§ 7. Действия над линейными отображениями

Б.М.Верников

Уральский федеральный университет,

Институт естественных наук и математики,

кафедра алгебры и фундаментальной информатики

7.1. Линейные операции над линейными отображениями

Определения

Пусть V и W — векторные пространства над полем F. Обозначим множество всех линейных отображений из V в W через $\mathrm{Hom}(V, W)$, а множество всех линейных операторов в пространстве V — через $\mathrm{Hom}(V)$. Пусть $\mathcal{A}, \mathcal{B} \in \mathrm{Hom}(V, W)$. *Суммой* отображений $\mathcal{A}$ и $\mathcal{B}$ называется отображение $\mathcal{S}$ из V в W, задаваемое правилом $\mathcal{S}(\mathbf{x}) = \mathcal{A}(\mathbf{x}) + \mathcal{B}(\mathbf{x})$ для всякого $\mathbf{x} \in V$. Сумма отображений $\mathcal{A}$ и $\mathcal{B}$ обозначается через $\mathcal{A} + \mathcal{B}$. Далее, *произведением отображения* $\mathcal{A}$ *на скаляр* $t \in F$ называется отображение $\mathcal{C}$ из V в W, задаваемое правилом $\mathcal{C}(\mathbf{x}) = t\mathcal{A}(\mathbf{x})$ для всякого $\mathbf{x} \in V$. Произведение отображения $\mathcal{A}$ на скаляр t обозначается через $t\mathcal{A}$. Операции сложения линейных отображений и умножения линейного отображения на скаляр часто объединяют термином *линейные операции над линейными отображениями*.

Поскольку линейный оператор — частный случай линейного отображения, сумма линейных операторов и произведение линейного оператора на скаляр не нуждаются в специальных определениях.

Предложение 7.1

Сумма линейных отображений и произведение линейного отображения на скаляр являются линейными отображениями. Множество $\mathrm{Hom}(V, W)$ *с операциями сложения операторов и умножения оператора на скаляр является векторным пространством. В частности,* $\langle \mathrm{Hom}(V, W); + \rangle$ *— абелева группа.* $\square$

Мы не приводим доказательство этого утверждения, поскольку оно очевидным образом вытекает из определений. Отметим только, что роль нулевого вектора пространства линейных отображений играет нулевое отображение.

Изоморфизм векторных пространств линейных отображений и матриц (1)

Теорема 7.1

Если V и W — векторные пространства над полем F, $\dim V = n$ и $\dim W = m$, то векторные пространства $\mathrm{Hom}(V, W)$ и $F^{m \times n}$ изоморфны.

Доказательство. Зафиксируем в V базис $P = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$, а в W — базис $Q = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_m\}$. Определим отображение φ из пространства $\mathrm{Hom}(V, W)$ в пространство $F^{m \times n}$ правилом: если $\mathcal{A}$ — линейное отображение из V в W, то $\varphi(\mathcal{A})$ — матрица отображения A в базисах P и Q. Пусть $\mathcal{A}, \mathcal{B} \in \mathrm{Hom}(V, W)$ и $t \in F$. Надо проверить, что отображение φ биективно и выполнены равенства

$$\varphi(\mathcal{A} + \mathcal{B}) = \varphi(\mathcal{A}) + \varphi(\mathcal{B}) \text{ и } \varphi(t\mathcal{A}) = t\varphi(\mathcal{A}).$$

В матрице $\varphi(\mathcal{A} + \mathcal{B})$ по столбцам записаны координаты векторов вида $(\mathcal{A} + \mathcal{B})(\mathbf{p}_i)$ в базисе Q, а в матрицах $\varphi(\mathcal{A})$ и $\varphi(\mathcal{B})$ — координаты векторов $\mathcal{A}(\mathbf{p}_i)$ и $\mathcal{B}(\mathbf{p}_i)$ соответственно в том же базисе ($i = 1, 2, \ldots, n$). Поскольку $(\mathcal{A} + \mathcal{B})(\mathbf{p}_i) = \mathcal{A}(\mathbf{p}_i) + \mathcal{B}(\mathbf{p}_i)$, координаты вектора $(\mathcal{A} + \mathcal{B})(\mathbf{p}_i)$ равны сумме координат векторов $\mathcal{A}(\mathbf{p}_i)$ и $\mathcal{B}(\mathbf{p}_i)$. Первое из требуемых равенств доказано. Второе проверяется вполне аналогично.

Проверим, что отображение φ биективно. Если $\mathcal{A}, \mathcal{B} \in \mathrm{Hom}(V, W)$ и $\varphi(\mathcal{A}) = \varphi(\mathcal{B})$, то из определения матрицы линейного отображения в базисах вытекает, что отображения $\mathcal{A}$ и $\mathcal{B}$ одинаково действуют на базисных векторах пространства V. Но тогда $\mathcal{A} = \mathcal{B}$, так как линейное отображение однозначно определяется своим действием на базисных векторах (см. теорему 6.1). Следовательно, отображение φ инъективно.

Осталось доказать, что φ сюръективно. Пусть $A = (a_{ij})$ — произвольная матрица размера $m \times n$. Для всякого $j = 1, 2, \ldots, n$ положим $\mathbf{w}_j = a_{1j}\mathbf{p}_1 + a_{2j}\mathbf{p}_2 + \cdots + a_{nj}\mathbf{p}_n$. В силу теоремы 6.1 существует линейное отображение $\mathcal{A}$ из V в W такое, что $\mathcal{A}(\mathbf{p}_i) = \mathbf{w}_i$ для всех $i = 1, 2, \ldots, n$. Из определения матрицы отображения в базисах вытекает, что $A_{P,Q} = A$, т. е. $\varphi(\mathcal{A}) = A$. Следовательно, отображение φ сюръективно. $\square$

Из теоремы 7.1 непосредственно вытекает следующее утверждение.

Следствие 7.1

Если V — n-мерное векторное пространство над полем F, то векторные пространства $\mathrm{Hom}(V)$ *и* $F^{n\times n}$ *изоморфны.* □

Как отмечалось в § 2, размерность пространства $F^{m\times n}$ равна mn. Поэтому из теоремы 7.1 вытекает следующее утверждение.

Следствие 7.2

Если V и W — векторные пространства, $\dim V = n$ *и* $\dim W = m$*, то* $\dim \mathrm{Hom}(V, W) = mn$ *и* $\dim \mathrm{Hom}(V) = n^2$. □

7.2. Произведение линейных отображений. Обратное отображение

Пусть U, V и W — векторные пространства над одним и тем же полем F, $\mathcal{A}$ — линейное отображение из U в V, а $\mathcal{B}$ — линейное отображение из V в W. Тогда можно рассмотреть произведение отображений $\mathcal{A}$ и $\mathcal{B}$ в обычном смысле этого слова. Легко проверяется, что $\mathcal{AB}$ — линейное отображение из U в W. В самом деле, если $\mathbf{x}, \mathbf{y} \in U$, а $t \in F$, то

$$\begin{aligned}(\mathcal{AB})(\mathbf{x}+\mathbf{y}) &= \mathcal{B}\big(\mathcal{A}(\mathbf{x}+\mathbf{y})\big) = \mathcal{B}\big(\mathcal{A}(\mathbf{x})+\mathcal{A}(\mathbf{y})\big) = \mathcal{B}\big(\mathcal{A}(\mathbf{x})\big)+\mathcal{B}\big(\mathcal{A}(\mathbf{y})\big) = \\ &= (\mathcal{AB})(\mathbf{x})+(\mathcal{AB})(\mathbf{y}), \\ (\mathcal{AB})(t\mathbf{x}) &= \mathcal{B}\big(\mathcal{A}(t\mathbf{x})\big) = \mathcal{B}\big(t\mathcal{A}(\mathbf{x})\big) = t\mathcal{B}\big(\mathcal{A}(\mathbf{x})\big) = t(\mathcal{AB})(\mathbf{x}).\end{aligned}$$

Лемма 7.1

Пусть U, V и W — векторные пространства над полем F, P, Q и R — базисы пространств U, V и W соответственно, $\mathcal{A}$ — линейное отображение из U в V, а $\mathcal{B}$ — линейное отображение из V в W. Если отображение $\mathcal{A}$ имеет в базисах P и Q матрицу A, а отображение $\mathcal{B}$ имеет в базисах Q и R матрицу B, то отображение $\mathcal{AB}$ имеет в базисах P и R матрицу BA.

Произведение линейных отображений (2)

Доказательство. Используя формулу (1) из §6 и учитывая, что $A = A_{P,Q}$, а $B = B_{Q,R}$, имеем

$$\big[(\mathcal{AB})(\mathbf{x})\big]_R = \big[\mathcal{B}(\mathcal{A}(\mathbf{x}))\big]_R = B \cdot \big[\mathcal{A}(\mathbf{x})\big]_Q = BA \cdot [\mathbf{x}]_P.$$

Обозначим матрицу отображения $\mathcal{AB}$ в базисах P и R через C. В силу формулы (1) из §6 $\big[(\mathcal{AB})(\mathbf{x})\big]_R = C \cdot [\mathbf{x}]_P$. Таким образом, $C \cdot [\mathbf{x}]_P = BA \cdot [\mathbf{x}]_P$. Поскольку это равенство выполнено для любого столбца $[\mathbf{x}]_P$, из ослабленного закона сокращения для матриц вытекает, что $C = BA$. $\square$

Определение

Линейные операторы $\mathcal{A}$ и $\mathcal{B}$ называются *перестановочными*, если $\mathcal{AB} = \mathcal{BA}$.

Замечание 7.1

Линейные операторы $\mathcal{A}$ и $\mathcal{B}$ перестановочны тогда и только тогда, когда их матрицы в некотором базисе перестановочны.

Доказательство. *Необходимость* следует из леммы 7.1, а *достаточность* — из того, что линейный оператор однозначно определяется своей матрицей. $\square$

Если линейное отображение $\mathcal{A}$ из пространства V в пространство W инъективно, то существует обратное к нему отображение, которое обозначается через $\mathcal{A}^{-1}$.

Замечание 7.2

Пусть V и W — векторные пространства над полем F, а $\mathcal{A}$ — линейное отображение из V в W. Если отображение, обратное $\mathcal{A}$, существует, то оно линейно.

Доказательство. Пусть $\mathbf{x}, \mathbf{y} \in W$ и $t \in F$. Положим $\mathbf{x}' = \mathcal{A}^{-1}(\mathbf{x})$ и $\mathbf{y}' = \mathcal{A}^{-1}(\mathbf{y})$. Тогда $\mathcal{A}(\mathbf{x}' + \mathbf{y}') = \mathcal{A}(\mathbf{x}') + \mathcal{A}(\mathbf{y}') = \mathbf{x} + \mathbf{y}$, и потому $\mathcal{A}^{-1}(\mathbf{x} + \mathbf{y}) = \mathbf{x}' + \mathbf{y}' = \mathcal{A}^{-1}(\mathbf{x}) + \mathcal{A}^{-1}(\mathbf{y})$. Далее, $\mathcal{A}(t\mathbf{x}') = t\mathcal{A}(\mathbf{x}') = t\mathbf{x}$, и потому $\mathcal{A}^{-1}(t\mathbf{x}) = t\mathbf{x}' = t\mathcal{A}^{-1}(\mathbf{x})$. □

Из леммы 7.1 вытекает

Замечание 7.3

Если линейный оператор $\mathcal{A}$ имеет в некотором базисе F матрицу A и существует оператор $\mathcal{A}^{-1}$, то $\mathcal{A}^{-1}$ имеет в базисе F матрицу A^{-1}. □

Невырожденный линейный оператор

Определение

Если оператор, обратный к $\mathcal{A}$ существует, то оператор $\mathcal{A}$ называется *обратимым* или *невырожденным*.

Следующее утверждение, вытекающее из замечания 7.3, объясняет происхождение термина «невырожденный оператор».

Лемма 7.2

Для линейного оператора $\mathcal{A}$ следующие условия эквивалентны:

а) *оператор $\mathcal{A}$ невырожден;*

б) *матрица оператора $\mathcal{A}$ в некотором базисе невырожденна;*

в) *матрица оператора $\mathcal{A}$ в любом базисе невырожденна.* $\square$

7.3. Линейные операторы и многочлены

Определение

Для любого оператора $\mathcal{A} \in \mathrm{Hom}(V)$ и любого натурального n определим по индукции оператор $\mathcal{A}^n$ в V: $\mathcal{A}^1 = \mathcal{A}$, а если $n > 1$, то $\mathcal{A}^n(\mathbf{x}) = \mathcal{A}\big(\mathcal{A}^{n-1}(\mathbf{x})\big)$ для любого вектора $\mathbf{x} \in V$. Кроме того, положим $\mathcal{A}^0 = \mathcal{E}$. Это позволяет определить *значение многочлена от линейного оператора* подобно тому, как это было сделано в § 14 в курсе «Основы алгебры» для квадратных матриц: если $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in F[x]$, то, по определению,

$$f(\mathcal{A}) = a_n\mathcal{A}^n + a_{n-1}\mathcal{A}^{n-1} + \cdots + a_0\mathcal{E}.$$

Предложение 7.2

Если линейный оператор $\mathcal{A}$ имеет в некотором базисе матрицу A, то оператор $f(\mathcal{A})$ имеет в том же базисе матрицу $f(A)$.

Доказательство. Из доказательства теоремы 7.1 вытекает, что если операторы $\mathcal{X}$ и $\mathcal{Y}$ имеют в некотором базисе матрицы X и Y соответственно, а t — произвольный скаляр, то операторы $\mathcal{X} + \mathcal{Y}$ и $t\mathcal{X}$ имеют в том же базисе матрицы $X + Y$ и tX соответственно.

Поэтому достаточно доказать, что если n — натуральное число, то оператор $\mathcal{A}^n$ имеет в базисе P матрицу A^n. Обозначим матрицу оператора $\mathcal{A}^n$ в базисе P через B. В силу формулы (2) из § 6 $[\mathcal{A}^n(\mathbf{x})]_P = B \cdot [\mathbf{x}]_P$. С другой стороны, с помощью очевидной индукции по n можно доказать, что $[\mathcal{A}^n(\mathbf{x})]_P = A^n \cdot [\mathbf{x}]_P$. Таким образом, $B \cdot [\mathbf{x}]_P = A^n \cdot [\mathbf{x}]_P$. Поскольку это равенство выполнено для любого вектора $\mathbf{x}$, из ослабленного закона сокращения для матриц вытекает, что $B = A^n$. $\square$

Введем полезное для дальнейшего понятие.

Определение

Квадратные матрицы A и B над некоторым полем F называются *подобными*, если существует невырожденная квадратная матрица T над F такая, что $B = T^{-1}AT$.

Таким образом, матрицы одного и того же линейного оператора в двух различных базисах подобны.

Отметим, что отношение подобия на множестве всех квадратных матриц одного и того же порядка является отношением эквивалентности.

Многочлены, аннулирующие линейный оператор (1)

Определение

Пусть $\mathcal{A}$ — линейный оператор в векторном пространстве V. Говорят, что многочлен f *аннулирует* оператор $\mathcal{A}$, если $f(\mathcal{A}) = \mathcal{O}$.

Предложение 7.3

Для многочлена $f(x)$ следующие условия эквивалентны:

а) *$f(x)$ аннулирует линейный оператор $\mathcal{A}$;*

б) *$f(x)$ аннулирует матрицу линейного оператора $\mathcal{A}$ в некотором базисе;*

в) *$f(x)$ аннулирует матрицу линейного оператора $\mathcal{A}$ в любом базисе.*

Доказательство. Эквивалентность условий а) и в) вытекает из предложения 7.2, а импликация в) $\Longrightarrow$ б) очевидна. Осталось доказать импликацию б) $\Longrightarrow$ в). Матрицы оператора в двух разных базисах подобны. Поэтому достаточно убедиться в том, что если матрицы A и B подобны и $f(A) = O$, то $f(B) = O$. В самом деле, пусть матрицы A и B подобны, т. е. $B = T^{-1}AT$ для некоторой невырожденной квадратной матрицы T, и $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$.

Многочлены, аннулирующие линейный оператор (2)

Тогда

$$B^k = \underbrace{(T^{-1}AT)(T^{-1}AT)\cdots(T^{-1}AT)}_{k \text{ скобок}} =$$
$$= T^{-1}A(TT^{-1})A\cdots(TT^{-1})AT = T^{-1}A^kT,$$

для всякого натурального k, и потому

$$\begin{aligned} f(B) &= a_nB^n + a_{n-1}B^{n-1} + \cdots + a_0E = \\ &= a_nT^{-1}A^nT + a_{n-1}T^{-1}A^{n-1}T + \cdots + a_0T^{-1}ET = \\ &= T^{-1}(a_nA^n + a_{n-1}A^{n-1} + \cdots + a_0E)T = T^{-1}f(A)T. \end{aligned}$$

В частности, если $f(A) = O$, то $f(B) = O$. □

В процессе доказательства последнего предложения проверен следующий факт, который пригодится нам в дальнейшем.

Замечание 7.4

Если A и B — квадратные матрицы одного и того же порядка и $B = T^{-1}AT$ для некоторой невырожденной квадратной матрицы T, а k — натуральное число, то $B^k = T^{-1}A^kT$. $\square$

Предложение 7.4

Характеристические многочлены подобных матриц совпадают.

Доказательство. В самом деле, пусть $B = T^{-1}AT$. Используя свойства произведения матриц и свойства определителей, имеем

$$\chi_B(x) = |B - xE| = |T^{-1}AT - xT^{-1}ET| = |T^{-1}AT - T^{-1}(xE)T| =$$
$$= |T^{-1}(A - xE)T| = |T^{-1}| \cdot |A - xE| \cdot |T| =$$
$$= \frac{1}{|T|} \cdot |A - xE| \cdot |T| = |A - xE| = \chi_A(x).$$

Предложение доказано. $\square$

Предложение 7.4 показывает, что если A и B — матрицы одного и того же линейного оператора в разных базисах, то $\chi_A(x) = \chi_B(x)$. Это делает корректным следующее определение.

Определение

Характеристическим многочленом линейного оператора называется характеристический многочлен его матрицы в любом базисе. Характеристический многочлен оператора $\mathcal{A}$ обозначается через $\chi_{\mathcal{A}}(x)$.

Из теоремы Гамильтона–Кэли непосредственно вытекает ее «операторная версия».

Теорема 7.2 (теорема Гамильтона–Кэли для линейных операторов)

Характеристический многочлен произвольного линейного оператора аннулирует этот оператор. $\square$

Минимальный многочлен линейного оператора (1)

Дальнейшие рассуждения во многом повторяют рассуждения, проведенные в конце § 14 в курсе «Основы алгебры».

Теорема Гамильтона–Кэли показывает, что для всякого линейного оператора в n-мерном векторном пространстве существует аннулирующий его многочлен степени n. Но некоторые операторы могут аннулироваться многочленами меньшей степени. Например, единичный оператор, очевидно, аннулируется линейным многочленом $f(x) = x - 1$.

Определение

Минимальным многочленом линейного оператора $\mathcal{A}$ называется унитарный многочлен, который аннулирует $\mathcal{A}$ и имеет наименьшую степени среди многочленов с этим свойством.

Лемма 7.3

Любой линейный оператор имеет минимальный многочлен, и притом только один. $\square$

Мы не приводим доказательство этого утверждения, поскольку оно почти дословно совпадает с доказательством леммы 14.1 из курса «Основы алгебры».

Предложение 7.5

Пусть $\mathcal{A}$ — линейный оператор в векторном пространстве V над полем F, а A — его матрица в некотором базисе. Многочлен f над полем F является минимальным многочленом оператора $\mathcal{A}$ тогда и только тогда, когда он является минимальным многочленом матрицы A.

Доказательство. *Необходимость.* Пусть f — минимальный многочлен оператора $\mathcal{A}$. Тогда f унитарен и $f(\mathcal{A}) = \mathcal{O}$. Из последнего равенства и предложения 7.3 вытекает, что $f(A) = O$. Если $g(A) = O$ для некоторого многочлена g такого, что $\deg g < \deg f$, то, в силу предложения 7.3, $g(\mathcal{A}) = \mathcal{O}$. Но это противоречит тому, что f — минимальный многочлен оператора $\mathcal{A}$.

Достаточность. Пусть f — минимальный многочлен матрицы A. Тогда f унитарен и $f(A) = O$. Из последнего равенства и предложения 7.3 вытекает, что $f(\mathcal{A}) = \mathcal{O}$. Если $g(\mathcal{A}) = \mathcal{O}$ для некоторого многочлена g такого, что $\deg g < \deg f$, то, в силу предложения 7.3, $g(A) = O$. Но это противоречит тому, что f — минимальный многочлен матрицы $\mathcal{A}$. $\square$

Минимальный многочлен линейного оператора (3)

Следующее утверждение указывает важное свойство минимального многочлена. Оно доказывается вполне аналогично предложению 14.2 из курса «Основы алгебры».

Предложение 7.6

Минимальный многочлен произвольного линейного оператора делит любой аннулирующий его многочлен. □

Из этого предложения и теоремы Гамильтона–Кэли для линейных операторов непосредственно вытекает

Следствие 7.3

Минимальный многочлен произвольного линейного оператора делит его характеристический многочлен. □

Явной формулы для вычисления минимального многочлена линейного оператора не существует. В дальнейшем мы укажем один из возможных способов его нахождения.