Lecture 02

Vector II: Norm and Inner Product

Instructor: Ruoyu Sun
Cosme Louart

香港中文大學(深圳) | 数据科学学院
The Chinese University of Hong Kong, Shenzhen | School of Data Science

Recall

In the last lecture …

- Definition of **(column) vectors**
- Basic vector operations (**addition, multiplication, linear combination**)

Linear Algebra Terminology and Conventions

- A **vector** is an ordered list of numbers (often a column) written as

$$\begin{bmatrix} -1.1 \\ 0.0 \\ 3.6 \\ -7.2 \end{bmatrix} \quad \text{or} \quad (-1.1, 0, 3.6, -7.2)$$

- Numbers in the list/array are the elements (entries, coefficients, components)
- Number of elements is the size (dimension, length) of the vector (n-length) (n-dimensional)

  [*Warning*]: In some textbooks, like Gilbert Strang, length refers to something else

- The vector above has dimension 4; its third entry is 3.6
- A vector of size n is also called an n-vector
- Numbers are called scalars (compared to vectors)

Linear Algebra Terminology and Conventions

- Vectors are often represented by mathematical symbols, e.g.,

$$x, y, u, v$$

  It could be either in boldface or not, depending on contexts

  Other conventions for beginners and engineers: $\mathbf{u}, \mathbf{v}$ or $\vec{\mathbf{u}}, \vec{\mathbf{v}}$

- The i-th element of an n-vector $\mathbf{u}$ is denoted as u_i
- In u_i, i is the index
- For an n-vector, indexes run from $i = 1$ to $i = n$

  [*Warning*]: sometime u_i may be used to denote the i-th vector in a list of vectors

- Two vectors $\mathbf{u}, \mathbf{v}$ are equal iff $\mathbf{u} - \mathbf{v} = 0$, written as $\mathbf{u} = \mathbf{v}$

Linear Algebra Terminology and Conventions

- Zero vector: All elements are zeros, e.g., $(0,0,0,0)$ denoted by 0, $\mathbf{0}$, or $\mathbf{0}_n$
- One vector: All elements are ones, e.g., $(1,1,1,1)$ denoted by 1, $\mathbf{1}$, or $\mathbf{1}_n$
- Non-zero vector: Not a zero vector
- Non-one vector? Rarely seen!

Learning Goals Today

Today: More vector operations (1) Vector Norm; (2) Inner Product.

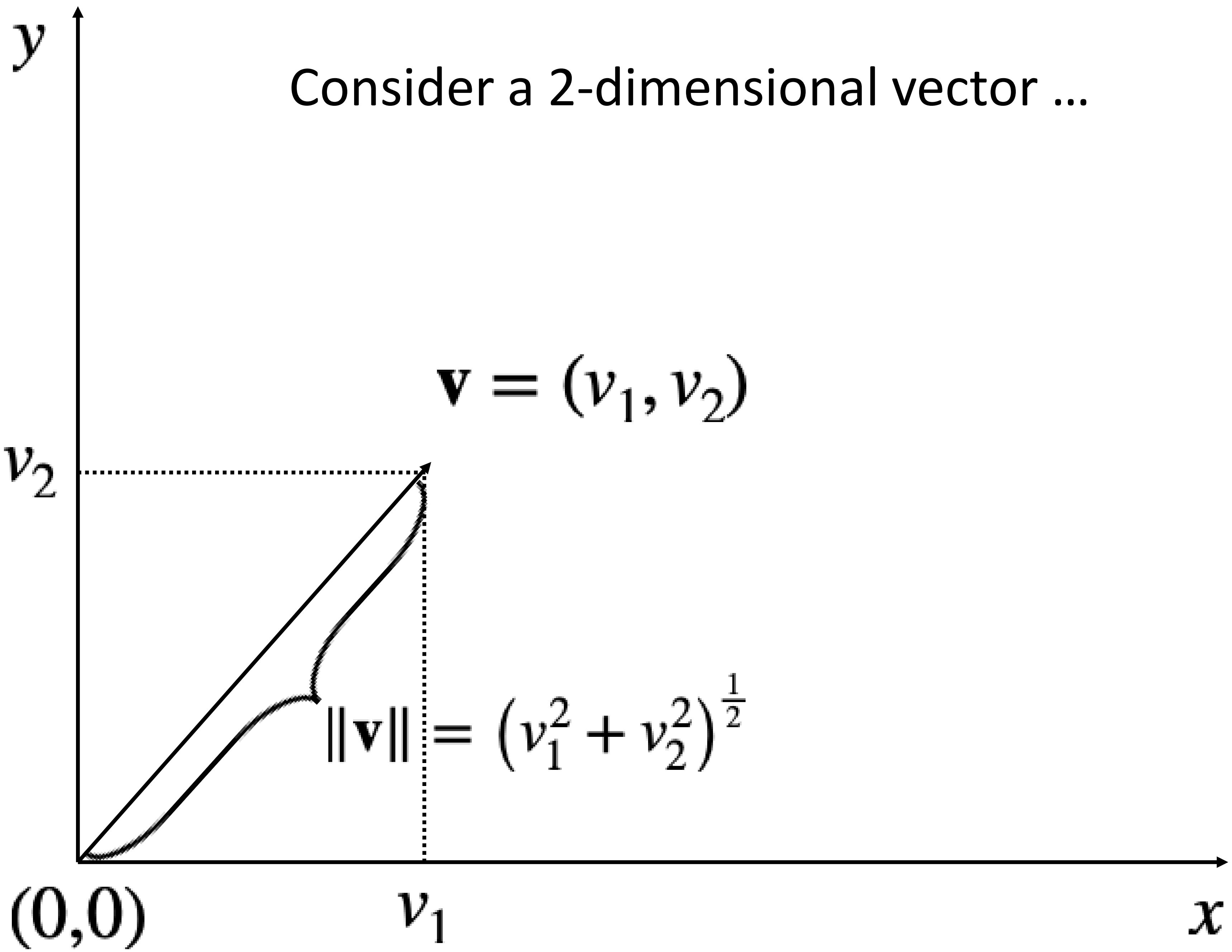
After this lecture, you should be able to:

- calculate the **norm** of a vector
- calculate the **inner product** of a vector
- utilize the **Cauchy Schwarz inequality** and **triangular inequality**
- provide real-world examples of "**inner product**”

Part I Vector Norm

Vector Norm ("Geometric Length")

Consider a 2-dimensional vector …



Vector Norm

A generalized notion of "*absolute value*" ...

Definition (ℓ_2 -norm)

Let $\mathbf{v} = (v_1, \ldots, v_n)$ be a n-length vector. The ℓ_2**-norm** of $\mathbf{v}$, denoted by $\|\mathbf{v}\|_2$ is defined as

$$\|\mathbf{v}\|_2 := \left(v_1^2 + \cdots + v_n^2\right)^{\frac{1}{2}}$$

The ℓ_2-norm is also called an *Euclidean norm* or geometric length

We often abbreviate $\|\cdot\|_2$ as $\|\cdot\|$

Properties of Vector Norms

Properties of Vector Norms

P1 $\|\mathbf{v}\| \geq 0$ $\|\mathbf{v}\| = 0$ iff $\mathbf{v} = 0$

P2 $\|c\mathbf{v}\| = \,?$

P3 $\|\mathbf{u} + \mathbf{v}\| \quad ? \quad \|\mathbf{u}\| + \|\mathbf{v}\|$

Unit Vector

Definition (Unit Vector)

A vector $\mathbf{v}$ is called a **unit vector** if $\|\mathbf{v}\| = 1$

For any non-zero vector $\mathbf{v}$, $\dfrac{\mathbf{v}}{\|\mathbf{v}\|}$ is a unit vector

Unit Vector: Examples

Examples (Unit Vector)

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Standard Unit Vectors in a Cartesian coordinate system

[*Warning*]: In some textbooks, like Stephen Boyd, only above vectors are called unit vectors

Part II Inner Product

Vector Operations

Question1: Can you think about any other vector operations?

Vector Addition

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{bmatrix}$$

Vector Operations

Question1: Can you think about any other vector operations?

Vector Addition

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{bmatrix}$$

How about

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \times \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 w_1 \\ v_2 w_2 \\ \vdots \\ v_n w_n \end{bmatrix}$$

?

This is a “point-wise product”.

Dot Product or inner product

Definition (Dot Product)

A *dot product* between two vectors (of the same size)

$$\mathbf{v} = (v_1, \ldots, v_n)$$

$$\mathbf{w} = (w_1, \ldots, w_n)$$

is defined as

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{n} v_i w_i$$

Example: Calculate the dot product of (-1,2,2) and (1,0,-3)

Dot Product or Inner Product

Sometime we also write a dot product as $\mathbf{v} \cdot \mathbf{w}$ or $\mathbf{v}^\top \mathbf{w}$

Recommend writing as $\langle \mathbf{v}, \mathbf{w} \rangle$ to avoid confusion.

A *dot product* is called an *inner product* in more general settings

Properties of Inner Products

Properties of Inner Products

P1 Linearity $\langle a\mathbf{v} + b\mathbf{u}, \mathbf{w} \rangle = a\langle \mathbf{v}, \mathbf{w} \rangle + b\langle \mathbf{u}, \mathbf{w} \rangle$ for all $\mathbf{u}$ $\mathbf{v}$ $\mathbf{w}$

Properties of Inner Products

Properties of Inner Products

P1 Linearity $\langle a\mathbf{v} + b\mathbf{u}, \mathbf{w}\rangle = a\langle \mathbf{v}, \mathbf{w}\rangle + b\langle \mathbf{u}, \mathbf{w}\rangle$ for all $\mathbf{u}\ \mathbf{v}\ \mathbf{w}$

P2 Symmetry $\langle \mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{w}, \mathbf{v}\rangle$ for all $\mathbf{v}\ \mathbf{w}$

Properties of Inner Products

Properties of Inner Products

P1 Linearity $\langle a\mathbf{v} + b\mathbf{u}, \mathbf{w}\rangle = a\langle \mathbf{v}, \mathbf{w}\rangle + b\langle \mathbf{u}, \mathbf{w}\rangle$ for all $\mathbf{u}\ \mathbf{v}\ \mathbf{w}$

P2 Symmetry $\langle \mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{w}, \mathbf{v}\rangle$ for all $\mathbf{v}\ \mathbf{w}$

P3 Positivity $\langle \mathbf{v}, \mathbf{v}\rangle \geq 0$ for all $\mathbf{v}$

$\langle \mathbf{v}, \mathbf{v}\rangle = 0$ iff $\mathbf{v} = 0$

Properties of Inner Products

Properties of Inner Products

P1 Linearity $\langle a\mathbf{v} + b\mathbf{u}, \mathbf{w}\rangle = a\langle \mathbf{v}, \mathbf{w}\rangle + b\langle \mathbf{u}, \mathbf{w}\rangle$ for all $\mathbf{u}$ $\mathbf{v}$ $\mathbf{w}$

P2 Symmetry $\langle \mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{w}, \mathbf{v}\rangle$ for all $\mathbf{v}$ $\mathbf{w}$

P3 Positivity $\langle \mathbf{v}, \mathbf{v}\rangle \geq 0$ for all $\mathbf{v}$

$\langle \mathbf{v}, \mathbf{v}\rangle = 0$ iff $\mathbf{v} = 0$

verify $\langle \mathbf{v}, \mathbf{v}\rangle = \|v\|^2$!

Reading: Inner product naturally induces a norm and every "inner product space" is a "normed vector space"

Application 1: Weight, Feature and Score

Example 1.1: My MAT2041 final score

	Assignment	Midterm	Final	Total
Weight	0.35	0.3	0.35	
Score	90	85	95	
Cost	31.5	25.5	33.25	90.25

"weight vector"

"Feature vector"

"Score"

$\mathbf{w}$ = (0.35, 0.3, 0.35): weight vector

$\mathbf{v}$ = (90, 85, 95): feature vector

$\langle \mathbf{v}, \mathbf{w} \rangle$ = 90.25 =: score

Application 1: Weight, Feature and Score

Example 1.2: My movie preference

Movie 1: Moon Man

	Action Film	Hollywood	Comedy	Total
Weight	10%	10%	80%	
Value	0.5	0.5	10	
Score	0.05	0.05	8	8.1

"weight"

"Feature"

User 1

"Score"

Application 1: Weight, Feature and Score

Example 1.2: My movie preference

Movie 2: The Matrix

	Action Film	Hollywood	Comedy	Total
Weight	70%	20%	10%	
Value	10	8	2	
Score	7	1.6	0.2	8.8

"weight"

"Feature"

User 2

"Score"

Application 1: Weight, Feature and Score

Application 1: (general)

$\mathbf{w}$ is a vector of the same size (often called a weight vector),

$\mathbf{v}$ represents a set of "features" of an object,

Score: inner product $\langle \mathbf{v}, \mathbf{w} \rangle$ is a weighted sum of the feature values.

Examples:

	Grading	**Movie**	**Your example?**
"Weight"	Grading Scheme	Personal Preference	?
"Feature"	Scores	Movie Features	?
"Score"	Total Score	**Score**	?

Application 1: Weight, Feature and Score

Meta-application: evaluation.
(Useful for ranking, comparing, etc.)

Goal: To evaluate a city, a university, an employee, a basketball player, etc.

Step 1: Set up "features" (indicators)

Step 2: Provide a score for each feature, obtaining a Feature vector

Step 3: Provide weights to the features; get weight vector

Step 4: Compute the inner product, to get the "score"

Remark: "Score" can be used in other areas, e.g. machine learning

Dot Product is a Primary-School Topic?

You may ask: we have known this in preliminary school, why learning it again?

Reason 1: High dimension

Because of the high dimension of data, need computer to process.

Need more formal treatment and techniques.

What if there are 1000,000 users and 10,000 movies?

How can you sort and recommend movies to users?

Dot Product is a Primary-School Topic?

You may ask: we have known this in preliminary school, why learning it again?

Reason 1: High dimension

Because of the high dimension of data, need computer to process.

Need more formal treatment and techniques.

Reason 2: Abstract mathematical representation

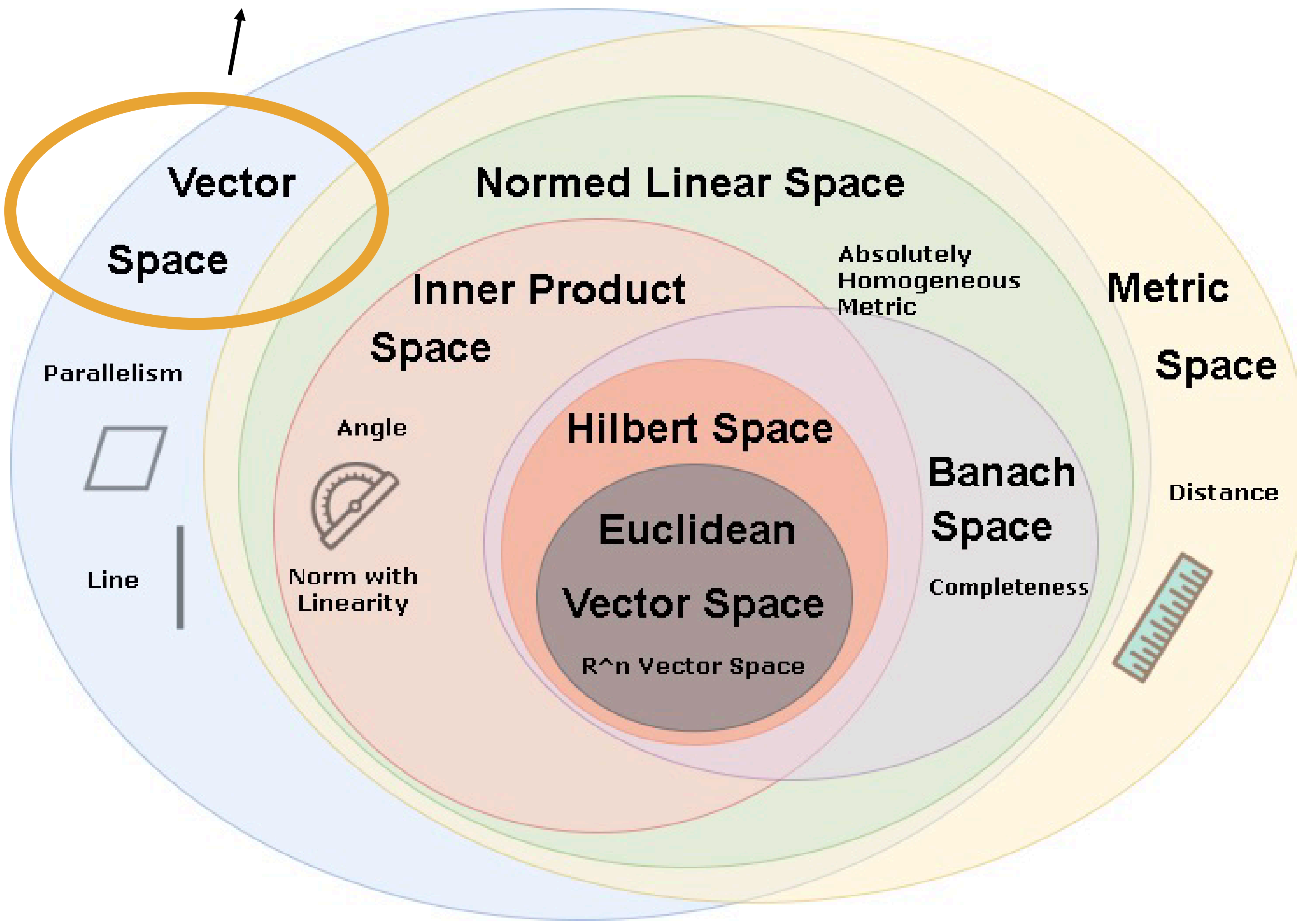
View vector as a “whole”, NOT separate numbers

Then, new properties emerge

Handle/understand inner product and many other things better

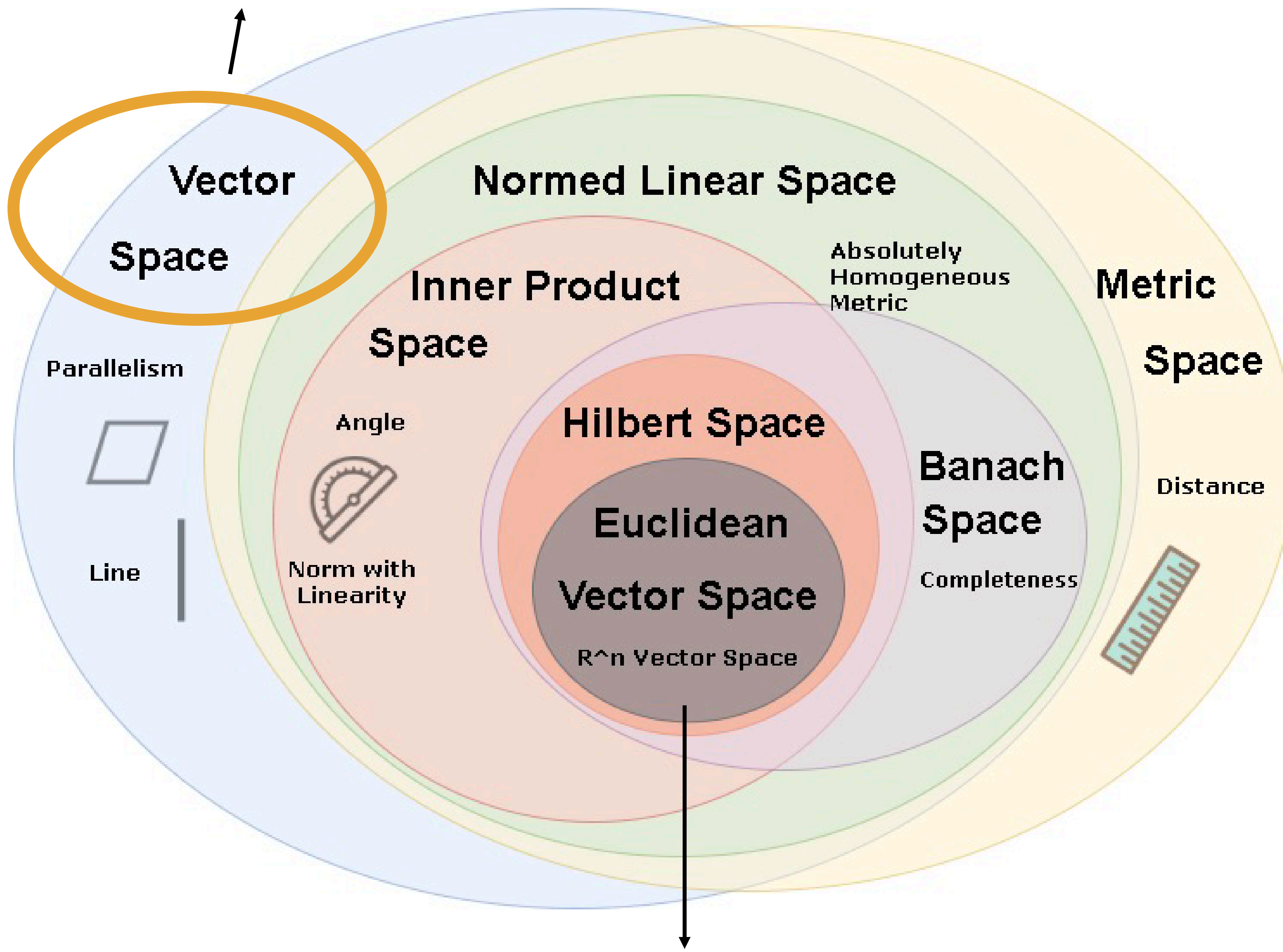
Reading: Dot Product and Spaces

We will mostly study this space throughout this course without digging into more specific spaces (Lecture 9)



Reading: Dot Product and Spaces

We will mostly study this space throughout this course without digging into more specific spaces (Lecture 9)



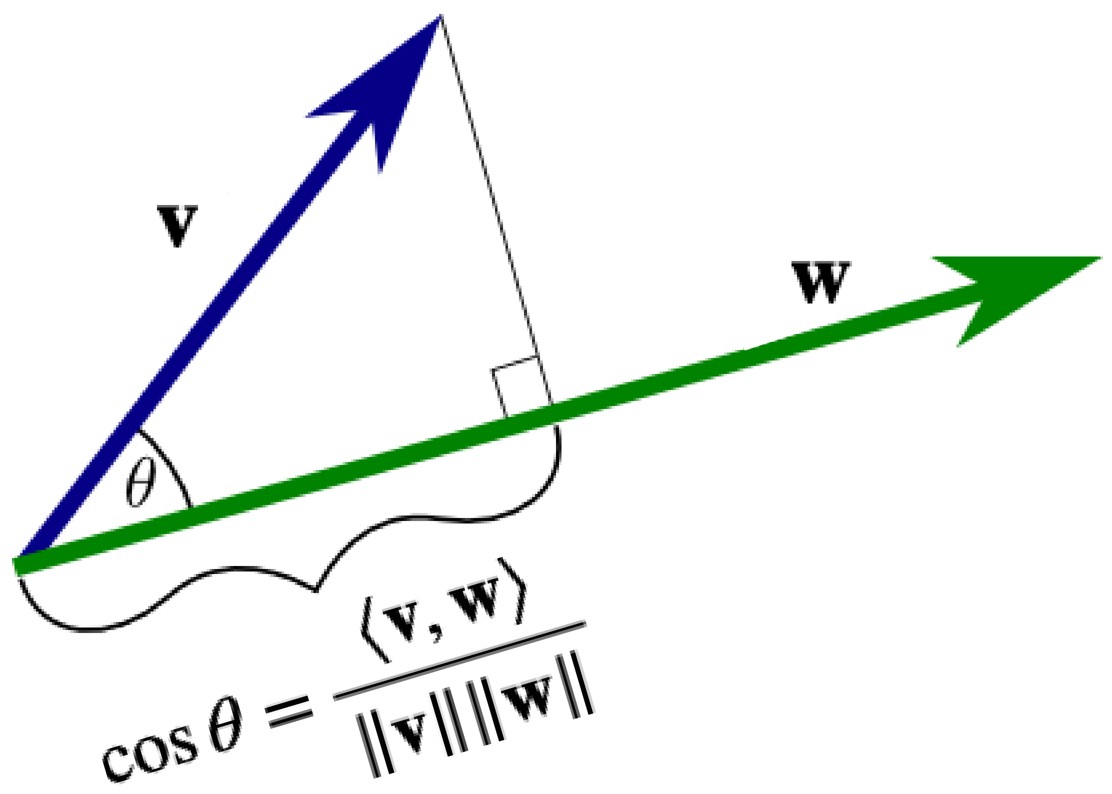
Euclidean space $\mathbb{R}^n$ is the most fundamental and widely used space in our daily life (n is finite)

Part III Properties of Inner Product and Norm

Property 1: Cosine Similarity

Euclidean dot product formula

Cosine Similarity ⟶ $\cos\theta = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\| \|\mathbf{w}\|}$



Property 2: Relation of Dot Product and Norm

Cauchy–Schwarz Inequality

$$|\langle \mathbf{v} \cdot \mathbf{w} \rangle| \leq \|\mathbf{v}\| \|\mathbf{w}\|$$

An "incorrect" explanation:

$$\cos\theta = \frac{\langle \mathbf{v} \cdot \mathbf{w} \rangle}{\|\mathbf{v}\| \|\mathbf{w}\|} \leq 1$$

We will formally prove Cauchy–Schwarz inequality in approx. Week 8!

Property 3: Triangle Inequality

Triangle inequality

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$$

Proof:

Property 4: Pythagoras Law

Pythagoras Law

$$\|v\|^2 + \|w\|^2 = \|v - w\|^2 \quad \text{iff} \quad \langle \mathbf{v}, \mathbf{w} \rangle = 0$$

Proof:

Practice Problem

Problem (Dot Product)

Can we write find three vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ such that in a 2D plane

$$\langle \mathbf{u}, \mathbf{v} \rangle < 0 \qquad \langle \mathbf{u}, \mathbf{w} \rangle < 0 \qquad \langle \mathbf{v}, \mathbf{w} \rangle < 0?$$

Practice Problem

Problem (Dot Product)

Can we write find three vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ such that in a 2D plane

$$\langle \mathbf{u}, \mathbf{v} \rangle < 0 \qquad \langle \mathbf{u}, \mathbf{w} \rangle < 0 \qquad \langle \mathbf{v}, \mathbf{w} \rangle < 0?$$

Answer:

$$\mathbf{u} = (1,0) \quad \mathbf{v} = (-1,4) \quad \mathbf{w} = (-1,-4)$$

How about four vectors?

Summary Today

Today, we have learned: $\|\mathbf{v}\| = \|\mathbf{v}\|_2 := \left(v_1^2 + \cdots + v_n^2\right)^{\frac{1}{2}}$

— Norm of vector

ℓ_2 norm a.k.a. (also known as) Euclidean norm

— Inner product of vector

— Applications of inner product: “feature” and “score”

More Examples

Question: What if we have three vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$

Can we write $\mathbf{x}$ as a linear combination of $\mathbf{y}, \mathbf{z}$?

Why do we need to do this in our real-world?

The next lecture!