3-2. Inner products

내적은 벡터의 길이나 두 벡터 간의 각도 또는 거리 등과 같은 직관적인 기하학적 개념을 도입하기 위한 것입니다.

내적의 주요 목적은 벡터가 서로 직교(orthogonal)인지 확인하는 것입니다.

3.2.1 Dot Product

- 결과 값이 스칼라이기 때문에 이렇게 **scalar product** 라고 부릅니다.
- $\mathbb{R}^n$, n차원 실수 공간

$$\boldsymbol{x}^\top \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i \,.$$

3.2.2 General Inner Products

지금까지 한 개의 벡터를 다뤘습니다. → Linear Mapping

bilinear mapping : Ω 는 두 인자(벡터)에 대한 mapping 이며, 각 인자에 대해 **linear** 합니다.

vector space V 와 모든 $x, y, z \in V$ 와 $\lambda, \psi \in \mathbb{R}$ 에 대해 다음이 성립합니다.

$$\Omega(\lambda \boldsymbol{x} + \psi \boldsymbol{y}, \boldsymbol{z}) = \lambda \Omega(\boldsymbol{x}, \boldsymbol{z}) + \psi \Omega(\boldsymbol{y}, \boldsymbol{z}) \tag{3.6}$$
$$\Omega(\boldsymbol{x}, \lambda \boldsymbol{y} + \psi \boldsymbol{z}) = \lambda \Omega(\boldsymbol{x}, \boldsymbol{y}) + \psi \Omega(\boldsymbol{x}, \boldsymbol{z}) \,. \tag{3.7}$$

- (3.6) → 첫 번째 인자에 대한 Linear한지를 확인
- (3.7) → 두 번째 인자에 대한 Linear한지를 확인

정의 3.2 (Symmetric and Positive definite)

vector space V

$\Omega : V \times V \rightarrow \mathbb{R}$, 두 개의 벡터를 인자로 취해 하나의 실수로 매핑하는 bilinear mapping

- Symmetric (대칭성)
  - 모든 $x, y \in V$ 에 대해 다음을 만족한다면, symmetric(대칭성)이라고 합니다.
  - $\Omega(x, y) = \Omega(y, x)$
- Positive definite (양의 정부호)
  - 모든 $x \in V \setminus \{0\}$ 에 대해 다음을 만족한다면, positive definite(양의 정부호)라고 합니다.
  - $\Omega(x, x) = 0, \;\; \Omega(0, 0) = 0$

정의 3.3 (Inner product)

vector space V

$\Omega : V \times V \rightarrow \mathbb{R}$, 두 개의 벡터를 인자로 취해 하나의 실수로 매핑하는 bilinear mapping

- 위 두가지 조건(정의 3.2)을 모두 만족하는 bilinear mapping을 **Inner Product (내적)**이라고 부릅니다.
  - 기호로는 보통 $\langle x, y \rangle$을 사용합니다.
  - 이때, 벡터 공간 V는 **Inner Product Space(내적 공간)** 라고 부릅니다.
    - $(V, \langle \cdot, \cdot \rangle)$

예시 3.3 (Inner Product That is Not the Dot Product)

벡터 공간 $V = \mathbb{R}^2$ 에서, Dot product가 아닌 내적을 정의했습니다.

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := x_1y_1 - (x_1y_2 + x_2y_1) + 2x_2y_2 \tag{3.9}$$

우선 내적(inner product)가 되기위해서는 세 가지 조건을 만족해야합니다.

1. **Bilinearlity**

   a. $x = (x_1, x_2), \;\; y = (y_1, y_2), \;\; z = (z_1, z_2), \;\; \lambda, \psi \in \mathbb{R}$

   ▼ 첫 번째 인자

   > $$\langle \lambda x + \psi z, \; y \rangle = \lambda \langle x, y \rangle + \psi \langle z, y \rangle$$

> $\langle x, \lambda y + \psi z\rangle \ = x_1(\lambda y_1 + \psi z_1) \ - [x_1(\lambda y_2 + \psi z_2) + x_2(\lambda y_1 + \psi z_1)] \ + 2x_2(\lambda y_2 + \psi z_2)$
>
> $= \lambda(x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2) \ + \psi(x_1z_1 - x_1z_2 - x_2z_1 + 2x_2z_2)$
>
> $= \lambda\langle x, y\rangle + \psi\langle x, z\rangle$

▼ 두 번째 인자

> $\langle x,\ \lambda y + \psi z\rangle = \lambda\langle x, y\rangle + \psi\langle x, z\rangle$
>
> $\langle x,\ \lambda y + \psi z\rangle \ = x_1(\lambda y_1 + \psi z_1) - [x_1(\lambda y_2 + \psi z_2) + x_2(\lambda y_1 + \psi z_1)] \ + 2x_2(\lambda y_2 + \psi z_2)$
>
> $= \lambda(x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2) \ + \psi(x_1z_1 - x_1z_2 - x_2z_1 + 2x_2z_2)$
>
> $= \lambda\langle x, y\rangle + \psi\langle x, z\rangle$

2. **Symmetry**

   $\langle x, y\rangle \ = x_1y_1 - (x_1y_2 + x_2y_1) + 2x_2y_2$

   $\langle y, x\rangle \ = y_1x_1 - (y_1x_2 + y_2x_1) + 2y_2x_2$

   $= x_1y_1 - (x_1y_2 + x_2y_1) + 2x_2y_2 = \langle x, y\rangle$

3. **Positive Definite**

$$
\begin{aligned}
&\langle x, x\rangle \ = x_1^2 - 2x_1x_2 + 2x_2^2 \ = (x_1 - x_2)^2 + x_2^2 \\
&(x_1 - x_2)^2 + x_2^2 \geq 0 \\
&\therefore \quad \langle x, x\rangle > 0 \quad \text{if } x \neq 0 \quad \text{and} \quad \langle 0, 0\rangle = 0 \quad \checkmark \text{성립}
\end{aligned}
$$

3.2.3 Symmetric, Positive Definite Matrices

머신러닝에서 Symmetric, positive definite matrices는 중요한 역할을 합니다.

이 행렬은 내적(inner product)을 기반으로 정의되며, 커널 이론이나 행렬 분해에서 쓰입니다.

어떤 n-차원 벡터 공간 V에서 내적은 $\langle\cdot,\cdot\rangle : V \times V \rightarrow \mathbb{R}$ 로 정의되어 있다고 가정합시다.

V의 ordered Basis $B = (b_1, \ldots, b_n)$에 대해, 모든 벡터 $x, y \in V$는 다음과 같이 표현할 수 있습니다.

$x = \sum_{i=1}^{n} \psi_i b_i, \quad y = \sum_{j=1}^{n} \lambda_j b_j$

내적의 bilinearity 에 의해 다음과 같은 성질이 성립합니다.

$$
\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \left\langle \sum_{i=1}^{n} \psi_i \boldsymbol{b}_i, \sum_{j=1}^{n} \lambda_j \boldsymbol{b}_j \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \psi_i \langle \boldsymbol{b}_i, \boldsymbol{b}_j\rangle \lambda_j = \hat{\boldsymbol{x}}^\top \boldsymbol{A} \hat{\boldsymbol{y}}, \quad (3.10)
$$

- $\hat{x}, \hat{y}$: 벡터 x, y의 기저 B에 대한 좌표 벡터
- $A_{ij} := \langle b_i, b_j\rangle$: 기저 벡터들끼리의 내적값으로 구성된 행렬

또한 만약 내적이 대칭(Symmetric)이라면 $A = A^T$, positive definite 라면 $\forall x \in V \setminus \{0\} : x^T A x > 0$ 입니다.

만약 위 조건에서 $\forall x \in V \setminus \{0\} : x^T A x \geq 0$ 이라면, A는 symmetirc, positive semidefinite라고 부릅니다.

예시 3.4 (Symmetric, Positive Definite Matrices)

Example 3.4 (Symmetric, Positive Definite Matrices)
Consider the matrices

$$\boldsymbol{A}_1 = \begin{bmatrix} 9 & 6 \\ 6 & 5 \end{bmatrix}, \quad \boldsymbol{A}_2 = \begin{bmatrix} 9 & 6 \\ 6 & 3 \end{bmatrix}. \tag{3.12}$$

$\boldsymbol{A}_1$ is positive definite because it is symmetric and

$$\boldsymbol{x}^\top \boldsymbol{A}_1 \boldsymbol{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 9 & 6 \\ 6 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{3.13a}$$

$$= 9x_1^2 + 12x_1x_2 + 5x_2^2 = (3x_1 + 2x_2)^2 + x_2^2 > 0 \tag{3.13b}$$

for all $\boldsymbol{x} \in V \backslash \{\boldsymbol{0}\}$. In contrast, $\boldsymbol{A}_2$ is symmetric but not positive definite because $\boldsymbol{x}^\top \boldsymbol{A}_2 \boldsymbol{x} = 9x_1^2 + 12x_1x_2 + 3x_2^2 = (3x_1 + 2x_2)^2 - x_2^2$ can be less than 0, e.g., for $\boldsymbol{x} = [2, -3]^\top$.

- A_2의 경우 $x = [2, -3]^T$ 일 경우에 positive definite를 만족하지 않습니다.

출처

- Mathmatics for Machine Learning (https://github.com/mml-book/mml-book.github.io)
- https://blog.naver.com/walk_along/222187729940
- https://junstar92.github.io/mml-study-note/2022/07/07/ch3-2.html