Math Refresher for DS

Lecture 1

girafe
ai

Today

1. Course overview
2. Linear Algebra
   - Core objects
   - Vector spaces
3. A bit of Analytic Geometry
   - Orthogonal projections
   - Hyperplanes
   - Normals

About this course

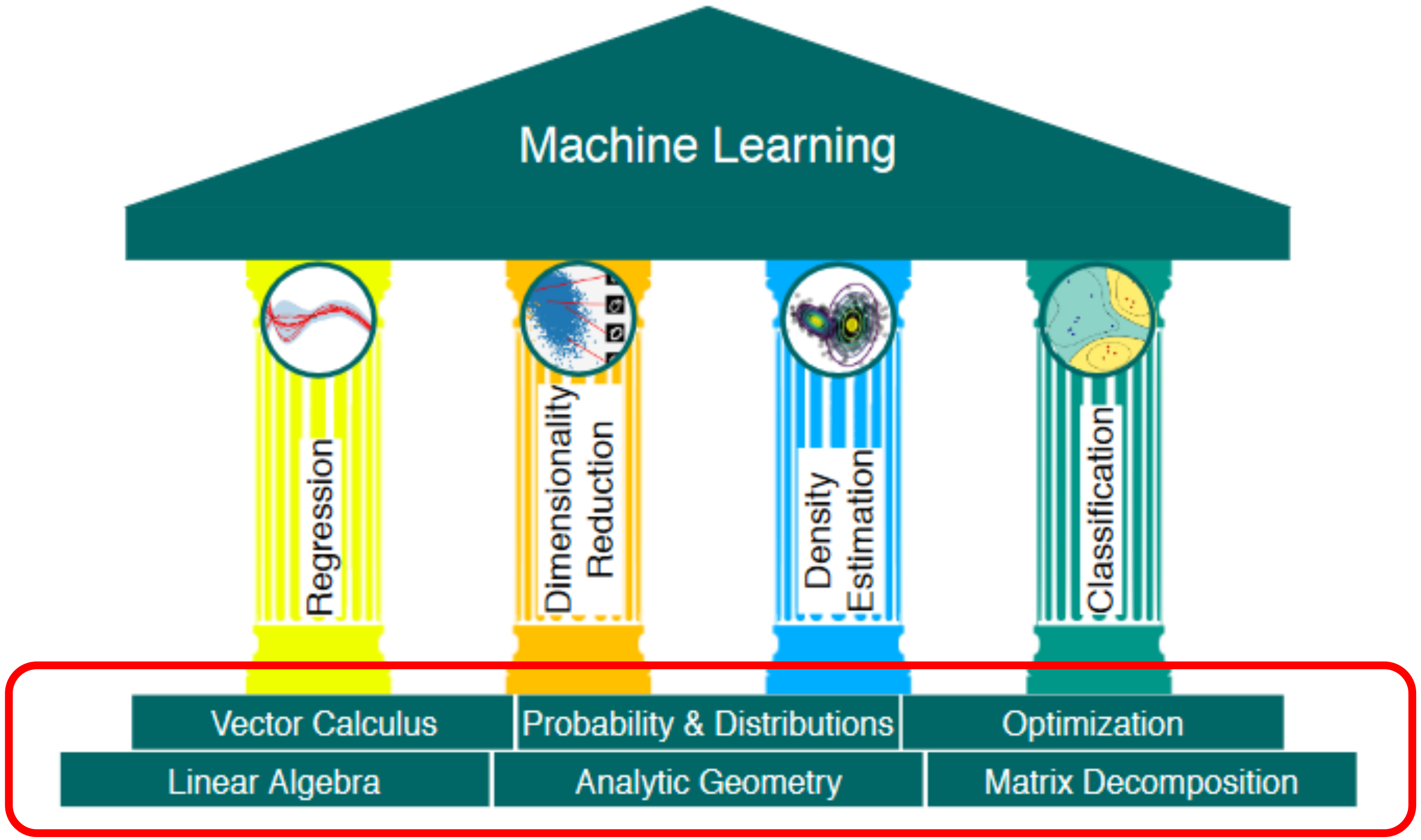


Image source: Mathematics for Machine Learning, p. 14 (https://mml-book.github.io/book/mml-book.pdf)

About this course

In this course:

1. Linear algebra
2. Calculus
3. (Basic) optimization

Prerequisites:

- basic knowledge of math;
- some Python.

About this course

Logistics

- Pre-recorded lectures
- Online practical sessions
  - → Tuesdays & Fridays
    19:00 Moscow time
- 5 graded assignments
- 2 exams
- Final grade:
  - 30% Linear Algebra exam
  - 30% Calculus & Optimization exam
  - 40% graded assignments

About me

Evgeniya Korneva

- evgeniakorneva@gmail.com
- evgeniyako

- PhD researcher KU LEUVEN
- Lecturer HARBOUR SPACE
- DS content lead JetBrains Academy
- *(Soon)* Data Scientists acmetric

Linear Algebra: the Basics

Linear Algebra: Core Objects

- $\alpha \in \mathbb{R}$ - a scalar *Example:* -2

girafe ai

Linear Algebra: Core Objects

- $\alpha \in \mathbb{R}$ - a **scalar** *Example:* -2
- $x \in \mathbb{R}^n$ - a **vector** with n entries

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Example: $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in \mathbb{R}^3$, $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} \in \mathbb{R}^5$

Linear Algebra: Core Objects

- $\alpha \in \mathbb{R}$ - a **scalar** *Example:* -2
- $x \in \mathbb{R}^n$ - a **vector** with n entries

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Example: $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in \mathbb{R}^3, \quad \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} \in \mathbb{R}^5$

- $A \in \mathbb{R}^{m \times n}$ - a **matrix** with m rows and n columns

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Example: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \in \mathbb{R}^{2\times 3}$

girafe ai

Linear Algebra: Core Objects

- $\alpha \in \mathbb{R}$ - a scalar *Example:* -2
- $x \in \mathbb{R}^n$ - a **vector** with n entries

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Example: $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in \mathbb{R}^3, \quad \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} \in \mathbb{R}^5$

- $A \in \mathbb{R}^{m \times n}$ - a **matrix** with m rows and n columns

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

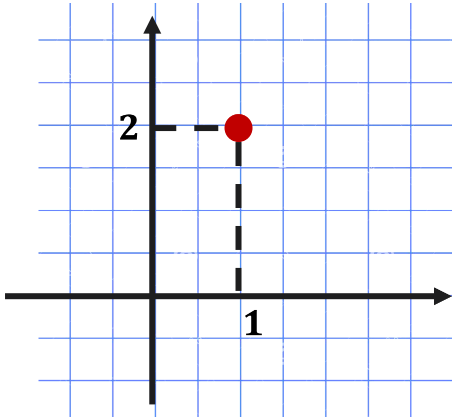
Example: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \in \mathbb{R}^{2\times 3}$

What are Vectors?

- Ordered sets of numbers: $x = [1, 2]$

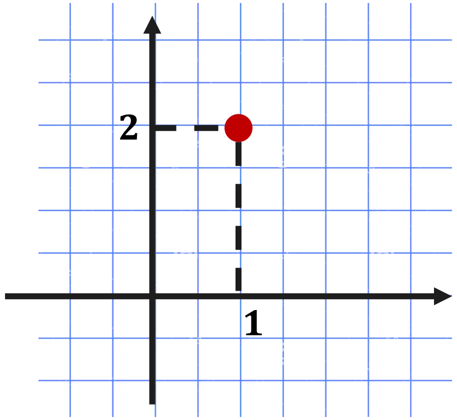
What are Vectors?

- Ordered sets of numbers: $x = [1, 2]$
- A point with Cartesian coordinates

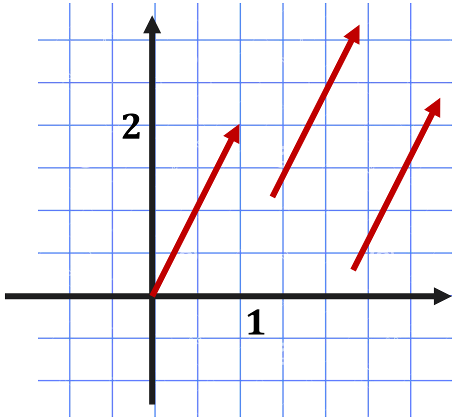


What are Vectors?

- Ordered sets of numbers: $x = [1, 2]$
- A point with Cartesian coordinates



- Direction + length



girafe ai

Vector Spaces

Vector Space: Definition

- A real-valued **vector space** $(V, +, \cdot)$ is a set of vectors V with two operations

$$(1)\ +: \ V \times V \to V, \qquad (2)\ \cdot: \ \mathbb{R} \times V \to V$$

Vector Space: Definition

- A real-valued **vector space** $(V, +, \cdot)$ is a set of vectors V with two operations

$$(1) +: \ V \times V \to V, \qquad (2) \cdot: \ \mathbb{R} \times V \to V$$

that satisfy the following properties (axioms):

	Property	Meaning
1.	**Associativity** of addition	$x + (y + z) = (x + y) + z$
2.	**Commutativity** of addition	$x + y = y + x$
3.	**Identity element** of addition	$\exists 0 \in V: \ \forall x \in V \quad 0 + x = x$
4.	**Identity element** of scalar multiplication	$\forall x \in V \quad 1 \cdot x = x$
5.	**Inverse element** of addition	$\forall x \in V \ \exists -x \in V: \ x + (-x) = 0$
6.	**Compatibility** of scalar multiplication	$\alpha(\beta x) = (\alpha\beta)x$
7.	**Distributivity**	$(\alpha + \beta)x = \alpha x + \beta x$
8.		$\alpha(x + y) = \alpha x + \alpha y$

girafe ai

Let’s define vector operations!

Operations with Vectors

1. Sum of two vectors:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n, \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n, \qquad x + y = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_2 \end{bmatrix} \in \mathbb{R}^n$$

Operations with Vectors

1. Sum of two vectors:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n, \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n, \qquad x + y = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_2 \end{bmatrix} \in \mathbb{R}^n$$

2. Multiplying by a scalar:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n, \qquad \alpha \in \mathbb{R}, \qquad \alpha x = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix} \in \mathbb{R}^n$$

Operations with Vectors: Example

$x, y \in \mathbb{R}^3$:

$$x = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad y = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Operations with Vectors: Example

$x, y \in \mathbb{R}^3$:

$$x = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad y = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Multiplying by a scalar:

$$10x = \begin{bmatrix} 10 \\ 0 \\ 10 \end{bmatrix}$$

girafe ai

Operations with Vectors: Example

$x, y \in \mathbb{R}^3$:

$$x = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad y = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Multiplying by a scalar:

$$10x = \begin{bmatrix} 10 \\ 0 \\ 10 \end{bmatrix}$$

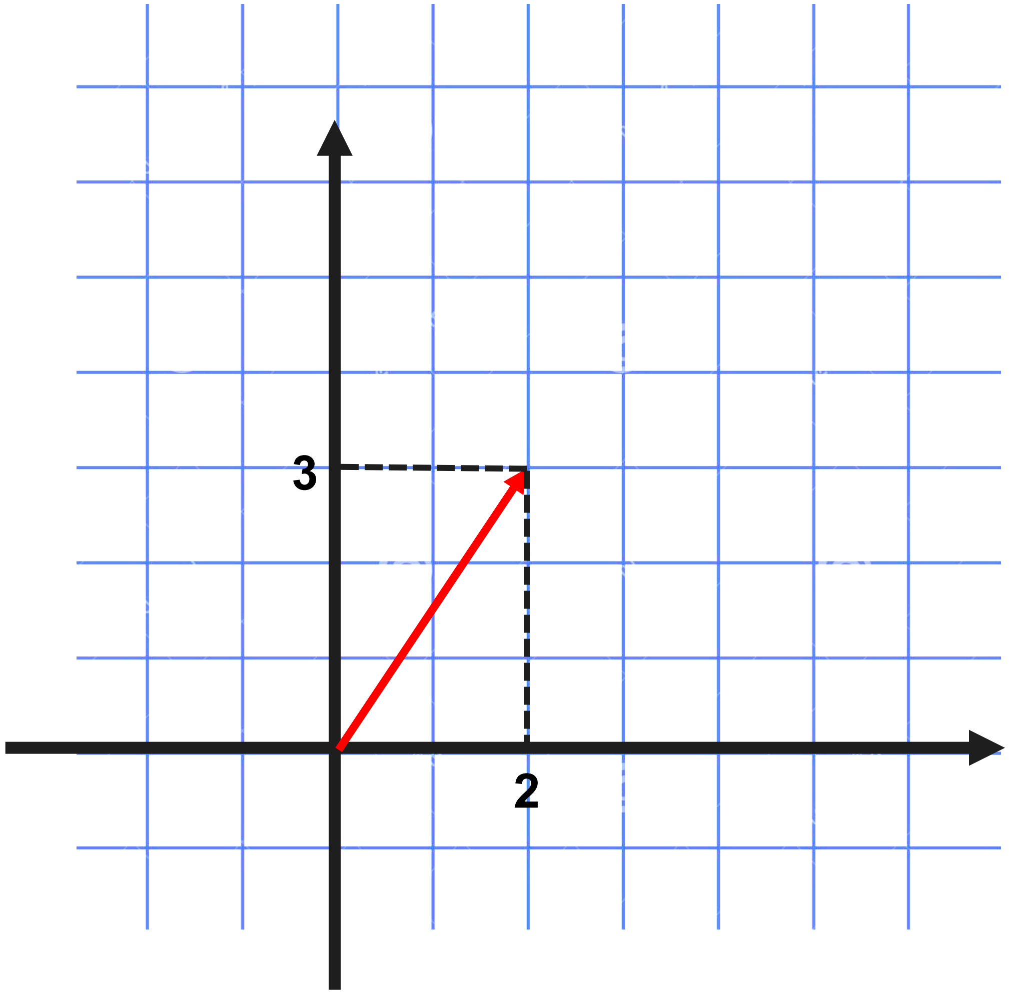
Sum:

$$x + y = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

Vector Operations: Geometrical Interpretation

Vectors: Geometrical Interpretation

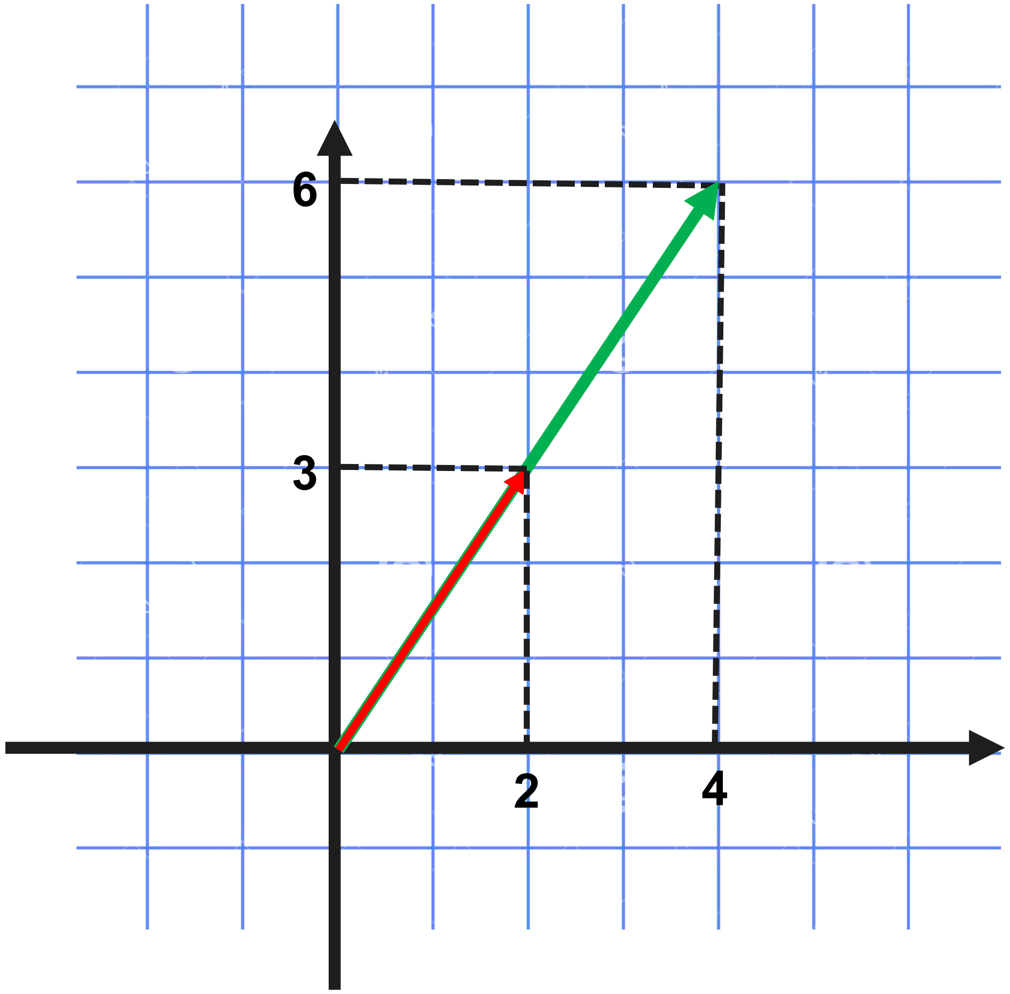
$\vec{a} = [2, 3]$



Vectors: Geometrical Interpretation

$\vec{a} = [2, 3]$

$2\vec{a} = [4, 6]$

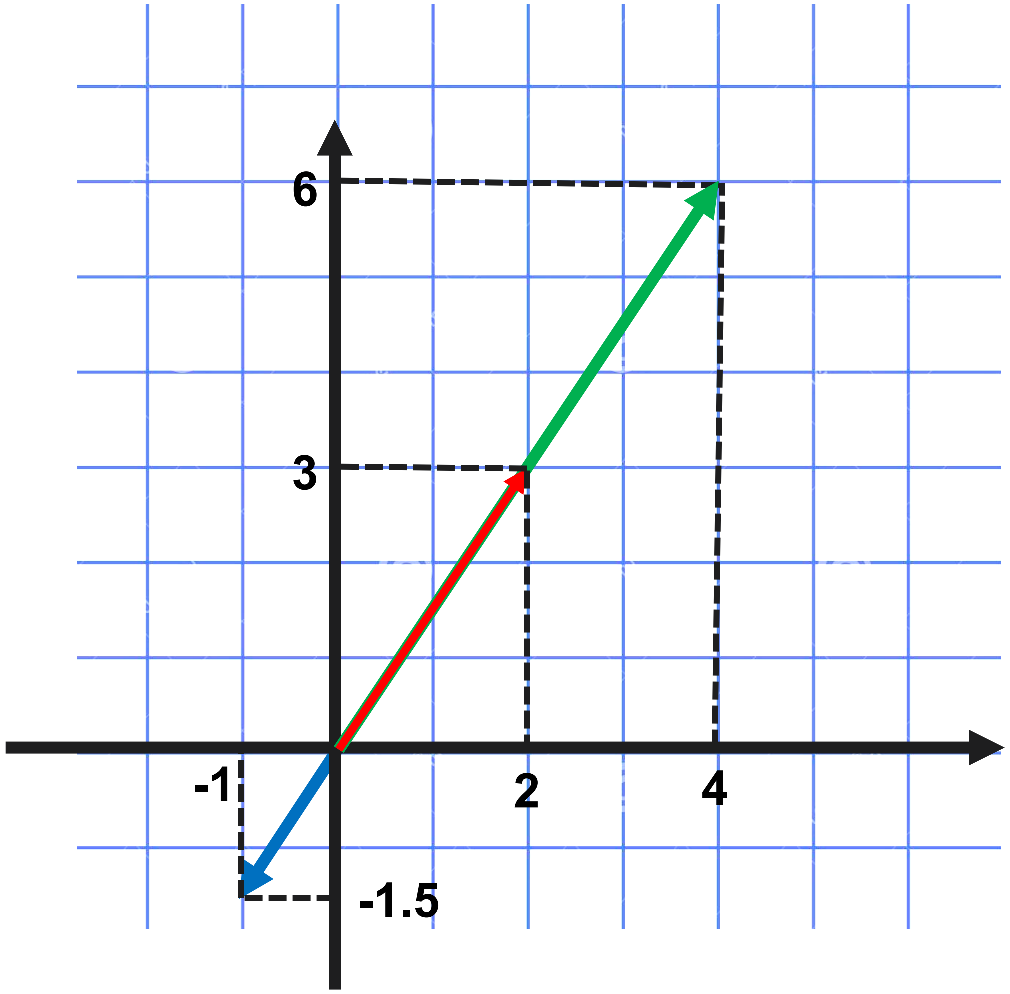


Vectors: Geometrical Interpretation

$$\vec{a} = [2, 3]$$

$$2\vec{a} = [4, 6]$$

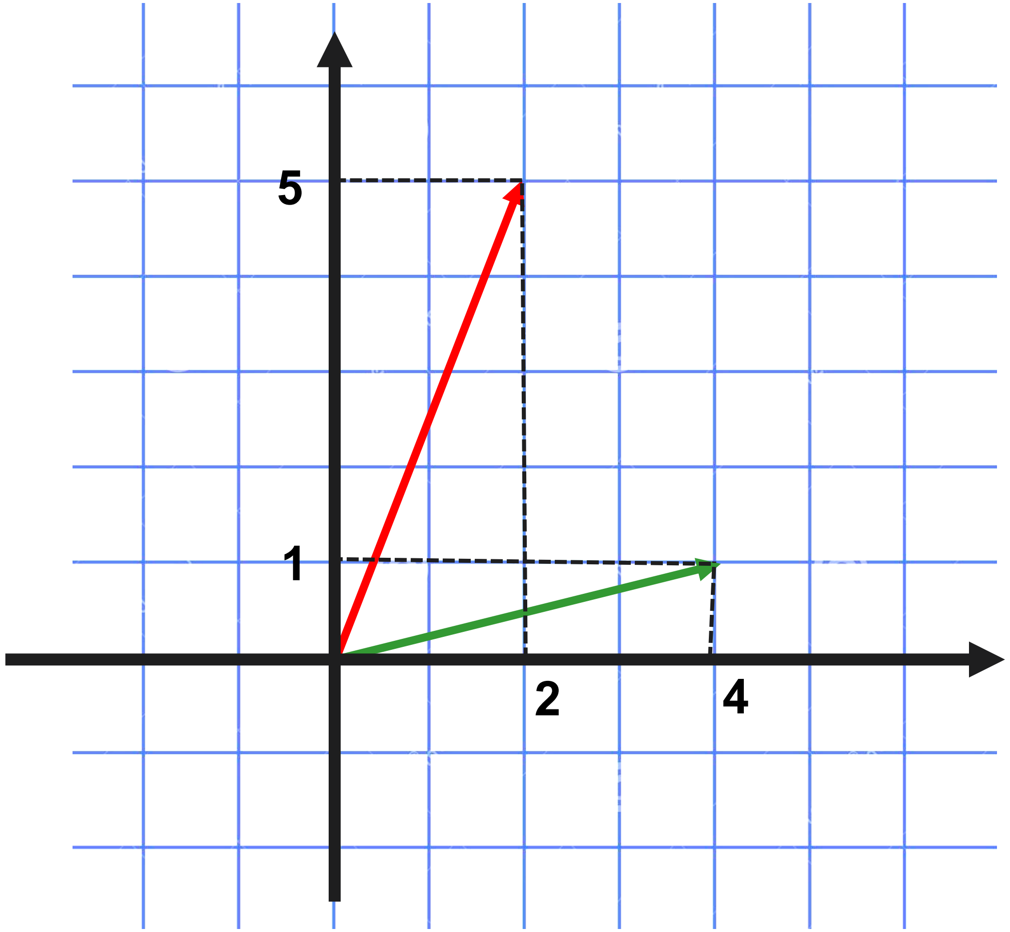
$$-0.5\vec{a} = [-1, -1.5]$$



Vectors: Geometrical Interpretation

$\vec{a} = [2, 5]$

$\vec{b} = [4, 1]$

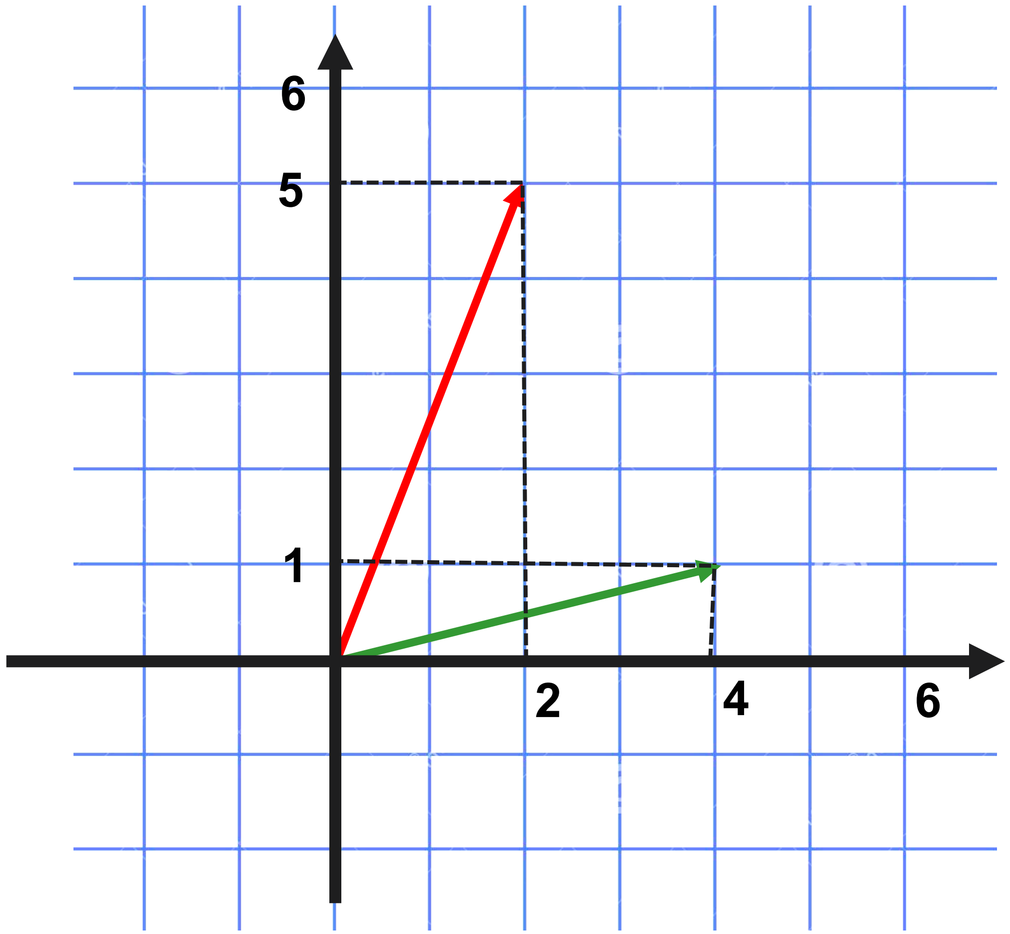


Vectors: Geometrical Interpretation

$\vec{a} = [2, 5]$

$\vec{b} = [4, 1]$

$\vec{a} + \vec{b} = [6, 6]$

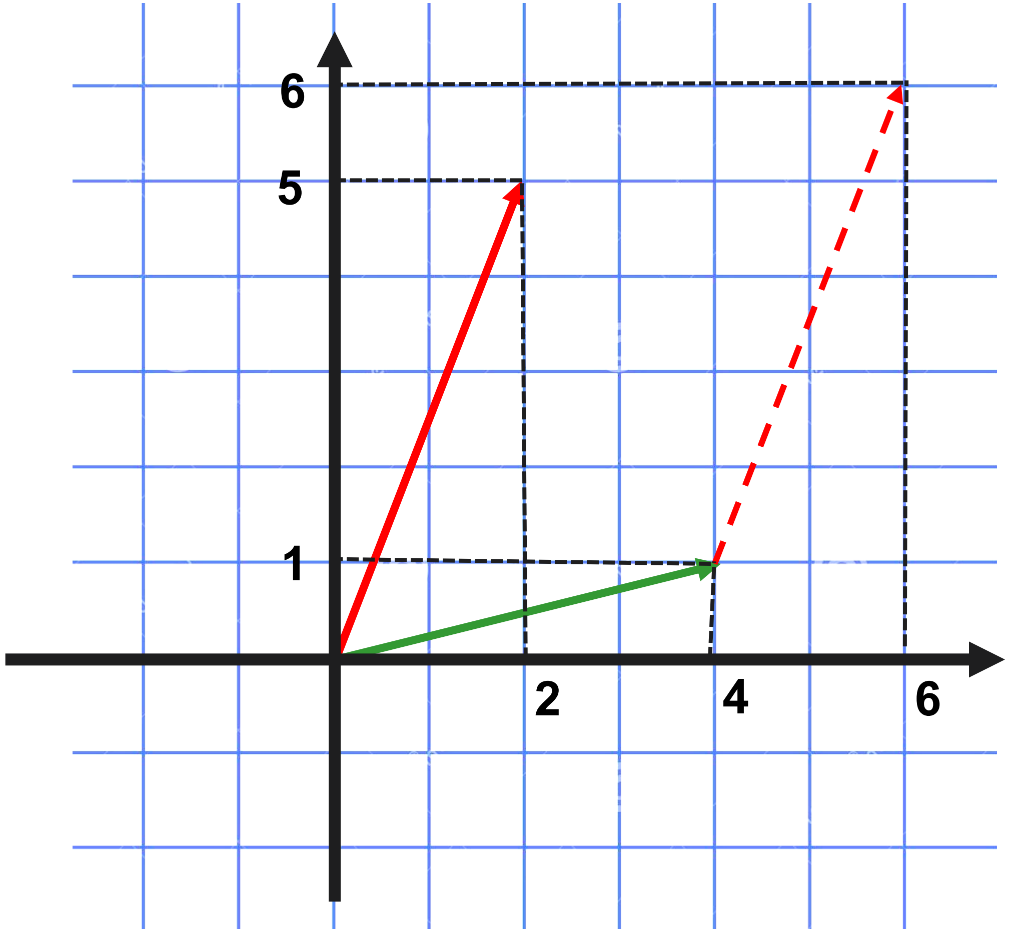


Vectors: Geometrical Interpretation

$\vec{a} = [2, 5]$

$\vec{b} = [4, 1]$

$\vec{a} + \vec{b} = [6, 6]$

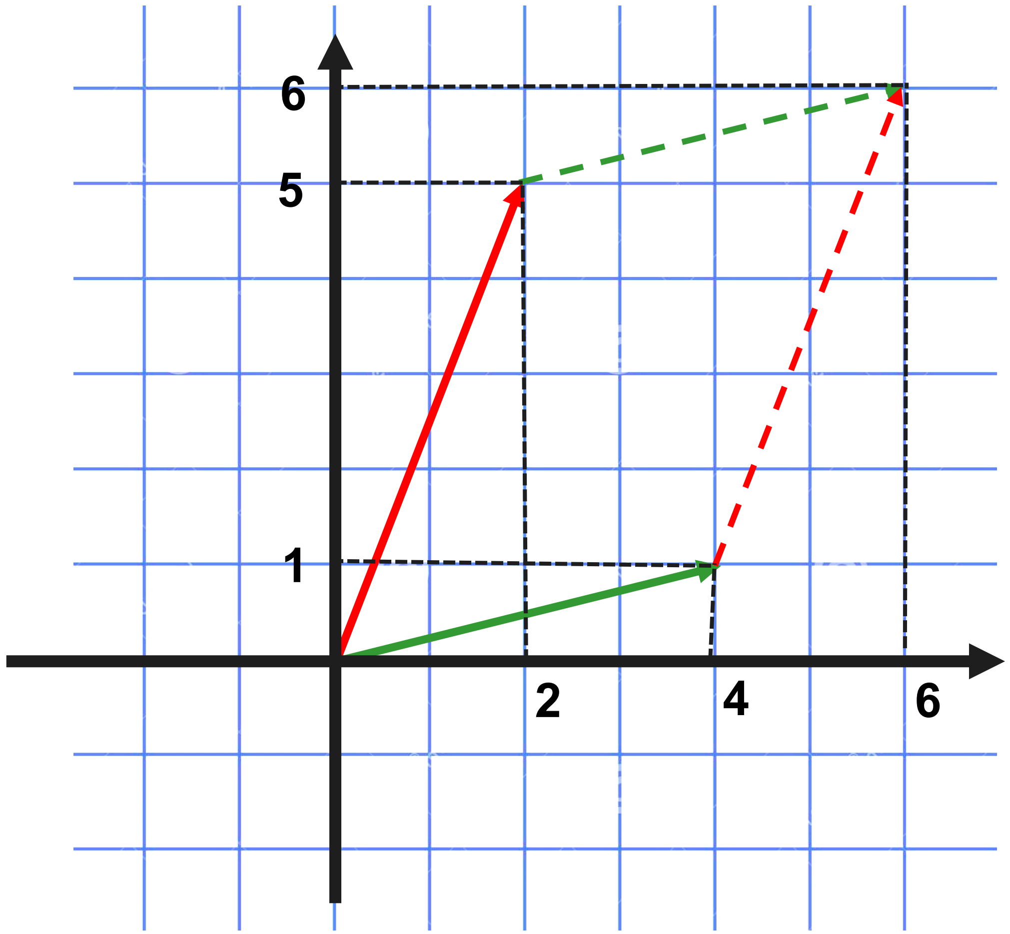


Vectors: Geometrical Interpretation

$\vec{a} = [2, 5]$

$\vec{b} = [4, 1]$

$\vec{a} + \vec{b} = [6, 6]$

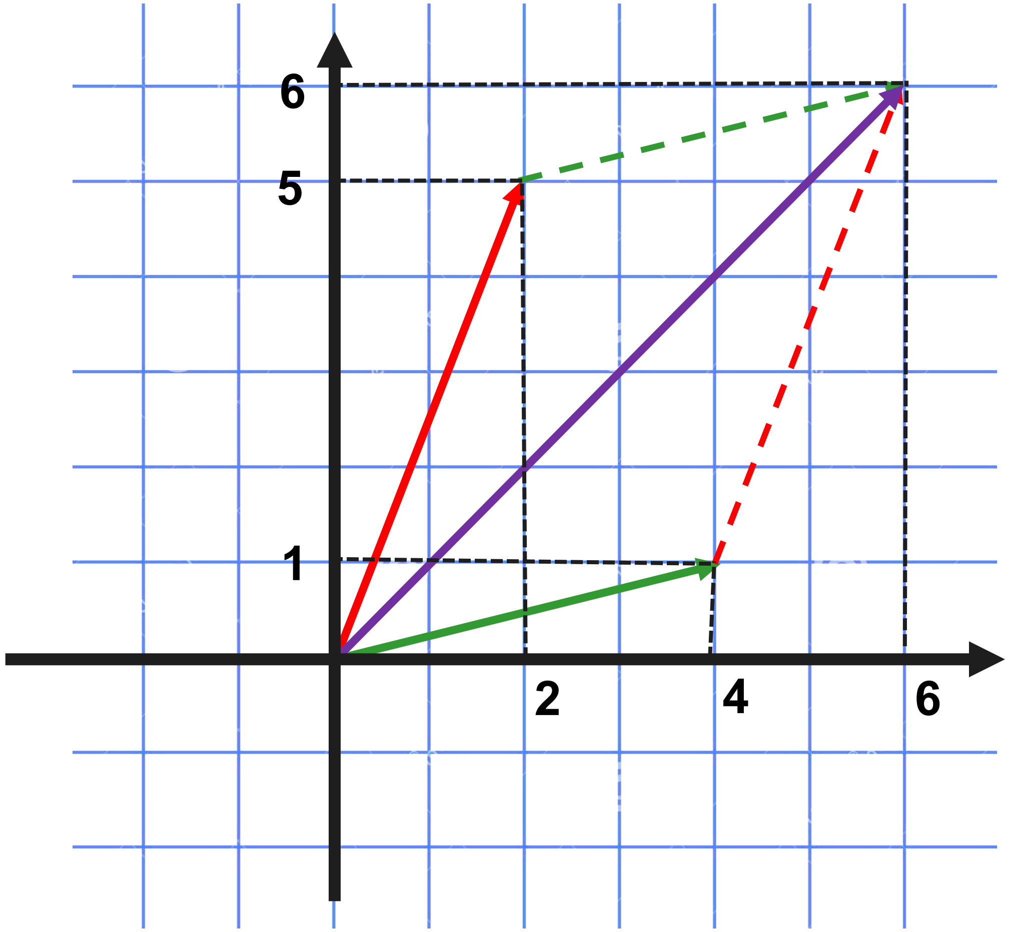


Vectors: Geometrical Interpretation

$$\vec{a} = [2, 5]$$

$$\vec{b} = [4, 1]$$

$$\vec{a} + \vec{b} = [6, 6]$$



Back to Vector Spaces

Operations with Vectors

1. Sum of two vectors:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \qquad x + y = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_2 \end{bmatrix} \in \mathbb{R}^n$$

2. Multiplying by a scalar:

satisfy axioms (1) – (8)
(check it yourself)

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \alpha \in \mathbb{R}, \qquad \alpha x = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix} \in \mathbb{R}^n$$

Vector Spaces

$(\mathbb{R}^n, +, \cdot), n \in \mathbb{N}$ - a vector space with operations

1. vector addition:

$$x + y = (x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

2. multiplication by a scalar:

$$\alpha x = \alpha(x_1, x_2, \dots, x_n) = (\alpha x_1, \alpha x_2, \dots, \alpha x_n)$$

Vector Spaces: Another Example

Vector Spaces: Another Example

- $\mathbb{P}^n$ - a set of polynomials of degree $\leq n$ with real coefficients
  - *Example:* $\mathbb{P}^3 = \{ax^3 + bx^2 + cx + d \mid a, b, c, d \in \mathbb{R}\}$

Vector Spaces: Another Example

- $\mathbb{P}^n$ - a set of polynomials of degree $\leq n$ with real coefficients
  - *Example:* $\mathbb{P}^3 = \{ax^3 + bx^2 + cx + d \mid a, b, c, d \in \mathbb{R}\}$
- We can add up polynomials:

$$(a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0) + (b_n x^n + b_{n-1} x^{n-1} + \cdots + b_0) =$$
$$= (a_n + b_n) x^n + (a_{n-1} + b_{n-1}) x^{n-1} + \cdots + (a_0 + b_0) \in \mathbb{P}^n$$

Vector Spaces: Another Example

- $\mathbb{P}^n$ - a set of polynomials of degree $\leq n$ with real coefficients
  - *Example:* $\mathbb{P}^3 = \{ax^3 + bx^2 + cx + d \mid a, b, c, d \in \mathbb{R}\}$
- We can add up polynomials:

$$(a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0) + (b_n x^n + b_{n-1} x^{n-1} + \cdots + b_0) =$$
$$= (a_n + b_n) x^n + (a_{n-1} + b_{n-1}) x^{n-1} + \cdots + (a_0 + b_0) \in \mathbb{P}^n$$

- And we can multiply them by a real number:

$$\lambda (a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0) = \lambda a_n x^n + \lambda a_{n-1} x^{n-1} + \cdots + \lambda a_0 \in \mathbb{P}^n$$

Vector Spaces: Another Example

- $\mathbb{P}^n$ - a set of polynomials of degree $\leq n$ with real coefficients
  - *Example:* $\mathbb{P}^3 = \{ax^3 + bx^2 + cx + d \mid a, b, c, d \in \mathbb{R}\}$
- We can add up polynomials:

$$(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0) + (b_n x^n + b_{n-1}x^{n-1} + \cdots + b_0) =$$
$$= (a_n + b_n)x^n + (a_{n-1} + b_{n-1})x^{n-1} + \cdots + (a_0 + b_0) \in \mathbb{P}^n$$

- And we can multiply them by a real number:

$$\lambda(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0) = \lambda a_n x^n + \lambda a_{n-1}x^{n-1} + \cdots + \lambda a_0 \in \mathbb{P}^n$$

- These operations satisfy axioms (1) – (8)!

$\rightarrow (\mathbb{P}^n, +, \cdot)$ is also a vector space!

girafe ai

Inner Product

Inner Product

- Inner product is a function $\langle \cdot , \cdot \rangle : V \times V \to \mathbb{R}$ that satisfies the following properties:
    - *Symmetric:* $\forall x, y \in V \;\; \langle x, y \rangle = \langle y, x \rangle$
    - *Positive definite:* $\forall x \in V \setminus \{0\} \;\; \langle x, x \rangle > 0$ and $\langle x, 0 \rangle = 0$.

Dot Product

- A particular type of inner product.

Dot Product

- A particular type of inner product.

- For $x = [x_1, \dots, x_n],\ y = [y_1, \dots, y_n] \in \mathbb{R}^n$

$$(x, y) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Dot Product

- A particular type of inner product.
- For $x = [x_1, \dots, x_n],\ y = [y_1, \dots, y_n] \in \mathbb{R}^n$

$$(x, y) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

- Example:

$$x = [1, 2, 3, 4], \qquad y = [-1, 0, 1, 2]$$

$$(x, y) = 1 \cdot (-1) + 2 \cdot 0 + 3 \cdot 1 + 4 \cdot 2 = -1 + 0 + 3 + 8 = 10$$

Euclidian Vector Space

- A vector space with inner product is called an *inner product space*.

Euclidian Vector Space

- A vector space with inner product is called an *inner product space*.

- $(\mathbb{R}^n, +, \cdot)$ with dot product $\langle \cdot, \cdot \rangle$ is called a *Euclidian vector space*.

girafe ai

Euclidian Vector Space

- A vector space with inner product is called an *inner product space*.

- $(\mathbb{R}^n, +, \cdot)$ with dot product $(\cdot,\cdot)$ is called a *Euclidian vector space*.

- *Note: there're inner products different from dot product.*

Norms

Norm

- A **norm** on a vector space $\mathbb{V}$ is a function $\|\cdot\|: \mathbb{V} \to \mathbb{R}$ which assigns each vector $x \in \mathbb{V}$ its *length* $\|x\| \in \mathbb{R}$.

Norm

- A **norm** on a vector space $\mathbb{V}$ is a function $\|\cdot\|: \mathbb{V} \to \mathbb{R}$ which assigns each vector $x \in \mathbb{V}$ its *length* $\|x\| \in \mathbb{R}$.

- A norm must satisfy the following properties:

Norm

- A **norm** on a vector space $\mathbb{V}$ is a function $\|\cdot\|: \mathbb{V} \to \mathbb{R}$ which assigns each vector $x \in \mathbb{V}$ its *length* $\|x\| \in \mathbb{R}$.

- A norm must satisfy the following properties:
  - *Absolutely homogeneous:* $\forall x \in \mathbb{V},\ \forall \lambda \in \mathbb{R}\quad \|\lambda x\| = |\lambda|\|x\|$

girafe ai

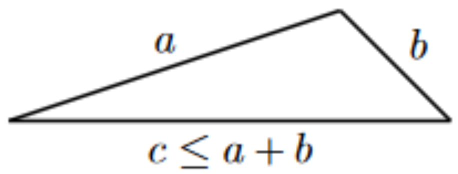
Norm

- A **norm** on a vector space $\mathbb{V}$ is a function $\|\cdot\|: \mathbb{V} \to \mathbb{R}$ which assigns each vector $x \in \mathbb{V}$ its *length* $\|x\| \in \mathbb{R}$.
- A norm must satisfy the following properties:
  - *Absolutely homogeneous:* $\forall x \in \mathbb{V},\ \forall \lambda \in \mathbb{R} \quad \|\lambda x\| = |\lambda|\|x\|$
  - *Positive definite:* $\forall x \in \mathbb{V} \quad \|x\| \geq 0$ and $\|x\| = 0 \Leftrightarrow x = 0$

Norm

- A **norm** on a vector space $\mathbb{V}$ is a function $\|\cdot\|: \mathbb{V} \to \mathbb{R}$ which assigns each vector $x \in \mathbb{V}$ its *length* $\|x\| \in \mathbb{R}$.

- A norm must satisfy the following properties:
  - *Absolutely homogeneous:* $\forall x \in \mathbb{V},\ \forall \lambda \in \mathbb{R} \quad \|\lambda x\| = |\lambda|\|x\|$
  - *Positive definite:* $\forall x \in \mathbb{V} \quad \|x\| \geq 0$ and $\|x\| = 0 \Leftrightarrow x = 0$
  - *Triangle inequality:* $\forall x, y \in \mathbb{V} \ \|x + y\| \leq \|x\| + \|y\|$



girafe ai

Examples of Norms

Manhattan Norm

- A norm for $x \in \mathbb{R}^n$:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

Manhattan Norm

- A norm for $x \in \mathbb{R}^n$:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

- Examples:

$$\|[1, 2, 3]\|_1 = 1 + 2 + 3 = 6$$

$$\|[1, 0]\|_1 = 1 + 0 = 1$$

$$\|[-1, 0]\|_1 = |-1| + 0 = 1$$

Manhattan Norm

- A norm for $x \in \mathbb{R}^n$:

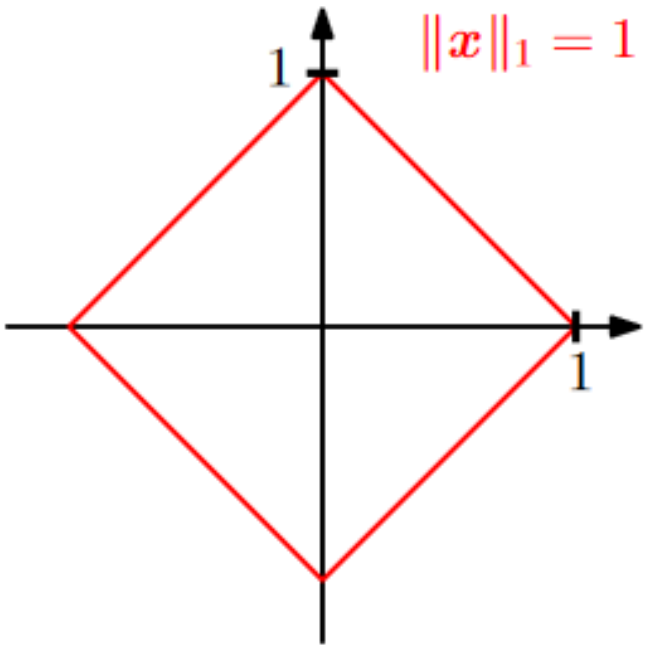
$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

- Examples:

$$\|[1,2,3]\|_1 = 1 + 2 + 3 = 5$$

$$\|[1,0]\|_1 = 1 + 0 = 1$$

$$\|[-1,0]\|_1 = |-1| + 0 = 1$$



Euclidian Norm

- A norm for $x \in \mathbb{R}^n$:

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

- Computes *Euclidian distance* from the origin.

Euclidian Norm

- A norm for $x \in \mathbb{R}^n$:

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

- Computes *Euclidian distance* from the origin.

Euclidian Norm

- A norm for $x \in \mathbb{R}^n$:

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

- Computes *Euclidian distance* from the origin.

- Examples:

$$\|[1,2,3]\|_2 = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$$

$$\|[1,0]\|_2 = \sqrt{1^2 + 0} = 1$$

$$\|[-1,0]\|_2 = \sqrt{(-1)^2 + 0} = 1$$

Euclidian Norm

- A norm for $x \in \mathbb{R}^n$:

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

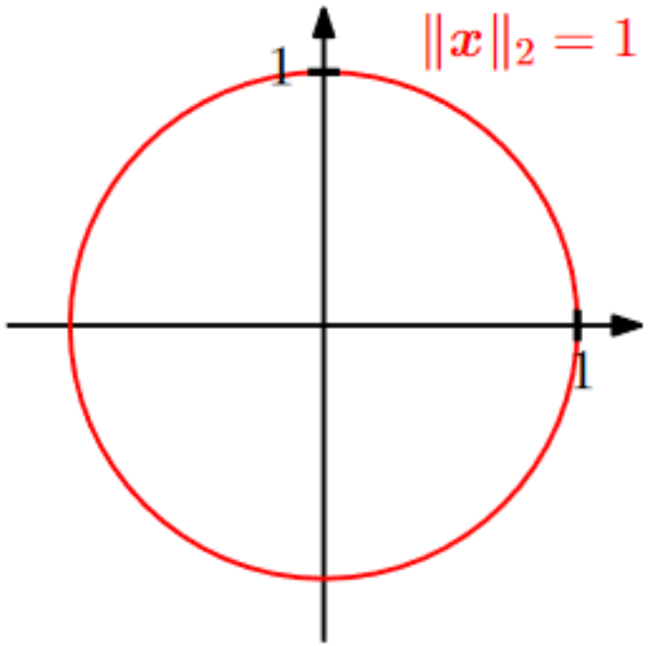
- Computes *Euclidian distance* from the origin.

- Examples:

$$\|[1,2,3]\|_2 = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$$

$$\|[1,0]\|_2 = \sqrt{1^2 + 0} = 1$$

$$\|[-1,0]\|_2 = \sqrt{(-1)^2 + 0} = 1$$



Other norms

- In general, for $x = [x_1, \dots, x_n] \in \mathbb{R}^n$ an $\boldsymbol{\ell_p}$-norm is defined as follows:

$$\|x\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}\,.$$

Other norms

- In general, for $x = [x_1, \dots, x_n] \in \mathbb{R}^n$ an $\boldsymbol{\ell_p}$-norm is defined as follows:

$$\|x\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}\,.$$

  - ℓ_1 - Manhattan norm $\|\cdot\|_1$;
  - ℓ_2 - Euclidian norm $\|\cdot\|$ (default);

Other norms

- In general, for $x = [x_1, \dots, x_n] \in \mathbb{R}^n$ an $\boldsymbol{\ell_p}$-norm is defined as follows:

$$\|x\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}\,.$$

  - ℓ_1 - Manhattan norm $\|\cdot\|_1$;
  - ℓ_2 - Euclidian norm $\|\cdot\|$ (default);
  - ℓ_∞: $\| x \|_\infty = \max_i |x_i|$

    Example: $\|[1,2,3]\|_\infty = 3,\quad \|[1,0]\|_\infty = 1,\quad \|[-1,0.5]\|_\infty = 1.$

Inner Product and Norms

- Any inner product induces a norm:

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

Inner Product and Norms

- Any inner product induces a norm:

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

Check yourself that this is indeed a norm.

Inner Product and Norms

- Any inner product induces a norm:

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

Check yourself that this is indeed a norm.

- Example: dot product induces Euclidian norm:

$$\sqrt{(x, x)} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \|x\|_2$$

girafe ai

Inner Product and Norms

- Any inner product induces a norm:

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

Check yourself that this is indeed a norm.

- Example: dot product induces Euclidian norm:

$$\sqrt{(x, x)} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \|x\|_2$$

- (!) Not every norm is induced by an inner product.
  Example: Manhattan norm.

Cauchy-Schwarz Inequality

- For an inner product vector space, the induced norm satisfies the inequality:

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$$

Cauchy-Schwarz Inequality

- For an inner product vector space, the induced norm satisfies the inequality:

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$$

- For dot product and Euclidian norm:

$$|(x, y)| \leq \|x\|_2 \cdot \|y\|_2$$

Distance between Vectors

- Distance between two vectors x and y is defined as

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

Distance between Vectors

- Distance between two vectors x and y is defined as

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

- For dot product and Euclidian norm, we get *Euclidian distance*:

$$d(x, y) = \|x - y\|_2 = \sqrt{(x - y, x - y)} =$$

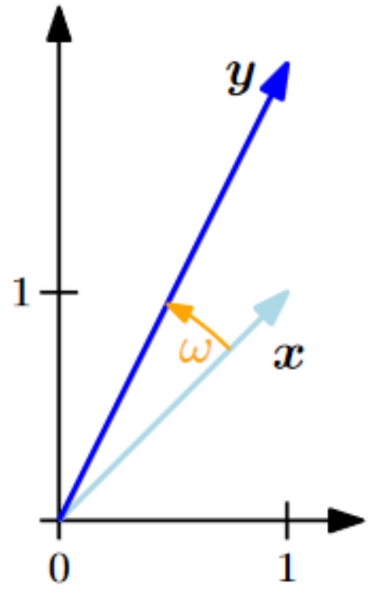
$$= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}.$$

Angles and Orthogonality

Angle between Two Vectors

- Inner product also captures the geometry of vector space by defining the angle between two vectors.

- Remember Cauchy-Schwarz inequality:



$$|(x, y)| \leq \|x\| \cdot \|y\|$$

$$-1 \leq \frac{(x, y)}{\|x\| \cdot \|y\|} \leq 1$$

$\omega\colon\ \cos\omega = \frac{(x,y)}{\|x\|\cdot\|y\|}$ - angle between x and y.

Angle between Two Vectors: Example

- What is the angle ω between $x = [5, 0]$ and $y = [1, 1]$?

$$\omega = \arccos \frac{(x, y)}{\|x\| \|y\|} = \arccos \frac{5 \cdot 1 + 0 \cdot 1}{\sqrt{5^2 + 0^2} \cdot \sqrt{1^2 + 1^2}} = \arccos \frac{5}{5\sqrt{2}} = \arccos \frac{\sqrt{2}}{4} = \frac{\pi}{4}.$$

Angle between Two Vectors: Example

- What is the angle ω between $x = [5, 0]$ and $y = [1, 1]$?

$$\omega = \arccos \frac{(x, y)}{\|x\|\|y\|} = \arccos \frac{5 \cdot 1 + 0 \cdot 1}{\sqrt{5^2 + 0^2} \cdot \sqrt{1^2 + 1^2}} = \arccos \frac{5}{5\sqrt{2}} = \arccos \frac{\sqrt{2}}{4} = \frac{\pi}{4}.$$

- What is the angle ω between $x = \left[\sqrt{3}, 0{,}1\right]$ and $y = [1, 0, 0]$?

$$\omega = \arccos \frac{(x, y)}{\|x\|\|y\|} = \arccos \frac{\sqrt{3}}{\sqrt{3 + 1} \cdot \sqrt{1}} = \arccos \frac{\sqrt{3}}{2} = \frac{\pi}{6}.$$

Angle between Two Vectors: Example

- What is the angle ω between $x = [5, 0]$ and $y = [1, 1]$?

$$\omega = \arccos \frac{(x, y)}{\|x\|\|y\|} = \arccos \frac{5 \cdot 1 + 0 \cdot 1}{\sqrt{5^2 + 0^2} \cdot \sqrt{1^2 + 1^2}} = \arccos \frac{5}{5\sqrt{2}} = \arccos \frac{\sqrt{2}}{4} = \frac{\pi}{4}.$$

- What is the angle ω between $x = \left[\sqrt{3}, 0, 1\right]$ and $y = [1, 0, 0]$?

$$\omega = \arccos \frac{(x, y)}{\|x\|\|y\|} = \arccos \frac{\sqrt{3}}{\sqrt{3 + 1} \cdot \sqrt{1}} = \arccos \frac{\sqrt{3}}{2} = \frac{\pi}{6}.$$

- What is the angle ω between $x = [1, 0, 0, 0, 1]$ and $y = [0, 1, 1, 0, 0]$?

$$\omega = \arccos \frac{(x, y)}{\|x\|\|y\|} = \arccos \frac{0}{\sqrt{2} \cdot \sqrt{2}} = \arccos 0 = \frac{\pi}{2}.$$

Orthogonality

- x and y are orthogonal if and only if $\langle x, y \rangle = 0$.

Orthogonality

- x and y are orthogonal if and only if $\langle x, y \rangle = 0$.
- Examples in $\mathbb{R}^n$ with dot product:

$$x = [2, 3], \qquad y = [1, 0], \qquad (x, y) = 2 \neq 0 \quad \rightarrow$$

x and y are not orthogonal.

Orthogonality

- x and y are orthogonal if and only if $\langle x, y \rangle = 0$.
- Examples in $\mathbb{R}^n$ with dot product:

$$x = [2, 3], \qquad y = [1, 0], \qquad (x, y) = 2 \neq 0 \;\; \rightarrow$$

x and y are not orthogonal.

$$x = [1, 2, 3], \qquad y = [-2,1, 0], \qquad (x, y) = -2 + 2 + 0 = 0 \;\; \rightarrow$$

x and y are orthogonal.

Orthogonality

- x and y are orthogonal if and only if $\langle x, y\rangle = 0$.
- Examples in $\mathbb{R}^n$ with dot product:

$$x = [2, 3], \qquad y = [1, 0], \qquad (x, y) = 2 \neq 0 \;\; \rightarrow$$

x and y are not orthogonal.

$$x = [1, 2, 3], \qquad y = [-2,1, 0], \qquad (x, y) = -2 + 2 + 0 = 0 \;\; \rightarrow$$

x and y are orthogonal.

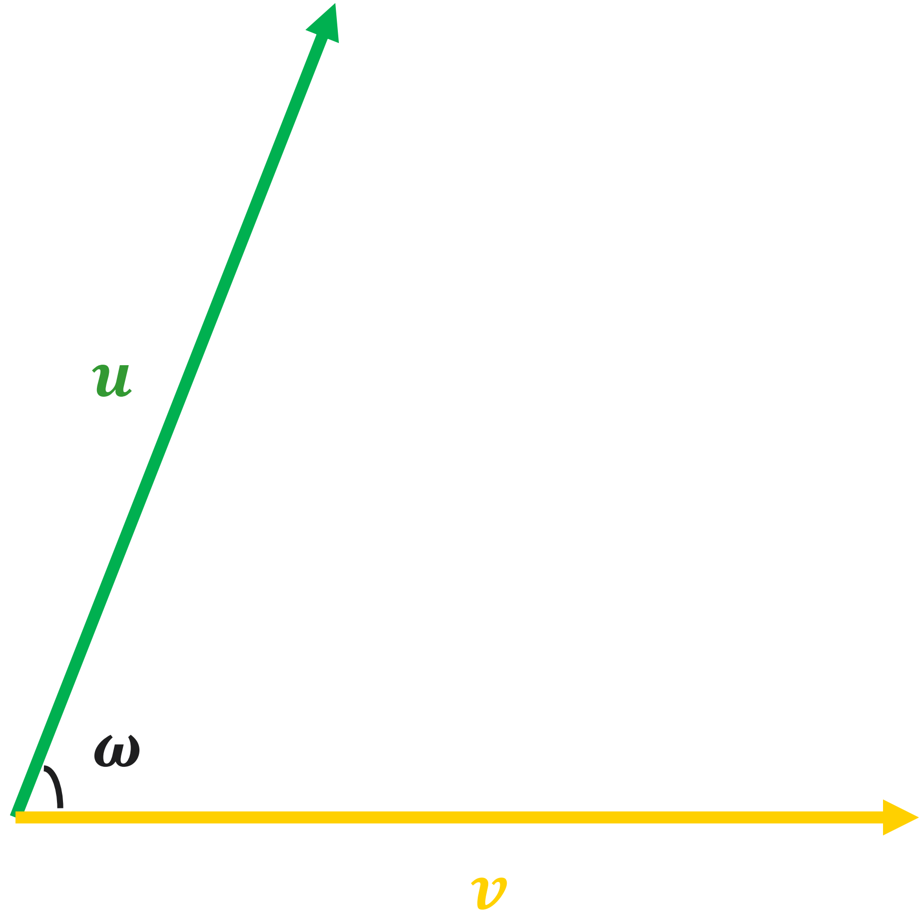
$$x = [1, 0], \qquad y = [0,1], \qquad (x, y) = 0, \qquad \|x\| = \|y\| = 1 \;\; \rightarrow$$

x and y are *orthonormal*.

girafe ai

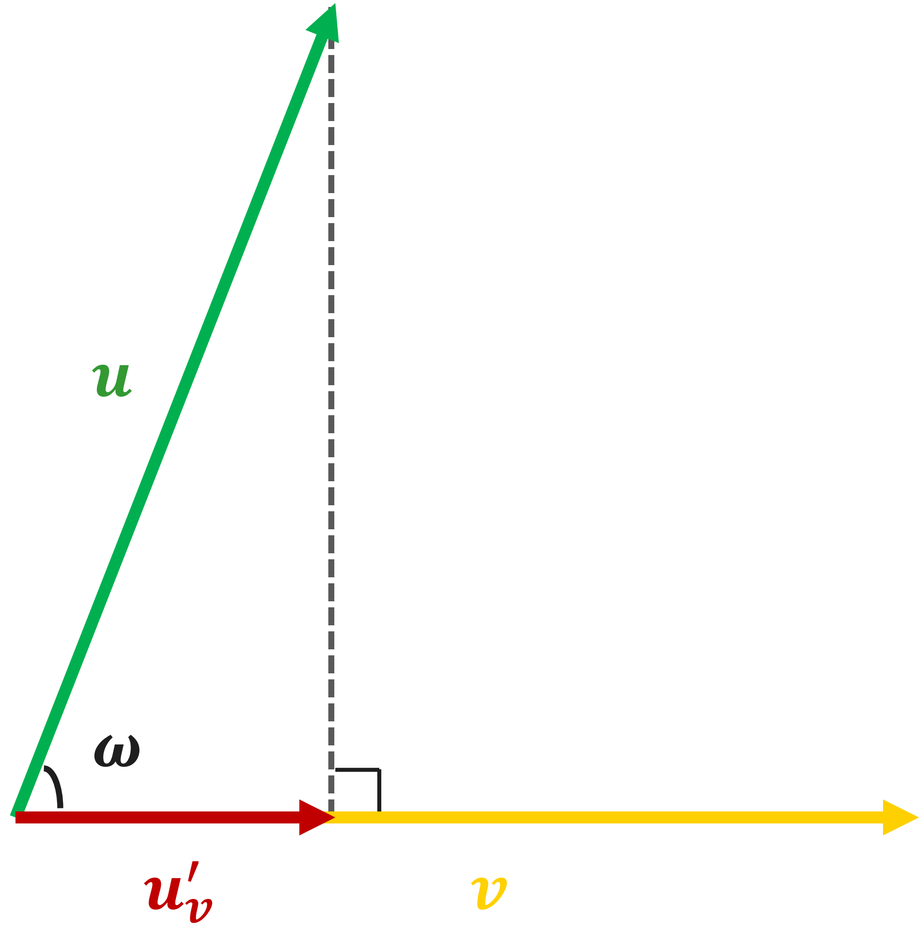
Orthogonal Projection

- Suppose we have two vectors u and v.



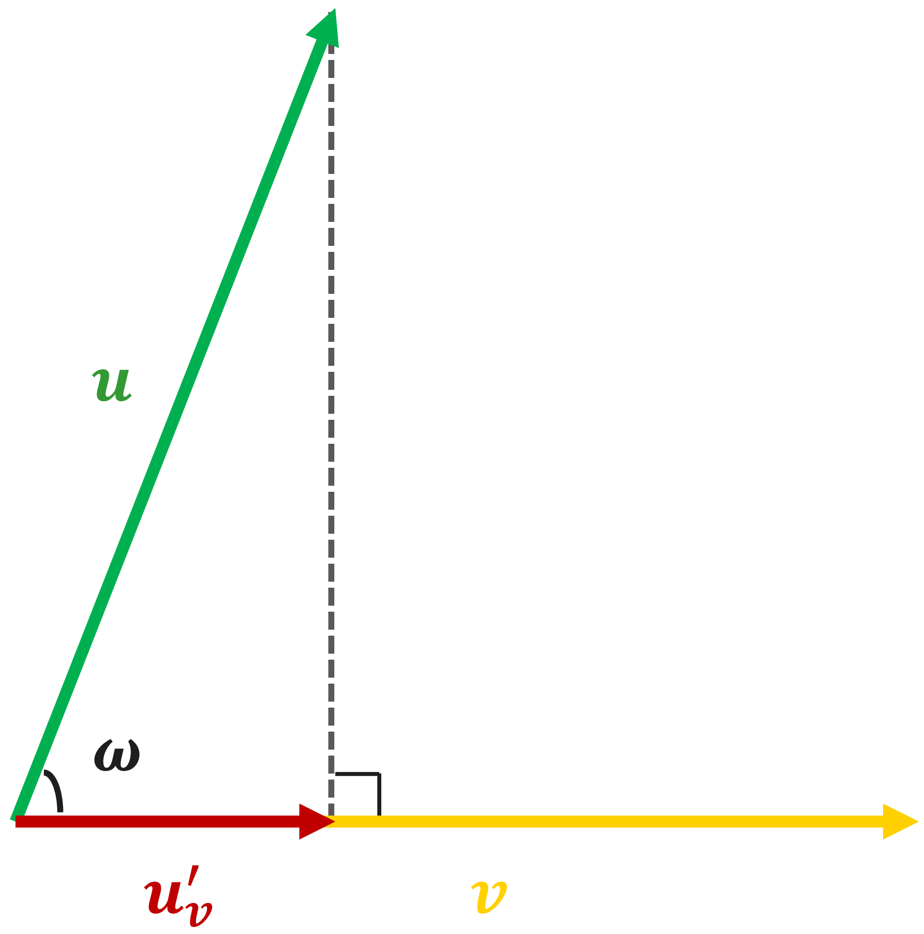
Orthogonal Projection

- Suppose we have two vectors u and v.
- u'_v - *orthogonal projection* of u on v.



Orthogonal Projection

- Suppose we have two vectors u and v.
- u'_v - *orthogonal projection* of u on v.
- What is the length of u'_v?



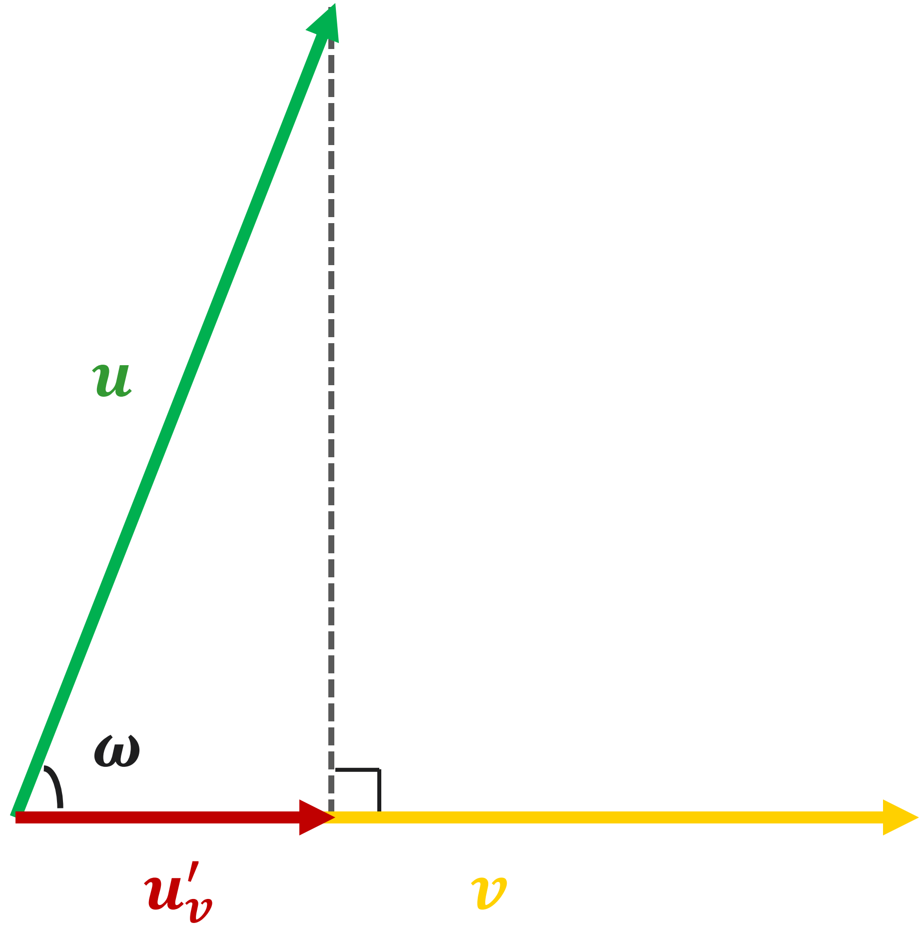
Orthogonal Projection

- Suppose we have two vectors u and v.
- u'_v - *orthogonal projection* of u on v.
- What is the length of u'_v?

If $\mathbf{0} \leq \boldsymbol{\omega} \leq \mathbf{90}$

$$(u, v) = \|u\|\|v\| \cos \boldsymbol{\omega} = \|u\|\|v\| \frac{\|\boldsymbol{u'_v}\|}{\|\boldsymbol{u}\|} =$$

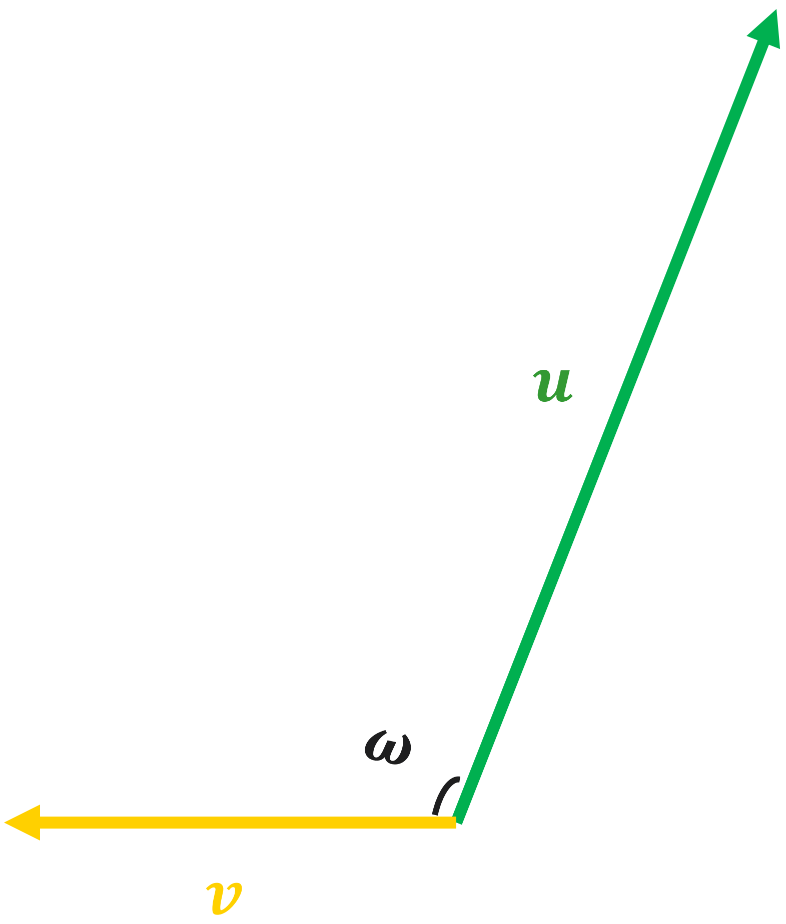
$$= \|u'_v\|\|v\|$$



Orthogonal Projection

- Suppose we have two vectors u and v.
- u'_v - *orthogonal projection* of u on v.
- What is the length of u'_v?

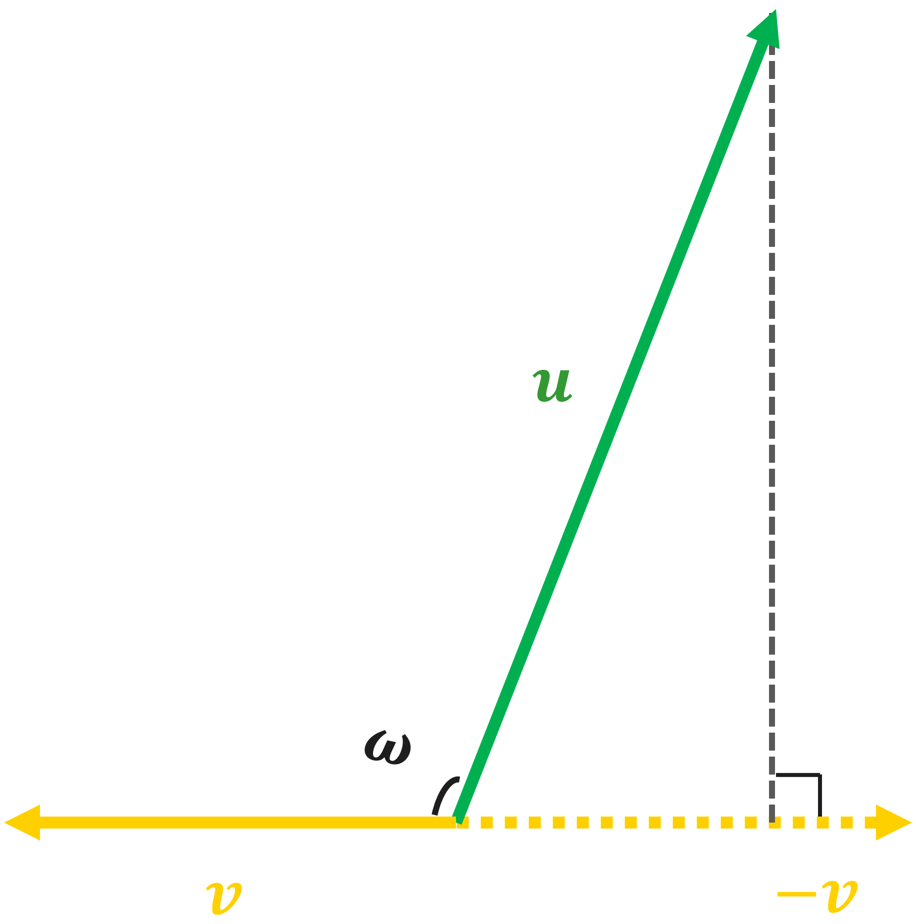
If $\mathbf{90 \leq \omega \leq 180}$



Orthogonal Projection

- Suppose we have two vectors u and v.
- u'_v - *orthogonal projection* of u on v.
- What is the length of u'_v?

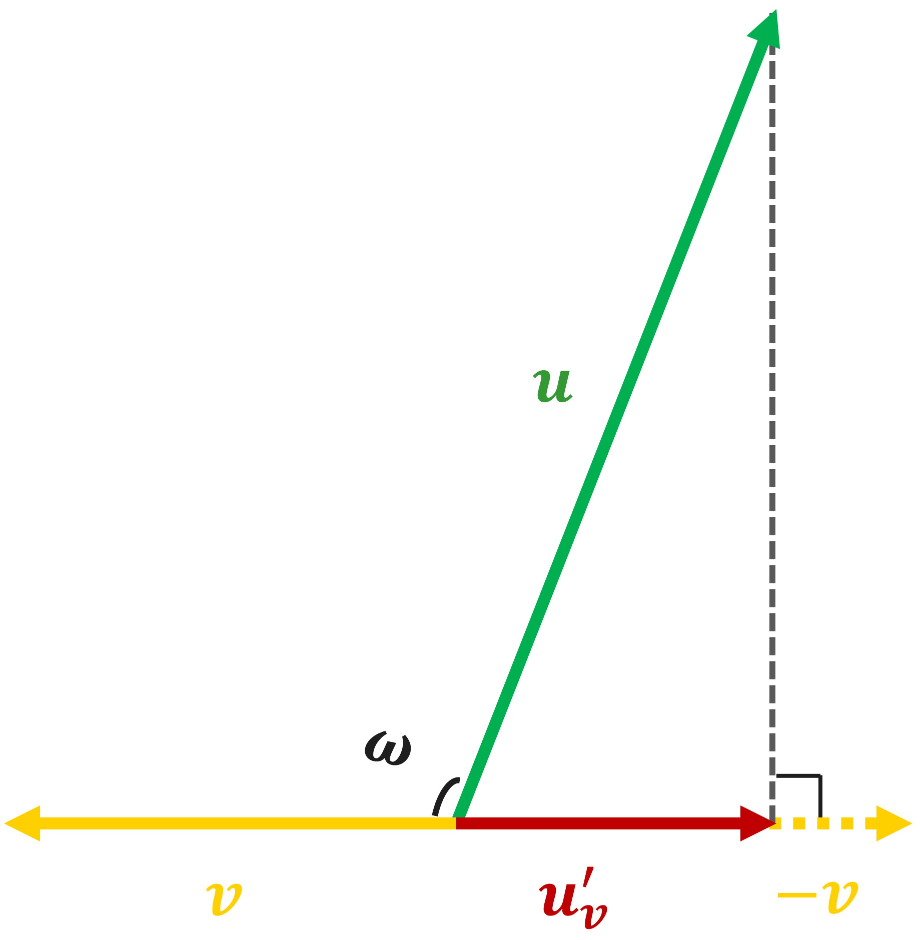
If $\mathbf{90 \leq \omega \leq 180}$



Orthogonal Projection

- Suppose we have two vectors u and v.
- u'_v - *orthogonal projection* of u on v.
- What is the length of u'_v?

If $\mathbf{90 \leq \omega \leq 180}$

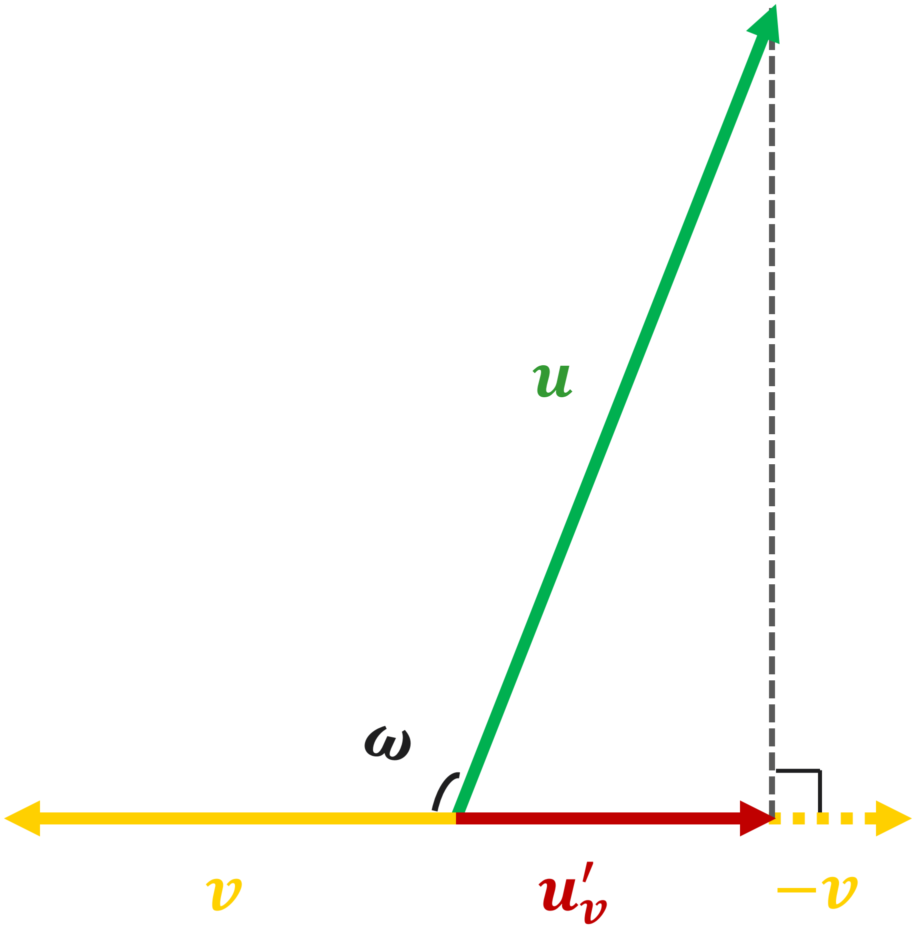


Orthogonal Projection

- Suppose we have two vectors u and v.
- u'_v - *orthogonal projection* of u on v.
- What is the length of u'_v?

If $\mathbf{90 \leq \omega \leq 180}$

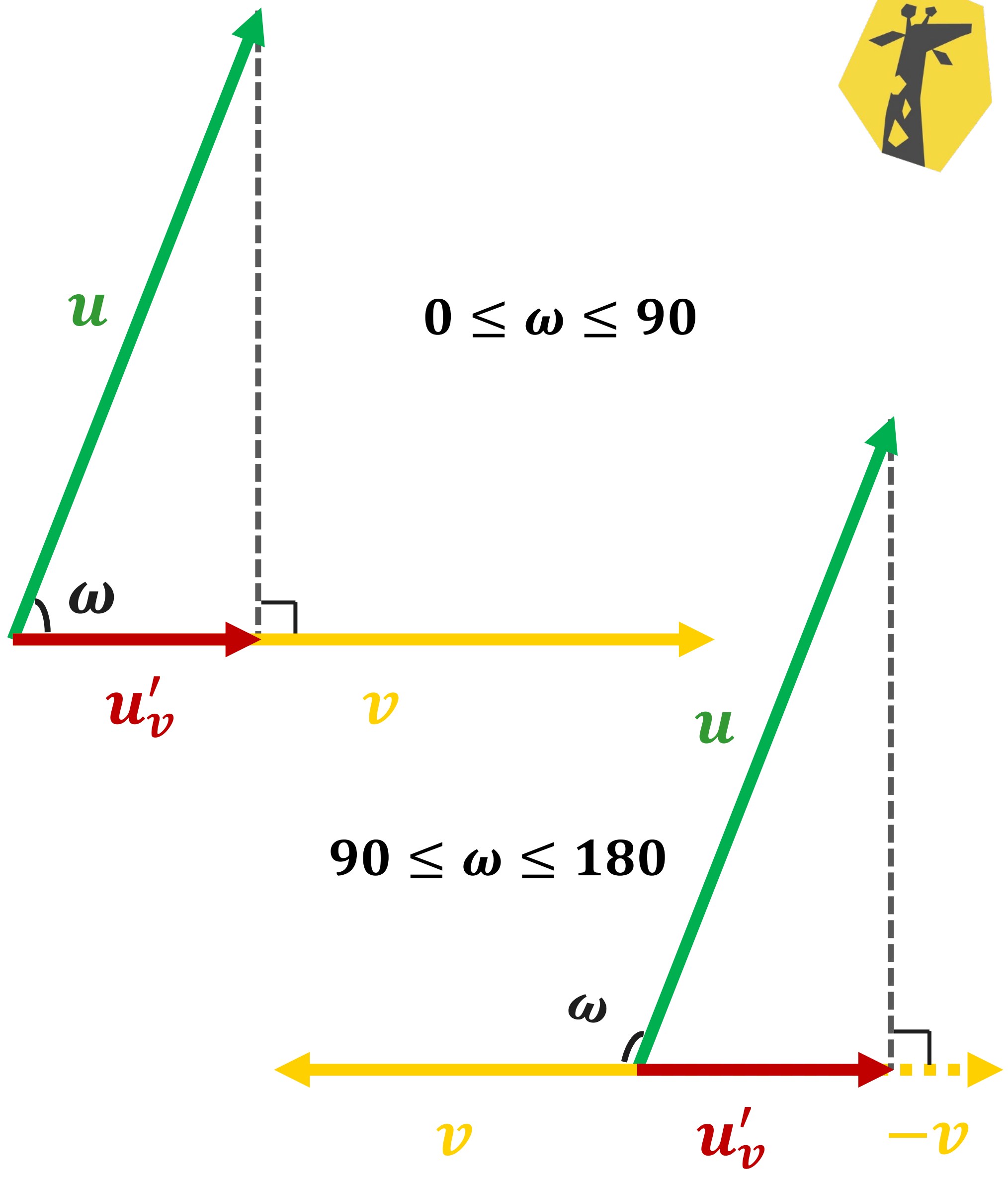
$$-(u, v) = \|u\|\|v\| \cos\omega = \|u\|\|v\| \frac{\|u'_v\|}{\|u\|} =$$
$$= \|u'_v\|\|v\|$$



Orthogonal Projection

- Suppose we have two vectors u and v.
- u'_v - *orthogonal projection* of u on v.
- What is the length of u'_v?

$$|(\boldsymbol{u},\boldsymbol{v})| = \|\boldsymbol{u}'_v\|\|\boldsymbol{v}\| \quad \leftrightarrow \quad \|\boldsymbol{u}'_v\| = \frac{|(\boldsymbol{u},\boldsymbol{v})|}{\|\boldsymbol{v}\|}$$

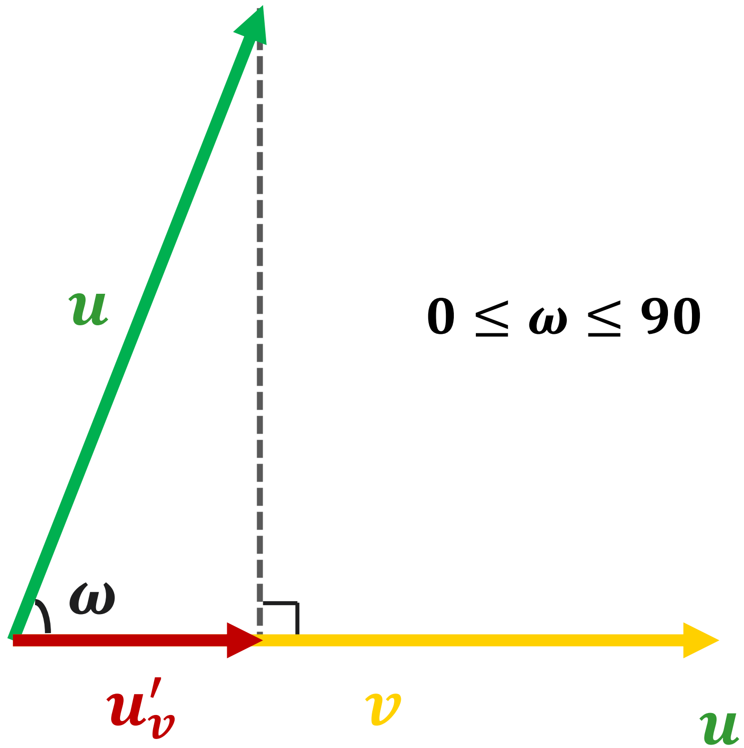


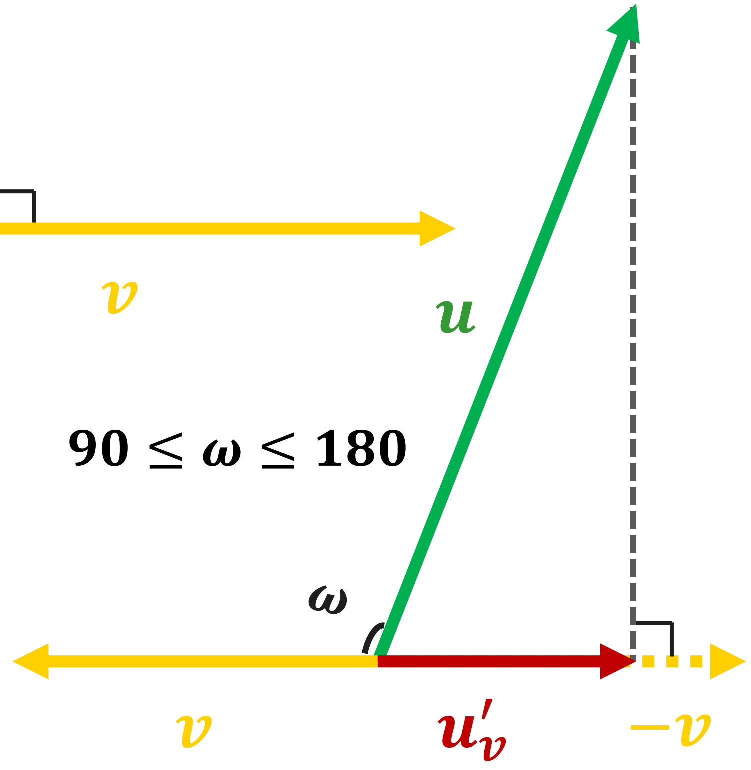
Orthogonal Projection

- Suppose we have two vectors u and v.
- u'_v - *orthogonal projection* of u on v.
- What is the length of u'_v?

$$|(u, v)| = \|u'_v\| \|v\| \quad \leftrightarrow \quad \|u'_v\| = \frac{|(u, v)|}{\|v\|}$$

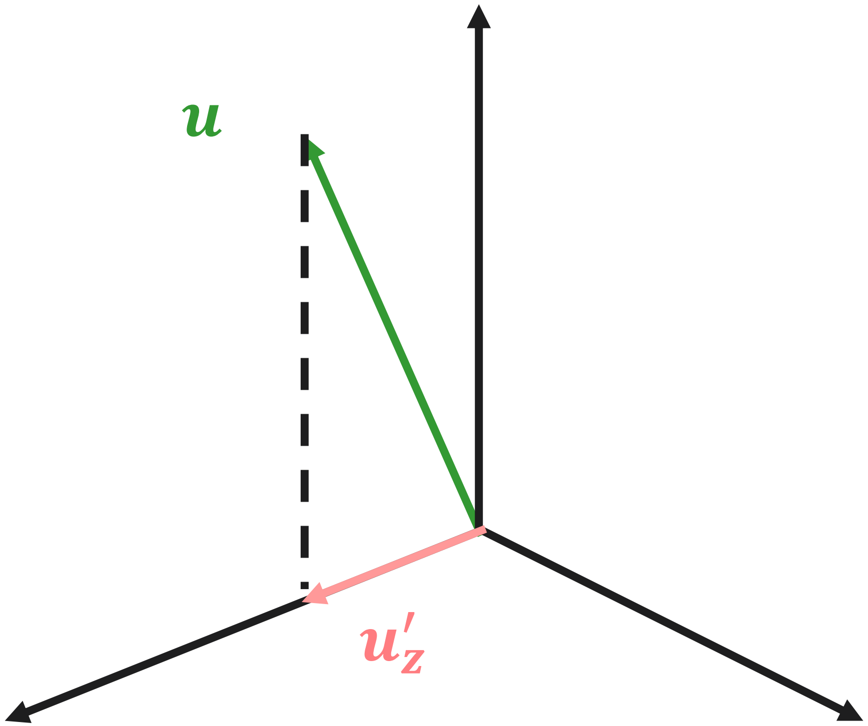
$$u'_v = \frac{(u, v)}{(v, v)} v.$$





Orthogonal Projection: Example

- What's projection of $u = [1, 3, 2]$ on $z = [0, 0, 1]$?

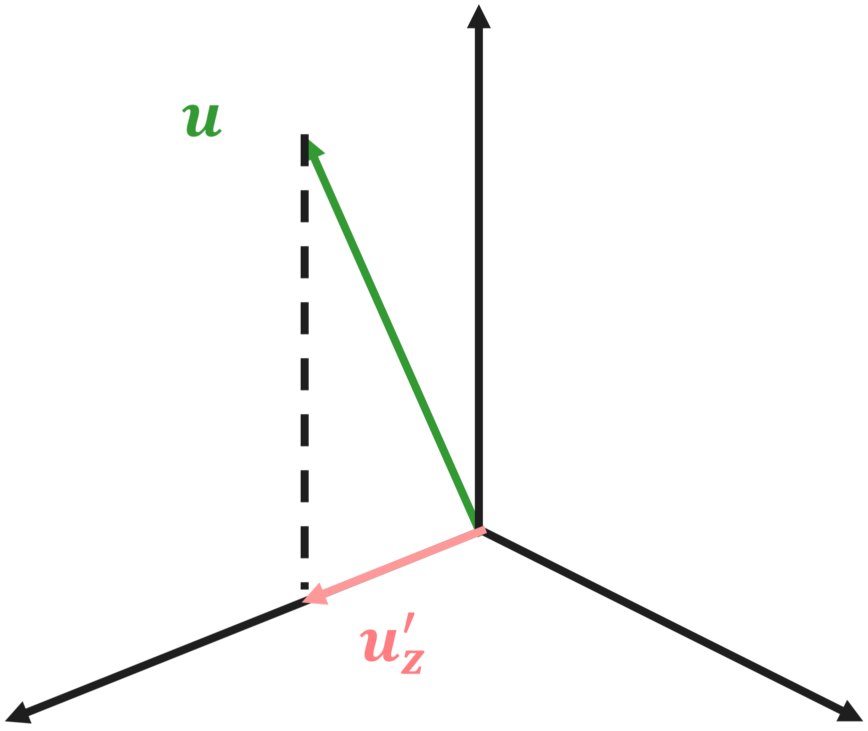


Orthogonal Projection: Example

- What's projection of $u = [1, 3, 2]$ on $z = [0, 0, 1]$?

$$u'_z = \frac{(u, z)}{(z, z)} z = 2z = [0, 0, 2]$$

(Projection on the axis = drop other coordinates)



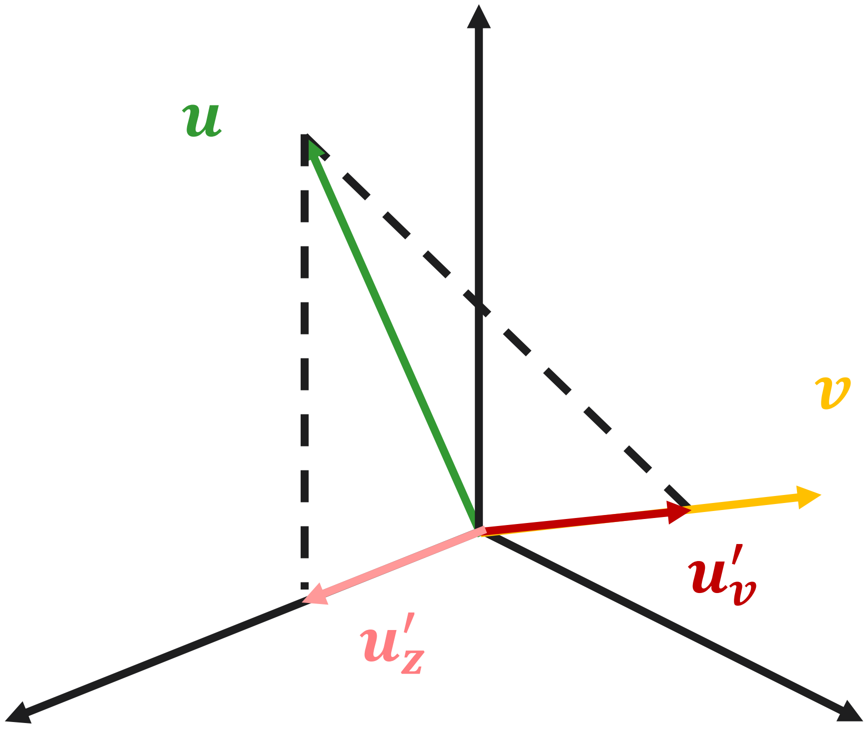
Orthogonal Projection: Example

- What's projection of $u = [1, 3, 2]$ on $z = [0, 0, 1]$?

$$u'_z = \frac{(u, z)}{(z, z)} z = 2z = [0, 0, 2]$$

(Projection on the axis = drop other coordinates)

- What's projection of $u = [1, 3, 2]$ on $\mathrm{v} = [4, 1, 3]$?



girafe ai

Orthogonal Projection: Example

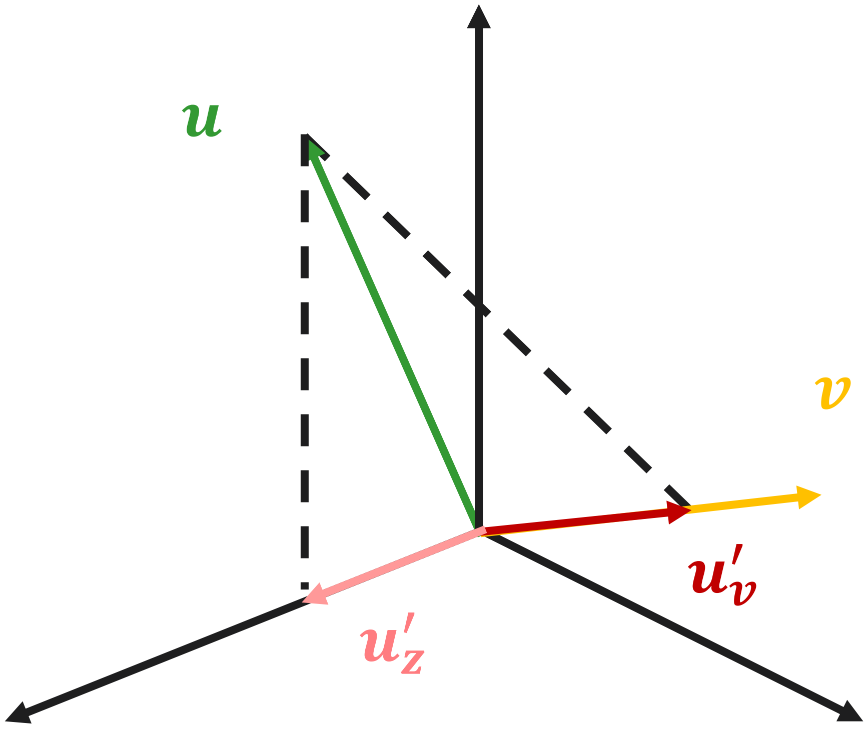
- What's projection of $u = [1, 3, 2]$ on $z = [0, 0, 1]$?

$$u'_z = \frac{(u, z)}{(z, z)} z = 2z = [0, 0, 2]$$

(Projection on the axis = drop other coordinates)

- What's projection of $u = [1, 3, 2]$ on $\mathrm{v} = [4, 1, 3]$?

$$u'_v = \frac{(u, v)}{(v, v)} v = \frac{4 + 3 + 6}{16 + 1 + 9} v = \frac{1}{2} v = [2, 0.5, 1.5].$$



Hyperplanes

- A hyperplane is described by equation

$$w_1 x_1 + w_2 x_2 + \cdots + w_n x_n + b = 0$$

where at least one $w_i \neq 0$.

- A more compact notation:

$$(w, x) + b = 0, \qquad w = (w_1, \dots, w_n)$$

Hyperplanes

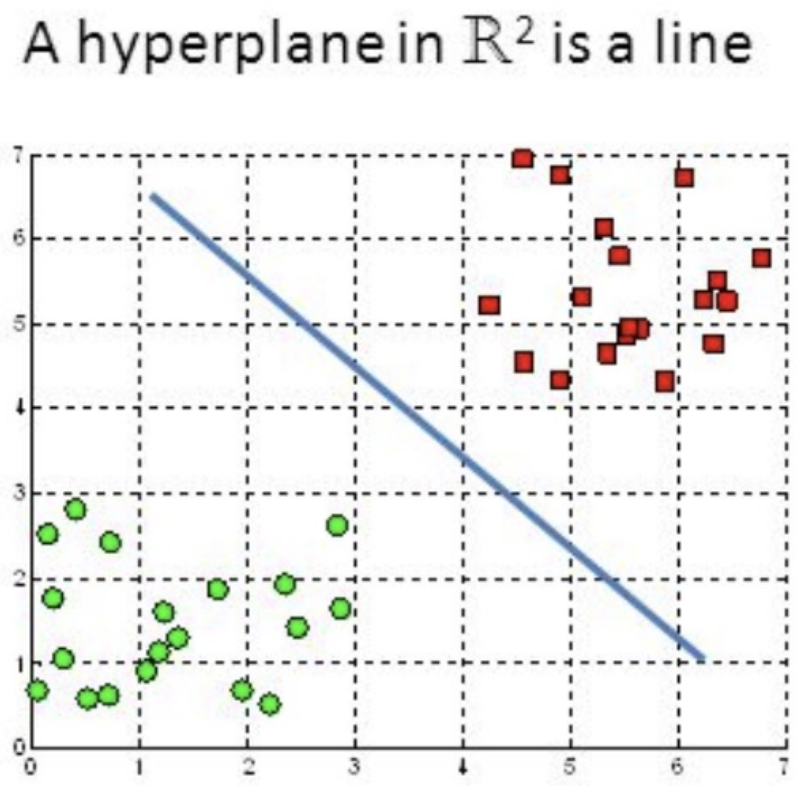
Hyperplanes

- A hyperplane in $\mathbb{R}^n$ is described by equation

$$w_1 x_1 + w_2 x_2 + \cdots + w_n x_n + b = 0$$

where at least one $w_i \neq 0$.

Hyperplanes



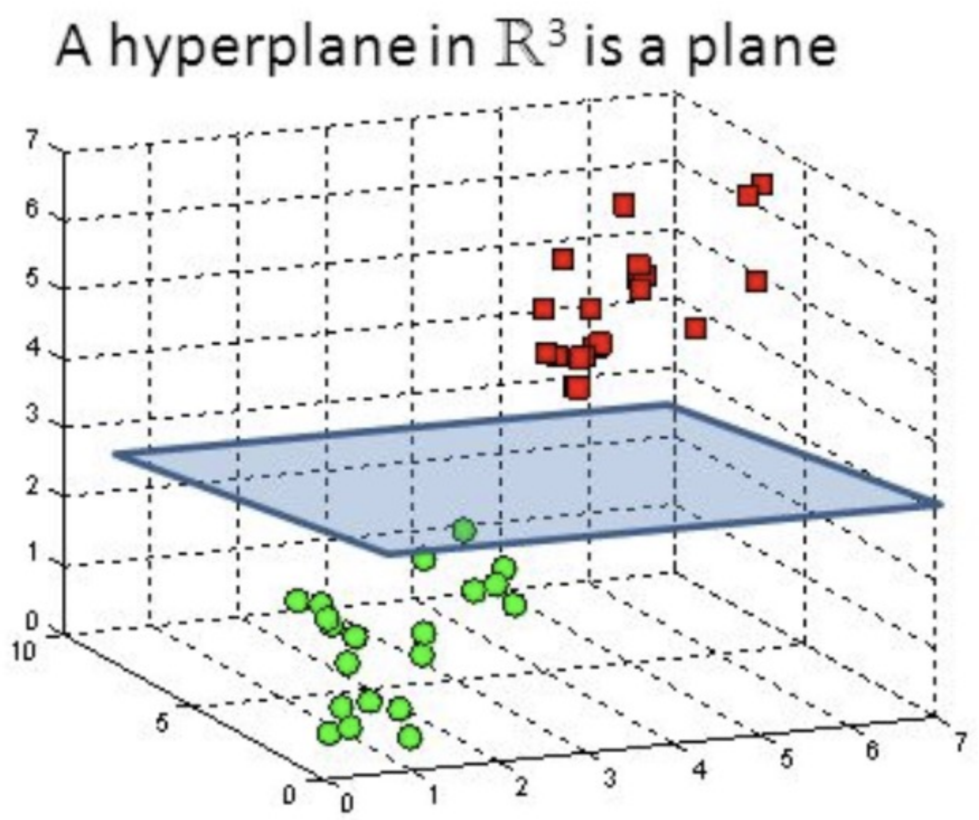


Image source: https://deepai.org/machine-learning-glossary-and-terms/hyperplane

girafe ai

Hyperplanes

- A hyperplane in $\mathbb{R}^n$ is described by equation

$$w_1 x_1 + w_2 x_2 + \cdots + w_n x_n + b = 0$$

where at least one $w_i \neq 0$.

- A more compact notation:

$$(w, x) + b = 0, \qquad w = (w_1, \dots, w_n)$$

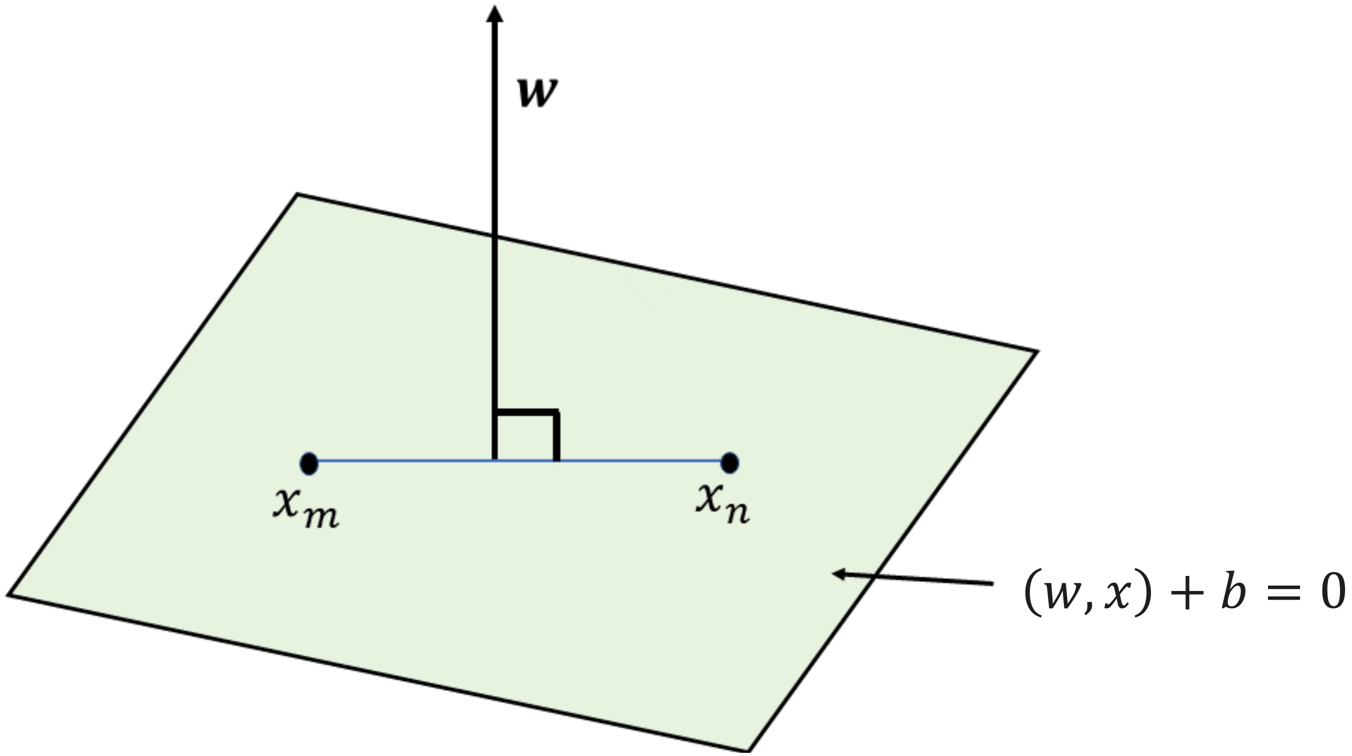
Normal to a Hyperplane

- Consider a hyperplane $(w, x) + b = 0$.

- Vector $w = (w_1, \dots, w_n)$ defines the hyperplane.

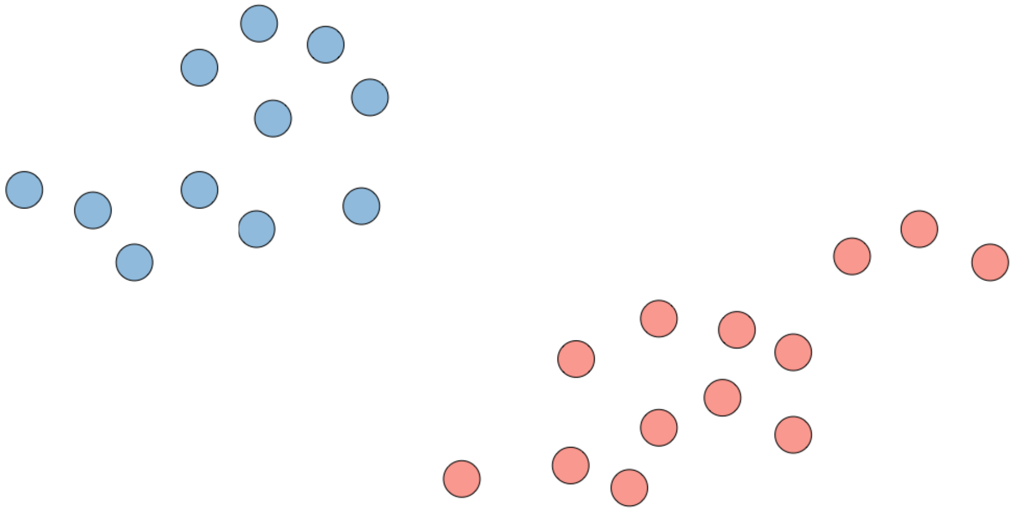
Normal to a Hyperplane

- Consider a hyperplane $(w, x) + b = 0$.
- Vector $w = (w_1, \dots, w_n)$ defines the hyperplane.
- w is a *normal vector* to this hyperplane: it's orthogonal to every vector on it.



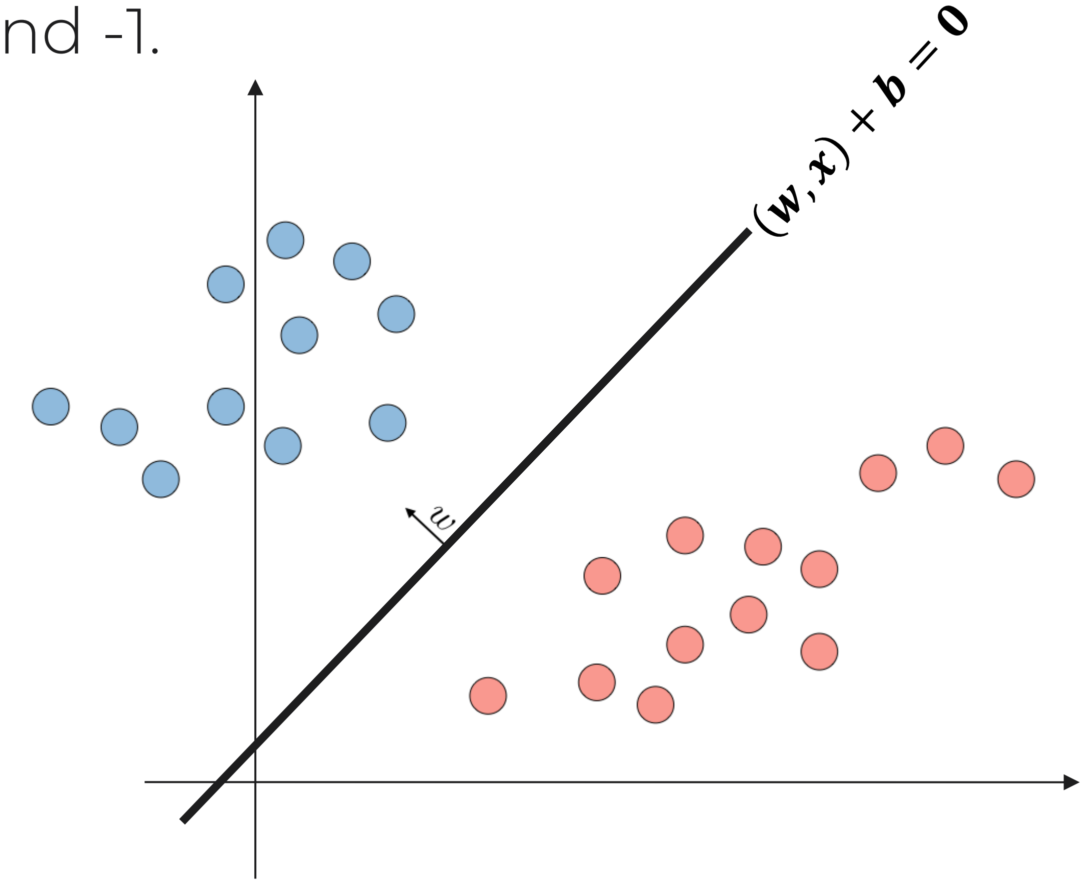
ML Example: Linear Classifier

- Objects = 2D vectors
- Binary classification: classes +1 and -1.



ML Example: Linear Classifier

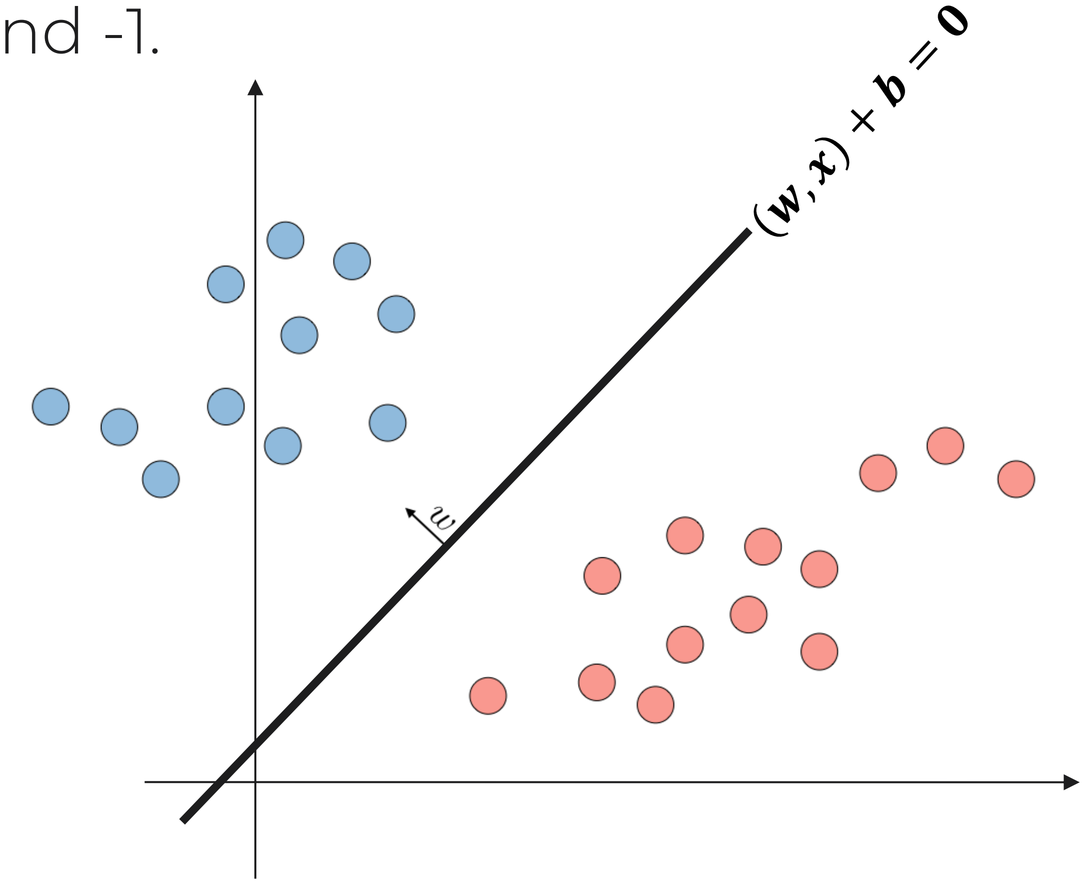
- Objects = 2D vectors
- Binary classification: classes +1 and -1.
- Linear classifier: separating hyperplane $(w, x) + b = 0$
  - objects "above" are class +1
  - objects "below" are class -1.



ML Example: Linear Classifier

- Objects = 2D vectors
- Binary classification: classes +1 and -1.
- Linear classifier: separating hyperplane $(w, x) + b = 0$
  - objects "above" are class +1
  - objects "below" are class -1.

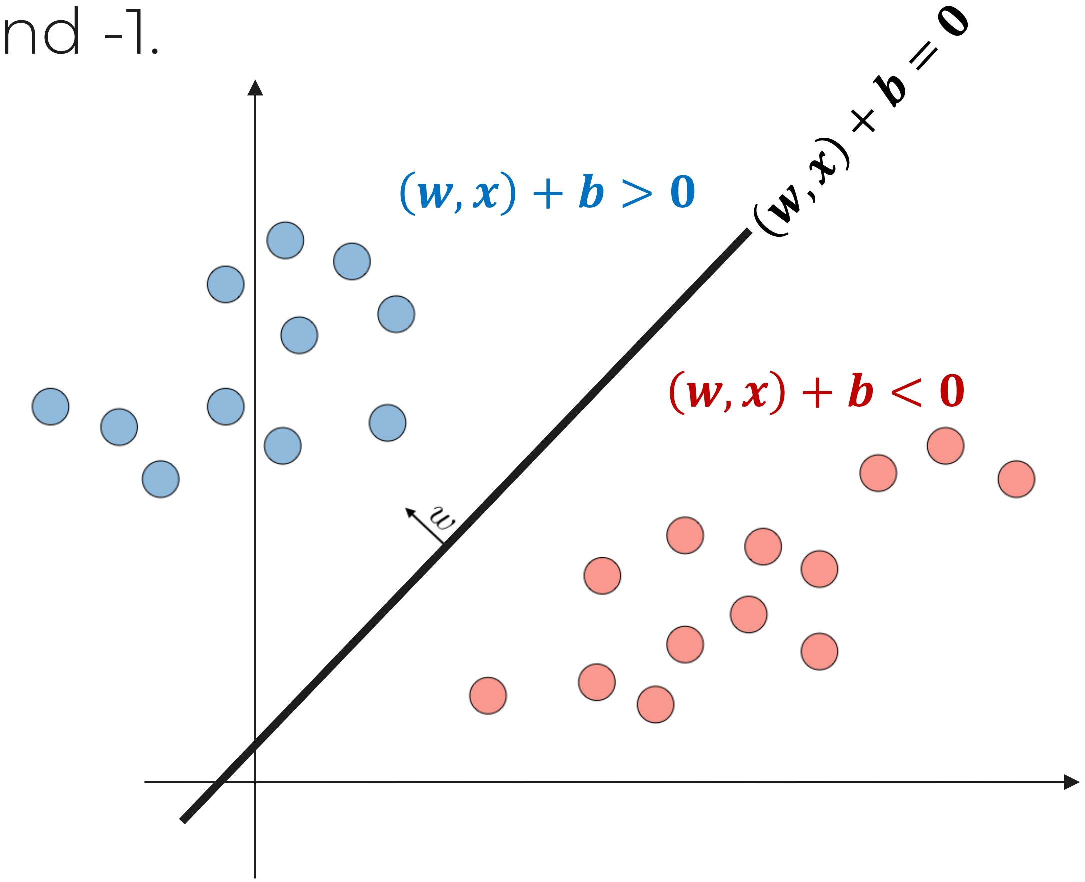
- How can we formalize this?



ML Example: Linear Classifier

- Objects = 2D vectors
- Binary classification: classes +1 and -1.
- Linear classifier: separating hyperplane $(w, x) + b = 0$
    - objects "above" are class +1
    - objects "below" are class -1.

- How can we formalize this?
    - objects "above": $(w, x) + b > 0$
    - objects "below": $(w, x) + b < 0$



girafe ai

To sum up

- Vectors
    - Vector spaces
    - Inner products
    - Lengths
    - Distances
    - Angles
- Analytic Geometry
    - Projections
    - Hyperplanes
    - Normal vector

Next Time

- More on vector spaces.