Общий вид линейного отображения

Оказывается, что любое линейное отображение можно представить в виде *матрицы* – таблицы, заполненной числами. Это мы и будем формулировать и доказывать.

Цели на этот и следующий урок

На этом уроке мы докажем первые три утверждения из списка. А на следующем уроке мы докажем четвёртое утверждение.

1. любое линейное отображение $f : \mathbb{R} \to \mathbb{R}$ задаётся одним числом,
2. любое линейное отображение $f : \mathbb{R} \to \mathbb{R}^m$ задаётся столбцом высоты m, состоящим из чисел,
3. любое линейное отображение $f : \mathbb{R}^n \to \mathbb{R}$ задаётся строкой длины n, состоящей из чисел,
4. любое линейное отображение $f : \mathbb{R}^n \to \mathbb{R}^m$ задаётся таблицей с m строками и n столбцами, заполненной числами.

Утверждения 1-3 это частные случаи утверждения 4. Действительно, число можно воспринимать как таблицу размера 1 на 1. Столбец – как таблицу ширины 1. А строку – как таблицу высоты 1. Поэтому можно было бы сразу доказать утверждение 4, и из него бы следовали утверждения 1 - 3.

Но мы сначала докажем утверждения 1 - 3 отдельно, чтобы лучше понять утверждение 4, которое мы будем доказывать на следующем уроке.

Линейное отображение $f : \mathbb{R} \to \mathbb{R}$

Дано линейное отображение $f : \mathbb{R} \to \mathbb{R}$. Пусть мы знаем, чему равно $f(1)$. Докажем, что тогда мы можем узнать, чему равно $f(x)$ для любого $x \in \mathbb{R}$.

Действительно, так как x это число, мы можем воспользоваться вторым условием линейности: $f(x) = f(x \cdot 1) = x \cdot f(1)$. Например, если $f(1) = -\frac{1}{2}$, то $f(42) = f(42 \cdot 1) = 42f(1) = 42 \cdot (-\frac{1}{2}) = -21$.

Тем самым, по значению f в единице мы можем восстановить всё линейное отображение. То есть линейное отображение из $f : \mathbb{R} \to \mathbb{R}$ можно воспринимать как число $f(1)$, раз уж f им задаётся.

Геометрически можно представлять отображение f, как растягивание прямой в $f(1)$ раз (воспринимая множество $\mathbb{R}$ как прямую).

Пусть дано линейное отображение $f : \mathbb{R} \to \mathbb{R}$, и известно, что $f(1) = 12$. Чему равно $f(5)$?

Введите численный ответ

Введите число

Задача. Линейные отображения 4.

Линейное отображение $f:\mathbb{R}\to\mathbb{R}^m$

Обозначение. Когда мы будем оперировать векторами в вычислениях с матрицами, нам будет удобно представлять некоторые векторы не как строку $\vec{y}=(y_1,\ldots,y_m)$, а как столбец. Вот так:

$$\vec{y}=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Докажите, что любое линейное отображение $f:\mathbb{R}\to\mathbb{R}^m$ задаётся столбцом высоты m.

Подсказка. Доказательство аналогично доказательству с последнего шага с теорией.

Как можно геометрически представить отображение $f:\mathbb{R}\to\mathbb{R}^m$?

Пусть дано линейное отображение $f : \mathbb{R} \to \mathbb{R}^3$, и известно, что $f(1) = \begin{pmatrix} 6 \\ -2 \\ 0.5 \end{pmatrix}$. Чему равно $f(4)$?

Выберите один вариант из списка

$\begin{pmatrix} 6 \\ -2 \\ 0.5 \end{pmatrix}$

$\begin{pmatrix} 24 \\ -8 \\ 2 \end{pmatrix}$

$\begin{pmatrix} 12 \\ -4 \\ 1 \end{pmatrix}$

$\begin{pmatrix} 10 \\ 2 \\ 4.5 \end{pmatrix}$

Линейное отображение $f : \mathbb{R}^n \to \mathbb{R}$

На ближайших шагах мы будем понимать линейные отображения $\mathbb{R}^n \to \mathbb{R}$.

Применяем первое условие линейности

Для начала, пусть $n = 2$. То есть мы работаем с некоторым отображением $f : \mathbb{R}^2 \to \mathbb{R}$. Пусть известно, что $f(1, 0) = 10$ и $f(0, 1) = -7$. Попробуем исходя из этих данных понять, чему равно $f(1, 1)$.

Так как $(1, 1) = (1, 0) + (0, 1)$, по первому условию линейности выполнено $f(1, 1) = f(1, 0) + f(0, 1)$. Значит,

$$f(1, 1) = f(1, 0) + f(0, 1) = 10 - 7 = 3.$$

Мы нашли $f(1, 1) = 3$. По сути, мы воспользовались тем, что линейные отображения согласованы с операцией сложения векторов.

Пусть известно, что $f(1,0) = 10$ и $f(0,1) = -7$. Вычислите, чему равно $f(1,-1)$.

Введите численный ответ

Введите число

Линейное отображение $f : \mathbb{R}^n \to \mathbb{R}$

Применяем второе условие линейности

Дано линейное отображение $f : \mathbb{R}^3 \to \mathbb{R}$. Известно, что $f(1,0,0) = 8$, $f(0,1,0) = 6$ и $f(0,0,1) = 3$.

Попробуем исходя из этих данных понять, чему равно $f(5,0,0)$. Так как $(5,0,0) = 5 \cdot (1,0,0)$, по второму условию линейности $f(5,0,0) = 5 \cdot f(1,0,0)$. Значит,

$$f(5,0,0) = 5 \cdot f(1,0,0) = 5 \cdot 8 = 40.$$

Мы нашли $f(5,0,0) = 40$. По сути, мы воспользовались тем, что линейные отображения согласованы с операцией умножения вектора на число.

Пусть известно, что $f(1,0,0)=8$, $f(0,1,0)=6$ и $f(0,0,1)=3$. Вычислите, чему равно $f(0,-4,0)$.

Введите численный ответ

Введите число

Линейное отображение $f : \mathbb{R}^n \to \mathbb{R}$

Применяем оба условия линейности

Дано линейное отображение $f : \mathbb{R}^3 \to \mathbb{R}$.

Известно, что $f(1,0,0) = 3$, $f(0,1,0) = 5$ и $f(0,0,1) = 7$. Попробуем исходя из этих данных понять, чему равно $f(2,3,4)$. Применяя сначала первое, а потом второе условие линейности, получаем

$$f(2,3,4) = f(2,0,0) + f(0,3,0) + f(0,0,4) = 2f(1,0,0) + 3f(0,1,0) + 4f(0,0,1).$$

Теперь нам осталось только подставить значения $f(1,0,0)$, $f(0,1,0)$, $f(0,0,1)$ и сделать арифметические вычисления:

$$2f(1,0,0) + 3f(0,1,0) + 4f(0,0,1) = 2 \cdot 3 + 3 \cdot 5 + 4 \cdot 7 = 6 + 15 + 28 = 49.$$

Мы нашли $f(2,3,4) = 49$.

Известно, что $f(1,0,0)=3$, $f(0,1,0)=5$ и $f(0,0,1)=7$. Вычислите, чему равно $f(4,-2,5)$.

Введите численный ответ

Введите число

Линейное отображение $f:\mathbb{R}^n\to\mathbb{R}$

Как мы увидели на предыдущих шагах, линейное отображение $f:\mathbb{R}^3\to\mathbb{R}$ определяется своими значениями на векторах $(1,0,0)$, $(0,1,0)$ и $(0,0,1)$. То есть зная $f(1,0,0)$, $f(0,1,0)$ и $f(0,0,1)$ мы можем найти значение f на любом векторе.

Обобщим этот результат. Сначала введём имена для векторов из $\mathbb{R}^n$, являющихся аналогами векторов $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ в $\mathbb{R}^3$

Обозначение. Для удобства введём следующие обозначения векторов:

- $\vec{e}_1 := (1,0,0,\dots,0)$
- $\vec{e}_2 := (0,1,0,\dots,0)$
- $\vdots$
- $\vec{e}_n := (0,0,0,\dots,1)$

Комментарий. Знак «$:=$» здесь и далее в курсе означает «по определению равно».

Другими словами, у вектора $\vec{e}_i$ равна единице i-ая координата, а все остальные координаты равны нулю. Также про $\vec{e}_i$ можно думать, как про единичный вектор, направленный вдоль i-ой координатной оси. Например, в $\mathbb{R}^3$ вектор $\vec{e}_1$ направлен вдоль OX, $\vec{e}_2$ вдоль OY, и $\vec{e}_3$ вдоль OZ.

Пусть дано линейное отображение $f:\mathbb{R}^n\to\mathbb{R}$. И для каждого i известно значение f на векторе $\vec{e}_i$. То есть $f(\vec{e}_i)=a_i$ для некоторого числа $a_i\in\mathbb{R}$. Мы хотим доказать, что f определяется набором чисел $a_1,\dots,a_n$.

Пример. На прошлом шаге было $a_1=3, a_2=5, a_3=7$, и по этим числам мы могли вычислить $f(\vec{x})$ для любого $\vec{x}\in\mathbb{R}^3$.

Рассмотрим любой вектор $\vec{x}=(x_1,\dots,x_n)$. По определению векторного пространства $\mathbb{R}^n$, выполнено:

$$(x_1,\dots,x_n)=(x_1,0,\dots,0)+(0,x_2,\dots,0)+\dots+(0,0,\ \dots,x_n)=x_1\vec{e}_1+x_2\vec{e}_2+\dots+x_n\vec{e}_n.$$

Обозначение. В таких случаях говорят, что вектор $\vec{x}$ *выражается* через векторы $\vec{e}_1,\dots,\vec{e}_n$. То есть представляется в виде суммы векторов $\vec{e}_1,\dots,\vec{e}_n$ с некоторыми коэффициентами.

Применив сначала первое, а затем второе условие линейности, получаем

$$f(\vec{x})=f(x_1\vec{e}_1+\dots+x_n\vec{e}_n)=f(x_1\vec{e}_1)+\dots+f(x_n\vec{e}_n)=x_1f(\vec{e}_1)+\dots+x_nf(\vec{e}_n)=x_1a_1+\dots+x_na_n.$$

Нам будет удобнее поменять порядок множителей: вместо $x_i\ a_i$ будем писать a_ix_i. Итак, мы получили, что $f(\vec{x})=a_1x_1+\dots+a_nx_n$.

Строка на столбец

Итак, $f(\vec{x}) = a_1x_1 + \dots + a_nx_n$, где $a_i = f(\vec{e}_i)$. Объединим $a_1, \dots, a_n$ в вектор $(a_1, \dots, a_n)$. Этот вектор нам удобнее будет записывать как строку. А вычисление значения функции f мы будем записывать как *действие* строки $(a_1, \dots, a_n)$ на столбец $\vec{x}$. Вот так:

$$f(\vec{x}) = (a_1, \dots, a_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := a_1x_1 + \dots + a_nx_n = \sum_{i=1}^{n} a_ix_i.$$

Можно говорить, что мы *умножаем* строку $(a_1, \dots, a_n)$ на столбец $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.

Заметьте, что здесь мы используем слово "умножаем" в новом смысле, не имеющем отношения к умножению чисел. По сути, мы придумали операцию между строкой и столбцом, и дали ей такое же название, как и для операции умножения чисел.

Мы поняли, что любое линейное отображение $f : \mathbb{R}^n \to \mathbb{R}$ задаётся строкой из n чисел.

Комментарии

Комментарий 1. Порядок в нашей операции важен: строка должна быть записана *слева* от столбца. Почему – узнаем, когда будем учиться умножать матрицы друг на друга. Почему мы хотим записывать $(a_1, \ldots, a_n)$ как строку? Это мы тоже узнаем в следующем уроке, когда будем составлять из нескольких строк матрицу.

Комментарий 2. Выражение $\sum\limits_{i=1}^{n} a_i x_i$ также называют *скалярным произведением* векторов $(a_1, \ldots, a_n)$ и $(x_1, \ldots, x_n)$. У скалярного произведения есть много полезных свойств, о них мы поговорим в более поздних уроках.

Комментарий 3. Мы написали, что умножение строки на столбец не имеет отношения к умножению чисел. Это не совсем так. Пусть даны числа p и q. Тогда результат умножения p и q как чисел равен числу pq. Теперь будем смотреть на p как на строку длины 1, а на q как на столбец высоты 1. Результат умножения p на q как строки на столбец тоже равен числу pq (легко видеть из формулы для умножения строки на столбец). Так что в каком-то смысле умножение чисел – частный случай умножения строки на столбец.

Пусть линейное отображение $f:\mathbb{R}^4 \to \mathbb{R}$ задано строкой $(-7, 4,\ 2, 10)$. Найдите значение f от следующих столбцов:

Заполните пропуски

1. $f\begin{pmatrix}1\\1\\1\\1\end{pmatrix} =$ ______ .

2. $f\begin{pmatrix}3\\0\\5\\1\end{pmatrix} =$ ______ .

3. $f\begin{pmatrix}8\\2\\7\\4\end{pmatrix} =$ ______ .

Что мы прошли на этом уроке

- Мы узнали, что линейное отображение — это функция из одного векторного пространства в другое, которая уважает операции сложения и умножения на число. А ещё линейные отображения это основные компоненты *нейронной сети*.
- Мы поняли, как:

1. одним числом задать линейное отображение $f : \mathbb{R} \to \mathbb{R}$
2. столбцом из m чисел задать $f : \mathbb{R} \to \mathbb{R}^m$
3. строкой из n чисел задать $f : \mathbb{R}^n \to \mathbb{R}$. В частности, мы научились умножать строку на столбец.

Что нас ждёт на следующем уроке

На следующем уроке мы

- научимся по *матрице* строить линейное отображение из $\mathbb{R}^n$ в $\mathbb{R}^m$
- поймём, как перемножать *матрицу* и вектор