Векторы — инструмент для описания и анализа пространства. Векторы это направленные отрезки, которые характеризуются величиной (длиной) и направлением. Векторы позволяют решать задачи в физике, а главное в Машинном Обучении (классификация данных, рекомендательные системы, обработка естественнго языка).

Основные элементы, связанные с векторами:

- Длина (модуль): численное значение, определяющее размер вектора.
- Направление: указывает, куда «смотрит» вектор.
- Операции с векторами: сложение, вычитание, умножение на число, скалярное и векторное произведение, которые упрощают изучение взаимосвязей между объектами.

Векторы находят применение как в школьной геометрии для решения задач на плоскости, так и в более сложных разделах математики, таких как аналитическая геометрия, физика и компьютерная графика.

Человеческим языком: векторы это передвижение всего пространства в каком-то направлении на какое-то расстояние.

Разложение вектора

Каждый вектор можно представить в виде суммы единичных векторов.

Например вектор из точки A в точку B можно описать как, из точки A сделать i шагов вправо и j шага вверх и мы окажемся в точке B. Любой вектор $\vec{v}$ плоскости единственным образом выражается в виде:

$$\vec{v} = v_1\vec{i} + v_2\vec{j}$$

где v_1, v_2 — числа, которые называются координатами вектора.

Координаты вектора и длина отрезка

Пусть заданы точки $A(x_1, y_1)$ и $B(x_2, y_2)$.

- Вектор $\overrightarrow{AB}$ имеет следующие координаты:

$$\overrightarrow{AB} = (x_2 - x_1, y_2 - y_1)$$

- Длину отрезка:

$$|\overrightarrow{AB}| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Правило сложения векторов

Пусть $\vec{v} = (v_1; v_2)$ и $\vec{w} = (w_1; w_2)$. Тогда:

$$\vec{v} + \vec{w} = (v_1 + w_1; v_2 + w_2).$$

Правило умножения вектора на число

Пусть вектор $\vec{v} = (v_1, v_2)$. Умножим его на число λ:

$$\lambda\vec{v} = (\lambda v_1, \lambda v_2).$$

Скалярное произведение

Скалярное произведение двух векторов $\vec{a}$ и $\vec{b}$ определяется как:

$$\vec{a} \cdot \vec{b} = |\vec{a}||\vec{b}| \cos\angle(\vec{a}, \vec{b}).$$

Задачи для закрепления (на бумаге)

1. Даны две точки плоскости $A(2,1)$ и $B(-2,3)$. Найти координаты вектора $\overrightarrow{AB}$.
2. Даны точки $A(-3,5)$ и $B(1,-3)$. Найти длину отрезка AB.
3. Даны векторы $\vec{a}(1;-2)$ и $\vec{b}(2;3)$. Найти:
   - $2\vec{a}$,
   - $\vec{a}+\vec{b}$,
   - $\vec{a}-\vec{b}$.

Задачи на написание кода

1. Вводится n, в следующих n строках вводятся через пробел 2d координаты векторов. Выведите сумму всех векторов
2. Вводится n, в следующих n строках вводятся через пробел 2d координаты векторов. Известно, что какая-то часть векторов направлены вверх с погрешностью 30 градусов, другая часть направлены вниз с погрешностью 30 градусов. Выведите номера (начиная с 0) векторов сначала всех, кто смотрят в вверх, потом тех, кто смотрят вниз.
3. Вводится n, в следующих n строках даются координаты 2 точек определяющие прямую через пробелы:

   x_1 y_1 x_2 y_2

   Известно, что все прямые параллельны друг другу, найдите самые близкие друг к другу два прямяых и выведите их номера (начиная с 0) через пробел.
4. Вводится n, в следующих n строках вводятся через пробел 2d координаты векторов. Известно, что какая-то часть векторов направлены в одну сторону с погрешностью 30 градусов другая часть направлены в противоположную сторону с погрешностью 30 градусов Выведите номера (начиная с 0) вектороов сначала всех, кто смотрят в одну, потом тех, кто смотрят во вторую сторону.
5. (Задача на машинное обучение) Вводится n, в следующих n строках вводятся через пробел 2d координаты векторов. Известно, что какая-то часть векторов направлены в одну сторону с погрешностью 30 градусов другая часть направлены в другую сторону с погрешностью 30 градусов, разница между двумя направлениями $>$ 40 градусов. Выведите номера (начиная с 0) векторов сначала всех, кто смотрят в одну, потом тех, кто смотрят во вторую сторону.

Косинусное расстояние между векторами

Косинусное расстояние — это метрика, которая показывает степень различия между двумя векторами, основываясь на угле между ними. В отличие от эвклидового расстояния, косинусное расстояние оценивает, насколько похожи направления векторов, а не их длины.

Определение

Косинусное расстояние вычисляется на основе косинуса угла между двумя векторами. Если даны два вектора $\vec{a}$ и $\vec{b}$, то косинусное расстояние определяется следующим образом:

$$\text{Cosine Distance} = 1 - \cos(\theta),$$

где $\cos(\theta)$ — это косинус угла θ между векторами, который вычисляется как:

$$\cos(\theta) = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|},$$

где:

- $\vec{a} \cdot \vec{b}$ — скалярное произведение векторов,
- $\|\vec{a}\|$ и $\|\vec{b}\|$ — длины (модули) векторов.

Интерпретация

- $\text{Cosine Distance} = 0$ — векторы полностью совпадают по направлению ($\theta = 0^\circ$).
- $\text{Cosine Distance} = 1$ — векторы полностью противоположны ($\theta = 180^\circ$).
- Значения между 0 и 1 показывают степень различия в направлениях векторов.

Применение

Косинусное расстояние часто используется в задачах машинного обучения и анализа данных, например:

- Сравнение текстовых данных в задачах обработки естественного языка (например, сходство текстов на основе частот слов).
- Кластеризация и поиск ближайших соседей в задачах с высокоразмерными данными.
- Оценка схожести рекомендаций или предпочтений пользователей в рекомендательных системах.

Важно отметить, что косинусное расстояние не учитывает длины векторов, а только угол между ними, что делает его особенно полезным для задач, где важны пропорции, а не абсолютные величины.