

### 3-8. 코사인 유사도

벡터와 벡터 사이의 각을 구하는 법

↳  $L_2$  Norm과 내적의 정의를 통해

#### 내적의 정의식 ①

$$\hookrightarrow \langle a, b \rangle = \|a\| \|b\| \cos \theta$$

$$\cos \theta = \frac{\langle a, b \rangle}{\|a\| \|b\|}$$

#### 내적의 정의식 ②

$$\langle a, b \rangle = a_1 b_1 + a_2 b_2 + \dots + a_n b_n = \sum_{i=1}^n a_i b_i$$

#### $L_2$ 의 정의식

$$\|a\|_2 = \sqrt{\sum_{i=1}^n a_i^2} = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}$$

이 식들을 조합하면  $\cos \theta$ 를 구할 수 있음

↳ Cosine Similarity라고 함

$$\hookrightarrow \cos(a, b)$$

#### 정의

$$\cos(a, b) = \frac{\langle a, b \rangle}{\|a\| \|b\|} = \frac{\sum_{i=1}^n a_i b_i}{\sqrt{\sum_{i=1}^n a_i^2} \sqrt{\sum_{i=1}^n b_i^2}}$$

$$-1 \leq \text{Cosine Similarity} \leq 1$$

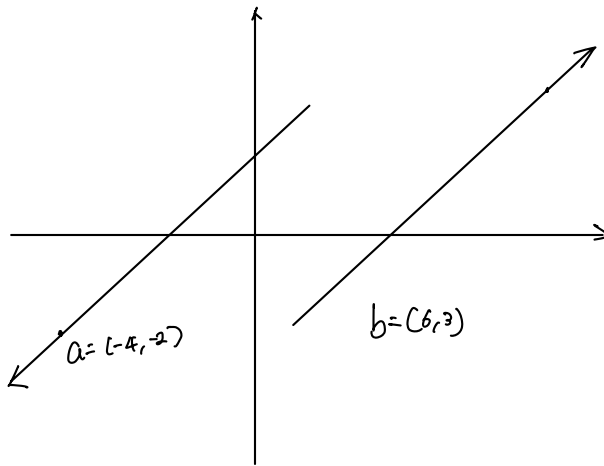
유사도가 -1 일 때는 두 벡터가 서로 반대 방향으로 평행.

유사도가 0 일 때는 두 벡터가 직교

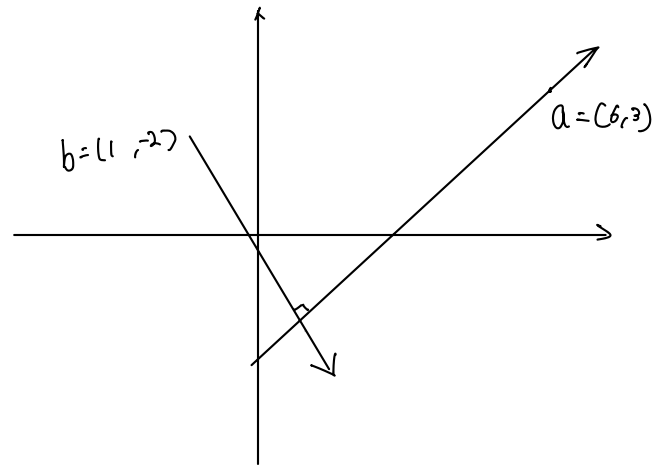
유사도가 1 일 때는 두 벡터가 같은 방향으로 평행.

Q1(1)

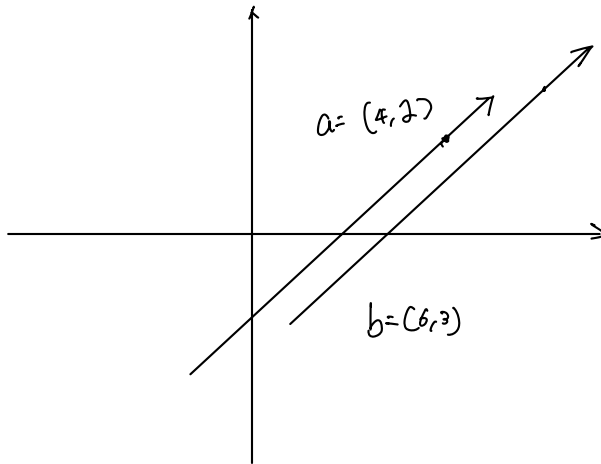
①  $\cos(a,b) = -1$



②  $\cos(a,b) = 0$



③  $\cos(a,b) = 1$



Cosine Similarity 가 높을까?

그 유사하다.

인공지능예제?

인공지능이 텍스트를 분석할 때 어휘적으로 단어 분량을 파악해 처리함

- 어휘로 만들어진 단어나 문장들은 서로의 관계성을 파악할 때 코사인 유사도를 사용  
코사인 유사도가 높을 수록 단어나 문장들은 더욱 가까운 관계임