Ch3. Analytic Geometry

Geometric vector(기하학적 벡터)로 두 벡터 사이의 길이, 거리, 각을 계산함

이를 위해서 벡터공간에 내적을 준비하여 벡터공간을 기하학적으로 알아봄

내적과 상응하는 놈, 기준들은 유사성과 거리의 직관적인 개념을 포착하는데, SVM 에서 활용됨

3.1 Norms

Norm(놈): 벡터의 크기 또는 길이를 측정하는 방법

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 $\|\boldsymbol{x}\| = 0 \iff \boldsymbol{x} = \boldsymbol{0}$

-완전히 동차

-코시 슈바르츠 부등식(Cauchy - Schwarz Inequality) 기반, 어느 삼각형이든 두 변의 길이는 나머지 변의 길이보다 같거나 큼

-양의 정부호 행렬. 양수이자 실수인 고유값을 갖는 행렬

Example 3.1 (Manhattan Norm)
The *Manhattan norm* on $\mathbb{R}^n$ is defined for $\boldsymbol{x} \in \mathbb{R}^n$ as

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

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*.

벡터의 요소에 대한 절댓값의 합. 요소의 값 변화를 정확하게 파악 가능.

예시. X = [1, 2, 3, 4, 5], ||x||1 = (|1|+|2|+|3|+|4|+|5|)=15

사용 영역. L1 정규화, 컴퓨터비전

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*.

n 차원 좌표평면(유클리드 공간)에서의 벡터의 크기를 계산함

예시. 피타고라스의 정리. 2 차원 좌표 평면 상의 최단 거리를 계산함

사용 영역. L2 정규화, kNN 알고리즘, kmean 알고리즘

참고: http://taewan.kim/post/norm/

3.2 Inner Products(내적)

내적은 직관적인 기하학 개념, 벡터의 길이와 두 벡터 간의 거리 등을 소개함

주요 목적은 벡터가 직교하는지 결정하는 것임

3.2.1 Dot Product

Scalar product/dot product 는 이미 익숙함

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

이러한 product 를 dot product 라고 부르겠음

3.2.2 General Inner Products

Scalar 로 더하기 곱하기를 하면서 mapping(매핑, 사상. 함수를 일반화한 것)을 재배열함

Binary mapping 은 두 개의 인수로 매핑하고 각각의 인수에서 선형적임

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

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

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

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.

양의 정부호, 대칭 쌍선형 매핑

3.2. 3 Symmetric, Positive Definite Matrices 대칭 양 정부호 행렬

Symmetric matrix(대칭행렬) : A = A transpose, square matrix 임

Positive definite matrix(정정 행렬) : 성분이 모두 실수이고 대칭인 n*n 정방(square) 행렬 A= A transpose 가 x ! = 0 인 모든 벡터에 대해서 다음 부등식을 만족하면 행렬 A 를 정정 행렬이라고 정의함

X_transpose*A*x 〉 0, ∀x ! =0 ⇔ A〉0

만약 〉=라면 semi-positive

정정 행렬의 예시.

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

벡터를 $x = [x_1 \ x_2]^T$, $x_1 \neq 0$, $x_2 \neq 0$일 때

$$x^T A x = [x_1 \ x_2] \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= [2x_1 + x_2 \quad x_1 + 2x_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 2x_1^2 + x_1x_2 + x_1x_2 + 2x_2^2$$

$$= (x_1 + x_2)^2 + x_1^2 + x_2^2 > 0$$

예시 출처: https://pasus.tistory.com/10

정정 행렬인지 아닌지 여부는 정의대로 판별하기는 힘들고, 행렬식(determinant)나 고유값(eigenvalue)을 구해서 판별하는 방법이 있음

Theorem 3.5. *For a real-valued, finite-dimensional vector space V and an ordered basis B of V, it holds that $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ is an inner product if and only if there exists a symmetric, positive definite matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with*

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \hat{\boldsymbol{x}}^\top \boldsymbol{A} \hat{\boldsymbol{y}} \,. \tag{3.15}$$

The following properties hold if $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is symmetric and positive definite:

- The null space (kernel) of $\boldsymbol{A}$ consists only of $\boldsymbol{0}$ because $\boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x} > 0$ for all $\boldsymbol{x} \neq \boldsymbol{0}$. This implies that $\boldsymbol{A}\boldsymbol{x} \neq \boldsymbol{0}$ if $\boldsymbol{x} \neq \boldsymbol{0}$.
- The diagonal elements a_{ii} of $\boldsymbol{A}$ are positive because $a_{ii} = \boldsymbol{e}_i^\top \boldsymbol{A} \boldsymbol{e}_i > 0$, where $\boldsymbol{e}_i$ is the ith vector of the standard basis in $\mathbb{R}^n$.

3.3 Lengths and Distances

내적은 놈을 유도함

||x|| := sqrt(⟨x, x⟩)

내적으로 벡터의 길이를 계산할 수 있음

하지만 모든 놈이 내적으로부터 유도되는 것은 아님. Ex. Manhattan norm

Remark. 내적 벡터 공간 (V, ⟨ ·,·⟩)에서 유도된 놈 || ·||은 코시 슈바르츠 부등식(Cauchy - Schwarz Inequality)을 만족함

Definition 3.6 (Distance and Metric)**.** Consider an inner product space $(V, \langle\cdot,\cdot\rangle)$. Then

$$d(\boldsymbol{x},\boldsymbol{y}) := \|\boldsymbol{x}-\boldsymbol{y}\| = \sqrt{\langle \boldsymbol{x}-\boldsymbol{y}, \boldsymbol{x}-\boldsymbol{y}\rangle} \tag{3.21}$$

is called the *distance* between $\boldsymbol{x}$ and $\boldsymbol{y}$ for $\boldsymbol{x},\boldsymbol{y} \in V$. If we use the dot product as the inner product, then the distance is called *Euclidean distance*.

The mapping

$$d : V \times V \to \mathbb{R} \tag{3.22}$$

$$(\boldsymbol{x},\boldsymbol{y}) \mapsto d(\boldsymbol{x},\boldsymbol{y}) \tag{3.23}$$

is called a *metric*.

Remark. 벡터의 길이처럼 벡터 간 거리는 내적이 필요 없고 놈으로도 충분함. 거리는 내적의 선택에 따라 달라질 수 있음

Metric d 는 다음을 만족함

Remark. ⟨x y⟩, d(x, y)는 반대 방향

비슷한 x, y -> 내적이 크고 metric 은 작은 값

3.4 Angles and Orthogonality

-내적은 두 벡터 간의 각을 정의하면서 벡터 공간의 기하학도 포착함

-코시 슈바르츠 부등식(Cauchy - Schwarz Inequality)을 사용하여 두 벡터 x, y 사이의 각 w 를 정의함

intuition in $\mathbb{R}^2$ and $\mathbb{R}^3$. Assume that $\boldsymbol{x} \neq \boldsymbol{0}, \boldsymbol{y} \neq \boldsymbol{0}$. Then

$$-1 \leqslant \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\| \, \|\boldsymbol{y}\|} \leqslant 1 \,. \tag{3.24}$$

Therefore, there exists a unique $\omega \in [0, \pi]$, illustrated in Figure 3.4, with

$$\cos \omega = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\| \, \|\boldsymbol{y}\|} \,. \tag{3.25}$$

The number ω is the *angle* between the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. Intuitively, the angle between two vectors tells us how similar their orientations are. For example, using the dot product, the angle between $\boldsymbol{x}$ and $\boldsymbol{y} = 4\boldsymbol{x}$, i.e., $\boldsymbol{y}$ is a scaled version of $\boldsymbol{x}$, is 0: Their orientation is the same.

-내적의 주요 특징 중 하나는 직교하는 벡터를 특징지을 수 있다는 것임

Definition 3.7 (Orthogonality)**.** Two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are *orthogonal* if and only if $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$, and we write $\boldsymbol{x} \perp \boldsymbol{y}$. If additionally $\|\boldsymbol{x}\| = 1 = \|\boldsymbol{y}\|$, i.e., the vectors are unit vectors, then $\boldsymbol{x}$ and $\boldsymbol{y}$ are *orthonormal*.

-암시: 0-벡터는 벡터 공간의 모든 벡터에 직교함

Remark: Orthogonaltiy(직교성)은 perpendicularity(수직)을 dot product 일 필요 없는 두 줄의 형태로 일반화한 것임

기하학적으로 직교하는 벡터를 특정한 내적에 관련하여 90 도를 가지는 것으로 생각할 수 있음

Definition 3.8 (Orthogonal Matrix)**.** A square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is an *orthogonal matrix* if and only if its columns are orthonormal so that

$$\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{I} = \boldsymbol{A}^\top \boldsymbol{A} \,, \tag{3.29}$$

which implies that

$$\boldsymbol{A}^{-1} = \boldsymbol{A}^\top \,, \tag{3.30}$$

i.e., the inverse is obtained by simply transposing the matrix.

Transformations by orthogonal matrices are special because the length of a vector $\boldsymbol{x}$ is not changed when transforming it using an orthogonal matrix $\boldsymbol{A}$. For the dot product, we obtain

$$\|\boldsymbol{A}\boldsymbol{x}\|^2 = (\boldsymbol{A}\boldsymbol{x})^\top (\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^\top \boldsymbol{I}\boldsymbol{x} = \boldsymbol{x}^\top \boldsymbol{x} = \|\boldsymbol{x}\|^2 \,. \tag{3.31}$$

Moreover, the angle between any two vectors $\boldsymbol{x}, \boldsymbol{y}$, as measured by their inner product, is also unchanged when transforming both of them using an orthogonal matrix $\boldsymbol{A}$. Assuming the dot product as the inner product, the angle of the images $\boldsymbol{A}\boldsymbol{x}$ and $\boldsymbol{A}\boldsymbol{y}$ is given as

$$\cos \omega = \frac{(\boldsymbol{A}\boldsymbol{x})^\top (\boldsymbol{A}\boldsymbol{y})}{\|\boldsymbol{A}\boldsymbol{x}\| \, \|\boldsymbol{A}\boldsymbol{y}\|} = \frac{\boldsymbol{x}^\top \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{y}}{\sqrt{\boldsymbol{x}^\top \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{x}\boldsymbol{y}^\top \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{y}}} = \frac{\boldsymbol{x}^\top \boldsymbol{y}}{\|\boldsymbol{x}\| \, \|\boldsymbol{y}\|} \,, \tag{3.32}$$

which gives exactly the angle between $\boldsymbol{x}$ and $\boldsymbol{y}$. This means that orthogonal matrices $\boldsymbol{A}$ with $\boldsymbol{A}^\top = \boldsymbol{A}^{-1}$ preserve both angles and distances. It

3.5 Orthonormal Basis 정규직교 기저

Orthonormal basis: 기저 벡터는 직교하고 각각의 기저 벡터 길이는 1

Definition 3.9 (Orthonormal Basis)**.** Consider an n-dimensional vector space V and a basis $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ 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 $i, j = 1, \ldots, n$ 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.

-Gram Schmidt process(그람 슈미트 과정)

Section 2.6.1 에서 벡터의 세트로 포괄되는(spanned) 벡터 공간을 위한 기저를 찾기 위해 가우스 소거법(Gaussian elimination)을 사용할 수 있음

세트 {b1, …, bn}의 직교하지 않고 정규화되지 않은 기저 벡터가 있다고 하자.

결합하여 하나의 행렬로 만들고, 가우스 소거법을 이용해서 정규직교 기조가 될 수 있음

3.6 Orthogonal Complement 직교 여공간: 주어진 부분공간과 수직인 벡터들의 공간

서로 직교하는 벡터 공간에 대해 살펴봄

Consider a D-dimensional vector space V and an M-dimensional subspace $U \subseteq V$. Then its *orthogonal complement* $U^\perp$ is a $(D-M)$-dimensional subspace of V and contains all vectors in V that are orthogonal to every vector in U. Furthermore, $U \cap U^\perp = \{\boldsymbol{0}\}$ so that any vector $\boldsymbol{x} \in V$ can be

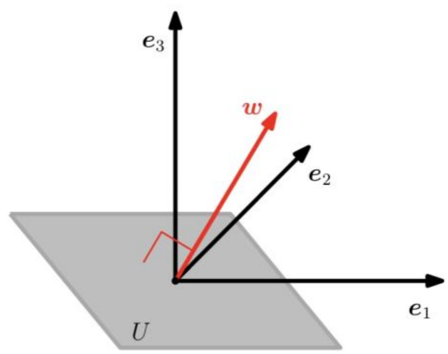
uniquely decomposed into

$$\boldsymbol{x} = \sum_{m=1}^{M} \lambda_m \boldsymbol{b}_m + \sum_{j=1}^{D-M} \psi_j \boldsymbol{b}_j^\perp, \quad \lambda_m,\ \psi_j \in \mathbb{R}, \tag{3.36}$$

where $(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_M)$ is a basis of U and $(\boldsymbol{b}_1^\perp, \ldots, \boldsymbol{b}_{D-M}^\perp)$ is a basis of $U^\perp$.

직교 여공간은 3 차원의 벡터 공간에서 2 차원의 부분공간을 설명하는데 사용될 수 있음

벡터 w with ||w||=1 은 plane(평면) U 에 직교하는데, 이것은 기저 벡터임



w 에 직교하는 모든 벡터는 plane(평면) U 에 놓여있음

w 는 U 의 normal vector(법선)이라고 부름

일반적으로, 직교 여공간은 n-차원의 벡터와 affine space(아핀 공간)에 있는 hyperplane(초평면)을 설명하는 데 사용될 수 있음