Линейные отображения

Зачем изучать линейные отображения?

Нейронные сети

Линейные отображения это основные компоненты *нейронной сети*. Обычная нейронная сеть состоит из нескольких *линейных отображений* и нескольких *функций активации*. Когда мы тренируем нейронную сеть, мы ищем наилучшие *коэффициенты* для линейных отображений, из которых она состоит. Понять функции активации просто, мы это сделаем на первом уроке второй недели. Понимать линейные отображения мы будем всю первую неделю курса.

Преобразование векторных данных

Очищенные данные обычно представлены в виде набора векторов. Поэтому многие преобразования данных задаются в виде линейных отображений. Например, *метод главных компонент* – один из основных способов уменьшить размерность данных – использует линейное отображение.

Дифференциал

Дифференциал, с которым мы познакомились в курсе матана, это тоже линейное отображение.

Дополнительный материал – 3blue1brown. Видео Chapter 3. Linear transformations and matrices (длина 11 минут). В конце видео Грант уже начинает рассказывать про матрицы; к ним мы перейдем в следующем уроке.

Определение

Слово "отображение" это синоним слова "функция". В нашем случае это будет функция из одного векторного пространства в другое. Например, линейное отображение f может быть функцией из векторного пространства $\mathbb{R}^3$ в векторное пространство $\mathbb{R}^2$. По-другому это можно записать так:

$$f : \mathbb{R}^3 \to \mathbb{R}^2.$$

То есть на вход f принимает вектор $\vec{x} \in \mathbb{R}^3$, а на выход даёт вектор $f(\vec{x}) \in \mathbb{R}^2$. В общем случае линейное отображение будет функцией из какого-то векторного пространства V в какое-то векторное пространство W, то есть $f : V \to W$.

Определение. *Линейным отображением* векторного пространства V в векторное пространство W называется функция $f : V \to W$ удовлетворяющая следующим двум условиям:

1. $f(\vec{x}) + f(\vec{y}) = f(\vec{x} + \vec{y})$ для любых $\vec{x}, \vec{y} \in V$,
2. $f(c\vec{x}) = cf(\vec{x})$ для любого $\vec{x} \in V$ и любого числа $c \in \mathbb{R}$.

Другими словами, что означает первое условие?

- Посмотрим на левую сторону равенства. Мы сначала применяем f к $\vec{x}$ и f к $\vec{y}$, а потом складываем результаты.
- Посмотрим на правую сторону равенства. Мы сначала складываем $\vec{x}$ и $\vec{y}$, а потом применяем к результату f.

В обоих случаях мы должны получить один и тот же вектор. Заметим, что в левой части равенства сложение векторов происходит в векторном пространстве W, а в правой части сложение происходит в V. То есть про первое условие следует думать как "f согласовано с операцией сложения векторов". Или "f уважает операцию сложения".

Комментарий. Слово "уважает" это дословный перевод (f respects addition).

Аналогично, про второе условие следует думать как "f согласовано с операцией умножения на число". Или "f уважает операцию умножения на число" (f respects multiplication by scalar).

То есть линейное отображение – это функция $f : V \to W$, которая согласована с операциями сложения векторов и умножения вектора на число.

Как мы помним, векторное пространство это множество векторов, которые можно складывать друг с другом и умножать на числа. По сути, операция сложения векторов и операция умножения на число – это те две операции, которые и задают векторное пространство. То есть линейное отображение это отображение, которое уважает структуру векторного пространства.

На следующем шаге мы разберём несколько примеров отображений и посмотрим, линейные они или нет.

Пример 1

Пусть $f : \mathbb{R}^2 \to \mathbb{R}$. И f отображает вектор $(x_1, x_2) \in \mathbb{R}^2$ в вектор $(x_1 + x_2) \in \mathbb{R}$. То есть, например, $f(3, 6) = 9$ и $f(-4, 5) = 1$.

Является ли это отображение линейным? Да. Проверим это. То есть покажем, что выполняются оба условия из определения линейного отображения: сохранение операции сложения и операции умножения на число.

Первое условие. Рассмотрим любые два вектора $\vec{x},\ \vec{y} \in \mathbb{R}^2$. Для них выполнено:

- $f(\vec{x}) + f(\vec{y}) = (x_1 + x_2) + (y_1 + y_2)$,
- $f(\vec{x} + \vec{y}) = f(x_1 + y_1, x_2 + y_2) = (x_1 + y_1) + (x_2 + y_2)$

Видно, что результаты этих строк отличаются только порядком слагаемых, то есть $f(\vec{x}) + f(\vec{y}) = f(\vec{x} + \vec{y})$.

Второе условие.

$$f(c\vec{x}) = f(cx_1, cx_2) = cx_1 + cx_2 = c(x_1 + x_2) = cf(\vec{x}).$$

То есть второе условие тоже выполнено. Значит, отображение $f(x_1, x_2) = x_1 + x_2$ является линейным.

Пример 2

Пусть $f : \mathbb{R}^2 \to \mathbb{R}$ отображает все векторы в вектор $7 \in \mathbb{R}$. Является ли это отображение линейным? Нет. Докажем это методом от противного.

Пусть это отображение линейно. Подставим во второе свойство вектор $\vec{x} = (1, 0)$ и $c = 2$. Получим, что $f(2, 0)$ должен быть равен $2 \cdot f(1, 0)$. Но это противоречие, ведь $f(2, 0) = 7$, а $2 \cdot f(1, 0) = 2 \cdot 7 = 14 \neq 7$.

Значит, отображение $f(\vec{x}) = 7$ не является линейным.

Задача с проверкой. Линейные отображения 1.

Какие из этих отображений линейны?

Если отображение линейное, то докажите, что первое и второе условие выполнены. Если отображение не линейное, то приведите пример, на котором одно из условий не выполняется.

Комментарий. Эта задача выглядит короткой, но её решение может занять некоторое ненулевое время. Для каждого из отображений ниже формально проверьте, выполняются ли условия линейности (как мы сделали в примерах с предыдущего шага). По сути этот шаг состоит из четырёх подзадач, каждая из которых не очевидна.

Выберите все подходящие ответы из списка

$f : \mathbb{R}^2 \to \mathbb{R}$, где $f(x_1, x_2) = x_1 x_2$
$f : \mathbb{R}^2 \to \mathbb{R}$, где $f(x_1, x_2) = x_1$
$f : \mathbb{R}^2 \to \mathbb{R}^2$, где $f(x_1, x_2) = (x_2, x_1)$
$f : \mathbb{R} \to \mathbb{R}$, где $f(x_1) = x_1^2$

Задача с проверкой. Линейные отображения 2.

Какие из этих отображений линейны?

Если отображение линейное, то докажите, что первое и второе условие выполнены. Если отображение не линейное, то приведите пример, на котором одно из условий не выполняется.

Комментарий. Эта задача выглядит короткой, но её решение может занять некоторое ненулевое время. Для каждого из отображений ниже формально проверьте, выполняются ли условия линейности (как мы сделали в примерах с пред-предыдущего шага). По сути этот шаг состоит из четырёх подзадач, каждая из которых не очевидна.

Выберите все подходящие ответы из списка

$f : \mathbb{R}^2 \to \mathbb{R}$, где $f(x_1, x_2) = 3x_1 - 5x_2$
$f : \mathbb{R}^2 \to \mathbb{R}^4$, где $f(x_1, x_2) = (x_2, x_2, x_2, x_2)$
$f : \mathbb{R}^3 \to \mathbb{R}$, где $f(x_1, x_2, x_3) = 0$
$f : \mathbb{R} \to \mathbb{R}$, где $f(x) = |x|$

Задача. Линейные отображения 3.

Докажите, что если $f : V \to W$ это линейное отображение, то

1. $f(\vec{0}_V) = \vec{0}_W$. В левой части используется нулевой вектор из пространства V, а в правой нулевой вектор из пространства W. Мы добавили нижний индекс, чтобы как-то различать эти два нулевых вектора.
2. $f(-\vec{x}) = -f(\vec{x})$. В левой части используется вектор, противоположный $\vec{x} \in V$, а в правой — вектор, противоположный $f(\vec{x}) \in W$.

Попробуйте не пользоваться покоординатным представлением векторов в своём доказательстве. Например, использовать утверждение $(-\vec{x}) + \vec{x} = \vec{0}$, но не использовать утверждение $(-\vec{x}) = (-x_1, -x_2, \ldots, -x_n)$.

Какие из этих отображений линейны?

Выберите все подходящие ответы из списка

$f(\overrightarrow{x}) = -\overrightarrow{x}$ для $\overrightarrow{x} \in \mathbb{R}^{100}$

$f(x_1, x_2, x_3) = x_1 x_2 x_3$

$f(x_1) = (x_1, x_1^2, x_1^3)$

$f(x_1, x_2) = (x_1 + 3x_2, 2x_1 - 5x_2)$

$f(x_1, x_2, x_3) = 18x_1 + \frac{2}{7}x_3$