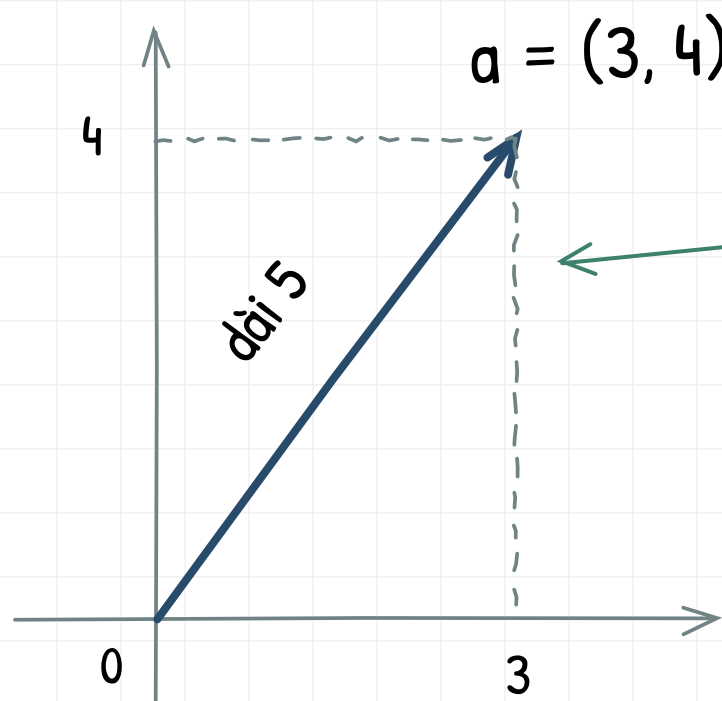


Vector & tích vô hướng

Một mảng số... nhưng hãy nhìn thành mũi tên!



sang phải 3
đi lên 4

$$\begin{aligned}\text{Độ dài} &= \text{norm} \\ |a| &= \sqrt{3^2 + 4^2} \\ &= 5\end{aligned}$$

($|a|$ ở đây là độ dài vector)

Dot: nhân từng cặp, rồi cộng

$$a = (3, 4) \quad b = (4, 3)$$

$$a \cdot b = 3 \times 4 + 4 \times 3 = 24$$

Kết quả là một SỐ.

Cây cầu từ số sang hình

$$a \cdot b = |a| |b| \cos \theta$$

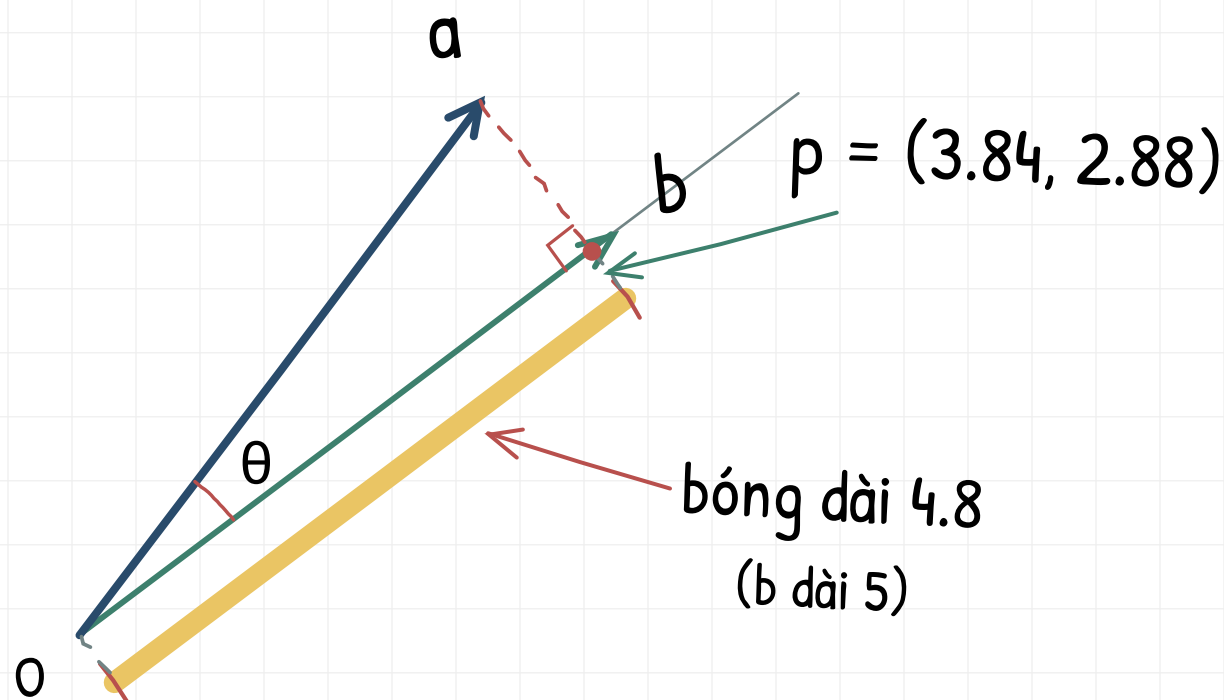
Dot trộn cả độ dài lẫn hướng.

Muốn chỉ so hướng? Chia cho hai độ dài.

☆ Nhớ hình trước, công thức sẽ có chỗ bám.

Phép chiếu = “cái bóng”

Thả vuông góc từ đầu a xuống đường chứa b .



1. Bóng dưới dạng SỐ

$$s = |a| \cos \theta = (a \cdot b) / |b|$$

$$= 24 / 5 = 4.8$$

Có dấu! Góc tù thì s âm.

2. Bóng dưới dạng VECTOR

$$p = [(a \cdot b) / (b \cdot b)] \times b$$

$$= (24 / 25) \times (4, 3) = (3.84, 2.88)$$

Chia $|b|$ để lấy s , rồi nhân $b / |b|$.

Vậy mẫu số có hai lần $|b|$: $b \cdot b = |b|^2$.

☆ Phần dư $a - p = (-0.84, 1.12)$ vuông góc b .

Kiểm: $(-0.84) \times 4 + 1.12 \times 3 = 0$.

Cosine: chỉ giữ phần hướng

$$\cos(a, b) = (a \cdot b) / (|a| |b|)$$

a, b khác vector 0

$$a = (3, 4), b = (4, 3), c = (-4, 3)$$

Cả ba dài 5 $\Rightarrow$ mẫu số luôn là 25.

cặp	dot	cos	đọc là
a, b	24	0.96	gần cùng hướng
a, c	0	0	vuông góc
b, c	-7	-0.28	góc tù

Luôn trong $[-1, 1]$ vì $|a \cdot b| \leq |a| |b|$.

Cauchy-Schwarz: đúng cả ở nhiều chiều.

Kéo dài không đổi cosine

$$q = (1, 0) \quad d = (10, 0)$$

$q \rightarrow$

$d \rightarrow$

$$\cos(q, d) = 1$$

$$\text{khoảng cách} = 9$$

Cùng hướng, dù khác độ dài.

Nối sang retrieval

Chuẩn hóa query và tài liệu về độ dài 1:

cosine = dot. Euclid cũng cho cùng thứ hạng.

Chưa chuẩn hóa thì dot $\neq$ cosine: kiểm tra mô hình embedding trước.

Nhớ đúng & tự kiểm

Ba chỗ dễ trượt

- 01 Dot lớn chưa chắc gần hướng hơn.
Độ dài tăng cũng có thể làm dot tăng.
- 02 Vuông góc: hình chiếu bằng 0.
Chưa đủ để kết luận “không chia sẻ thông tin”.
- 03 Cosine thô không phải tương quan Pearson.
 $\text{Pearson} = \text{cosine}$ sau khi trừ trung bình riêng mỗi dãy (phương sai phải khác 0).

Che phần dưới, trả lời thử

1. Nhân cả a và b với 100: dot, cosine đổi sao?
2. Vì sao phép chiếu chia cho $b \cdot b$?
3. Phần dư sau phép chiếu có hướng thế nào?
4. Vector 0 có tính cosine được không?

Đáp nhanh

1. $\text{Dot} \times 10.000$; cosine giữ nguyên.
2. Chia độ dài hai lần: lấy s, rồi lấy hướng đơn vị.
3. Vuông góc với hướng được chiếu lên.
4. Không: mẫu số bằng 0.