

Mathematics for Machine Learning

— Linear Algebra: Norms, Inner Products & Orthogonality

Joseph Chuang-Chieh Lin

Department of Computer Science & Engineering,
National Taiwan Ocean University

Fall 2025

Credits for the resource

- The slides are based on the textbooks:
 - *Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong: Mathematics for Machine Learning. Cambridge University Press. 2020.*
 - *Howard Anton, Chris Rorres, Anton Kaul: Elementary Linear Algebra. Wiley. 2019.*
- We could partially refer to the monograph:
Francesco Orabona: A Modern Introduction to Online Learning.
<https://arxiv.org/abs/1912.13213>

Outline

- 1 Norms
- 2 Inner Products
- 3 Lengths & Distances
- 4 Angles and Orthogonality
- 5 Orthonormal Basis
- 6 Inner Product of Functions

Outline

- 1 Norms
- 2 Inner Products
- 3 Lengths & Distances
- 4 Angles and Orthogonality
- 5 Orthonormal Basis
- 6 Inner Product of Functions

Norm

Norm

A norm on a vector space V is a function

$$\begin{aligned}\|\cdot\| : V &\rightarrow \mathbb{R} \\ \mathbf{x} &\rightarrow \|\mathbf{x}\|\end{aligned}$$

such that for $\lambda \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in V$ the following conditions hold:

- $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$.
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.
- $\|\mathbf{x}\| \geq 0$ and $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$.

ℓ_1 norm, ℓ_2 norm & ℓ_∞ norm ℓ_1 norm (Manhattan Norm)

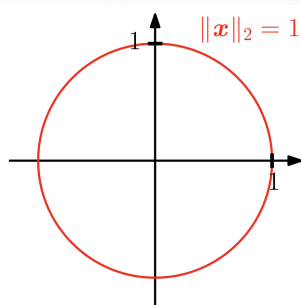
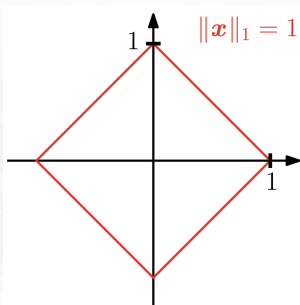
For $\mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{x}\|_1 := \sum_{i=1}^n |x_i|$.

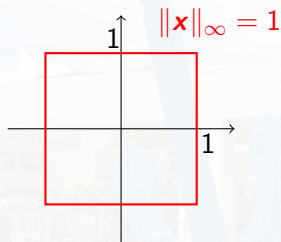
 ℓ_2 norm

For $\mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{x}\|_2 := \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{\mathbf{x}^\top \mathbf{x}}$.

 ℓ_∞ norm

For $\mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{x}\|_\infty := \max_{i \in n} |x_i|$.





Outline

- 1 Norms
- 2 Inner Products**
- 3 Lengths & Distances
- 4 Angles and Orthogonality
- 5 Orthonormal Basis
- 6 Inner Product of Functions

Dot Product

Dot Product

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

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

General Inner Products

Bilinear Mapping f

Given a vector space V . For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $\lambda, \psi \in \mathbb{R}$, such that

$$f(\lambda \mathbf{x} + \psi \mathbf{y}, \mathbf{z}) = \lambda f(\mathbf{x}, \mathbf{z}) + \psi f(\mathbf{y}, \mathbf{z})$$

$$f(\mathbf{x}, \lambda \mathbf{y} + \psi \mathbf{z}) = \lambda f(\mathbf{x}, \mathbf{y}) + \psi f(\mathbf{x}, \mathbf{z})$$

General Inner Products

Bilinear Mapping f

Given a vector space V . For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $\lambda, \psi \in \mathbb{R}$, such that

$$f(\lambda \mathbf{x} + \psi \mathbf{y}, \mathbf{z}) = \lambda f(\mathbf{x}, \mathbf{z}) + \psi f(\mathbf{y}, \mathbf{z}) \quad (\text{linear in the 1st argument})$$

$$f(\mathbf{x}, \lambda \mathbf{y} + \psi \mathbf{z}) = \lambda f(\mathbf{x}, \mathbf{y}) + \psi f(\mathbf{x}, \mathbf{z}) \quad (\text{linear in the 2nd argument})$$

Symmetric & Positive Definite (1/6)

Symmetric

Let V be a vector space and $f : V \times V \rightarrow \mathbb{R}$ be a bilinear mapping. Then f is **symmetric** if $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$.

Positive Definite

Let V be a vector space and $f : V \times V \rightarrow \mathbb{R}$ be a bilinear mapping. Then f is **positive definite** if $\forall \mathbf{x} \in V \setminus \{\mathbf{0}\}$, we have

$$f(\mathbf{x}, \mathbf{x}) > 0 \quad \text{and} \quad f(\mathbf{0}, \mathbf{0}) = 0.$$

Symmetric & Positive Definite (1/6)

Symmetric

Let V be a vector space and $f : V \times V \rightarrow \mathbb{R}$ be a bilinear mapping. Then f is **symmetric** if $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$.

Positive Definite

Let V be a vector space and $f : V \times V \rightarrow \mathbb{R}$ be a bilinear mapping. Then f is **positive definite** if $\forall \mathbf{x} \in V \setminus \{\mathbf{0}\}$, we have

$$f(\mathbf{x}, \mathbf{x}) > 0 \quad \text{and} \quad f(\mathbf{0}, \mathbf{0}) = 0.$$

Inner Product

A **positive definite & symmetric bilinear** mapping $f : V \times V \rightarrow \mathbb{R}$ is called an **inner product** on V and we write $f(\mathbf{x}, \mathbf{y})$ as $\langle \mathbf{x}, \mathbf{y} \rangle$.

Symmetric & Positive Definite (2/6)

- Important in machine learning.
 - Matrix decompositions.
 - Key in defining kernels in the SVM (support vector machine).

An Exercise

Exercise

Consider $V = \mathbb{R}^2$. Define that

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

Show that $\langle \cdot, \cdot \rangle$ is an inner product.

Symmetric & Positive Definite (3/6)

Consider an n -dimensional vector space V with an inner product $\langle \cdot \rangle : V \times V \rightarrow \mathbb{R}$ and an ordered basis $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ of V .

- Assume that for $\mathbf{x}, \mathbf{y} \in V$,
 - $\mathbf{x} = \sum_{i=1}^n \psi_i \mathbf{b}_i$
 - $\mathbf{y} = \sum_{j=1}^n \lambda_j \mathbf{b}_j$for suitable $\psi_i, \lambda_j \in \mathbb{R}$.
- By the bilinearity of the inner product, we have

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

Symmetric & Positive Definite (3/6)

Consider an n -dimensional vector space V with an inner product $\langle \cdot \rangle : V \times V \rightarrow \mathbb{R}$ and an ordered basis $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ of V .

- Assume that for $\mathbf{x}, \mathbf{y} \in V$,

- $\mathbf{x} = \sum_{i=1}^n \psi_i \mathbf{b}_i$

- $\mathbf{y} = \sum_{j=1}^n \lambda_j \mathbf{b}_j$

for suitable $\psi_i, \lambda_j \in \mathbb{R}$.

- By the bilinearity of the inner product, we have

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

where $A_{ij} := \langle \mathbf{b}_i, \mathbf{b}_j \rangle$, $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are the coordinates of \mathbf{x} and \mathbf{y} w.r.t. the basis B .

Symmetric & Positive Definite (3/6)

Consider an n -dimensional vector space V with an inner product $\langle \cdot \rangle : V \times V \rightarrow \mathbb{R}$ and an ordered basis $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ of V .

- Assume that for $\mathbf{x}, \mathbf{y} \in V$,
 - $\mathbf{x} = \sum_{i=1}^n \psi_i \mathbf{b}_i$
 - $\mathbf{y} = \sum_{j=1}^n \lambda_j \mathbf{b}_j$for suitable $\psi_i, \lambda_j \in \mathbb{R}$.
- By the bilinearity of the inner product, we have

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

where $A_{ij} := \langle \mathbf{b}_i, \mathbf{b}_j \rangle$, $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are the coordinates of \mathbf{x} and \mathbf{y} w.r.t. the basis B .

- ★ Note that the symmetry of the inner product implies that \mathbf{A} is symmetric.

Example

Consider $V = \mathbb{R}^2$ with an inner product $\langle \cdot \rangle : V \times V \rightarrow \mathbb{R}$ and an ordered basis $B = (\mathbf{q}_1, \mathbf{q}_2)$ of V , where $\mathbf{q}_1 = [1, 1]^\top$, $\mathbf{q}_2 = [1, -2]^\top$.

Compute $\langle \mathbf{x}, \mathbf{y} \rangle$, where

$$\mathbf{x} = 2\mathbf{q}_1 + 3\mathbf{q}_2$$

$$\mathbf{y} = -\mathbf{q}_1 + 2\mathbf{q}_2$$

- $\langle \mathbf{x}, \mathbf{y} \rangle =$

Example

Consider $V = \mathbb{R}^2$ with an inner product $\langle \cdot \rangle : V \times V \rightarrow \mathbb{R}$ and an ordered basis $B = (\mathbf{q}_1, \mathbf{q}_2)$ of V , where $\mathbf{q}_1 = [1, 1]^\top$, $\mathbf{q}_2 = [1, -2]^\top$.

Compute $\langle \mathbf{x}, \mathbf{y} \rangle$, where

$$\mathbf{x} = 2\mathbf{q}_1 + 3\mathbf{q}_2$$

$$\mathbf{y} = -\mathbf{q}_1 + 2\mathbf{q}_2$$

- $\langle \mathbf{x}, \mathbf{y} \rangle = -2\langle \mathbf{q}_1, \mathbf{q}_1 \rangle - 3\langle \mathbf{q}_2, \mathbf{q}_1 \rangle + 4\langle \mathbf{q}_1, \mathbf{q}_2 \rangle + 6\langle \mathbf{q}_2, \mathbf{q}_2 \rangle$

Example

Consider $V = \mathbb{R}^2$ with an inner product $\langle \cdot \rangle : V \times V \rightarrow \mathbb{R}$ and an ordered basis $B = (\mathbf{q}_1, \mathbf{q}_2)$ of V , where $\mathbf{q}_1 = [1, 1]^\top$, $\mathbf{q}_2 = [1, -2]^\top$.

Compute $\langle \mathbf{x}, \mathbf{y} \rangle$, where

$$\mathbf{x} = 2\mathbf{q}_1 + 3\mathbf{q}_2$$

$$\mathbf{y} = -\mathbf{q}_1 + 2\mathbf{q}_2$$

- $\langle \mathbf{x}, \mathbf{y} \rangle = -2\langle \mathbf{q}_1, \mathbf{q}_1 \rangle - 3\langle \mathbf{q}_2, \mathbf{q}_1 \rangle + 4\langle \mathbf{q}_1, \mathbf{q}_2 \rangle + 6\langle \mathbf{q}_2, \mathbf{q}_2 \rangle = 25.$
- W.r.t. the standard basis,

Example

Consider $V = \mathbb{R}^2$ with an inner product $\langle \cdot \rangle : V \times V \rightarrow \mathbb{R}$ and an ordered basis $B = (\mathbf{q}_1, \mathbf{q}_2)$ of V , where $\mathbf{q}_1 = [1, 1]^\top$, $\mathbf{q}_2 = [1, -2]^\top$.

Compute $\langle \mathbf{x}, \mathbf{y} \rangle$, where

$$\mathbf{x} = 2\mathbf{q}_1 + 3\mathbf{q}_2$$

$$\mathbf{y} = -\mathbf{q}_1 + 2\mathbf{q}_2$$

- $\langle \mathbf{x}, \mathbf{y} \rangle = -2\langle \mathbf{q}_1, \mathbf{q}_1 \rangle - 3\langle \mathbf{q}_2, \mathbf{q}_1 \rangle + 4\langle \mathbf{q}_1, \mathbf{q}_2 \rangle + 6\langle \mathbf{q}_2, \mathbf{q}_2 \rangle = 25.$
- W.r.t. the standard basis,

$$\mathbf{x} = 5\mathbf{e}_1 - 4\mathbf{e}_2$$

$$\mathbf{y} = \mathbf{e}_1 - 5\mathbf{e}_2$$

Example

Consider $V = \mathbb{R}^2$ with an inner product $\langle \cdot \rangle : V \times V \rightarrow \mathbb{R}$ and an ordered basis $B = (\mathbf{q}_1, \mathbf{q}_2)$ of V , where $\mathbf{q}_1 = [1, 1]^\top$, $\mathbf{q}_2 = [1, -2]^\top$.

Compute $\langle \mathbf{x}, \mathbf{y} \rangle$, where

$$\mathbf{x} = 2\mathbf{q}_1 + 3\mathbf{q}_2$$

$$\mathbf{y} = -\mathbf{q}_1 + 2\mathbf{q}_2$$

- $\langle \mathbf{x}, \mathbf{y} \rangle = -2\langle \mathbf{q}_1, \mathbf{q}_1 \rangle - 3\langle \mathbf{q}_2, \mathbf{q}_1 \rangle + 4\langle \mathbf{q}_1, \mathbf{q}_2 \rangle + 6\langle \mathbf{q}_2, \mathbf{q}_2 \rangle = 25.$
- W.r.t. the standard basis,

$$\mathbf{x} = 5\mathbf{e}_1 - 4\mathbf{e}_2 \implies \hat{\mathbf{x}} = [5, -4]^\top$$

$$\mathbf{y} = \mathbf{e}_1 - 5\mathbf{e}_2 \implies \hat{\mathbf{y}} = [1, -5]^\top$$

Example

Consider $V = \mathbb{R}^2$ with an inner product $\langle \cdot \rangle : V \times V \rightarrow \mathbb{R}$ and an ordered basis $B = (\mathbf{q}_1, \mathbf{q}_2)$ of V , where $\mathbf{q}_1 = [1, 1]^\top$, $\mathbf{q}_2 = [1, -2]^\top$.

Compute $\langle \mathbf{x}, \mathbf{y} \rangle$, where

$$\mathbf{x} = 2\mathbf{q}_1 + 3\mathbf{q}_2$$

$$\mathbf{y} = -\mathbf{q}_1 + 2\mathbf{q}_2$$

- $\langle \mathbf{x}, \mathbf{y} \rangle = -2\langle \mathbf{q}_1, \mathbf{q}_1 \rangle - 3\langle \mathbf{q}_2, \mathbf{q}_1 \rangle + 4\langle \mathbf{q}_1, \mathbf{q}_2 \rangle + 6\langle \mathbf{q}_2, \mathbf{q}_2 \rangle = 25.$
- W.r.t. the standard basis,

$$\mathbf{x} = 5\mathbf{e}_1 - 4\mathbf{e}_2 \implies \hat{\mathbf{x}} = [5, -4]^\top$$

$$\mathbf{y} = \mathbf{e}_1 - 5\mathbf{e}_2 \implies \hat{\mathbf{y}} = [1, -5]^\top$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Example

Consider $V = \mathbb{R}^2$ with an inner product $\langle \cdot \rangle : V \times V \rightarrow \mathbb{R}$ and an ordered basis $B = (\mathbf{q}_1, \mathbf{q}_2)$ of V , where $\mathbf{q}_1 = [1, 1]^\top$, $\mathbf{q}_2 = [1, -2]^\top$.

Compute $\langle \mathbf{x}, \mathbf{y} \rangle$, where

$$\mathbf{x} = 2\mathbf{q}_1 + 3\mathbf{q}_2$$

$$\mathbf{y} = -\mathbf{q}_1 + 2\mathbf{q}_2$$

- $\langle \mathbf{x}, \mathbf{y} \rangle = -2\langle \mathbf{q}_1, \mathbf{q}_1 \rangle - 3\langle \mathbf{q}_2, \mathbf{q}_1 \rangle + 4\langle \mathbf{q}_1, \mathbf{q}_2 \rangle + 6\langle \mathbf{q}_2, \mathbf{q}_2 \rangle = 25.$
- W.r.t. the standard basis,

$$\mathbf{x} = 5\mathbf{e}_1 - 4\mathbf{e}_2 \implies \hat{\mathbf{x}} = [5, -4]^\top$$

$$\mathbf{y} = \mathbf{e}_1 - 5\mathbf{e}_2 \implies \hat{\mathbf{y}} = [1, -5]^\top$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \implies \hat{\mathbf{x}}^\top \mathbf{A} \hat{\mathbf{y}} = [5, -4] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -5 \end{bmatrix} = 25.$$

Symmetric & Positive Definite (4/6)

The positive definiteness of the inner product implies that

$$\forall \mathbf{x} \in V \setminus \{\mathbf{0}\} : \mathbf{x}^\top \mathbf{A} \mathbf{x} > 0.$$

Symmetric, Positive Definite Matrix

A symmetric matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ that satisfies the property:

$$\forall \mathbf{x} \in V \setminus \{\mathbf{0}\} : \mathbf{x}^\top \mathbf{A} \mathbf{x} > 0.$$

is called **symmetric, positive definite** (or just **positive definite**).

If only \geq holds, then \mathbf{A} is called symmetric, **positive semidefinite**.

Example

Consider the matrices $\mathbf{A}_1 = \begin{bmatrix} 9 & 6 \\ 6 & 5 \end{bmatrix}$, $\mathbf{A}_2 = \begin{bmatrix} 9 & 6 \\ 6 & 3 \end{bmatrix}$

- \mathbf{A}_1 is positive definite (why?)

Example

Consider the matrices $\mathbf{A}_1 = \begin{bmatrix} 9 & 6 \\ 6 & 5 \end{bmatrix}$, $\mathbf{A}_2 = \begin{bmatrix} 9 & 6 \\ 6 & 3 \end{bmatrix}$

- \mathbf{A}_1 is positive definite (why?)
- \mathbf{A}_2 is NOT positive definite (why?)

Symmetric & Positive Definite (5/6)

If $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric, positive definite, then

$$\langle \mathbf{x}, \mathbf{y} \rangle = \hat{\mathbf{x}}^\top \mathbf{A} \hat{\mathbf{y}}.$$

Symmetric & Positive Definite (5/6)

If $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric, positive definite, then

$$\langle \mathbf{x}, \mathbf{y} \rangle = \hat{\mathbf{x}}^\top \mathbf{A} \hat{\mathbf{y}}.$$

This defines an inner product w.r.t. an ordered basis B , where $\hat{\mathbf{x}}, \hat{\mathbf{y}}$ are the coordinates of \mathbf{x}, \mathbf{y} w.r.t. B .

Remark

Semidefinite Matrix

If $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric and for all \mathbf{x} we have $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$, we call \mathbf{A} a semidefinite matrix.

Remark: If $\mathbf{A} \in \mathbb{R}^{n \times n}$ is not necessarily symmetric & positive definite:

Remark

Semidefinite Matrix

If $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric and for all \mathbf{x} we have $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$, we call \mathbf{A} a semidefinite matrix.

Remark: If $\mathbf{A} \in \mathbb{R}^{n \times n}$ is not necessarily symmetric & positive definite:

- Try $\hat{\mathbf{A}} := \mathbf{A} \mathbf{A}^\top$.

Remark

Semidefinite Matrix

If $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric and for all \mathbf{x} we have $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$, we call \mathbf{A} a semidefinite matrix.

Remark: If $\mathbf{A} \in \mathbb{R}^{n \times n}$ is not necessarily symmetric & positive definite:

- Try $\hat{\mathbf{A}} := \mathbf{A} \mathbf{A}^\top$.
- $\hat{\mathbf{A}}$ must be semidefinite

Remark

Semidefinite Matrix

If $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric and for all \mathbf{x} we have $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$, we call \mathbf{A} a semidefinite matrix.

Remark: If $\mathbf{A} \in \mathbb{R}^{n \times n}$ is not necessarily symmetric & positive definite:

- Try $\hat{\mathbf{A}} := \mathbf{A} \mathbf{A}^\top$.
- $\hat{\mathbf{A}}$ must be semidefinite (why?).

Symmetric & Positive Definite (6/6)

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

- $\text{null}(\mathbf{A}) = \{\mathbf{0}\}$.

Symmetric & Positive Definite (6/6)

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

- $\text{null}(\mathbf{A}) = \{\mathbf{0}\}$.
 - Since $\mathbf{x}^\top \mathbf{A} \mathbf{x} > 0$ for all $\mathbf{x} > 0 \Rightarrow \mathbf{A} \mathbf{x} \neq \mathbf{0}$ if $\mathbf{x} \neq \mathbf{0}$.
- For the diagonal elements a_{ii} of \mathbf{A} , $a_{ii} = \mathbf{e}_i^\top \mathbf{A} \mathbf{e}_i > 0$.
 - \mathbf{e}_i : the i th vector of the standard basis of \mathbb{R}^n .

Outline

- 1 Norms
- 2 Inner Products
- 3 Lengths & Distances**
- 4 Angles and Orthogonality
- 5 Orthonormal Basis
- 6 Inner Product of Functions

Remark

- Note that **any inner product induces a norm**:

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

Cauchy-Schwarz Inequality

For an inner product vector space $(V, \langle \cdot \rangle)$, the induced norm $\|\cdot\|$ satisfies the Cauchy-Schwarz inequality

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|.$$

Lengths of Vectors

Example

Compute the length of a vector $\mathbf{x} = [1, 1]^\top \in \mathbb{R}^2$ using

- Dot product

- $\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^\top \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \mathbf{y} = x_1 y_1 - \frac{1}{2}(x_1 y_2 + x_2 y_1) + x_2 y_2.$

Distance & Metric

Distance

Consider an inner product space $(V, \langle \cdot \rangle)$. Then, the **distance** between \mathbf{x} and \mathbf{y} for $\mathbf{x}, \mathbf{y} \in V$ is

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

- The mapping $d : V \times V \rightarrow \mathbb{R}$ for which (\mathbf{x}, \mathbf{y}) maps to $d(\mathbf{x}, \mathbf{y})$ is called a **metric**

Distance & Metric

Distance

Consider an inner product space $(V, \langle \cdot \rangle)$. Then, the **distance** between \mathbf{x} and \mathbf{y} for $\mathbf{x}, \mathbf{y} \in V$ is

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

- The mapping $d : V \times V \rightarrow \mathbb{R}$ for which (\mathbf{x}, \mathbf{y}) maps to $d(\mathbf{x}, \mathbf{y})$ is called a **metric**, which satisfies:

Distance & Metric

Distance

Consider an inner product space $(V, \langle \cdot \rangle)$. Then, the **distance** between \mathbf{x} and \mathbf{y} for $\mathbf{x}, \mathbf{y} \in V$ is

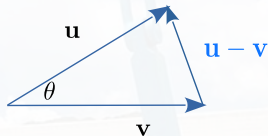
$$d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\| = \sqrt{\langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle}.$$

- The mapping $d : V \times V \rightarrow \mathbb{R}$ for which (\mathbf{x}, \mathbf{y}) maps to $d(\mathbf{x}, \mathbf{y})$ is called a **metric**, which satisfies:
 - *positive definite*: $d(\mathbf{x}, \mathbf{y}) \geq 0$ for all $\mathbf{x}, \mathbf{y} \in V$ and $d(\mathbf{x}, \mathbf{y}) = 0$ iff $\mathbf{x} = \mathbf{y}$.
 - *symmetric*: $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in V$.
 - *triangular inequality*: $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$.

Outline

- 1 Norms
- 2 Inner Products
- 3 Lengths & Distances
- 4 Angles and Orthogonality**
- 5 Orthonormal Basis
- 6 Inner Product of Functions

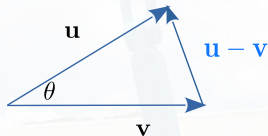
Recall from Senior High School Math



Law of Cosines

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\| \|\mathbf{v}\| \cos \theta$$

Recall from Senior High School Math



Law of Cosines

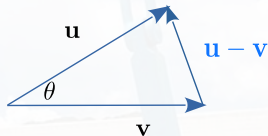
$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\| \|\mathbf{v}\| \cos \theta$$

Note:

$$\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\langle \mathbf{u}, \mathbf{v} \rangle.$$

Thus,

Recall from Senior High School Math



Law of Cosines

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\| \|\mathbf{v}\| \cos \theta$$

Note:

$$\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\langle \mathbf{u}, \mathbf{v} \rangle.$$

Thus,

$$\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\| \cdot \|\mathbf{v}\| \cos \theta.$$

Angles

Assume that $\mathbf{x} \neq \mathbf{0}, \mathbf{y} \neq \mathbf{0}$. Then by the Cauchy-Schwarz inequality,

$$-1 \leq \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|} \leq 1.$$

Angles

Assume that $\mathbf{x} \neq \mathbf{0}, \mathbf{y} \neq \mathbf{0}$. Then by the Cauchy-Schwarz inequality,

$$-1 \leq \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|} \leq 1.$$

Thus, there exists a unique $\theta \in [0, \pi]$, such that

$$\cos(\theta) = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}.$$

Angles

Assume that $\mathbf{x} \neq \mathbf{0}, \mathbf{y} \neq \mathbf{0}$. Then by the Cauchy-Schwarz inequality,

$$-1 \leq \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|} \leq 1.$$

Thus, there exists a unique $\theta \in [0, \pi]$, such that

$$\cos(\theta) = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}.$$

We call θ the **angle** between \mathbf{x} and \mathbf{y} .

Orthogonality

Orthogonality

- Two vectors \mathbf{x} and \mathbf{y} are **orthogonal** if and only if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.
 - We write $\mathbf{x} \perp \mathbf{y}$.
- If \mathbf{x} and \mathbf{y} are orthogonal and $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$, then \mathbf{x} and \mathbf{y} are both **orthonormal**.

Orthogonal Matrix

Orthogonal Matrix

A square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is an orthogonal matrix iff its columns are orthonormal so that

$$\mathbf{A}\mathbf{A}^\top = \mathbf{I} = \mathbf{A}^\top \mathbf{A},$$

which implies

$$\mathbf{A}^{-1} = \mathbf{A}^\top.$$

Remark

Transformations by orthogonal matrices do NOT change the length of a vector.

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

Orthogonal Matrix

Orthogonal Matrix

A square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is an orthogonal matrix iff its columns are orthonormal so that

$$\mathbf{A}\mathbf{A}^\top = \mathbf{I} = \mathbf{A}^\top \mathbf{A},$$

which implies

$$\mathbf{A}^{-1} = \mathbf{A}^\top.$$

Remark

Transformations by orthogonal matrices do NOT change the length of a vector.

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

Orthogonal Matrix

Orthogonal Matrix

A square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is an orthogonal matrix iff its columns are orthonormal so that

$$\mathbf{A}\mathbf{A}^\top = \mathbf{I} = \mathbf{A}^\top \mathbf{A},$$

which implies

$$\mathbf{A}^{-1} = \mathbf{A}^\top.$$

Remark

Transformations by orthogonal matrices do NOT change the length of a vector.

$$\|\mathbf{Ax}\|^2 = (\mathbf{Ax})^\top (\mathbf{Ax}) = \mathbf{x}^\top \mathbf{A}^\top \mathbf{Ax} = \mathbf{x}^\top \mathbf{I} \mathbf{x} = \mathbf{x}^\top \mathbf{x} = \|\mathbf{x}\|^2.$$

Let θ be the angle between \mathbf{Ax} and \mathbf{Ay} , what is $\cos \theta$?

Outline

- 1 Norms
- 2 Inner Products
- 3 Lengths & Distances
- 4 Angles and Orthogonality
- 5 Orthonormal Basis**
- 6 Inner Product of Functions

Orthonormal Basis

Orthonormal Basis

Consider an n -dimensional vector space V and a basis $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of V .
If for all $i, j = 1, \dots, n$

$$\langle \mathbf{b}_i, \mathbf{b}_j \rangle = 0 \quad \text{for } i \neq j \quad (1)$$

$$\langle \mathbf{b}_i, \mathbf{b}_i \rangle = 1, \quad (2)$$

then the basis is called an **orthonormal basis**.

- Only (1) is satisfied \Rightarrow orthogonal basis.

Example

- The standard basis for \mathbb{R}^n .
- $\mathbf{b}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$

Outline

- 1 Norms
- 2 Inner Products
- 3 Lengths & Distances
- 4 Angles and Orthogonality
- 5 Orthonormal Basis
- 6 Inner Product of Functions**

Inner Product of Functions

Inner Product of Functions

Given two functions $u, v : \mathbb{R} \rightarrow \mathbb{R}$, the inner product of u and v can be defined as

$$\langle u, v \rangle := \int_a^b u(x)v(x)dx$$

for lower and upper limits $a, b < \infty$.

Example

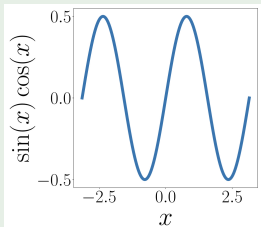
Example (Exercise)

- Choose $u(x) = \sin(x)$ and $v(x) = \cos(x)$.
- Define $f(x) = u(x)v(x)$.

Example

Example (Exercise)

- Choose $u(x) = \sin(x)$ and $v(x) = \cos(x)$.
- Define $f(x) = u(x)v(x)$.

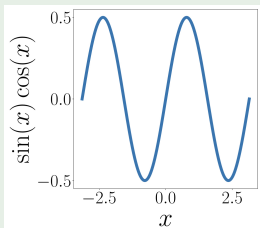


- We can observe that $f(-x) = -f(x)$
- $\int_{-\pi}^{\pi} u(x)v(x)dx = 0$.
- ★ **Note:** $\int \sin(x)\cos(x)dx =$

Example

Example (Exercise)

- Choose $u(x) = \sin(x)$ and $v(x) = \cos(x)$.
- Define $f(x) = u(x)v(x)$.

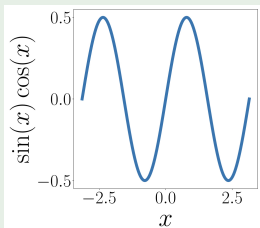


- We can observe that $f(-x) = -f(x)$
- $\int_{-\pi}^{\pi} u(x)v(x)dx = 0$.
- ★ **Note:** $\int \sin(x)\cos(x)dx = \int udu =$

Example

Example (Exercise)

- Choose $u(x) = \sin(x)$ and $v(x) = \cos(x)$.
- Define $f(x) = u(x)v(x)$.



- We can observe that $f(-x) = -f(x)$
- $\int_{-\pi}^{\pi} u(x)v(x)dx = 0$.
- ★ **Note:** $\int \sin(x)\cos(x)dx = \int udu = \frac{1}{2}u^2$, where $u = \sin(x)$.

Discussions