Глава VIII. Линейные операторы

§ 29. Линейный оператор, матрица оператора в базисе

Б.М.Верников

Уральский федеральный университет,
Институт естественных наук и математики,
кафедра алгебры и фундаментальной информатики

Определение

Пусть V — векторное пространство над полем F. Отображение $\mathcal{A}\colon V \longrightarrow V$ называется *линейным оператором*, если для любых векторов $\mathbf{x}_1, \mathbf{x}_2 \in V$ и любого скаляра $t \in F$ выполняются равенства $\mathcal{A}(\mathbf{x}_1 + \mathbf{x}_2) = \mathcal{A}(\mathbf{x}_1) + \mathcal{A}(\mathbf{x}_2)$ и $\mathcal{A}(t\mathbf{x}_1) = t\mathcal{A}(\mathbf{x}_1)$.

- Относительно первого равенства из определения говорят, что $\mathcal{A}$ *сохраняет сумму векторов*, относительно второго — что $\mathcal{A}$ *сохраняет произведение вектора на скаляр*.
- Как отмечалось в § 22, векторное пространство над полем F можно рассматривать как универсальную алгебру в сигнатуре, состоящей из бинарной операции сложения векторов и набора унарных операций умножения на элементы поля F (по одной операции для каждого $t \in F$). Из определения эндоморфизма универсальной алгебры (см. § 4) вытекает, что линейные операторы — это в точности эндоморфизмы векторного пространства.

Приведем примеры линейных операторов.

Пример 1. Представим пространство $\mathbb{R}_2$ как множество направленных отрезков на плоскости, выходящих из начала координат O. Тогда поворот на угол α, симметрия относительно любой прямой, проходящей через начало координат (в частности, относительно любой из осей координат), симметрия относительно точки O, проекция на любую из осей координат — примеры линейных операторов в пространстве $\mathbb{R}_2$. Если интерпретировать $\mathbb{R}_3$ как множество направленных отрезков трехмерного пространства, выходящих из начала координат O, то симметрия относительно любой прямой или плоскости, проходящей через точку O, симметрия относительно этой точки, проекция на любую из координатных плоскостей — примеры линейных операторов в пространстве $\mathbb{R}_3$.

Пример 2. Зафиксируем произвольный скаляр t и зададим оператор $\mathcal{A}$ следующим правилом: $\mathcal{A}(\mathbf{x}) = t\mathbf{x}$ для всякого вектора $\mathbf{x} \in V$. Этот оператор называется *оператором растяжения в t раз*. Линейность оператора растяжения с очевидностью вытекает из определения векторного пространства (см. § 22). Особо отметим два частных случая оператора растяжения. Первый из них — это оператор растяжения при $t = 0$. Он обозначается буквой $\mathcal{O}$ и называется *нулевым*. Ясно, что нулевой оператор переводит произвольный вектор из V в нулевой вектор. Второй частный случай оператора растяжения возникает при $t = 1$. Соответствующий оператор обозначается буквой $\mathcal{E}$ и называется *тождественным* или *единичным*. Этот оператор переводит произвольный вектор из V в себя.

Пример 3. Зафиксируем в пространстве V некоторое подпространство M. В силу предложения о дополняющем подпространстве (см. § 24) существует подпространство M' в V такое, что $V = M \oplus M'$. Следовательно, произвольный вектор $\mathbf{x} \in V$ можно, и притом единственным образом, представить в виде $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$, где $\mathbf{x}_1 \in M$ и $\mathbf{x}_2 \in M'$ (см. теорему о прямой сумме подпространств в § 24). Рассмотрим оператор $\mathcal{P}$ в пространстве V, задаваемый правилом $\mathcal{P}(\mathbf{x}) = \mathbf{x}_1$. Легко проверяется, что этот оператор — линейный. Он называется *оператором проектирования на подпространство* M *параллельно* M'.

Пример 4. В пространстве $F[x]$ всех многочленов над полем F определим оператор $\mathcal{D}$ правилом: $\mathcal{D}(p) = p'$, где p' — производная многочлена p. Этот оператор называется *оператором дифференцирования*. Из свойств производной вытекает, что этот оператор линеен. Точно так же определяется оператор дифференцирования в пространстве $F_n[x]$ всех многочленов степени $\leqslant n$ над F.

Пример 5. Пусть F — поле, n — натуральное число, а A — квадратная матрица порядка n над полем F. Определим оператор $\mathcal{A}$ в F_n правилом $\mathcal{A}(\mathbf{x}) = \mathbf{x} \cdot A$ для всякого вектора $\mathbf{x} \in F_n$. Вектор $\mathbf{x}$ рассматривается здесь как матрица размера $1 \times n$, и потому $\mathbf{x} \cdot A$ — это обычное произведение матриц, результатом которого является матрица размера $1 \times n$, т. е. вектор из F_n. Из свойств произведения матриц вытекает, что этот оператор линеен (см. § 4 и 25).

Пример 6. Пусть F — поле, n — натуральное числа, а A и B — квадратные матрицы порядка n над полем F. Определим оператор $\mathcal{M}$ в пространстве $F^{n \times n}$ правилом: $\mathcal{M}(X) = AXB$ для всякой матрицы $X \in F^{n \times n}$. Как и в предыдущем примере, линейность этого оператора вытекает из свойств произведения матриц.

Следующими свойствами линейных операторов мы часто будем пользоваться в дальнейшем.

Замечание о свойствах линейного оператора

Пусть V — векторное пространство над полем F, а $\mathcal{A}$ — линейный оператор в V. Тогда:

1) $\mathcal{A}(\mathbf{0}) = \mathbf{0}$.
2) $\mathcal{A}(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \dots + \lambda_m\mathbf{v}_m) = \lambda_1\mathcal{A}(\mathbf{v}_1) + \lambda_2\mathcal{A}(\mathbf{v}_2) + \dots + \lambda_m\mathcal{A}(\mathbf{v}_m)$ *для любых векторов* $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m \in V$ *и любых скаляров* $\lambda_1, \lambda_2, \dots, \lambda_m \in F$.

Доказательство. Первое свойство вытекает из того, что

$$\mathcal{A}(\mathbf{0}) = \mathcal{A}(0 \cdot \mathbf{0}) = 0 \cdot \mathcal{A}(\mathbf{0}) = \mathbf{0}.$$

Второе свойство непосредственно вытекает из определения линейного оператора. $\square$

Теорема существования и единственности линейного оператора

Пусть V — векторное пространство над полем F, $P = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ — базис пространства V, а $\mathbf{w}_1$, $\mathbf{w}_2$, …, $\mathbf{w}_n$ — произвольные векторы из V. Тогда существует, и притом только один, линейный оператор $\mathcal{A}$ в V такой, что $\mathcal{A}(\mathbf{p}_i) = \mathbf{w}_i$ для всех $i = 1, 2, \ldots, n$.

Доказательство. Существование. Пусть $\mathbf{x} \in V$, а $(x_1, x_2, \ldots, x_n)$ — координаты вектора $\mathbf{x}$ в базисе P. Определим оператор $\mathcal{A}$ в V правилом: $\mathcal{A}(\mathbf{x}) = x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \cdots + x_n\mathbf{w}_n$. В силу единственности координат вектора в базисе это определение корректно (т. е. образ вектора $\mathbf{x}$ под действием $\mathcal{A}$ определен однозначно). А из замечания о координатах суммы векторов и произведения вектора на скаляр (см. § 23) вытекает, что этот оператор линеен. Осталось заметить, что, для всякого $i = 1, 2, \ldots, n$, вектор $\mathbf{p}_i$ имеет в базисе P координаты $(0, \ldots, 0, 1, 0, \ldots, 0)$, где 1 стоит на i-м месте, и потому $\mathcal{A}(\mathbf{p}_i) = \mathbf{w}_i$.

Единственность. Пусть $\mathcal{B}$ — линейный оператор в V такой, что $\mathcal{B}(\mathbf{p}_i) = \mathbf{w}_i$ для всех $i = 1, 2, \ldots, n$. Пусть $\mathbf{x} \in V$, а $(x_1, x_2, \ldots, x_n)$ — координаты вектора $\mathbf{x}$ в базисе P. Тогда $\mathbf{x} = x_1\mathbf{p}_1 + x_2\mathbf{p}_2 + \cdots + x_n\mathbf{p}_n$. В силу замечания о свойствах линейного оператора имеем

$$\begin{aligned}\mathcal{B}(\mathbf{x}) = \mathcal{B}(x_1\mathbf{p}_1 + x_2\mathbf{p}_2 + \cdots + x_n\mathbf{p}_n) = x_1\mathcal{B}(\mathbf{p}_1) + x_2\mathcal{B}(\mathbf{p}_2) + \cdots + x_n\mathcal{B}(\mathbf{p}_n) = \\ = x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \cdots + x_n\mathbf{w}_n = \mathcal{A}(\mathbf{x}).\end{aligned}$$

Следовательно, $\mathcal{B} = \mathcal{A}$. $\square$

- Доказанная только что теорема устанавливает, что линейный оператор в V однозначно определяется тем, как он действует на базисных векторах пространства V. Поэтому для того, чтобы иметь полную информацию о линейном операторе, достаточно знать координаты образов этих базисных векторов в каком-нибудь базисе пространства V. Это делает естественным определение матрицы линейного оператора в базисе, которое дано на следующем слайде.

Определение

Пусть $\mathcal{A}$ — линейный оператор в пространстве V, а $P = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ — базис этого пространства. Квадратная матрица порядка n, i-й столбец которой состоит из координат вектора $\mathcal{A}(\mathbf{p}_i)$ в базисе P (для всех $i = 1, 2, \ldots, n$), называется *матрицей оператора $\mathcal{A}$ в базисе P*. Эта матрица обозначается через A_P, а если базис однозначно определяется из контекста или неважен — то просто через A .

Обозначение

Если V — векторное пространство, $\dim V = n$, P — произвольный базис в V, а $\mathbf{x} \in V$, будем обозначать через $[\mathbf{x}]_P$ матрицу размера $n \times 1$ (т. е. столбец длины n), в которой записаны координаты вектора $\mathbf{x}$ в базисе P.

Пусть $\mathcal{A}$ — линейный оператор в пространстве V, который в базисе $P = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ имеет матрицу $A_P = (a_{ij})$, а $\mathbf{x} \in V$. Обозначим координаты вектора $\mathbf{x}$ в базисе P через $(x_1, x_2, \ldots, x_n)$. Как найти координаты вектора $\mathbf{y} = \mathcal{A}(\mathbf{x})$ в базисе P? Обозначим эти координаты через $(y_1, y_2, \ldots, y_n)$. Тогда

$$\begin{aligned} y_1\mathbf{p}_1 + y_2\mathbf{p}_2 + \cdots + y_n\mathbf{p}_n = \mathbf{y} = \mathcal{A}(\mathbf{x}) = \mathcal{A}(x_1\mathbf{p}_1 + x_2\mathbf{p}_2 + \cdots + x_n\mathbf{p}_n) = \\ = x_1\mathcal{A}(\mathbf{p}_1) + x_2\mathcal{A}(\mathbf{p}_2) + \cdots + x_n\mathcal{A}(\mathbf{p}_n). \end{aligned}$$

Поскольку столбцы матрицы A_P — это координаты векторов $\mathcal{A}(\mathbf{p}_1)$, $\mathcal{A}(\mathbf{p}_2)$, $\ldots$, $\mathcal{A}(\mathbf{p}_n)$ в базисе P, выполнены равенства

$$\begin{array}{rcl} \mathcal{A}(\mathbf{p}_1) & = & a_{11}\mathbf{p}_1 + a_{21}\mathbf{p}_2 + \cdots + a_{n1}\mathbf{p}_n, \\ \mathcal{A}(\mathbf{p}_2) & = & a_{12}\mathbf{p}_1 + a_{22}\mathbf{p}_2 + \cdots + a_{n2}\mathbf{p}_n, \\ \multicolumn{3}{c}{\dotfill} \\ \mathcal{A}(\mathbf{p}_n) & = & a_{1n}\mathbf{p}_1 + a_{2n}\mathbf{p}_2 + \cdots + a_{nn}\mathbf{p}_n. \end{array}$$

Следовательно,

$$\begin{aligned}
y_1\mathbf{p}_1 + y_2\mathbf{p}_2 + \cdots + y_n\mathbf{p}_n &= x_1\mathcal{A}(\mathbf{p}_1) + x_2\mathcal{A}(\mathbf{p}_2) + \cdots + x_n\mathcal{A}(\mathbf{p}_n) = \\
&= x_1\big(a_{11}\mathbf{p}_1 + a_{21}\mathbf{p}_2 + \cdots + a_{n1}\mathbf{p}_n\big) + \\
&+ x_2\big(a_{12}\mathbf{p}_1 + a_{22}\mathbf{p}_2 + \cdots + a_{n2}\mathbf{p}_n\big) + \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&+ x_n\big(a_{1n}\mathbf{p}_1 + a_{2n}\mathbf{p}_2 + \cdots + a_{nn}\mathbf{p}_n\big) = \\
&= \big(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n\big)\mathbf{p}_1 + \\
&+ \big(a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n\big)\mathbf{p}_2 + \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&+ \big(a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n\big)\mathbf{p}_n .
\end{aligned}$$

В силу единственности разложения вектора по базису это означает, что

$$\begin{cases}
y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\
y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
y_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n.
\end{cases}$$

Эту систему равенств можно переписать в виде

$$\big[\mathcal{A}(\mathbf{x})\big]_P = A_P \cdot [\mathbf{x}]_P. \qquad (1)$$

Нахождение координат образа вектора с помощью матрицы оператора (3)

Если нужно записать координаты вектора $\mathcal{A}(\mathbf{x})$ в строку, а не в столбец, то можно использовать следующее равенство, которое получится, если транспонировать обе части равенства (1) и воспользоваться свойствами транспонирования матриц:

$$[\mathcal{A}(\mathbf{x})]_P^\top = [\mathbf{x}]_P^\top \cdot A_P^\top. \tag{2}$$

Ясно, что матрицы одного и того же линейного оператора в разных базисах могут быть различными. Чтобы выяснить, как они связаны между собой, нам понадобится понятие матрицы перехода от одного базиса к другому.

Определение

Пусть V — векторное пространство над полем F, а $P = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ и $Q = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ — базисы этого пространства. *Матрицей перехода от базиса P к базису Q* называется квадратная матрица порядка n, в которой, для всякого $i = 1, 2, \ldots, n$, в i-м столбце записаны координаты вектора q_i в базисе P. Эта матрица обозначается через T_{PQ}. При этом базис P мы будем называть *старым*, а базис Q — *новым*.

Отметим, что в случае трехмерного физического пространства матрица перехода от одного базиса к другому уже возникала в § 14.

Пусть P и Q — базисы векторного пространства V. Чтобы найти матрицу T_{PQ}, надо найти координаты каждого из векторов $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ в базисе P. Для этого можно воспользоваться изложенным в § 23 алгоритмом нахождения координат вектора в базисе. В соответствии с этим алгоритмом, чтобы найти координаты вектора $\mathbf{q}_i$ в базисе P (где $1 \leqslant i \leqslant n$), надо решить методом Гаусса–Жордана крамеровскую систему линейных уравнений, у которой в основной матрице по столбцам записаны координаты векторов $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ в некотором базисе, а в последнем столбце — координаты вектора $\mathbf{q}_i$ в том же базисе. Иными словами, надо решить n систем линейных уравнений, которые имеют одну и ту же основную матрицу и отличаются только столбцом свободных членов. Ясно, что при решении всех этих систем почти все преобразования будут одинаковыми, по разному будет меняться лишь последний столбец. Для того, чтобы сэкономить время, можно решать все эти системы одновременно в соответствии с алгоритмом, который сформулирован на следующем слайде.

Алгоритм нахождения матрицы перехода от одного базиса к другому

Чтобы найти матрицу перехода от базиса $P = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ векторного пространства V к базису $Q = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ того же пространства, надо составить матрицу размера $n \times 2n$, записав в ее левую часть (первые n столбцов) координаты векторов $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ в некотором базисе, а в ее правую часть (последние n столбцов) координаты векторов $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ в том же базисе. После этого надо элементарными преобразованиями всей матрицы привести ее левую часть к единичному виду. В этот момент в правой части по столбцам будут записаны координаты векторов $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ в базисе P, т. е. матрица T_{PQ}.

В случае, когда $V = F_n$, а в качестве старого базиса выступает стандартный базис пространства F_n, матрицу перехода от него к любому другому можно выписать сразу, без всяких вычислений. В самом деле, сравнивая определение матрицы перехода от одного базиса к другому с замечанием о компонентах вектора (см. § 23), мы получаем

Замечание о матрице перехода от стандартного базиса к произвольному

Матрицей перехода от стандартного базиса пространства F_n к произвольному базису Q этого пространства является матрица, в которой по столбцам записаны векторы базиса Q. □

Предложение о матрице перехода

Пусть V — векторное пространство, а $P = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ и $Q = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ — базисы этого пространства. Матрица T_{PQ} обратима и обратной к ней является матрица T_{QP}.

Доказательство. Положим $T_{PQ} = (t_{ij})$ и $T_{QP} = (s_{ij})$. Достаточно доказать, что $T_{PQ}T_{QP} = E$. Пусть $T_{PQ}T_{QP} = (x_{ij})$. Используя определение матрицы перехода от одного базиса к другому, имеем

$$\begin{aligned}
\mathbf{p}_j &= s_{1j}\mathbf{q}_1 + s_{2j}\mathbf{q}_2 + \cdots + s_{nj}\mathbf{q}_n = \\
&= s_{1j}(t_{11}\mathbf{p}_1 + t_{21}\mathbf{p}_2 + \cdots + t_{n1}\mathbf{p}_n) + \\
&+ s_{2j}(t_{12}\mathbf{p}_1 + t_{22}\mathbf{p}_2 + \cdots + t_{n2}\mathbf{p}_n) + \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&+ s_{nj}(t_{1n}\mathbf{p}_1 + t_{2n}\mathbf{p}_2 + \cdots + t_{nn}\mathbf{p}_n).
\end{aligned}$$

Раскрывая скобки, перегруппировывая слагаемые и учитывая определение произведения матриц, имеем

$$\begin{aligned}
\mathbf{p}_j &= (t_{11}s_{1j} + t_{12}s_{2j} + \cdots + t_{1n}s_{nj})\mathbf{p}_1 + \\
&+ (t_{21}s_{1j} + t_{22}s_{2j} + \cdots + t_{2n}s_{nj})\mathbf{p}_2 + \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&+ (t_{n1}s_{1j} + t_{n2}s_{2j} + \cdots + t_{nn}s_{nj})\mathbf{p}_n = \\
&= x_{1j}\mathbf{p}_1 + x_{2j}\mathbf{p}_2 + \cdots + x_{nj}\mathbf{p}_n.
\end{aligned}$$

С другой стороны, $\mathbf{p}_j = 0 \cdot \mathbf{p}_1 + \cdots + 0 \cdot p_{j-1} + 1 \cdot \mathbf{p}_j + 0 \cdot p_{j+1} + \cdots + 0 \cdot \mathbf{p}_n$. В силу единственности разложения вектора по базису получаем, что

$$x_{ij} = \begin{cases} 1, & \text{если } i = j, \\ 0, & \text{если } i \neq j. \end{cases}$$

Это и означает, что $T_{PQ} T_{QP} = E$. $\square$

Выясним, как связаны между собой координаты одного и того же вектора в разных базисах.

Лемма о замене координат вектора

Пусть V — векторное пространство, $\mathbf{x} \in V$, а $P = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ и $Q = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ — базисы пространства V. Тогда

$$[\mathbf{x}]_P = T_{PQ}[\mathbf{x}]_Q. \tag{3}$$

Доказательство. Обозначим координаты вектора $\mathbf{x}$ в базисе P через $(x_1, x_2, \ldots, x_n)$, а его координаты в базисе Q — через $(x'_1, x'_2, \ldots, x'_n)$. Положим $T_{PQ} = (t_{ij})$.

Вычислим двумя способами вектор $\mathbf{x}$. С одной стороны, $\mathbf{x} = x_1\mathbf{p}_1 + x_2\mathbf{p}_2 + \cdots + x_n\mathbf{p}_n$. С другой,

$$\begin{aligned}
\mathbf{x} &= x_1'\mathbf{q}_1 + x_2'\mathbf{q}_2 + \cdots + x_n'\mathbf{q}_n = \\
&= x_1'(t_{11}\mathbf{p}_1 + t_{21}\mathbf{p}_2 + \cdots + t_{n1}\mathbf{p}_n) + \\
&+ x_2'(t_{12}\mathbf{p}_1 + t_{22}\mathbf{p}_2 + \cdots + t_{n2}\mathbf{p}_n) + \\
&\quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&+ x_n'(t_{1n}\mathbf{p}_1 + t_{2n}\mathbf{p}_2 + \cdots + t_{nn}\mathbf{p}_n) = \\
&= (t_{11}x_1' + t_{12}x_2' + \cdots + t_{1n}x_n')\mathbf{p}_1 + \\
&+ (t_{21}x_1' + t_{22}x_2' + \cdots + t_{2n}x_n')\mathbf{p}_2 + \\
&\quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&+ (t_{n1}x_1' + t_{n2}x_2' + \cdots + t_{nn}x_n')\mathbf{p}_n.
\end{aligned}$$

В силу единственности разложения вектора по базису, имеем

$$\begin{cases}
x_1 = t_{11}x_1' + t_{12}x_2' + \cdots + t_{1n}x_n', \\
x_2 = t_{21}x_1' + t_{22}x_2' + \cdots + t_{2n}x_n', \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
x_n = t_{n1}x_1' + t_{n2}x_2' + \cdots + t_{nn}x_n'.
\end{cases}$$

Ясно, что этот набор равенств эквивалентен матричному равенству (3). □

Теперь мы уже готовы доказать следующее утверждение.

Теорема о замене матрицы оператора

Пусть V — векторное пространство, $\mathcal{A}$ — линейный оператор в V, а $P = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ и $Q = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ — базисы пространства V. Тогда

$$A_Q = T_{QP} A_P T_{PQ}. \tag{4}$$

Доказательство. В силу формулы (1) выполнены равенства $[\mathcal{A}(\mathbf{x})]_P = A_P \cdot [\mathbf{x}]_P$ и $[\mathcal{A}(\mathbf{x})]_Q = A_Q \cdot [\mathbf{x}]_Q$. Кроме того, в силу формулы (3), верно равенство $[\mathbf{x}]_P = T_{PQ} \cdot [\mathbf{x}]_Q$. Следовательно,

$$[\mathcal{A}(\mathbf{x})]_P = A_P \cdot [\mathbf{x}]_P = A_P T_{PQ} \cdot [\mathbf{x}]_Q.$$

С другой стороны, из равенства (3), примененного к вектору $\mathcal{A}(\mathbf{x})$, и равенства (1), примененного к базису Q, вытекает, что

$$[\mathcal{A}(\mathbf{x})]_P = T_{PQ}[\mathcal{A}(\mathbf{x})]_Q = T_{PQ}A_Q \cdot [\mathbf{x}]_Q.$$

Сравнивая две последние цепочки равенств, получаем, что

$$A_P T_{PQ} \cdot [\mathbf{x}]_Q = T_{PQ}A_Q \cdot [\mathbf{x}]_Q.$$

Ясно, что любая матрица размера $n \times 1$, где $n = \dim V$, совпадает со столбцом вида $[\mathbf{x}]_Q$ для подходящего вектора $\mathbf{x}$. Поэтому мы можем применить ослабленный закон сокращения для матриц (см. § 25) и получить, что $A_P T_{PQ} = T_{PQ}A_Q$. Принимая во внимание предложение о матрице перехода, умножим обе части последнего равенство слева на матрицу $T_{PQ}^{-1} = T_{QP}$. В результате получится равенство (4). $\square$

Положим $A = A_P$, $A' = A_Q$ и $T = T_{PQ}$. Тогда, с учетом предложения о матрице перехода, формулу (4) можно переписать в виде $A' = T^{-1}AT$.

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 любых двух базисах подобны.* $\square$

Лемма об отношении подобия матриц

Отношение подобия на множестве всех квадратных матриц одного и того же порядка является отношением эквивалентности.

Доказательство. В самом деле, это отношение рефлексивно (т. е. всякая квадратная матрица подобна сама себе), поскольку $A = EAE = E^{-1}AE$, симметрично, поскольку умножая равенство $B = T^{-1}AT$ на T слева и на T^{-1} справа, мы получим $A = TBT^{-1} = (T^{-1})^{-1}BT^{-1}$, и транзитивно, поскольку из равенств $B = T^{-1}AT$ и $C = S^{-1}BS$ вытекает, что $C = S^{-1}T^{-1}ATS = (TS)^{-1}A(TS)$. $\square$

Оставшаяся часть параграфа посвящена рассмотрению действий над линейными операторами.

Определение

Пусть V — векторное пространство над полем F, а $\mathcal{A}$ и $\mathcal{B}$ — линейные операторы в V. *Суммой* операторов $\mathcal{A}$ и $\mathcal{B}$ называется оператор $\mathcal{S}$ в V, задаваемый правилом $\mathcal{S}(\mathbf{x}) = \mathcal{A}(\mathbf{x}) + \mathcal{B}(\mathbf{x})$ для всякого $\mathbf{x} \in V$. Сумма операторов $\mathcal{A}$ и $\mathcal{B}$ обозначается через $\mathcal{A} + \mathcal{B}$.

Множество всех линейных операторов в V обозначим через $\mathrm{Hom}(V)$.

Предложение о свойствах суммы операторов

Сумма линейных операторов является линейным оператором. Множество $\mathrm{Hom}(V)$ *с операцией сложения операторов является абелевой группой.*

Доказательство. Пусть $\mathcal{A}, \mathcal{B} \in \mathrm{Hom}(V)$ и $\mathcal{S} = \mathcal{A} + \mathcal{B}$. Для любых $\mathbf{x}, \mathbf{y} \in V$ и $t \in F$ имеем

$$\begin{aligned} \mathcal{S}(\mathbf{x}+\mathbf{y}) &= \mathcal{A}(\mathbf{x}+\mathbf{y}) + \mathcal{B}(\mathbf{x}+\mathbf{y}) = \mathcal{A}(\mathbf{x}) + \mathcal{A}(\mathbf{y}) + \mathcal{B}(\mathbf{x}) + \mathcal{B}(\mathbf{y}) = \\ &= \big(\mathcal{A}(\mathbf{x}) + \mathcal{B}(\mathbf{x})\big) + \big(\mathcal{A}(\mathbf{y}) + \mathcal{B}(\mathbf{y})\big) = \mathcal{S}(\mathbf{x}) + \mathcal{S}(\mathbf{y}) \quad \text{и} \\ \mathcal{S}(t\mathbf{x}) &= \mathcal{A}(t\mathbf{x}) + \mathcal{B}(t\mathbf{x}) = t\mathcal{A}(\mathbf{x}) + t\mathcal{B}(\mathbf{x}) = t\big(\mathcal{A}(\mathbf{x}) + \mathcal{B}(\mathbf{x})\big) = t\mathcal{S}(\mathbf{x}). \end{aligned}$$

Следовательно, оператор $\mathcal{S}$ линеен.

Далее, если $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathrm{Hom}(V)$, то

$$
\begin{aligned}
(\mathcal{A}+\mathcal{B})(\mathbf{x}) &= \mathcal{A}(\mathbf{x})+\mathcal{B}(\mathbf{x}) = \mathcal{B}(\mathbf{x})+\mathcal{A}(\mathbf{x}) = (\mathcal{B}+\mathcal{A})(\mathbf{x}) \quad \text{и} \\
((\mathcal{A}+\mathcal{B})+\mathcal{C})(\mathbf{x}) &= (\mathcal{A}+\mathcal{B})(\mathbf{x})+\mathcal{C}(\mathbf{x}) = \big(\mathcal{A}(\mathbf{x})+\mathcal{B}(\mathbf{x})\big)+\mathcal{C}(\mathbf{x}) = \\
&= \mathcal{A}(\mathbf{x})+\big(\mathcal{B}(\mathbf{x})+\mathcal{C}(\mathbf{x})\big) = \mathcal{A}(\mathbf{x})+(\mathcal{B}+\mathcal{C})(\mathbf{x}) = \\
&= \big(\mathcal{A}+(\mathcal{B}+\mathcal{C})\big)(\mathbf{x}),
\end{aligned}
$$

откуда $\mathcal{A}+\mathcal{B}=\mathcal{B}+\mathcal{A}$ и $(\mathcal{A}+\mathcal{B})+\mathcal{C}=\mathcal{A}+(\mathcal{B}+\mathcal{C})$. Нейтральным элементом по сложению является нулевой оператор $\mathcal{O}$, поскольку

$$(\mathcal{A}+\mathcal{O})(\mathbf{x}) = \mathcal{A}(\mathbf{x})+\mathcal{O}(\mathbf{x}) = \mathcal{A}(\mathbf{x})+\mathbf{0} = \mathcal{A}(\mathbf{x}).$$

Обратным по сложению элементом к оператору $\mathcal{A} \in \mathrm{Hom}(V)$ является оператор $-\mathcal{A}$, определяемый правилом $(-\mathcal{A})(\mathbf{x}) = -\mathcal{A}(\mathbf{x})$, поскольку

$$
\begin{aligned}
\big(\mathcal{A}+(-\mathcal{A})\big)(\mathbf{x}) &= \mathcal{A}(\mathbf{x})+(-\mathcal{A})(\mathbf{x}) = \mathcal{A}(\mathbf{x})+\big(-\mathcal{A}(\mathbf{x})\big) = \\
&= \mathcal{A}(\mathbf{x})-\mathcal{A}(\mathbf{x}) = \mathbf{0} = \mathcal{O}(\mathbf{x}).
\end{aligned}
$$

Предложение доказано. $\square$

Определение

Пусть V — векторное пространство над полем F, $\mathcal{A}$ — линейный оператор в V, а $t \in F$. *Произведением оператора* $\mathcal{A}$ *на скаляр* t называется оператор $\mathcal{B}$ в V, задаваемый правилом $\mathcal{B}(\mathbf{x}) = t\mathcal{A}(\mathbf{x})$ для всякого $\mathbf{x} \in V$. Произведение оператора $\mathcal{A}$ на скаляр t обозначается через $t\mathcal{A}$.

Предложение о пространстве линейных операторов

Произведение линейного оператора на скаляр является линейным оператором. Множество $\mathrm{Hom}(V)$ *с операциями сложения операторов и умножения оператора на скаляр является векторным пространством.*

Доказательство. Пусть $\mathcal{A}, \mathcal{B} \in \mathrm{Hom}(V)$, $\mathbf{x}, \mathbf{y} \in V$ и $t, s \in F$. Тогда:

$$(t\mathcal{A})(\mathbf{x}+\mathbf{y}) = t\bigl(\mathcal{A}(\mathbf{x}+\mathbf{y})\bigr) = t(\mathcal{A}(\mathbf{x})+\mathcal{A}(\mathbf{y})) = t\mathcal{A}(\mathbf{x}) + t\mathcal{A}(\mathbf{y}) =$$
$$= (t\mathcal{A})(\mathbf{x}) + (t\mathcal{A})(\mathbf{y}) \text{ и}$$
$$(t\mathcal{A})(s\mathbf{x}) = t\bigl(\mathcal{A}(s\mathbf{x})\bigr) = t\bigl(s\mathcal{A}(\mathbf{x})\bigr) = (ts)\bigl(\mathcal{A}(\mathbf{x})\bigr) = s\bigl(t\mathcal{A}(\mathbf{x})\bigr) = s\bigl((t\mathcal{A})(\mathbf{x})\bigr).$$

Следовательно, $t\mathcal{A}$ — линейный оператор.

Далее,

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

т. е. $t(\mathcal{A}+\mathcal{B}) = t\mathcal{A}+t\mathcal{B}$;

$$\begin{aligned}\big((t+s)\mathcal{A}\big)(\mathbf{x}) &= (t+s)\mathcal{A}(\mathbf{x}) = t\mathcal{A}(\mathbf{x})+s\mathcal{A}(\mathbf{x}) = \\ &= (t\mathcal{A})(\mathbf{x})+(s\mathcal{A})(\mathbf{x}) = (t\mathcal{A}+s\mathcal{A})(\mathbf{x}),\end{aligned}$$

т. е. $t\mathcal{A}+s\mathcal{A} = (t+s)\mathcal{A}$;

$$\big(t(s\mathcal{A})\big)(\mathbf{x}) = t\big((s\mathcal{A})(\mathbf{x})\big) = (ts)\big(\mathcal{A}(\mathbf{x})\big) = \big((ts)\mathcal{A}\big)(\mathbf{x}),$$

т. е. $t(s\mathcal{A}) = (ts)\mathcal{A}$; наконец,

$$(1\cdot\mathcal{A})(\mathbf{x}) = 1\cdot\big(\mathcal{A}(\mathbf{x})\big) = \mathcal{A}(\mathbf{x}),$$

т. е. $1\cdot\mathcal{A} = \mathcal{A}$. С учетом предложения о свойствах суммы операторов, мы получаем, что в $\mathrm{Hom}(V)$ выполнены все аксиомы векторного пространства. $\square$

Теорема о пространствах линейных операторов и матриц

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

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

$$\varphi(\mathcal{A}+\mathcal{B}) = \varphi(\mathcal{A}) + \varphi(\mathcal{B}) \text{ и } \varphi(t\mathcal{A}) = t\varphi(\mathcal{A}). \tag{5}$$

В матрице $\varphi(\mathcal{A}+\mathcal{B})$ по столбцам записаны координаты векторов вида $(\mathcal{A}+\mathcal{B})(\mathbf{p}_i)$ в базисе P, 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 равенств (5) доказано. Второе из них проверяется вполне аналогично.

Проверим, что отображение φ биективно. Если $\mathcal{A}, \mathcal{B} \in \mathrm{Hom}(V)$ и $\varphi(\mathcal{A}) = \varphi(\mathcal{B})$, то из определения матрицы линейного оператора вытекает, что операторы $\mathcal{A}$ и $\mathcal{B}$ одинаково действуют на базисных векторах пространства V. Но тогда $\mathcal{A} = \mathcal{B}$, так как линейный оператор однозначно определяется своим действием на базисных векторах. Следовательно, отображение φ инъективно.

Осталось доказать, что φ сюръективно. Пусть $A = (a_{ij})$ — произвольная квадратная матрица порядка 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 существования и единственности линейного оператора существует линейный оператор $\mathcal{A}$ в V такой, что $\mathcal{A}(\mathbf{p}_i) = \mathbf{w}_i$ для всех $i = 1, 2, \ldots, n$. Из определения матрицы оператора в базисе вытекает, что $A_P = A$, т.е. $\varphi(\mathcal{A}) = A$. Следовательно, отображение φ сюръективно. □

Как отмечалось в § 23, размерность пространства матриц размера $m \times n$ равна mn. Поэтому из доказанной теоремы вытекает

Следствие о размерности пространства линейных операторов

Если V — векторное пространство и $\dim V = n$, то $\dim \mathrm{Hom}(V) = n^2$. □

Для любого оператора $\mathcal{A} \in \mathrm{Hom}(V)$ и любого натурального n определим по индукции оператор $\mathcal{A}^n$: $\mathcal{A}^1 = \mathcal{A}$, а если $n > 1$, то $\mathcal{A}^n(\mathbf{x}) = \mathcal{A}(\mathcal{A}^{n-1}(\mathbf{x}))$ 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 § 25 для квадратных матриц: если $f(x) = a_n x^n + a_{n-1}x^{n-1} + \dots + a_0 \in F[x]$, то, по определению,

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

Предложение о матрице оператора $f(\mathcal{A})$

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

Доказательство. Из доказательства теоремы о пространствах линейных операторов и матриц 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 (1), $[\mathcal{A}^n(\mathbf{x})]_P = B \cdot [\mathbf{x}]_P$. 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 $B = A^n$. $\square$

Определение

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

Предложение о многочленах, аннулирующих линейный оператор

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 о матрице оператора $f(\mathcal{A})$, 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 $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$.

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 $f(B) = O$. $\square$

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

Замечание о степенях подобных матриц

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 о характеристических многочленах подобных матриц

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. □

Это предложение 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

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 (см. § 25) непосредственно вытекает ее «операторная версия».

Теорема Гамильтона–Кэли для линейных операторов

Характеристический многочлен произвольного линейного оператора аннулирует этот оператор. □