $\mathcal{A}$ вместо x обращаются в нуль. Оказывается, что многочлен $r(x)$ ненулевой, $r(A) = 0$, но степень r строго меньше, чем степень h. Отсюда следует, что единственная непротиворечивая возможность — это равенство $r = 0$.

Пусть $h(x) = b_0 + b_1x + \cdots + b_mx^m$. Допустим, что $b_0 \neq 0$. Тогда, поскольку $h(x)$ ненулевой, по крайней мере один из коэффициентов b_i

при $i > 0$ также должен быть ненулевым. Если бы это было не так, то оказалось бы, что $h = b_0 \neq 0$, но тогда $h(A) = b_0 E \neq 0$. Рассмотрим равенство:

$$0 = h(A) = b_0 E + b_1 A + b_2 A^2 + \cdots + b_m A^m = b_0 E + A(b_1 E + b_2 A + \cdots + b_m A^{m-1}).$$

Так как $b_0 \neq 0$, то на этот элемент поля K можно разделить обе части предыдущего равенства. В результате получим:

$$0 = E + A(b_0^{-1} b_1 E + b_0^{-1} b_2 A + \cdots + b_0^{-1} b_m A^{m-1}).$$

Положим $u(x) = (-b_0^{-1} b_1) + (-b_0^{-1} b_2)x + \cdots + (-b_0^{-1} b_m)x^{m-1}$. Тогда из последнего полученного равенства следует, что

$$E = Au(A) = u(A)A.$$

Но это и означает, что у матрицы A есть обратная матрица, равная $u(A)$.

С другой стороны, если $b_0 = 0$, то из $h(A) = 0$ следует, что

$$0 = b_1 A + b_2 A^2 + \cdots + b_m A^m = A(b_1 E + b_2 A + \cdots + b_m A^{m-1}).$$

Так как многочлен $h(x)$ отличен от нуля, у него должен быть по крайней мере один ненулевой коэффициент. (Можно считать, что $b_m \neq 0$). Поэтому многочлен $w(x) = b_1 + b_2 x + \cdots + b_m x^{m-1}$ не равен нулю. Так как степень w строго меньше степени h, то $w(A) \neq 0$ (это следует из определения h). Таким образом, получено равенство:

$$0 = Aw(A),$$

где $A \neq 0$ и $w(A) \neq 0$. Теперь, если существует A^{-1}, то немедленно получается противоречие:

$$0 = A^{-1} \cdot 0 = A^{-1} A w(A) = w(A).$$

Пусть теперь λ — собственное значение A, $Av = \lambda v$ для некоторого столбца $v \neq 0$. Тогда $A^i v = \lambda^i v$ для всех $i \geq 0$, и

$$h(A)v = \sum_{i=0}^{m} b_i A^i v = \sum_{i=0}^{m} b_i \lambda^i v = h(\lambda)v.$$

Это равенство справедливо для любого многочлена $h(x)$, но если $h(A) = 0$, значит $h(A)v = 0$ (здесь справа от знака равенства стоит нулевой вектор!), откуда получаем $h(\lambda)v = 0$. Так как собственный вектор-столбец v отличен от нуля, то это возможно лишь в случае, когда $h(\lambda) = 0$.

Наконец, пусть $h(\lambda) = 0$ для некоторого $\lambda \in K$. Так как по теореме Гамильтона-Кэли $\chi_A(A) = 0$, то из уже доказанного пункта 2) следует, что $\chi_A(x) = h(x)g(x)$ для какого-то многочлена $g(x)$. Подставляя λ вместо x, получаем, что

$$\chi_A(\lambda) = h(\lambda)g(\lambda) = 0 \cdot g(\lambda) = 0.$$

Но, как известно, каждый корень характеристического многочлена матрицы является собственным значением этогй матрицы (теорема 1.4.1 из выпуска I). □

Определение 2.4.1. Многочлен $h(x)$, существование которого устанавливается предыдущей теоремой, называется *минимальным многочленом* матрицы A, и обозначается через $\mu_A(x)$.

Если имеются два многочлена, $h_1(x)$ и $h_2(x)$, удовлетворяющие свойствам, определяющим минимальный многочлен матрицы A (то есть $h_i(A) = 0$ и степени h_i минимальны, $i = 1, 2$), то, согласно пункту 2) теоремы 2.4.1, многочлен $h_1(x)$ делится на $h_2(x)$, а $h_2(x)$ делится на $h_1(x)$. Отсюда следует, что $h_1(x) = \alpha h_2(x)$, где α — ненулевой элемент поля K. Таким образом, минимальный многочлен линейного оператора определен с точностью до скалярного множителя. Но если потребовать,

чтобы старший коэффициент этого многочлена был равен единице, то такой мноочлен уже определяется полностью однозначно.

Наиболее интересные свойства минимального многочлена мы сможем доказать лишь для тех случаев, когда выполняется теорема Гамильтона-Кэли, которая утверждает, что $\chi_A(A) = 0$. Это влечет, в частности, что характеристический многочлен без остатка делится на минимальный многочлен. Теорема Гамильтона-Кэли верна для любой матрицы в случае, когда поле K есть поле комплексных чисел $\mathbb{C}$, или поле действительных чиселℝ. В случае произвольного поля она тоже верна, но мы пока не располагаем средствами, чтобы это доказать. В некоторых случаях (но далеко не во всех) минимальный многочлен совпадает с характеристическим многочленом оператора.

Лемма 2.4.1. *Минимальный многочлен матрицы A равен минимальному многочлену матрицы $B^{-1}AB$.*

Доказательство. Утверждение леммы непосредственно следует из тождества:

$$f(B^{-1}AB) = B^{-1}f(A)B \tag{2.4.1}$$

Чтобы доказать это тождество, надо сначала установить его в частном случае, при $f(x) = x^m$. Если $f(x) = x^m$, то

$$f(B^{-1}AB) = (B^{-1}AB)^m = B^{-1}ABB^{-1}ABB^{-1}AB\dots B^{-1}AB.$$

После того, как все рядом стоящие B и B^{-1} сократятся, правая часть окажется равной $B^{-1}A^mB$, то есть $B^{-1}f(A)B$. Если теперь $f(x) = \sum\limits_{i=0}^{k} a_ix^i$, то

$$f(B^{-1}AB) = \sum_{i=0}^{k} a_i(B^{-1}AB)^i = \sum_{i=0}^{k} a_i(B^{-1}A^iB) = B^{-1}(\sum_{i=0}^{k} a_iA^i)B = B^{-1}f(A)B.$$

Теперь очевидно, что если ненулевой многочлен $f(x)$, таков, что $f(A) = 0$, то и $f(B^{-1}AB) = 0$, и наоборот, если $f(B^{-1}AB) = 0$, то $f(A) = 0$. Это означает, что минимальные многочлены матриц A и $B^{-1}AB$ определяются как многочлены наименьшей степени в одном и том же множестве ненулевых многочленов, и можно считать, что это один и тот же (с точностью до скалярного множителя) многочлен. □

Эта лемма позволяет сделать вывод, что, как и в случае характеристического многочлена, можно определить *минимальный многочлен* линейного оператора $\mathcal{A}$ как минимальный многочлен матрицы этого оператора в любом базисе. Из леммы 2.4.1 следует, что при переходе к другому базису этот многочлен не изменится. Будем обозначать минимальный многочлен оператора $\mathcal{A}$ через $\mu_{\mathcal{A}}(x)$. Таким образом, если A есть матрица $\mathcal{A}$ в некотором (каком угодно) базисе, то $\mu_{\mathcal{A}}(x) = \mu_A(x)$.

Теорему 2.4.1 иможно было бы доказать (практически теми же рассуждениями) сразу для минимального многочлена линейного оператора (а не матрицы).

Лемма 2.4.2. *Пусть матрица оператора $\mathcal{A}$ в некотором базисе имеет блочно-диагональный вид:*

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_k \end{pmatrix}$$

Тогда $\mu_A(x) = \text{НОК}(\mu_{A_1}(x), \mu_{A_2}(x), \dots, \mu_{A_k}(x))$.

Доказательство. Начать можно с очевидного равенства:

$$A^n = \begin{pmatrix} A_1^n & 0 & \dots & 0 \\ 0 & A_2^n & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_k^n \end{pmatrix}$$

для каждого $n \geq 0$. Отсюда легко следует, что для любого многочлена $f(x)$ выполняется равенство:

$$f(A) = \begin{pmatrix} f(A_1) & 0 & \dots & 0 \\ 0 & f(A_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f(A_k) \end{pmatrix}$$

Если $f(x) = \text{НОК}(\mu_{A_1}(x), \dots, \mu_{A_k}(x))$, то $f(x) = \mu_{A_i}(x)g_i(x)$ для каждого i. И тогда для каждого i выполняется равенство $f(A_i) = \mu_{A_i}(A_i)g_i(A_i) = 0$. Поэтому $f(A) = 0$. Допустим, что для какого-то многочлена $g(x)$ выполнно равенство $g(A) = 0$. Отсюда следует, что $g(A_i) = 0$ для каждого i. А отсюда, ввиду пункта 2) теоремы 2.4.1, следует, что g делится на $\mu_{A_i}(x)$ для всех i. А значит, $g(x)$ делится на $f(x) = \text{НОК}(\mu_{A_1}(x), \dots, \mu_{A_k}(x))$. Следовательно, $f(x)$ удовлетворяет определению минимального многочлена матрицы A. □

Лемма 2.4.3. *Пусть $f(x)$ — многочлен с коэффициентами из поля K, $J_n(\lambda)$ — жорданова клетка порядка n. Тогда*

$$f(J_n(\lambda)) = \begin{pmatrix} f(\lambda) & f'(\lambda)/1! & f''(\lambda)/2! & f'''(\lambda)/3! & \dots & f^{(n-1)}(\lambda)/(n-1)! \\ 0 & f(\lambda) & f'(\lambda)/1! & f''(\lambda)/2! & \dots & f^{(n-2)}(\lambda)/(n-2)! \\ 0 & 0 & f(\lambda) & f'(\lambda)/1! & \dots & f^{(n-3)}(\lambda)/(n-3)! \\ 0 & 0 & 0 & f(\lambda) & \dots & f^{(n-4)}(\lambda)/(n-4)! \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & f(\lambda) \end{pmatrix}$$

Доказательство. Заметим, что i, j-й элемент матрицы в правой части равенства есть $f^{(j-i)}(\lambda)/(j-i)!$.

Основная часть работы будет состоять в доказательстве равенства для $f(x) = x^m$ при произвольном m. В этом случае $f^{(k)}(x)/k! = C_m^k x^{m-k}$ при $k \leq m$, и $f^{(k)}(x)/k! = 0$ при $k > m$. Напомним, что $C_m^k = \dfrac{m!}{k!(m-k)!}$.

В частности, так как $0! = 1$, то $C_m^0 = C_m^m = 1$. Удобно считать, что $C_m^k = 0$ при $k > m$. В дальнейшем нам понадобится тождество:

$$C_m^{k-1} + C_m^k = C_{m+1}^k \tag{2.4.2}$$

Это одно из важнейших свойств биномиальных коэффициентов. Если оно почему-то окажется неизвестным к тому моменту, когда материал этого параграфа будет изучаться, то его можно рассматривать как не слишком трудное упражнение.

Таким образом, то, что нам необходимо сейчас доказать — это равенство:

$$J_n(\lambda)^m = \begin{pmatrix} C_m^0\lambda^m & C_m^1\lambda^{m-1} & C_m^2\lambda^{m-2} & C_m^3\lambda^{m-3} & \dots & C_m^{n-1}\lambda^{m-n+1} \\ 0 & C_m^0\lambda^m & C_m^1\lambda^{m-1} & C_m^2\lambda^{m-2} & \dots & C_m^{n-2}\lambda^{m-n+2} \\ 0 & 0 & C_m^0\lambda^m & C_m^1\lambda^{m-1} & \dots & C_m^{n-3}\lambda^{m-n+3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & C_m^0\lambda^m \end{pmatrix}$$

Проведем индукцию по m. В случае $m = 0$ и $m = 1$ утверждение очевидно. Допустим, что оно верно для $J_n(\lambda)^m$. Рассмотрим

$$J_n(\lambda)^{m+1} = J_n(\lambda) \cdot J_n(\lambda)^m.$$

Вычисляя правую часть, и умножая матрицы, получим в i-й строке и j-м столбце при $i \le j$ произведение строки

$$(0, \dots, \lambda, 1, 0, \dots, 0),$$

в которой λ располагается на i-м месте, на столбец

$$(C_m^{j-1}\lambda^{m-j+1}, C_m^{j-2}\lambda^{m-j+2}, C_m^{j-3}\lambda^{m-j+3}, \dots, C_m^0\lambda^m, 0, \dots, 0)^{\mathrm{T}},$$

где $C_m^0\lambda^m = \lambda^m$ располагается на j-м месте, а на i-м и $i+1$-м местах находятся соответственно $C_m^{j-i}\lambda^{m-j+i}$ и $C_m^{j-i-1}\lambda^{m-j+i+1}$. В результате

получим:

$$\lambda \cdot C_m^{j-i}\lambda^{m-j+i} + 1 \cdot C_m^{j-i-1}\lambda^{m-j+i+1} = (C_m^{j-i} + C_m^{j-i-1})\lambda^{m-j+i+1} = C_{m+1}^{j-i}\lambda^{m-j+i+1}.$$

Последнее равенство вытекает из тождества (2.4.2). Но именно элемент $C_{m+1}^{j-i}\lambda^{m-j+i+1}$ должен находиться на i,j-месте в матрице $J_n(\lambda)^{m+1}$, сделанное предположение о виде матрицы $J_n(\lambda)^{m+1}$ справедливо. Таким образом, индуктивное рассуждение проведено успешно, и равенство для $J_n(\lambda)^m$ установлено.

Случай произвольного $f(x)$ легко следует из уже доказаного. Во-первых, если $f(x) = \sum\limits_{m\geq 0} a_m x^m$, то $f(J_n(\lambda)) = \sum\limits_{m\geq 0} a_m J_n(\lambda)^m$. Отсюда следует, что каждый i,j-й элемент матрицы $f(J_n(\lambda))$ есть линейная комбинация i,j-х элементов матриц $J_n(\lambda)^m$ с коэффициентами a_m, то есть это $\sum\limits_{m\geq 0} a_m(\lambda^m)^{(j-i)}/(j-i)!$. Здесь через $(\lambda^m)^{(j-i)}$ обозначен результат подстановки элемента λ вместо x в многочлен $(x^m)^{(j-i)}$. Но так как

$$\left(\sum_{n\geq 0} a_m x^m\right)^{(k)} = \sum_{n\geq 0} a_m (x^m)^{(k)},$$

то

$$\sum_{m\geq 0} a_m(\lambda^m)^{(j-i)}/(j-i)! = f^{(j-i)}(\lambda)/(j-i)!,$$

что и требовалось доказать. □

Теорема 2.4.2. *Минимальный многочлен жордановой клетки $J_n(\lambda)$ равен $(x-\lambda)^n$.*

Доказательство. Пусть $f(x) = (x-\lambda)^n$. Тогда λ будет корнем всех производных $f^{(k)}(x)$ при $0 \leq k \leq n-1$. Согласно предыдущей лемме 2.4.3, отсюда следует, что $f(J_n(\lambda)) = 0$. С другой стороны, из леммы 2.4.3 также следует, что если для какого-то многочлена $g(x)$ выполняется равенство $g(J_n(\lambda)) = 0$, то λ является корнем всех производных

$g^{(k)}(x)$ при $0 \leq k \leq n-1$ (так как все компоненты матрицы $g(J_n(\lambda))$ должны быть равными нулю). Отсюда, согласно известному критерию для кратности корня, следует, что λ есть корень $g(x)$ кратности n, то есть $g(x) = (x-\lambda)^n u(x)$. Таким образом, многочлен $(x-\lambda)^n$ удовлетворяет всем условиям, определяющим минимальный многочлен жордановой клетки. □

Предыдущие утверждения позволяют сформулировать следующий алгоритм вычисления минимального многочлена произвольной матрицы A, для которой можно найти жорданову нормальную форму.

Теорема 2.4.3. *Если известна жорданова нормальная форма матрицы A, то ее минимальный многочлен есть*

$$\mu_A(x) = (x-\lambda_1)^{s_1}(x-\lambda_2)^{s_2}\dots(x-\lambda_m)^{s_m},$$

где $\lambda_1, \dots, \lambda_m$ — собственные значения A, и для каждого j, $1 \leq j \leq m$, показатель степени s_j есть наивысший порядок жордановых клеток, входящих в жорданову нормальную форму матрицы A, и отвечающих собственному значению λ_j.

Доказательство. Приведение матрицы к жордановой нормальной форме равносильно выбору жорданова базиса, а базисные векторы можно упорядочивать разными способами. Отсюда следует, что жордановы клетки в жордановой нормальной форме матрицы можно располагать в любом порядке. Нам удобно выбрать такой способ упорядочения, когда клетки, отвечающие одному и тому же собственному занчению, следуют подряд одна за другой. Это значит, что жорданву нормальную форму можно представить в виде:

$$\begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_k \end{pmatrix} \tag{2.4.3}$$

где каждая матрица A_i есть блочно-диагональная матрица: состоящая из жордановых клеток вида $J_t(\lambda_i)$ с, возможно, различными t, но с одним и тем же значением λ_i. Значения же $\lambda_1, \dots, \lambda_k$ будем предполагать попарно различными. Таким образом, для каждого i

$$A_i = \begin{pmatrix} J_{t_1}(\lambda_i) & 0 & \dots & 0 \\ 0 & J_{t_2}(\lambda_i) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{t_p}(\lambda_i) \end{pmatrix}$$

где числа p, $t_1, \dots, t_p$ зависят от индекса i. Теперь можно использовать лемму 2.4.2 и теорему 2.4.2, чтобы найти минимальный многочлен этой матрицы. Он будет равен

$$\text{НОК}((x-\lambda_i)^{t_1}, \dots, (x-\lambda_i)^{t_p}),$$

а этот многочлен, очевидно, равен $(x-\lambda_i)^{s_i}$, где $s_i = \max(t_1, \dots, t_p)$. (Напомним еще раз, что числа $p, t_1, \dots, t_p$ зависят от i, и для другого индекса могут быть совершенно иными.)

Теперь снова применим лемму 2.4.2 для матрицы (2.4.3). Так как все собственные значения λ_i различны, то наименьшее общее кратное многочленов $(x-\lambda_i)^{s_i}$ при $i = 1, \dots, k$, равно их произведению. Это и есть то утверждение, котороетребовалось доказать. □

ГЛАВА III. ПРИЛОЖЕНИЯ. СПРАВОЧНАЯ ИНФОРМАЦИЯ

В этой главе собраны некоторые сведения, которые могут понадобиться для понимания материала предшествующих глав (а также материала последующих выпусков данного учебного пособия). В основном речь идет о материале, который должен быть известен по первому семестру. В некоторых случаях, однако, приводятся подробности, которые могли не встретиться в лекционном курсе первого семестра. Разумеется, эти сведения не используются в основном материале нашего курса, и приведены только для "полноты картины". Доказательства большей частью отсутствуют. Исключение сделано лишь для нескольких исключительно важных для нашго курса утверждений, касающихся линйной независимости и разложения в прямую сумму.

3.1. Векторные пространства

Предположим, что нам известно определение поля (оно напоминается в третьем параграфе этого Приложения). Пусть K — некоторое поле. *Векторным пространством* над полем K называется множество V, для элементов которого (*векторов*) определены операции сложения $v+u \in V$ и умножения слева на элементы поля (скаляры): если $\alpha \in K$, $v \in V$, то $\alpha v \in V$. Должен существовать также особый нулевой вектор $0 \in V$ (не путать с нулем — элементом поля). Требуется, чтобы выполнялись следующие условия (аксиомы векторного пространства):

1. $v_1 + v_2 = v_2 + v_1$ для любых векторов $v_1, v_2 \in V$;
2. $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$ для любых векторов $v_1, v_2, v_3 \in V$;
3. $v + 0 = 0 + v = v$ для любго вектора $v \in V$;

4. $v + (-1)v = 0$ для любого вектора $v \in V$. Здесь $-1 \in K$, и элемент $(-1)v$ обычно обозначается через $-v$. В частности, $v - u = v + (-1)u$.

5. $1v = v$ для любого вектора $v \in V$. Здесь $1 \in K$.

6. $\alpha(v_1 + v_2) = \alpha v_1 + \alpha v_2$ для каждого $\alpha \in K$ и любых векторов $v_1, v_2 \in V$.

7. $(\alpha\beta)v = \alpha(\beta v)$ для любых $\alpha, \beta \in K$ и каждого $v \in V$. Отсюда, в частности, следует, что $(-\alpha)v = -(\alpha v)$.

8. $(\alpha + \beta)v = \alpha v + \beta v$ для любых $\alpha, \beta \in K$ и каждого $v \in V$. Отсюда, в частности, следует, что $0v = 0$, причем тот нуль, который стоит слева от знака равенства, есть элемент поля K, а тот, который стоит справа есть нулевой вектор.

Выражение $\sum\limits_{i=1}^{k} \lambda_i u_i$ называется *линейной комбинацией* векторов $v_1, \ldots, v_k$. Такая линейная комбинация называется *тривиальной*, если все коэффициенты λ_i равны нулю, и *нетривиальной*, если есть хотя бы один ненулевой коэффициент. Будем говорить, что вектор v *линейно зависит* от векторов $u_1, \ldots, u_k$, если найдутся такие $\alpha_1, \ldots, \alpha_k$ из поля K, что $v = \sum\limits_{i=1}^{k} \alpha_i u_i$. Если X — не обязательно конечное подмножество V, то будем говорить, что v *линейно зависит* от X, если v линейно зависит от некоторого конечного подмножества множества X.

Пусть V — векторное пространство над полем K. Векторное *подпространство* W пространства V — это подмножество $W \subseteq V$, обладающее следующим свойством: $\alpha_1 w_1 + \alpha_2 w_2 \in W$ для любых $w_1, w_2 \in W$ и произвольных $\alpha_1, \alpha_2 \in K$. Из этого определения следует, что множество W само является векторным пространством над полем K с теми же, что и в V, операциями сложения и умножения на элементы поля.

Пусть X — подмножество векторного пространства V. *Линейной оболочкой* X, или *подпространством, порожденным* множеством X (обозначение — $\langle X \rangle$) называется множество всех линейных комбинаций элементов X с коэффициентами из K:

$$\langle X \rangle = \{ \sum_{x \in X} r_x x \mid r_x \in K, \text{ почти все } r_x = 0 \}$$

В случае, если X конечно, например, если $X = \{u_1, \ldots, u_k\}$, то

$$\langle u_1, \ldots, u_k \rangle = \{\sum_{i=1}^{k} \lambda_i u_i | \lambda_i \in K, 1 \leq i \leq k\}$$

Говорят, что множество X, $X \subseteq V$, *порождает* векторное пространство V (или что X есть *множество образующих* для V), если $V = \langle X \rangle$. В случае произвольного X линейная оболочка $\langle X \rangle$ является векторным подпространством пространства V. Множество $W \subseteq V$ является подпространством V тогда и только тогда, если $W = \langle W \rangle$. Операция взятия линейной оболочки обладает следующими свойствами:

1. $X \subseteq \langle X \rangle$;
2. $X \subseteq Y \Longrightarrow \langle X \rangle \subseteq \langle Y \rangle$;
3. $\langle \langle X \rangle \rangle = \langle X \rangle$;
4. Если $v \in \langle X \rangle$, то найдется конечное подмножество $\{u_1, \ldots, u_k\} \subseteq X$ такое, что $v \in \langle u_1, \ldots, u_k \rangle$;
5. Если $v \in \langle X \cup \{u\} \rangle$, и $v \notin \langle X \rangle$, то $u \in \langle X \cup \{v\} \rangle$;
6. Если $v_1, \ldots, v_m$ — векторы пространства V, и $\alpha_1, \ldots, \alpha_m$ — ненулевые элементы поля (скаляры), то $\langle v_1, \ldots, v_m \rangle = \langle \alpha_1 v_1, \ldots, \alpha_m v_m \rangle$.

Отметим еще, что $\langle \varnothing \rangle = \{0\}$ (это можно даже принять за определение линейной оболочки пустого множества), $\langle 0 \rangle = \{0\}$, $\langle v \rangle = \{\alpha v | \alpha \in K\}$.

Конечномерные векторные пространства можно определить условием: $V = \langle X \rangle$ для некоторого конечного множества X. В этом случае конечномерными будут и все подпространства V.

Множество векторов $X \subseteq V$ называется *линейно зависимым*, если в нем можно выбрать непустое конечное подмножество векторов $u_1, \ldots, u_k$, для которого существует нетривиальная линейная комбинация, равная нулю, т.е.

$$\sum_{i=1}^{k} \alpha_i u_i = 0,$$

и есть такие коэффициенты α_i, которые не равны нулю. Из этого определения следует, что каждое множество X, которое содержит нулевой вектор, является линейно зависимым, так как $1 \cdot 0 = 0$ есть нетривиальная линейная комбинация, равная нулю. Из определения также следует, что если X линейно зависимо, и $X \subseteq Y$, то и Y будет линейно зависимым.

Отметим еще, что если $v \in \langle X \rangle$, но $v \notin X$, то множество $X \cup \{v\}$ линейно зависимо. Линейная оболочка $\langle X \rangle$ состоит из X, и из всех векторов, которые линейно зависят от всевозможных конечных подмножеств множества X (если X конечно, то только от самого X). Если множество X линейно зависимо, то найдется такой вектор $v \in X$, что $v \in \langle X \setminus \{v\} \rangle$.

Множество $X \subset V$ называется *линейно независимым*, если оно не является линейно зависимым. Это равносильно тому, что для каждого конечного подмножества $\{u_1, \ldots, u_k\} \subseteq X$ из равенства нулю некоторой линейной комбинации

$$\sum_{i=1}^{k} \alpha_i u_i = 0$$

непременно следует, что все коэффициенты α_i должны быть равны нулю. Если множество X само конечно, например, $X = \{u_1, \ldots, u_k\}$, то для проверки его линейной независимости достаточно брать только од-

но его подмножество — само X.

Лемма 3.1.1. *Эквивалентны следующие условия:*

1) *Множество векторов $u_1, \ldots, u_k$ линейно независимо;*

2) *Для каждого вектора $v \in \langle u_1, \ldots, u_k \rangle$, если $v = \sum\limits_{i=1}^{k} \alpha_i u_i$, то коэффициенты α_i определяются однозначно. Иными словами, если есть другая запись, $v = \sum\limits_{i=1}^{k} \beta_i u_i$, то $\alpha_i = \beta_i$ для всех i.*

Доказательство. Будем рассуждать от противного.

1) $\Rightarrow$ 2). Допустим, что векторы $u_1, \ldots, u_k$ линейно независимы, но для какого-то вектора v существуют два способа записи: $v = \sum\limits_{i=1}^{k} \alpha_i u_i$, $v = \sum\limits_{i=1}^{k} \beta_i u_i$, причем $\alpha_j \neq \beta_j$ для некоторого j. Вычтем из первого равенства второе, и получим:

$$\sum_{i=1}^{k} \alpha_i u_i - \sum_{i=1}^{k} \beta_i u_i = \sum_{i=1}^{k} (\alpha_i - \beta_i) u_i = v - v = 0.$$

Но если $\alpha_j \neq \beta_j$, то $\alpha_j - \beta_j \neq 0$, а это означает, что мы получили нетривиальную линейную комбинацию векторов $u_1, \ldots, u_k$, равную нулю, что противоречит линейной независимости этих векторов.

2) $\Rightarrow$ 1). Допустим, что $u_1, \ldots, u_k$ линейно зависимы. Это значит, что существует нетривиальная линйная комбинация этих векторов, равная нулю: $\sum\limits_{i=1}^{k} \alpha_i u_i = 0$. Нетривиальность означает, что $\alpha_j \neq 0$ для какого-то j, $1 \leq j \leq k$. Однако мы всегда можем представить нулевой вектор в виде линейной комбинации векторов $u_1, \ldots, u_k$ с нулевыми коэффициентами: $0 \cdot u_1 + \cdots + 0 \cdot u_k = 0$. Но тогда, согласно условию 2), примененного к вектору $v = 0$, должны выполняться равенства $\alpha_i = 0$ для всех i, в том числе и для $i = j$. Получено противоречие. □

Аналогичное утверждение справедливо также для бесконечных линейно независимых множеств. Доказательство остается практически

тем же самым, так как в любых двух записях вектора v в виде линейных комбинаций векторов из данного бесконечного множества используется только конечное число векторов этого множества. Если бесконечное множество уже было линейно независимым, то любое его конечное подмножество линейно независимо. При рассуждении в другую сторону фактически доказывается, что любое конечное подмножество данного бесконечного множества линейно независимо, а это эквивалентно определению линейной независимости всего множества.

Сформулируем некоторые полезные свойства линейно независимых множеств.

1. Пустое множество линейно независимо;
2. Каждое подмножество линейно независимого множества линейно независимо;
3. Множество X линейно независимо тогда и только тогда, если $v \notin \langle X \setminus \{v\} \rangle$ для каждого $v \in X$;
4. Пусть $\{v_1, \ldots, v_m\}$ и $\{u_1, \ldots, u_k\}$ — два линейно независимых подмножества векторного пространства V. Если $k > m$, то найдется вектор u_j, не равный никакому v_i и такой, что множество $\{v_1, \ldots, v_m, u_j\}$ линейно независимо;
5. Если $v \notin \langle X \rangle$, и множество X линейно независимо, то $X \cup \{v\}$ линейно независимо;
6. Если $v_1, \ldots, v_m$ — линейно независимые векторы, и $\alpha_1, \ldots, \alpha_m$ — ненулевые элементы поля (скаляры), то $\alpha_1 v_1, \ldots, \alpha_m v_m$ — также линейно независимые векторы.

Базисом векторного пространства V называется такое его подмножество X, что X, во-первых, порождает V (является для V множеством

образующих), а во-вторых, линейно независимо.

На языке линейных оболочек это можно выразить так: $V = \langle X \rangle$, и $v \notin \langle X \setminus \{v\} \rangle$ для каждого $v \in X$.

В случае конечного базиса $X = \{u_1, \ldots, u_n\}$ (если такой существует), определение базиса означает, что, во-первых, любой вектор $v \in V$ можно представить в виде линейной комбинации

$$v = \sum_{i=1}^{n} \alpha_i u_i \tag{3.1.1}$$

а во-вторых, векторы $u_1, \ldots, u_n$ линейно независимы. По лемме 3.1.1 это равносильно тому, что каждый вектор v записывается в виде (3.1.1) единственным способом.

Таким образом, по вектору v однозначно определяется упорядоченный набор элементов поля $(\alpha_1, \ldots, \alpha_n)$. Элементы α_i называются *координатами* вектора v в базисе (или относительно базиса) $u_1, \ldots, u_n$. С другой стороны, если взять какой угодно упорядоченный набор элементов поля $(\alpha_1, \ldots, \alpha_n)$, то по нему, используя имеющийся базис, можно построить вектор v по формуле (3.1.1). Все это означает, что справедливо следующий важный факт:

Теорема 3.1.1. *По заданном (конечному) базису $u_1, \ldots, u_n$ векторного пространства V можно построить взаимно-однозначное соответствие между V и множеством K^n упорядоченных наборов из n элеменотов поля K.*

Базис произвольного (не обязательно конечномерного) пространства V можно охарактеризовать еще следующим образом:

Теорема 3.1.2. *Эквивалентны следующие утверждения:*

1) *Множество $X \subset V$ является базисом векторного пространства V;*

2) *Множество $X \subset V$ является линейно независимым, но если добавить к нему хотя бы еще один элемент, то оно перестанет быть линейно зависимым (это еще выражают так: X — максимальное линейно независимое подмножество V);*

3) *Множество $X \subset V$ порождает пространство V (т.е. $V = \langle X \rangle$), но если удалить из него хотя бы один любой элемент, то оно перестанет обладать этим свойством (более короткая формулировка: X — минимальное порождающее подмножество V);*

Отметим также одно простое свойство, которое, однако, часто бывает необходимо использовать:ъ

Лемма 3.1.2. *Если $v_1, \ldots, v_m$ — базис пространства V, и $\alpha_1, \ldots, \alpha_m$ — ненулевые элементы поля (скаляры), то $\alpha_1 v_1, \ldots, \alpha_m v_m$ — также базис V.*

Следующая теорема справедлива для произвольных векторных пространств (не обязательно конечномерных) над любыми полями, так что множества X, Y, Z из ее формулировки могут иметь любую мощность.

Теорема 3.1.3. *Пусть V есть векторное пространство над полем K.*

(1) *Если в V существует базис X, то любой другой базис Y имеет ту же мощность, что и X. В частности, если существует конечный базис из n элементов, то все базисы содержат ровно n элементов.*

(2) *Если дано линейно независимое подмножество $X \subset V$ (возможно, даже пустое), и любое порождающее V множество Y (т.е. $\langle Y \rangle = V$, так что не исключен и случай $Y = V$), то существует*

подмножество $Z \subseteq Y$, такое, что $X \cap Z = \varnothing$, и $X \cup Z$ есть базис V.

(3) *В частности, полагая X пустым, получим, что из любого множества образующих V всегда можно выбрать базис (т.е. для любого $Y \subseteq V$ такого, что $\langle Y \rangle = V$, найдется базис $Z \subset V$. В частности, любое векторное пространство над полем обладает базисом.*

(4) *Другой частный случай пункта* (2)*: любое линейно независимое подмножество содержится в некотором базисе.*

Согласно первому пункту данной теоремы, мощность базиса не зависит от того, какой базис выбран. В случае, если один из базисов пространства V конечен, конечен и любой другой базис, и в нем содержится то же самое количество элементов. Это количество обозначается через $\dim(V)$ и называется *размерностью* векторного пространства V. Если базис бесконечнен, то его мощность не является числом в обычном смысле этого слова, и мы не будем подробно рассматривать эту ситуацию. Ограничимся пока лишь тем, что будем говорить в случае необходимости, что пространство бесконечномерно.

Сформулируем простейшие свойства размерности.

1. $\dim(V) = 0 \Leftrightarrow V = \{0\}$;
2. $v \neq 0 \Longrightarrow \dim(\langle v \rangle) = 1$;
3. $U \subseteq V \Longrightarrow \dim(U) \leq \dim(V)$;
4. Если $U \subseteq V$, и $\dim(U) = \dim(V)$, то $U = V$.

Еще раз отметим, что любое линейно независимое подмножество всегда содержится в некотором базисе векторного пространства. По-

этому мощность любого линейно независимого подмножества не превосходит размерности пространства. Если же мощность линейно независимого подмножества равна размерности всего пространства, то это подмножество само является базисом пространства.

Важнейшим примером n-мерного векторного пространста является векторное пространство K^n, элементами которого (векторами) являются столбцы высоты n, компонентами которых являются всевозможные элементы поля K. Операции сложения векторов и умножения вектора на скаляр (элемент поля) на K^n — это частные случаи сложения матриц и умножения матрицы на скаляр. Будем предполагать известными все свойства операций с матрицами. Пример базиса в K^n:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \; \ldots, \; e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Этот базис частно называют *стандартным базисом* K^n, или *базисом из единичных векторов*. Если

$$v = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

есть произвольный элемент K^n, то

$$v = \alpha_1 e_1 + \cdots + \alpha_n e_n.$$

Легко проверяется, что векторы $e_1, \ldots, e_n$ линейно независимы. Таким образом, $\dim(K^n) = n$.

Пусть U, W — два подпространства векторного пространства V. Тогда $U \cap W$ — также подпространство. Рассмотрим множество $U + W =$

$\{u+w|u \in U, w \in W\}$. Легко убедиться, что и это множество будет подпространством V. Оно называется *суммой* подпространств U и W. Все это можно обобщить на случай любого конечного количества слагаемых (можно и на бесконечный случай, но мы его не будем рассматривать). Пусть $U_1, \ldots, U_m$ — подпространства векторного пространства V. Назовем *суммой* этих подпространств множество

$$\{u_1 + \cdots + u_m | u_1 \in U_1, \ldots, u_m \in U_m\} \tag{3.1.2}$$

Обозначим это множество через $U_1 + \cdots + U_m$ или через $\sum\limits_{i=1}^{m} U_i$.

Лемма 3.1.3. *Множества $U_1 \cap \cdots \cap U_m$ и $U_1 + \cdots + U_m$ являются подпространствами векторного пространства V.*

Для каждого j имеет место включение

$$U_j \subseteq \sum_{i=1}^{m} U_i.$$

Это следует из того, что каждый вектор $u \in U_j$ можно формально представить в виде суммы m слагаемых: $u = 0 + \cdots + 0 + u + 0 + \cdots + 0$, в которой сам u располагается на j-м месте, а на каждом другом i-м месте $(i \neq j)$ расположен нулевой вектор, принадлежащий подпространству U_i.

Лемма 3.1.4. *Пусть X_1, $\ldots$, X_k — подмножества произвольного векторного пространства V. Тогда*

$$\langle X_1 \rangle + \cdots + \langle X_k \rangle = \langle X_1 \cup \cdots \cup X_k \rangle \tag{3.1.3}$$

Теорема 3.1.4. *Пусть даны два конечномерных подпространства U_1 и U_2 векторного пространства V. Тогда*

$$\dim(U_1) + \dim(U_2) = \dim(U_1 + U_2) + \dim(U_1 \cap U_2) \tag{3.1.4}$$

Пусть $U_1, \ldots, U_m$ — подпространства некоторого векторного пространства V.

Будем говорить, что сумма $\sum\limits_{i=1}^{m} U_i$ подпространств пространства V является *прямой суммой*, если для любого вектора u, представимого в виде $u = u_1 + \cdots + u_m$, где $u_1 \in U_1, \ldots, u_m \in U_m$, это представление определено однозначно. Иными словами, если $u = u_1 + \cdots + u_m$ и $u = u'_1 + \cdots + u'_m$, где $u_i, u'_i \in U_i$ для всех i, то $u_1 = u'_1$, $\ldots$, $u_m = u'_m$.

Лемма 3.1.5. *Это условие равносильно тому, что* $0 = u_1 + \cdots + u_m$ *тогда и только тогда, если все* $u_i = 0$.

Смысл понятия прямой суммы проясняет следующая лемма:

Лемма 3.1.6. *Пусть* $U_1, \ldots, U_m$ *— подпространства векторного пространства* V. *Сумма подпространств* $U_1 + \cdots + U_m$ *будет прямой тогда и только тогда, если для каждого набора индексов* $1 \leq i_1 < \cdots < i_k \leq m$ *и любых ненулевых элементов* $u_1 \in U_{i_1}, \ldots, u_k \in U_{i_k}$ *множество векторов* $u_1, \ldots, u_k$ *будет линейно независимым.*

Подчеркнем, что выбираются только ненулевые векторы, иначе нельзя говорить о линейной независимости. Таким образом, понятие прямой суммы выражает интуитивное представление о “линейно независимых” подпространствах.

Прямая сумма обозначается следующим образом: $U_1 \oplus \cdots \oplus U_m$ или $\bigoplus\limits_{i=1}^{m} U_i$.

Пусть $W_1, \ldots, W_m$ — подпространства некоторого векторного пространства W. Тогда определено подпространство $W_1 + \cdots + W_m = \sum\limits_{i=1}^{m} W_i$, состоящее из всевозможных сумм $w_1 + \cdots + w_m$, в которых $w_i \in W_i$ для каждого i. Это подпространство называется *суммой* подпространств $W_1, \ldots, W_m$. Заметим, что для каждого j имеет место включение $W_j \subseteq$

$\sum\limits_{i=1}^{m} W_i$. Это следует из того, что каждый вектор $w \in W_j$ можно формально представить в виде суммы m слагаемых: $w = 0+\cdots+0+w+0+\cdots+0$, в которой сам w располагается на j-м месте, а на каждом другом i-м месте ($i \neq j$) расположен нулевой вектор, который, разумеется, принадлежит подпространству W_i.

Будем говорить, что сумма $\sum\limits_{i=1}^{m} W_i$ подпространств является *прямой суммой*, если для любого вектора w, представимого в виде $w = w_1 + \cdots + w_m$, где $w_1 \in W_1, \ldots, w_m \in W_m$, это представление определено однозначно. Иными словами, если $w = w'_1 + \cdots + w'_m$ и $w = w''_1 + \cdots + w''_m$, где $w'_i, w''_i \in W_i$ для всех i, то $w'_1 = w''_1$, $\ldots$, $w'_m = w''_m$.

Это условие равносильно тому, что $0 = w_1 + \cdots + w_m$ тогда и только тогда, если все $w_i = 0$. В самом деле, если для каждого $w \in W_1 + \cdots + W_m$ представление его в виде $w = w_1 + \cdots + w_m$ является единственным, то оно является единственным и для нулевого вектора. Но для нулевого вектора одно такое представление мы знаем: $0 = 0 + \cdots + 0$, где стоящий в сумме на j-м месте нуль понимается как нулевой вектор подпространства W_j (на самом деле, конечно, все это один и тот же нулевой вектор пространства V). Тогда, если имеется еще какое-либо выражение $0 = w_1 + \cdots + w_m$, где $w_j \in W_j$ для каждого j, то из единственности следует, что для каждого j имеется равенство $w_j = 0$.

Обратно, если запись $0 = w_1 + \cdots + w_m$ возможна только когда все $w_j = 0$, то из $w = w'_1 + \cdots + w'_m$ и $w = w''_1 + \cdots + w''_m$ следует $0 = w - w = (w'_1 - w''_1) + \cdots + (w'_m - w''_m)$, и так как $w'_i - w''_i \in W_i$ для всех i, то $w'_i - w''_i = 0$, то есть $w'_i = w''_i$, и это значит, что любой вектор из $\sum\limits_{i=1}^{m} W_i$ можно представить в виде $w_1 + \cdots + w_m$ одним единственным способом.

Прямая сумма обозначается следующим образом: $W_1 \oplus \cdots \oplus W_m$ или $\bigoplus\limits_{i=1}^{m} W_i$.

Для работы с прямыми суммами очень полезно следующее утверж-

дение.

Теорема 3.1.5. *Пусть даны подпространства $W_1, \dots, W_m$ некоторого векторного пространства, и для каждого i, $1 \le i \le m$, дан базис $\{w_{i,1}, \dots, w_{i,k_i}\}$ подпространства W_i. Тогда, если сумма $W = W_1 + \dots + W_m$ является прямой суммой, то множество $\{w_{i,j} | 1 \le i \le m, 1 \le j \le k_i\}$ является базисом пространства W.*

Обратно, пусть дан базис некоторого векторного пространства W, представленный в виде $\{w_{i,j} | 1 \le i \le m, 1 \le j \le k_i\}$. Для каждого i, $1 \le i \le m$, определим подпространство $W_i = \langle w_{i,1}, \dots, w_{i,k_i} \rangle$. Тогда $W = W_1 \oplus \dots \oplus W_m$.

Доказательство. Проверим свойства базиса для $\{w_{i,j} | 1 \le i \le m, 1 \le j \le k_i\}$. Каждый вектор $w \in W_1 \oplus \dots \oplus W_m$ можно представить в виде $w = w_1 + \dots + w_m$, где $w_i \in W_i$ для всех i. Так как даны базисы пространств W_i, то каждый такой w_i можно записать в виде $w_i = \sum\limits_{j=1}^{k_i} \alpha_{i,j} w_{i,j}$, где $\alpha_{i,j} \in K$. Тогда для w получаем запись:

$$w = \sum_{i=1}^{m} \sum_{j=1}^{k_i} \alpha_{i,j} w_{i,j},$$

и это означает, что выполнено первое свойство базиса: каждый вектор можно представить в виде линейной кобинации его элементов. Теперь покажем линейную независимость элементов предполагаемого базиса. Пусть

$$\sum_{i=1}^{m} \sum_{j=1}^{k_i} \alpha_{i,j} w_{i,j} = 0.$$

Вспомним, что векторы $\sum\limits_{j=1}^{k_i} \alpha_{i,j} w_{i,j}$ принадлежат пространствам W_i (как линейные комбинации базисных элементов этих подпространств). Тогда по определению прямой суммы подпространств каждая такой вектор

должен быть равен нулю:

$$\sum_{j=1}^{k_i} \alpha_{i,j} w_{i,j} = 0.$$

Но по условию, входящие в эту линейную комбинацию векторы $w_{i,j}$ линейно независимы (ибо это базис в W_i). Отсюда следует, что $\alpha_{i,j} = 0$ для всех возможных i и j.

Обратно, пусть $\{w_{i,j} | 1 \leq i \leq m, 1 \leq j \leq k_i\}$ — базис в W, и W_i для каждого i есть линейная оболочка множества векторов $\{w_{i,j} | 1 \leq j \leq k_i\}$. Заметим, что так как это множество линейно независимо (как подмножество базиса), то оно будет базисом W_i. Каждый вектор $w \in W$ можно выразить как линейную комбинацию элементов базиса:

$$w = \sum_{i=1}^{m} \sum_{j=1}^{k_i} \alpha_{i,j} w_{i,j}.$$

Так как суммы $w_i = \sum\limits_{j=1}^{k_i} \alpha_{i,j} w_{i,j}$ принадлежат подпространствам W_i по самому определению этих подпространств, то мы получаем запись:

$$w = w + \cdots + w_m,$$

откуда следует, что $W = W_1 + \cdots + W_m$. Осталось убедиться, что сумма прямая. Пусть $w_1 + \cdots + w_m = 0$. Записывая w_i в виде $w_i = \sum\limits_{j=1}^{k_i} \alpha_{i,j} w_{i,j}$, преобразуем сумму $w + \cdots + w_m$ в выражение $\sum\limits_{i=1}^{m} \sum\limits_{j=1}^{k_i} \alpha_{i,j} w_{i,j}$. Эта линейная комбинация по предположению равна нулю, но так как в ней участвуют (без повторений) векторы базиса W, то все коэффициенты $\alpha_{i,j}$ должны быть нулями. Но тогда и $w_i = 0$ для каждого i. □

Следствие 3.1.1.

$$\dim(U_1 \oplus \cdots \oplus U_m) = \dim(U_1) + \cdots + \dim(U_m).$$

Обратно, если

$$\dim(U_1 + \cdots + U_m) = \dim(U_1) + \cdots + \dim(U_m),$$

то сумма подпространств $U_1 + \cdots + U_m$ является прямой суммой.

Следствие 3.1.2. *Векторы $v_1, \ldots, v_n$ образуют базис векторного пространства V тогда и только тогда, если*

$$V = \langle v_1 \rangle \oplus \cdots \oplus \langle v_n \rangle.$$

Теорема 3.1.6. *Пусть $U_1, \ldots U_m$ — подпространства векторного пространства V. Тогда равносильны следующие свойства:*

1) $V = U_1 \oplus \cdots \oplus U_m$;

2) $V = U_1 + \cdots + U_m$ *и* $U_j \cap (U_1 + \cdots + U_{j-1} + U_{j+1} + \cdots + U_m) = \varnothing$ *для каждого индекса* j, $1 \leq j \leq m$.

В случае двух слагаемых условие 2) *означает, что* $V = U_1 + U_2$, *и* $U_1 \cap U_2 = \varnothing$.

Пусть X — некоторое подмножество конечномерного векторного пространства V. *Рангом* множества векторов X называется число $\dim(\langle X \rangle)$. Выше уже было отмечено, что из множества образующих X векторного пространства $\langle X \rangle$ всегда можно выбрать базис этого пространства. Вспоминая свойства базиса, выводим отсюда следующее утверждение.

Лемма 3.1.7. *Ранг множества векторов X равен мощности (любого) максимального линейно независимого подмножества, содержащегося в X.*

Рассмотрим множество $M_{n,m}(K)$ матриц над полем K с n строками и m столбцами ($n \times m$-матриц). Строки матрицы $A \in M_{n,m}(K)$ можно

рассматривать как векторы из m-мерного пространства над полем K. Ранг множества строк матрицы A временно назовем рангом A по строкам. По предыдущей лемме ранг по строкам равен количеству строк в максимальном линейно независимом подмножестве строк матрицы A. Аналогично можно определить ранг по столбцам матрицы A как ранг множества векторов-столбцов этой матрицы (векторов пространства K^n). Он равен максимальному количеству линейно независимых столбцов A. Таким образом, ранги по строкам и по столбцам определяются как размерности двух различных подпространств в двух разных векторных пространствах. Тем не менее справедлива следующая важная теорема:

Теорема 3.1.7. *Ранг матрицы по строкам равен ее рангу по столбцам.*

Число, равное рангу по строкам и рангу по столбцам матрицы A, разывается *рангом* этой матрицы, и обозначается через $\mathrm{rk}(A)$. Встречается также обозначение $\mathrm{rank}(A)$.

Механизм возможного доказательства предыдущей теоремы объясняетсz следующим фактом.

Теорема 3.1.8. *Пусть $A_1, \ldots, A_k$ — векторы-столбцы из пространства K^n. Составим из них блочную $n \times k$-матрицу $A = (A_1, \ldots, A_k)$. Векторы $A_1, \ldots, A_k$ будут линейно независимыми в том и только в том случае, если в матрице A найдется минор порядка k, не равный нулю. Это также равносильно тому, что ранг матрицы A равен k.*

Аналогичное утверждение справедливо для линейно независимых векторов-строк.

Следствие 3.1.3. *Ранг матрицы A равен наибольшему среди порядков отличных от нуля миноров этой матрицы.*

Теорема 3.1.9. *Ранг матрицы не изменится, если со строками или столбцами этой матрицы провести элементарные преобразования.*

Довольно часто бывает нужно найти не только ранг матрицы, но и некоторый максимальный линейно независимый набор столбцов (или строк) этой матрицы. В этом случае помогает следующий факт.

Лемма 3.1.8. *Пусть A есть $n \times m$-матрица, и ее столбцы с номерами $j_1, \ldots, j_k$ линейно независимы. Тогда любая совокупность элементарных преобразований со* **строками** *этой матрицы дает матрицу, в которой столбцы с теми же номерами $j_1, \ldots, j_k$ остаются линейно независимыми.*

Аналогично и для строк: если в A строки с номерами $i_1, \ldots, i_k$ были линейно независимыми, то после любого набора элементарных преобразований со **столбцами** *матрицы мы получим матрицу, в которой строки с теми же номерами $i_1, \ldots, i_k$ будут линейно независимыми.*

Напомним еще одну из основных теорем из теории линейных уравнений. Пусть A есть $n \times m$-матрица. Рассмотрим систему однородных линейных уравнений:

$$Ax = 0$$

В более подробной записи это выглядит так:

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_m &= 0 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,m}x_m &= 0 \\ \qquad \cdots\cdots\cdots\cdots\cdots & \cdots \quad \cdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,m}x_m &= 0 \end{cases}$$

Теорема 3.1.10. *Множество решений однородной системы линейных уравнений является векторным подпространством пространства K^m. Размерность этого подпространства равна $m - \mathrm{rk}(A)$ (количество переменных минус ранг матрицы системы).*

Напомним, что базис пространства решений однородной системы линейных уравнений называется *фундаментальной системой решений* данной системы линейных уравнений.

Сформулируем, наконец, следующий алгоритм нахождения базиса пересечения двух конечномерных подпространств векторного пространства:

Теорема 3.1.11. *Пусть $U, W \subseteq V$ — подпространства векторного пространства V над K. Пусть $u_1, \ldots, u_n$ — базис U, $w_1, \ldots, w_m$ — базис W, и пусть $u_1, \ldots, u_n, w_1, \ldots, w_k$ — базис подпространства $U + W$. Выразим элементы $w_{k+1}, \ldots, w_m$ через этот базис:*

$$w_{k+s} = \sum_{i=1}^{n} a_{i,k+s} u_i + \sum_{j=1}^{k} b_{j,k+s} w_j, \quad s = 1, \ldots, m-k.$$

Тогда элементы $v_s = \sum\limits_{i=1}^{n} a_{i,k+s} u_i$, $s = 1, \ldots, m-k$, образуют базис $U \cap W$.

3.2. Группы

Группой называется множество G, на котором определена бинарная операция (умножение):

$$G \times G \longrightarrow G, \quad (g_1, g_2) \mapsto g_1 g_2,$$

такая, что выполняются следующие свойства:

1) (асоциативность) $(g_1 g_2) g_3 = g_1 (g_2 g_3)$ для любых $g_1, g_2, g_3 \in G$;

2) существует $e \in G$, такой, что для всех $g \in G$ имеют место равенства: $ge = eg = e$;

3) для каждого $x \in G$ найдется $y \in G$ такой, что $xy = yx = e$.

Покажем, что элемент y из свойства 3) определяется однозначно. Допустим, что для данного x нашлось два обратных элемента y_1 и y_2. Тогда $(y_1x)y_2 = ey_2 = y_2$. Но, с другой стороны, $(y_1x)y_2 = y_1(xy_2) = y_1e = y_1$. Итак, $y_1 = y_2$. Так как обратный к $g \in G$ элемент определяется однозначно, его обозначают как g^{-1}. Свойство единственности g^{-1} используется при доказательстве некоторых важных соотношений. Покажем, например, что в любой группе G для всех $x, y \in G$ имеет место равенство:

$$(xy)^{-1} = y^{-1}x^{-1}$$

Для этого достаточно проверить, что $(xy)(y^{-1}x^{-1}) = (y^{-1}x^{-1})(xy) = e$, что не должно вызывать затруднений. Отсюда следует, что элемент $y^{-1}x^{-1}$ обладает в точности теми же самыми свойствами, которые характеризуют $(xy)^{-1}$. Ввиду единственности обратного элемента заключаем, что $(xy)^{-1} = y^{-1}x^{-1}$. Индукцией нетрудно показать, что

$$(x_1x_2 \dots x_n)^{-1} = x_n^{-1} \dots x_2^{-1}x_1^{-1}$$

для всех n.

Отметим еще, что $(x^{-1})^{-1} = x$. Это также можно установить, используя свойство единственности обратного элемента. Положим $y = x^{-1}$, и найдем y^{-1}. Для этого достаточно заметить, что равенства $xy = yx = e$ могут служить не только определением обратного элемента для x, но и обратного элемента для y, а этим элементом оказывается именно x, и только он, ввиду единственности обратного для y.

И еще одно (возможно, тривиальное) замечание. Элемент

$$\overbrace{xx \dots x}^{n}$$

(n-кратное произведение x на x) принято обозначать через x^n. Будем считать очевидным, что, ввиду ассоциативности умножения, $x^nx^m =$

x^{n+m} (в некоторых книгах это равенство доказывается!). Будем также полагать по определению, что

$$x^{-n} = \overbrace{x^{-1}x^{-1}\dots x^{-1}}^{n}.$$

Проверьте, что $(x^n)^{-1} = x^{-n}$. Для групп, в которых вместо умножения пишется сложение, вместо x^{-1} надо писать $-x$, вместо x^n должно стоять $x+\dots+x = nx$, и соответственно вместо x^{-n} используется запись $-nx$.

Следующий пример является одним из центральных во всей теории групп.

Пусть F — поле. Например, это может быть любое из полей $\mathbb{Q}$ (рациональные числа), $\mathbb{R}$ (действительные числа), $\mathbb{C}$ (комплексные числа). Обозначим через $GL_n(F)$ множество всех невырожденных $n \times n$-матриц с компонентами из поля F. Напомним, что матрица A называется *невырожденной*, если ее определитель $det(A)$ не равен нулю. Это эквивалентно существованию обратной к A матрицы, то есть такой матрицы A^{-1}, что

$$AA^{-1} = A^{-1}A = E_n.$$

Здесь E_n — единичная $n \times n$-матрица. Хорошо известно, что произведение невырожденных матриц является невырожденной матрицей. Следовательно, произведение матриц определяет бинарную операцию

$$GL_n(F) \times GL_n(F) \longrightarrow GL_n(F), \quad (A, B) \mapsto AB.$$

Известно, что произведение матриц ассоциативно, а матрица E_n обладает свойством нейтрального элемента: $AE_n = E_nA = A$. Все это показывает, что $GL_n(F)$ является группой. Группа $GL_n(F)$ называется *общей линейной группой* степени n над полем F. В группе $GL_n(F)$ определена операция транспонирования: $A \mapsto A^{\mathrm{T}}$, где i,j-й элемент матрицы A^{T}

равен j, i-му элементу A для всех $1 \leq i, j \leq n$. Одно из свойств операции транспонирования таково: $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$. Кроме того $(A^{\mathrm{T}})^{\mathrm{T}} = A$. Это показывает, что операция транспонирования походит на операцию взятия обратного элемента. Докажем, что

$$(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}.$$

Пусть $X = A^{\mathrm{T}}$. Любая матрица Y, такая, что $XY = YX = E_n$, будет обратной к X. Покажем, что в качестве Y можно взять $(A^{-1})^{\mathrm{T}}$. В самом деле, используя свойства транспонирования, получим:

$$XY = A^{\mathrm{T}}(A^{-1})^{\mathrm{T}} = (A^{-1}A)^{\mathrm{T}} = E_n^{\mathrm{T}} = E_n,$$

и точно так же проверяется, что $YX = E_n$. Ввиду единственности обратного элемента в группе требуемое равенство доказано.

Отметим, что при $n = 1$ группа $GL_n(F)$ является множеством всех ненулевых элементов поля F, а операция умножения 1×1-матриц сводится к операции умножения элементов поля.

Гомоморфизм h из группы G_1 в группу G_2 — это отображение $h : G_1 \longrightarrow G_2$, удовлетворяющее следующим двум свойствам. Во-первых, для любых $x, y \in G_1$ имеет место равенство $h(xy) = h(x)h(y)$. Во-вторых, нейтральный элемент группы G_1 должен отображаться в нейтральный элемент группы G_2, то есть $h(e) = e$, или $h(1) = 1$, если нейтральные элементы обозначены символом 1.

Если из контекста не будет ясно, к какой группе принадлежит тот или иной нейтральный элемент,то надо использовать обозначения вида e_{G_1} или e_1 для нейтрального элемента G_1, и т.п.

Докажем, что для каждого $g \in G_1$ имеет место равенство:

$$h(g^{-1}) = h(g)^{-1}.$$

Положим $x = h(g)$, и пусть $y = h(g^{-1})$. Тогда

$$xy = h(g)h(g^{-1}) = h(gg^{-1}) = h(e) = e,$$
$$yx = h(g^{-1})h(g) = h(g^{-1}g) = h(e) = e.$$

Таким образом, $y = x^{-1}$, что и утверждалось.

Гомоморфизм h называется *изоморфизмом*, если существует гомоморфизм $f : G_2 \longrightarrow G_1$, такой, что $hf = 1_{G_2}$ и $fh = 1_{G_1}$. Здесь через 1_{G_1} и 1_{G_2} обозначаются тождественные отображения G_1 и G_2.

Иными словами, для каждого $x \in G_1$ имеет место равенство $f(h(x)) = x$, а для каждого $y \in G_2$ — равенство $h(f(y)) = y$.

Подгруппой G' группы G называется подмножество $G' \subseteq G$, обладающее следующими свойствами:

1) нейтральный элемент (единица) группы G принадлежит G';

2) из $x, y \in G'$ следует $xy \in G'$;

3) если $x \in G'$, то $x^{-1} \in G'$.

Это определение означает, что, если взять ограничение бинарной операции для G на $G' \times G' \subseteq G \times G$, то его можно рассматривать как отображение в G', и относительно этой бинарной операции множество G' само становится группой, причем отображение включения $G' \subseteq G$ оказывается гомоморфизмом групп. Сама группа G и множество $\{e\}$ являются подгруппами G. Эти подгруппы принято называть тривиальными.

Вот важный пример подгруппы: легко проверить, что множество $SL_n(F) = \{A \in GL_n(F) | det(A) = 1\}$ является подгруппой группы $GL_n(F)$. $SL_n(F)$ называется *специальной линейной группой* n-й степени над полем F.

Сформулируем в явном виде аксиомы группы для случая, когда групповая операция записывается как $x + y$ (сложение). Операция сложения должна удовлетворять следующим свойствам:

1) (ассоциативность) $(g_1 + g_2) + g_3 = g_1 + (g_2 + g_3)$ для любых $g_1, g_2, g_3 \in G$;

2) существует элемент $0 \in G$, такой, что для всех $g \in G$ имеют место равенства: $g + 0 = 0 + g = e$;

3) для каждого $x \in G$ найдется $y \in G$ такой, что $x + y = y + x = 0$. В аддитивной записи обратный элемент y обозначается как $-x$, при этом используется также обозначение $a - b = a + (-b)$.

Аддитивная запись групповой операции чаще всего используется для коммутативных групп, то есть групп, в которых

4) $x + y = y + x$ для всех $x, y \in G$.

Такие группы часто называются *абелевыми*. Простейший пример такой группы — группа $\mathbb{E}$ всех целых чисел.

Гомоморфизм абелевых групп $h : G_1 \longrightarrow G_2$ должен удовлетворять свойствам:

1) $h(x + y) = h(x) + h(y)$ для всех $x, y \in G_1$;

2) $h(0) = 0$.

Отсюда следует, что $h(-x) = -h(x)$.

Каждое векторное пространство является абелевой группой по сложению. Каждое линейное отображение векторных пространств является гомоморфизмом абелевых групп.

Таким образом, теорию групп можно считать обобщением теории векторных пространств и линейных отображений.

3.3. Кольца и поля

Кольцом называется множество R с двумя бинарными операциями, сложением и умножением, такими, что относительно сложения R есть абелева группа, и выполняется свойство дистрибутивности (билинейности) умножения: для всех $x, y, z \in R$

$$x(y+z) = xy + xz \qquad , \qquad (x+y)z = xz + yz \ .$$

Кольцо называется *ассоциативным*, если операция умножения в R ассоциативна, то есть для всех $x, y, z \in R$ имеет место равенство $x(yz) = (xy)z$. Кольцо называется *коммутативным*, если $xy = yx$ для всех $x, y \in R$. Говорят, что R есть *кольцо с единицей*, если в нем есть элемент, обычно обозначаемый как 1, и обладающий свойством: $1 \cdot a = a \cdot 1 = a$ для каждого $a \in R$. В дальнейшем рассматриваются только ассоциативные кольца с единицей, называемые просто *кольцами*.

Коммутативное кольцо с единицей называется *полем*, если для каждого ненулевого элемента $a \in R$ найдется элемент $b \in R$ такой, что $ab = ba = 1$. В предыдущем параграфе, где речь шла о группах, было показано, что такой элемент b определен однозначно (доказательство в случае элементов, принадлежащих не группе, а полю, остается тем же самым). Этот элемент называется *обратным элементом* к элементу a и обозначается через a^{-1}. Свойства обратных элементов в полях те же самые, что и у обратных элементов в группах, но ввиду коммутативности, например, $(ab)^{-1} = a^{-1}b^{-1}$. Легко убедиться, что все ненулевые элементы поля образуют группу по умножению. У нуля нет обратного элемента, если не считать случая поля из одного элемента, в котором нуль совпадает с единицей. Однако такое поле мы рассматривать не будем.

Пусть R, S — кольца. *Гомоморфизмом* колец называется отображе-

ние $f : R \longrightarrow S$, обладающее следующими свойствами:

$$f(x \pm y) = f(x) \pm f(y), \quad f(0) = 0,$$
$$f(xy) = f(x)f(y), \qquad f(1) = 1.$$

Напомним некоторые примеры колец и полей.

Пример 3.3.1. Числовые кольца и поля: $\mathbb{Z}$ — кольцо целых чисел (оно не является полем), $\mathbb{Q}$ — поле рациональных чисел, $\mathbb{R}$ — поле действительных чисел, $\mathbb{C}$ — поле комплексных чисел.

Пример 3.3.2. Множество многочленов от n переменных $R[x_1, \ldots, x_n]$ с коэффициентами из кольца R является кольцом относительно обычных операций сложения и умножения многочленов. При этом кольцо коэффициентов R может и не быть коммутативным, но требуется, чтобы переменные x_i коммутировали между собой и со всеми элементами из R.

Пример 3.3.3. Пусть R — некоторое кольцо. Множество всех квадратных $n \times n$-матриц с компонентами из R образует кольцо относительно обычных операций сложения и умножения матриц. Коммутативность R не обязательна. Кольцо $n \times n$-матриц над R обозначается через $M_n(R)$.

В теории линейных операторов на векторных пространствах большую роль играют кольца вида $M_n(K[x])$, где $K[x]$ — кольцо многочленов над полем K. Иногда бывает необходимо рассматривать кольцо многочленов вида $M_n(K)[x]$, т.е. многочлены, коэффициентами которых являются матрицы n-го порядка над полем K. Можно показать, что $M_n(K[x])$ и $M_n(K)[x]$ — это, по сути, одно и то же кольцо.

Пример 3.3.4. В параграфе 1.6 фактически было построено кольцо, состоящее из всех выражений вида $f(\mathcal{A})$, где $\mathcal{A}$ — нкоторый фиксированный линейный оператор, а в качестве $f(x)$ берутся всевозможные

многочлены из $K[x]$. Это кольцо можно обозначить по аналогии с кольцом многочленов через $K[\mathcal{A}]$. В дальнейшем, когда будет изучено понятие факторкольца по идеалу, станет понятно, что $K[\mathcal{A}]$ являтся факторкольцом кольца многочленов по идеалу, порожденному минимальным многочленом оператора $\mathcal{A}$ (точнее, изоморфно этому факторкольцу).

Пример 3.3.5. Рассмотрим множество из двух элементов 0 и 1, и определим операции сложения и умножения с этими элементами следующим образом:

$0+0=0, 0+1=1+0=1, 1+1=0,$

$0\cdot 0=0, 1\cdot 0=0\cdot 1=0, 1\cdot 1=1.$

Легко проверяется, что выполнены все свойства поля. Это поле называется еще *полем Галуа* из двух элементов и обозначается через $\mathbb{F}_2$ или через $GF(2)$.

Опишем вкратце, как устроены другие конечные поля. Возьмем произвольное простое число p и рассмотрим множество $\{0, 1, \ldots, p-1\}$, которое обозначим через $\mathbb{F}_p$ или через $GF(p)$. (GF означает "Galois Field", т.е. поле Галуа.) Определим на этом множестве новые операции "сложения" и "умножения" следующим образом. Полагаем

$$a \boxplus b = a + b \pmod p, \quad a \boxtimes b = a \times b \pmod p,$$

где $\pmod p$ означает взятие остатка от деления на p. Можно показать, что новые операции определяют на $GF(p)$ структуру поля, в котором 0 и 1 обладают обычными свойствами. Обратный по умножению элемент к элементу $a \neq 0$ определяется так. Находится число c такое, что $1 \leq c \leq p-1$ и $ac = 1+pl$ (это делается с помощью алгоритма Евклида). Число c и будет обратным к a относительно новой операции умножения. Заметим, что на практике новые обозначения для введенных выше новых сложения и умножения не используются. Употребляются обычные

обозначения $a + b$ и ab, смысл определяется из контекста.

Таким способом строятся, однако, далеко не все конечные поля. Общий случай выглядит так. Пусть K есть одно из полей $GF(p)$ для некоторого простого p, построенное выше. Рассмотрим кольцо многочленов $K[x]$. Для многочленов с коэффициентами из конечного поля справедливы многие свойства многочленов с действительными коэффициентами (и доказательства у них те же самые). В частности, это относится к свойствам степеней при умножении, к делимости, к наличию алгоритма Евклида, и к однозначности разложения многочлена в произведение неприводимых многочленов. Выберем в $K[x]$ некоторый неприводимый многочлен f и пусть n есть его степень. Рассмотрим в $K[x]$, рассматриваемом как векторное пространство над K, векторное подпространство F, состоящее из всех многочленов, степени которых строго меньше n. Определим на этом векторном пространстве новую операцию умножения, полагая произведение двух многочленов $u(x)$ и $v(x)$ равным остатку от деления обычного произведения $u(x)v(x)$ на неприводимый многочлен f. Можно показать, что эта операция (вместе с уже определенными в F сложением и вычитанием) превращает F в поле, содержащее в качестве подполя поле K. Обратный элемент для $v(x) \neq 0$ определяется следующим образом. Находится многочлен $w \in K[x]$ с условием $deg(w) < n$ и такой, что $vw = 1 + fh$ (это опять делается с помощью алгоритма Евклида). Многочлен $w \in F$ и будет обратным к u относительно новой операции умножения.

Легко заметить, что если рассматривать F как векторное пространство над K, то в качестве его базиса можно выбрать элементы $1, x, x^2, \ldots, x^{n-1}$. Поскольку количество элементов в поле K равно p, отсюда следует, что количество элементов в поле F равно p^n.

Можно показать, что если мы выберем другой неприводимый много-

член той же степени n, то построенное с его помощью новое поле будет изоморфно полю F и как кольцо, и как векторное пространство над K. Поэтому числа p и n полностью определяют конечное поле (с точностью до изоморфизма, как это обычно и бывает в математике). Это поле обозначается через $GF(p^n)$. При $n = 1$ (случай неприводимого многочлена вида $x - a$) все сводится к уже построенному выше $GF(p)$.

Можно показать, что описанными выше способами получаются все конечные поля. Можно также показать, что для любого $n \geq 1$ и каждого простого p всегда существует поле $GF(p^n)$. Таким образом, если где-то встречается обозначение вида $GF(q)$ или $\mathbb{F}_q$, то это означает, что $q = p^n$, где p — простое число.

3.4. Многочлены

Обозначим через $K[x]$ кольцо многочленов (или полиномов) от одной переменной x над полем K.

Напомним, что если $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ — многочлен с коэффициентами из поля K, и $a_n \neq 0$, то число n называется *степенью* многочлена f и обозначается через $\deg(f)$. Многочлены степени нуль — это в точности ненулевые элементы поля K. По определению, степень нуля равна $-\infty$. С учетом этого справедлив следующий факт:

$$\deg(fg) = \deg(f) + \deg(g)$$

для любых многочленов f и g из $K[x]$.

Теорема 3.4.1. (Алгоритм Евклида) *Если $f(x)$ и $g(x)$ — два многочлена над полем K, и $g \neq 0$, то найдутся такие многочлены $h(x)$ и $r(x)$, что $f(x) = h(x)g(x) + r(x)$, причем $\deg(r) < \deg(g)$. Многочлены h и r, удовлетворяющие этим условиям, определяются однозначно.*

Многочлен $r(x)$ называется *остатком* от деления f на g, а многочлен $h(x)$ — *частным* от деления f на g. Если $r(x) = 0$, то говорят, что f *делится на* g *без остатка*, или просто делится на g. В этом случае говорят также, что g делит f.

Отметим одно важное обстоятельство. Если $f = hg$, то для каждого ненулевого $\lambda \in K$ имеется также равенство $f = (\lambda^{-1}h)(\lambda g)$. Поэтому, когда рассматривают делимость многочленов, приходится часто предполагать, что делители данного многочлена f определены с точностью до умножения на ненулевой элемент поля (на скаляр).

Наибольшим общим делителем семейства многочленов $f_1, \ldots, f_k$, ни один из которых не равен нулю, называется многочлен $d(x)$, обладающий следующими свойствами: каждый многочлен f_i делится на d без остатка и если какой-то многочлен g делит все многочлены $f_1, \ldots, f_k$, то он является делителем многочлена d. Обозначение для наибольшего общего делителя (НОД):

$$d = \text{НОД}(f_1, \ldots, f_k), \quad \text{или} \quad d = \gcd(f_1, \ldots, f_k).$$

Теорема 3.4.2. *Наибольший общий делитель многочленов существует и определен однозначно с точностью до умножения на ненулевой элемент поля.*

Многочлены называются *взаимно простыми*, если их наибольший общий делитель является ненулевым элементом поля. В этом случае его можно считать равным единице.

Теорема 3.4.3. (Тождество Безу) *Пусть* $d = \text{НОД}(f_1, \ldots, f_k)$. *Тогда найдутся такие многочлены* $w_1(x), \ldots, w_k(x)$, *что*

$$d = w_1 f_1 + w_2 f_2 + \cdots + w_k f_k.$$

Следствие 3.4.1. *Многочлены* $f_1, \ldots, f_k$ *являются взаимно простыми в том и только в том случае, если найдутся такие многочлены*

$w_1(x), \ldots, w_k(x)$, *что*

$$w_1 f_1 + w_2 f_2 + \cdots + w_k f_k = 1.$$

Многочлен $f(x) \in K[x]$ называется *неприводимым*, если f нельзя представить в виде $f = f_1 f_2$, где $\deg(f_1) \geq 1$ и $\deg(f_2) \geq 1$. Неприводимые многочлены являются аналогами простых чисел в кольце $K[x]$. Теория делимости в кольце многочленов почти полностью аналогична теории делимости для целых чисел.

Заметим, что любой многочлен вида $ax + b$ является неприводимым, а многочлен вида $f(x) = ax^2 + bx + c$ над полем действительных чисел неприводим в том и только в том случае, если его дискриминант отрицателен: $b^2 - 4ac < 0$. Если при этом $a > 0$, то $f(x) > 0$ для всех действительных x.

Теорема 3.4.4. *Каждый отличный от нуля многочлен f из $K[x]$ может быть разложен в произведение*

$$f = a f_1^{n_1} f_2^{n_2} \ldots f_m^{n_m} \tag{3.4.1}$$

где $a \in K$, $a \neq 0$, а $f_1, \ldots, f_m$ — неприводимые многочлены. При этом каждый неприводимый многочлен из этого набора определяется однозначно с точностью до ненулевого скалярного множителя (элемента поля), а степени $n_1, \ldots, n_m$ определены строго однозначно. В случае, если f есть многочлен со старшим коэффициентом, равным a, можно выбрать все f_i так, чтобы их старшие коэффициенты были равны единице, и тогда представление f в виде (3.4.1) *полностью однозначно.*

Если дано разложение (3.4.1) многочлена f в произведение неприводимых множителей $f = a f_1^{n_1} f_2^{n_2} \ldots f_k^{n_k}$, то положительные целые числа

$n_1, \ldots, n_m$ называются *кратностями*, с которыми неприводимые многочлены $f_1, \ldots, f_m$ входят в разложение f в произведение неприводимых множителей.

Режде, чем сформулировать критерий для определения кратности вхождения неприводимого множителя, напомним, что если многочлен f имеет вид $f = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = \sum\limits_{j=0}^{n} a_j x^j$, то его *производная* — это многочлен $f' = a_1 + 2a_2 x + \cdots + n a_n x^{n-1} = \sum\limits_{j=1}^{n} j a_j x^{j-1}$. Вторая производная f'' определяется как $(f')'$, третья f''' — как $(f'')'$, и если уже определена k-я производная $f^{(k)}$, то $k+1$-я производная $f^{(k+1)}$ есть $(f^{(k)})'$.

Лемма 3.4.1. *Неприводимый многочлен $g(x)$ входит в разложение (3.4.1) многочлена $f(x)$ с кратностью, большнй или равной двум, тогда и только тогда, если $g(x)$ является делителем многочлена НОД(f, f').*

Корень многочлена $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \in K[x]$ — это такой элемент α поля K, что выполнено равенство: $f(\alpha) = a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_n\alpha^n = 0$.

Лемма 3.4.2. $f(\alpha) = 0 \Longleftrightarrow f(x) = (x - \alpha)g(x)$.

Если α — корень $f(x)$, и можно представить f в виде $f(x) = (x - \alpha)^k g(x)$, причем $g(\alpha) \neq 0$, то говорят, что корень α имеет *кратность* k.

Лемма 3.4.3. *Корень α многочлена $f(x)$ имет кратность k тогда и только тогда, если он является также корнем первых $k-1$ производных многочлена f:*

$$f(\alpha) = f'(\alpha) = f''(\alpha) = \cdots f^{(k-1)}(\alpha) = 0.$$

Следующая теорема когда-то называлась “основной теоремой алгебры”.

Теорема 3.4.5. 1) *Каждый неприводимый многочлен над полем* $\mathbb{R}$ *имеет либо степень, равную единице, либо степень, равную двум.*

2) *Каждый неприводимый многочлен над нолем* $\mathbb{C}$ *имеет степень единица.*

3) *У каждого многочлена нечетной степени из* $\mathbb{R}[x]$ *имеется действительный корень.*

4) *Каждый многочлен над полем комплексных чисел, имеющий положительную степень, имеет корень в поле комплексных чисел.*

5) *Из этого следует, что каждый многочлен над полем комплексных чисел можно представить в виде:*

$$f(x) = a(x-\alpha_1)^{k_1}(x-\alpha_2)^{k_2}\dots(x-\alpha_m)^{k_m}.$$

Здесь $\deg(f) = k_1+\dots+k_m$, $\alpha_1,\dots,\alpha_m$ *— корни* f, $a \in K$ *— старший коэффициент* f. *Представление* f *в таком виде определено однозначно.*

Поля, для многочленов над которыми справедливы утверждения пунктов 4) и 5) этой теоремы, называются *алгебраически замкнутыми.* Поле $\mathbb{C}$ алгебраически замкнуто, поле $\mathbb{R}$ нет. Из общей теории полей известно, что каждое поле K можно сделать подполем некоторого алгебраически замкнутого поля.

Так как $\mathbb{R} \subset \mathbb{C}$, то каждый многочлен с действительными коэффициентами можно считать и многочленом с комплексными коэффициентами. Поэтому у каждого могочлена положительной степени из $\mathbb{R}[x]$ обязательно есть комплексный корень.

Лемма 3.4.4. *Пусть* $f(x) \in \mathbb{R}[x]$, $z \in \mathbb{C}$ *и* $f(z) = 0$. *Тогда* $f(\overline{z}) = 0$.

Таким образом, если $z = a + bi$ — комплексный корень многочлена f, имеющего действительные коэффициенты, и $b \neq 0$, то f делится на многочлен с действительными коэффициентами $(x - z)(x - \overline{z}) = x^2 - 2ax + a^2 + b^2$. Этот многочлен неприводим, и с точностью до скалярного множителя такой вид имеют все неприводимые многочлены степени два из $\mathbb{R}[x]$.

Литература

[1] Артамонов В.А. Введение в высшую алгебру и аналитическую геометрию. — М.: Факториал Пресс, 2007.

[2] Винберг Э.Б. Курс алгебры. — 3-е изд., испр. и доп. — М.: Изд-во “Факториал Пресс”, 2002.

[3] Гельфанд И.М. Лекции по линейной алгебре. — 5-е изд., исправленное. — М.: Добросвет, Московский центр непрерывного математического образования, 1998. — 320 с.

[4] Методические указания к курсу “Линейная алгебра и геометрия” по теме “Линейные преобразования” / Составитель Ю.Б. Ермолаев. — Казань: Казанский государственный университет, 1987.

[5] Ильин В.А., Позняк Э.Г. Линейная алгебра. — М.: Наука. Гл. ред. физ.-мат. лит., 1974.

[6] Ильин С.Н. Элементы алгебры: комплексные числа, системы линейных уравненйи, многочлены: Учебное пособие. — Казань: Казанский государственный университет, 2006.

[7] Корешков Н.А. Линейные операторы: Учебное пособие. — Казань: Казанский государственный университет, 2004.

[8] Кострикин А.И., Манин Ю.И. Линейная алгебра и геометрия. — 2-е изд., перераб. — М.: Наука. Гл. ред. физ.-мат. лит., 1986.

[9] Кострикин А.И. Введение в алгебру. Часть I. Основы алгебры. — 2-е изд., исправл. — М.: Физико-математическая литература, 2001.

[10] Кострикин А.И. Введение в алгебру. Часть II. Линейная алгебра. — 2-е изд., исправл. — М.: Физико-математическая литература, 2001.

[11] Кострикин А.И. Введение в алгебру. Часть III. Основные структуры. — 2-е изд., исправл. — М.: Физико-математическая литература, 2001.

[12] Сборник задач по алгебре / Под ред. А.И.Кострикина. — Ид. 3-е, испр. и доп. — М.:ФИЗМАТЛИТ, 2001.

[13] Курош А.Г. Курс высшей алгебры. — Изд. десятое, стереотипное. — М.: Наука. Гл. ред. физ.-мат. лит., 1971.

[14] Мальцев А.И. Основы линейной алгебры. — Изд. третье, перераб. — М.: Наука. Гл. ред. физ.-мат. лит., 1979.

[15] Шевцов Г.С. Линейная алгебра: теория и прикладные аспекты. — 2-е изд., испр. и доп. — М.: Магистр: ИНФРА-М, 2011.

[16] Лекции по алгебре. Семестр 2. Выпуск I. Линейные отображения и линейные операторы: Учебно-методическое пособие / С.Н. Тронин. — Казань: Казанский (Приволжский) федеральный университет, 2012.

[17] Тронин С.Н. Введение в теорию групп. Задачи и теоремы. Часть 1: Учебное пособие. — Казань: Казанский государственный университет, 2006.