С.Н.Тронин

ЛЕКЦИИ ПО АЛГЕБРЕ

Семестр 2

Выпуск I

Линейные отображения и линейные операторы

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, доцент А.Н. Абызов

Лекции по алгебре. Семестр 2. Выпуск I. **Линейные отображения и линейные операторы:** Учебно-методическое пособие / С.Н. Тронин. — Казань: Казанский (Приволжский) федеральный университет, 2012. — 67 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. В первом выпуске излагаются основы теории линейных отображений и линейных операторов. 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. Второй выпуск посвящен жордановой нормальной форме линейного оператора, и, кроме того, содержит справочную информацию по материалу первого семестра. В третьем выпуске излагается теория евклидовых и унитарных пространств, а также теория линейных операторов, действующих на таких пространствах.

Векторные пространства и линейные отображения, появившиеся в середине XIX века, к настоящему времени стали важнейшими понятиями не только для алгебры, но и для большинства прочих разделов математики. С одной стороны, речь, конечно, идет о геометрии многомерных пространств. С другой стороны, например, в курсе математического анализа показывается, что производная векторной функции от векторного аргумента представляет собой линейное отображение. В частности, линейным отображением является операция взятия обычной производной, изучаемая в первом семестре. Интегрирование также является линейным отображением. Огромное количество примеров линейных отображений (и их частного случая — линейных операторов) можно найти в физике. Например, современная квантовая теория исходит из того, что физическая величина — это линейный оператор (а не скаляр, как в старой ньютоновской физике), а то, что измеряет прибор, есть собственное значение данного оператора. Ввиду всего этого овладение основами теории векторных пространств и линейных отображений для студента-математика не менее важно, чем, например, овладение основами дифференциального и интегрального исчисления.

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 [10], [11], [12]. Второй том этого учебника [11] — это та книга, которую можно рекомендовать в первую очередь для основательного изучения материала, относящегося ко всем выпускам нашего пособия. Необходимо также отметить предшествовавшую [11] книгу высокого уровня [9], из которой можно узнать, например, много интересного 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 [4], В.А.Ильина и Э.Г.Позняка [6], А.Г. Куроша [14], и А.И. Мальцева [15]. Из более современных учебников отметим книги Э.Б. Винберга [2], В.А. Артамонова [1] и Г.С. Шевцова [16]. Наконец, в список литературы включено несколько учебных пособий сотрудников кафедры алгебры и математической логики Казанского федерального университета [5], [7], [8].

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 1.8 первой главы может присутствовать определение 1.8.2, теорема 1.8.2 и формула (1.8.2). Опыт показывает, что к этому можно легко привыкнуть. Доказательства заканчиваются символом $\Box$.

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

ГЛАВА I. ЛИНЕЙНЫЕ ОТОБРАЖЕНИЯ

1.1. Основная теорема

Через K в дальнейшем будет обозначаться некоторое поле. Основные интересующие нас примеры полей таковы: $\mathbb{Q}$ (поле рациональных чисел), $\mathbb{R}$ (поле действительных чисел), $\mathbb{C}$ (поле комплексныз чисел).

Определение векторного пространства над полем K, и базиса векторного пространства предполагается известным. Определения и основные свойства можно найти также в параграфе 3.1 второго выпуска данного пособия.

Напомним некоторые примеры векторных пространств.

Пример 1.1.1. Само поле K является одномерным векторным пространством над K. В качестве базиса можно взять любой ннулевой элемент поля K.

Пример 1.1.2. Поле $\mathbb{C}$ есть векторное пространство над полем $\mathbb{R}$ с базисом 1, i.

Пример 1.1.3. Множество $m \times n$-матриц $M_{m,n}(K)$ над полем K есть векторное пространство над K с базисом, состоящим из *матричных единиц* — матриц $E_{i,j}$, i, j-е элементы которых равны единице, а все остальные компоненты нулевые. Если $A \in M_{m,n}(K)$, $a_{k,l}$ — элементы A, то

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j} E_{i,j}.$$

Пример 1.1.4. Пусть в предыдущем примере $n = 1$, т.е. рассматриваются матрицы с m строками и одним столбцом. В этом случае вместо $M_{m,1}(K)$ используется обозначение K^m. Это векторное пространство называется пространством столбцов высоты m. Его базис (так называемый *стандартный базис*) состоит из столбцов e_i, в которых i-я компонента равна единице, а все остальные нулевые.

Определение 1.1.1. Пусть V и W — векторные пространства над полем K. Отображение $f : V \longrightarrow W$ называется *линейным*, если выполнены следующие свойства:

1. $f(v_1 + v_2) = f(v_1) + f(v_2)$;
2. $f(\lambda v) = \lambda f(v)$.

для произвольных $v, v_1, v_2 \in V$, $\lambda \in K$.

Все эти условия можно заменить одним: $f(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 f(v_1) + \alpha_2 f(v_2)$ для любых векторов $v_1, v_2 \in V$ и произвольных элементов $\alpha_1, \alpha_2 \in K$. Заметим, что из опредления следует $f(0) = 0$, $f(-v) = -f(v)$.

Когда надо подчеркнуть, над каким полем определены векторные пространства, говорят о K-*линейном отображении.*

Из определения линейного отображения следует также, что

$$f(\sum_{i=1}^{n} \lambda_i v_i) = \sum_{i=1}^{n} \lambda_i f(v_i)$$

для любого $m \geq 1$, каких угодно векторов $v_1, \ldots, v_n \in V$, и любых $\lambda_1, \ldots, \lambda_n \in K$. Это легко доказывается индукцией по n.

Определение 1.1.2. Линейное отображение вида $f : V \longrightarrow V$ называется *линейным оператором.*

Рассмотрим несколько примеров линейных отображений.

Пример 1.1.5. Пусть V и W — любые векторные пространства. Рассмотрим отображение $f : V \to W$ такое, что $f(v) = 0$ для каждого вектора $v \in V$. Легко проверяется, что это линейное отображение.

Пример 1.1.6. Пусть V — произвольное векторное пространство, $\alpha \in K$ — фиксированный элемент поля. Отображение $f : V \to V$ такое, что $f(v) = \alpha v$, будет линейным. Если взять $\alpha = 1$, то получим тождественное (единичное) отображение $V \to V$, отображающее каждый вектор в самого себя. Оно, таким образом, тоже будет линейным.

Если пространство V одномерно, то каждое линейное отображение $f : V \to V$ имеет вид $f(v) = \alpha v$ для какого-то $\alpha \in K$. В самом деле, пусть e — базисный вектор V. Тогда $f(e) = \alpha e$ для некоторого $\alpha \in K$, поскольку по условию (одномерность V) каждый вектор $v \in V$ (а значит, и $f(e)$) имеет вид $v = \lambda e$, где $\lambda \in K$. Элемент α определен однозначно и не зависит от выбора базиса (это надо обосновать!). Для произвольного $v = \lambda e$ теперь получаем, что

$$f(v) = f(\lambda e) = \lambda f(e) = \lambda \alpha e = \alpha(\lambda e) = \alpha v,$$

что и требовалось доказать.

Пример 1.1.7. Рассмотрим $\mathbb{C}$ как векторное пространство над $\mathbb{R}$, и пусть $\alpha \in \mathbb{C}$ — фиксированный элемент. Отображение $f : \mathbb{C} \to \mathbb{C}$ такое, что $f(v) = \alpha v$, будет $\mathbb{R}$-линейным. Если $\alpha = \cos\varphi + i \sin\varphi$, а элементы $\mathbb{C}$ представлять как векторы на плоскости, то отображение f можно геометрически описать как поворот на угол φ против часовой стрелки.

Пример 1.1.8. Пусть $V = K^n$, $W = K^m$, и A — произвольная матрица из $M_{m,n}(K)$. Тогда отображение $f : K^n \longrightarrow K^m$, такое, что $f(x) = Ax$ (матричное умножение матрицы на столбец), будет линейным отображением.

Следующую теорему ввиду ее важности можно назвать *“Основной теоремой о линейных отображениях”*.

Теорема 1.1.1. *Пусть* V *и* W *— векторные пространства над полем* K, $v_1, \dots v_n$ *— базис* V.

1) *Пусть* $f, g : V \longrightarrow W$ *— линейные отображения, и* $f(v_i) = g(v_i)$ *для каждого* i, $1 \leq i \leq n$. *Тогда* $f = g$.

   Отметим, что тут достаточно предполагать, что вектора $v_1, \dots, v_n$ *только порождают* V *(т.е. их линейная оболочка совпадает с* V*), и не обязательно линейно независимы.*

2) *Пусть выбраны какие угодно элементы* $u_1, \dots, u_n \in W$. *Тогда существует линейное отображение* $f : V \longrightarrow W$, *такое, что* $f(v_i) = u_i$ *для всех* i, *причем такое отображение* f *определено этим условием однозначно.*

Отображение f из пункта 2) строится следующим образом. Пусть $v \in V$ — произвольный вектор. Тогда его единственным способом можно представить в виде $v = \sum\limits_{i=1}^{n} \lambda_i v_i$. Коэфициенты $\lambda_i \in K$ определяются по v однозначно. Полагаем

$$f(v) = \sum_{i=1}^{n} \lambda_i u_i \tag{1.1.1}$$

Доказательство. Докажем пункт 1). Два отображения f и g равны по определению тогда и только тогда, если $f(v) = g(v)$ для любого $v \in V$. Но так как у нас есть базис $v_1, \dots, v_n$ пространства V, можно представить v в виде $v = \sum\limits_{i=1}^{n} \alpha_i v_i$, где $\alpha_i \in K$ для всех i. Вычислим $f(v)$ с учетом того, что f — линейное отображение:

$$f(v) = f(\sum_{i=1}^{n} \alpha_i v_i) = \sum_{i=1}^{n} \alpha_i f(v_i).$$

Аналогично,

$$g(v) = \sum_{i=1}^{n} \alpha_i g(v_i).$$

Но так как по условию $f(v_i) = g(v_i)$ для всех i, то правые части этих выражений равны, а значит, равны и левые, то есть $f(v) = g(v)$. Ввиду произвольности v отсюда следует $f = g$.

Отметим, что в приведенном выше рассуждении нигде не использовалась линейная независимость векторов $v_1, \dots, v_n$, достаточно было того, что каждый вектор $v \in V$ можно хотя бы одним способом представить в виде линейной комбинации $\sum\limits_{i=1}^{n} \alpha_i v_i$. А это и означает, что вектора $v_1, \dots, v_n$ порождают V.

Докажем пункт 2). Предположим, что линейное отображение f с указанным в пункте 2) свойством (т.е. $f(v_i) = u_i$ для всех i) существует. Пусть v — некоторый вектор из V. Запишем его в виде $v = \sum\limits_{i=1}^{n} \lambda_i v_i$. Поробуем выяснить, каким должно быть значение $f(v)$:

$$f(v) = f(\sum_{i=1}^{n} \lambda_i v_i) = \sum_{i=1}^{n} \lambda_i f(v_i) = \sum_{i=1}^{n} \lambda_i u_i.$$

Последнее равенство справедливо ввиду условия $f(v_i) = u_i$ для каждого i. Полученная формула совпадает с (1.1.1), и это означает, что если f вообще существует, то для произвольного v значение $f(v)$ определяется по формуле (1.1.1). Остается доказать, что формула (1.1.1) определяет отображение $f : V \to W$, и это отображение линейно.

Первое утверждение кажется тривиальным, но это не совсем так. Дело в том, что значение выражения, стоящего справа от знака равенства в формуле (1.1.1), зависит от коэффициентов λ_i в записи вектора v в виде линейной комбинации векторов $v_1, \dots, v_n$. Но если бы тот же вектор v можно было представить еще одним способом, например, $v = \sum\limits_{i=1}^{n} \alpha_i v_i$, и $\lambda_j \neq \alpha_j$ для некоторого j, то возник бы вопрос, как вычислять $f(v)$: то ли $f(v) = \sum\limits_{i=1}^{n} \lambda_i u_i$, то ли $f(v) = \sum\limits_{i=1}^{n} \alpha_i u_i$. Это означало бы, что отображение не определено: нет способа *однозначно* сопоставить аргументу v значение $f(v)$.

Однако в данном случае векторы $v_1, \ldots, v_n$ образуют базис, и поэтому коэффициенты λ_i определяются по v единственно возможным образом. Поэтому той неоднозначности, о которой говорилось в предыдущем абзаце, появиться не может, и формула (1.1.1) корректно определяет отображение.

Покажем, что это отображение линейно. Пусть v' и v'' — два вектора из V, $\alpha', \alpha'' \in K$. Запишем v' и v'' в виде линейных комбинаций базисных векторов:

$$v' = \sum_{i=1}^{n} \lambda_i' v_i, \quad v'' = \sum_{i=1}^{n} \lambda_i'' v_i.$$

Тогда

$$\alpha' v' + \alpha'' v'' = \sum_{i=1}^{n} (\alpha' \lambda_i' + \alpha'' \lambda_i'') v_i,$$

причем коэффициенты $\alpha' \lambda_i' + \alpha'' \lambda_i''$ определяются по вектору $\alpha' v' + \alpha'' v''$ однозначно. Согласно формуле (1.1.1),

$$f(\alpha' v' + \alpha'' v'') = \sum_{i=1}^{n} (\alpha' \lambda_i' + \alpha'' \lambda_i'') u_i,$$

по этой же формуле

$$\alpha' f(v') + \alpha'' f(v'') = \alpha' \sum_{i=1}^{n} \lambda_i' u_i + \alpha'' \sum_{i=1}^{n} \lambda_i'' u_i.$$

Ясно, что правые части равны, поэтому равны и левые, то есть

$$f(\alpha' v' + \alpha'' v'') = \alpha' f(v') + \alpha'' f(v'').$$

Теорема доказана. $\square$

1.2. Матрицы линейных отображений

Допустим, что в векторном пространстве V выбран базис $v_1, \ldots, v_n$, а в векторном пространстве W выбран базис $w_1, \ldots, w_m$. Пусть $f : V \longrightarrow$

W — линейное отображение. Тогда для каждого j, $1 \leq j \leq n$, существуют такие элементы $a_{i,j} \in K$, что имеют место равенства:

$$f(v_j) = \sum_{i=1}^{m} a_{i,j} w_i \tag{1.2.1}$$

Элементы $a_{i,j}$ определены однозначно. Составим из них матрицу с m строками и n столбцами:

$$M_f = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}$$

Определение 1.2.1. Построенная выше матрица M_f называется *матрицей линейного отображения* f относительно базисов $v_1, \dots, v_n$, $w_1, \dots, w_m$ (или в базисах $v_1, \dots, v_n$, $w_1, \dots, w_m$).

Теорема 1.2.1. *При выбранных и зафиксированных базисах в V и W соответствие $f \mapsto M_f$ задает взаимно-однозначное соответствие между линейными отображениями из V в W, и $m \times n$-матрицами с компонентами из K.*

Доказательство. Обозначим временно через $\mathcal{L}(V, W)$ множество всех линейных отображений из V в W. Когда мы сопоставляем отображению $f \in \mathcal{L}(V, W)$ матрицу M_f, мы тем самым задаем отображение из $\mathcal{L}(V, W)$ в $M_{m,n}(K)$. Чтобы доказать теорему, необходимо построить обратное отображение $M_{m,n}(K) \to \mathcal{L}(V, W)$.

Пусть дана матрица $A \in M_{m,n}(K)$ с компонентами $a_{i,j}$, $1 \leq i \leq m$, $1 \leq j \leq n$. Сначала образуем элементы $u_j = \sum\limits_{i=1}^{m} a_{i,j} w_i \in W$. Теперь, по теореме 1.1.1, существует, притом только одно, линейное отображение $h_A : V \to W$, такое, что $h_A(v_j) = u_j$ для каждого j, $1 \leq j \leq n$. Соответствие $A \mapsto h_A$ задает отображение $M_{m,n}(K) \to \mathcal{L}(V, W)$.

Проверим, что построенные отображения взаимно обратны. Пусть дано линейное отображение $f : V \to W$, т.е. элемент множества $\mathcal{L}(V, W)$. Обозначим через A его матрицу M_f. Вспоминая, как определяется эта матрица, видим, что для каждого базисного вектора v_j имеется равенство: $f(v_j) = h_A(v_j)$. Но тогда по теореме 1.1.1 $f = h_A$. Таким образом, суперпозиция отображений $\mathcal{L}(V, W) \to M_{m,n}(K) \to \mathcal{L}(V, W)$, $f \mapsto M_f = A \mapsto h_A$, отображает f в f, т.е. это тождественное отображение.

С другой стороны, пусть дана матрица A. Обозначим h_A через f, и построим по f матрицу $B = M_f$. С одной стороны, $h_A(v_j) = \sum\limits_{i=1}^{m} a_{i,j} w_i$. С другой стороны, $f(v_j) = \sum\limits_{i=1}^{m} b_{i,j} w_i$. Но так как $h_A = f$, то $h_A(v_j) = f(v_j)$, то

$$\sum_{i=1}^{m} a_{i,j} w_i = \sum_{i=1}^{m} b_{i,j} w_i.$$

А поскольку элементы $w_1, \ldots, w_m$ являются базисом в W, то отсюда следует, что $a_{i,j} = b_{i,j}$ для всех i, j. Таким образом, $B = A$. Это означает, при суперпозиции отображений $M_{m,n}(K) \to L \to M_{m,n}(K)$, $A \mapsto h_A = f \mapsto M_f$, матрица A переходит сама в себя, то есть эта суперпозиция — тождественное отображение. Это означает, что построенные нами отображения $\mathcal{L}(V, W) \to M_{m,n}(K)$, $M_{m,n}(K) \to \mathcal{L}(V, W)$ являются взаимно обратными биекциями. $\square$

Замечание. Легко заметить, что если V и W — не обязательно совпадающие векторные пространства, то при любом выборе базисов нулевому линейному отображению из V в W будет соответствовать нулевая матрица. Позднее, когда будет определена структура векторного пространства на множестве $\mathcal{L}(V, W)$, будет показано, что биекция, построенная в теореме 1.2.1, сама является линейным отображением.

Если рассматривается линейный оператор $f : V \to V$, то чаще всего для построения его матрицы используется не два различных базиса, а один и тот же, например, $v_1, \dots, v_n$. В этом случае матрица f определяется из равенств:

$$f(v_j) = \sum_{i=1}^{n} a_{i,j} v_i \tag{1.2.2}$$

Это равенство есть важный частный случай равенства (1.2.1).

Найдем матрицы линейных отображений из примеров 1.1.6, 1.1.7, 1.1.8.

Пример 1.2.1. Пусть $f : V \to V$ таково, что $f(v) = \alpha v$. Пусть $v_1, \dots, v_n$ — произвольный базис V. Тогда матрица f имеет следующий вид:

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

В самом деле, пусть $f(v_i) = \alpha v_i$ для всех i, $1 \leq i \leq n$. Эти равенства надо представлять в следующем виде:

$$f(v_i) = 0 \cdot v_1 + \dots + 0 \cdot v_{i-1} + \alpha \cdot v_i + 0 \cdot v_{i+1} + \dots + 0 \cdot v_n,$$

и тогда коэффициенты при векторах из базиса в правой части равенства образуют i-й столбец матрицы M_f. Но очевидно, что это в точности i-й столбец $n \times n$-матрицы, показанной выше.

В частности, полагая $\alpha = 1$, делаем вывод, что матрицей тождественного (единичного) линейного отображения в любом базисе является единичная матрица.

В случае, если пространство V одномерно, матрица линейного отображения f имеет размер 1×1 и поэтому фактически совпадает с элементом α.

Пример 1.2.2. Рассмотрим линейное отображение из примера 1.1.7. В качества базиса $\mathbb{C}$ над $\mathbb{R}$ возьмем $1, i$. Если $f(z) = \alpha z$, и $\alpha = \cos\varphi + i\sin\varphi$, то матрица f такова:

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

Это следует из равенств:

$$f(1) = \alpha \cdot 1 = \ \cos\varphi \cdot 1 + \sin\varphi \cdot i,$$
$$f(i) = \alpha \cdot i = -\sin\varphi \cdot 1 + \cos\varphi \cdot i,$$

и определения матрицы линейного отображения. Коэффициенты при базисных векторах $1, i$ в правых частях выписанных выше равенств должны располагаться в столбцах матрицы данного линейного отображения.

Пример 1.2.3. Если в примере 1.1.8 в качестве базисов K^n и K^m выбрать стандартные базисы (см. пример 1.1.4), то матрицей линейного отображения f, действующего из K^n в K^m по правилу $f(x) = Ax$, будет сама матрица A. В самом деле, пусть $a_{i,j}$ — i, j-й элемент A для всех i, j, и пусть для каждого i через e_i обозначен столбец

$$\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

где единица располагается в i-й строке, а во всех остальных строках стоят нули. Тогда легко проверить, что Ae_j — это j-й столбец матрицы A, и он записывается в виде

$$Ae_j = a_{1,j} \cdot e_1 + a_{2,j} \cdot e_2 + \cdots + a_{i,j} \cdot e_i + \cdots + a_{n,j} \cdot e_n.$$

Коэффициенты при векторах базиса $e_1, \dots, e_n$ в правой части этого равенства образуют j-й столбец матрицы рассматриваемого линейного отображения. Но очевидно, что это в точности j-й столбец самой матрицы A. Это и означает, что матрицей линейного отображения является матрица A.

Пример 1.2.4. Рассмотрим векторное пространство $K[x]_n$, состоящее из всех многочленов над полем K ($K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$), имеющих степень не более n. Базисом этого пространства являются многочлены $1, x, x^2, \dots, x^{n-1}$. Отображение $\psi : K[x]_n \to K[x]_n$, сопоставляющее многочлену $f = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1}$ его производную $\psi(f) = a_1 + 2a_2x \dots + (n-1)a_{n-1}x^{n-2}$, является линейным оператором, имеющим в указанном выше базисе следующую матрицу:

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

Однако в пространстве $K[x]_n$ существуют и другие базисы. Одним из таких базисов является множество: $1, x/1!, x^2/2!, \dots, x^{n-1}/(n-1)!$. В этом базисе у того же отображения ψ матрица будет такой:

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

Напомним некоторые факты из материала первого семестра. Пусть $v_1, \dots, v_n$ и $v'_1, \dots, v'_n$ — два базиса векторного пространства V (заметим, что здесь нумерация векторов базиса имеет существенное значение). Тогда для каждого i имеют место равенства:

$$v'_i = \sum_{j=1}^{n} b_{j,i} v_j,$$

где элементы $b_{j,i} \in K$ определены однозначно. Матрица $B = (b_{j,i})$ называется *матрицей перехода* от базиса $v_1, \dots, v_n$ к базису $v'_1, \dots, v'_n$. Матрица B невырождена, и $C = B^{-1}$ есть матрица прехода от базиса $v'_1, \dots, v'_n$ к базису $v_1, \dots, v_n$.

Пусть $v = \sum\limits_{i=1}^{n} x_i v_i$, $v = \sum\limits_{i=1}^{n} x'_i v'_i$, и

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

Тогда $x = Bx'$.

Теорема 1.2.2. *Пусть $f : V \to V$ — линейный оператор, $v_1, \dots, v_n$ — базис V, и $A = M_f$ — матрица f в этом базисе, т.е. $f(v_j) = \sum\limits_{i=1}^{n} a_{i,j} v_i$. Если v и x имеют тот же смысл, что и выше, $f(v) = \sum\limits_{i=1}^{n} y_i v_i$, $y_i \in K$, $y = (y_1, \dots, y_n)^{\mathrm{T}}$, то*

$$y = Ax \tag{1.2.3}$$

Рассмотрим другой базис $v'_1, \dots, v'_n$ пространства V, и пусть B есть матрица перехода от $v_1, \dots, v_n$ к $v'_1, \dots, v'_n$. Тогда матрицей f в базисе $v'_1, \dots, v'_n$ будет матрица

$$A' = B^{-1} A B \tag{1.2.4}$$

Напомним, что матрицы A и $B^{-1}AB$ называются *подобными*.

Доказательство. Применим оператор f к вектору $v = \sum\limits_{j=1}^{n} x_j v_j$:

$$f(v) = f(\sum_{j=1}^{n} x_j v_j) = \sum_{j=1}^{n} x_j f(v_j).$$

Но $f(v_j) = \sum\limits_{i=1}^{n} a_{i,j}v_i$ для всех j. Делая подстановку, и группируя слагаемые, получаем:

$$f(v) = \sum_{j=1}^{n} x_j(\sum_{i=1}^{n} a_{i,j}v_i) = \sum_{j=1}^{n}\sum_{i=1}^{n} x_j a_{i,j} v_i = \sum_{i=1}^{n}(\sum_{j=1}^{n} a_{i,j}x_j)v_i$$

Сравнивая правую часть полученного равенства с выражением $f(v) = \sum\limits_{i=1}^{n} y_i v_i$, и приравнивая кожффициенты при всех элементах базиса v_i (так как коэффициенты при базисных векторах в записи данного вектора $f(v)$ должны быть определены однозначно, то это законная операция) получим для каждого i равенства:

$$y_i = \sum_{j=1}^{n} a_{i,j}x_j.$$

Формула (1.2.3) представляет собой матричный способ записи этих полученных равенств.

Пусть теперь $v_1, \ldots, v_n$ и $v'_1, \ldots, v'_n$ — два базиса в пространстве V, $v'_j = \sum\limits_{i=1}^{n} b_{i,j}v_i$, $v_k = \sum\limits_{l=1}^{n} b'_{l,k}v'_l$ для всех j, k, $1 \leq j, k \leq n$. Элементы $b_{i,j}$ образуют матрицу B, элементы $b'_{l,k}$ образуют матрицу $B' = B^{-1}$.

Рассмотрим равенства $f(v'_j) = \sum\limits_{l=1}^{n} a'_{l,j}v'_l$, и вычислим $f(v'_j)$ еще одним способом. Для этого применим оператор f к равенству $v'_j = \sum\limits_{i=1}^{n} b_{i,j}v_i$:

$$f(v'_j) = \sum_{i=1}^{n} b_{i,j}f(v_i).$$

Подставляя сюда выражения $f(v_i) = \sum\limits_{k=1}^{n} a_{k,i}v_k$, получаем:

$$f(v'_j) = \sum_{i=1}^{n} b_{i,j}\sum_{k=1}^{n} a_{k,i}v_k = \sum_{k=1}^{n}(\sum_{i=1}^{n} a_{k,i}b_{i,j})v_k.$$

Наконец, подставим вместо v_k выражение $\sum\limits_{l=1}^{n} b'_{l,k} v'_l$. Получим

$$\begin{aligned} f(v'_j) = & \sum_{k=1}^{n} (\sum_{i=1}^{n} a_{k,i} b_{i,j})(\sum_{l=1}^{n} b'_{l,k} v'_l) = \\ & \sum_{k=1}^{n} \sum_{i=1}^{n} \sum_{l=1}^{n} a_{k,i} b_{i,j} b'_{l,k} v'_l = \\ & \sum_{l=1}^{n} (\sum_{k=1}^{n} \sum_{i=1}^{n} b'_{l,k} a_{k,i} b_{i,j}) v'_l. \end{aligned}$$

Таким образом,

$$f(v'_j) = \sum_{l=1}^{n} a'_{l,j} v'_l = \sum_{l=1}^{n} (\sum_{k=1}^{n} \sum_{i=1}^{n} b'_{l,k} a_{k,i} b_{i,j}) v'_l.$$

Приравнивая коэффициенты при каждом базисном векторе v'_l, получаем набор равенств (для всех возможных l, j, $1 \leq l, j \leq n$):

$$a'_{l,j} = \sum_{k=1}^{n} \sum_{i=1}^{n} b'_{l,k} a_{k,i} b_{i,j}.$$

Вспоминая, как умножаются матрицы, видим, что в матричной форме все эти равенства записываются так:

$$A' = B'AB = B^{-1}AB.$$

Это и есть искомое равенство (1.2.4). □

Теорема 1.2.3. *Пусть даны два линейных отображения $f : V \to W$, $g : W \to U$, и базисы: $v_1, \dots, v_n$ в V, $w_1, \dots, w_m$ в W, $u_1, \dots, u_k$ в U. Суперпозиция линейных отображений gf также является линейным отображением, и если M_f, M_g — матрицы f и g в указанных базисах, то $M_{gf} = M_g M_f$.*

Доказательство. Пусть $f(v_j) = \sum\limits_{i=1}^{m} a_{i,j} w_i$, $g(w_i) = \sum\limits_{l=1}^{k} b_{l,i} u_l$. Таким образом, $M_f = A$ есть матрица с компонентами $a_{i,j}$, $M_g = B$ — матрица с компонентами $b_{l,i}$. Чтобы вычислить матрицу M_{gf}, надо вычислить

$gf(v_j)$, и записать эти векторы через базис $u_1, \dots, u_k$. Приступим к вычислениям.

$$gf(v_j) = g(f(v_j)) = g(\sum_{i=1}^{m} a_{i,j}w_i) = \sum_{i=1}^{m} a_{i,j}g(w_i).$$

Теперь подставим вместо каждого $g(w_i)$ выражение $\sum\limits_{l=1}^{k} b_{l,i}u_l$. В результате получим:

$$gf(v_j) = \sum_{i=1}^{m} a_{i,j}(\sum_{l=1}^{k} b_{l,i}u_l) = \sum_{i=1}^{m}\sum_{l=1}^{k} a_{i,j}b_{l,i}u_l = \sum_{l=1}^{k}(\sum_{i=1}^{m} b_{l,i}a_{i,j})u_l.$$

Коэффициенты при базисных векторах u_l обязательно должны быть компонентами матрицы линейного отображения gf (виду единственности представления векторов в виде линейной комбинации элементов базиса). Поэтому l, j- я компонента матрицы M_{gf} равна

$$\sum_{i=1}^{m} b_{l,i}a_{i,j}.$$

Очевидно, что это выражение является также l, j-й компонентой матрицы BA. Таким образом, $M_{gf} = BA = M_gM_f$. □

1.3. Изоморфизмы. Ядро и образ

Лемма 1.3.1. *Допустим, что линейное отображение $f : V \to W$ биективно. Тогда обратное к f отображение $g = f^{-1}$ также является линейным.*

Доказательство. Пусть $w_1, w_2 \in W$, $\alpha_1, \alpha_2 \in K$. Необходимо показать, что $g(\alpha_1w_1 + \alpha_2w_2) = \alpha_1g(w_1) + \alpha_2g(w_2)$.

Так как отображение f биективно, то существуют $v_1, v_2 \in V$ такие, что $w_1 = f(v_1)$ и $w_2 = f(v_2)$. Тогда

$$\alpha_1w_1 + \alpha_2w_2 = \alpha_1f(v_1) + \alpha_2f(v_2) = f(\alpha_1v_1 + \alpha_2v_2).$$

Последнее равенство выполняется ввиду линейности f. Так как $g(f(v)) = v$ для каждого $v \in V$, то

$$g(\alpha_1 w_1 + \alpha_2 w_2) = g(f(\alpha_1 v_1 + \alpha_2 v_2)) = \alpha_1 v_1 + \alpha_2 v_2.$$

Применяя к левой и правой частям равенств $w_i = f(v_i)$ отображение g, получаем $g(w_i) = g(f(v_i)) = v_i$, $i = 1, 2$. Таким образом, $\alpha_1 v_1 + \alpha_2 v_2 = \alpha_1 g(w_1) + \alpha_2 g(w_2)$, что и доказывает лемму. □

Определение 1.3.1. Биективное линейное отображение $f : V \to W$ называется *изоморфизмом* между V и W. Если между V и W существует изоморфизм, то говорят также, что пространства V и W изоморфны. Это обозначается следующим образом: $V \cong W$.

Теорема 1.3.1. 1) *Линейное отображение $f : V \to W$ является биективным тогда и только тогда, когда для любого базиса $v_1, \dots, v_n$ пространства V элементы $f(v_1), \dots, f(v_n)$ являются базисом W.*

2) *Векторные пространства V и W изоморфны тогда и только тогда, если $\dim(V) = \dim(W)$.*

Доказательство. Докажем пункт 1). Допустим, что линейное отображение f биективно. Тогда, как уже показано выше, обратное к нему отображение g линейно. Пусть $v_1, \dots, v_n$ — базис V. Ввиду биективности f каждый вектор $w \in W$ можно представить в виде $w = f(v)$ для некоторого $v \in V$. Вектор v запишем через выбранный базис: $v = \sum\limits_{i=1}^{n} \alpha_i v_i$. Тогда

$$w = f(v) = f(\sum_{i=1}^{n} \alpha_i v_i) = \sum_{i=1}^{n} \alpha_i f(v_i).$$

Теперь, чтобы убедиться в том, что векторы $f(v_1), \dots, f(v_n)$ образуют базис пространства W, осталось доказать, что эти векторы линейно независимы. Рассмотрим линейную комбинацию $\sum\limits_{i=1}^{n} \alpha_i f(v_i) = 0$. Используя линейность отображения f, получаем отсюда:

$$0 = \sum_{i=1}^{n} \alpha_i f(v_i) = f(\sum_{i=1}^{n} \alpha_i v_i).$$

Применим к левой и правой частям полученного равенства линейное отображение g, и воспользуемся тем, что $g(0) = 0$ и $g(f(v)) = v$ для любого $v \in V$:

$$0 = g(0) = g(f(\sum_{i=1}^{n} \alpha_i v_i)) = \sum_{i=1}^{n} \alpha_i v_i.$$

Получена равная нулю линейная комбинация векторов $v_1, \dots, v_n$, которые линейно независимы по определению базиса. Это означает, что $\alpha_i = 0$ для всех индексов i, $1 \leq i \leq n$. Но тем самым показано, что произвольная линейная комбинация векторов $f(v_1), \dots, f(v_n)$, которая равна нулю, является тривиальной. То есть векторы $f(v_1), \dots, f(v_n)$ линейно независимы, и поэтому являются базисом пространства W.

Обратно, пусть для базиса $v_1, \dots, v_n$ пространства V векторы $f(v_1), \dots, f(v_n)$ являются базисом пространства W. Обозначим для каждого i через w_i базисный вектор $f(v_i)$. Тогда, согласно теореме 1.1.1, существует (притом однозначно определенное) линейное отображение $g : W \to V$, такое, что $g(w_i) = v_i$ для всех индексов i. Покажем, что f и g — взаимно-обратные отображения. Прежде всего, по самому построению, $g(w_i) = g(f(v_i)) = v_i$ и $f(g(w_i)) = f(v_i) = w_i$ для всех i. Рассмотрим тождественное отображение $1_V : V \to V$, которое переводит каждый вектор $v \in V$ в самого себя. Выше уже было показано, что это линейное отображение. Но так как $1_V(v_i) = v_i = g(f(v_i))$ для всех i, то по первой части теоремы 1.1.1 получаем $gf = 1_V$. Точно так же рассмотрим тождественное отображение $1_W : W \to W$, для которого

$1_W(w_i) = w_i = f(g(w_i))$, откуда следует, что $fg = 1_W$. Это означает, что f и g — взаимно обратные линейные отображения.

Докажем пункт 2). Если существует изоморфизм между V и W, то есть биективное линейное отображение $f : V \to W$ то, согласно пункту 1), базис $v_1, \dots, v_n$ отображается в базис $f(v_1), \dots, f(v_n)$. Это означает, что $n = \dim(V) = \dim(W)$.

Обратно, пусть $\dim(V) = \dim(W) = n$. Выберем базис $v_1, \dots, v_n$ в пространстве V, и базис $w_1, \dots, w_n$ пространства W. Тогда, по теореме 1.1.1, существуют линейное отображение $f : V \to W$ такое, что $f(v_i) = w_i$, и линейное отображение $g : W \to V$ такое, что $g(w_i) = v_i$ для всех индексов i. Таким образом, $f(g(w_i)) = w_i$ и $g(f(v_i)) = v_i$. Теперь, рассуждая так же, как при доказательстве пункта 1), выводим отсюда, что f и g являются взаимно обратными биективными отображениями. То есть пространства V и W изоморфны. $\square$

Теорема 1.3.2. *Линейное отображение $f : V \to W$ является биективным тогда и только тогда, когда его матрица M_f является квадратной и невырожденной (обратимой). При этом, если $g = f^{-1}$, то $M_g = M_f^{-1}$. Данное свойство не зависит от выбора базисов, в которых вычисляется матрица M_f.*

Доказательство. Если f — биекция, то $n = \dim(V) = \dim(W) = m$, и это значит, что $m \times n$-матрица M_f является квадратной при любом выборе базисов в V и W. Выберем какие угодно базисы в V и в W, и будем считать, что матрица M_f вычислена с помощью этих базисов. Если $g = f^{-1}$, то, как уже известно, g также является линейным отображением. По определению g, имеют место равенства $fg = 1_W$ и $gf = 1_V$. Применим к ним формулу из теоремы 1.2.3, предполагая, что матрица M_g вычислена с помощью уже выбранных базисов. Поскольку матрицы тождественных отображений при любом выборе базиса являются единичными, получаем отсюда, что

$$M_{fg} = M_f M_g = E_n, \quad M_{gf} = M_g M_f = E_n.$$

Обратно, пусть дано линейное отображение $f : V \to W$, и его матрица $A = M_f$, вычисленная с помощью базиса $v_1, \dots, v_n$ пространства V и базиса $w_1, \dots, w_m$ пространства W, является квадратной (то есть $n = m$) и невырожденной (то есть существует $B = A^{-1}$). По теореме 1.2.1 найдется линейное отображение $g : W \to V$, матрицей M_g которого в выбранных базисах будет матрица B. При этом такое линейное отображение g определено однозначно. Снова применяя теорему 1.2.3, получаем равенства:

$$M_f M_g = AB = E_n = M_{fg}, \quad M_g M_f = BA = E_n = M_{gf}.$$

Но ввиду примера 1.2.1, единичной матрице при выбранном базисе пространства может соответствовать только тождественное отображение. Таким образом,

$$fg = 1_W, \quad gf = 1_V,$$

и это означает, что f есть биекция, то есть изоморфизм. □

Следующий пример является *крайне* важным.

Пример 1.3.1. Если $\dim(V) = n$, то $V \cong K^n$. Если $v_1, \dots, v_n$ — некоторый упорядочнный базис V, то изоморфизм строится следующим образом. Вектору $v = \sum\limits_{n=1}^{n} x_i v_i$, где $x_i \in K$, сопоставляется столбец $f(v) = (x_1, \dots, x_n)^{\mathrm{T}}$. Изоморфизмы между V и K^n взаимно-однозначно соответствуют (упорядоченным) базисам V.

Определение 1.3.2. Пусть $f : V \to W$ — линейное отображение. *Ядром* f называется множество $\mathrm{Ker}(f) = \{v | v \in V, f(v) = 0\}$. *Образом* f называется множество $\mathrm{Im}(f) = \{w | w = f(v), v \in V\}$.

Лемма 1.3.2. $\mathrm{Ker}(f)$ — *векторное подпространство пространства* V, $\mathrm{Im}(f)$ — *векторное подпространство пространства* W.

Доказательство. Пусть $v_1, v_2 \in \mathrm{Ker}(f)$, и $\alpha_1, \alpha_2 \in K$. Покажем, что $\alpha_1 v_1 + \alpha_2 v_2 \in \mathrm{Ker}(f)$. Во-первых, заметим, что условия $v_i \in \mathrm{Ker}(f)$ $(i = 1, 2)$ означают, что $f(v_i) = 0$. Отсюда следует равенство:

$$f(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 f(v_1) + \alpha_2 f(v_2) = 0.$$

Но по определению ядра это и означает, что $\alpha_1 v_1 + \alpha_2 v_2 \in \mathrm{Ker}(f)$.

В случае $\mathrm{Im}(f)$ надо также взять $w_1, w_2 \in \mathrm{Im}(f)$, $\alpha_1, \alpha_2 \in K$, и показать, что $\alpha_1 w_1 + \alpha_2 w_2 \in \mathrm{Im}(f)$. По определению $\mathrm{Im}(f)$, если $w_i \in \mathrm{Im}(f)$ $(i = 1, 2)$, то это означает, что $w_i = f(v_i)$ для некоторых $v_i \in V$. Тогда

$$\alpha_1 w_1 + \alpha_2 w_2 = \alpha_1 f(v_1) + \alpha_2 f(v_2) = f(\alpha_1 v_1 + \alpha_2 v_2).$$

Вектор $f(\alpha_1 v_1 + \alpha_2 v_2)$ принадлежит множеству $\mathrm{Im}(f)$ по определению.

$\square$

Теорема 1.3.3. *Линейное отображение f инъективно тогда и только тогда, если* $\mathrm{Ker}(f) = \{0\}$.

Доказательство. Инъективность любого отображения f означает по определению, что из $f(v_1) = f(v_2)$ всегда следует $v_1 = v_2$. Если f инъективно и линейно, то для $v \in \mathrm{Ker}(f)$ будем иметь $f(v) = 0 = f(0)$. Отсюда ввиду инъективности следует, что $v = 0$, то есть $\mathrm{Ker}(f)$ состоит из одного нулевого вектора.

Обратно, пусть f — линейное отображение и $\mathrm{Ker}(f) = \{0\}$. Если $f(v_1) = f(v_2)$, то $f(v_1 - v_2) = f(v_1) - f(v_2) = 0$, и поэтому $v_1 - v_2 \in \mathrm{Ker}(f)$. Однако единственным векторм, содержащимся в ядре, является нулевой вектор. Следовательно, $v_1 - v_2 = 0$, $v_1 = v_2$, и это означает, что отображение f биективно. $\square$

Теорема 1.3.4. *Пусть* $f : V \to W$ *— линейное отображение, и* M_f *— его матрица (вычисленная относительно произвольных базисов* V *и* W*). Тогда* $\dim(\mathrm{Im}(f)) = \mathrm{rk}(M_f)$.

Доказательство. Разобъем доказательство на следующие этапы (некоторые из них имеют самостоятельное значение и будут использоваться в дальнейшем).

1. Пусть $v_1, \dots, v_n$ — базис V, $w_1, \dots, w_m$ — базис W, и $f(v_i) = \sum\limits_{j=1}^{m} a_{j,i} w_j$. Таким образом, матрица A, составленная из элементов $a_{j,i}$ есть матрица M_f. Рассмотрим линейное отображение $g : K^n \to K^m$, действующее по правилу $g(x) = Ax$, и пусть $\varphi : V \to K^n$ и $\psi : W \to K^m$ — изоморфизмы, построенные по базисам $v_1, \dots, v_n$ и $w_1, \dots, w_m$ так же, как в примере 1.3.1.

Все это можно наглядно представить в виде следующего рисунка (такие рисунки называются *диаграммами*):

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle\varphi}\big\downarrow & & {\scriptstyle\psi}\big\downarrow \\ K^n & \xrightarrow{g} & K^m \end{array} \tag{1.3.1}$$

Утверждение, о котором идет речь, заключается в том, что $\psi f = g\varphi$ (в этом случае говорят, что диаграмма *коммутативна*). В самом деле, из теоремы 1.2.3 (точнее, из равенства (1.2.4)) следует, что если $v = \sum\limits_{i=1}^{n} x_i v_i \in V$, $x = (x_1, \dots, x_n)^{\mathrm{T}} \in K^n$, $f(v) = \sum\limits_{j=1}^{m} y_j w_j \in W$, $y = (y_1, \dots, y_m)^{\mathrm{T}} \in K^m$, то $y = Ax$. Но по построению $x = \varphi(v)$, $y = \psi(f(v))$, и равенство $y = Ax = g(x)$ записывается в виде $\psi(f(v)) = g(\varphi(v))$. А это и есть коммутативность диаграммы (1.3.1).

2. Из (1.3.1) следует, что ограничение ψ на подпространство $\mathrm{Im}(f)$ является изоморфизмом между подпространствами $\mathrm{Im}(f)$ и $\mathrm{Im}(g)$.

В самом деле, ограничение линейного отображения на подпространство есть линейное отображение, областью определения которого является данное подпространство (это следует из самого определения линейного отображения, так как линейная комбинация векторов некоторого подпространства принадлежит тому же подпространству). Далее, ограничение инъективного отображения, очевидно, также инъективно. Теперь, если $w \in \mathrm{Im}(f)$, то $w = f(v)$ для некоторого $v \in V$, то $\psi(w) = \psi(f(v)) = g(\varphi(v)) \in \mathrm{Im}(g)$. Таким образом, ограничение ψ на подпространство $\mathrm{Im}(f)$ можно рассматривать как линейное отображение из $\mathrm{Im}(f)$ в $\mathrm{Im}(g)$. Чтобы установить его сюръективность, выберем произвольный вектор $y \in \mathrm{Im}(g) \subseteq K^m$. Так как ψ является изоморфизмом, то найдется вектор $w \in W$ такой, что $y = \psi(w)$. Но из того, что $y \in \mathrm{Im}(g)$ следует, что $y = g(x)$, $x \in K^n$, а так как φ является изоморфизмом, то $x = \varphi(v)$ для некоторого вектора $v \in V$. Таким образом, $y = g(\varphi(v)) = \psi(f(v)) = \psi(w)$. Так как ψ есть инъективное отображение, то $w = f(v) \in \mathrm{Im}(f)$. Но это и означает, что отображение ψ сюръективно. Итак, ψ есть линейное отображение из $\mathrm{Im}(f)$ в $\mathrm{Im}(g)$, инъективное и сюръективное. Следовательно, это изоморфизм.

Изоморфные векторные пространства имеют одинаковые размерности, $\dim(\mathrm{Im}(f)) = \dim(\mathrm{Im}(g))$. Таким образом, все свелось к доказательству равенства $\dim(\mathrm{Im}(g)) = \mathrm{rk}(A)$ (напомним, что $A = M_f$).

3. Пусть $A = (A_1, \dots, A_n)$, где A_i есть i-й столбец A. Тогда $\mathrm{Im}(g) = \langle A_1, \dots, A_n \rangle$.

В самом деле, пусть $e_i \in K^n$ есть столбец, в i-й строке которого располагается единица, а все остальные элементы равны нулю. Тогда $e_1, \dots, e_n$ является базисом пространства K^n, и легко заметить, что для каждого i имеется равенство $A_i = g(e_i) = Ae_i$. Для произвольного $x \in K^n$, $x = (x_1, \dots, x_n)^{\mathrm{T}}$, получаем $x = \sum\limits_{i=1}^{n} x_i e_i$. Отсюда

$$g(x) = Ax = \sum_{i=1}^{n} x_i g(e_i) = \sum_{i=1}^{n} x_i A_i \in \langle A_1, \dots, A_n \rangle.$$

Это означает, что $\mathrm{Im}(g) \subseteq \langle A_1, \dots , A_n\rangle$. Чтобы показать обратное включение, рассмотрим произвольную линейную комбинацию $y = \sum\limits_{i=1}^{n} \alpha_i A_i \in \langle A_1, \dots , A_n\rangle$. Положим $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in K^n$. Легко убедиться, что $y = g(x) \in \mathrm{Im}(g)$.

4. Если $U \subseteq W$ — подпространство W, и $U = \langle u_1, \dots , u_n\rangle$ (в этом случае говорят, что $u_1, \dots , u_n$ — множество *образующих*, или *порождающих* элементов подпространства U), то из $u_1, \dots , u_n$ можно выбрать базис U. Это — максимальное линейно независимое подмножество множества $u_1, \dots , u_n$.

В самом деле, пусть максимальное линейно независимое подмножество среди $u{-}1, \dots , u_n$ есть $u_1, \dots , u_k$. Покажем, что каждый вектор $u \in U$ есть линейная комбинация векторов $u_1, \dots , u_k$. По условию u есть линейная комбинация векторов $u_1, \dots , u_k, u_{k+1}, \dots , u_n$. Запишем это так:

$$u = \sum_{i=1}^{k} \alpha_i u_i + \sum_{j=k+1}^{n} \beta_j u_j. \tag{1.3.2}$$

Условие максимальности при выборе $u_1, \dots , u_k$ означает, что добавление к этому множеству любого вектора v_j, $j = k + 1, \dots , n$ делает множество векторов $u_1, \dots , u_k, u_j$ линейно зависимым. Это значит, что есть нетривиальная линейная комбинация вида

$$\sum_{i=1}^{k} \gamma_{i,j} u_i + \gamma_{j,j} u_j = 0. \tag{1.3.3}$$

Если бы в этой линейной комбинации было $\gamma_{j,j} = 0$, то мы получили бы линейную зависимость между векторами $u_1, \dots , u_k$ (так среди $\gamma_{i,j}$ по предположению обязательно есть хотя бы одно ненулевое число). Поскольку это невозможно, то $\gamma_{j,j} \neq 0$. Разделив обе части равенства (1.3.2) на $\gamma_{j,j}$, мы получим возможность выразить u_j в виде линейной комбинации векторов $u_1, \dots, u_k$:

$$u_j = \sum_{i=1}^{k} \xi_{i,j} u_i,$$

где $\xi_{i,j} = -\gamma_{i,j}/\gamma_{j,j}$. Подставим полученные выражения для u_j в равенство (1.3.2), и получим:

$$u = \sum_{i=1}^{k} \alpha_i u_i + \sum_{j=k+1}^{n} \beta_j \sum_{i=1}^{k} \xi_{i,j} u_i = \sum_{i=1}^{k} (\alpha_i + \sum_{j=k+1}^{n} \beta_j \xi_{i,j}) u_i$$

Таким образом, каждый вектор из U выражается в виде линейной комбинации векторов $u_1, \dots, u_k$. Так как эти векторы выбраны линейно независимыми, то они образуют базис подпространства U.

5. Теперь остается только вспомнить свойство, дающее выражение для ранга матрицы A. Ранг A равен максимальному числу линейно независимых столбцов, то есть максимальному линейно независимому подмножеству среди $A_1, \dots, A_n$. Согласно четвертому этапу доказательства, это число равно размерности подпространства, порожденного столбцами матрицы A. Согласно третьему этапу, эта размерность равна $\dim(\mathrm{Im}(g))$, а еще раньше было показано, что на самом деле это число равно $\dim(\mathrm{Im}(f))$. □

Теорема 1.3.5. *Пусть $f : V \to W$ — линейное отображение. Тогда*

$$\dim(V) = \dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)) \qquad (1.3.4)$$

Доказательство. Сначала выберем некоторый базис $w_1, \dots, w_k$ подпространства $\mathrm{Im}(f)$. По определению пространства $Im(f)$, можно найти векторы $v_1, \dots, v_k$ из V, такие, что $f(v_1) = w_1, \dots, f(v_k) = w_k$. Заметим, что ни один из векторов $v_1, \dots, v_k$ не принадлежит ядру f. Выберем теперь базис $v_{k+1}, \dots, v_n$ в подпространстве $\mathrm{Ker}(f)$. Утверждается, что векторы $v_1, \dots, v_n$ образуют базис пространства V. Выберем в V впроизвольный вектор v, и рассмотрим вектор $f(v)$. Этот вектор принадлежит подпространству $\mathrm{Im}(f)$, и поэтому его можно выразить через базис этого подпространства:

$$f(v) = \sum_{i=1}^{k} \alpha_i w_i.$$

Но так как $w_i = f(v_i)$ для всех i, то правую часть можно представить в виде:

$$\sum_{i=1}^{k} \alpha_i w_i = \sum_{i=1}^{k} \alpha_i f(v_i) = f(\sum_{i=1}^{k} \alpha_i v_i).$$

Отсюда следует, что

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

Это означает, что $v - \sum_{i=1}^{k} \alpha_i v_i \in \mathrm{Ker}(f)$. Таким образом, этот вектор можно выразить через выбранный базис подпространства $\mathrm{Ker}(f)$:

$$v - \sum_{i=1}^{k} \alpha_i v_i = \sum_{j=k+1}^{n} \alpha_i v_i.$$

Но тогда

$$v = \sum_{i=1}^{n} \alpha_i v_i.$$

Остается показать, что векторы $v_1, \ldots, v_n$ линейно независимы. Пусть

$$\sum_{i=1}^{n} \lambda_i v_i = 0.$$

Применим к левой и правой частям этого равенства линейное отображение f. С одной стороны, $f(0) = 0$. С другой стороны,

$$f(\sum_{i=1}^{n} \lambda_i v_i) = f(\sum_{i=1}^{k} \lambda_i v_i),$$

так как $\sum\limits_{i=k+1}^{n} \lambda_i v_i \in \mathrm{Ker}(f)$. Далее,

$$f(\sum_{i=1}^{k} \lambda_i v_i) = \sum_{i=1}^{k} \lambda_i f(v_i) = \sum_{i=1}^{k} \lambda_i w_i.$$

Так как векторы $w_1, \ldots, w_k$ линейно независимы, то $\lambda_1 = \cdots = \lambda_k = 0$. Таким образом, исходная линейная зависимость сводится к равенству:

$$\sum_{i=k+1}^{n} \lambda_i v_i = 0.$$

Но это — линейная зависимость между векторами базиса подпространства $\mathrm{Ker}(f)$, которые линейно независимы. Отсюда следует, что $\lambda_{k+1} = \cdots = \lambda_n = 0$, и это означает, что векторы $v_1, \ldots, v_n$ линйно независимы, и, таким образом, являются базисом пространства V.

Но по самому пострению этого базиса, $\dim(V) = n = k + (n - k)$, $k = \dim(\mathrm{Im}(f))$, $n - k = \dim(\mathrm{Ker}(f))$. Отсюда следует утверждение теоремы. $\square$

Теорема 1.3.6. *Пусть $f : V \to W$ — линейное отображение, M_f — его матрица (вычисленная относительно произвольных базисов V и W), $n = \dim(V)$, $m = \dim(W)$ (так что M_f является $m \times n$-матрицей). Отображение f инъективно тогда и только тогда, когда $n = \mathrm{rk}(M_f)$. Отображение f сюръективно тогда и только тогда, когда $m = \mathrm{rk}(M_f)$.*

Доказательство. Применим предыдущую теорему (равенство (1.3.4):

$$\dim(V) = \dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)).$$

Так как $\dim(V) = n$, а $\dim(\mathrm{Im}(f)) = \mathrm{rk}(M_f)$, получаем равенство:

$$n = \dim(\mathrm{Ker}(f)) + \mathrm{rk}(M_f).$$

Мы уже знаем, что инъективность f равносильна равенству $\mathrm{Ker}(f) = \{0\}$, что эквивалентно равенству $\dim(\mathrm{Ker}(f)) = 0$. Поэтому из $\dim(\mathrm{Ker}(f)) = 0$ следует $n = \mathrm{rk}(M_f)$, и наоборот, из $n = \mathrm{rk}(M_f)$ следует $\dim(\mathrm{Ker}(f)) = 0$, то есть инъективность f.

Сюръективность f означает, что $\mathrm{Im}(f) = W$, что равносильно равенству $\dim(\mathrm{Im}(f)) = \dim(W) = m$, или $\mathrm{rk}(M_f) = m$. □

Теорема 1.3.7. *Пусть $f : V \to V$ — линейный оператор, $\dim(V) = n < \infty$. Для того, чтобы оператор f был биективным (т.е. изоморфизмом), необходимо и достаточно, чтобы $\mathrm{Ker}(f) = \{0\}$.*

Доказательство. Если оператор f биективен, то он и инъективен. Это влечет равенство нулю ядра оператора.

Обратно, пусть $\mathrm{Ker}(f) = \{0\}$. Это значит, что отображение f инъективно. Это значит также, что $\dim(\mathrm{Ker}(f)) = 0$. Из равенства

$$\dim(V) = \dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f))$$

теперь следует, что $\dim(\mathrm{Im}(f)) = \dim(V)$. Но так как $\mathrm{Im}(f) \subseteq V$, то из равенства размерностей следует $\mathrm{Im}(V) = V$. Таким образом, отображение f является также и сюръективным. □

1.4. Собственные векторы и собственные значения

Определение 1.4.1. Рассмотрим линейный оператор $f : V \to V$. Элемент $\lambda \in K$ называется *собственным значением* линейного оператора f, если существует $v \in V$, $v \neq 0$, такой, что $f(v) = \lambda v$. В этом случае вектор v называется *собственным вектором* оператора f, отвечающим (или соответствующим) собственному значению λ.

В некоторых книгах собственные значения линейного оператора называются также его *характеристическими числами.*

Отметим, что условие $v \neq 0$ является совершенно обязательным.

Пример 1.4.1. Элемент $0 \in K$ является собственным значением f тогда и только тогда, если $\mathrm{Ker}(f) \neq \{0\}$. В этом случае ненулевые векторы из $\mathrm{Ker}(f)$ — это в точности все собственные векторы f, отвечающие собственному значению 0.

Пусть $f : V \to V$ — некоторый линейный оператор. Для произвольного $\lambda \in K$ обозначим через V^λ множество *всех* векторов из V, таких, что $f(v) = \lambda v$. Ясно, что $V^\lambda \neq 0$ в том и только в том случае, если λ есть собственное значение f. В этом случае V^λ называется *пространством (всех) собственных векторов, отвечающих собственному значению* λ.

Так как $f(0) = 0 = \lambda \cdot 0$, то нулевой вектор принадлежит множеству V^λ. Собственные векторы (отвечающие собственному значению λ) — это *ненулевые* векторы из V^λ. Если таких векторов нет, то λ не является собственным значением оператора f.

Лемма 1.4.1. *V^λ является векторным подпространством пространства V (для любого $\lambda \in K$).*

Доказательство. Пусть $v_1, v_2 \in V^\lambda$, $\alpha - 1, \alpha_2 \in K$. Покажем, что $v = \alpha_1 v_1 + \alpha_2 v_2 \in V^\lambda$. В самом деле,

$$f(v) = f(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 f(v_1) + \alpha_2 f(v_2) =$$
$$\alpha_1(\lambda v_1) + \alpha_2(\lambda v_2) = \lambda(\alpha_1 v_1 + \alpha_2 v_2) = \lambda v.$$

□

Если A есть матрица оператора f в некотором базисе $v_1, \ldots, v_n$, $v = \sum\limits_{i=1}^{n} x_i v_i$, и $x = (x_1, \ldots, x_n)^{\mathrm{T}}$, то $x \neq 0$ и $Ax = \lambda x$. Обратно, если для некоторого ненулевого столбца x имеет место равенство $Ax = \lambda x$, то $f(v) = \lambda v$ для $v = \sum\limits_{i=1}^{n} x_i v_i$. Таким образом, можно без потери общности говорить о *собственных значениях матрицы* A: это такие $\lambda \in K$, для которых существует ненулевые столбцы x со свойством $Ax = \lambda x$. Собственные значения матрицы A оказываются собственными значениями оператора f такого, что $A = M_f$, и наоборот.

Лемма 1.4.2. *Пусть A есть некоторая $n \times n$-матрица над полем K, E — единичная $n \times n$-матрица. Элемент $\lambda \in K$ будет собственным значением A тогда и только тогда, если матрица $A - \lambda E$ является вырожденной.*

Доказательство. Утверждение следует из того, что условие $Ax = \lambda x$ можно переписать в виде $(A - \lambda E)x = 0$. Таким образом, собственные векторы (столбцы), отвечающие собственному значению λ матрицы A вместе с нулевым столбцом образуют ядро оператора $f : K^n \to K^n$, действующего по правилу $f(x) = (A - \lambda E)x$. Ненулевые собственнные векторы существуют, таким образом, в том и только в том случае, когда оператор f не является инъективным, а это возможно тогда и только тогда, если $\mathrm{rk}(A - \lambda E) < n$, т.е. если матрица $A - \lambda E$ является вырожденной. $\square$

Замечание. В дальнейшем, когда будут определены линейные комбинации произвольных линейных операторов (а не только матриц), станет понятно, что подпространство V^λ является ядром оператора $f - \lambda e$, где через e обозначен тождественный оператор (его матрица в любом базисе является единичной).

Определение 1.4.2. Пусть A есть некоторая $n \times n$-матрица с компонентами из K. Рассмотрим многочлен $\chi_A(x)$, равный определителю матрицы $A - xE_n$. Этот многочлен называется *характеристическим многочленом* матрицы A.

Более подробно, это определитель матрицы:

$$A - xE_n = \begin{pmatrix} a_{1,1} - x & a_{2,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} - x & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} - x \end{pmatrix}$$

Лемма 1.4.3. *Для любой невырожденной $n \times n$-матрицы B имеет место равенство:* $\chi_A(x) = \chi_{B^{-1}AB}(x)$.

Доказательство. Обозначим определитель матрицы C через $|C|$. Вспомним, что $|C_1C_2| = |C_1| \cdot |C_2|$, что $\lambda E = \lambda(B^{-1}B) = B^{-1}(\lambda E)B$, и что $|B^{-1}| = |B|^{-1}$. Поэтому

$$\begin{aligned} |B^{-1}AB - \lambda E| = \ & |B^{-1}AB - B^{-1}(\lambda E)B| = \\ & |B^{-1}(A - \lambda E)B| = \\ & |B^{-1}| \cdot |A - \lambda E| \cdot |B| = |A - \lambda E|. \end{aligned}$$

□

Определение 1.4.3. Рассмотрим линейное отображение $f : V \to V$, и пусть $v_1, \dots, v_n$ — некоторый базис V, а матрица $A = (a_{i,j})$ есть матрица f в этом базисе. Многочлен $\chi_A(x)$ называется *характеристическим многочленом* линейного оператора f, и обозначается через $\chi_f(x)$. Это определение корректно, так как по предыдущей лемме многочлен $\chi_A(x)$ не зависит от выбора базиса.

Замечание. Аналогичным образом можно определить *определитель* $det(f)$ произвольного линейного оператора f как определитель матрицы A этого оператора в произвольно выбранном базисе. При переходе к другому базису матрица f примет вид $B^{-1}AB$, и определитель не изменится. Кроме того, так как $tr(A) = tr(B^{-1}AB)$, то можно определить *след* $tr(f)$ линейного оператора f, полагая его равным следу матрицы

f в произвольном базисе. Как и выше, от выбора базиса результат не зависит. Заметим еще, что

$$\chi_f(x) = (-1)^n x^n + (-1)^{n-1} tr(f) x^{n-1} + \cdots + det(f).$$

Теорема 1.4.1. *Элемент $\lambda \in K$ является собственным значением линейного оператора f тогда и только тогда, если λ есть корень многочлена $\chi_f(x)$.*

Доказательство. Пусть $\lambda \in K$ есть собственное значсение f, отвечающее собственному вектору $v \in V$. Выберем какой-нибудь базис $v_1, \dots, v_n$ пространства V, и пусть $A = M_f$ — матрица f в этом базисе, $v = \sum\limits_{i=1}^{n} x_i v_i$, $x = (x_1, \dots, x_n)^{\mathrm{T}} \in K^n$. Так как $f(v) = \lambda v$, то по теореме 1.2.2 $Ax = \lambda x$, откуда следует, что $Ax - \lambda x = (A - \lambda E)x = 0$. Столбец x является ненулевым по условию (так как v — собственный вектор), и поэтому мы получаем систему линейных однородных уравнений $(A - \lambda E)x = 0$, имеющую ненулевое решение. Определитель такой системы должен равняться нулю, но этот определитель есть $|A - \lambda E| = \chi_f(\lambda)$. Таким образом, собственное значение оператора f является корнем его характеристического многочлена.

Обратно, пусть λ есть корень характеристического многочлена оператора f. Выберем снова какой-нибудь базис $v_1, \dots, v_n$ пространства V, и вычислим матрицу $A = M_f$ в этом базисе. Так как определитель $|A - \lambda E| = \chi_f(\lambda)$ по условию равняется нулю, то у системы линейных однородных уравнений $(A - \lambda E)x = 0$ должно существовать ненулевое решение $x = (x_1, \dots, x_n)^{\mathrm{T}}$. Равенство $(A - \lambda E)x = 0$ можно переписать в равносильной форме как $Ax = \lambda x$. Рассмотрим вектор $v = \sum\limits_{i=1}^{n} x_i v_i$. Так как не все его координаты равны нулю, то и он тоже ненулевой. Остается заметить, что $f(v) = \lambda v$. Это снова следует из теоремы 1.2.2, с учетом того, что при выбранном базисе соответствие между векторами из V и столбцами из их координат в данном базисе является взаимно-однозначным. Более подробно это можно показать так. Рассмотрим коммутативную диаграмму (1.3.1):

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ {\scriptstyle\varphi}\big\downarrow & & {\scriptstyle\psi}\big\downarrow \\ K^n & \xrightarrow{g} & K^m \end{array}$$

Здесь, как и выше, $g(x) = Ax$ и $\varphi(v) = x$. Тогда

$$\psi(f(v)) = g(\varphi(v)) = Ax = \lambda x = \lambda\psi(v) = \psi(\lambda v).$$

Но отображение ψ является биекцией. Отсюда следует, что $f(v) = \lambda v$. $\square$

Теорема 1.4.2. *Пусть $\lambda_1, \dots, \lambda_m$ — различные собственные значения оператора f (не обязательно все имеющиеся). Выберем ненулевые векторы $v_1 \in V^{\lambda_1}$, …, $v_m \in V^{\lambda_m}$. Тогда они линейно независимы.*

Доказательство. Проведем индукцию по m. При $m = 1$ утверждение очевидно: множество, состоящее из одного ненулевого вектора v_1 линейно независимо. Допустим, что утверждение теоремы справедливо для любых различных $m - 1$ собственных значений и соответствующих им ненулевых собственных векторов. Предположим, что в случае m значений имеется нетривиальная линейная зависимость следующего вида:

$$\alpha_1 v_1 + \dots + \alpha_{m-1} v_{m-1} + \alpha_m v_m = 0 \tag{1.4.1}$$

Если допустить, что $\alpha_m = 0$, то тогда какой-то из элементов поля $\alpha_1, \dots, \alpha_{m-1}$ должен быть ненулевым, так как по предположению линейная зависимость (1.4.1) нетривиальна. Но тогда получилась бы нетривиальная линейная зависимость

$$\alpha_1 v_1 + \dots + \alpha_{m-1} v_{m-1} = 0,$$

а это противоречит предположению индукции. Следовательно, $\alpha_m \neq 0$, и тогда на этот элемент можно разделить обе части равенства (1.4.1). Перенеся затем вектор v_m в правую часть полученного равенства, и изменив знаки, приходим к равенству вида:

$$\beta_1 v_1 + \cdots + \beta_{m-1} v_{m-1} = v_m \tag{1.4.2}$$

В этом равенстве не все $\beta_1, \ldots, \beta_m$ равны нулю, так как иначе было бы $0 = v_m$, что не так по условию. Подействуем на левую и правую части равенства (1.4.2) линейным оператором f. Так как $f(v_i) = \lambda_i v_i$ для всех i, то получается следующее равенство:

$$\beta_1 \lambda_1 v_1 + \cdots + \beta_{m-1} \lambda_{m-1} v_{m-1} = \lambda_m v_m \tag{1.4.3}$$

Если $\lambda_m = 0$, то все прочие λ_i должны быть ненулевыми, так как по условию все рассматриваемые собственные значения попарно различны. Поэтому получается нетривиальная линейная зависимость:

$$\beta_1 \lambda_1 v_1 + \cdots + \beta_{m-1} \lambda_{m-1} v_{m-1} = 0.$$

Нетривиальна она потому, что для какого-то j элемент поля β_j отличен от нуля, а потому и $\beta_j \lambda_j \neq 0$. Таким образом, снова получается противоречие с предположением индукции. Допустим, что $\lambda_m \neq 0$. Тогда левую и правую части равенства (1.4.3) можно разделить на λ_m. Выполнив это деление, получим равенство

$$\gamma_1 \lambda_1 v_1 + \cdots + \gamma_{m-1} \lambda_{m-1} v_{m-1} = v_m \tag{1.4.4}$$

Здесь $\gamma_i = \beta_i \lambda_m^{-1}$ для всех i, $1 \leq i \leq m-1$. Снова в левой части этого равенства найдется индекс j, для которого $\gamma_j \neq 0$ (потому что существует индекс j, для которого $\beta_j \neq 0$). Теперь, вычитая из равенства (1.4.4) равенство (1.4.2), получим

$$(\gamma_1 \lambda_1 - \beta_1) v_1 + \cdots + (\gamma_{m-1} \lambda_{m-1} - \beta_{m-1}) v_{m-1} = 0$$

Так как по предположению индукции векторы $v_1, \dots, v_{m-1}$ линейно независимы, то отсюда следует, что для всех i, $1 \le i \le m-1$ коэффициенты $\gamma_i \lambda_i - \beta_i$ должны быть равны нулю. Вспоминая, что $\gamma_i = \beta_i \lambda_m^{-1}$, приходим к равенствам

$$\beta_i \lambda_m^{-1} \lambda_i - \beta_i = 0,$$

или

$$\beta_i \lambda_i = \beta_i \lambda_m.$$

Выберем тот индекс j, для которого $\beta_j \neq 0$. Тогда соответствующее равенство можно разделить на β_j. В результате получим, что

$$\lambda_j = \lambda_m,$$

где $j < m$. Но это противоречит условиям теоремы. Следовательно, предположение о линейной зависимости векторов $v_1, \dots, v_m$ приводит к противоречию. $\square$

Определение 1.4.4. Пусть дан линейный оператор $f : V \to V$. Говорят, что он является *диагонализируемым*, если существует базис пространства V, состоящий из собственных векторов оператора f.

Матрица диагонализируемого оператора f в этом базисе имеет вид:

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

При этом $\chi_f(t) = (-1)^n (t - \lambda_1) \dots (t - \lambda_n)$, $\lambda_1, \dots, \lambda_n$ — все собственные значения f (с учетом кратностей).

Диагонализируемость оператора f равносильна тому, что для любой его матрицы A (матрицы в произвольно выбранном базисе) существует невырожденная матрица B такая, что $B^{-1}AB$ является диагональной матрицей.

Определение 1.4.5. Пусть дан линейный оператор $f : V \to V$. Множество всех его собственных значений называется *спектром* оператора f. Говорят, что оператор f имеет *простой спектр*, если количество различных собственных значений f равно размерности пространства V.

Теорема 1.4.3. *Линейный оператор, имеющий простой спектр, диагонализируем.*

Доказательство. Пусть $n = \dim(V)$, и $\lambda_1, \dots , \lambda_n$ — все различные собственные значения оператора $f : V \to V$. Тогда в каждом из подпространств V^{λ_j} найдется по ненулевому вектору v_j, $1 \leq j \leq n$ (если λ_j есть собственное значение f, то $V^{\lambda_j} \neq \{0\}$ по определению собственного вектора). Согласно предыдущей теореме множество векторов $v_1, \dots , v_n$ является линейно независимым. Но так как количество этих линейно независимых векторов равно размерности пространства V, то эти вектора образуют базис V, причем базис, состоящий из собственных векторов, что и требовалось доказать. □

Лемма 1.4.4. *Пусть $f : V \to V$ — линейный оператор, и $\lambda_1, \dots , \lambda_m$ — его различные собственные значения. Тогда сумма подпространств $V^{\lambda_1} + \dots + V^{\lambda_m}$ является прямой суммой.*

Доказательство. Пусть $v \in V^{\lambda_1}+\dots+V^{\lambda_m}$. Допустим, что имеется два представления v в виде:

$$v = v_1 + \dots + v_m = v'_1 + \dots + v'_m,$$

где $v_i, v'_i \in V^{\lambda_i}$ для каждого i. Допустим, что $v_{i_1} \neq v'_{i_1}, \dots , v_{i_k} \neq v'_{i_k}$, и $v_j = v'_j$ для прочих индексов j. Тогда

$$0 = v - v = (v_{i_1} - v'_{i_1}) + \dots + (v_{i_k} - v'_{i_k}).$$

Все разности $v_{i_j} - v'_{i_j}$ являются ненулевыми векторами из V^{λ_j}. Таким образом, получена нетривиальная линейная зависимость между ненулевыми собственными векторами оператора f, отвечающими различным собственным значениям. Это противоречит доказанному выше утверждению о линейной независимости таких векторов. $\square$

Теорема 1.4.4. *Пусть $f : V \to V$ — линейный оператор, и $\lambda_1, \ldots, \lambda_m$ — все различные собственные значения f. Эквивалентны следующие утверждения:*

1) *Оператор f диагонализируем;*

2) $V = V^{\lambda_1} \oplus \cdots \oplus V^{\lambda_m}$.

3) $\dim(V) = \dim(V^{\lambda_1}) + \cdots + \dim(V^{\lambda_m})$.

Доказательство. 1) $\Longrightarrow$ 2). Пусть в пространстве V существует базис $v_1, \ldots, v_n$, состоящий из собственных векторов оператора f. Можно считать, что векторы в этом базисе пронумерованы так, что $v_1, \ldots, v_{n_1}$ — собственные векторы, отвечающие собственному значению λ_1, $v_{n_1+1}, \ldots, v_{n_1+n_2}$ — собственные векторы, отвечающие собственному значению λ_2, и далее, $v_{n_1+\cdots+n_{i-1}+1}, \ldots, v_{n_1+\cdots+n_{i-1}+v_{n_i}}$ — собственные векторы, отвечающие собственному значению λ_i для каждого i, $1 \leq i \leq m$, так что $n_1 + \cdots + n_m = n = \dim(V)$. Отсюда следует, что $\langle v_{n_1+\cdots+n_{i-1}+1}, \ldots, v_{n_1+\cdots+n_{i-1}+v_{n_i}} \rangle \subseteq V^{\lambda_i}$. Отсюда следует, что $n_i \leq \dim(V^{\lambda_i})$. Но так как сумма подпространств $V^{\lambda_1} + \cdots + V^{\lambda_m}$ является прямой суммой, то

$$\dim(V^{\lambda_1}+\cdots+V^{\lambda_m}) = \dim(V^{\lambda_1})+\cdots+\dim(V^{\lambda_m}) \geq n_1+\cdots+n_m = n = \dim(V).$$

Однако $V^{\lambda_1} + \cdots + V^{\lambda_m}$ является подпространством пространства V, и поэтому его размерность не превосходит числа n. Следовательно, все отношения $\geq$ на самом деле являются равенствами, откуда следует, что

$n_i = \dim(V^{\lambda_i})$ для всех i (если $n_i < \dim(V^{\lambda_i})$ хотя бы для одного i, то $n \geq \dim(V^{\lambda_1}) + \cdots + \dim(V^{\lambda_m}) > n_1 + \cdots + n_m = n$.).

Равенства $n_i = \dim(V^{\lambda_i})$ означают, ввиду $\langle v_{n_1+\cdots+n_{i-1}+1}, \ldots , v_{n_1+\cdots+n_{i-1}+v_{n_i}} \rangle \subseteq V^{\lambda_i}$, что это включение является равенством и, таким образом, векторы $v_{n_1+\cdots+n_{i-1}+1}, \ldots , v_{n_1+\cdots+n_{i-1}+v_{n_i}}$ являются базисом подпространства V^{λ_i}. Так как $\bigcup\limits_{i=1}^{m} \{v_{n_1+\cdots+n_{i-1}+1}, \ldots , v_{n_1+\cdots+n_{i-1}+v_{n_i}}\} = \{v_1, \ldots , v_n\}$ есть базис V, то $V = V^{\lambda_1} \oplus \cdots \oplus V^{\lambda_m}$.

2) $\Longrightarrow$ 1) следует из известного свойства базиса прямой суммы (см. теорему 3.1.5), так как, выбрав базис в каждом из прямых слагаемых $V^{\lambda_1}, \ldots , V^{\lambda_m}$, мы получим базис пространства V, состоящий по построению из собственных векторов.

2) $\Longrightarrow$ 3). Если $V = V^{\lambda_1} \oplus \cdots \oplus V^{\lambda_m}$, то согласно известному свойству размерности прямой суммы $\dim(V) = \dim(V^{\lambda_1}) + \cdots + \dim(V^{\lambda_m})$.

3) $\Longrightarrow$ 2). По лемме 1.4.4 сумма подпространств $V' = V^{\lambda_1} + \cdots + V^{\lambda_m}$ является прямой суммой. Отсюда следует, что $\dim(V') = \dim(V^{\lambda_1}) + \cdots + \dim(V^{\lambda_m})$. Но тогда, согласно условию 3), $\dim(V) = \dim(V')$. Так как V' есть подпространство пространства V, отсюда следует, что $V = V' = V^{\lambda_1} \oplus \cdots \oplus V^{\lambda_m}$. $\square$

1.5. Инвариантные подпространства

Определение 1.5.1. Пусть дан линейный оператор $f : V \to V$. Подпространство U пространства V называется *инвариантным* относительно f (или *инвариантным подпространством оператора* f), если для каждого вектора $u \in U$ имеет место включение $f(u) \in U$.

Это свойство записывается также в виде $f(U) \subseteq U$. В этом случае $f|_U$ есть линейный оператор, действующий из U в U.

Пример 1.5.1. Тривиальными примерами инвариантных подпространств являются нулевое подпространство и все пространство V. $\mathrm{Ker}(f)$ и $\mathrm{Im}(f)$ являются инвариантными подпространствами f. Если λ — собственное значение f, то V^λ — инвариантное подпространство f. Сумма и пересечение инвариантных подпространств являются инвариантными подпространствами.

Если $f : V \to V$ — линейный оператор, и $W \subseteq V$ — его инвариантное подпространство, то ограничение оператора f на подпространство W (обозначение $f|_W$) можно рассматривать как линейное отображение (линейный оператор) из W в W.

Лемма 1.5.1. *Пусть $f : V \to V$ — линейный оператор, и $U = \langle v_1, \ldots, v_m \rangle$ — некоторое подпространство. Оно будет инвариантным относительно f тогда и только тогда, если для каждого i, $1 \leq i \leq m$, будет выполняться включение $f(v_i) \in U$.*

Доказательство. Если U инвариантно относительно f, то утверждение леммы очевидно. Обратно, пусть выполняется утверждение леммы. Рассмотрим произвольный вектор $u \in U$. Так как $U = \langle v_1, \ldots, v_m \rangle$, то $u = \sum\limits_{j=1}^{m} \alpha_j v_j$. Тогда $f(u) = \sum\limits_{j=1}^{m} \alpha_j f(v_j)$. Так как по условию $f(v_j) \in U$ для всех j, то и линейная комбинация векторов $f(v_j)$ принадлежит подпространству U. Отсюда следует, что $u \in U$. □

Теорема 1.5.1. *Пусть $f : V \to V$ — линейный оператор, $U \subseteq V$ — инвариантное относительно f подпространство, $v_1, \ldots, v_m$ — базис U. Дополним его до базиса всего пространства V векторами $v_{m+1}, \ldots, v_m$. Тогда матрица A оператора f в этом базисе будет иметь следующий блочный вид:*

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ 0 & A_{2,2} \end{pmatrix} \tag{1.5.1}$$

При этом блок $A_{1,1}$ *есть* $m \times m$*-матрица, и является матрицей ограничения* f *на подпространство* U *в базисе* $v_1, \dots, v_m$.

Обратно, если матрица оператора f *в некотором базисе* $v_1, \dots, v_m, v_{m+1}, \dots, v_n$ *имеет вид* (1.5.1) *и* $A_{1,1}$ *есть блок размером* $m \times m$*, то подпространство* $\langle v_1, \dots, v_m \rangle$ *инвариантно относительно* f.

Доказательство. Из того, что подпространство $U = \langle v_1, \dots, v_m \rangle$ инвариантно относительно f, следует, что $f(v_j) = \sum\limits_{k=1}^{m} a_{k,j} v_k$ при $1 \le j \le m$. В то же время при $j > m$ должны выполняться равенства: $f(v_j) = \sum\limits_{k=1}^{m} a_{k,j} v_k + \sum\limits_{k=m+1}^{n} a_{k,j} v_k$. Если записать полученную таким образом матрицу оператора f, то получится матрица вида:

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

При этом матрица $A_{1,1}$ имеет размер $m \times m$, и состоит из элементов $a_{k,j}$, где $1 \le k, j \le m$. По определению матрицы линейного оператора $f|_U : U \to U$, вычисленной в базисе $v_1, \dots, v_m$, эта матрица совпадает с $A_{1,1}$. Матрица $A_{1,2}$ имеет размер $m \times (n-m)$, и состоит из элементов $a_{k,j}$, где $1 \le k \le m$, $m+1 \le j \le n$, матрица $A_{2,2}$ имеет размер $(n-m) \times (n-m)$, и состоит из элементов $a_{k,j}$, где $m+1 \le k, j \le n$.

Обратно, пусть в некотором базисе $v_1, \dots, v_m, v_{m+1}, \dots, v_n$ матрица оператора f имеет вид (1.5.1), причем блок $A_{1,1}$ — это матрица размером $m \times m$. Так как в матрице (1.5.1) в левом нижнем углу стоит блок из нулей, то для базисных элементов $v_1, \dots, v_m$ должны выполняться равенства: $f(v_j) = \sum\limits_{k=1}^{m} a_{k,j} v_k$. Если $U = \langle v_1, \langle, v_m \rangle$, то $f(v_j) \in U$ для всех j. По предыдущей лемме отсюда следует, что подпространство U инвариантно относительно f. □

Теорема 1.5.2. *Пусть* $f : V \to V$ *— линейный оператор, и* $V =$

$V_1 \oplus \dots \oplus V_m$, где каждое V_i есть инвариантное относительно f подпространство. Выберем в каждом подпространстве V_i некоторый базис $v_{i,1}, \dots, v_{i,n_i}$, $1 \le i \le m$, Тогда множество $\bigcup_{i=1}^{m}\{v_{i,1}, \dots, v_{i,n_i}\}$ есть базис всего пространства V, и матрица f в этом базисе имеет следующий блочно-диагональный вид:

$$\begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_m \end{pmatrix} \tag{1.5.2}$$

Каждый блок A_i есть квадратная $n_1 \times n_i$-матрица, которая является матрицей ограничения f на подпространство V_i в базисе $v_{i,1}, \dots, v_{i,n_i}$, $1 \le i \le m$.

Обратно, пусть матрицей f в некотором базисе вида $\bigcup_{i=1}^{m}\{v_{i,1}, \dots, v_{i,n_i}\}$ является матрица вида (1.5.2), причем блоки A_i имеют размер $n_i \times n_i$. Тогда все подпространства $V_i = \langle v_{i,1}, \dots, v_{i,n_i}\rangle$ инвариантны относительно f, и $V = V_1 \oplus \dots \oplus V_m$.

Доказательство. Из того, что все подпространства V_i инвариантны относительно f, следует, что для каждого j имеется включение $f(v_{i,j}) \in V_i$, а так как $V_i = \langle v_{i,1}, \dots, v_{i,n_i}\rangle$, то

$$f(v_{i,j}) = \sum_{k=1}^{n_i} a_{i,j}^{i,k} v_{i,k} \tag{1.5.3}$$

Матрица A_i размером $n_i \times n_i$, k, j-й элемент которой есть $a_{i,j}^{i,k}$, есть матрица линейного оператора $f|_{V_i} : V_i \to V_i$. Если же рассматривать соотношения (1.5.3) как определение матрицы оператора на всем пространстве V, и выписать подробно эту матрицу, то это и будет матрица (1.5.2).

Обратно, пусть матрица f имеет вид (1.5.2). Это означает, что выполняются соотношения (1.5.3) для всех i и j. Положим $V_i =$

$\langle v_{i,1}, \dots, v_{i,n_i}\rangle$. Тогда из (1.5.3) следует, что $f(v_{i,j}) \in V_i$ для всех i и j. Это означает, что подпространствва V_i инвариантны относительно f для всех i (лемма 1.5.1). По самому построению сумма этих подпространств является прямой (материал первого семестра; см. также теорему 3.1.5 в третьей главе) и совпадает со всем пространством V. $\square$

Снова рассмотрим линейный оператор $f : V \to V$, и разложение $V = V_1 \oplus \dots \oplus V_m$ в прямую сумму инвариантных подпространств. Тогда $f_i = f|_{V_i}$ есть линейный оператор $f_i : V_i \to V_i$, и если, как в предыдущей теореме, в каждом пространстве V_i заданы базисы $\{v_{i,1}, \dots, v_{i,n_i}\}$, то матрицы A_i из (1.5.1) являются матрицами операторов f_i в указанных базисах.

Легко убедиться, что набор операторов $f_1, \dots, f_m$ полностью определяет оператор f. Чтобы проверить это, запишем произвольный вектор $v \in V$ в виде $v = v_1 + \dots + v_m$, где $v_i \in V_i$ для каждого i, $1 \leq i \leq m$. Из определения прямой суммы следует, что такую запись можно осуществить только одним способом. Применяя оператор f, получаем $f(v) = f(v_1) + \dots + f(v_m)$. Но так как V_i инвариантны, то для всех i будем иметь $f(v_i) \in V_i$, и, более того, $f(v_i) = f_i(v_i)$. Таким образом,

$$f(v) = f_1(v_1) + \dots + f_m(v_m) \tag{1.5.4}$$

Из этого соотношения следует, что если известны $f_1, \dots, f_m$, то известен и f.

Обратно, пусть дано какое-то разложение пространства V в прямую сумму $V = V_1 \oplus \dots \oplus V_m$, и для всех i, $1 \leq i \leq m$, заданы линейные операторы $f_i : V_i \to V_i$. Тогда можно построить линейный оператор $f : V \to V$ по формуле (1.5.4). Нетрудно убедиться, что все подпространства V_i будут инвариантными относительно f, и $f|_{V_i} = f_i$ для каждого i.

Построенный таким образом линейный оператор f называется прямой суммой линейных операторов $f_1, \dots, f_m$ и обозначается через $f = f_1 \oplus \dots \oplus f_m$. Матрицы операторов такого вида в базисах, составленных из базисов прямых слагаемых V_i, это матрицы вида (1.5.2).

Теорема 1.5.3. *Пусть V — векторное пространство над полем $\mathbb{C}$. У любого линейного оператора $f : V \to V$ существует одномерное инвариантное подпространство. Если же V — векторное пространство над $\mathbb{R}$, то каждый линейный оператор f обладает либо одномерным, либо двумерным инвариантным подпространством.*

Доказательство. В случае, когда V — векторное пространство над полем $\mathbb{C}$, у характеристического многочлена $\chi_f(x)$ обязательно есть корень $\lambda \in \mathbb{C}$. Этот корень, как уже известно, является собственным значением f, и поэтому должен существовать ненулевой вектор v (собственный вектор) со свойством $f(v) = \lambda v$. Тогда $\langle v \rangle$ — одномерное инвариантное подпространство оператора f.

Рассмотрим случай векторного пространства над полем действительных чисел. Коэффициенты многочлена $\chi_f(x)$ теперь будут действительными числами, и наличие у $\chi_f(x)$ действительного корня не гарантировано. Однако если такой корень имеется, то можно расуждать так же, как и в комплексном случае, откуда следует существование одномерного инвариантного подпространства.

Допустим теперь, что у многочлена $\chi_f(x)$ нет действительных корней. Это означает, в частности, что у оператора f нет собственных значений и собственных векторов. Выберем некоторый базис $v_1, \dots, v_n$ пространства V, и пусть A — матрица оператора f в этом базисе. Таким образом, $f(v_j) = \sum\limits_{k=1}^{n} a_{k,j} v_k$ для всех j, и $\chi_f(x) = |A - xE|$. Выберем некоторый комплексный корень $\lambda = \alpha + i\beta$ этого многочлена, рассматриваемого уже как многочлен над полем комплексных чисел. Так как определитель матрицы (с комплексными компонентами) $A - \lambda E$ равен нулю, должно существовать ненулевое (комплексное) решение системы линейных однородных уравнений $(A - \lambda E)y = 0$, где $y = (y_1, \dots, y_n)^{\mathrm{T}}$. Пусть $y_1 = \xi_1 + i\eta_1, \dots, y_n = \xi_n + i\eta_n$ — это ненулевое решение. Представим равенство $(A - \lambda E)y = 0$ в виде $Ay = \lambda y$, и запишем его более подробно:

$$\begin{aligned}
&a_{1,1}y_1 + a_{1,2}y_2 + \dots + a_{1,n}y_n = \lambda y_1\\
&a_{2,1}y_1 + a_{2,2}y_2 + \dots + a_{2,n}y_n = \lambda y_2\\
&\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots\\
&a_{n,1}y_1 + a_{n,2}y_2 + \dots + a_{n,n}y_n = \lambda y_n
\end{aligned}$$

Отделяя вещественную часть от комплексной, получим два набора равенств:

$$\begin{aligned}
&a_{1,1}\xi_1 + a_{1,2}\xi_2 + \dots + a_{1,n}\xi_n = \alpha\xi_1 - \beta\eta_1\\
&a_{2,1}\xi_1 + a_{2,2}\xi_2 + \dots + a_{2,n}\xi_n = \alpha\xi_2 - \beta\eta_2\\
&\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots\\
&a_{n,1}\xi_1 + a_{n,2}\xi_2 + \dots + a_{n,n}\xi_n = \alpha\xi_n - \beta\eta_n
\end{aligned} \tag{1.5.5}$$

и

$$\begin{aligned}
&a_{1,1}\eta_1 + a_{1,2}\eta_2 + \dots + a_{1,n}\eta_n = \alpha\eta_1 + \beta\xi_1\\
&a_{2,1}\eta_1 + a_{2,2}\eta_2 + \dots + a_{2,n}\eta_n = \alpha\eta_2 + \beta\xi_2\\
&\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots\\
&a_{n,1}\eta_1 + a_{n,2}\eta_2 + \dots + a_{n,n}\eta_n = \alpha\eta_n + \beta\xi_n
\end{aligned} \tag{1.5.6}$$

Обозначим через ξ столбец $(\xi_1, \dots, \xi_n)^{\mathrm{T}}$, а через η — столбец $(\eta_1, \dots, \eta_n)^{\mathrm{T}}$. Тогда (1.5.5) запишется в виде $A\xi = \alpha\xi - \beta\eta$, а (1.5.6) — в виде $A\eta = \beta\xi + \alpha\eta$. Покажем, что столбцы ξ и η не нулевые, и не пропорциональны друг другу. Напомним, что число $\lambda = \alpha + i\beta$ не является действительным, так что $\beta \neq 0$. Кроме того, столбцы ξ и η не могут быть равны нулю одновременно (одновременное равенство нулю означало бы, что все y_j равны нулю).

Допустим, что $\xi = 0$. Тогда $\eta \neq 0$, и из $A\eta = \beta\xi + \alpha\eta$ получаем $A\eta = \alpha\eta$, то есть имеется ненулевое решение системы уравнений $(A - \alpha E)z =$

0, а это означает, что определитель $|A - \alpha E|$ равен нулю. Но тогда α является действительным корнем $\chi_f(x) = |A - xE|$, что противоречит сделанному выше предположению.

Допустим, что $\eta = 0$. Тогда $\xi \neq 0$, и из $A\xi = \alpha\xi - \beta\eta$ получим $A\xi = \alpha\xi$. Таким образом, ξ оказывается ненулевым решением системы уравнений $(A - \alpha E)z = 0$, а это снова влечет, что α есть действительный корень $\chi_f(x)$.

Теперь предположим, что $\eta = \gamma\xi$ для некоторого действительного числа γ. Так как уже показано, что столбцы ξ и η ненулевые, то $\gamma \neq 0$. Теперь из $A\eta = \beta\xi + \alpha\eta$ следует

$$A\eta = (\beta\gamma + \alpha)\eta.$$

Отсюда, как и выше, делаем вывод, что η есть ненулевое решение системы уравнений $(A - (\beta\gamma + \alpha)E)z = 0$, а это означает, что определитель матрицы $A - (\beta\gamma + \alpha)E$ равен нулю. Поэтому действительное число $\beta\gamma + \alpha$ является корнем многочлена $\chi_f(x) = |A - xE|$, а это противоречит сделанному предположению о том, что действительных корней нет.

Рассмотрим векторы $u = \xi_1 v_1 + \dots + \xi_n v_n$ и $w = \eta_1 v_1 + \dots + \eta_n v_n$, и выясним, как выглядят векторы $f(u)$ и $f(w)$. Используем теорему 1.2.2 (равенство (1.2.3)). Согласно этой теореме, координатами вектора $f(u)$ в базисе $v_1, \dots, v_n$ являются компоненты столбца $A\xi = \alpha\xi - \beta\eta$. Таким образом, $f(u)$ есть вектор

$$\begin{aligned}(\alpha\xi_1 - \beta\eta_1)v_1 + \dots + (\alpha\xi_n - \beta\eta_n)v_n &= \\ \alpha(\xi_1 v_1 + \dots + \xi_n v_n) - \beta(\eta_1 v_1 + \dots + \eta_n v_n) &= \alpha u - \beta w.\end{aligned}$$

Точно так же получаем равенство $f(w) = \beta u + \alpha w$. Рассмотрим подпространство $\langle u, w\rangle$. Векторы u и w ненулевые (иначе были бы нулевыми столбцы их координат ξ, η), и не пропорциональны (иначе пропорциональными были бы столбцы их координат, а выше было показано, что это не так). Следовательно, эти векторы линейно независимы, и подпространство $\langle u, w \rangle$ двумерно. Из равенств $f(u) = \alpha u - \beta w$ и $f(w) = \beta u + \alpha w$ и из леммы 1.5.1 следует, что это подпространство является инвариантным относительно оператора f. $\square$

1.6. Векторные пространства линейных отображений

Для доказательства более глубоких теорем об инвариантных подпространствах будут необходимы некоторые сведения о структуре множества всех линейных отображений из одного пространства в другое.

Определение 1.6.1. Множество всех линейных отображений из некоторого множества X в вектороное пространство W будем обозначать через $\mathrm{Map}(X, W)$ (в литературе можно встретить и другие обозначения, например, W^X).

Лемма 1.6.1. *Множество* $\mathrm{Map}(X, W)$ *обладает структурой векторного пространства над тем же полем* K*, над которым определено и пространство* W*. При этом операции в пространстве* $\mathrm{Map}(X, W)$ *определяются следующим образом:*

$$(f+g)(x) = f(x) + g(x), \qquad (\alpha f)(x) = \alpha f(x).$$

Доказательство. Поясним обозначения. Пусть f и g — произвольные отображения из X в W. Через $f+g$ обозначается новое отображение, которое переводит элемент $x \in X$ в вектор $f(x)+g(x) \in W$. Если $\alpha \in K$, то через αf обозначается новое отображение, которое переводит элемент x в вектор $\alpha f(x) \in W$, то есть в произведение элемента поля α на вектор $f(x)$. Проверим, что эти операции превращают $\mathrm{Map}(X, W)$ в векторное пространство над полем K.

Необходимо проверить аксиомы, определяющие векторное пространство. Все проверки производятся по одной схеме. А именно, каждый раз необходимо установить равенство двух функций, а для этого надо убедиться, что значения этих функций совпадают при любом значении аргумента $x \in X$. Например, если мы проверяем равенство $(f+g)+h = f+(g+h)$, то необходимо вычислить значения $((f+g)+h)(x)$ и $(f+(g+h))(x)$. Вычисления в данном случае состоят в применении определений. С одной стороны,

$$((f+g)+h)(x) = (f+g)(x) + h(x) = (f(x)+g(x)) + h(x).$$

С другой стороны,

$$(f+(g+h))(x) = f(x) + (g+h)(x) = f(x) + (g(x)+h(x)).$$

В правых частях этих равенств находятся векторы из векторного пространства W, а для векторов справедливо равенство $(f(x)+g(x)) + h(x) = f(x) + (g(x)+h(x))$. Тем самым показано, что значения функций $(f+g)+h$ и $f+(g+h)$ совпадают для каждого значения аргумента $x \in X$, но это равносильно тому, что $(f+g)+h = f+(g+h)$.

Аналогично проверяется коммутативность сложения: для каждого $x \in X$

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

Отсюда следует, что $f+g = g+f$. Далее, определим отображение $e : X \to W$ такое, что $e(x) = 0 \in W$ для всех $x \in X$. Тогда для каждой функции $f : X \to W$ будем иметь равенства:

$$(f+e)(x) = f(x) + e(x) = f(x) + 0 = f(x),$$
$$(e+f)(x) = e(x) + f(x) = 0 + f(x) = f(x).$$

Отсюда следует, что $f+e = f = e+f$. Таким образом, функция e играет роль нулевого элемента для операции сложения в $Map(X,W)$, и ее можно обозначить символом 0. Это общепринятое обозначение, но надо иметь в виду, что одним и тем же символом "0" будет обозначаться нулевой элемент поля K, нулевой вектор векторного пространства W, и нулевая функция, то есть отображение из X в W, обозначаемое тем же символом 0, и такое, что $0(x) = 0$, где справа от знака равенства стоит нулевой вектор пространства W. В каком значении употребляется символ "0" в той или иной конкретной ситуации, должно быть видно из контекста.

Покажем теперь, что у каждого элемента $f \in \mathrm{Map}(X, W)$ есть обратный по сложению элемент. Определим отображение $-f$, полагая $(-f)(x) = -f(x)$. Слева $-f$ есть символ, обозначающий некоторое новое отображение, а справа $-f(x) \in W$ усть вектор из пространства W, обратный по сложению к вектору $f(x)$. Проверим, что $f + (-f) = (-f) + f = 0$. Для какого угодно $x \in X$ получаем:

$$(f+(-f))(x) = f(x)+(-f)(x) = f(x)+(-f(x)) = f(x)-f(x) = 0 = 0(x),$$

и точно так же $((-f) + f)(x) = 0 = 0(x)$.

Теперь займемся операцией αf, где $\alpha \in K$, $f \in \mathrm{Map}(X, W)$. Во-первых, заметим, что $-f = (-1)f$, так как для произвольного $x \in X$ по определению имеются равенства: $(-f)(x) = -f(x) = (-1)f(x) = ((-1)f)(x)$. Равенство функций $(\alpha\beta)f = \alpha(\beta f)$ следует из того, что

$$((\alpha\beta)f)(x) = (\alpha\beta)f(x) = \alpha(\beta f(x)) = \alpha(\beta f)(x) = (\alpha(\beta f))(x).$$

Равенство функций $(\alpha + \beta)f = (\alpha f) + (\beta f)$ следует из того, что

$$((\alpha+\beta)f)(x) = (\alpha+\beta)f(x) = \alpha f(x)+\beta f(x) = (\alpha f)(x)+(\beta f)(x) = (\alpha f+\beta f)(x).$$

Равенство $\alpha(f + g) = \alpha f + \alpha g$ следует из того, что

$$\begin{aligned}(\alpha(f + g))(x) = \ & \alpha(f + g)(x) = \alpha(f(x) + g(x)) = \\ & \alpha f(x) + \alpha g(x) = (\alpha f)(x) + (\alpha g)(x) = \\ & (\alpha f + \alpha g)(x).\end{aligned}$$

Остается убедиться, что $1f = f$ и $0f = 0$. Проверка первого из этих равенств тривиальна: $(1f)(x) = 1 \cdot f(x) = f(x)$. Во втором надо иметь в виду, что в выражении $0f$ символ 0 обозначает элемент поля K, а справ от знака равенства стоит символ 0, обозначающий нулевую функцию (нулевое отображение). Проверка такова: $(0f)(x) = 0 \cdot f(x) = 0 = 0(x)$. Здесь символ 0 появляется еще в одном значении, как нулевой вектор пространства W. $\square$

Определение 1.6.2. Множество всех линейных отображений из векторного пространства V в векторное пространство W (над одним и тем же полем K) обозначается через $\mathcal{L}(V, W)$. Если $V = W$, то это множество обозначается также через $\mathcal{L}(V)$. Если $W = K$, то множество $\mathcal{L}(V, K)$ обозначается через V^*, и называется *пространством, двойственным* (или *сопряженным*) *к пространству* V.

Отметим, что $\mathcal{L}(V, W) \subset \mathrm{Map}(V, W)$ (справа — множество *всех*, а не только линейных, отображений из V в W).

Теорема 1.6.1. *Множество $\mathcal{L}(V, W)$ есть векторное подпространство векторного пространства* $\mathrm{Map}(X, W)$. *При любом выборе базисов в V и W взаимно-однозначное соответствие между $\mathcal{L}(V, W)$ и $M_{m,n}(K)$ оказывается изоморфизмом векторных пространств. Если $W = V$, то суперпозиция линейных операторов определяет на $\mathcal{L}(V) = \mathcal{L}(V, V)$ структуру ассоциативного кольца с единицей, изоморфного кольцу квадратных матриц $M_n(K)$. Этот изоморфизм одновременно является изоморфизмом векторных пространств над полем K.*

Доказательство. Проверим, что $\mathcal{L}(V, W)$ есть векторное подпространство векторного пространства $\mathrm{Map}(V, W)$. Пусть $f, g : V \to W$ — линейные отображения, и $\alpha, \beta \in K$. Определим отображение $h : V \to W$

по формуле: $h(v) = \alpha f(v) + \beta g(v)$. Покажем, что это отображение линейно. Пусть $v_1, v_2 \in V$, $\lambda_1, \lambda_2 \in K$. Тогда

$$\begin{aligned} h(\lambda_1 v_1 + \lambda_2 v_2) = \ & \alpha f(\lambda_1 v_1 + \lambda_2 v_2) + \beta g(\lambda_1 v_1 + \lambda_2 v_2) = \\ & \alpha(\lambda_1 f(v_1) + \lambda_2 f(v_2)) + \beta(\lambda_1 g(v_1) + \lambda_2 g(v_2)) = \\ & \lambda_1(\alpha f(v_1) + \beta g(v_1)) + \lambda_2(\alpha f(v_2) + \beta g(v_2)) = \\ & \lambda_1 h(v_1) + \lambda_2 h(v_2). \end{aligned}$$

То, что $h = \alpha f + \beta g \in \mathcal{L}(V, W)$, и означает, что $\mathcal{L}(V, W)$ есть векторное подпространство пространства $\mathrm{Map}(V, W)$.

Выберем в пространствах V и W базисы $v_1, \ldots, v_n$ и $w_1, \ldots, w_m$. Рассмотрим $f, g \in \mathcal{L}(V, W)$, $\alpha, \beta \in K$, и покажем, что сопоставление линейному отображению его матрицы (вычисленной в данных фиксированных базисах) $f \mapsto M_f$ есть линейное отображение. Таким образом, надо доказать равенство $M_{\alpha f + \beta g} = \alpha M_f + \beta M_g$.

Положим $A = (a_{i,j}) = M_f$, $B = (b_{i,j}) = M_g$. По определению,

$$f(v_j) = \sum_{i=1}^{m} a_{i,j} w_i, \quad g(v_j) = \sum_{i=1}^{m} b_{i,j} w_i.$$

Отсюда

$$(\alpha f + \beta g)(v_j) = \alpha f(v_j) + \beta g(v_j) = \sum_{i=1}^{m} (\alpha a_{i,j} + \beta b_{i,j}) w_i.$$

Отсюда сразу следует, что матрицей линейного отображения $\alpha f + \beta g$ будет матрица $\alpha A + \beta B$. Поскольку уже известно, что соответствие $f \mapsto M_f$ есть биекция, то из доказаного следует, что это соответствие есть изоморфизм векторных пространств.

Теперь рассмотрим три векторных пространства V, W, U. Пусть $f, f_1, f_2 \in \mathcal{L}(V, W)$, $g, g_1, g_2 \in \mathcal{L}(W, U)$, $\alpha_1, \alpha_2, \beta_1, \beta_2 \in K$. Определены суперпозиции отображений вида gf_i, fg_i и т.п., и это также линейные отображения. Покажем, что $g(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1(gf_1) + \alpha_2(gf_2)$ и

$(\beta_1 g_1 + \beta_2 g_2)f = \beta_1(g_1 f) + \beta_2(g_2 f)$. Как и выше, необходимо взять произвольный вектор $v \in V$, вычислить значения функций в левых и правых частях доказываемых равенств для аргумента v, и убедиться, что эти значения совпадают.

$$\begin{aligned}(g(\alpha_1 f_1 + \alpha_2 f_2))(v) = \ & g((\alpha_1 f_1 + \alpha_2 f_2)(v)) = \\ & g(\alpha_1 f_1(v) + \alpha_2 f_2(v)) = \\ & \alpha_1 g(f_1(v)) + \alpha_2 g(f_2(v)) = \\ & \alpha_1 (g f_1)(v) + \alpha_2 (g f_2)(v) = \\ & (\alpha_1 (g f_1) + \alpha_2 (g f_2))(v).\end{aligned}$$

Второе равенство доказывается примерно так же.

Теперь рассмотрим случай $V = W$. Множество $\mathcal{L}(V) = \mathcal{L}(V, V)$, как уже показано, есть векторное пространство. Для любых $f_1, f_2 \in \mathcal{L}(V)$ определена суперпозиция $f_1 f_2 \in \mathcal{L}(V)$, причем $(f_1 f_2) f_3 = f_1(f_2 f_3)$. Так как $g(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1(g f_1) + \alpha_2(g f_2)$ и $(\beta_1 g_1 + \beta_2 g_2)f = \beta_1(g_1 f) + \beta_2(g_2 f)$, то отсюда следует, что $\mathcal{L}(V)$ является ассоциативным кольцом с определенной выше операцией сложения $f + g$ и с операцией умножения — суперпозицией fg. В этом кольце есть нулевой элемент 0_V и единица 1_V (тождественное отображение).

Выберем и зафиксируем базис $v_1, \dots, v_n$ в пространстве V, и будем рассматривать матрицы линейных отображений из $\mathcal{L}(V)$ в этом базисе. Ранее уже было доказано, что $M_{fg} = M_f M_g$, и что матрица тождественного (единичного) отображения $1_V : V \to V$ есть единичноая матрица. Все это означает, что биекция между $\mathcal{L}(V)$ и $M_n(K)$, которая строится по данному базису, есть изоморфизм *колец*. $\square$

Из этой теоремы следует, то если $\dim(V) = n, \dim(W) = m$, то $\dim(\mathcal{L}(V, W)) = \dim(M_{m,n}(K)) = nm$. В частности, $\dim(V^*) = n = \dim(V)$. По произвольному базису $v_1, \dots, v_n$ пространства V можно построить семейство линейных отображений $\nu_i : V \to K$, $1 \le i \le n$, такиих, что $\nu_i(v_j) = \delta_{i,j}$.

Лемма 1.6.2. *Семейство $\nu_1, \dots, \nu_n$ является базисом в V^*. Для каждого линейного отображения $\xi : V \to K$ имеет место равенство:*

$$\xi = \sum_{i=1}^{n} \xi(v_i)\nu_i$$

Доказательство. Дадим два доказательства первого утверждения. Первое доказательство использует построенный в предыдущей теореме изоморфизм между пространствами $\mathcal{L}(V, K)$ и $M_{1,n}(K)$. Этот изоморфизм сопоставляет линейному отображению $f : V \to K$ его матрицу M_f, вычисленную с помощью базиса $v_1, \dots, v_n$ пространства V, и базиса 1 одномерного пространства K. Эта матрица с одной строкой и n столбцами, т.е. строка $(a_1, \dots, a_n)$, где $a_j = f(v_j)$ для каждого j (это сразу следует из определения матрицы линейного отображения). Отсюда следует, что отображению ν_i соответствует матрица — строка $e_i = (0, \dots, 0, 1, 0, \dots, 0)$, где единица располагается на i-м месте (в i-м столбце), а на остальных местах располагаются нули. Но элементы $e_1, \dots, e_n$ образуют базис пространства $M_{1,n}(K)$. Соответствующие им при изоморфизме двух пространств элементы $\nu_1, \dots, \nu_n$ поэтому также являются базисом.

Второе доказательство таково. Покажем сначала, что

$$\xi = \sum_{i=1}^{n} \xi(v_i)\nu_i$$

для каждого линейного отображения $\xi : V \to K$. Слева и справа от знака равенства расположены здесь линейные отображения, а для того, чтобы доказать равенство двух линейных отображений, достаточно (и необходимо) показать, что значения этих отображений совпадают на каждом базисном векторе v_j. Таким образом, следует убедиться, что значение отображения в правой части доказываемого равенства на векторе v_j равно $\xi(v_j)$. Проделаем это вычисление:

$$(\sum_{i=1}^{n}\xi(v_i)\nu_i)(v_j) = \sum_{i=1}^{n}\xi(v_i)\nu_i(v_j) = \sum_{i=1}^{n}\xi(v_i)\delta_{i,j} = \xi(v_j).$$

Покажем теперь, что отображения $\nu_1, \dots, \nu_n$ как векторы пространства V^* являются линейно независимыми. Допустим, что некоторая линейная комбинация этих отображений равна нулю:

$$\sum_{i=1}^{n}\alpha_i\nu_i = 0.$$

Покажем, что в этом случае равны нулю все коэффициенты α_i. Рассматриваемое равенство является равенством двух линейных отображений, поэтому должны быть равны значения этих отображений при любом значении аргумента. В частности, для произвольного v_j получаем:

$$(\sum_{i=1}^{n}\alpha_i\nu_i)(v_j) = \sum_{i=1}^{n}\alpha_i\nu_i(v_j) = \sum_{i=1}^{n}\alpha_i\delta_{i,j} = \alpha_j = 0(v_j) = 0.$$

Таким образом, выполнены оба свойства из определения базиса пространства V^*. □

Построенный выше базис $\nu_1, \dots, \nu_n$ пространства V^* называется базисом, *дуальным* к базису $v_1, \dots, v_n$.

Лемма 1.6.3. *Пусть $\varphi : V \to W$ — линейное отображение. Тогда определено линейное отображение $\varphi^* : W^* \to V^*$, переводящее $\xi \in W^*$ в $\varphi^*(\alpha) = \xi\varphi$.*

Тут надо иметь в виду, что φ и ξ — отображения следующего вида:

$$V \xrightarrow{\varphi} W \xrightarrow{\xi} K,$$

и поэтому суперпозиция $\xi\varphi$ определена.

Доказательство. Необходимо убедиться в том, что $\varphi^*(\lambda_1\xi_1 + \lambda_2\xi_2) = \lambda_1\varphi^*(\xi_1) + \lambda_2\varphi^*(\xi_2)$ для произвольных $\lambda_1, \lambda_1 \in K$ и $\xi_1, \xi_2 \in W^*$. Выписывая явный вид всех входящих в это равенство выражений, получаем следующее соотношение (именно его и надо установить):

$$(\lambda_1\xi_1 + \lambda_2\xi_2)\varphi = \lambda_1(\xi_1\varphi) + \lambda_2(\xi_2\varphi).$$

Но это — частный случай равенства, уже полученного при доказательстве теоремы 1.6.1. □

Полагая $V = W$, получаем, что для каждого линейного оператора $\varphi : V \to V$ определен линейный оператор $\varphi^* : V^* \to V^*$.

1.7. Снова инвариантные подпространства

Теорема 1.7.1. *Пусть V — векторное пространство размерности n над полем комплексных чисел, $f : V \to V$ — линейный оператор. Тогда существует подпространство $W \subset V$, инвариантное относительно f и имеющее размерность $n-1$.*

Доказательство. Рассмотрим линейный оператор $f^* : V^* \to V^*$, переводящий линейное отображение $\xi : V \to K$ в линейное отображение $\xi f : V \to K$. Мы уже знаем, что в случае, когда рассматривается векторное пространство над полем комплексных чисел $\mathbb{C}$, у каждого такого линейного оператора есть одномерное инвариантное подпространство, то есть некоторый собственный вектор. Пусть $\xi \in V^*$ — это такой собственный вектор, и $\lambda \in \mathbb{C}$ — это соответствующее ему собственное значение. По определению это означает, что $\xi \neq 0$, и

$$f^*(\xi) = \xi f = \lambda\xi \tag{1.7.1}$$

Так как $\xi \neq 0$, то существует вектор $v \in V$, такой, что $\xi(v) = \alpha \neq 0$, $\alpha \in K$. Но тогда для каждого $\lambda \in K$ получаем равенство $\lambda = \xi(\lambda\alpha^{-1}v)$.

Это означает, что образ $\mathrm{Im}(\xi)$ линейного отображения ξ совпадает с K, то есть имеет размерность, равную единице. Применяя равенство

$$n = \dim(V) = \dim(\mathrm{Ker}(\xi)) + \dim(\mathrm{Im}(\xi)),$$

получаем, что $\dim(\mathrm{Ker}(\xi)) = n - 1$. Осталось убедиться, что подпространство $W = \mathrm{Ker}(\xi)$ является инвариантным относительно f. В самом деле, пусть $w \in W$. Чтобы убедиться в том, что $f(w) \in W$, надо использовать условие, определяющее подпространство W, а именно, $W = \mathrm{Ker}(\xi)$. Иными словами, $f(w) \in W$ тогда и только тогда, если $\xi(f(w)) = 0$. Проверим, что это так и есть. Для этого используем равенство (1.7.1):

$$\xi(f(w)) = (\xi f)(w) = (\lambda\xi)(w) = \lambda\xi(w) = 0,$$

так как $\xi(w) = 0$ ввиду $w \in W = \mathrm{Ker}(\xi)$. Теорема доказана. $\square$

В дальнейшем будут использоваться следующие обозначения. Линейные операторы будут обозначаться буквами “рукописного” алфавита $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$, причем буквой $\mathcal{E}$ будт обозначаться тождественный оператор. Далее, там, где жто возможно, будут опускаться скобки при записи аргумента функции. Например, вместо $\mathcal{A}(x)$ будем писать $\mathcal{A}x$. Если $\mathcal{A}$ — линейный оператор, действующий из V в V, то его образ будет обозначаться через $\mathcal{A}V$. Итак, $\mathcal{A}V = \mathrm{Im}(\mathcal{A}) = \{\mathcal{A}v | v \in V\}$. Наконец, если $\mathcal{A}$ есть линейный оператор, обозначенный буквой “рукописного” шрифта, то через A будет обозначаться его матрица в каком-либо базисе (каком именно — должно следовать из контекста). То есть матрица обозначается той же буквой, но в “печатном” шрифте. Эти обозначения весьма распостранены в теории линейных операторов.

Теорема 1.7.2. *Пусть V — векторное пространство над $\mathbb{C}$, $\dim(V) = n$, $\mathcal{A} : V \to V$ — произвольный линейный оператор. Тог-*

да существуют инвариантные относительно $\mathcal{A}$ подпространства

$$V = V_n \supset V_{n-1} \supset \ldots \supset V_j \supset \ldots \supset V_1 \supset V_0 = \{0\},$$

такие, что $\dim(V_j) = j$ *для каждого* j, $0 \leq j \leq n$.

Доказательство. Эта теорема легко выводится из предыдущей теоремы с учетом следующего замечания. Пусть $U \subseteq V$ — инвариантное относительно $\mathcal{A}$ подпространство, и пусть $\mathcal{B} = \mathcal{A}|_U : U \to U$ — ограничение A на U. Если подпространство $W \subset U$ инвариантно относительно $\mathcal{B}$, то оно инвариантно и относительно $\mathcal{A}$.

В самом деле, если $w \in W$, то $\mathcal{A}w = \mathcal{B}w \in W$. Первое равенство имеет место по определению оператора $\mathcal{B} = \mathcal{A}|_U$, а второе — ввиду инвариантости W относительно $\mathcal{B}$. □

Следствие 1.7.1. *Для любого линейного оператора $\mathcal{A}$ можно выбрать базис таким образом, что его матрица в этом базисе будет иметь верхнетреугольный вид. Иными словами, для любой матрицы A с элементами из поля $\mathbb{C}$ существует невырожденная матрица B такая, что матрица $B^{-1}AB$ является верхнетреугольной.*

Доказательство. Рассмотрим инвариантные подпространства, существование которых гарантирует предыдущая теорема:

$$V = V_n \supset V_{n-1} \supset \ldots \supset V_j \supset \ldots \supset V_1 \supset V_0 = \{0\},$$

где $dim(V_j) = j$ для всех j. Выберем базисный вектор v_1 в одномерном подпространстве V_1. Рассуждая по индукции, предположим, что уже выбраны векторы $v_1, \ldots, v_j$ так , что $V_1 = \langle v_1 \rangle$, $V_2 = \langle v_1, v_2 \rangle$, $\ldots$, $V_j = \langle v_1, \ldots, v_j \rangle$. Рассмотрим V_{j+1}. Подпространство V_j содержится в V_{j+1}. V_{j+1} имеет на единицу большую размерность, чем V_j, поэтому базис V_j можно дополнить одним вектором v_{j+1} до базиса V_{j+1}. Таким образом, $V_{j+1} = \langle v_1, \dots, v_j, v_{j+1} \rangle$. В частности, $V_n = \langle v_1, \dots, v_n \rangle$. Вычислим матрицу $\mathcal{A}$ в этом базисе. Поскольку V_j является инвариантным подпространством для каждого $j = 1, \dots, n$. В частности, $\mathcal{A}v_j \in V_j$, то есть вектор $\mathcal{A}v_j$ есть линейная комбинация векторов $v_1, \dots, v_j$. Отсюда получим:

$$
\begin{aligned}
&\mathcal{A}v_1 = a_{1,1}v_1 \\
&\mathcal{A}v_2 = a_{1,2}v_1 + a_{2,2}v_2 \\
&\quad \dots \quad \dots \quad \dots \quad \dots \\
&\mathcal{A}v_j = a_{1,j}v_1 + a_{2,j}v_2 + \dots + a_{j,j}v_j \\
&\quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \\
&\mathcal{A}v_n = a_{1,n}v_1 + a_{2,n}v_2 + \dots + a_{j,n}v_j + \dots + a_{n,n}v_n
\end{aligned}
\tag{1.7.2}
$$

Таким образом, получилась верхнетреугольная матрица

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

$\square$

Отметим, что если матрица A верхнетреугольна, то на диагонали у нее располагаются ее собственные значения, и только они.

Тот факт, что пространство всех линейных операторов из V в V обладает структурой кольца с единицей $\mathcal{E}$, означает, что для каждого многочлена $f(x) = a_0 + a_1x + \dots + a_kx^k$ с коэффициентами из K, и для каждого линейного оператора $\mathcal{A}$ определен линейный оператор $a_0\mathcal{E} + a_1\mathcal{A} + \dots + a_k\mathcal{A}^k$, который обозначается через $f(\mathcal{A})$. При этом, если $f(x) = f_1(x) \dots f_l(x)$, то $f(\mathcal{A}) = f_1(\mathcal{A}) \dots f_l(\mathcal{A})$. Если $f(x) = f_1(x)f_2(x)$, то $f(\mathcal{A}) = f_1(\mathcal{A})f_2(\mathcal{A}) = f_2(\mathcal{A})f_1(\mathcal{A})$. Кроме того, если $f(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x)$, где $\alpha_1, \alpha_2 \in K$, а f_1 и f_2 — многочлены, то $f(\mathcal{A}) = \alpha_1 f_1(\mathcal{A}) +$

$\alpha_2 f_2(\mathcal{A})$. Точно таким же образом можно определить многочлены вида $f(A)$, где A есть квадратная $n \times n$-матрица с компонентами из поля K. Выполняются те же свойства, что и для $f(\mathcal{A})$. Отсюда следует, что, если выбран некоторый базис пространства V, и $M_{\mathcal{A}}$ — матрица $\mathcal{A}$ в этом базисе, то для любого многочлена $f(x)$ будет выполняться равенство: $M_{f(\mathcal{A})} = f(M_{\mathcal{A}})$.

Ввиду того, что множество линейных операторов является конечномерным векторным пространством, бесконечная последовательность $\mathcal{E}, \mathcal{A}, \mathcal{A}^2, \dots, \mathcal{A}^k, \dots$ не может быть линейно независимым множеством. Должна существовать нетривиальная линейная зависимость вида $a_{i_1}\mathcal{A}^{i_1} + \dots + a_{i_m}\mathcal{A}^{i_m} = 0$. Это означает, что найдется ненулевой многочлен $f(x) = a_{i_1}x^{i_1} + \dots + a_{i_m}x^{i_m}$, такой, что $f(\mathcal{A}) = 0$. Этот многочлен опредлен далеко не однозначно. Например, если $g(x) = f(x)f_1(x)$, то и $g(\mathcal{A}) = 0$ при любом $f_1(x)$. Описание всей совокупности таких многочленов будет дано в последнем параграфе второй главы. Явный вид для одного из них дает следующая теорема.

Теорема 1.7.3. (Гамильтон–Кэли) *Пусть V — векторное пространство над $\mathbb{C}$,*
$\mathcal{A} : V \to V$ — линейный оператор, $\chi_{\mathcal{A}}(x)$ — его характеристический многочлен. Тогда $\chi_{\mathcal{A}}(\mathcal{A}) = 0$.

При доказательстве теоремы используется следующая лемма.

Лемма 1.7.1. *Пусть $\mathcal{A} : V \to V$ — некоторый линейный оператор, и $U \subseteq V$ — его инвариантное подпространство. Тогда U будет также инвариантным относительно оператора $\mathcal{A} - \lambda\mathcal{E}$ для любого λ из поля K.*

Заметим, что $\mathcal{A} = (\mathcal{A} - \lambda\mathcal{E}) - (-\lambda)\mathcal{E}$, так что если подпространство U инвариантно относительно $\mathcal{A} - \lambda\mathcal{E}$, то оно инвариантно и относительно $\mathcal{A}$.

Доказательство леммы. Пусть $u \in U$. Тогда по условию $\mathcal{A}u \in U$. Рассмотрим $(\mathcal{A} - \lambda\mathcal{E})u = \mathcal{A}u - \lambda u$. Векторы $\mathcal{A}u$ и λu принадлежат подпространству U, поэтому их разность также принадлежит U. $\square$

Доказательство теоремы Гамильтона-Кэли. Рассмотрим инвариантные подпространства $V_j = \langle v_1, \dots, v_j \rangle$, $1 \le j \le n$ из теоремы 1.7.2 и следствия 1.7.1. Из (1.7.2) следует, что элементы $a_{1,1}, \dots, a_{n,n}$ — это все корни характеристического многочлена $\chi_{\mathcal{A}}(x)$. Точнее, $\chi_{\mathcal{A}}(x) = (-1)^n(x - a_{1,1})(x - a_{2,2}) \dots (x - a_{n,n})$. Положим $h(x) = (x - a_{1,1})(x - a_{2,2}) \dots (x - a_{n,n})$. Необходимо показать, что линейный оператор $h(\mathcal{A}) = (\mathcal{A} - a_{1,1}\mathcal{E})(\mathcal{A} - a_{2,2}\mathcal{E}) \dots (\mathcal{A} - a_{n,n}\mathcal{E})$ является нулевым. Это значит, что надо проверить равенство $h(\mathcal{A})v = 0$ для любого вектора $v \in V$. Для этого (по теореме 1.1.1) достаточно убедиться, что $h(\mathcal{A})v_j = 0$ для всех элементов базиса v_j.

Из (1.7.2) следует, что

$$\mathcal{A}v_j - a_{j,j}v_j = (\mathcal{A} - a_{j,j}\mathcal{E})v_j = a_{1,j}v_1 + \dots + a_{j-1,j}v_{j-1} \in \langle v_1, \dots, v_{j-1} \rangle = V_{j-1}.$$

Так как $(\mathcal{A} - a_{j,j}\mathcal{E})v_k \in V_{j-1}$ при $1 \le k \le j-1$ (по лемме, ввиду инвариантности V_{j-1} относительно $\mathcal{A} - a_{j,j}\mathcal{E}$), то отсюда следует, что $(\mathcal{A} - a_{j,j}\mathcal{E})v \in V_{j-1}$ для каждого $v \in V_j$. Это можно записать в виде $\mathcal{A}V_j \subseteq V_{j-1}$. Заметим, что при $j = 1$ из (1.7.2) следует, что $(\mathcal{A} - a_{1,1}\mathcal{E})v_1 = 0$, и поэтому $(\mathcal{A} - a_{1,1}\mathcal{E})v = 0$ для каждого $v \in V_1 = \langle v_1 \rangle$.

Рассмотрим $h(\mathcal{A})v_j$. Так как $h(\mathcal{A}) = \prod\limits_{k=1}^{n}(\mathcal{A} - a_{k,k}\mathcal{E})$, то $h(\mathcal{A})v_j = (\prod\limits_{k=1}^{j}(\mathcal{A} - a_{k,k}\mathcal{E}))v$, где $v = (\prod\limits_{k=j+1}^{n}(\mathcal{A} - a_{k,k}\mathcal{E}))v_j \in V_j$ (это также следует из леммы). Поскольку произведение здесь — это суперпозиция линейных операторов, то вычисление $h(\mathcal{A})v$ будет состоять в последовательном вычислении $w_1 = (\mathcal{A} - a_{j,j}\mathcal{E})v \in V_{j-1}$, $w_2 = (\mathcal{A} - a_{j-1,j-1}\mathcal{E})w_1 \in V_{j-2}$, $w_3 = (\mathcal{A} - a_{j-2,j-2}\mathcal{E})w_2 \in V_{j-3}$, $\dots$, $w_j = (\mathcal{A} - a_{1,1}\mathcal{E})w_{j-1} = 0$. $\square$

Если проанализировать доказательство теоремы Гамильтона-Кэли, а также доказательства тех предшествущих теорем, которые в нем используются, то станет ясно, что условие $K = \mathbb{C}$ можно заменить на более слабое условие: характеристический многочлен $\chi_{\mathcal{A}}(x)$ должен раскладываться на линейные множители над полем K.

Следствие 1.7.2. *Пусть V — векторное пространство над $\mathbb{R}$, $\mathcal{A} : V \to V$ — линейный оператор, $\chi_{\mathcal{A}}(x)$ — его характеристический многочлен. Тогда $\chi_{\mathcal{A}}(\mathcal{A}) = 0$.*

Доказательство основано на том, что, во-первых, можно свести вопрос о линейном операторе к вопросу о его матрице. Как уже отмечено, $f(M_{\mathcal{A}}) = M_{f(\mathcal{A})}$ для любого многочлена f, и поэтому, если A есть матрица $\mathcal{A}$, то $\chi_{\mathcal{A}}(A) = \chi_{A}(A)$ есть матрица оператора $\chi_{\mathcal{A}}(\mathcal{A})$, и она равна нулю тогда и только тогда, если равен нулю оператор сам $\chi_{\mathcal{A}}(\mathcal{A})$.

Во вторых, матрицу A с действительными компонентами можно рассматривать как матрицу некоторого оператора, действующего в $n-$мерном векторном пространстве над полем *комплексных* чисел. А к такому оператору применима доказанная выше теорема Гамильтона-Кэли. □

Существуют другие доказательства теоремы Гамильтона-Кэли, пригодные для любого поля K. Одно из таких доказательств можно найти в классической книге Ф.Р. Гантмахера «Теория матриц» [3, с. 87]. Мы используем приведенное выше доказательство потому, что оно, во-первых, интересно и поучительно само по себе, а во-вторых, другие доказательства (например, то, которое дано у Гантмахера), потребовали бы на лекциях дополнительного времени, что при ограниченном количестве лекций едва ли возможно.

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] Гантмахер Ф.Р. Теория матриц. — 4-е изд. — М.: Наука. Гл. ред. физ.-мат. лит., 1988.

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

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

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

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

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

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

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

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

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

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

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

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

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