

2차시

인공지능 & 머신 러닝 개론 2

20 . . .

지은이: 이은재

I. 머신러닝 입문에 필요한 수학 (1)

1. 절댓값(absolute value)과 유클리드 거리(Euclidean distance)

- **절댓값**: 수직선 상의 어떤 수가 원점으로부터 떨어진 거리
- **유클리드 거리**: 유클리드 공간 상의 어느 한 점과 다른 점 사이의 직선 거리
- **1차원(선)**에서의 유클리드 거리

$$D = |\text{시작점} - \text{도착점}|$$

- **2차원(평면)**에서의 유클리드 거리

두 점 $A(a_1, a_2)$ 와 $B(b_1, b_2)$ 가 있다고 할 때,

$$D = \|A - B\| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}$$

- **3차원(공간)**에서의 유클리드 거리

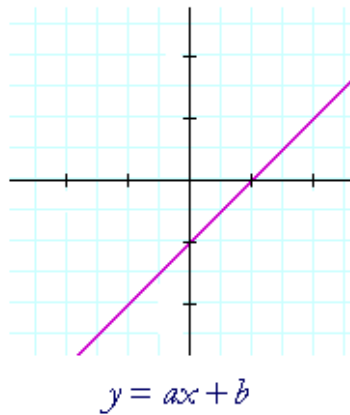
두 점 $A(a_1, a_2, a_3)$ 와 $B(b_1, b_2, b_3)$ 가 있다고 할 때,

$$D = \|A - B\| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2}$$

※ 거리를 구하는 공식에 왜 제곱근이 들어가는지 꼭 기억해두시다!

2. 그래프 그리기

x, y에 대한 1차 방정식 $y = ax + b$ 가 있다고 할 때, 그래프는 다음과 같이 직선으로 그려진다.



- 기울기

: 기울기 a 는 x 의 변화량에 따른 y 의 변화량을 알려준다.

- 즉, 1차 함수 그래프의 기울기는 x 의 변화에 y 가 얼마나 큰 영향을 받는지를 결정한다.
- 인공지능을 공부할 때 기울기는 어떤 ‘특성’의 ‘가중치’를 의미하며 기호로 a 대신 w 를 사용해 표기한다.

- 절편

: 절편 b 는 1차 함수 그래프가 x 축, y 축과 만나는 지점을 알려준다.

- 즉, 1차 함수 그래프의 절편은 x 의 변화와 무관하게 y 의 값이 치우친 정도를 결정한다.
- 인공지능을 공부할 때 절편은 y 가 가지는 ‘편향’ 또는 ‘선입견’을 의미한다.

3. 벡터

‘벡터’의 물리학적 의미는 ‘크기와 방향을 가지는 물리량’이지만, 수학이나 프로그래밍에서의 벡터는 ‘여러 개의 데이터를 하나의 열 또는 행에 담아둔 배열’을 말한다. 벡터의 성분 수를 ‘차원’이라고 한다. n개의 성분을 가지는 벡터를 ‘n-차원 벡터’라고 부른다.

ex)

$$a = (a_1, a_2, \dots, a_n) \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix} \quad c = ([a_1, b_1], [a_2, b_2], \dots, [a_n, b_n])$$

4. 벡터의 덧셈과 뺄셈, 스칼라곱

차원이 같은 벡터끼리는 서로 더하고, 뺄 수 있다.

(나누는 연산은 없음 / 벡터의 곱셈은 두 종류로, 추후에 다룹니다)

ex)

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \\ 9 \end{pmatrix} \quad (1, 2, 3) - (4, 5, 6) = (-3, -3, -3) \quad \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 5 \\ 6 \\ 7 \end{pmatrix} = \text{계산불가}$$

실수와 벡터 사이에는 곱셈이 가능한데,

이것을 스칼라곱(scalar product)이라고 부른다.

$$5 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \times 1 \\ 5 \times 2 \\ 5 \times 3 \end{pmatrix} = \begin{pmatrix} 5 \\ 10 \\ 15 \end{pmatrix}$$

ex) Word2Vec



The fat cat sat on the mat

The fat cat sat on the mat

The fat cat sat on the mat

The fat cat sat on the mat

The fat cat sat on the mat

The fat cat sat on the mat

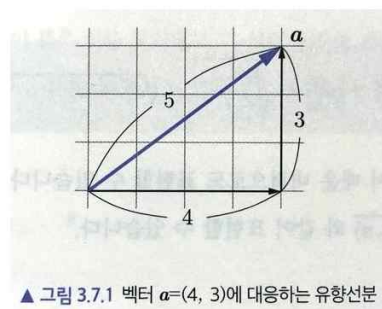
The fat cat sat on the mat

중심 단어	주변 단어
[1, 0, 0, 0, 0, 0, 0]	[0, 1, 0, 0, 0, 0, 0], [0, 0, 1, 0, 0, 0, 0]
[0, 1, 0, 0, 0, 0, 0]	[1, 0, 0, 0, 0, 0, 0], [0, 0, 1, 0, 0, 0, 0], [0, 0, 0, 1, 0, 0, 0]
[0, 0, 1, 0, 0, 0, 0]	[1, 0, 0, 0, 0, 0, 0], [0, 1, 0, 0, 0, 0, 0], [0, 0, 0, 1, 0, 0, 0], [0, 0, 0, 0, 1, 0, 0]
[0, 0, 0, 1, 0, 0, 0]	[0, 1, 0, 0, 0, 0, 0], [0, 0, 1, 0, 0, 0, 0], [0, 0, 0, 0, 1, 0, 0], [0, 0, 0, 0, 0, 1, 0]
[0, 0, 0, 0, 1, 0, 0]	[0, 0, 1, 0, 0, 0, 0], [0, 0, 0, 1, 0, 0, 0], [0, 0, 0, 0, 0, 1, 0], [0, 0, 0, 0, 0, 0, 1]
[0, 0, 0, 0, 0, 1, 0]	[0, 0, 0, 1, 0, 0, 0], [0, 0, 0, 0, 1, 0, 0], [0, 0, 0, 0, 0, 0, 1]
[0, 0, 0, 0, 0, 0, 1]	[0, 0, 0, 0, 1, 0, 0], [0, 0, 0, 0, 0, 1, 0]

5. 벡터의 노름(norm)

벡터를 기하학적으로 표현하여 유향선분(방향이 있는 선분)으로 표시했을 때 벡터가 이동한 거리를 '노름(norm)'이라고 한다. 노름을 구하는 방법은 다양하지만 그중에서도 L2 노름이 유클리드 거리와 같아 가장 많이 사용된다.

x, y 좌표 상의 벡터 $a(4,3)$ 이 있다고 할 때, 다음과 같이 유향선분으로 나타낼 수 있다.



- L1 노름 (맨해튼 거리)

$$\|a\|_1 = |a_x| + |a_y| = 4 + 3 = 7$$

- $L2$ 노름 (유클리드 거리)

$$\|a\|_2 = \sqrt{\sum_{i=1}^n a_i^2} = \sqrt{a_x^2 + a_y^2} = \sqrt{4^2 + 3^2} = \sqrt{25} = 5$$