인공지능 수학

CONTENTS

CONTENTS

선형 대수 (Linear Algebra)

2. 벡터

선형 대수

- **선형 대수는 행렬(matrix)과 벡터(vector)를 다룬다.**
- 선형 대수는 어떤 함수가 선형(**linear**)일 때, 그 함수의 성질을 배우는 것이며, 선형은 다음과 같은 두개의 식으로 정의할 수 있습니다.

$$f(kx) = kf(x) \qquad\qquad f(x1 + x2) = f(x1) + f(x2)$$

- 우리가 알고 있는 대표적인 선형 함수는 $y = Wx$ 같은 정비례함수가 있다.

  또한 회전 변환, 확대, 축소 변환 역시 선형 함수이다.
- 비선형 함수는 이차 함수, 삼각함수, 로그 함수, 지수함수 등이 있다.
- 선형 대수 이론에 의하면 임의의 선형 함수는 행렬로 표현할 수 있으며,

  반대로 행렬은 어떤 선형 함수에 대응 한다.

선형 대수

- 그래서, 행렬에 대한 성질만 익힌다면 모든 선형 함수의 공통 성질을 알 수 있다. 일반적으로 비선형함수는 계산이 매우 복잡한데, 필요에 따라 선형으로 근사 시킬 수도 있습니다. 이렇게 하면 비선형함수에도 선형 대수를 적용할 수 있다.

행렬(Matrix)의 정의

행렬의 정의

- 아래와 같이 직사각형 모양으로 배열한 것이 행렬이며, 행렬의 각 수를 행렬의 성분이라 한다.

$$A = \begin{pmatrix} 80 & 90 & 85 \\ 95 & 70 & 75 \end{pmatrix}, \qquad \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

- 행렬의 가로줄을 행(**row**), 세로줄을 열(**column**)이라 한다. 위의 행렬은 행이 **2**개, 열이 **3**개 있는 행렬로 **2 × 3** 행렬이다. 행과 열의 개수가 모두 n개인 행렬을 n차 정사각 행렬이라고 한다.
- $(80 \quad 90 \quad 85)$ 처럼 하나의 행으로 이루어진 행렬을 행 벡터 (**row vector**),
- $\begin{pmatrix} 80 \\ 95 \end{pmatrix}$ 처럼 하나의 열로 이루어진 행렬을 열 벡터(**column vector**)라 한다.

행렬

$$A = \begin{pmatrix} 80 & 90 & 85 \\ 95 & 70 & 75 \end{pmatrix}, \qquad \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

- 위 행렬에서 **90**점과 같이 **1**행 **2**열에 있는 성분을 **1,2** 성분이라고 하고 a **12** 와 같이 표기 한다.

행렬의 상등

행렬의 상등

- $m \times n$ 행렬 A와 $p \times q$ 행렬 B에 대하여

  $m = p, n = q$일 때, (행렬 A와 B 의 행의 개수와 열의 개수가 같을 때)

  행렬 A와 B 는 같은 꼴이다.

  또 행렬 A와 B가 같은 꼴이고, 대응하는 성분도 각각 같을 때, 행렬 A와 B는 서로 같다고

  하며, 기호로 $A = B$ 로 표기한다.

이를테면 **2 × 2** 행렬이 서로 같을 조건은 다음과 같다.

두 행렬 $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$에 대하여

$$A = B \quad \Leftrightarrow \quad \begin{cases} a_{11} = b_{11}, & a_{12} = b_{12} \\ a_{21} = b_{21}, & a_{22} = b_{22} \end{cases}$$

행렬의 덧셈과 뺄셈

일반적으로 **두 행렬 A, B가 같은 꼴** 일 때,

A와 B의 대응하는 **각 성분의 합** 을 성분으로 하는 행렬을 **A와 B의 합** 이라 하고,

기호로 $A + B$ 와 같이 나타낸다.

또한 **두 행렬 A, B가 같은 꼴** 일 때,

A 의 각 성분에 대응하는 B의 성분을 뺀 결과를 가지는 행렬을 A와 B의 차라 하고,

기호로 $A - B$ 와 같이 나타낸다.

행렬의 덧셈과 뺄셈

이를테면 2×2 행렬의 덧셈과 뺄셈은 다음과 같으며,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$일 때

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix},$$

$$A - B = \begin{pmatrix} a_{11} - b_{11} & a_{12} - b_{12} \\ a_{21} - b_{21} & a_{22} - b_{22} \end{pmatrix}$$

행렬의 교환 법칙, 결합 법칙

행렬의 덧셈에서는 수의 덧셈에서 마찬가지로 교환 법칙, 결합법칙이 성립한다. 즉 같은 꼴의 행렬 A, B, C 에 대하여 다음과 같은 일반적인 법칙이 성립된다.

(1)교환 법칙 $A + B = B + A$

(2)결합 법칙 $(A + B) + C = A + (B + C)$

행렬의 실수 배

행렬 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$에 대해 다음과 같은 연산을 수행할 경우

$$A + A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2a & 2b \\ 2c & 2d \end{pmatrix}$$

$$A + A + A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 3a & 3b \\ 3c & 3d \end{pmatrix}$$

- 연산 결과인 $A + A$, $A + A + A$는 행렬 A 각 성분을 **2**배, **3**배 한 것임을 알 수 있다.
- 일반적으로 임의의 실수 k에 대하여 행렬 A의 각 성분을 k 배 한 것을 행렬 A를 k배 한 행렬이라 하고, 이것을 기호로 kA와 같이 나타낼 수 있다.

행렬의 실수 배

행렬 A가 2 × 2 행렬이고 k가 실수일 때, kA는 다음과 같이 나타낼 수 있다.

임의의 실수 k와 행렬 $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$에 대하여

$$kA = k\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} ka_{11} & ka_{12} \\ ka_{21} & ka_{22} \end{pmatrix}$$

또한, 임의의 행렬 A와 영행렬 O가 같은 꼴이고 k가 실수일 때, 행렬의 실수 배의 정의로부터 $1A = A$, $(-1)A = -A$, $0A = O$, $kO = O$ 같은 관계식이 성립하며, 일반적으로 행렬의 실수 배에 대하여 다음과 같은 성질이 있음을 알 수 있다. 같은 꼴의 행렬 A, B와 임의의 실수 k, l 에 대하여 다음의 관계식을 만족한다.

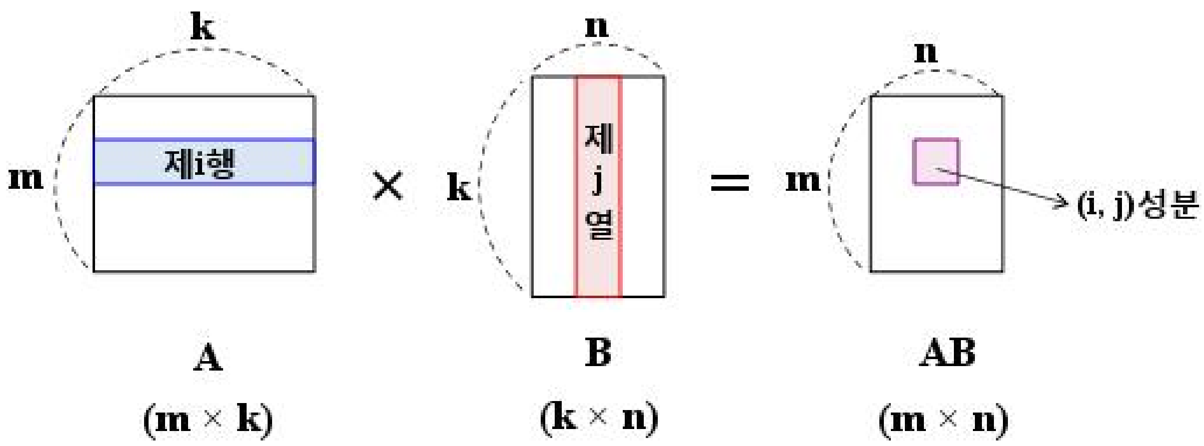
(1) $(kl)A = k(lA)$

(2) $(k + l)A = kA + lA$, $k(A + B) = kA + kB$

행렬의 곱

행렬의 곱

- 일반적으로 **m × k**행렬 A와 **k × n** 행렬 B에 대하여 행렬 A의 제i행의 성분과 행렬 B의 제j열의 성분을 차례로 곱하여 더한 값을 (i, j) 성분으로 하는 행렬을 행렬 A와 B의 곱이라 하고, 기호로 AB와 같이 나타냅니다.



벡터의 내적을 계산할 때 차원이 같아야만 내적을 계산할 수 있는 것처럼 행렬의 곱 **(dot product)**을 계산할 때는 앞 행렬의 열의 수와 뒤 행렬의 행의 수가 같아야만 계산을 할 수가 있다.

행렬의 곱

[참고] 행렬 A 열의 개수와 행렬 B 행의 개수가 같을 때에만 두 행렬의 곱 AB가 정의된다. 두 행렬 A, B가 모두 **2 × 2** 행렬일 때, 행렬 A와 B의 곱 AB는 다음과 같이 정의됨을 기억해야 합니다.

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \text{ 일 때}$$

$$AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

행렬의 곱

위 결과에서 볼 수 있듯이 일반적으로 행렬의 곱은 교환법칙이 성립하지 않는다.

행렬의 곱셈에서는 곱하는 순서를 함부로 바꾸면 안 된다는 것을 알 수 있다.

또한 n차원의 행 벡터와 열 벡터의 곱은 벡터의 내적과 같다는 것을 알 수 있다.

반대로 n 차원의 열 벡터와 행 벡터의 곱은 $n \times n$ 행렬이 됨을 알 수 있다.

$AB \neq BA$	**행렬의 곱셈은 교환법칙이 성립하지 않는다.** ★
$(AB)C = A(BC)$	행렬의 곱셈에 대한 결합법칙
$(A + B)C = AC + BC$	행렬의 덧셈과 곱셈에 대한 분배 법칙
$A(B + C) = AB + AC$	행렬의 곱셈과 덧셈에 대한 분배 법칙
$(kA)B = k(AB) = A(kB)$	행렬의 실수 배와 행렬의 곱에 대한 결합법칙

단위 행렬

- 실수 **1**은 곱셈에 대한 항등원이다. 즉 $a \cdot \mathbf{1} = \mathbf{1} \cdot a = a$ 이다.
- 이와 마찬가지로 n 차 정사각행렬은 행렬의 곱셈에 대한 항등원이 존재한다.
- 그 행렬을 단위 행렬이라고 하며 기호로 E, E_n 이나 I, I_n 혹은 **1** 등으로 나타낸다.

A가 정사각행렬일 때 $AE = EA = A$ 를 만족하는 단위 행렬이 존재한다. 이차 단위 행렬은 $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$이 되고, 삼차 단위 행렬 $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$이 된다. 여기서 반드시 기억해야 하는 것은 일반적으로 행렬의 곱은 교환법칙이 성립 하지 않지만, 단위 행렬과의 곱은 예외로 언제나 교환법칙이 성립한다는 사실이다. 이러한 특징은 꼭 기억해 두자.

대각합(trace)

행렬의 대각합(trace)은 정사각행렬의 주대각성분(\ )의 합으로 정의하고 있으며,

행렬 A의 대각합은 tr(A)로 표현하는데, 특히 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$일 경우 tr(A) = a + d 로

나타내고 있다.

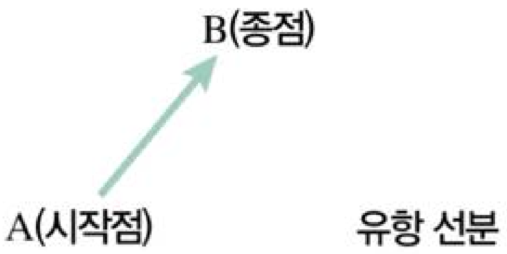
n차 정사각행렬 A의 i행 j열 성분을 a_{ij} 라 하면, 대각합은 $\sum$를 이용하여 다음과 같은 식으로 정의한다.

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i} a_{ii}$$

유향 성분과 벡터

두 점 A, B가 있을 때 A에서 B로 향하는 선분line segment을 생각해봅시다. 방향을 갖는 선분 AB를 유향선분directed segment이라고 합니다. 이때 A를 시작점, B를 종점이라고 합니다.

그림 2-13



유향선분 AB는 속성으로 점 A의 위치, B에 관한 방향, AB의 길이인 크기가 있습니다. 이 세 속성 중 방향과 크기만을 추상화한 양을 '벡터vector'라고 합니다. 보통 화살표선으로 표현합니다. 정리하면 다음과 같습니다.

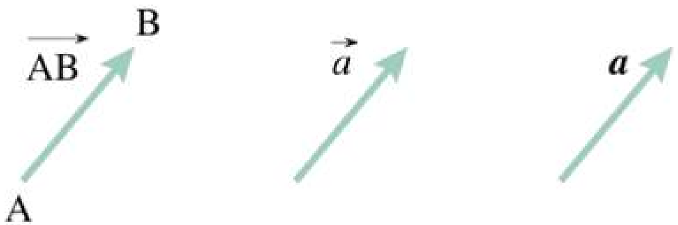
벡터는 방향과 크기를 갖는 양이며 화살표선으로 표현합니다.

유향 성분과 벡터

벡터는 방향과 크기를 갖는 양이며 화살표선으로 표현합니다.

유향선분 AB의 대표 벡터를 $(\overrightarrow{AB})$로 나타냅니다. 또한 화살을 딴 로마자인 $\vec{a}$나 굵은 $\boldsymbol{a}$ 문자 등으로 나타냅니다.

그림 2-14



벡터를 표시하는 기호는
여러 가지입니다.

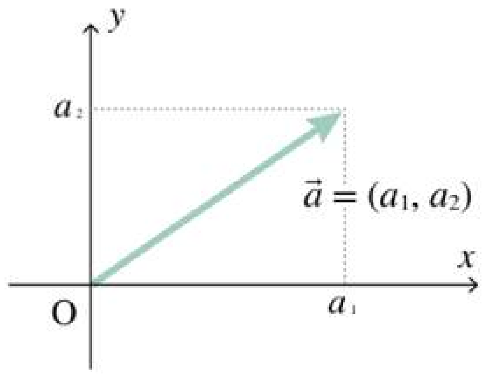
벡터의 성분 표시

벡터의 화살표선을 좌표 평면에 배치하여 표현할 수 있습니다. 화살의 시작점을 원점에 놓고 종점의 좌표로 벡터를 나타내는 것입니다. 이를 벡터의 성분 표시라고 합니다. 성분을 표시한 벡터 $\vec{a}$는 평면일 때 [식 2-10]처럼 표현합니다.

식 2-10

$$\vec{a} = (a_1, a_2)$$

그림 2-15



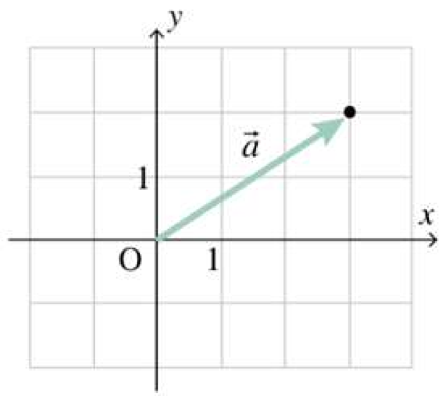
벡터의 성분 표시. '시작점을 원점으로 했을 때 종점의 좌표가 성분 표시'입니다.

벡터의 성분 표시

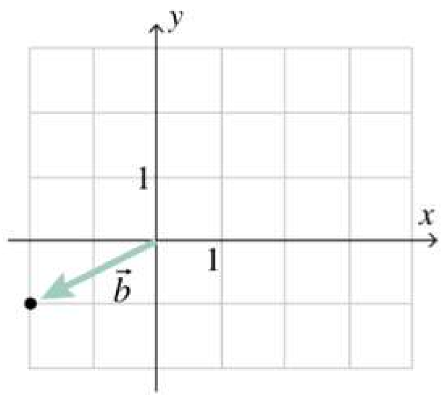
[그림 2−16]은 벡터 성분 표시의 다양한 예입니다.

그림 2-16

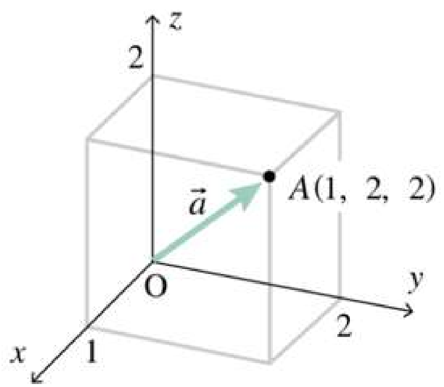
예 1) $\vec{a} = (3, 2)$를 나타내는 벡터



예 2) $\vec{b} = (-2, -1)$을 나타내는 벡터



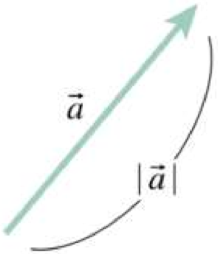
예 3) 3차원 공간의 경우도 마찬가지입니다.
예를 들어 $\vec{a} = (1, 2, 2)$는 오른쪽 그림의 벡터를 나타냅니다.



벡터의 크기 (2차원)

벡터를 나타내는 화살표선의 길이를 벡터의 크기라고 합니다. 벡터 $\vec{a}$의 크기는 $|\vec{a}|$[12]로 표현합니다.

그림 2-17

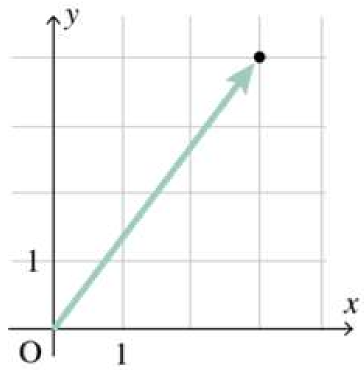


벡터의 크기를 계산하는 예는 [그림 2-18]과 같습니다.

그림 2-18[13]

예 1) $\vec{a} = (3, 4)$의 크기 $|\vec{a}|$는 오른쪽 그림을 참고해 다음처럼 계산할 수 있습니다.

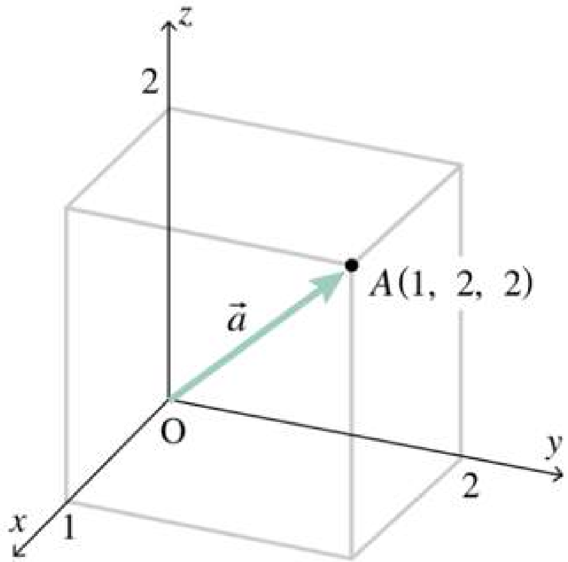
$$|\vec{a}| = \sqrt{3^2 + 4^2} = 5$$



벡터의 크기 (3차원)

예 2) 3차원 공간일 때도 계산 방식은 같습니다.
$\vec{a} = (1, 2, 2)$라는 벡터 공간이 있다면 크기 $|\vec{a}|$는 다음처럼 계산할 수 있습니다.

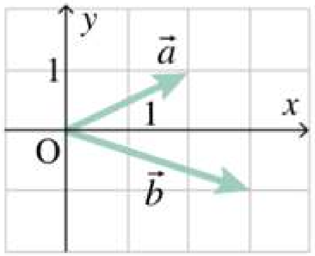
$$|\vec{a}| = \sqrt{1^2 + 2^2 + 2^2} = 3$$



벡터의 크기

또한 [그림 2-19]처럼 서로 다른 2개의 벡터가 있는 상황도 생각해봅시다.

그림 2-19



벡터 $\vec{a}$, $\vec{b}$ 각각의 크기를 다음처럼 계산할 수 있습니다.

$$|\vec{a}| = \sqrt{2^2 + 1^2} = \sqrt{5}$$

$$|\vec{b}| = \sqrt{3^2 + (-1)^2} = \sqrt{10}$$

벡터의 내적

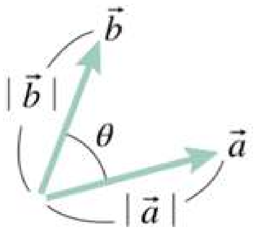
앞에서 벡터는 '크기와 방향을 갖는 양'이라고 정의했습니다. 따라서 벡터의 곱셈은 '크기와 방향'을 모두 고려해야 합니다. 이때 크기(스칼라)만 고려한 벡터의 곱셈을 내적이라고 합니다. 신경망의 입력과 출력은 모두 벡터 형태며 입력과 가중치의 내적을 출력으로 내보내므로 내적을 잘 알아두는 것이 좋습니다.

두 벡터 $\vec{a}$, $\vec{b}$의 내적은 $\vec{a} \cdot \vec{b}$로 표현하며 [식 2-11]로 구합니다. [그림 2-20]을 참고하면 더 이해하기 쉬울 겁니다.

식 2-11

$\vec{a} \cdot \vec{b} = |\vec{a}|\,|\vec{b}|\cos\theta$ (θ는 벡터 $\vec{a}$, $\vec{b}$가 구성하는 각도)

그림 2-20



벡터의 내적

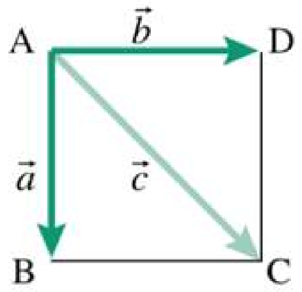
예를 들어 변의 길이가 1인 정사각형 ABCD가 있고, $\overrightarrow{AB}$를 벡터 $\vec{a}$, $\overrightarrow{AD}$를 벡터 $\vec{b}$, $\overrightarrow{AC}$를 벡터 $\vec{c}$라고 생각해봅시다. 벡터의 크기는 $|\vec{a}|=|\vec{b}|=1$, $|\vec{c}|=\sqrt{2}$ 입니다. 또한 벡터 $\vec{a}$와 $\vec{a}$가 이루는 각은 0°, $\vec{a}$와 $\vec{b}$가 이루는 각도는 90°, $\vec{a}$와 $\vec{c}$, $\vec{b}$와 $\vec{c}$가 이루는 각은 45°입니다. 이를 기반으로 다음처럼 벡터의 내적을 구할 수 있습니다.

$$\vec{a}\cdot\vec{a}=|\vec{a}||\vec{a}|\cos 0^\circ=|\vec{a}|^2=1^2=1$$

$$\vec{a}\cdot\vec{b}=|\vec{a}||\vec{b}|\cos 90^\circ=1\cdot1\cdot0=0$$

$$\vec{a}\cdot\vec{c}=|\vec{a}||\vec{c}|\cos 45^\circ=1\cdot\sqrt{2}\cdot\frac{1}{\sqrt{2}}=1$$

$$\vec{b}\cdot\vec{c}=|\vec{b}||\vec{c}|\cos 45^\circ=1\cdot\sqrt{2}\cdot\frac{1}{\sqrt{2}}=1$$



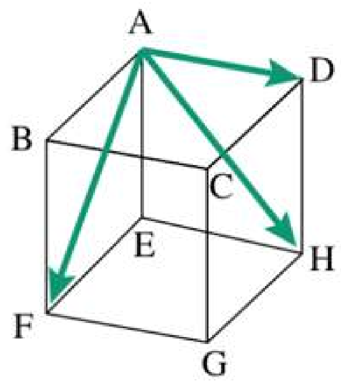
3차원일 때도 마찬가지입니다. 먼저 [그림 2-21]처럼 한 변의 길이가 3인 정육면체 ABCD-EFGH의 벡터 $\overrightarrow{AD}$, $\overrightarrow{AF}$, $\overrightarrow{AH}$가 있다고 생각해보겠습니다. 정사각형의 예처럼 $\overrightarrow{AD} \cdot \overrightarrow{AD}$ (0°), $\overrightarrow{AD} \cdot \overrightarrow{AF}$ (90°), $\overrightarrow{AF} \cdot \overrightarrow{AH}$ (60°)의 내적을 계산하면 다음과 같습니다.

그림 2-21

$$\overrightarrow{AD} \cdot \overrightarrow{AD} = \left|\overrightarrow{AD}\right|\left|\overrightarrow{AD}\right| \cos 0^\circ = 3 \cdot 3 \cdot 1 = 9$$

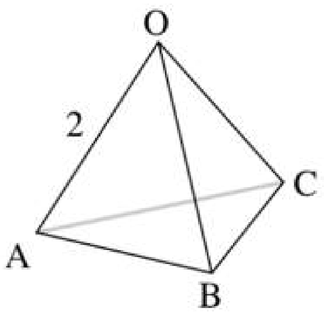
$$\overrightarrow{AD} \cdot \overrightarrow{AF} = \left|\overrightarrow{AD}\right|\left|\overrightarrow{AF}\right| \cos 90^\circ = 3 \cdot 3\sqrt{2} \cdot 0 = 0$$

$$\overrightarrow{AF} \cdot \overrightarrow{AH} = \left|\overrightarrow{AF}\right|\left|\overrightarrow{AH}\right| \cos 60^\circ = 3\sqrt{2} \cdot 3\sqrt{2} \cdot \frac{1}{2} = 9$$



이번에는 한 변의 길이가 2인 정사면체 OABC가 있을 때, 내적 $\overrightarrow{OA} \cdot \overrightarrow{OB}$ (60°)를 계산하겠습니다. [그림 2-22]와 같습니다.

그림 2-22



$\overrightarrow{OA} \cdot \overrightarrow{OB}$가 이루는 각도는 60°이므로

$$\overrightarrow{OA} \cdot \overrightarrow{OB} = \left|\overrightarrow{OA}\right|\left|\overrightarrow{OB}\right| \cos 60° = 2 \cdot 2 \cdot \frac{1}{2} = 2$$

코시-슈바르츠 부등식

[식 2-11]에서 중요한 공식을 도출할 수 있습니다. 임의의 $\cos\theta$는 $-1 \leqq \cos\theta \leqq 1$이라는 범위를 갖습니다. 이 범위에 벡터의 크기 $|\vec{a}||\vec{b}|$를 대입하면 다음 식이 성립합니다.

$$-|\vec{a}||\vec{b}| \leqq |\vec{a}||\vec{b}|\cos\theta \leqq |\vec{a}||\vec{b}|$$

이때 $|\vec{a}||\vec{b}|\cos\theta$는 벡터의 내적이므로 앞 식을 [식 2-12]로 바꿀 수 있습니다. 이를 코시-슈바르츠 부등식[14]이라고 합니다.

식 2-12

$$-|\vec{a}||\vec{b}| \leqq \vec{a} \cdot \vec{b} \leqq |\vec{a}||\vec{b}|$$

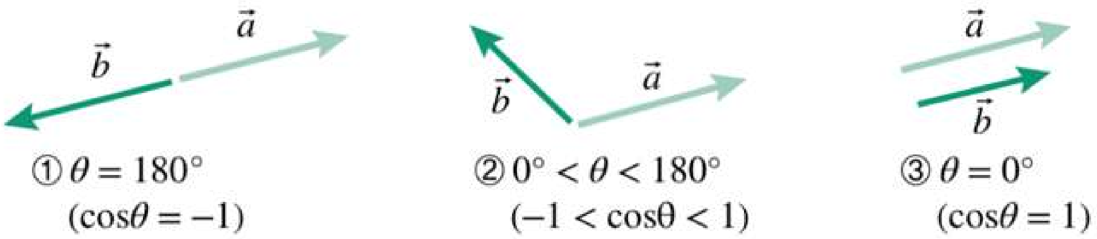
코시-슈바르츠 부등식

식 2-12

$$-|\vec{a}||\vec{b}| \leq \vec{a} \cdot \vec{b} \leq |\vec{a}||\vec{b}|$$

[식 2-12]를 그림으로 살펴봅시다. 2개의 벡터 $\vec{a}$, $\vec{b}$의 크기를 일정하게 하면 관계는 [그림 2-23]의 ①, ②, ③ 세 가지가 됩니다.

그림 2-23



[그림 2-23]의 ①~③은 다음과 같은 특징이 있습니다.

① 두 벡터가 반대 방향이라면 내적은 최솟값입니다.

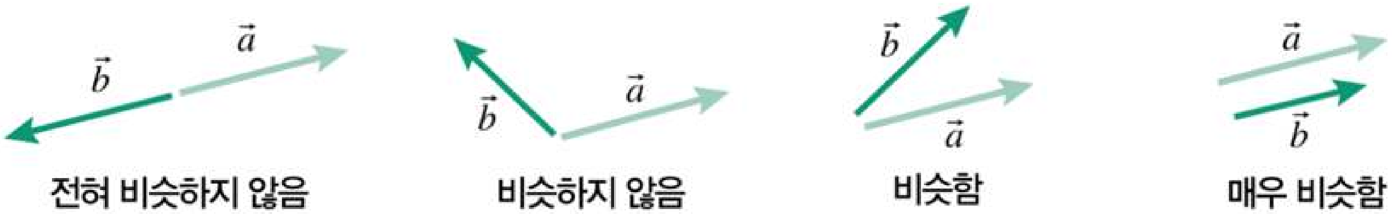
② 두 벡터가 평행하지 않다면 내적은 반대 방향일 때와 평행일 때 사이의 중간값입니다.

③ 두 벡터가 같은 방향이라면 내적은 최댓값입니다.

벡터의 내적의 의미

내적은 "두 벡터가 어느 정도로 같은 방향을 향하고 있는가?"를 나타냅니다. 벡터의 방향이 '비슷하다'라고 판단한다면 두 벡터의 내적은 커질 것입니다. 이러한 벡터의 방향별 내적의 크기를 나타내면 [그림 2-24]와 같습니다.

그림 2-24



내적으로 두 벡터의 상대적인 비슷함을 알 수 있습니다.

앞으로 합성곱 신경망을 살펴볼 때 이 개념이 중요합니다(부록 C 참고).

내적의 성분 표시

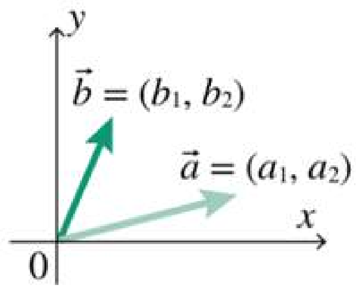
[식 2-10]을 내적의 성분 표시로 나타내보겠습니다. [그림 2-25] 같은 평면이라고 가정할 때 [식 2-13]이 성립합니다.

식 2-13

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2$$

$$(\vec{a} = (a_1, a_2), \vec{b} = (b_1, b_2))$$

그림 2-25



예를 들어 $\vec{a} = (2, 3)$이고 $\vec{b} = (5, 1)$이라면 내적 각각은 다음과 같습니다.

$$\vec{a} \cdot \vec{b} = 2 \cdot 5 + 3 \cdot 1 = 13,\ \vec{a} \cdot \vec{a} = 2 \cdot 2 + 3 \cdot 3 = 13,\ \vec{b} \cdot \vec{b} = 5 \cdot 5 + 1 \cdot 1 = 26$$

[그림 2-26] 같은 3차원 공간이라면 내적의 성분 표시는 [식 2-14]와 같습니다. 평면 벡터 내적 [식 2-13]에 z 성분을 덧붙였을 뿐입니다.[16]

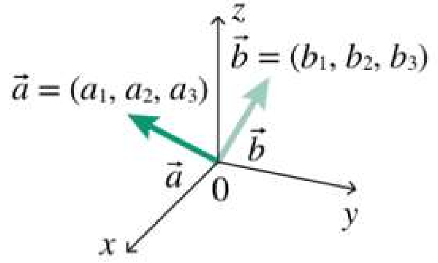
내적의 성분 표시

식 2-14

$$\vec{a} \cdot \vec{b} = a_1b_1+a_2b_2+a_3b_3$$

$$(\vec{a} = (a_1, a_2, a_3), \vec{b} = (b_1, b_2, b_3))$$

그림 2-26



그럼 해당 내적 계산의 예를 살펴보겠습니다. $\vec{a} = (2, 3, 2)$이고 $\vec{b} = (5, 1, -1)$이라면 내적 각각은 다음과 같습니다.

$$\vec{a} \cdot \vec{b} = 2 \cdot 5+3 \cdot 1+2 \cdot (-1) = 11,\ \vec{a} \cdot \vec{a} = 2 \cdot 2+3 \cdot 3+2 \cdot 2 = 17$$

또 다른 예 두 가지도 살펴보겠습니다. 이번에는 두 벡터 $\vec{a}$, $\vec{b}$의 내적 $\vec{a} \cdot \vec{b}$을 구하겠습니다. 첫 번째는 $\vec{a} = (2\sqrt{3}, 2)$, $\vec{b} = (1, \sqrt{3})$일 때의 내적 $\vec{a} \cdot \vec{b}$ 계산입니다.

$$\vec{a} \cdot \vec{b} = 2\sqrt{3} \cdot 1+2 \cdot \sqrt{3} = 4\sqrt{3}$$

두 번째는 $\vec{a} = (-3, 2, 1)$, $\vec{b} = (1, -3, 2)$일 때의 내적 $\vec{a} \cdot \vec{b}$ 계산입니다.

$$\vec{a} \cdot \vec{b} = -3 \cdot 1+2 \cdot (-3)+1 \cdot 2 = -7$$

벡터의 일반화

지금까지는 2차원 및 3차원 공간(평면 및 입체 공간)의 벡터를 살펴봤습니다. 벡터의 편리한 점은 평면과 입체 공간의 특징을 임의의 차원에 **그대로** 확장할 수 있다는 것입니다. 즉, 수만 차원의 공간을 처리하는 신경망도 2차원 및 3차원 벡터의 성질을 **그대로** 이용할 수 있습니다. 경사하강법(2장 10, 4장, 5장)에서 벡터를 사용하는 이유는 이러한 특징 때문입니다.

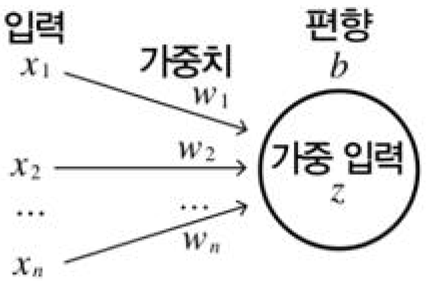
그럼 앞으로의 설명을 위해 지금까지 살펴본 2차원 및 3차원 공간 벡터의 식을 임의의 n차원으로 확장하겠습니다. 다음과 같습니다.

- 벡터 성분은 $\vec{a} = (a_1, a_2, \cdots, a_n)$로 표현할 수 있습니다.
- 두 벡터가 $\vec{a} = (a_1, a_2, \cdots, a_n)$, $\vec{b} = (b_1, b_2, \cdots, b_n)$일 때 내적의 성분 표시는 $\vec{a} \cdot \vec{b} = a_1b_1 + a_2b_2 + \cdots + a_nb_n$과 같습니다.
- 내적은 코시-슈바르츠 부등식 $-|\vec{a}||\vec{b}| \leq \vec{a} \cdot \vec{b} \leq |\vec{a}||\vec{b}|$이 성립해야 합니다.

신경망에 벡터의 내적 적용

그럼 벡터 이론을 실제 신경망에 적용하는 예를 살펴보겠습니다. [그림 2-27]처럼 복수의 입력 x_1, x_2, $\cdots$, x_n, 가중치 w_1, w_2, $\cdots$, w_n, 편향 b가 있을 때 가중 입력은 $z = w_1x_1+w_2x_2+ \cdots +w_nx_n+b$라고 설명했었습니다([식 1-8] 참고).

그림 2-27



이 가중 입력을 벡터로 간주해 다룰 수 있습니다. 입력과 가중치를 각각 벡터 성분으로 표시하면 $\vec{x} = (x_1, x_2, \cdots, x_n)$, $\vec{w} = (w_1, w_2, \cdots, w_n)$입니다. 따라서 가중 입력을 내적 형태로 바꾸면 $z = \vec{w} \cdot \vec{x}+b$가 됩니다.

이 예처럼 신경망에서 벡터는 많은 도움을 줄 수 있는 개념입니다.