Intro to Math and Python

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

based on [slides by Evgenia Korneva]

About this course

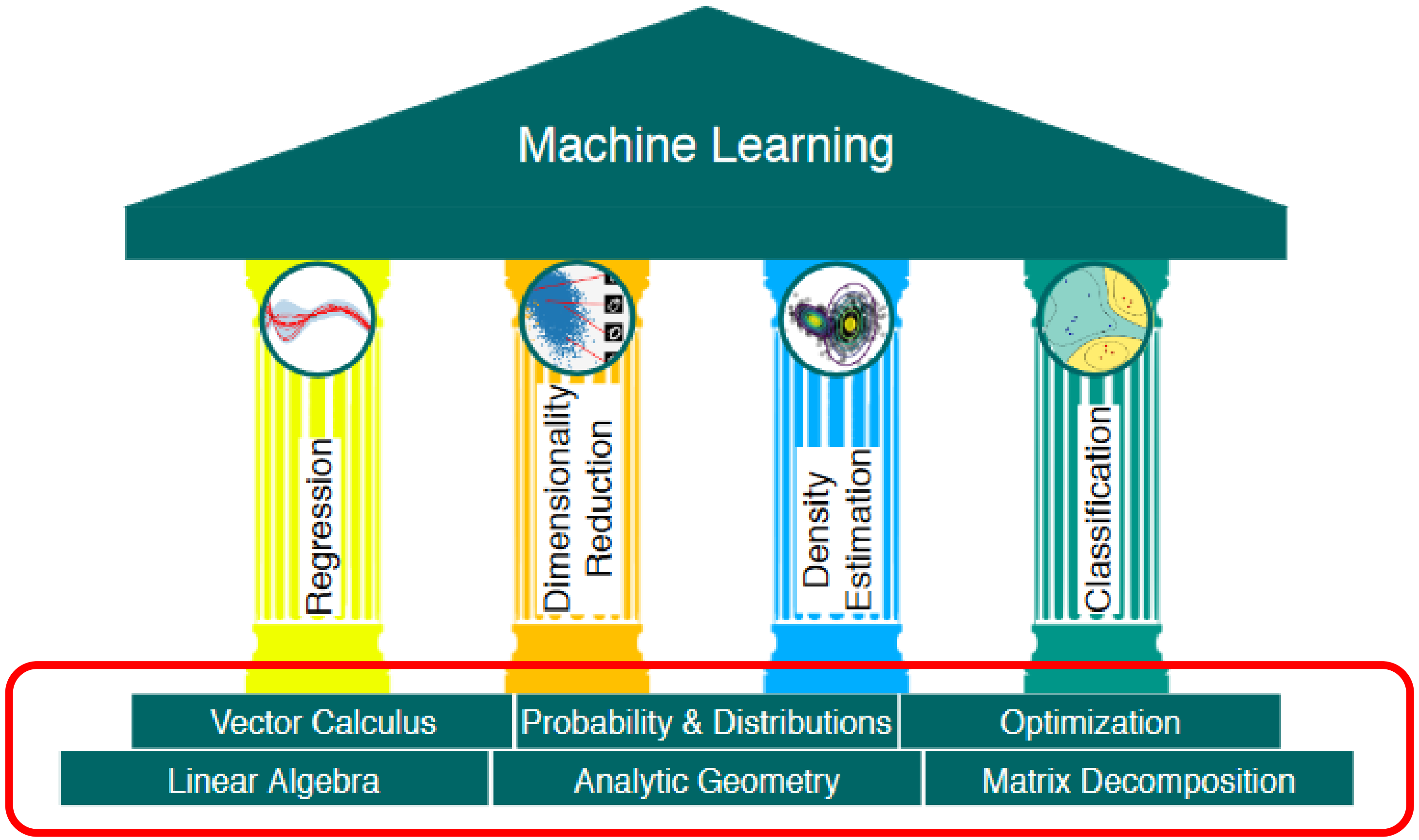


Image source: Mathematics for Machine Learning, p. 14 (https://mml-book.github.io/book/mml-book.pdf)

About this course

This week:

1. Linear algebra
2. Calculus
3. Probability theory

Prerequisites:

- basic knowledge of math;
- some Python.

About me

Iurii Efimov

yuri.efimov@phystech.edu

efimov-iurii

- PhD researcher
- Lecturer
- Artec3D DS Team

Linear Algebra: the Basics

Linear Algebra: Core Objects

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

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

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

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

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

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

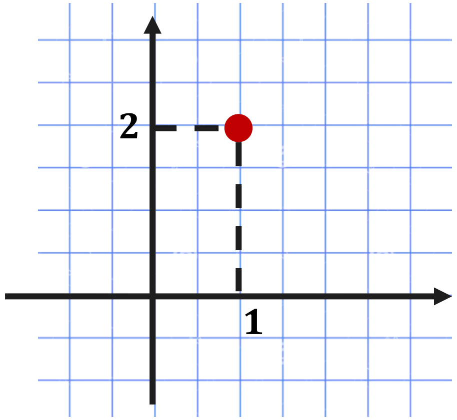
Example: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$, $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \in \mathbb{R}^{2\times 3}$

What are Vectors?

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

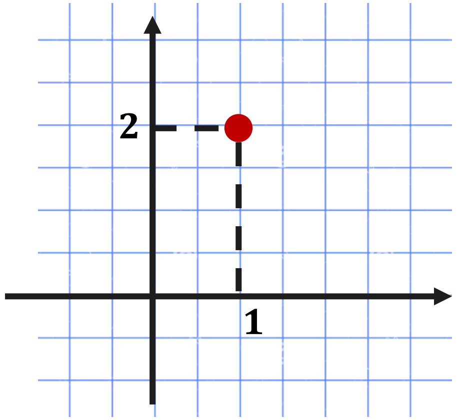
What are Vectors?

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

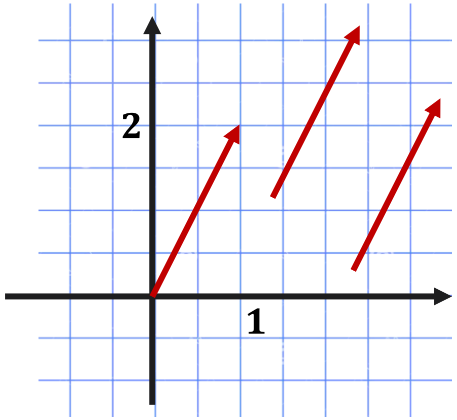


What are Vectors?

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



- Direction + length



Vector Spaces

Vector Space: Definition

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

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

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.		$\alpha(x + y) = \alpha x + \alpha y$

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:

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

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:

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

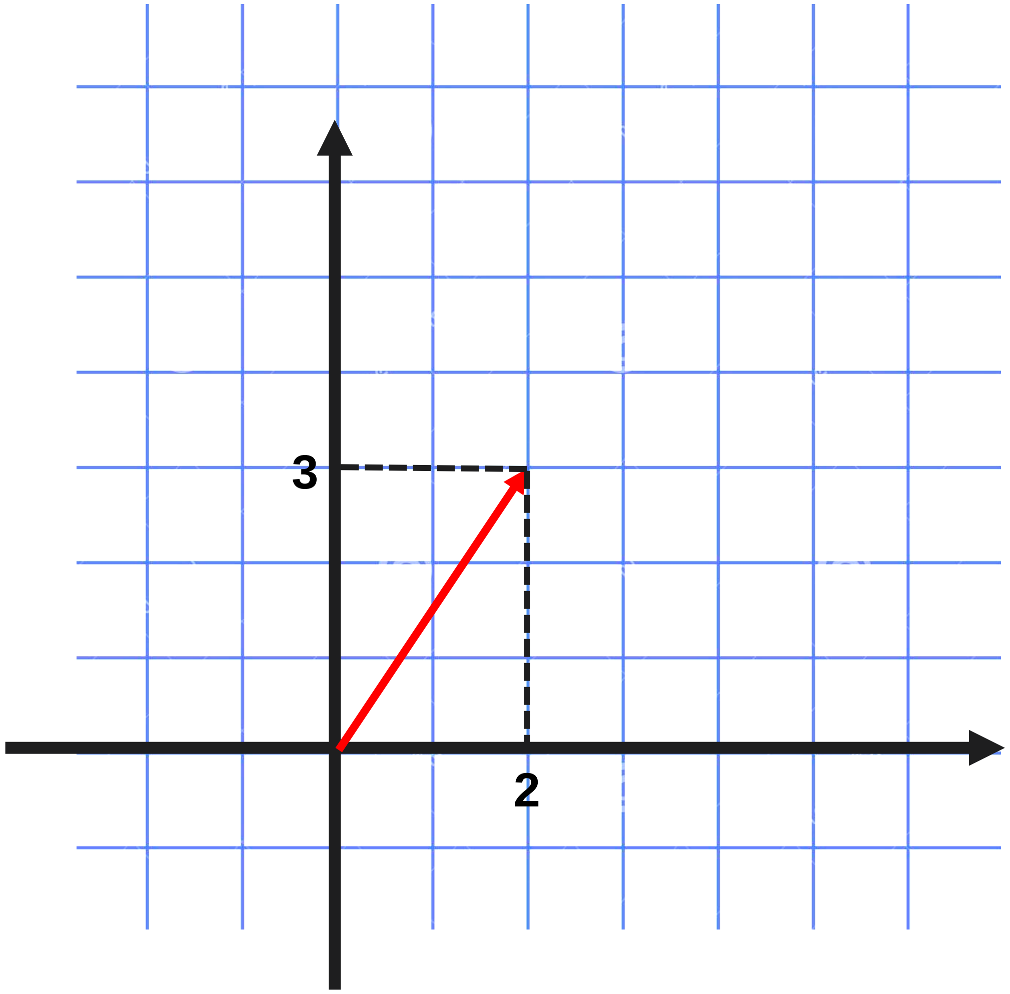
Sum:

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

Vector Operations: Geometrical Interpretation

Vectors: Geometrical Interpretation

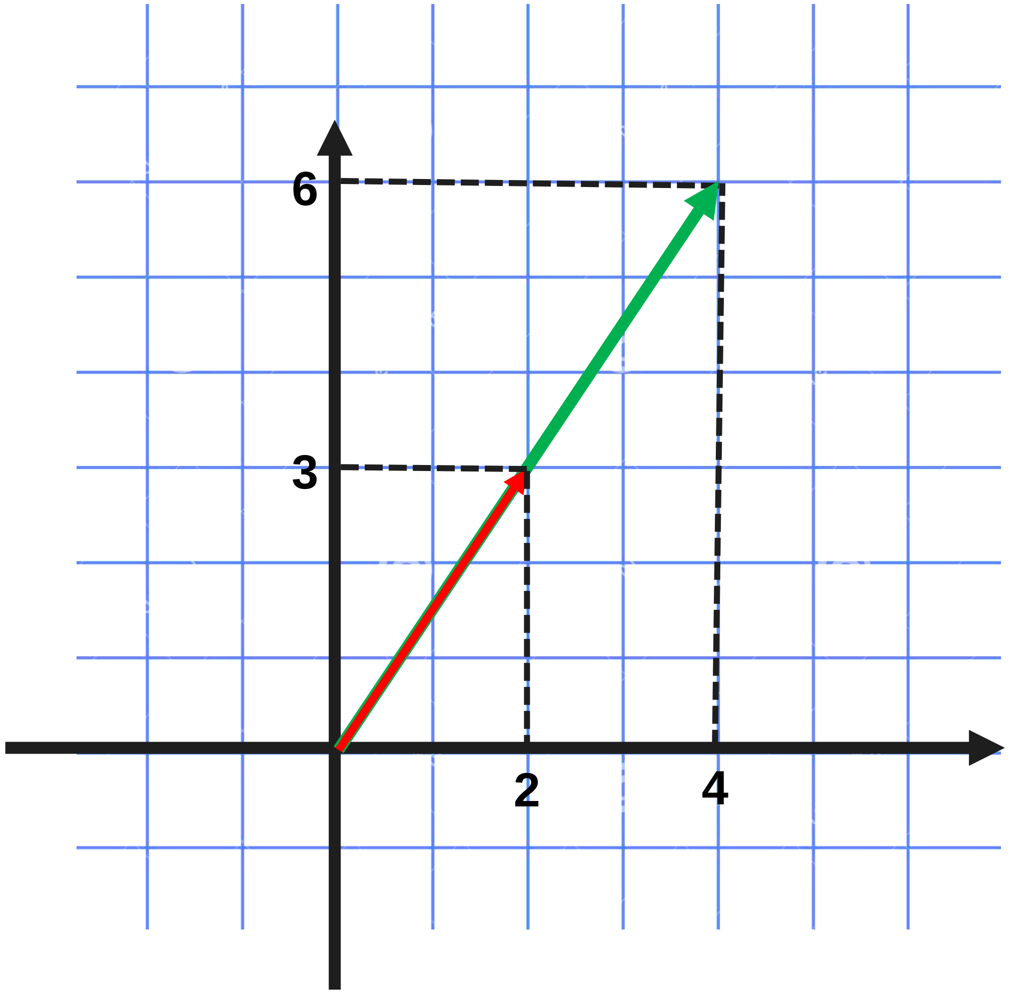
$$\vec{a} = [2, 3]$$



Vectors: Geometrical Interpretation

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

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

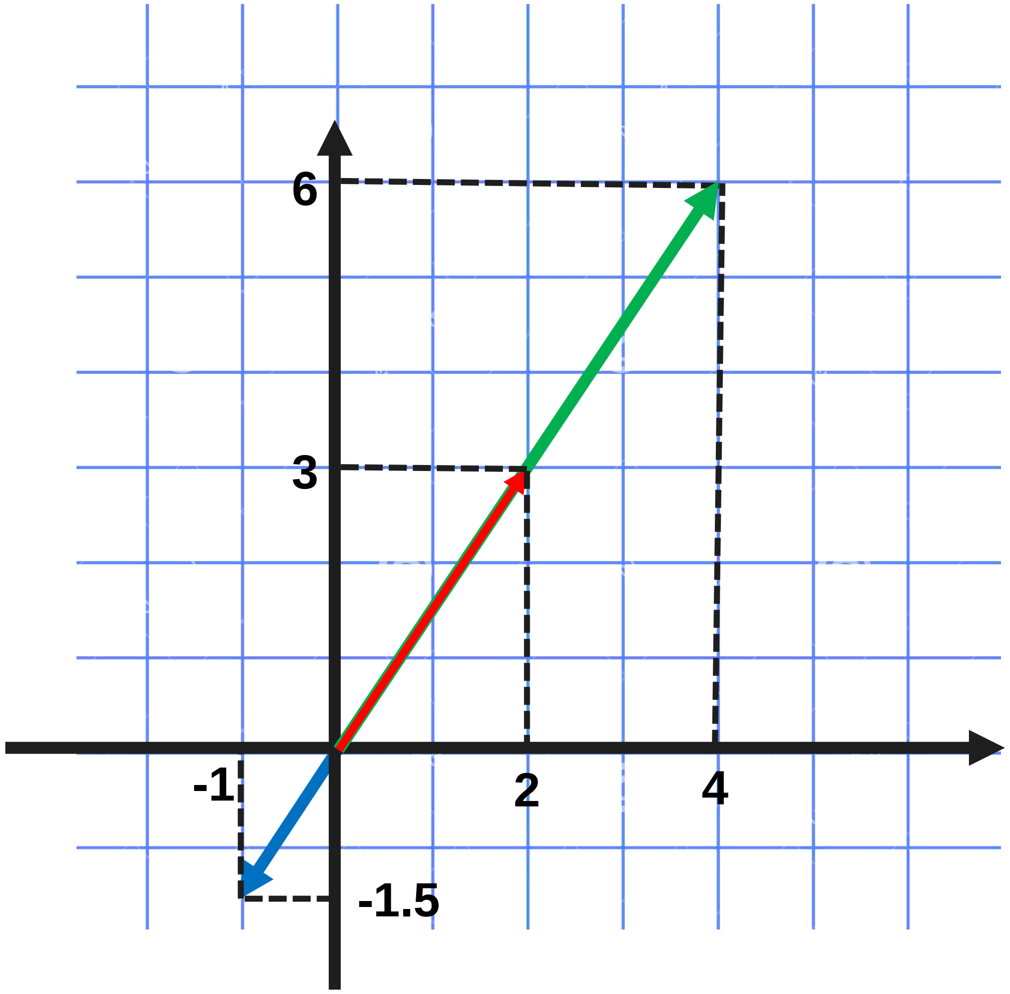


Vectors: Geometrical Interpretation

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

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

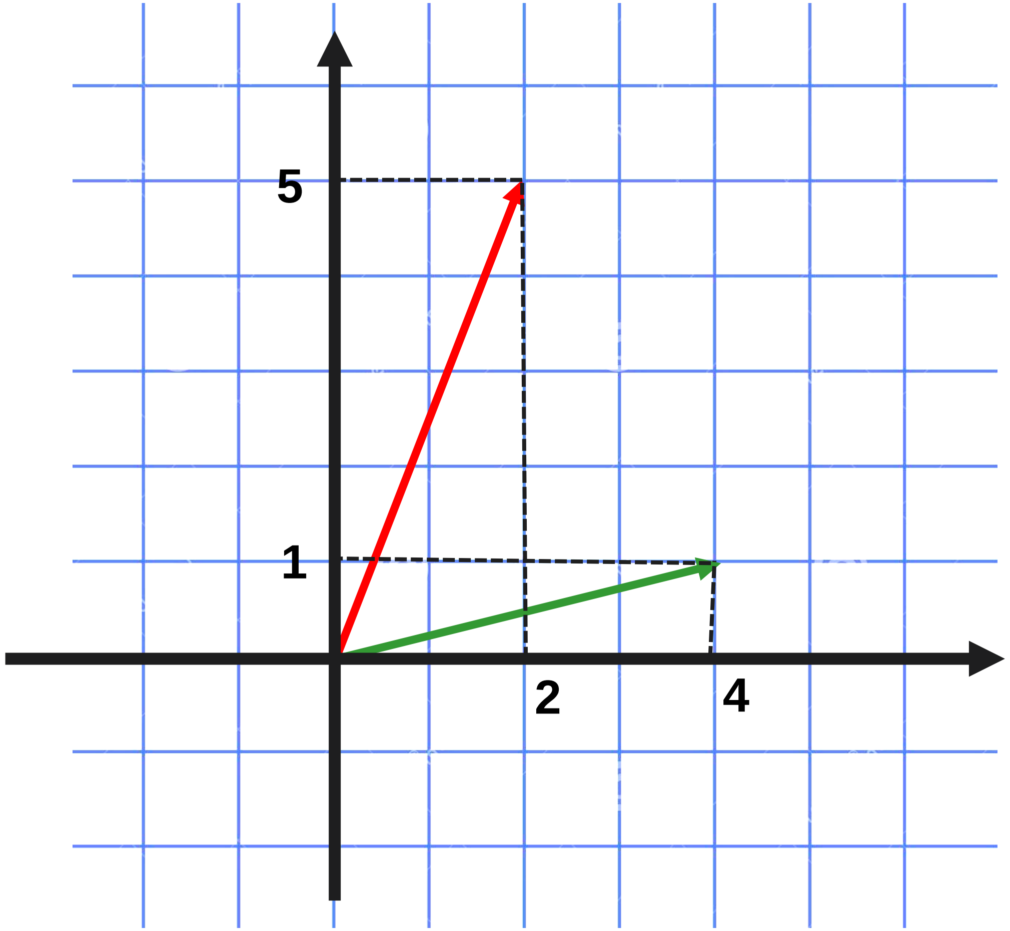
$-0.5\vec{a} = [-1, -1.5]$



Vectors: Geometrical Interpretation

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

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

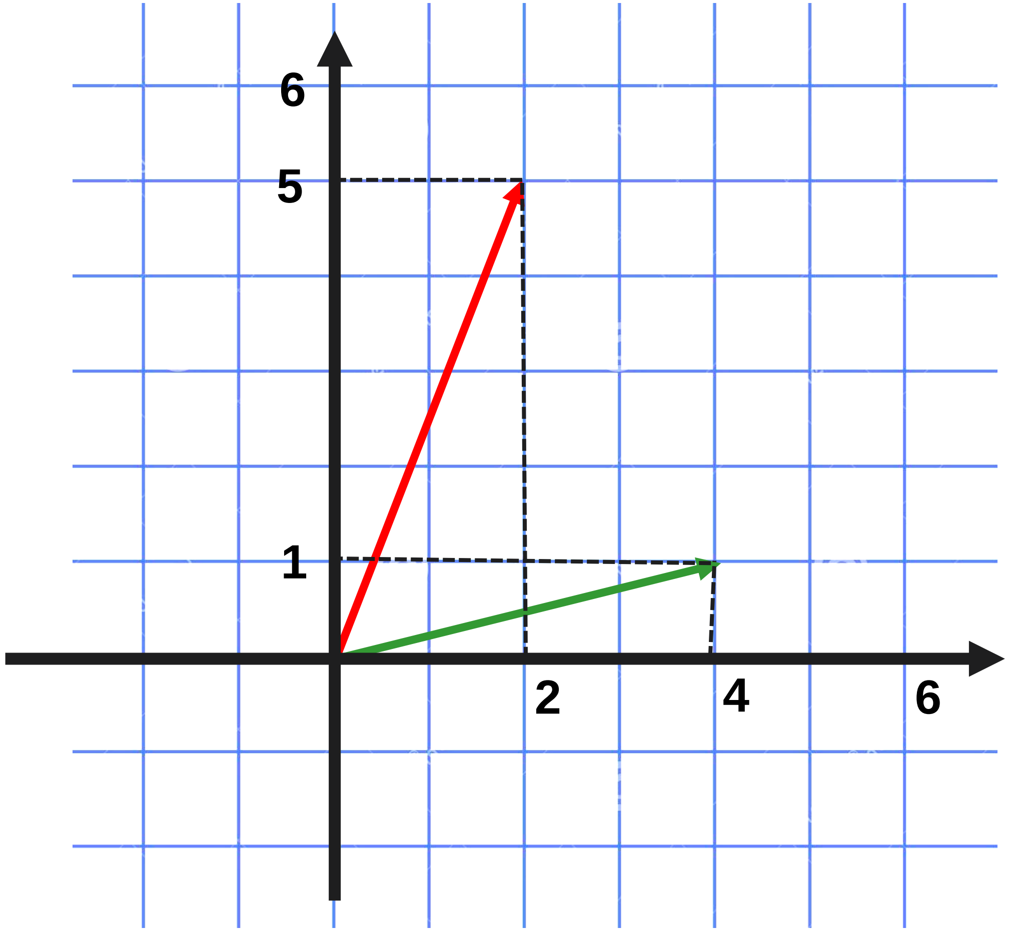


Vectors: Geometrical Interpretation

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

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

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

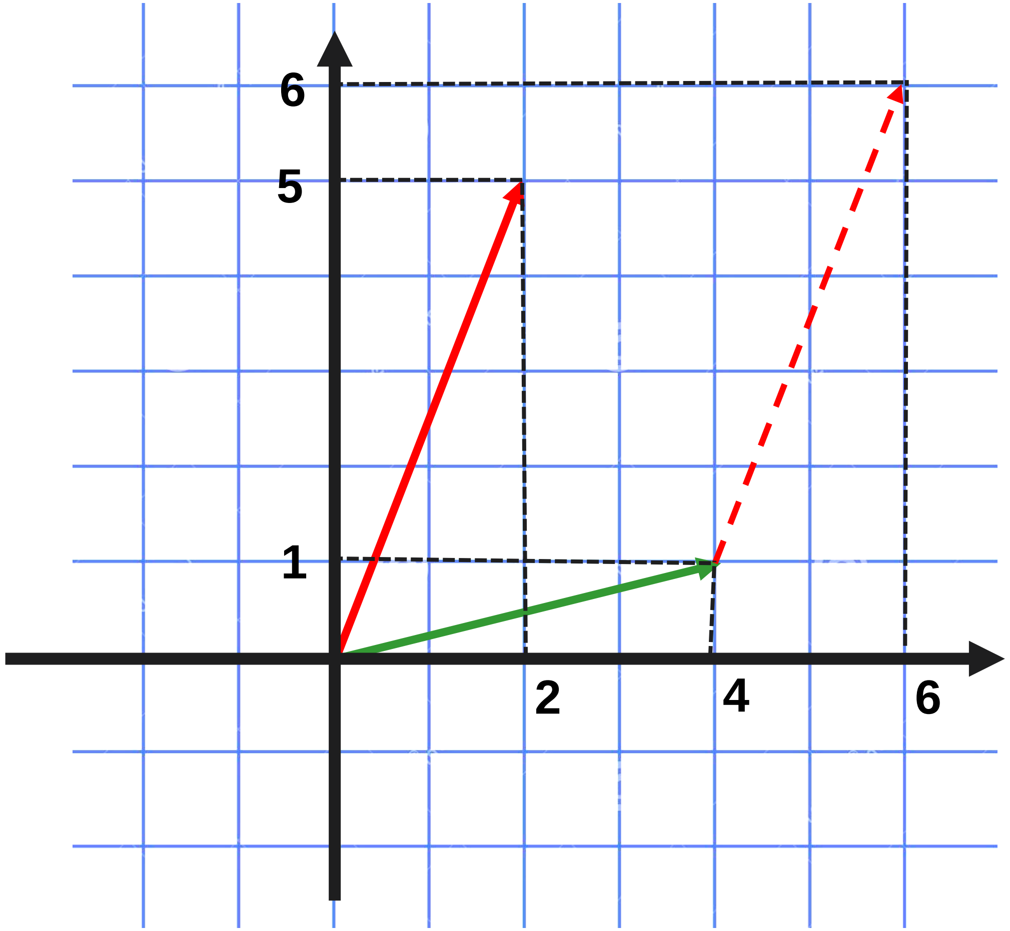


Vectors: Geometrical Interpretation

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

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

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

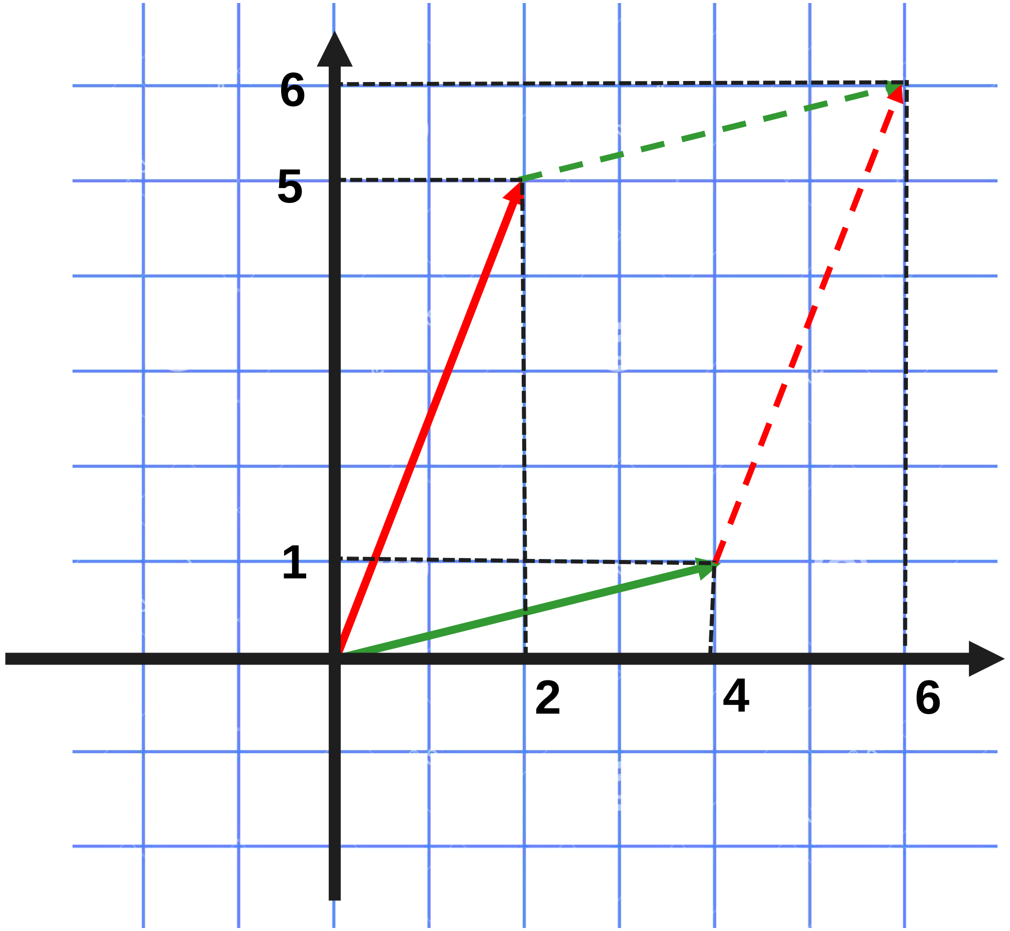


Vectors: Geometrical Interpretation

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

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

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

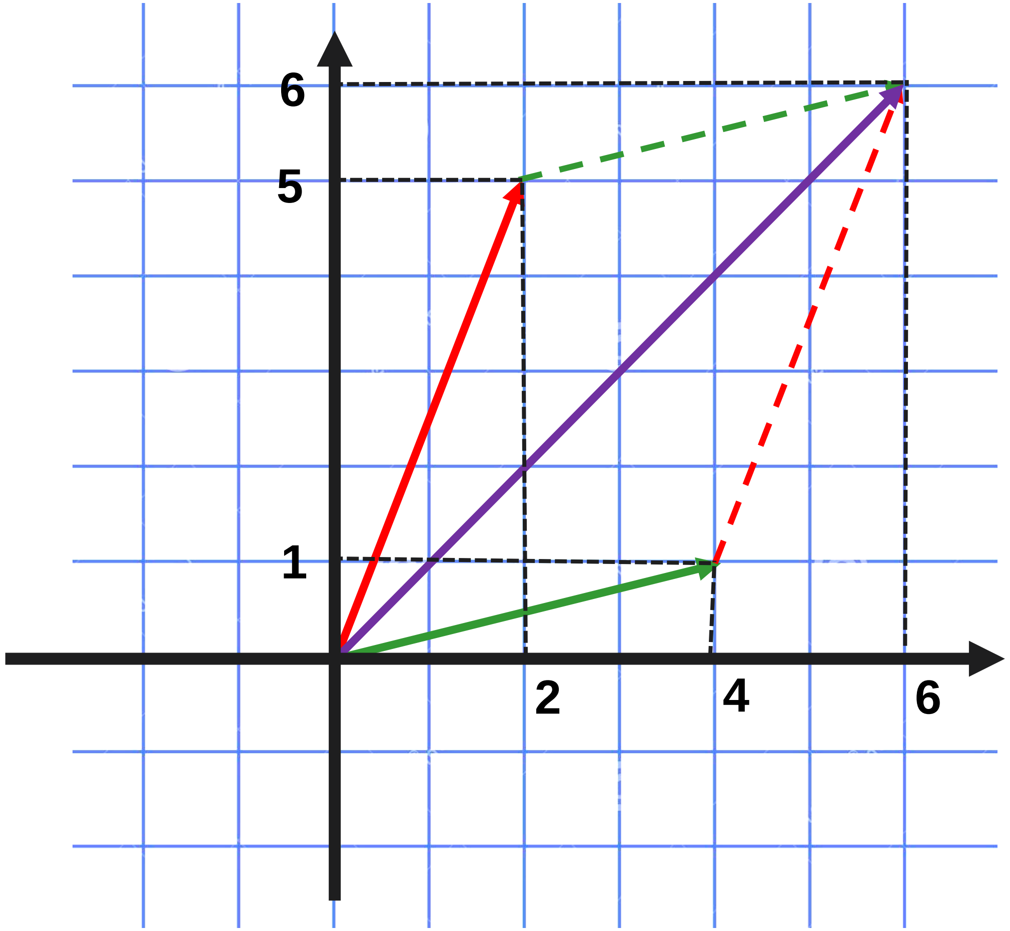


Vectors: Geometrical Interpretation

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

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

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



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)

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

girafe ai

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:

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

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

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

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 $(\,\cdot,\cdot\,)$ is called a *Euclidian vector space*.

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 $\|\cdot\|: \mathbb{V} \rightarrow \mathbb{R}$ 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 $\|\cdot\|: \mathbb{V} \rightarrow \mathbb{R}$ 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 $\|\cdot\|: \mathbb{V} \rightarrow \mathbb{R}$ 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* $\|x\| \in \mathbb{R}$.

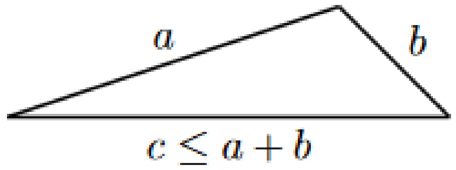
- A norm must satisfy the following properties:
  - *Absolutely homogeneous:* $\forall x \in \mathbb{V},\ \forall \lambda \in \mathbb{R}\quad \|\lambda x\| = |\lambda|\|x\|$
  - *Positive definite:* $\forall x \in \mathbb{V}\quad \|x\| \geq 0$ and $\|x\| = 0 \Leftrightarrow x = 0$

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 $\|x\| = 0 \Leftrightarrow x = 0$
  - *Triangle inequality:* $\forall x, y \in \mathbb{V}\ \ \|x + y\| \leq \|x\| + \|y\|$



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 $x \in \mathbb{R}^n$:

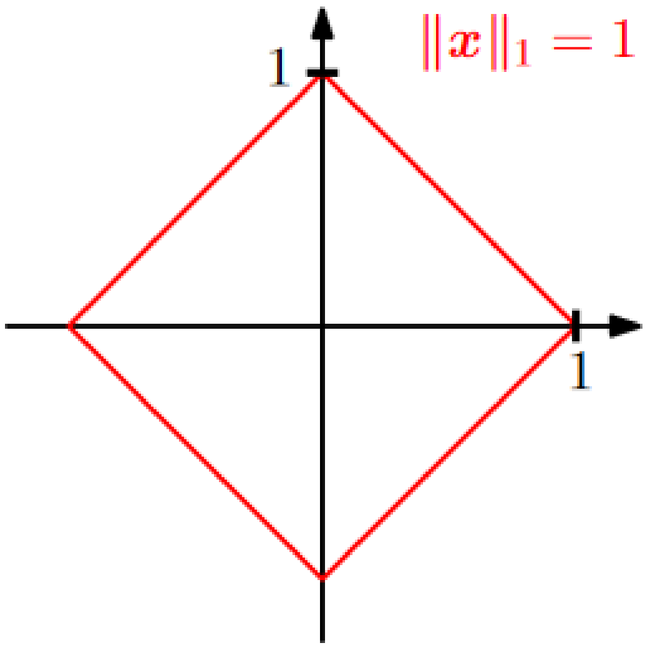
$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

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

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

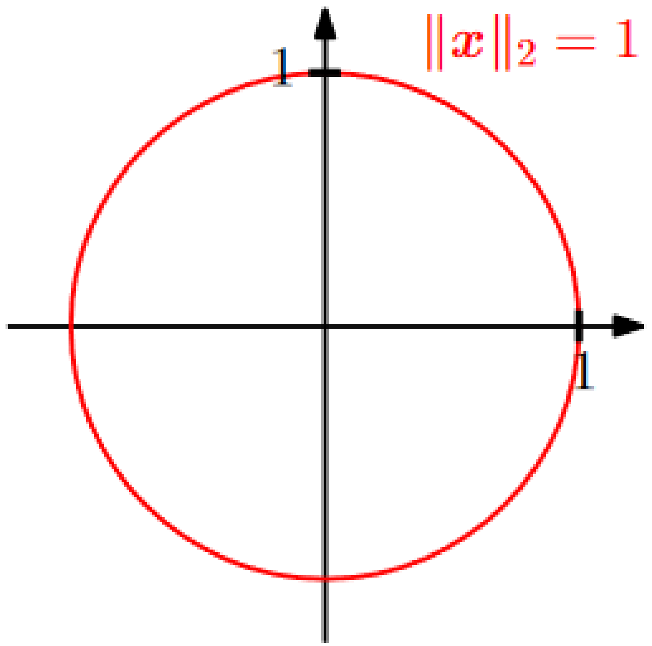
- Computes *Euclidian distance* from the origin.

- 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 $x = [x_1, \ldots, x_n] \in \mathbb{R}^n$ 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:

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

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:

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

- (!) 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 $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined as

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

girafe ai

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

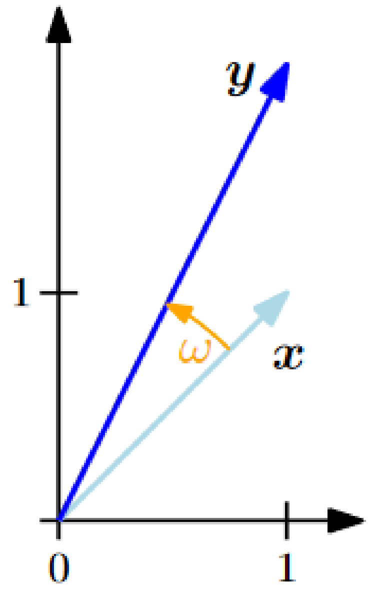
Angle between Two Vectors

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

- Remember Cauchy-Schwarz inequality:

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

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



ω: $\cos \omega = \frac{(x,y)}{\|x\| \cdot \|y\|}$ - 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 $y = [1, 0, 0]$?

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

girafe ai

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 $x = [\sqrt{3}, 0,1]$ 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 $y = [0, 1, 1, 0, 0]$?

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

girafe ai

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:

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

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.

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

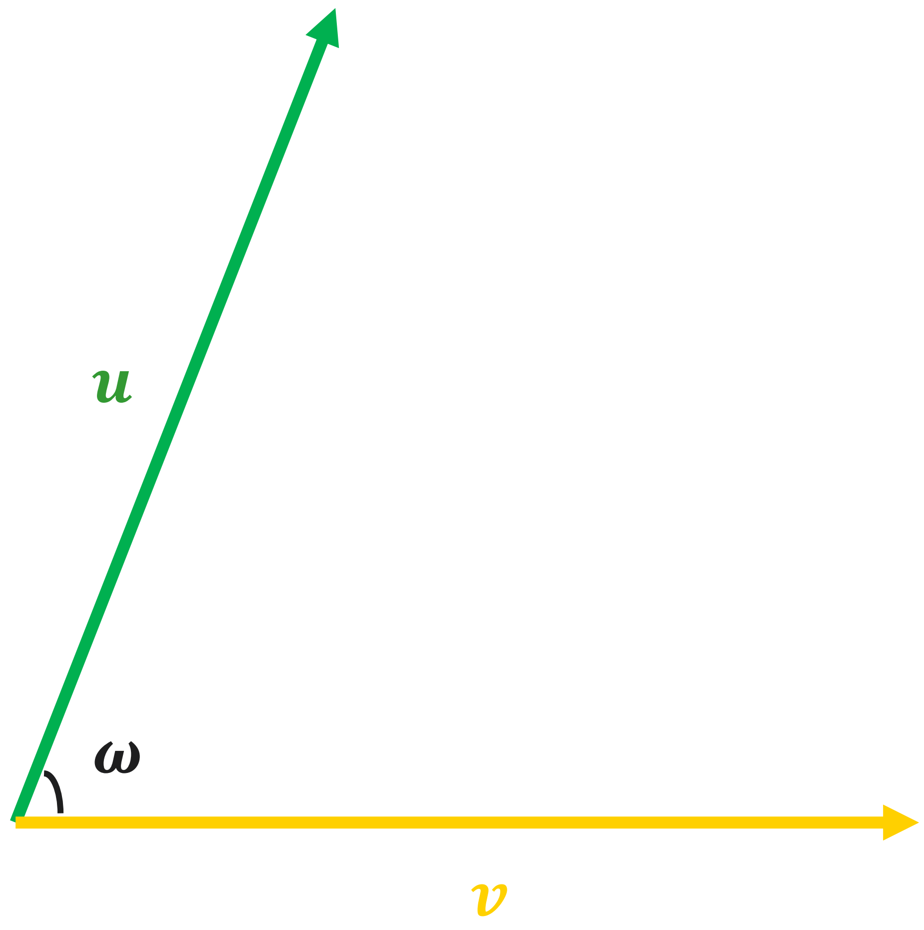
x and y are orthogonal.

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

x and y are *orthonormal*.

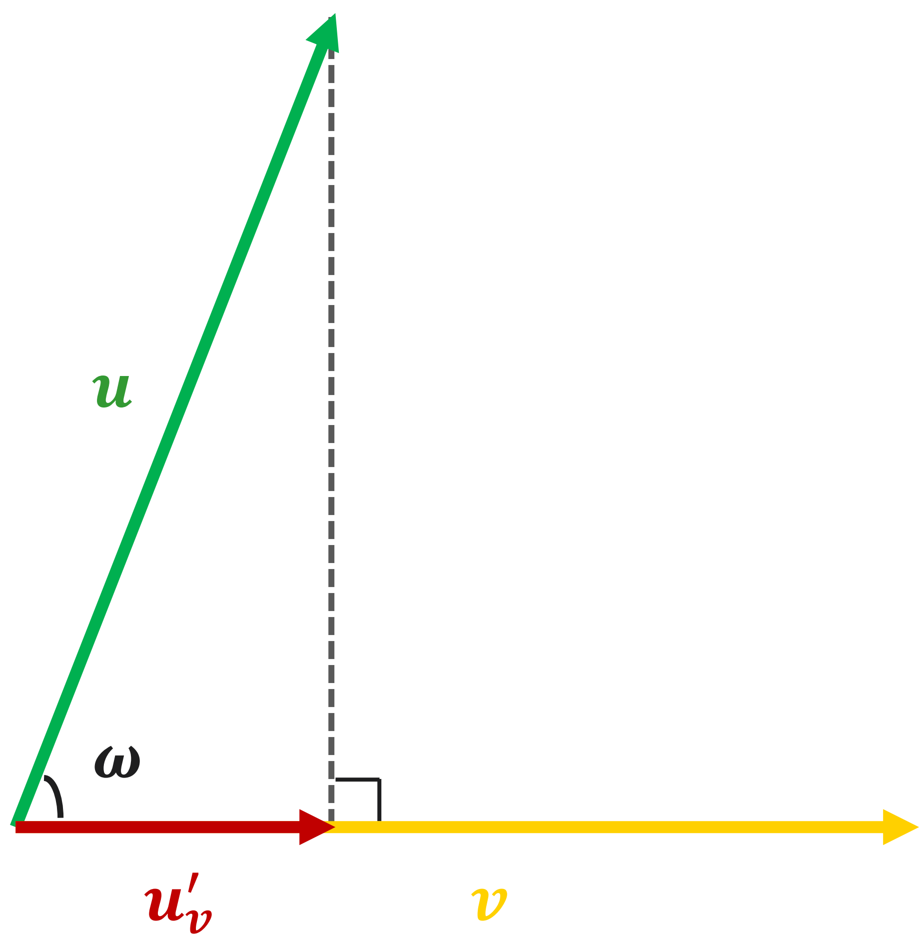
Orthogonal Projection

- Suppose we have two vectors u and v.



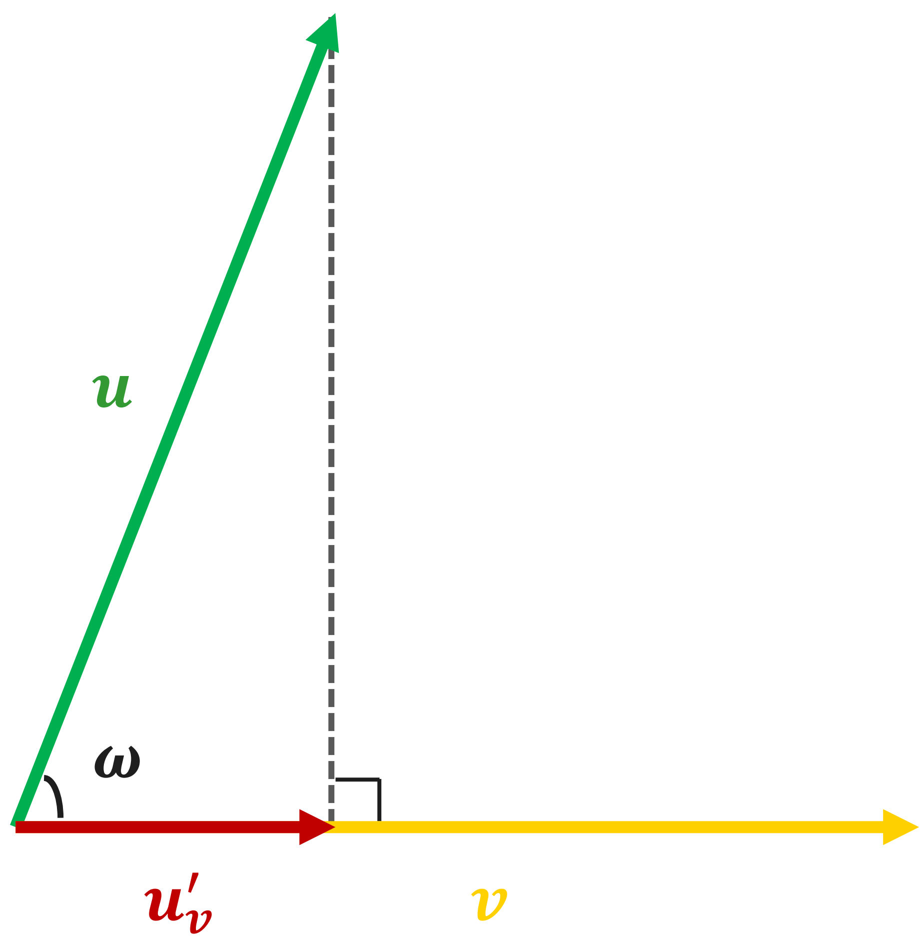
Orthogonal Projection

- 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?



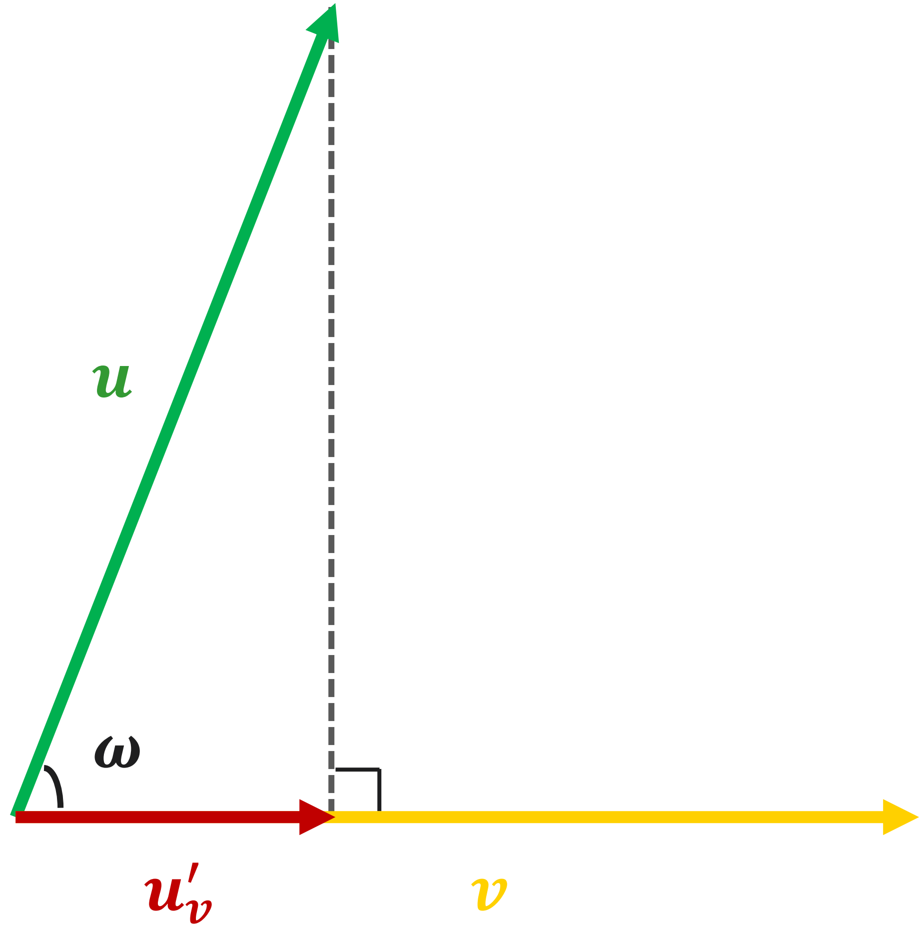
Orthogonal Projection

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

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

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

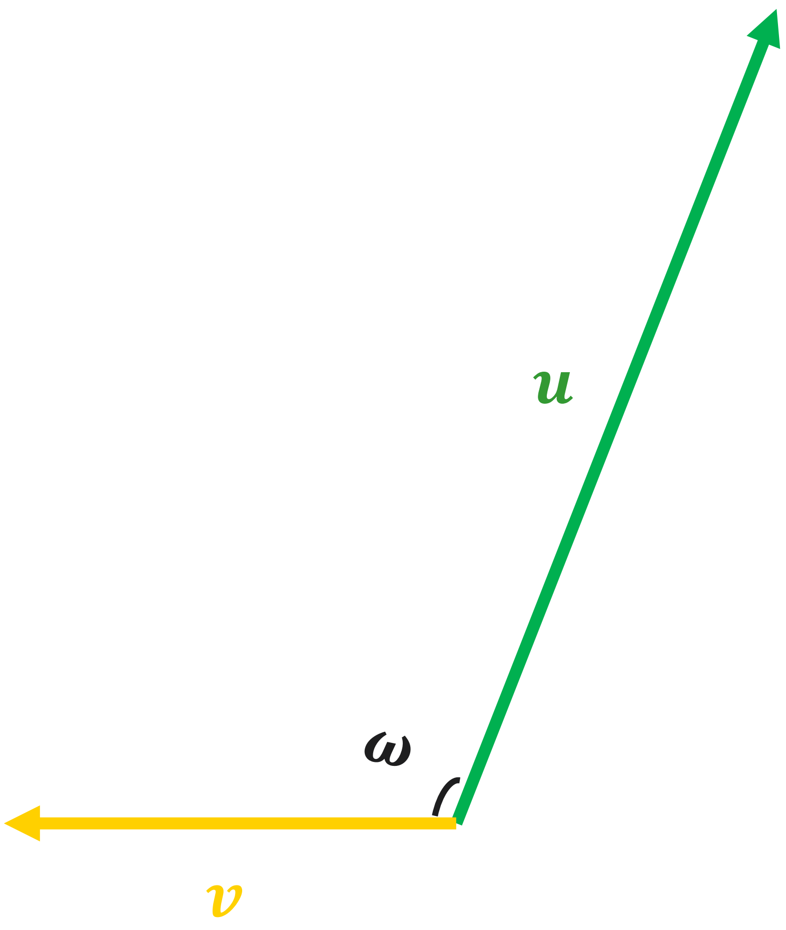
$$= \|u_v'\|\|v\|$$



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'?

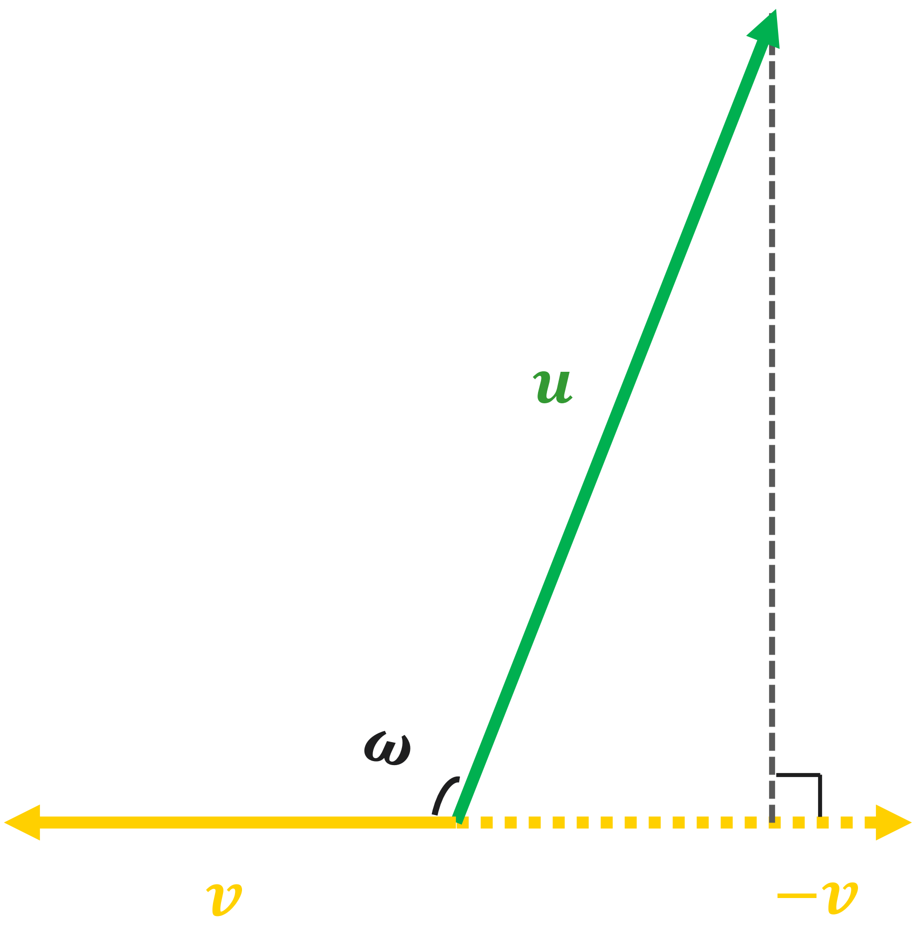
If $\mathbf{90 \leq \omega \leq 180}$



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?

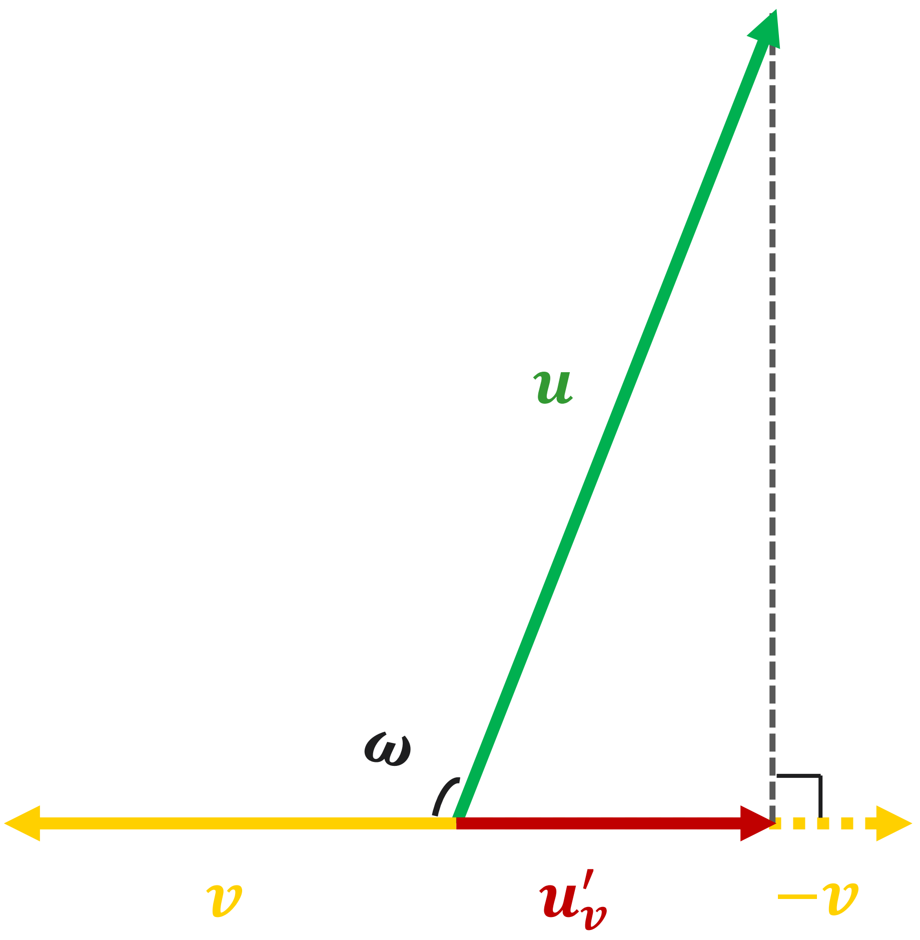
If $\mathbf{90 \leq \omega \leq 180}$



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



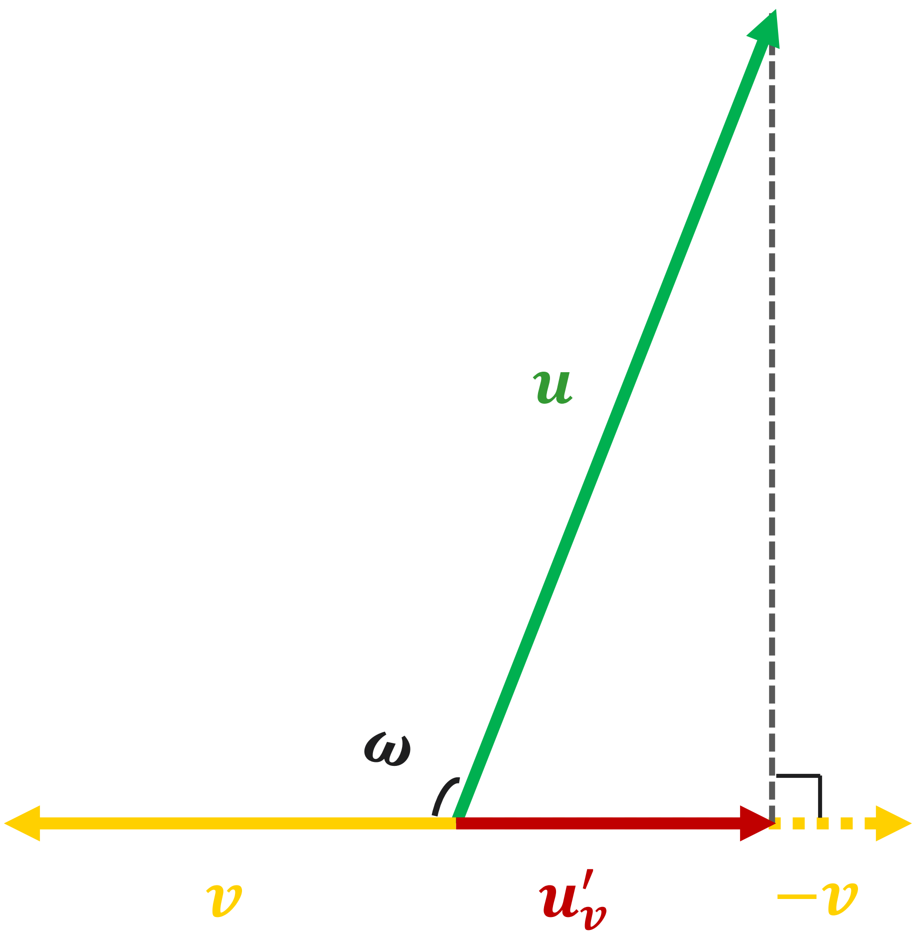
Orthogonal Projection

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

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

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

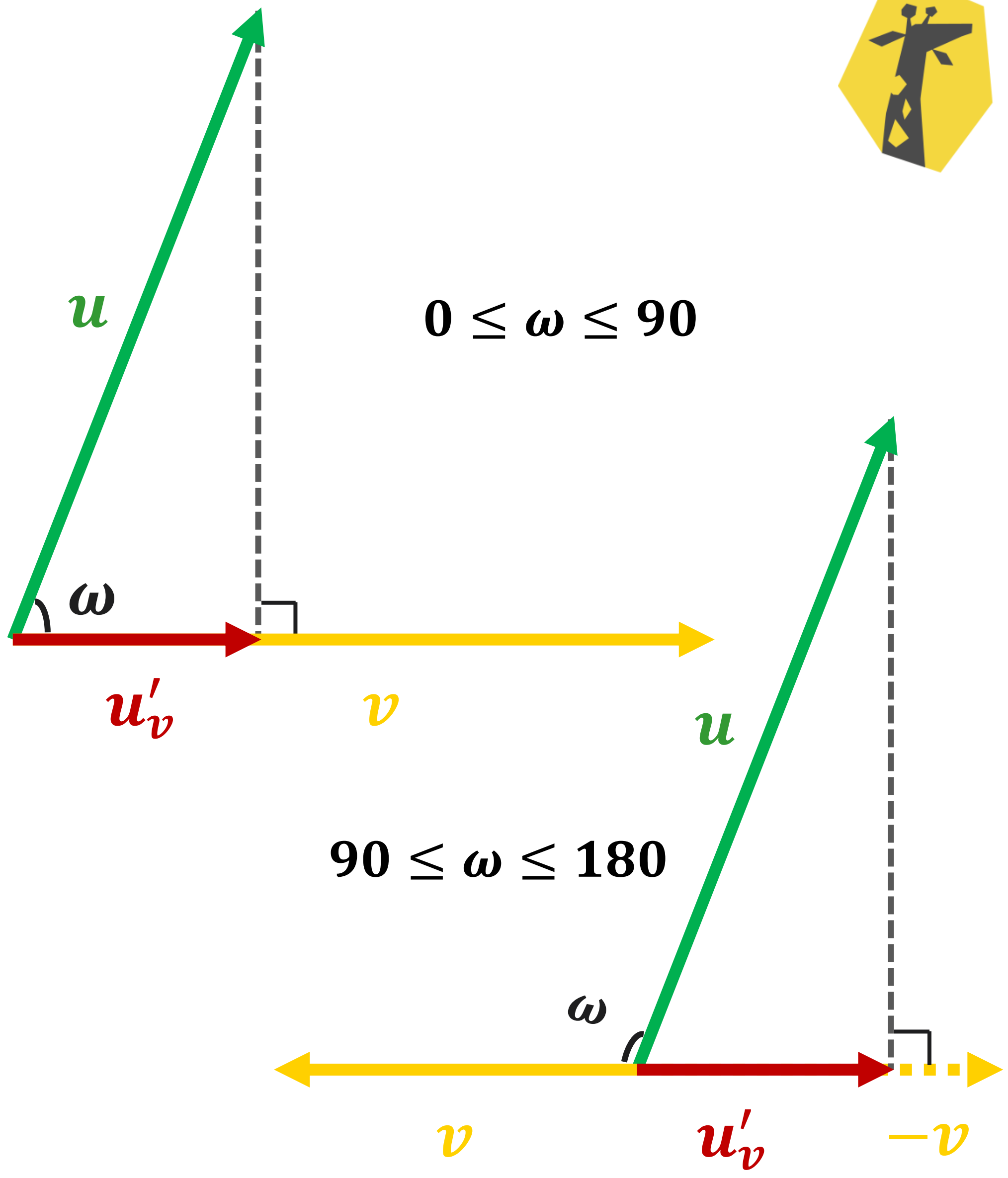
$$= \|u'_v\|\|v\|$$



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?

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

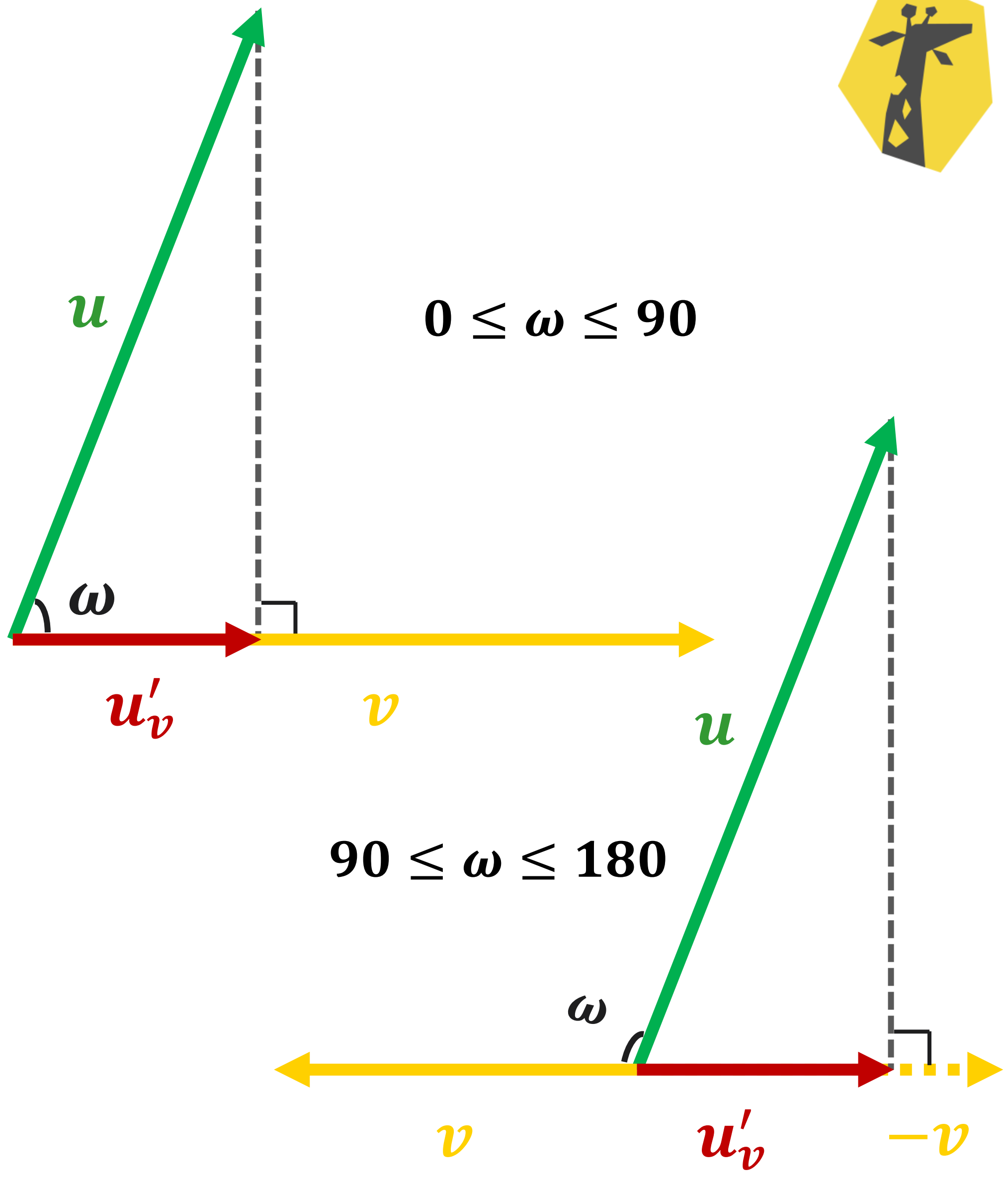


Orthogonal Projection

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

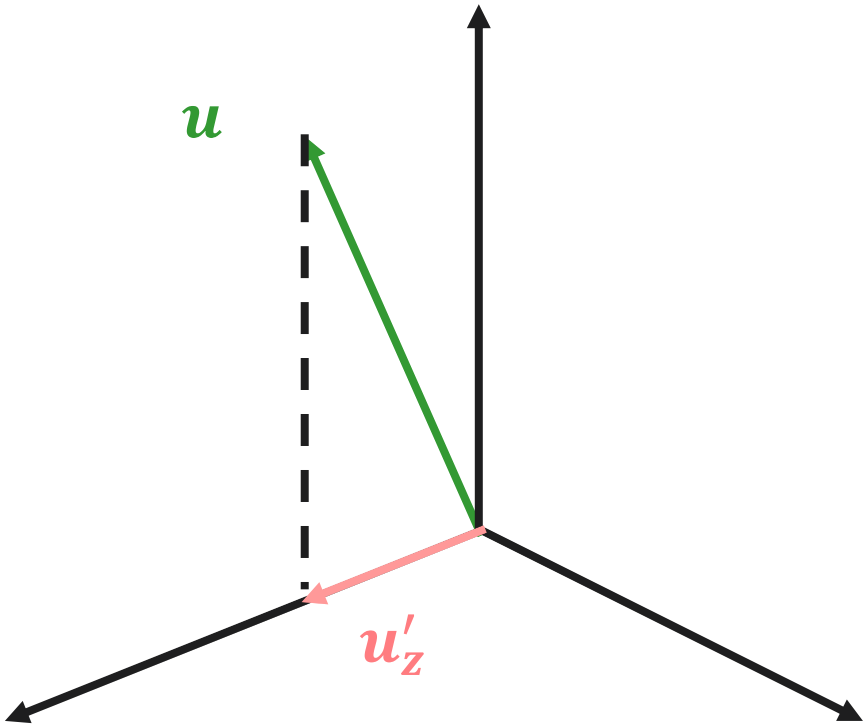
$$|(\boldsymbol{u}, \boldsymbol{v})| = \|\boldsymbol{u}'_v\| \|\boldsymbol{v}\| \quad \leftrightarrow \quad \|\boldsymbol{u}'_v\| = \frac{|(\boldsymbol{u}, \boldsymbol{v})|}{\|\boldsymbol{v}\|}$$

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



Orthogonal Projection: Example

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

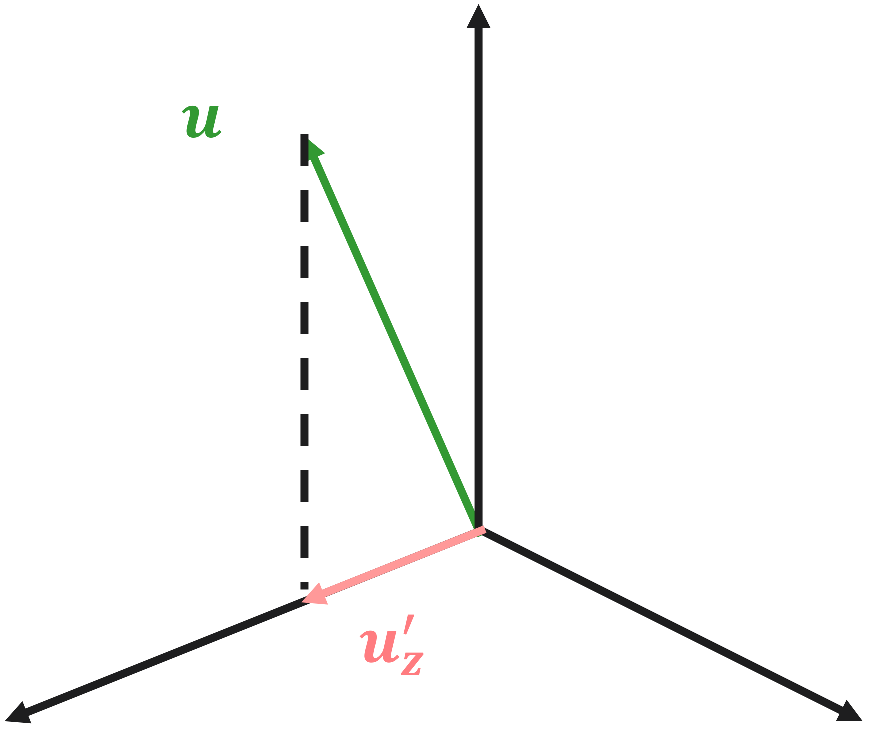


Orthogonal Projection: Example

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

$$u'_z = \frac{(u, z)}{(z, z)} z = 3z = [0, 0, 3]$$

(Projection on the axis = drop other coordinates)



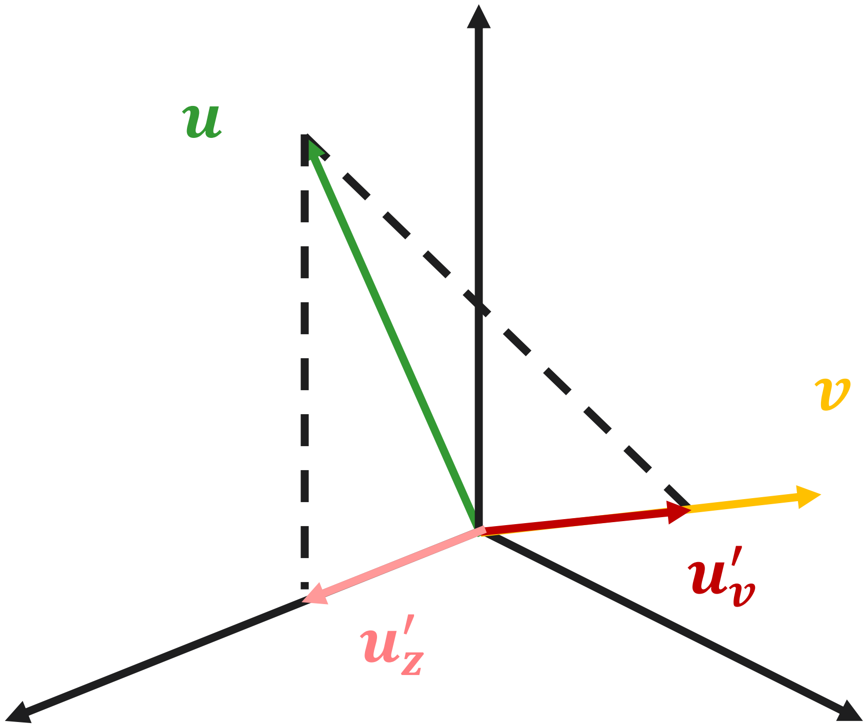
Orthogonal Projection: Example

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

$$u'_z = \frac{(u, z)}{(z, z)} z = 3z = [0, 0, 3]$$

(Projection on the axis = drop other coordinates)

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



Orthogonal Projection: Example

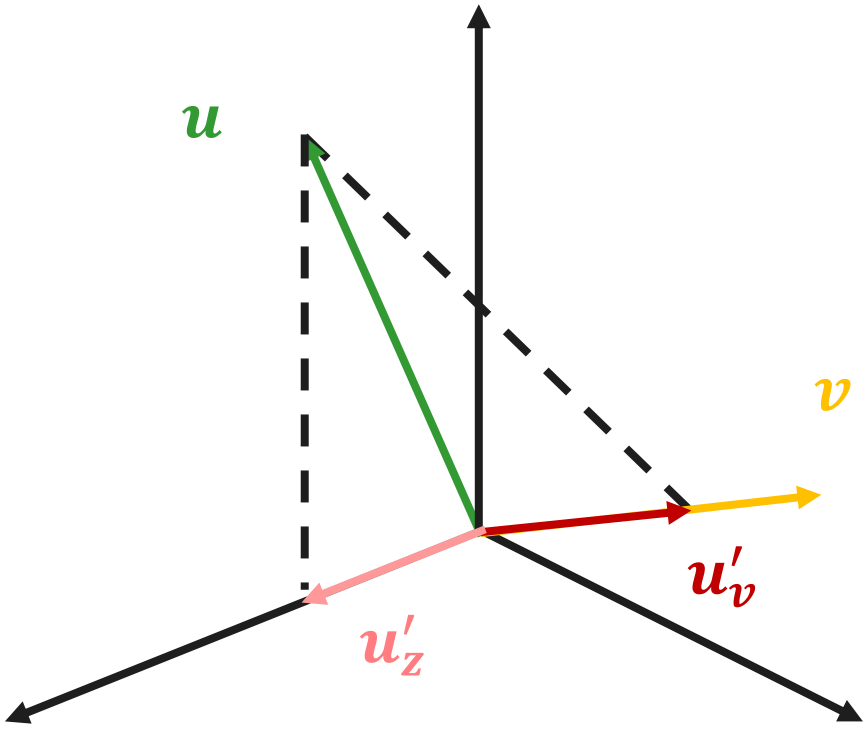
- What's projection of $u = [1, 3, 2]$ on $z = [0, 0, 1]$?

$$u'_z = \frac{(u, z)}{(z, z)} z = 3z = [0, 0, 3]$$

(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

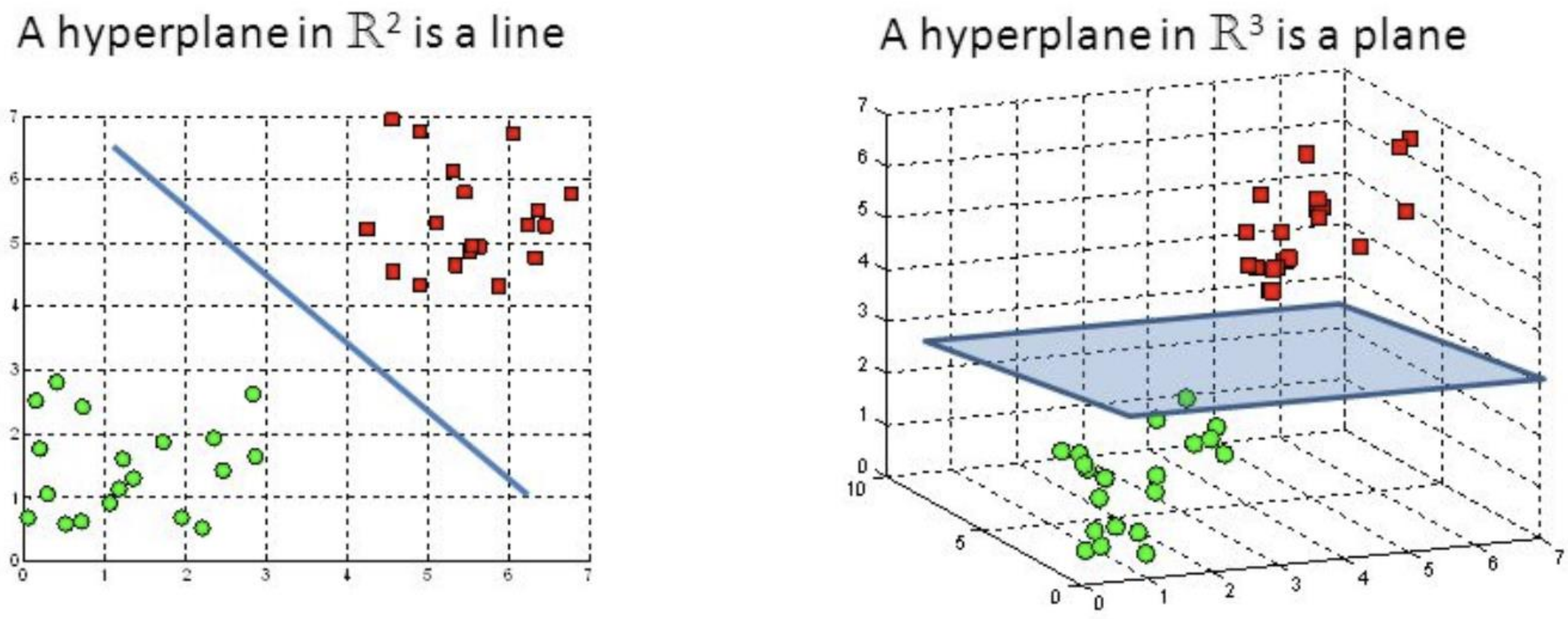


Image source: https://deepai.org/machine-learning-glossary-and-terms/hyperplane

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:

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

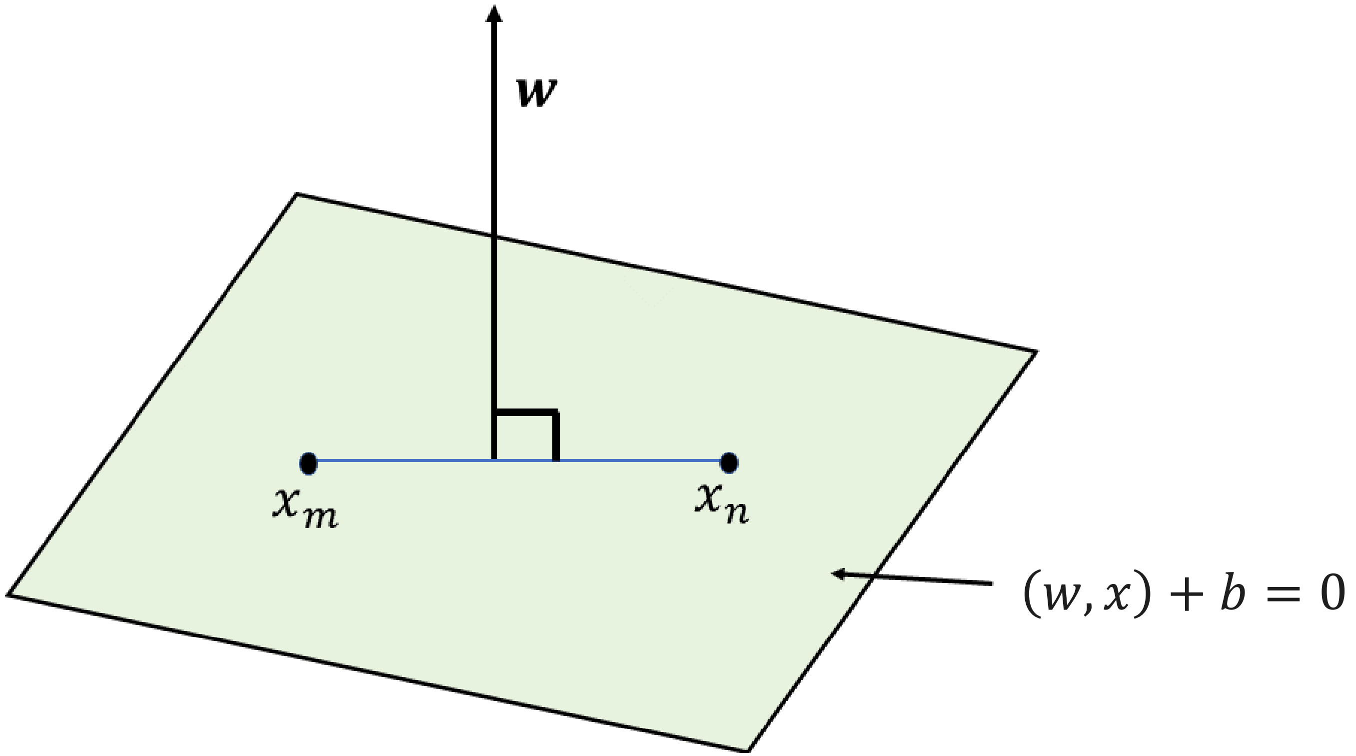
Normal to a Hyperplane

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

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

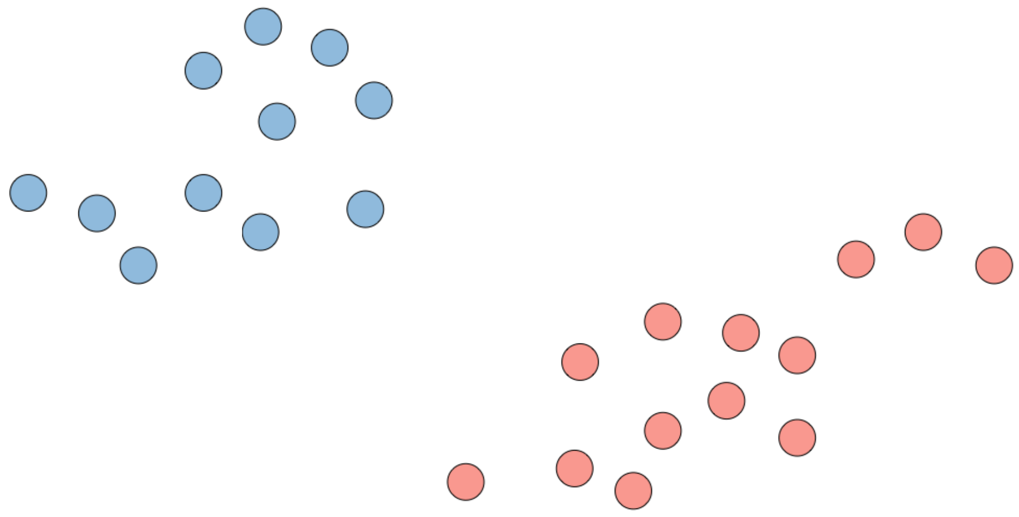
Normal to a Hyperplane

- 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.



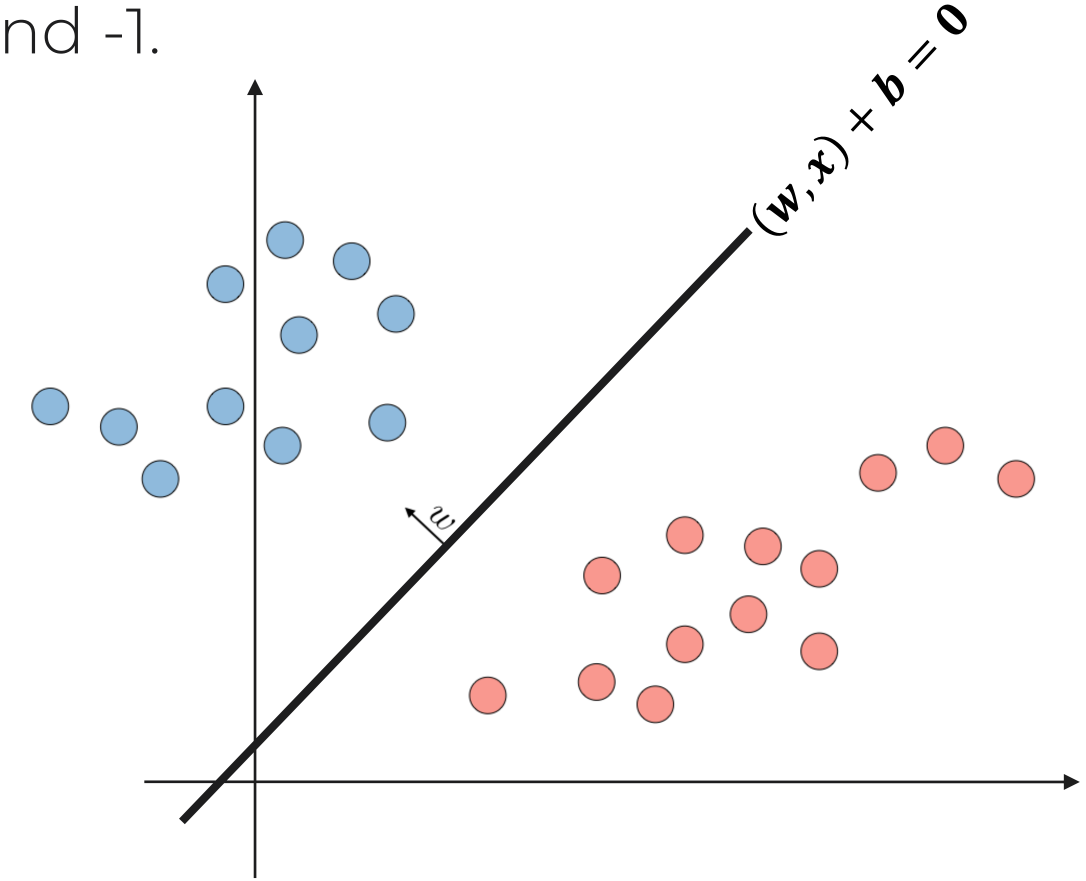
ML Example: Linear Classifier

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



ML Example: Linear Classifier

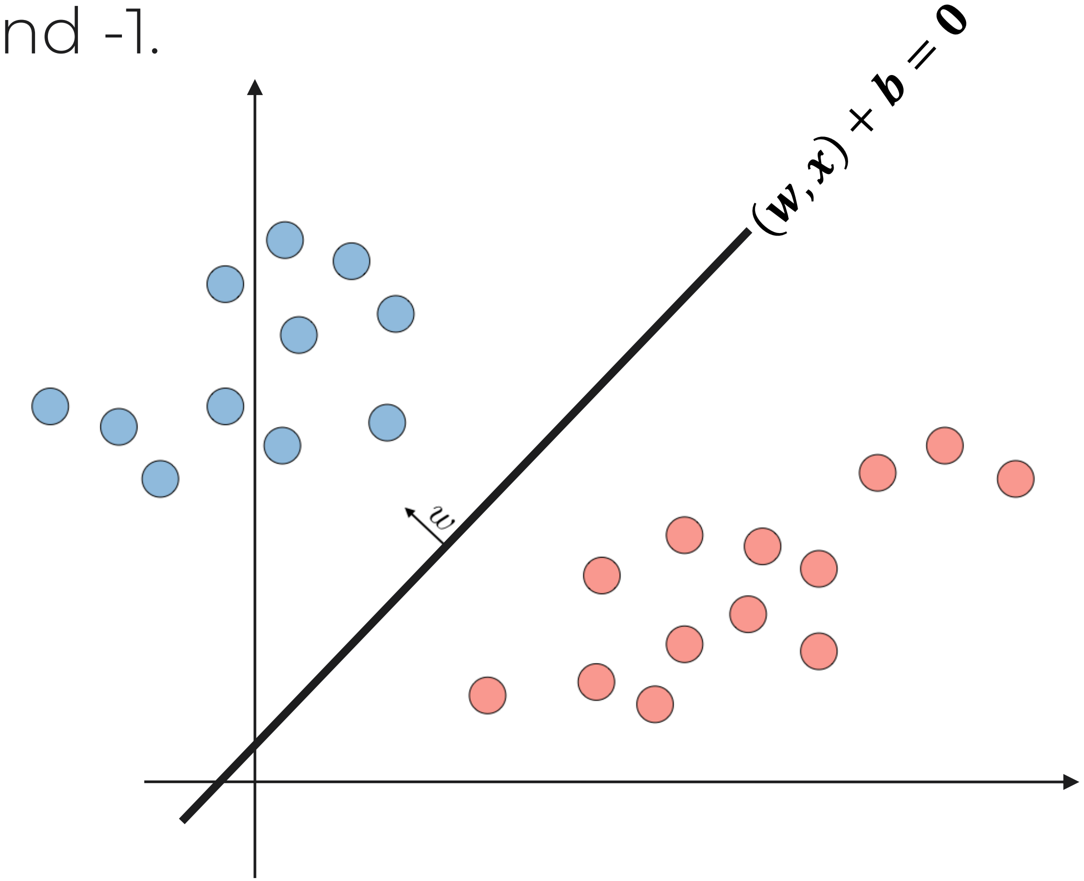
- Objects = 2D vectors
- Binary classification: classes +1 and -1.
- Linear classifier: separating hyperplane $(w, x) + b = 0$
    - objects "above" are class +1
    - objects "below" are class -1.



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.

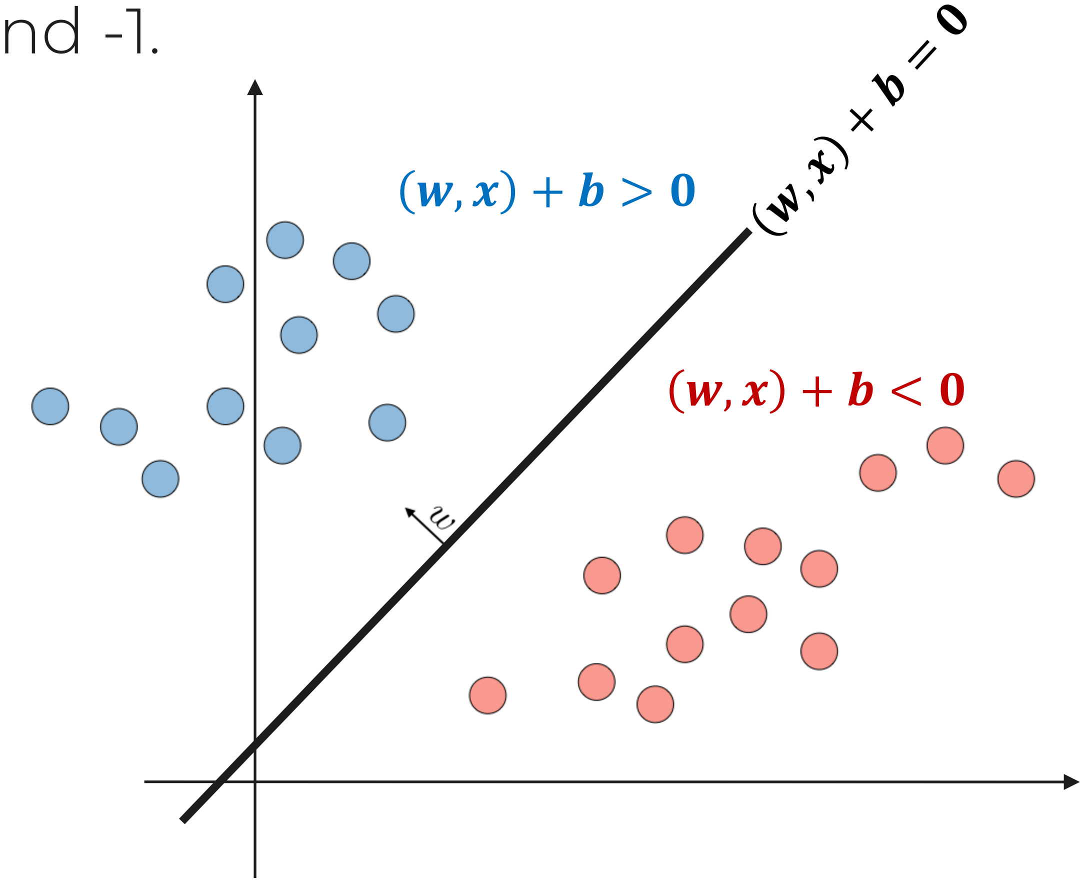
- How can we formalize this?



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

Linear Combinations

Linear Combinations

- We can add vectors together and multiply them by scalars. Let's combine these two operations!

Linear Combinations

- We can add vectors together and multiply them by scalars. Let's combine these two operations!

- V – a vector space.

Linear Combinations

- We can add vectors together and multiply them by scalars. Let's combine these two operations!

- V – a vector space.
- $x_1, x_2, \dots, x_k \in V$ – some vectors.

Linear Combinations

- We can add vectors together and multiply them by scalars. Let's combine these two operations!

- V – a vector space.
- $x_1, x_2, \dots, x_k \in V$ – some vectors.
- $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R}$ - some scalars.

Linear Combinations

- We can add vectors together and multiply them by scalars. Let's combine these two operations!

- V – a vector space.
- $x_1, x_2, \dots, x_k \in V$ – some vectors.
- $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R}$ - some scalars.

$$v = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k = \sum_{i=1}^{k} \lambda_i x_i \in V \text{ –}$$

a *linear combination* of $x_1, x_2, \dots, x_k$.

Linear Combinations: Examples

- In $(\mathbb{R}^2, +, \ \cdot)$, consider vectors $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Linear Combinations: Examples

- In $(\mathbb{R}^2, +, \cdot)$, consider vectors $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

  $u = 2e_1 - e_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ is a linear combination of e_1 and e_2.

Linear independence

Linear Combinations

- A zero vector can always be represented as a trivial linear combination of $x_1, x_2, \dots, x_k$:

$$0 = \sum_{i=1}^{k} 0 \cdot x_i$$

girafe ai

Linear Combinations

- A zero vector can always be represented as a trivial linear combination of $x_1, x_2, \dots, x_k$:

$$0 = \sum_{i=1}^{k} 0 \cdot x_i$$

- We are mostly interested in *non-trivial linear combinations* of $x_1, x_2, \dots, x_k$ where not all λ_i are $\mathbf{0}$.

Linear (In)dependence

- Consider a vector space V.
- $x_1, x_2, \dots, x_k \in V$ – some vectors.

Linear (In)dependence

- Consider a vector space V.
- $x_1, x_2, \dots, x_k \in V$ – some vectors.

- If there is a non-trivial linear combination of $x_1, x_2, \dots, x_k$ such that $\sum_{i=1}^{k} \lambda_i x_i = 0$ with at least one $\lambda_i \neq 0$, vectors $x_1, x_2, \dots, x_k$ are *linearly dependent.*

Linear (In)dependence

- Consider a vector space V.
- $x_1, x_2, \dots, x_k \in V$ – some vectors.

- If there is a <u>non-trivial</u> linear combination of $x_1, x_2, \dots, x_k$ such that $\sum_{i=1}^{k} \lambda_i x_i = 0$ with at least one $\lambda_i \neq 0$, vectors $x_1, x_2, \dots, x_k$ are *linearly dependent.*

- If only trivial solution exists, vectors $x_1, x_2, \dots, x_k$ are *linearly independent.*

Linear (In)dependence

- If there is a non-trivial linear combination of $x_1, x_2, \dots, x_k$ such that $\sum_{i=1}^{k} \lambda_i x_i = \mathbf{0}$ with at least one $\lambda_i \neq 0$, vectors $x_1, x_2, \dots, x_k$ are *linearly dependent.*

Linear (In)dependence

- If there is a non-trivial linear combination of $x_1, x_2, \dots, x_k$ such that $\sum_{i=1}^{k} \lambda_i x_i = 0$ with at least one $\lambda_i \neq 0$, vectors $x_1, x_2, \dots, x_k$ are *linearly dependent*.

$$\leftrightarrow$$

- A set of vectors $x_1, x_2, \dots, x_k$ is linearly dependent if and only if (at least) one of the vectors is a linear combination of the others

$$x_i = \alpha_1 x_1 + \cdots + \alpha_k x_k$$

Linear (In)dependence: Example 1

- Consider $\mathbb{R}^2$.

Linear (In)dependence: Example 1

- Consider $\mathbb{R}^2$.
- Vectors $u = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ and $v = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$ are not linearly independent: $u = -2v$.

Linear (In)dependence: Example 1

- Consider $\mathbb{R}^2$.
- Vectors $u = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ and $v = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$ are not linearly independent: $u = -2v$.
- Vectors $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are linearly independent:
  there are no $\lambda_1, \lambda_2 \in \mathbb{R}$ with at least one $\lambda_i \neq 0$ such that
  $\lambda_1 e_1 + \lambda_2 e_2 = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Linear (In)dependence: Example 1

- Consider $\mathbb{R}^2$.

- Vectors $u = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ and $v = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$ are not linearly independent: $u = -2v$.

- Vectors $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are linearly independent:
  there are no $\lambda_1, \lambda_2 \in \mathbb{R}$ with at least one $\lambda_i \neq 0$ such that
  $\lambda_1 e_1 + \lambda_2 e_2 = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

  (Or: you cannot represent e_1 as λe_2 or vice versa).

Dimension of a Linear Space

- The linear space V is called $\boldsymbol{n}$-*dimensional* if the maximal number of linearly independent vectors in it is equal to n.

Dimension of a Linear Space

- The linear space V is called ***n*-*dimensional*** if the maximal number of linearly independent vectors in it is equal to n.

  We denote this as $\mathbf{dim}(V) = n$.

Dimension: Example

- $\mathbb{R}^n$ is a n-dimensional vector space. Why?

- Consider n vectors $e_1, \dots, e_n$:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

- $e_1, \dots, e_n$ are linearly independent → $\dim(\mathbb{R}^n) \geq n$.

- Can there be more than n linearly independent vectors in $\mathbb{R}^n$?

Dimension: Example

- $\mathbb{R}^n$ is a n-dimensional vector space. Why?

Dimension: Example

- $\mathbb{R}^n$ is a n-dimensional vector space. Why?

- Consider n vectors $e_1, \ldots, e_n$:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \qquad \ldots, \qquad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Dimension: Example

- $\mathbb{R}^n$ is a n-dimensional vector space. Why?

- Consider n vectors $e_1, \dots, e_n$:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \qquad \dots, \qquad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

- $e_1, \dots, e_n$ are linearly independent $\rightarrow \dim(\mathbb{R}^n) \geq n$.

Dimension: Example

- Can there be more than n linearly independent vectors in $\mathbb{R}^n$?

Basis

Basis

- A set of n linearly independent vectors $e_1, e_2, \ldots, e_n$ in an n-dimensional space V is called a *basis* for V.

Basis

- A set of n linearly independent vectors $e_1, e_2, \dots, e_n$ in an n-dimensional space V is called a *basis* for V.

- Basis is A set of vectors with which we can represent every vector in the vector space by adding them together and scaling them.

Basis: Example

Basis: Example

- $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is a basis for $\mathbb{R}^2$.

girafe ai

Basis: Example

- $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is a basis for $\mathbb{R}^2$.

- $e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix},\ \dots,\ e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$ is a basis for $\mathbb{R}^n$.

Coordinates

- Let's fixe the order of the vectors in the basis:

$$e_1, e_2, \dots, e_n$$

Coordinates

- Let's fixe the order of the vectors in the basis:

$$e_1, e_2, \dots, e_n$$

- Every vector v in a vector space can be represented as a linear combination of $e_1, e_2, \dots, e_n$:

$$v = a_1 e_1 + a_2 e_2 + \dots + a_n e_n$$

Coordinates

- Let's fixe the order of the vectors in the basis:

$$e_1, e_2, \dots, e_n$$

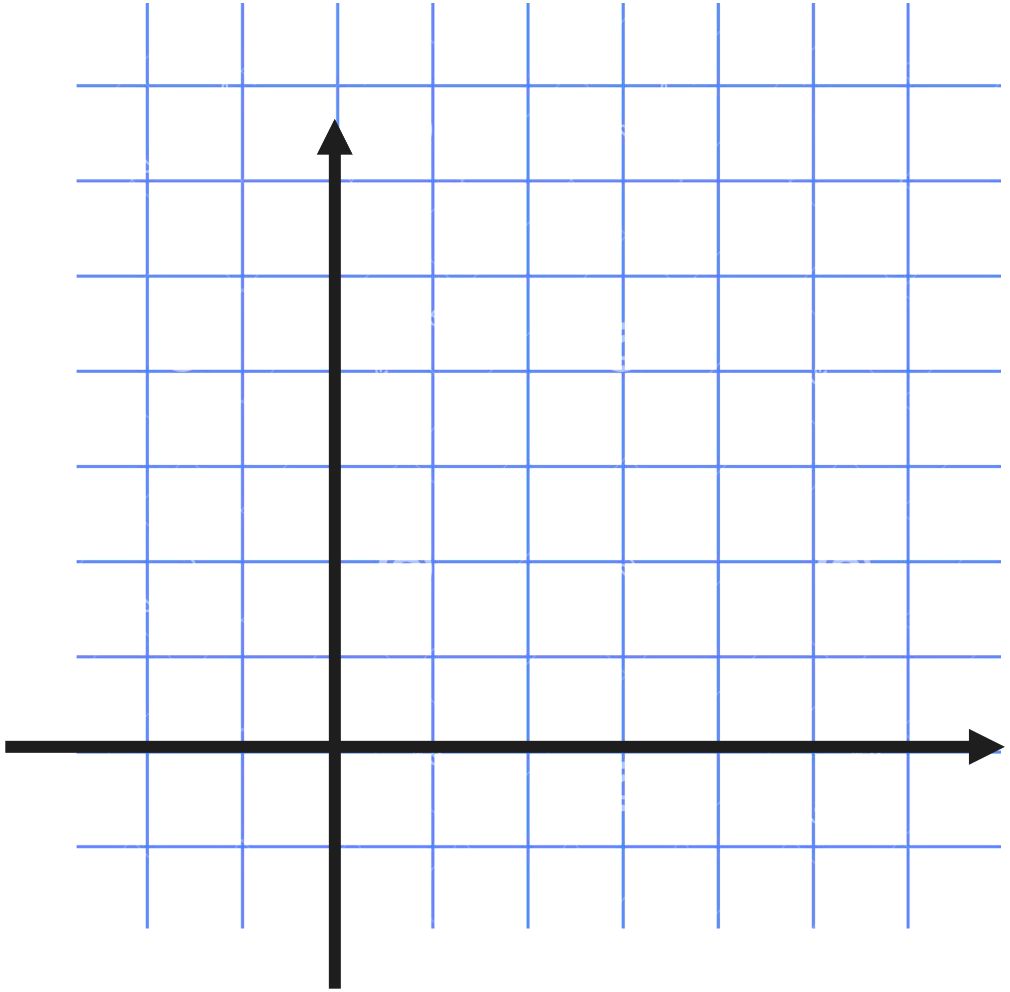
- Every vector v in a vector space can be represented as a linear combination of $e_1, e_2, \dots, e_n$:

$$v = a_1 e_1 + a_2 e_2 + \dots + a_n e_n$$

- $a_1, a_2, \dots, a_n$ - *coordinates* of the vector v in the basis $e_1, e_2, \dots, e_n$.

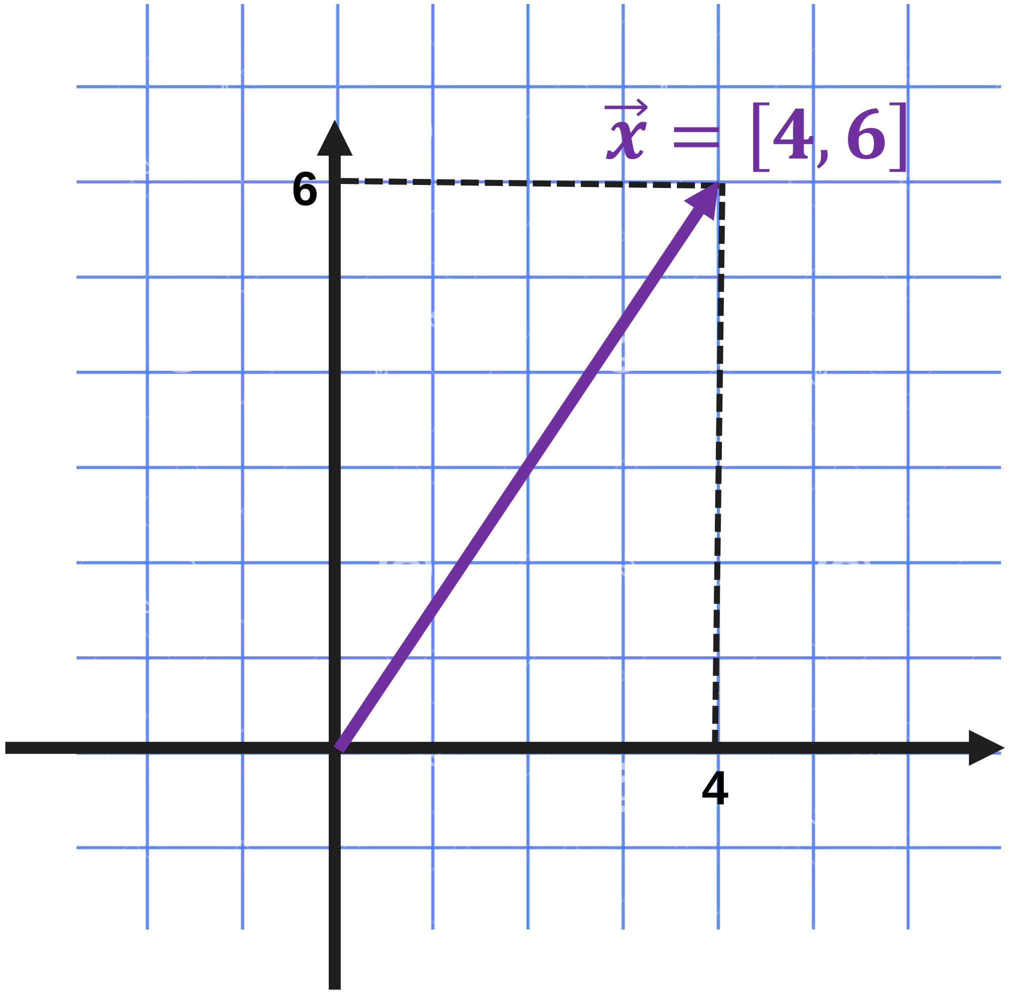
Coordinates: Example

- Consider $\mathbb{R}^2$.



Coordinates: Example

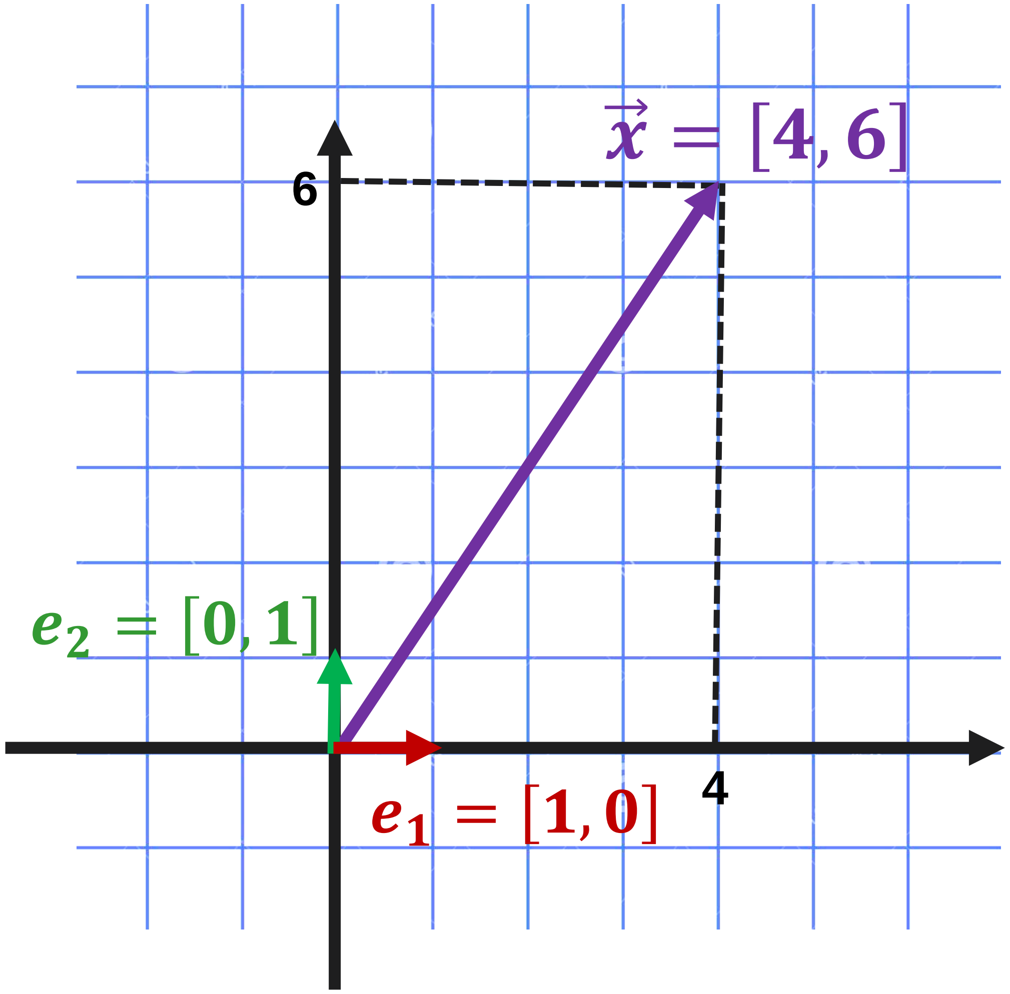
- Consider $\mathbb{R}^2$.
- $x = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$



Coordinates: Example

- Consider $\mathbb{R}^2$.
- $x = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$
- Canonical basis:

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



Coordinates: Example

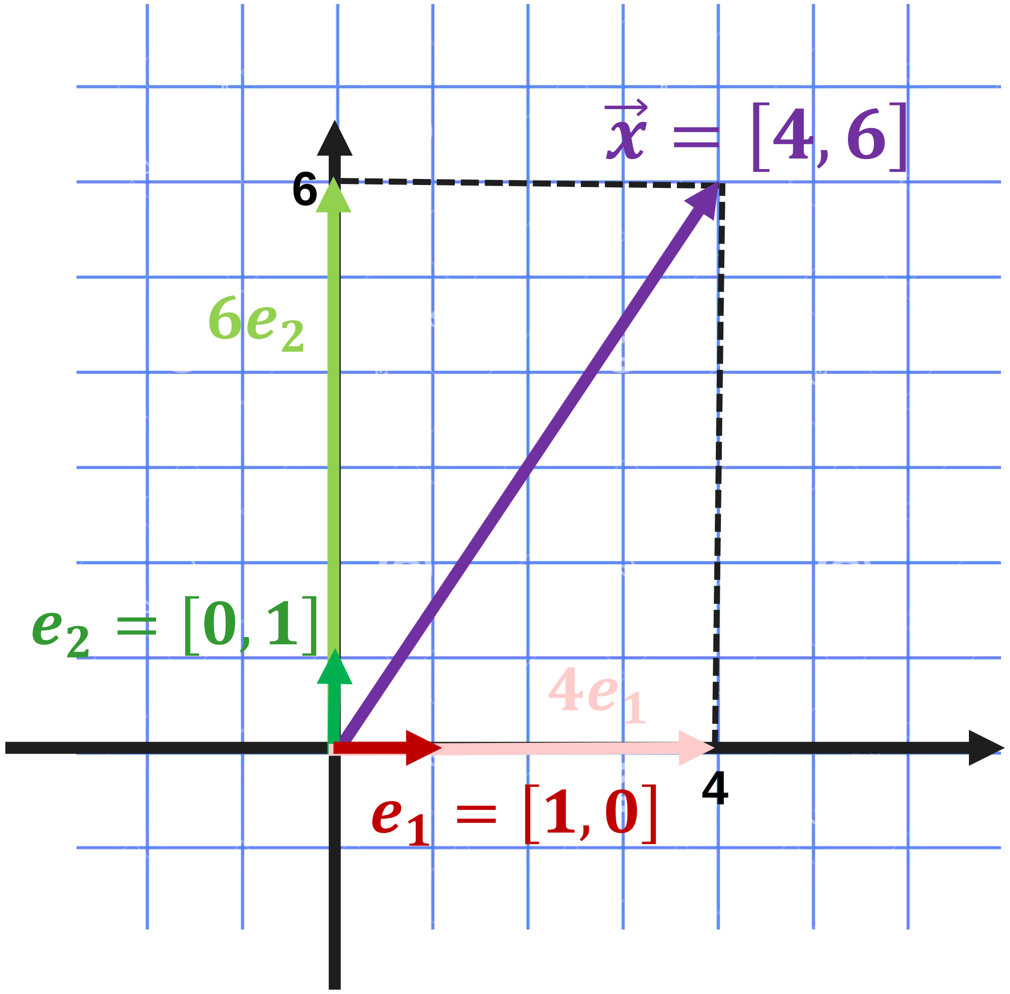
- Consider $\mathbb{R}^2$.

- $x = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$

- Canonical basis:

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

$$\boldsymbol{x} = 4\boldsymbol{e_1} + 6\boldsymbol{e_2}$$



Orthogonal Basis

- A basis $e_1, \dots, e_n$ where $\forall i, j \;\; e_i \perp e_j$ is an *orthogonal basis*.

Orthogonal Basis

- A basis $e_1, \dots, e_n$ where $\forall i, j \;\; e_i \perp e_j$ is an *orthogonal basis*.
- If, additionally, $\forall i \; \|e_i\| = 1$, then $e_1, \dots, e_n$ is an *orthonormal basis*.

Orthogonal Basis

- A basis $e_1, \dots, e_n$ where $\forall i, j \;\; e_i \perp e_j$ is an *orthogonal basis*.
- If, additionally, $\forall i \; \|e_i\| = 1$, then $e_1, \dots, e_n$ is an *orthonormal basis*.

Example:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \; e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \; \dots, \; e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$ is an orthonormal basis of $\mathbb{R}^n$.

Orthogonal Basis

- A basis $e_1, \dots, e_n$ where $\forall i, j \;\; e_i \perp e_j$ is an *orthogonal basis*.
- If, additionally, $\forall i \; \|e_i\| = 1$, then $e_1, \dots, e_n$ is an *orthonormal basis*.

Example:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \; e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \; \dots, \; e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \text{ is an orthonormal basis of } \mathbb{R}^n.$$

Change of Basis

Basis

- A vector space has more than one basis.

Basis

- A vector space has more than one basis.
- Example: $\mathbb{R}^2$
  - $e = \left\{ e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ - canonical basis;

Basis

- A vector space has more than one basis.
- Example: $\mathbb{R}^2$
    - $e = \left\{ e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ - canonical basis;
    - $a = \left\{ a_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, a_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ - another basis;

Basis

- A vector space has more than one basis.
- Example: $\mathbb{R}^2$
  - $e = \left\{ e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ - canonical basis;
  - $a = \left\{ a_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, a_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ - another basis;
  - $b = \left\{ b_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, b_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right\}$ - yet another one.

Basis

- A vector space has more than one basis.
- Example: $\mathbb{R}^2$
    - $e = \left\{e_1 = \begin{bmatrix}1\\0\end{bmatrix}, e_2 = \begin{bmatrix}0\\1\end{bmatrix}\right\}$ - canonical basis;
    - $a = \left\{a_1 = \begin{bmatrix}1\\1\end{bmatrix}, a_2 = \begin{bmatrix}1\\-1\end{bmatrix}\right\}$ - another basis;
    - $b = \left\{b_1 = \begin{bmatrix}2\\1\end{bmatrix}, b_2 = \begin{bmatrix}-1\\-1\end{bmatrix}\right\}$ - yet another one.
- Different basis = different coordinates.
  How exactly do they change?

Coordinate Change: Example

- Consider $\mathbb{R}^2$ with canonical basis

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

Coordinate Change: Example

- Consider $\mathbb{R}^2$ with canonical basis

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

- New basis:

$$e'_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \qquad e'_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

Coordinate Change: Example

- Consider $\mathbb{R}^2$ with canonical basis

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

- New basis:

$$e'_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \qquad e'_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

- $x_{old} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

Coordinate Change: Example

- Consider $\mathbb{R}^2$ with canonical basis

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

- New basis:

$$e'_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \qquad e'_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

- $x_{old} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

- What are the coordinates in the new basis?

$$x_{new} = ?$$

Coordinate Change: General Case

- Consider a vector space V with basis $e_1, e_2, \ldots e_n$.

Coordinate Change: General Case

- Consider a vector space V with basis $e_1, e_2, \dots e_n$.

- Imagine vector $x = [x_1, x_2, \dots, x_n] \in V$
  $x_1, x_2, \dots, x_n$ - coordinates in basis $e_1, e_2, \dots e_n$.

Coordinate Change: General Case

- Consider a vector space V with basis $e_1, e_2, \dots e_n$.
- Imagine vector $x = [x_1, x_2, \dots, x_n] \in V$
  $x_1, x_2, \dots, x_n$ - coordinates in basis $e_1, e_2, \dots e_n$.
- New basis: $e'_1, e'_2, \dots e'_n$.

Coordinate Change: General Case

- Consider a vector space V with basis $e_1, e_2, \dots e_n$.

- Imagine vector $x = [x_1, x_2, \dots, x_n] \in V$
  $x_1, x_2, \dots, x_n$ - coordinates in basis $e_1, e_2, \dots e_n$.

- New basis: $e'_1, e'_2, \dots e'_n$.

- What are the coordinates of x in this new basis?

$$x'_1, x'_2, \dots, x'_n = ?$$

girafe ai

Coordinate Change

- Old basis: $e_1, e_2, \dots e_n$
  New basis: $e'_1, e'_2, \dots e'_n$

- $x_{old} = [x_1, x_2, \dots, x_n]$, $\quad x_{new} = [x'_1, x'_2, \dots, x'_n] = ?$

- Coordinates of the new basis in the old one:

Coordinate Change

- Old basis: $e_1, e_2, \dots e_n$
  New basis: $e'_1, e'_2, \dots e'_n$

- $x_{old} = [x_1, x_2, \dots, x_n]$, $\quad x_{new} = [x'_1, x'_2, \dots, x'_n\ ] = ?$

- Coordinates of the new basis in the old one:

$$
\begin{gathered}
e'_1 = \boldsymbol{\alpha_{11}} e_1 + \boldsymbol{\alpha_{21}} e_2 + \cdots + \boldsymbol{\alpha_{n1}} e_n \\
e'_2 = \boldsymbol{\alpha_{12}} e_1 + \boldsymbol{\alpha_{22}} e_2 + \cdots + \boldsymbol{\alpha_{n2}} e_n \\
\vdots \\
e'_i = \boldsymbol{\alpha_{1i}} e_1 + \boldsymbol{\alpha_{2i}} e_2 + \cdots + \boldsymbol{\alpha_{ni}} e_n \\
\vdots \\
e'_n = \boldsymbol{\alpha_{1n}} e_1 + \boldsymbol{\alpha_{2n}} e_2 + \cdots + \boldsymbol{\alpha_{nn}} e_n
\end{gathered}
$$

Coordinate Change

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n = x'_1 \boldsymbol{e'_1} + x'_2 \boldsymbol{e'_2} + \cdots x'_n \boldsymbol{e'_n} =$$

Coordinate Change

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n = x'_1 \boldsymbol{e'_1} + x'_2 \boldsymbol{e'_2} + \cdots x'_n \boldsymbol{e'_n} =$$

Remember: $e'_i = \boldsymbol{\alpha_{1i}} e_1 + \boldsymbol{\alpha_{2i}} e_2 + \cdots + \boldsymbol{\alpha_{ni}} e_n$

girafe ai

Coordinate Change

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n = x'_1 \boldsymbol{e'_1} + x'_2 \boldsymbol{e'_2} + \cdots x'_n \boldsymbol{e'_n} =$$

Remember: $e'_i = \boldsymbol{\alpha_{1i}} e_1 + \boldsymbol{\alpha_{2i}} e_2 + \cdots + \boldsymbol{\alpha_{ni}} e_n$

$$= x'_1 \cdot (\boldsymbol{\alpha_{11} e_1} + \boldsymbol{\alpha_{21} e_2} + \cdots + \boldsymbol{\alpha_{n1} e_n}) + \cdots + x'_i \cdot (\boldsymbol{\alpha_{1i} e_1} + \boldsymbol{\alpha_{2i} e_2} + \cdots + \boldsymbol{\alpha_{ni} e_n}) + \cdots$$
$$+ x'_n (\boldsymbol{\alpha_{1n} e_1} + \boldsymbol{\alpha_{2n} e_2} + \cdots + \boldsymbol{\alpha_{nn} e_n}) =$$

Coordinate Change

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n = x'_1 \boldsymbol{e'_1} + x'_2 \boldsymbol{e'_2} + \cdots x'_n \boldsymbol{e'_n} =$$

Remember: $e_i' = \boldsymbol{\alpha_{1i}} e_1 + \boldsymbol{\alpha_{2i}} e_2 + \cdots + \boldsymbol{\alpha_{ni}} e_n$

$$= x'_1 \cdot (\boldsymbol{\alpha_{11} e_1} + \boldsymbol{\alpha_{21} e_2} + \cdots + \boldsymbol{\alpha_{n1} e_n}) + \cdots + x_i' \cdot (\boldsymbol{\alpha_{1i} e_1} + \boldsymbol{\alpha_{2i} e_2} + \cdots + \boldsymbol{\alpha_{ni} e_n}) + \cdots$$
$$+ x_n'(\boldsymbol{\alpha_{1n} e_1} + \boldsymbol{\alpha_{2n} e_2} + \cdots + \boldsymbol{\alpha_{nn} e_n}) =$$

$e_1, \ldots, e_n$ linearly independent -> coefficients in front of them should be the same on the both sides of the equality:

girafe ai

Coordinate Change

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n = x'_1 \boldsymbol{e'_1} + x'_2 \boldsymbol{e'_2} + \cdots x'_n \boldsymbol{e'_n} =$$

Remember: $e'_i = \boldsymbol{\alpha_{1i}} e_1 + \boldsymbol{\alpha_{2i}} e_2 + \cdots + \boldsymbol{\alpha_{ni}} e_n$

$$= x'_1 \cdot (\boldsymbol{\alpha_{11} e_1} + \boldsymbol{\alpha_{21} e_2} + \cdots + \boldsymbol{\alpha_{n1} e_n}) + \cdots + x'_i \cdot (\boldsymbol{\alpha_{1i} e_1} + \boldsymbol{\alpha_{2i} e_2} + \cdots + \boldsymbol{\alpha_{ni} e_n}) + \cdots$$
$$+ x'_n (\boldsymbol{\alpha_{1n} e_1} + \boldsymbol{\alpha_{2n} e_2} + \cdots + \boldsymbol{\alpha_{nn} e_n}) =$$

$e_1, \ldots, e_n$ linearly independent -> coefficients in front of them should be the same on the both sides of the equality:

$$\begin{aligned} x_1 &= x'_1 \alpha_{11} + \cdots + x'_i \alpha_{1i} + \cdots + x'_n \alpha_{1n} \\ x_2 &= x'_1 \alpha_{21} + \cdots + x'_i \alpha_{2i} + \cdots + x'_n \alpha_{2n} \\ &\vdots \\ x_n &= x'_1 \alpha_{n1} + \cdots + x'_i \alpha_{ni} + \cdots + x'_n \alpha_{nn} \end{aligned}$$

Coordinate Change

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n = x'_1 \boldsymbol{e'_1} + x'_2 \boldsymbol{e'_2} + \cdots x'_n \boldsymbol{e'_n} =$$

Remember: $e'_i = \boldsymbol{\alpha_{1i}} e_1 + \boldsymbol{\alpha_{2i}} e_2 + \cdots + \boldsymbol{\alpha_{ni}} e_n$

$$= x'_1 \cdot (\boldsymbol{\alpha_{11} e_1} + \boldsymbol{\alpha_{21} e_2} + \cdots + \boldsymbol{\alpha_{n1} e_n}) + \cdots + x'_i \cdot (\boldsymbol{\alpha_{1i} e_1} + \boldsymbol{\alpha_{2i} e_2} + \cdots + \boldsymbol{\alpha_{ni} e_n}) + \cdots$$
$$+ x'_n (\boldsymbol{\alpha_{1n} e_1} + \boldsymbol{\alpha_{2n} e_2} + \cdots + \boldsymbol{\alpha_{nn} e_n}) =$$

$$\boldsymbol{x_{old}} \quad \boxed{\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}} \begin{array}{l} = x'_1 \alpha_{11} + \cdots + x'_i \alpha_{1i} + \cdots + x'_n \alpha_{1n} \\ = x'_1 \alpha_{21} + \cdots + x'_i \alpha_{2i} + \cdots + x'_n \alpha_{2n} \\ \qquad\qquad\qquad \vdots \\ = x'_1 \alpha_{n1} + \cdots + x'_i \alpha_{ni} + \cdots + x'_n \alpha_{nn} \end{array}$$

Coordinate Change

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n = x'_1 \boldsymbol{e'_1} + x'_2 \boldsymbol{e'_2} + \cdots x'_n \boldsymbol{e'_n} =$$

Remember: $e'_i = \boldsymbol{\alpha_{1i}} e_1 + \boldsymbol{\alpha_{2i}} e_2 + \cdots + \boldsymbol{\alpha_{ni}} e_n$

$$= x'_1 \cdot (\boldsymbol{\alpha_{11} e_1} + \boldsymbol{\alpha_{21} e_2} + \cdots + \boldsymbol{\alpha_{n1} e_n}) + \cdots + x'_i \cdot (\boldsymbol{\alpha_{1i} e_1} + \boldsymbol{\alpha_{2i} e_2} + \cdots + \boldsymbol{\alpha_{ni} e_n}) + \cdots$$
$$+ x'_n (\boldsymbol{\alpha_{1n} e_1} + \boldsymbol{\alpha_{2n} e_2} + \cdots + \boldsymbol{\alpha_{nn} e_n}) =$$

$$\boldsymbol{x_{old}} \quad \begin{array}{l} \boxed{x_1} = \boxed{x'_1}\alpha_{11} + \cdots + \boxed{x'_i}\alpha_{1i} + \cdots + \boxed{x'_n}\alpha_{1n} \\ x_2 = x'_1\alpha_{21} + \cdots + x'_i\alpha_{2i} + \cdots + x'_n\alpha_{2n} \\ \vdots \\ x_n = x'_1\alpha_{n1} + \cdots + x'_i\alpha_{ni} + \cdots + x'_n\alpha_{nn} \end{array}$$

$\boldsymbol{x_{new}}$

Coordinate Change

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n = x'_1 \boldsymbol{e'_1} + x'_2 \boldsymbol{e'_2} + \cdots x'_n \boldsymbol{e'_n} =$$

Remember: $e'_i = \boldsymbol{\alpha_{1i}} e_1 + \boldsymbol{\alpha_{2i}} e_2 + \cdots + \boldsymbol{\alpha_{ni}} e_n$

$$= x'_1 \cdot (\boldsymbol{\alpha_{11} e_1} + \boldsymbol{\alpha_{21} e_2} + \cdots + \boldsymbol{\alpha_{n1} e_n}) + \cdots + x'_i \cdot (\boldsymbol{\alpha_{1i} e_1} + \boldsymbol{\alpha_{2i} e_2} + \cdots + \boldsymbol{\alpha_{ni} e_n}) + \cdots$$
$$+ x'_n (\boldsymbol{\alpha_{1n} e_1} + \boldsymbol{\alpha_{2n} e_2} + \cdots + \boldsymbol{\alpha_{nn} e_n}) =$$

$\boldsymbol{e'_i}$

$$\boldsymbol{x_{old}} \quad \begin{matrix} x_1 = x'_1 \alpha_{11} + \cdots + x'_i \alpha_{1i} + \cdots + x'_n \alpha_{1n} \\ x_2 = x'_1 \alpha_{21} + \cdots + x'_i \alpha_{2i} + \cdots + x'_n \alpha_{2n} \\ \vdots \\ x_n = x'_1 \alpha_{n1} + \cdots + x'_i \alpha_{ni} + \cdots + x'_n \alpha_{nn} \end{matrix}$$

$\boldsymbol{x_{new}}$

Coordinate Change: Example

- Consider $\mathbb{R}^2$ with basis $e_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $e_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$
- New basis: $e'_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $e_2' = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$
- $x_{old} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $x_{new} = \begin{bmatrix} x_1' \\ x_2' \end{bmatrix}$

girafe ai

Coordinate Change: Example

- Consider $\mathbb{R}^2$ with basis $e_1 = \begin{bmatrix}1\\1\end{bmatrix}$, $e_2 = \begin{bmatrix}1\\-1\end{bmatrix}$
- New basis: $e'_1 = \begin{bmatrix}2\\1\end{bmatrix}$, $e'_2 = \begin{bmatrix}-1\\-1\end{bmatrix}$
- $x_{old} = \begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}2\\-1\end{bmatrix}$, $x_{new} = \begin{bmatrix}x'_1\\x'_2\end{bmatrix}$

e'_i

$$x_{old} \qquad \begin{aligned} x_1 &= x'_1\alpha_{11} + \cdots + x'_i\alpha_{1i} + \cdots + x'_n\alpha_{1n} \\ x_2 &= x'_1\alpha_{21} + \cdots + x'_i\alpha_{2i} + \cdots + x'_n\alpha_{2n} \\ &\;\;\vdots \\ x_n &= x'_1\alpha_{n1} + \cdots + x'_i\alpha_{ni} + \cdots + x'_n\alpha_{nn} \end{aligned}$$

x_{new}

Coordinate Change: Example

- Consider $\mathbb{R}^2$ with basis $e_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},\ e_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$
- New basis: $e'_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix},\ e'_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$
- $x_{old} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix},\ \ x_{new} = \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix}$

e'_i

x_{old}

$$\begin{aligned} x_1 &= x'_1\alpha_{11} + \cdots + x'_i\alpha_{1i} + \cdots + x'_n\alpha_{1n} \\ x_2 &= x'_1\alpha_{21} + \cdots + x'_i\alpha_{2i} + \cdots + x'_n\alpha_{2n} \\ &\ \ \vdots \\ x_n &= x'_1\alpha_{n1} + \cdots + x'_i\alpha_{ni} + \cdots + x'_n\alpha_{nn} \end{aligned}$$

x_{new}

$$\begin{aligned} 2 &= 2x'_1 - 1x'_2 \\ -1 &= 1x'_1 - 1x'_2 \end{aligned}$$

Coordinate Change: Example

- Consider $\mathbb{R}^2$ with basis $e_1 = \begin{bmatrix}1\\1\end{bmatrix}$, $e_2 = \begin{bmatrix}1\\-1\end{bmatrix}$
- New basis: $e'_1 = \begin{bmatrix}2\\1\end{bmatrix}$, $e'_2 = \begin{bmatrix}-1\\-1\end{bmatrix}$
- $x_{old} = \begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}2\\-1\end{bmatrix}$, $x_{new} = \begin{bmatrix}x'_1\\x'_2\end{bmatrix}$

$\boldsymbol{e'_i}$

$\boldsymbol{x_{old}}$

$$\begin{matrix} x_1 = x'_1\alpha_{11} + \cdots + x'_i\alpha_{1i} + \cdots + x'_n\alpha_{1n} \\ x_2 = x'_1\alpha_{21} + \cdots + x'_i\alpha_{2i} + \cdots + x'_n\alpha_{2n} \\ \vdots \\ x_n = x'_1\alpha_{n1} + \cdots + x'_i\alpha_{ni} + \cdots + x'_n\alpha_{nn} \end{matrix}$$

$\boldsymbol{x_{new}}$

$$\begin{matrix} 2 = 2x'_1 - 1x'_2 \\ -1 = 1x'_1 - 1x'_2 \end{matrix} \iff \begin{matrix} x'_1 = 3 \\ x'_2 = 4 \end{matrix}$$

Coordinate Change: Example

- Consider $\mathbb{R}^2$ with basis $e_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $e_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$
- New basis: $e'_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $e'_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$
- $x_{old} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $x_{new} = \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix}$

$$e'_i$$

$$x_{old} \quad \begin{aligned} x_1 &= x'_1\alpha_{11} + \cdots + x'_i\alpha_{1i} + \cdots + x'_n\alpha_{1n} \\ x_2 &= x'_1\alpha_{21} + \cdots + x'_i\alpha_{2i} + \cdots + x'_n\alpha_{2n} \\ &\vdots \\ x_n &= x'_1\alpha_{n1} + \cdots + x'_i\alpha_{ni} + \cdots + x'_n\alpha_{nn} \end{aligned}$$

$$x_{new}$$

$$\begin{aligned} 2 &= 2x'_1 - 1x'_2 \\ -1 &= 1x'_1 - 1x'_2 \end{aligned} \iff \begin{aligned} x'_1 &= 3 \\ x'_2 &= 4 \end{aligned} \iff x_{new} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

Coordinate Change

- Going from one basis to the other:

$$\boldsymbol{e}'_i$$

$$\boldsymbol{x}_{old} \quad \begin{array}{l} x_1 = x'_1\alpha_{11} + \cdots + x'_i\alpha_{1i} + \cdots + x'_n\alpha_{1n} \\ x_2 = x'_1\alpha_{21} + \cdots + x'_i\alpha_{2i} + \cdots + x'_n\alpha_{2n} \\ \qquad\qquad\qquad\quad \vdots \\ x_n = x'_1\alpha_{n1} + \cdots + x'_i\alpha_{ni} + \cdots + x'_n\alpha_{nn} \end{array}$$

$$\boldsymbol{x}_{new}$$

Coordinate Change

- Going from one basis to the other:

$$\boldsymbol{e}'_{\boldsymbol{i}}$$

$$\boldsymbol{x_{old}} \quad \begin{aligned} x_1 &= x'_1\alpha_{11} + \cdots + x'_i\alpha_{1i} + \cdots + x'_n\alpha_{1n} \\ x_2 &= x'_1\alpha_{21} + \cdots + x'_i\alpha_{2i} + \cdots + x'_n\alpha_{2n} \\ &\ \ \vdots \\ x_n &= x'_1\alpha_{n1} + \cdots + x'_i\alpha_{ni} + \cdots + x'_n\alpha_{nn} \end{aligned}$$

$$\boldsymbol{x_{new}}$$

- There is a more compact way of writing this down using matrices.

Matrices

A Matrix

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

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

A Matrix

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

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

- *Examples:*

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

Special Matrices

- Diagonal matrix: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ $\left(a_{ii} \neq 0,\ a_{ij} = 0\ \forall i \neq j\right)$

Special Matrices

- Diagonal matrix: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ $\left(a_{ii} \neq 0,\ a_{ij} = 0\ \forall i \neq j\right)$
- Identity matrix: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $\left(a_{ii} = 1,\ a_{ij} = 0\ \forall i \neq j\right)$

Special Matrices

- Diagonal matrix: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ $(a_{ii} \neq 0,\ a_{ij} = 0\ \forall i \neq j)$
- Identity matrix: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $(a_{ii} = 1,\ a_{ij} = 0\ \forall i \neq j)$
- Symmetric matrix: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ $(a_{ij} = a_{ji})$

Special Matrices

- Diagonal matrix: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ $\left(a_{ii} \neq 0,\ a_{ij} = 0\ \forall i \neq j\right)$
- Identity matrix: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $\left(a_{ii} = 1,\ a_{ij} = 0\ \forall i \neq j\right)$
- Symmetric matrix: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ $\left(a_{ij} = a_{ji}\right)$
- Triangular matrix: $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{bmatrix}$ $\left(a_{ij} = 0\ \forall i > j \ or\ \ \forall i < j\right)$

Vectors vs Matrices

- An n-dimensional vector can be considered a $n \times 1$ matrix:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n \times 1}$$

Operations with Matrices

Transpose of a Matrix

- Consider a matrix

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

Transpose of a Matrix

- Consider a matrix

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

- Transpose = writing columns as rows:

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}, \qquad \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}^T = [x_1, \ldots, x_n]$$

Transpose of a Matrix: Example

- $\begin{bmatrix} 0 & 1 & 2 \\ 2 & 1 & 5 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 0 & 2 \\ 1 & 1 \\ 2 & 5 \end{bmatrix}$

girafe ai

Transpose of a Matrix: Example

- $$\begin{bmatrix} 0 & 1 & 2 \\ 2 & 1 & 5 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 0 & 2 \\ 1 & 1 \\ 2 & 5 \end{bmatrix}$$

- Transposing a symmetrical matrix = no changes:

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{bmatrix}^{T} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{bmatrix}$$

Multiplying by a Scalar

- We can multiply matrix by a scalar:

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

Multiplying by a Scalar

- We can multiply matrix by a scalar:

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

- Example:

$$5 \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$$

Sum of Two Matrices

- We can sum up matrices of the same size:

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

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

Sum of Two Matrices

- We can sum up matrices of the same size:

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

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

- Example:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 0 & 4 \end{bmatrix}$$

Matrices Also Form a Vector Space!

- $(\mathbb{R}^{m \times n}, +, \cdot)$ - a vector space.
  "Vectors" = matrices.

Matrices Also Form a Vector Space!

- $(\mathbb{R}^{m \times n}, +, \ \cdot)$ - a vector space.
  "Vectors" = matrices.

- You can check yourself that the necessary axioms hold.

Matrix Multiplication

- Consider two matrices $A = \{a_{ij}\}_{m\times n}$ and $b = \{b_{ij}\}_{n\times p}$.
- $C = AB$ – product of two matrices.

Matrix Multiplication

- Consider two matrices $A = \{a_{ij}\}_{m \times n}$ and $b = \{b_{ij}\}_{n \times p}$.
- $C = AB$ – product of two matrices.
- Each element c_{ij} of C is a dot product of i-th row of A and j-th column of B: $C = \{c_{ij}\}_{m \times p} = \{(A_i, B^j)\}_{m \times p}$

girafe ai

Matrix Multiplication

- Consider two matrices $A = \{a_{ij}\}_{m \times n}$ and $b = \{b_{ij}\}_{n \times p}$.
- $C = AB$ – product of two matrices.
- Each element c_{ij} of C is a dot product of i-th row of A and j-th column of B: $C = \{c_{ij}\}_{m \times p} = \{(A_i, B^j)\}_{m \times p}$
- Example $\mathbb{R}^{2\times 2}$: $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$

Matrix Multiplication: Example

$$\begin{bmatrix} 0 & 1 & 2 \\ 2 & 1 & 5 \end{bmatrix} \begin{bmatrix} 3 & 6 & 4 \\ 2 & 5 & 8 \\ 7 & 1 & 9 \end{bmatrix} =$$

Matrix Multiplication: Example

$$\begin{bmatrix} 0 & 1 & 2 \\ 2 & 1 & 5 \end{bmatrix} \begin{bmatrix} 3 & 6 & 4 \\ 2 & 5 & 8 \\ 7 & 1 & 9 \end{bmatrix} =$$

$$= \begin{bmatrix} 0+2+14 & 0+5+2 & 0+8+18 \\ 6+2+35 & 12+5+5 & 8+8+45 \end{bmatrix} =$$

Matrix Multiplication: Example

$$\begin{bmatrix} 0 & 1 & 2 \\ 2 & 1 & 5 \end{bmatrix} \begin{bmatrix} 3 & 6 & 4 \\ 2 & 5 & 8 \\ 7 & 1 & 9 \end{bmatrix} =$$

$$= \begin{bmatrix} 0 + 2 + 14 & 0 + 5 + 2 & 0 + 8 + 18 \\ 6 + 2 + 35 & 12 + 5 + 5 & 8 + 8 + 45 \end{bmatrix} =$$

$$= \begin{bmatrix} 16 & 7 & 26 \\ 43 & 22 & 61 \end{bmatrix}$$

To sum up

- Vectors
    - Vector spaces
    - Inner products
    - Lengths
    - Distances
    - Angles
- Analytic Geometry
    - Projections
    - Hyperplanes
    - Normal vector
- Vector spaces
    - Linear (in)dependence
    - Basis
- Matrices
    - Matrix operations

Next Time

- Calculus recap
- Probability theory recap