CH 3. Analytic Geomery

CH2 에서는 vector와 vector space, linear mapping을 추상적인 단계에서 배움.

CH3 에서는 geometric (기하) vector와 그 length, distance 에 대해 공부한다.

이걸 설명하기 위해

- inner product (내적)을 적용한 vector space
- norm
- metrics

를 공부한다.

3.1) Norms

↳ 이 개념을 통해 length of vectors를 알수 있다.

Definition 3.1 (Norm). A *norm* on a vector space V is a function

$$\|\cdot\| : V \to \mathbb{R}, \tag{3.1}$$

$$\boldsymbol{x} \mapsto \|\boldsymbol{x}\|, \tag{3.2}$$

which assigns each vector $\boldsymbol{x}$ its *length* $\|\boldsymbol{x}\| \in \mathbb{R}$, such that for all $\lambda \in \mathbb{R}$ and $\boldsymbol{x}, \boldsymbol{y} \in V$ the following hold:

- *Absolutely homogeneous:* $\|\lambda \boldsymbol{x}\| = |\lambda| \|\boldsymbol{x}\|$
- *Triangle inequality:* $\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$
- *Positive definite:* $\|\boldsymbol{x}\| \geqslant 0$ and $\|\boldsymbol{x}\| = 0 \iff \boldsymbol{x} = \boldsymbol{0}$

Example 3.1 (Manhattan Norm)

The *Manhattan norm* on $\mathbb{R}^n$ is defined for $\boldsymbol{x} \in \mathbb{R}^n$ as

Manhattan norm 기호 →

$$\|\boldsymbol{x}\|_1 := \sum_{i=1}^{n} |x_i|, \tag{3.3}$$

where $|\cdot|$ is the absolute value. The left panel of Figure 3.3 shows all vectors $\boldsymbol{x} \in \mathbb{R}^2$ with $\|\boldsymbol{x}\|_1 = 1$. The Manhattan norm is also called ℓ_1 *norm*.

Example 3.2 (Euclidean Norm)

The *Euclidean norm* of $\boldsymbol{x} \in \mathbb{R}^n$ is defined as

$$\|\boldsymbol{x}\|_2 := \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{\boldsymbol{x}^\top \boldsymbol{x}} \tag{3.4}$$

and computes the *Euclidean distance* of $\boldsymbol{x}$ from the origin. The right panel of Figure 3.3 shows all vectors $\boldsymbol{x} \in \mathbb{R}^2$ with $\|\boldsymbol{x}\|_2 = 1$. The Euclidean norm is also called ℓ_2 *norm*.

3.2) Inner products.

- geometrical concept 을 직관적으로 이해하기 위한 도입. → inner product.
  - length of vector, angle or distance between 2 vectors

- inner product의 주된 목적은. 두 vector가 직각인지 아닌지 여부를 알기 위해서이다.

3.2.1) Dot product.

우리에게 익숙한 particular 내적.

We may already be familiar with a particular type of inner product, the *scalar product/dot product* in $\mathbb{R}^n$, which is given by

$$\boldsymbol{x}^\top \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i \,. \tag{3.5}$$

We will refer to this particular inner product as the dot product in this book. However, inner products are more general concepts with specific properties, which we will now introduce.

1st

3.2.2) General Inner Products

좀 더 general 한 내적

- linear mapping : we can arrange the mapping with respect to addition & multiplication with scalar.
- bilinear mapping Ω : 2개의 벡터에 대하여 linear한 성질이 그대로 보존되는 mapping.

Recall the linear mapping from Section 2.7, where we can rearrange the mapping with respect to addition and multiplication with a scalar. A *bilinear mapping* Ω is a mapping with two arguments, and it is linear in each argument, i.e., when we look at a vector space V then it holds that for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V,\ \lambda, \psi \in \mathbb{R}$ that

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

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

Here, (3.6) asserts that Ω is linear in the first argument, and (3.7) asserts that Ω is linear in the second argument (see also (2.87)).

Ω is called

- symmetric : if) $\Omega(x,y) = \Omega(y,x)$ 순서 상관 X.
- positive definite : if)
  0이 아닌 벡터 x에 대해서
  $\Omega(x,x) > 0$ 이고 $\Omega(0,0)=0$ 을 만족할 경우.

Definition 3.2. Let V be a vector space and $\Omega : V \times V \to \mathbb{R}$ be a bilinear mapping that takes two vectors and maps them onto a real number. Then

- Ω is called *symmetric* if $\Omega(\boldsymbol{x}, \boldsymbol{y}) = \Omega(\boldsymbol{y}, \boldsymbol{x})$ for all $\boldsymbol{x}, \boldsymbol{y} \in V$, i.e., the order of the arguments does not matter.
- Ω is called *positive definite* if

$$\forall \boldsymbol{x} \in V \backslash \{\boldsymbol{0}\} : \Omega(\boldsymbol{x}, \boldsymbol{x}) > 0 \,, \quad \Omega(\boldsymbol{0}, \boldsymbol{0}) = 0 \,. \tag{3.8}$$

Definition 3.3. Let V be a vector space and $\Omega : V \times V \to \mathbb{R}$ be a bilinear mapping that takes two vectors and maps them onto a real number. Then

- A positive definite, symmetric bilinear mapping $\Omega : V \times V \to \mathbb{R}$ is called an *inner product* on V. We typically write $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ instead of $\Omega(\boldsymbol{x}, \boldsymbol{y})$.
- The pair $(V, \langle \cdot, \cdot \rangle)$ is called an *inner product space* or (real) *vector space with inner product*. If we use the dot product defined in (3.5), we call $(V, \langle \cdot, \cdot \rangle)$ a *Euclidean vector space*.

We will refer to these spaces as inner product spaces in this book.

Example 3.3 (Inner Product That Is Not the Dot Product)
Consider $V = \mathbb{R}^2$. If we define

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := x_1 y_1 - (x_1 y_2 + x_2 y_1) + 2 x_2 y_2 \tag{3.9}$$

then $\langle \cdot, \cdot \rangle$ is an inner product but different from the dot product. The proof will be an exercise.

3.2.3 Symmetric, Definite Matrices

Symmetric, positive definite 행렬은 내적으로 정의가 되며, 머신러닝에서 중요한 역할을 합니다. 뒤에 4.3장에서 다시 이와 관련된 내용으로 살펴볼 것입니다.
n차원의 inner product space (V, <·,·>)가 존재하고 또한 V의 ordered basis B=(b1, b2, ... , bn)가 있습니다. 그러면 V의 임의의 벡터 x, y는 b1, b2, ..., bn의 일차결합으로 각각 표현이 될 수 있습니다.

$$\boldsymbol{x} = \sum_{i=1}^{n} \psi_i \boldsymbol{b_i}, \quad \boldsymbol{y} = \sum_{i=1}^{n} \lambda_i \boldsymbol{b_i}.$$

이때 <x,y>를 다음과 같이 표현할 수 있는데요,

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

$$A_{ij} := \langle \boldsymbol{b}_i, \boldsymbol{b}_j \rangle$$

이를 통해서 살펴보면 inner product <x, y>는 행렬 A에 의해 유일하게 결정이 되며, 또한 A라는 행렬 또한 symmetric하다는 것을 알 수 있습니다. (왜냐면 <x,y>=<y,x>가 성립되는 inner product니까요) 또한 inner product의 posifive definiteness에 의해 다음이 성립합니다.

$$\forall \boldsymbol{x} \in V \backslash \{\boldsymbol{0}\} : \boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x} > 0. \tag{3.11}$$

3.3) Lengths & Distances

· 내적을 통해 norm을 정의하면

$$\|x\| := \sqrt{\langle x, x \rangle}$$

* 코시-슈바르츠 부등식

$$(a^2+b^2)(x^2+y^2) \ge (ax+by)^2$$

⇓

두 벡터 $v=(a,b)$, $w=(x,y)$가 있을 때,

$$\|v\| \cdot \|w\| \ge |\langle v, w \rangle|$$

norm(실수값)들의 곱

벡터의 내적(실수값)의 절대값.

[정의] 거리(distance)와 metric

inner product space인 (V, <·,·>)에 대해서 두 벡터 x, y 사이의 거리(distance)는

$$d(\boldsymbol{x}, \boldsymbol{y}) := \|\boldsymbol{x} - \boldsymbol{y}\| = \sqrt{\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle}$$

로 정의합니다. 만약 우리가 dot product를 내적으로 사용할 경우 Euclidean distance로 정의합니다.

이를 mapping으로 이해를 해보면

$$d : V \times V \to \mathbb{R}$$
$$(\boldsymbol{x}, \boldsymbol{y}) \mapsto d(\boldsymbol{x}, \boldsymbol{y})$$

거리는 두 벡터를 하나의 실수값으로 대응시키는 mapping이라고 할 수 있습니다. 이를 metric이라고 합니다.

고등학교 기하 시간에 벡터를 공간벡터로 좌표로 옮겨놓은 것을 생각해보면 좀 더 이해가 쉽습니다. Dot product로 정의된 Euclidean inner product의 정의된 경우 norm은 원점에서부터 벡터에 대응되는 점까지의 거리이고, 두 벡터 사이의 거리는 두 점 사이의 거리로 직관적으로 생각할 수 있습니다.

그리고 distance와 metric은 같은 내용이지만 관점의 차이가 조금 있습니다. distance는 측정값이라는 개념과 기하적인 의미를 가진 반면 metric은 우리가 2장에서 살펴봤던 것과 같이 관계적인 개념과 대수적인 의미가 있다는 것을 알 수 있습니다.

[참고-Remark]

벡터의 길이와 마찬가지로, 벡터 사이의 거리를 정의할 때는 norm의 개념을 반드시 필수적으로 필요로 합니다. (내적은 필수가 아닙니다. 두 벡터의 차의 norm로 정의를 합니다.) 만약 내적에 의해 norm을 정의했다면, 내적을 무엇으로 정의했느냐에 따라 distance(거리)는 달라집니다.

사실 metric d를 해석학이나 다른 과목에서 공부를 할 때는 다음의 조건들을 만족하는 함수라고 정의하기도 합니다. 필요충분조건이기도 하고, 어떤 관점에서 정의로 다가가느냐에 따라 다를 뿐 수학적인 본질은 같은 내용입니다.

(1) d는 positive definite : 모든 벡터 ***x***, ***y***에 대하여 $d(\boldsymbol{x},\boldsymbol{y}) \ge 0$ 입니다. $d(\boldsymbol{x},\boldsymbol{y})=0$과 ***x***=***y***는 동치조건입니다.

(2) d는 symmetric : 모든 벡터 ***x***, ***y***에 대하여 $d(\boldsymbol{x}, \boldsymbol{y})=d(\boldsymbol{y}, \boldsymbol{x})$가 성립합니다.

(3) triangular inequality : $d(\boldsymbol{x},\boldsymbol{z}) \le d(\boldsymbol{x},\boldsymbol{y}) + d(\boldsymbol{y},\boldsymbol{z})$

3.4) Angles and Orthogonality

* 코시-슈바르츠 부등식을 이용하여

두 벡터 사이의 각도 구하기

$\boldsymbol{x} \neq \boldsymbol{0}, \boldsymbol{y} \neq \boldsymbol{0}$ 인 벡터 $\boldsymbol{x}, \boldsymbol{y}$에 대하여

$$|<\boldsymbol{x}, \boldsymbol{y}>| \le \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|$$

$$\Rightarrow \quad -1 \le \frac{<\boldsymbol{x}, \boldsymbol{y}>}{\|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|} \le 1$$

따라서 $\omega \in [0, \pi]$에 대하여

$$\cos\omega = \frac{<\boldsymbol{x}, \boldsymbol{y}>}{\|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|}$$

를 만족하는 ω를 두 벡터 $\boldsymbol{x}, \boldsymbol{y}$ 사이의 각이라고 정의합니다.

→ 이렇게 각도를 정의하면,

두 벡터가 수직(Orthogonal)인지 아닌지를 쉽게 알 수 있다.

[수직 - Orthogonality]

두 벡터 ***x***,***y***가 <***x***, ***y***>=0 이면 orthogonal(수직)이라고 하고 $\boldsymbol{x} \perp \boldsymbol{y}$라고 표현합니다. 만약 $\|\boldsymbol{x}\|=1=\|\boldsymbol{y}\|$ (단위벡터-unit vector)이면서 수직이면 두 벡터는 orthonormal(정규직교)이라고 합니다.

만약 두 벡터 ***x***,***y***가 ***x***=(1,1)=[1,1]^T 이고 ***y***=(-1,1)=[-1, 1]^T 라고 한 번 가정을 해보면 두 벡터는 수직일까요? **수직 여부를 판단하기 전에 '내적'을 무엇을 정의했는지를 먼저 확인해야 합니다.** 만약에 dot product로 정의를 했다면 <***x***,***y***>=-1+1=0 이므로 수직이 된 것을 알 수 있습니다. 그러나 만약에

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^\top \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{y}$$

라고 정의했다면 상황이 달라집니다. 그 때는

$$<\boldsymbol{x}, \boldsymbol{y}> = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} [-1 \quad 1] = \begin{bmatrix} 1 \\ 1 \end{bmatrix} [-2 \quad 1] = -2+1 = -1$$

내적의 값이 이렇게 달라지고, 이 때 두 벡터 사이의 각은

$$\cos\omega = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\| \|\boldsymbol{y}\|} = -\frac{1}{3} \implies \omega \approx 1.91 \text{ rad} \approx 109.5^\circ$$

이 되므로 서로 수직이 아닙니다. 따라서 수직인지 아닌지는 어떤 내적을 적용했는지에 따라 달라진다는 것을 알 수 있습니다.

Orthogonal Matrix

수직을 정의했으니, Orthogonal Matrix라고 하는 특별한 행렬을 한 번 생각해보려고 합니다. 우선 정의는 아래와 같습니다.

[정의 - Orthogonal Matrix]

nxn 행렬 A에 대하여

$$\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{I} = \boldsymbol{A}^\top \boldsymbol{A}$$

를 만족하면 **orthogonal matrix(수직 행렬)**이라고 합니다. 즉

$$\boldsymbol{A}^{-1} = \boldsymbol{A}^\top$$

를 만족한다는 것을 알 수 있습니다. (대각성분끼리 대칭시키면 역행렬이라니!!)

이 행렬은 기하적으로 특별한 변환을 나타냅니다. 우리가 살펴본 norm은 하나의 벡터의 크기를 의미하는 개념이었고, 길이와 각은 두 벡터 사이에 적용되는 개념이었죠? 그러면 하나의 벡터 ***x***를 이 orthogonal 행렬에 의해 변환한 ***Ax***는 ***x***와 어떤 관계가 있을까요?

1) **이 orthogonal 행렬은 '길이'를 그대로 보존합니다.** 예를 들어 dot product에 대해서 살펴보면

$$\|\boldsymbol{A}\boldsymbol{x}\|^2 = (\boldsymbol{A}\boldsymbol{x})^\top(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^\top \boldsymbol{I}\boldsymbol{x} = \boldsymbol{x}^\top \boldsymbol{x} = \|\boldsymbol{x}\|^2$$

Ax의 길이와 ***x***의 길이가 같다는 것을 알 수 있습니다.

2) **이 orthogonal 행렬은 '두 벡터 사이의 각'을 그대로 보존합니다.** 예를 들어 ***x***, ***y***라는 벡터가 있을 때 이 벡터 사이의 각을 ω라고 할 때, orthogonal 행렬에 의해 옮겨진 ***Ax***, ***Ay*** 사이의 각을 살펴보면

$$\cos\omega = \frac{(\boldsymbol{A}\boldsymbol{x})^\top(\boldsymbol{A}\boldsymbol{y})}{\|\boldsymbol{A}\boldsymbol{x}\| \|\boldsymbol{A}\boldsymbol{y}\|} = \frac{\boldsymbol{x}^\top \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{y}}{\sqrt{\boldsymbol{x}^\top \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{x}\boldsymbol{y}^\top \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{y}}} = \frac{\boldsymbol{x}^\top \boldsymbol{y}}{\|\boldsymbol{x}\| \|\boldsymbol{y}\|}$$

원래의 각과 동일하다는 것을 알 수 있습니다. **즉 orthogonal matrix *A*는 각 벡터의 길이를 보존하고, 두 벡터 사이의 각을 유지하므로 두 벡터 사이의 거리도 동일하게 유지한다는 것을 알 수 있습니다.** 이러한 행렬은 rotation과 관련이 있는데, 이는 3.9장에서 좀 더 살펴볼 예정입니다.

3.5 Orthonormal Basis

· Orthonomal Basis

- 2개의 벡터가 서로 수직이고
- 둘다 길이가 1인 경우.

Definition 3.9 (Orthonormal Basis). Consider an n-dimensional vector space V and a basis $\{\boldsymbol{b}_1, \dots, \boldsymbol{b}_n\}$ of V. If

$$\langle \boldsymbol{b}_i, \boldsymbol{b}_j \rangle = 0 \quad \text{for } i \neq j \tag{3.33}$$

$$\langle \boldsymbol{b}_i, \boldsymbol{b}_i \rangle = 1 \tag{3.34}$$

for all $i, j = 1, \dots, n$ then the basis is called an *orthonormal basis* (*ONB*). If only (3.33) is satisfied, then the basis is called an *orthogonal basis*. Note that (3.34) implies that every basis vector has length/norm 1.

Recall from Section 2.6.1 that we can use Gaussian elimination to find a basis for a vector space spanned by a set of vectors. Assume we are given a set $\{\tilde{\boldsymbol{b}}_1, \dots, \tilde{\boldsymbol{b}}_n\}$ of non-orthogonal and unnormalized basis vectors. We concatenate them into a matrix $\tilde{\boldsymbol{B}} = [\tilde{\boldsymbol{b}}_1, \dots, \tilde{\boldsymbol{b}}_n]$ and apply Gaussian elimination to the augmented matrix (Section 2.3.2) $[\tilde{\boldsymbol{B}}\tilde{\boldsymbol{B}}^\top | \tilde{\boldsymbol{B}}]$ to obtain an orthonormal basis. This constructive way to iteratively build an orthonormal basis $\{\boldsymbol{b}_1, \dots, \boldsymbol{b}_n\}$ is called the *Gram-Schmidt process* (Strang, 2003).

* Gram-schmidt process

: Gaussian elimination을 이용해서 orthonormal basis 찾기.

a set of vectors로 표현되는 vector space의 basis를 찾는 방법.

given a set $\{\tilde{b}_1, \dots, \tilde{b}_n\}$ of non-orthogonal and unnormalized basis vectors

연결 시킴.

matrix $\tilde{B} = [\tilde{b}_1, \dots, \tilde{b}_n]$ and apply gaussian elimination to the augmented matrix $[\tilde{B}\tilde{B}^T | \tilde{B}]$ to obtain an orthonormal basis. $\{b_1, \dots, b_n\}$

3.6 Orthogonal Complement.

Orthogonality를 정의했으니, 이제 서로 수직인 vector spaces에 대해 알아보자.

D-dimensional vector space V &

M-dimensional sub-space $U \subseteq V$ 가 있다고 전제.

⇒ orthogonal complement $U^\perp$는

- $(D-M)$-dimensional subspace of V 이고.
- V 안에서 U에 존재하는 모든 vector에 대해 수직인 모든 vector들을 포함한다.

⇒ $U \cap U^\perp = \{0\}$

∴ $x \in V$ 모든 벡터 x는. 다음과 같이 유니크하게 분해될 수 있다.

$$\boldsymbol{x} = \sum_{m=1}^{M} \lambda_m \boldsymbol{b}_m + \sum_{j=1}^{D-M} \psi_j \boldsymbol{b}_j^\perp, \quad \lambda_m, \psi_j \in \mathbb{R},$$

이때 $(b_1, \dots, b_M)$은 U의 basis 이고.

$(b_1^\perp, \dots, b^\perp_{D-M})$은 $U^\perp$의 basis 이다.

⇒ 따라서 orthogonal complement는.

(two-dimensional subspace)인 plane U는.

3차원 벡터 공간에서 describe 하는게 가능하다.

+) $\|w\| = 1$ 이고. U에 수직인 벡터 w는.

$U^\perp$의 basis vector 이다.

⇒ vector w는. U의 normal vector라고 부른다.