3.1 Norm

Vector space V 의 norm 은 V 에서 실수 R 로 가는 함수로서 벡터 x 를 x 의 길이로 대응 시켜준다

$$\begin{aligned} \|\cdot\| : V &\to \mathbb{R}\,, \\ \boldsymbol{x} &\mapsto \|\boldsymbol{x}\|\,, \end{aligned}$$

또한 실수 와 벡터 x,y 에 대해 다음이 성립한다.

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

* 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*.

* 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

3.2.1 Dot product

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

3.2.2 General Inner Products

벡터에 대하여 linear 한 성질이 그대로 유지되는 mapping

$$\begin{aligned} \Omega(\lambda\boldsymbol{x} + \psi\boldsymbol{y}, \boldsymbol{z}) &= \lambda\Omega(\boldsymbol{x}, \boldsymbol{z}) + \psi\Omega(\boldsymbol{y}, \boldsymbol{z}) \\ \Omega(\boldsymbol{x}, \lambda\boldsymbol{y} + \psi\boldsymbol{z}) &= \lambda\Omega(\boldsymbol{x}, \boldsymbol{y}) + \psi\Omega(\boldsymbol{x}, \boldsymbol{z})\,. \end{aligned}$$

3.2.3 Symmetric, Positive Definite Matrices

Symmetric, Positive Definite 행렬은 내적으로 정의되며, ML 에서 중요한 역할

$\boldsymbol{x} = \sum_{i=1}^{n} \psi_i \boldsymbol{b}_i \in V$ 일 때, <x, y>를 다음과 같이 표현가능

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

Inner product <x, y>는 행렬 A 에 의해 유일하게 결정되며, A 라는 행렬 또한 symmetric 하다는 것을 알 수 있음

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

3.3 Lengths and Distances

3.2 에서 정의한 내적으로 다음과 같은 norm 을 정의할 수도 있다

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

하지만 모든 norm 이 내적을 통해 유도되어지는 건 아님

* Remark Cauchy-Schwarz Inequality

$$|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leqslant \|\boldsymbol{x}\|\|\boldsymbol{y}\|$$

좌변 = 벡터의 내적

우변: 벡터의 norm 의 곱

Definition 3.6 Distance and Metric

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

위의 식을 mapping 으로 이해해보면

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

거리는 두 벡터를 하나의 실수값으로 대응시키는 mapping 이라고 할 수 있다. 이를 metric 이라고 한다.

3.4 Angles and orthogonality

Definition 3.7 Orthogonality

두 벡터 x, y 가 <x, y> =0 이면 orthogonal(수직)이라고 하고 $\boldsymbol{x} \perp \boldsymbol{y}$라고 표현한다.

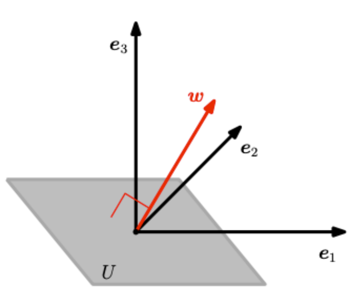
만약 $\|\boldsymbol{x}\| = 1 = \|\boldsymbol{y}\|$이면서 수직이면 두 벡터는 orthonormal(정규직교)라고 한다

3.5 Orthonormal Basis

The canonical/standard basis for a Euclidean vector space $\mathbb{R}^n$ is an orthonormal basis, where the inner product is the dot product of vectors.

$$\boldsymbol{b}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{b}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

3.6 Orthogonal Complement



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