Глава II. Линейные отображения и линейные операторы

§ 6. Линейное отображение и его матрица

Б.М.Верников

Уральский федеральный университет,

Институт естественных наук и математики,

кафедра алгебры и фундаментальной информатики

6.1. Определение, примеры, теорема существования и единственности

Определение

Пусть V и W — векторные пространства над одним и тем же полем F. Отображение $\mathcal{A}$ из V в W называется *линейным*, если для любых векторов $\mathbf{x}_1, \mathbf{x}_2 \in V$ и любого скаляра $t \in F$ выполняются равенства

$$\mathcal{A}(\mathbf{x}_1 + \mathbf{x}_2) = \mathcal{A}(\mathbf{x}_1) + \mathcal{A}(\mathbf{x}_2) \quad \text{и} \quad \mathcal{A}(t\mathbf{x}_1) = t\mathcal{A}(\mathbf{x}_1).$$

Относительно первого из этих равенств говорят, что $\mathcal{A}$ *сохраняет сумму векторов*, относительно второго — что $\mathcal{A}$ *сохраняет произведение вектора на скаляр*. Произвольное отображение из векторного пространства в себя называется *оператором* в этом пространстве. Если оператор является линейным отображением, то он называется *линейным оператором*. Иными словами, линейный оператор $\mathcal{A}$ в пространстве V — это линейное отображение из V в V.

Если рассматривать векторные пространства как универсальные алгебры, то линейные отображения являются гомоморфизмами из одного векторного пространства в другое, а линейные операторы — эндоморфизмами векторного пространства.

Приведем примеры линейных отображений и линейных операторов. Начнем с примеров линейных операторов.

Пример 1. Зафиксируем в физическом трехмерном пространстве некоторую плоскость π и обозначим через V множество всех векторов, коллинеарных π. Ясно, что множество V с обычными линейными операциями над векторами является векторным пространством. Будем считать, что в плоскости π задана прямоугольная декартова система координат. Тогда линейным оператором в V является, например:

а) отображение из V в V, которое произвольному вектору из V ставит в соответствие проекцию этого вектора на ось абсцисс;

б) отображение из V в V, которое произвольному вектору $\vec{x} \in V$ ставит в соответствие вектор, симметричный вектору $\vec{x}$ относительно оси ординат;

в) отображение из V в V, которое произвольному вектору $\vec{x} \in V$ ставит в соответствие вектор, полученный поворотом вектора $\vec{x}$ против часовой стрелки на фиксированный угол α.

Пример 2. В физическом трехмерном пространстве V зафиксируем вектор $\vec{a}$ и определим оператор $\mathcal{A}$ в V правилом: $\mathcal{A}(\vec{x}) = (\vec{a}\vec{x}) \cdot \vec{a}$ для всякого вектора $\vec{x} \in V$. Из свойств скалярного произведения вытекает, что этот оператор — линейный.

Примеры линейных операторов (2)

Пример 3. Вновь зафиксируем в физическом трехмерном пространстве V вектор $\vec{a}$ и определим оператор $\mathcal{A}$ в V правилом: $\mathcal{A}(\vec{x}) = \vec{a} \times \vec{x}$ для всякого вектора $\vec{x} \in V$. Из свойств векторного произведения вытекает, что этот оператор — линейный.

Пример 4. Пусть V — векторное пространство над полем F. Зафиксируем произвольный скаляр $t \in F$ и зададим оператор $\mathcal{A}$ в V 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. Особо отметим два частных случая оператора растяжения. Первый из них — это оператор растяжения при $t = 0$. Он обозначается буквой $\mathcal{O}$ и называется *нулевым оператором*. 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 оператором*. Этот оператор переводит произвольный вектор из V в себя.

Пример 5. Зафиксируем в пространстве V некоторое подпространство M. В силу предложения 3.2 существует подпространство M' в V такое, что $V = M \oplus M'$. Согласно теореме 3.2, 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 $\mathbf{x}_2 \in M'$. 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'. Частным случаем оператора проектирования является оператор, указанный в примере 1а).

Пример 6. В пространстве $F[x]$ всех многочленов над полем F определим оператор $\mathcal{D}$ правилом: $\mathcal{D}(f) = f'$. 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 полем F.

Пример отображения, не являющегося линейным оператором

А вот отображение, которое всякой дифференцируемой функции из $\mathbb{R}$ в $\mathbb{R}$ ставит в соответствие ее производную, не является линейным оператором в пространстве дифференцируемых функций. Более того, оно не является даже просто оператором в этом пространстве, поскольку производная дифференцируемой функции может не быть дифференцируемой.
Приведем подтверждающий это пример.

Пример 6.1

Рассмотрим функцию $f(x)$ из $\mathbb{R}$ в $\mathbb{R}$, задаваемую правилом

$$f(x) = x \cdot |x| = \begin{cases} x^2, & \text{если } x \geqslant 0; \\ -x^2, & \text{если } x < 0. \end{cases}$$

Тогда функция $f(x)$ дифференцируема, но функция $f'(x) = 2|x|$ не дифференцируема, поскольку не имеет производной в точке $x = 0$.

Пример отображения, не являющегося линейным оператором (рисунок)

На рис. 1 функция $f(x)$ из примера 6.1 изображена красным цветом, а ее производная — синим.

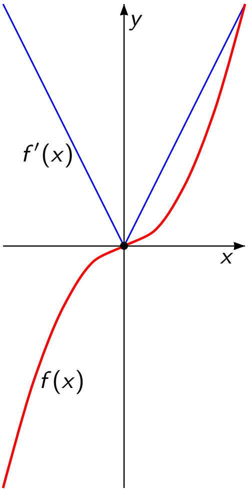

Рис. 1. Функция из примера 6.1 и ее производная

Некоторые из возникавших в примерах 1–6 линейных операторов можно рассматривать как линейные отображения. Например, нулевой оператор в пространстве V является линейным отображением из V в нулевое пространство, оператор проектирования пространства V на подпространство M параллельно M' — линейным отображением из V в M, а оператор дифференцирования в пространстве $F_n[x]$ — линейным отображением из $F_n[x]$ в $F_{n-1}[x]$. Во всех этих случаях векторное пространство отображается в его подпространство. Перейдем к примерам линейных отображений, не обладающих этим свойством.

Пример 7. Определим отображение $\mathcal{A}$ из пространства $F_n[x]$ в пространство F^{n+1} правилом: если $f = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$, то $\mathcal{A}(f) = (a_n, a_{n-1}, \ldots, a_0)$. Очевидно, что $\mathcal{A}$ — линейное отображение.

Пример 8. Определим отображение $\mathcal{A}$ из пространства $\mathbb{R}^n$ в пространство $\mathbb{C}^n$ правилом: если $\mathbf{x} \in \mathbb{R}^n$, то $\mathcal{A}(\mathbf{x}) = i\mathbf{x}$. Как и выше, линейность отображения $\mathcal{A}$ очевидна.

Пример 9. Пусть m и n — произвольные натуральные числа, а F — произвольное поле. Очевидно, что отображение из пространства $F^{m\times n}$ в пространство $F^{n\times m}$, которое всякой матрице $A \in F^{m\times n}$ ставит в соответствие матрицу $A^\top$, является линейным. Если $m = n$, то это отображение является линейным оператором в пространстве $F^{n\times n}$. Этот оператор называется *оператором транспонирования*.

Отметим, что отображения из примеров 7 и 9 являются изоморфизмами, а отображение из примера 8 — изоморфным вложением.

Пример 10. Пусть m, n и k — произвольные натуральные числа, а F — произвольное поле. Зафиксируем матрицу $A \in F^{n\times k}$ и определим отображение $\mathcal{M}$ из пространства $F^{m\times n}$ в пространство $F^{m\times k}$ правилом: $\mathcal{M}(X) = XA$ для всякой матрицы $X \in F^{m\times n}$. Из свойств произведения матриц вытекает, что это отображение линейно.

Напомним, что, как уже отмечалось в § 1, всякое поле можно рассматривать как векторное пространство над самим собой. Линейное отображение из векторного пространства V над полем F в пространство F называется *линейным функционалом*. Линейные функционалы — специфический, но важный тип линейных отображений, возникающий как в различных разделах математики, так и за ее пределами (например, в физике). Приведем несколько примеров линейных функционалов. В примерах 11–13 F — произвольное поле.

Пример 11. Определим отображение $\mathcal{A}$ из пространства F^n в F правилом: если $\mathbf{x} = (x_1, x_2, \dots, x_n) \in F^n$, то $\mathcal{A}(\mathbf{x}) = x_1 + x_2 + \cdots + x_n$. Очевидно, что $\mathcal{A}$ — линейный функционал.

Пример 12. Отображение из пространства $F[x]$ в F, ставящее в соответствие всякому многочлену над полем F его свободный член, является линейным функционалом.

Пример 13. Отображение из пространства $F^{n \times n}$ в F, ставящее в соответствие всякой квадратной матрице порядка n над полем F сумму элементов на ее главной диагонали, является линейным функционалом.

Пример 14. Пусть V — физическое трехмерное пространство. Зафиксируем вектор $\vec{a} \in V$ и рассмотрим отображение $\mathcal{A}$ из V в $\mathbb{R}$, задаваемое правилом: $\mathcal{A}(\vec{x}) = \vec{x}\vec{a}$ для всякого $\vec{x} \in V$. Ясно, что $\mathcal{A}$ — линейный функционал.

Пример 15. Пусть V — пространство всех функций из $\mathbb{R}$ в $\mathbb{R}$. Отображение из V в $\mathbb{R}$, ставящее в соответствие каждой функции $f(x)$ из V число $f(0)$, является линейным функционалом.

Перейдем к рассмотрению свойств линейных отображений. В дальнейших рассмотрениях в некоторых случаях речь будет одновременно идти о нулевых векторах двух различных векторных пространств. Во избежание путаницы договоримся, что в тех случаях, когда это необходимо, мы будем обозначать нулевой вектор векторного пространства U через $\mathbf{0}_U$.

Замечание 6.1

Пусть V и W — векторные пространства над полем F, а $\mathcal{A}$ — линейное отображение из V в W. Тогда:

1) $\mathcal{A}(\mathbf{0}_V) = \mathbf{0}_W$.
2) $\mathcal{A}(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_m\mathbf{v}_m) = \lambda_1\mathcal{A}(\mathbf{v}_1) + \lambda_2\mathcal{A}(\mathbf{v}_2) + \cdots + \lambda_m\mathcal{A}(\mathbf{v}_m)$ *для любых векторов* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ *из* V *и любых скаляров* λ_1, λ_2, ..., λ_m *из* F.

Доказательство. Первое свойство вытекает из того, что

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

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

Теорема 6.1

Пусть V и W — векторные пространства над полем F, $P = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ — базис пространства V, а $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ — произвольные векторы из W. Тогда существует, и притом только одно, линейное отображение $\mathcal{A}$ из V в W такое, что $\mathcal{A}(\mathbf{p}_i) = \mathbf{w}_i$ для всех $i = 1, 2, \ldots, n$.

Доказательство. *Существование.* Пусть $\mathbf{x} \in V$, а $(x_1, x_2, \ldots, x_n)$ — координаты вектора $\mathbf{x}$ в базисе P. Определим отображение $\mathcal{A}$ из V в W 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 2.5 вытекает, что это отображение линейно. 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 $\mathcal{B}$ — линейное отображение из V в W такое, что $\mathcal{B}(\mathbf{p}_i) = \mathbf{w}_i$ для всех $i = 1, 2, \dots, n$. Пусть $\mathbf{x} \in V$, а $(x_1, x_2, \dots, x_n)$ — координаты вектора $\mathbf{x}$ в базисе P. Тогда $\mathbf{x} = x_1\mathbf{p}_1 + x_2\mathbf{p}_2 + \dots + x_n\mathbf{p}_n$. В силу замечания 6.1 имеем

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

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

6.2. Матрица линейного отображения. Изменение матрицы оператора при замене базиса

Доказанная только что теорема устанавливает, что линейное отображение из V в W однозначно определяется тем, как оно действует на базисных векторах пространства V. Поэтому для того, чтобы иметь полную информацию о линейном отображении, достаточно знать координаты образов этих базисных векторов в каком-нибудь базисе пространства W. Это делает естественным следующее определение.

Определение

Пусть $\mathcal{A}$ — линейное отображение из V в W, $P = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ — базис пространства V, а $Q = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_m\}$ — базис пространства W. Матрица размера $m \times n$, i-й столбец которой состоит из координат вектора $\mathcal{A}(\mathbf{p}_i)$ в базисе Q (для всех $i = 1, 2, \ldots, n$), называется *матрицей отображения $\mathcal{A}$ в базисах P и Q*. Эта матрица обозначается через $A_{P,Q}$. Если $\mathcal{A}$ — линейный оператор в пространстве V, а P — базис V, то *матрицей оператора $\mathcal{A}$ в базисе P* называется квадратная матрица порядка n, i-й столбец которой состоит из координат вектора $\mathcal{A}(\mathbf{p}_i)$ в базисе P (для всех $i = 1, 2, \ldots, n$). Эта матрица обозначается через A_P. Таким образом, A_P — это матрица отображения $\mathcal{A}$ в базисах P и P.

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 $P = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ и $Q = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_m\}$ — базисы пространств V и W соответственно, $\mathcal{A}$ — линейное отображение из V в W, которое в базисах P и Q имеет матрицу $A_{P,Q} = (a_{ij})$, 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 Q? Обозначим эти координаты через $(y_1, y_2, \ldots, y_m)$. Тогда

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

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

$$\begin{array}{lcl} \mathcal{A}(\mathbf{p}_1) & = & a_{11}\mathbf{q}_1 + a_{21}\mathbf{q}_2 + \cdots + a_{m1}\mathbf{q}_m, \\ \mathcal{A}(\mathbf{p}_2) & = & a_{12}\mathbf{q}_1 + a_{22}\mathbf{q}_2 + \cdots + a_{m2}\mathbf{q}_m, \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ \mathcal{A}(\mathbf{p}_n) & = & a_{1n}\mathbf{q}_1 + a_{2n}\mathbf{q}_2 + \cdots + a_{mn}\mathbf{q}_m. \end{array}$$

Координаты образа вектора (2)

Следовательно,

$$
\begin{array}{rcl}
y_1\mathbf{q}_1 + y_2\mathbf{q}_2 + \cdots + y_m\mathbf{q}_m & = & x_1\mathcal{A}(\mathbf{p}_1) + x_2\mathcal{A}(\mathbf{p}_2) + \cdots + x_n\mathcal{A}(\mathbf{p}_n) \; = \\
& = & x_1\big(a_{11}\mathbf{q}_1 + a_{21}\mathbf{q}_2 + \cdots + a_{m1}\mathbf{q}_m\big) \; + \\
& + & x_2\big(a_{12}\mathbf{q}_1 + a_{22}\mathbf{q}_2 + \cdots + a_{m2}\mathbf{q}_m\big) \; + \\
& + & \dotfill \; + \\
& + & x_n\big(a_{1n}\mathbf{q}_1 + a_{2n}\mathbf{q}_2 + \cdots + a_{mn}\mathbf{q}_m\big) \; = \\
& = & \big(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n\big)\mathbf{q}_1 \; + \\
& + & \big(a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n\big)\mathbf{q}_2 \; + \\
& + & \dotfill \; + \\
& + & \big(a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n\big)\mathbf{q}_m.
\end{array}
$$

В силу единственности разложения вектора по базису это означает, что

$$
\begin{cases}
y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\
y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\
\dotfill \\
y_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n.
\end{cases}
$$

Эту систему равенств можно переписать в виде

$$
\big[\mathcal{A}(\mathbf{x})\big]_Q = A_{P,Q} \cdot [\mathbf{x}]_P. \qquad (1)
$$

Координаты образа вектора (3)

Применяя формулу (1) к случаю, когда $V = W$ (т. е. когда $\mathcal{A}$ — линейный оператор в V) и $P = Q$, получаем равенство

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

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 (2) и воспользоваться свойствами транспонирования матриц:

$$\big[\mathcal{A}(\mathbf{x})\big]_P^\top = [\mathbf{x}]_P^\top \cdot A_P^\top. \tag{3}$$

Ясно, что матрицы одного и того же линейного оператора в разных базисах могут быть различными. Возникает вопрос о том, как они связаны между собой. Ответ на него дает следующая теорема.

Теорема 6.2

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. Из формулы (2) и формулы (5) из § 2 вытекает, что

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

С другой стороны, из равенства (5) из § 2, примененного к вектору $\mathcal{A}(\mathbf{x})$, и равенства (2), 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 и получить, что $A_P T_{PQ} = T_{PQ} A_Q$. Принимая во внимание предложение 2.2, 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 2.2, формулу (4) можно переписать в виде $A' = T^{-1} A T$.

Таким образом, чтобы применить формулу (4) для решения конкретной задачи, достаточно сначала найти одну из матриц T_{PQ} и T_{QP} с помощью алгоритма 3.3, а затем — матрицу, обратную к найденной, с помощью алгоритма 15.1 из курса «Основы алгебры» или формулы (1) из § 15 того же курса.