Lectures 8-10

Linear algebra: Dot products and Projections

DSC 40A, Fall 2024

Agenda

- Recap: Simple linear regression and correlation.
- Connections to related models.
- Dot products.
- Spans and projections.

Dot product: coordinate definition

both vectors need to have same number of elements
cannot perform $\begin{bmatrix}1\\2\end{bmatrix} \cdot \begin{bmatrix}4\\5\\7\end{bmatrix}$

- The **dot product** of two vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$ is written as:

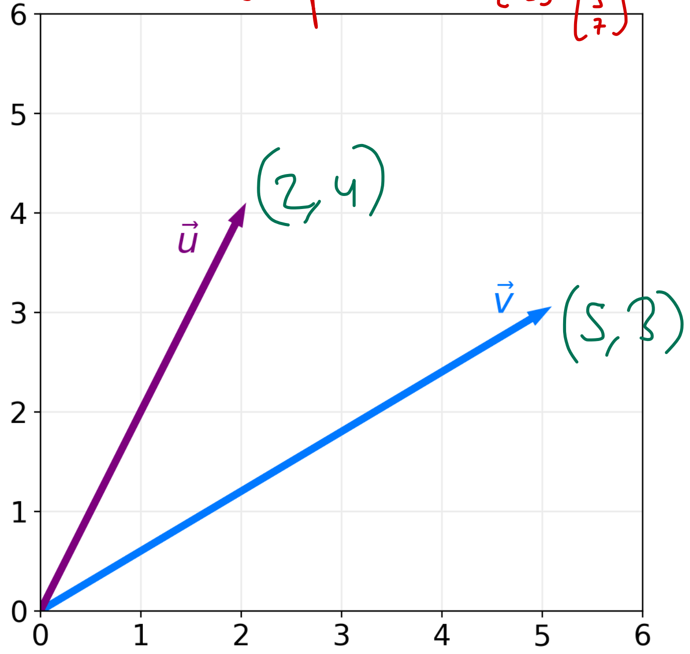
\cdot (latex)

$$\vec{u} \cdot \vec{v} = \vec{u}^\intercal \vec{v}$$

- The computational definition of the dot product:

$$\vec{u} \cdot \vec{v} = \sum_{i=1}^{n} u_i v_i = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$$

- The result is a **scalar**, i.e. a single number.



$\vec{u} \cdot \vec{v} = 2 \cdot 5 + 4 \cdot 3 = 10 + 12 = 22 \in \mathbb{R}$ ← scalar

$\vec{u}^T \vec{v} = \begin{bmatrix} 2 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \end{bmatrix} = 10 + 12 = 22$

$f(\vec{u}, \vec{v}) \in \mathbb{R}$

$\mathbb{R}^n$ $\mathbb{R}^n$

Dot product: geometric definition

- The computational definition of the dot product:

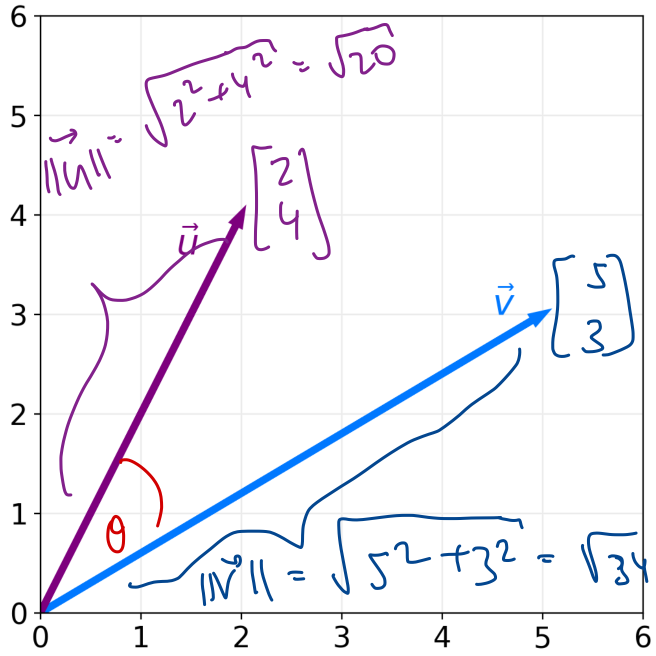
$$\vec{u} \cdot \vec{v} = \sum_{i=1}^{n} u_i v_i = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$$

- The geometric definition of the dot product:

$$\vec{u} \cdot \vec{v} = \|\vec{u}\| \|\vec{v}\| \cos\theta$$

where θ is the angle between $\vec{u}$ and $\vec{v}$.

- The two definitions are equivalent! This equivalence allows us to find the angle θ between two vectors.



$\vec{u} \cdot \vec{v} = 22$ (from previous slide)

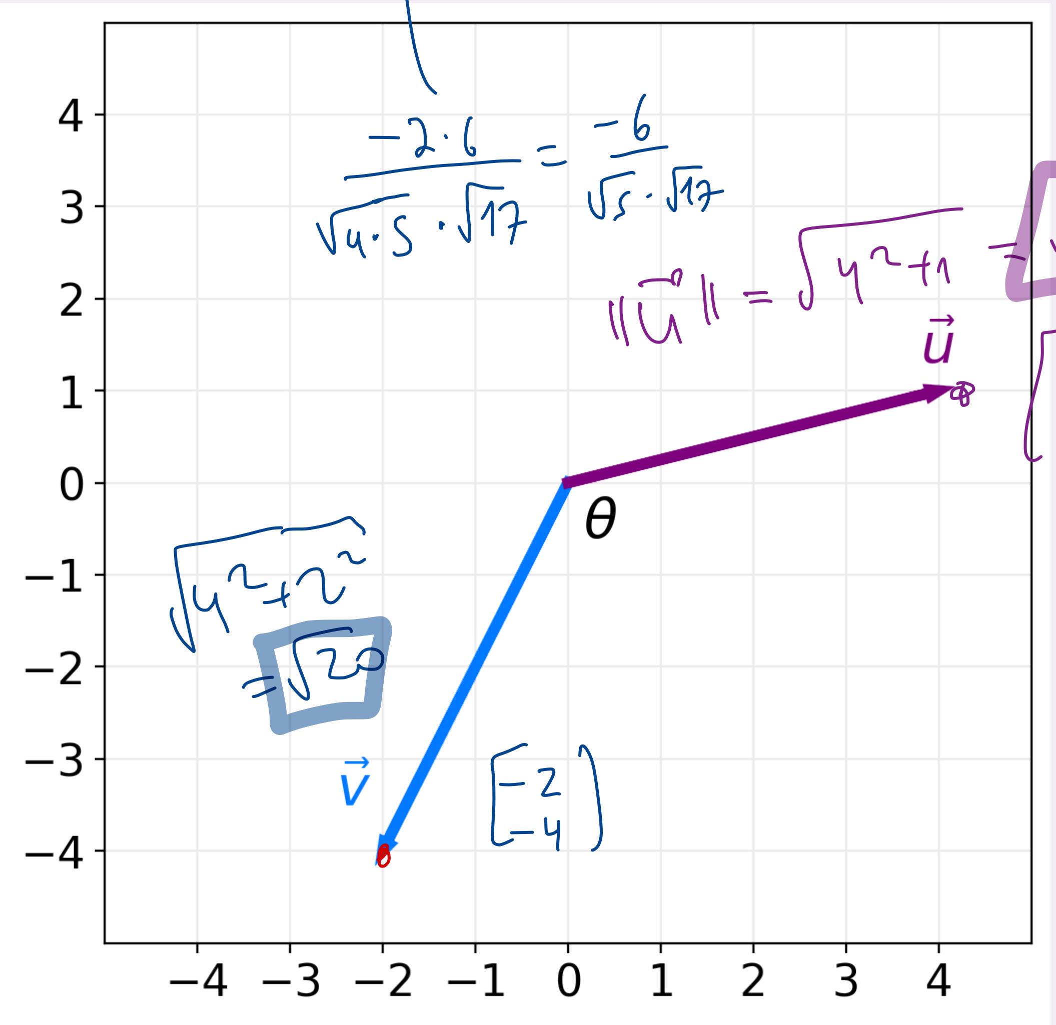
$$\cos\theta = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \|\vec{v}\|} = \frac{22}{\sqrt{20}\sqrt{34}} = \frac{11}{\sqrt{5} \cdot \sqrt{34}} = \frac{11}{\sqrt{170}}$$

Question 4: $\cos\theta$

What is $\cos\theta$?

- A. $\frac{6}{\sqrt{85}}$
- B. $\frac{-6}{\sqrt{85}}$
- C. $\frac{-3}{85}$
- D. $\frac{-2}{3}$

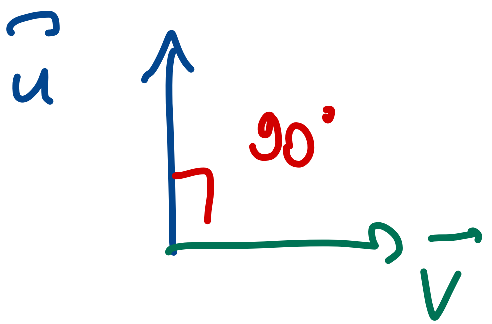
$$\cos\theta = \frac{-12}{\sqrt{20}\,\sqrt{17}} = \frac{-6}{\sqrt{5}\,\sqrt{17}} = \frac{-6}{\sqrt{85}}$$



$$\vec{u}\cdot\vec{v} = \|\vec{u}\|\cdot\|\vec{v}\|\cdot\cos\theta$$

$$\cos\theta = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}$$

$$\vec{u}\cdot\vec{v} = \vec{u}^T\vec{v} = \begin{bmatrix}4 & 1\end{bmatrix}\begin{bmatrix}-2\\-4\end{bmatrix} = -8-4 = -12$$



right angle
perpendicular

Orthogonal vectors

- Recall: $\cos 90^{\text{o}} = 0$.
- Since $\vec{u}\cdot\vec{v} = \|\vec{u}\|\,\|\vec{v}\|\cos\theta$, if the angle between two vectors is 90^{o} , their dot product is $\|\vec{u}\|\,\|\vec{v}\|\cos 90^{\text{o}} = 0$.
- If the angle between two vectors is 90^{o} , we say they are perpendicular, or more generally, **orthogonal**.
- **Key idea**:

$$\boxed{\text{two vectors are } \textbf{orthogonal} \iff \vec{u}\cdot\vec{v} = 0}$$

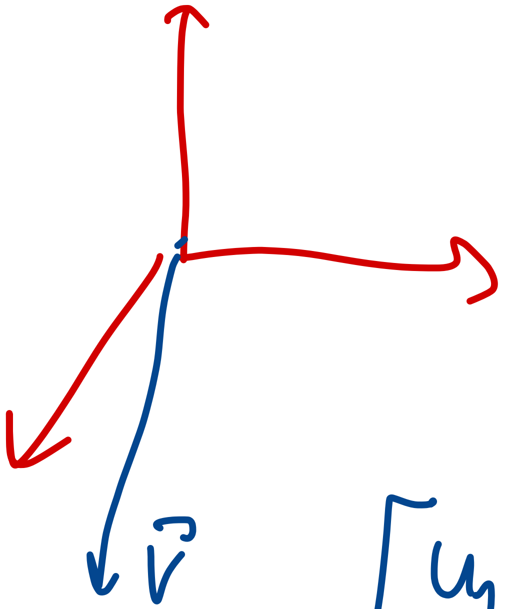
"if and only if"

Exercise

Find a non-zero vector in $\mathbb{R}^3$ orthogonal to:

$$\vec{v} = \begin{bmatrix} 2 \\ 5 \\ -8 \end{bmatrix}$$

Infinite possibilites



$$w = \begin{bmatrix} -5 \\ 2 \\ 0 \end{bmatrix} \qquad \vec{u}_i \cdot \vec{v} = 0$$

$$z = \begin{bmatrix} 0 \\ 8 \\ 5 \end{bmatrix}$$

$$\vec{u} \cdot \vec{v} = u_1 \cdot 2 + u_2 \cdot 5 + u_3 \cdot (-8) = 0$$

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

$$u_1 \cdot 2 = -5 \cdot u_2$$

Spans and projections

Adding and scaling vectors

- The sum of two vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$ is the **element-wise sum** of their components:

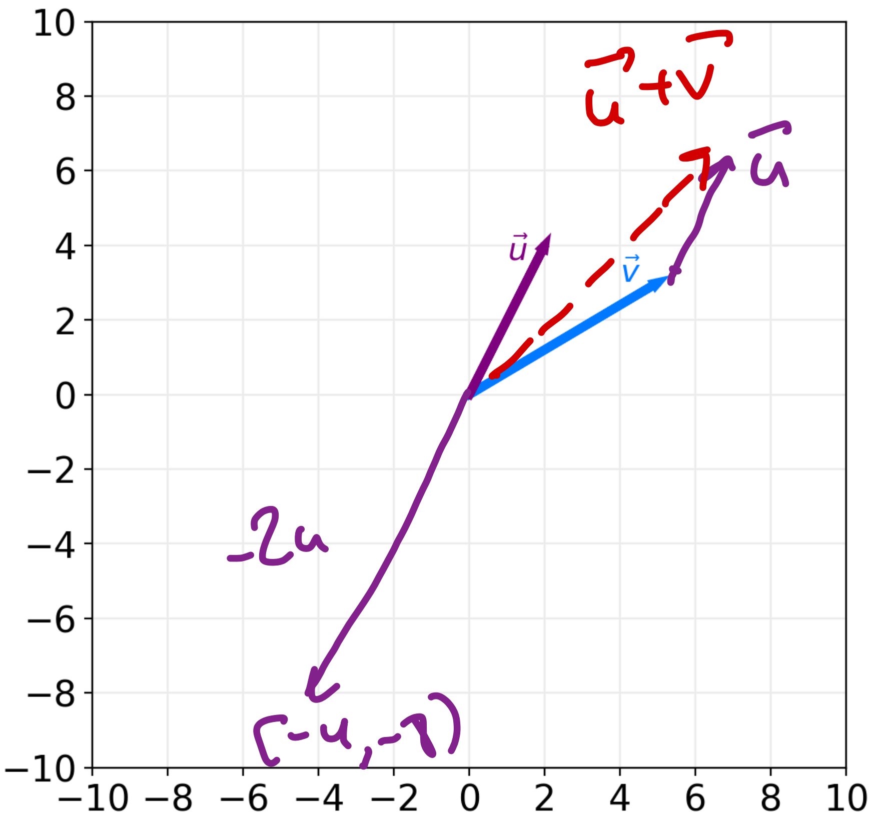
$$\vec{u} + \vec{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix} \in \mathbb{R}^n$$

- If c is a scalar, then:

$c \in \mathbb{R}$

$$c\vec{v} = \begin{bmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{bmatrix}$$

$\vec{u} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ $\quad -2\vec{u} = \begin{pmatrix} -4 \\ -8 \end{pmatrix}$



Linear combinations

Let $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_d$ all be vectors in $\mathbb{R}^n$.

A **linear combination** of $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_d$ is any vector of the form:

$$a_1\vec{v}_1 + a_2\vec{v}_2 + \ldots + a_d\vec{v}_d \in \mathbb{R}^n$$

where a_1, a_2, ..., a_d are all scalars.

$a_i \in \mathbb{R}$ for all $1 \leq i \leq d$

$\vec{v}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ $\vec{v}_2 = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$ $v_3 = \begin{bmatrix} 0 \\ 9 \end{bmatrix}$

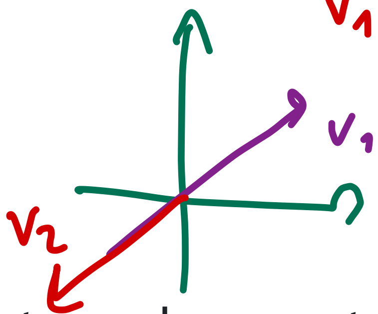
$v_1, v_2, v_3 \in \mathbb{R}^2$

$a_1\vec{v}_1 + a_2\vec{v}_2 + a_3\vec{v}_3 \in \mathbb{R}^2$

$a_1\vec{v}_1 + a_2\vec{v}_2 + a_3\vec{v}_3 = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$

$\vec{v}_1 = c\vec{v}_2$

span$\{\vec{v}_1, \vec{v}_2\}$ is one-dimensional



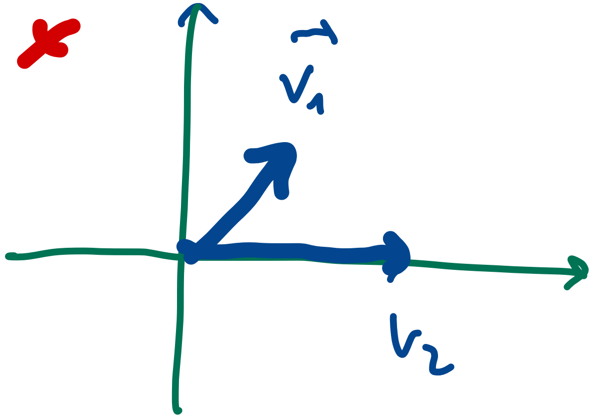
Span

- Let $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_d$ all be vectors in $\mathbb{R}^n$.
- The **span** of $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_d$ is the set of all vectors that can be created using linear combinations of those vectors.
- Formal definition:

$$\text{span}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_d) = \{a_1\vec{v}_1 + a_2\vec{v}_2 + \ldots + a_d\vec{v}_d : a_1, a_2, \ldots, a_n \in \mathbb{R}\}$$

Example

$\vec{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ $\vec{v}_2 = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$



span$\{\vec{v}_1, \vec{v}_2\}$ is all of $\mathbb{R}^2$

(the 2D plane)

Exercise

Let $\vec{v}_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$ and let $\vec{v}_2 = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$. Is $\vec{y} = \begin{bmatrix} 9 \\ 1 \end{bmatrix}$ in $\text{span}(\vec{v}_1, \vec{v}_2)$?

If so, write $\vec{y}$ as a linear combination of $\vec{v}_1$ and $\vec{v}_2$.

$\vec{v}_1 \neq c\,\vec{v}_2$ they "point" in different directions

$$w_1 \vec{v}_1 + w_2 \vec{v}_2 = \vec{y}$$

$$\begin{bmatrix} 2w_1 \\ -3w_1 \end{bmatrix} + \begin{bmatrix} -w_2 \\ 4w_2 \end{bmatrix} = \begin{bmatrix} 9 \\ 1 \end{bmatrix}$$

$$\Rightarrow \begin{matrix} 2w_1 - w_2 = 9 \\ -3w_1 + 4w_2 = 1 \end{matrix} \longrightarrow \text{solve for } w_1 \text{ and } w_2$$

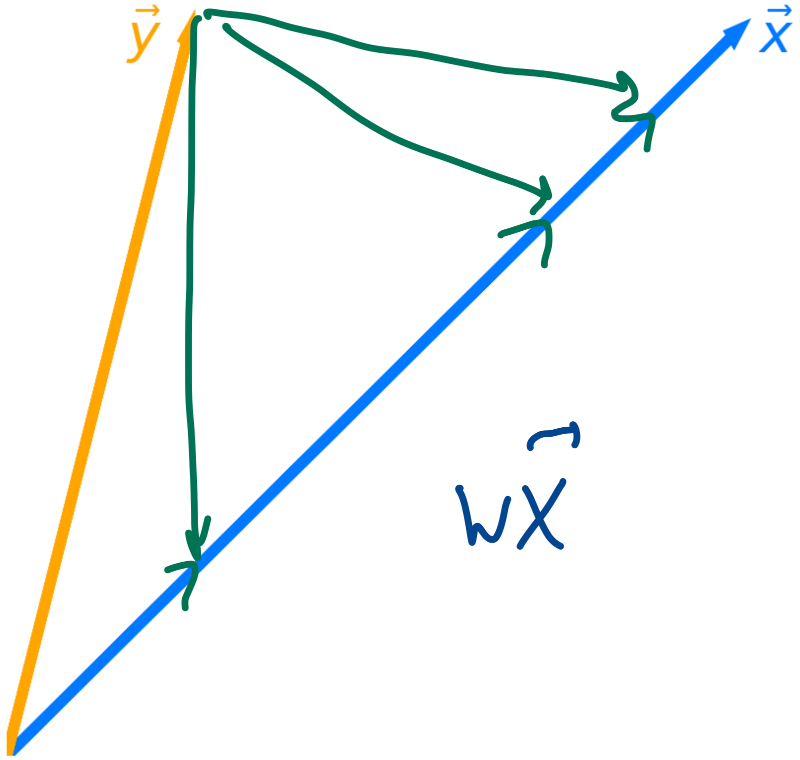
Projecting onto a single vector

- Let $\vec{x}$ and $\vec{y}$ be two vectors in $\mathbb{R}^n$.
- The span of $\vec{x}$ is the set of all vectors of the form:

$$w\vec{x}$$

where $w \in \mathbb{R}$ is a scalar.

- **Question**: What vector in $\text{span}(\vec{x})$ is closest to $\vec{y}$?
- The vector in $\text{span}(\vec{x})$ that is closest to $\vec{y}$ is the orthogonal **projection of $\vec{y}$ onto $\text{span}(\vec{x})$**.



Projection error

- Let $\vec{e} = \vec{y} - w\vec{x}$ be the **projection error**: that is, the vector that connects $\vec{y}$ to $\text{span}(\vec{x})$.
- **Goal: Find the w that makes $\vec{e}$ as short as possible.**
  - That is, minimize:

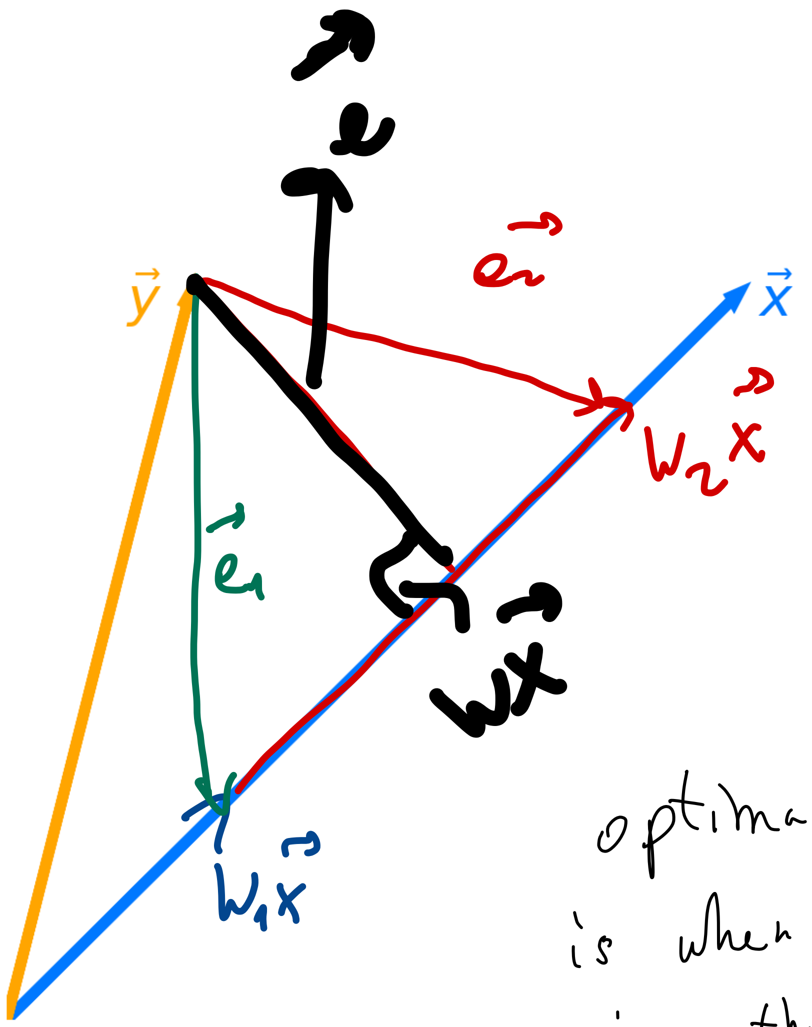
$$\|\vec{e}\| \rightarrow$$

length of the error

  - Equivalently, minimize:

$$\|\vec{y} - w\vec{x}\|$$

- **Idea**: To make $\vec{e}$ has short as possible, it should be **orthogonal to $w\vec{x}$**.

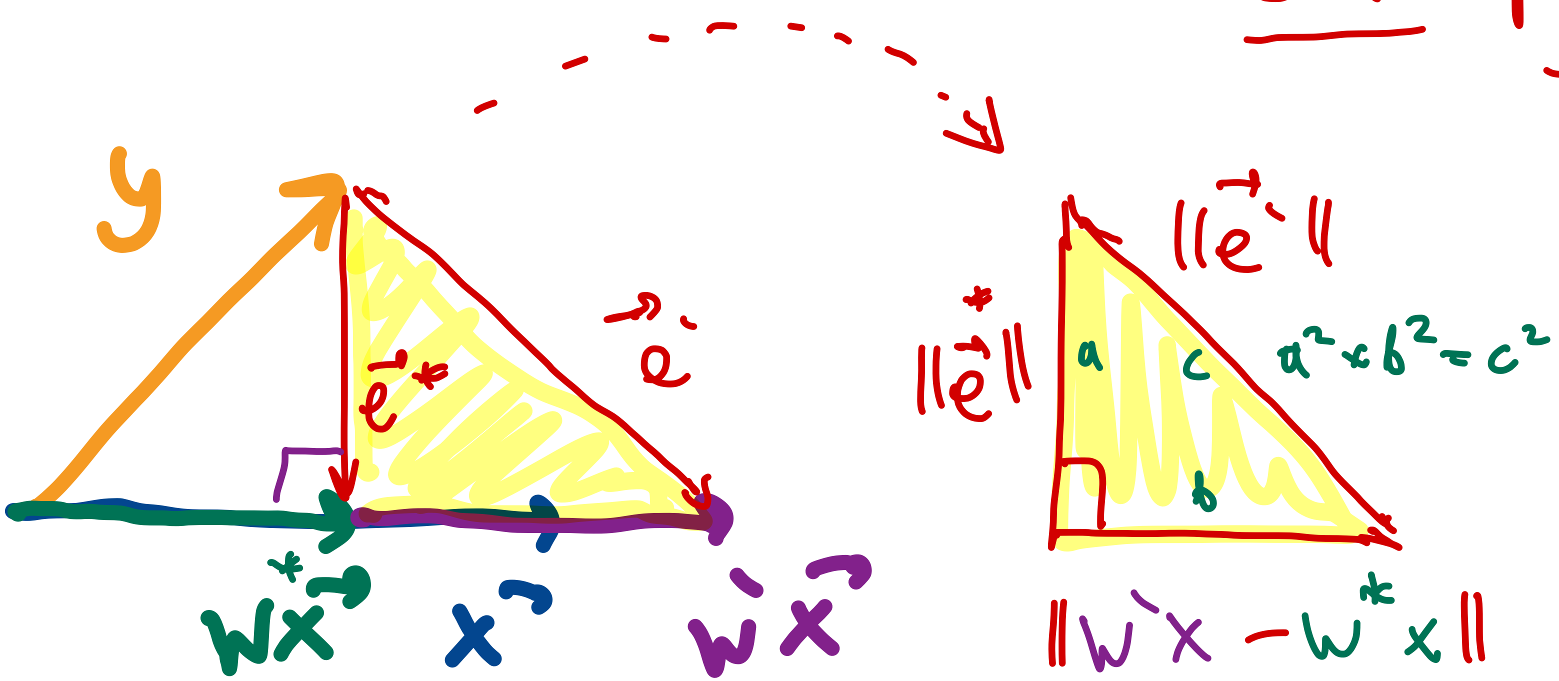


optimal w^* is when $\vec{e}$ is orthogonal to $\vec{x}$

Minimizing projection error

- **Goal: Find the w that makes $\vec{e} = \vec{y} - w\vec{x}$ as short as possible.**
- **Idea**: To make $\vec{e}$ as short as possible, it should be **orthogonal to** $w\vec{x}$.
- Can we prove that making $\vec{e}$ orthogonal to $w\vec{x}$ minimizes $\|\vec{e}\|$?

Goal: prove that $\vec{e}^*$ is the shortest possible error



Pythagorean theorem!

$$\|\vec{e}'\|^2 = \|\vec{e}^*\|^2 + \underbrace{\|w'\vec{x} - w^*\vec{x}\|^2}_{\geq 0}$$

$$\Rightarrow \|\vec{e}'\|^2 \geq \|\vec{e}^*\|^2$$

Minimizing projection error

- **Goal: Find the w that makes $\vec{e} = \vec{y} - w\vec{x}$ as short as possible.**
- Now we know that to minimize $\|\vec{e}\|$, $\vec{e}$ must be orthogonal to $w\vec{x}$.
- Given this fact, how can we solve for w?

$\vec{e}$ orthogonal to $\vec{x}$ $\quad \vec{e} \perp \vec{x} \Rightarrow w\vec{x} \cdot \vec{e} = 0$

$w\vec{x} \cdot (\vec{y} - w\vec{x}) = 0$ / (divide by w)

$\vec{x} \cdot (\vec{y} - w\vec{x}) = 0$

$\vec{x} \cdot \vec{y} - \vec{x} \cdot (w\vec{x}) = 0$

$\vec{x} \cdot \vec{y} = w(\vec{x} \cdot \vec{x})$

$$\Rightarrow w^* = \frac{\vec{x} \cdot \vec{y}}{\vec{x} \cdot \vec{x}} = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\|^2}$$

optimal w^* with smallest error

Orthogonal projection

- **Question**: What vector in $\text{span}(\vec{x})$ is closest to $\vec{y}$?
- **Answer**: It is the vector $w^* \vec{x}$, where:

$w^* \vec{x} \in \mathbb{R}^n$

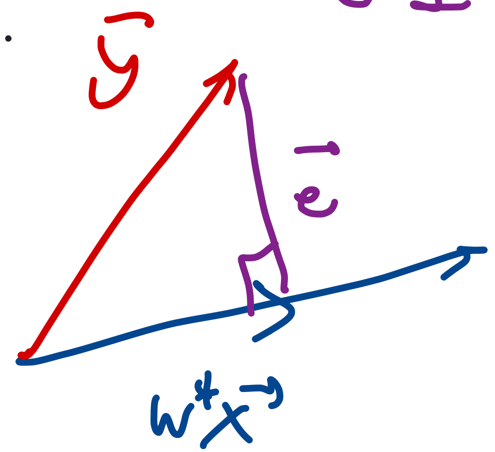
$$w^* = \frac{\vec{x} \cdot \vec{y}}{\vec{x} \cdot \vec{x}}$$

- Note that w^* is the solution to a minimization problem, specifically, this one:

$$\text{error}(w) = \|\vec{e}\| = \|\vec{y} - w\vec{x}\|$$

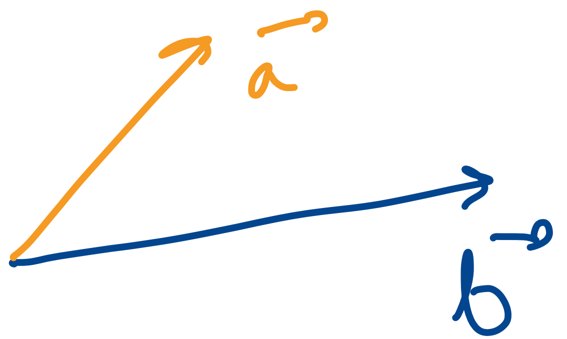
$\vec{e} \perp \vec{x}$

- We call $w^* \vec{x}$ the **orthogonal projection of** $\vec{y}$ **onto** $\text{span}(\vec{x})$.
    - Think of $w^* \vec{x}$ as the "shadow" of $\vec{y}$.



Exercise

Let $\vec{a} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} -1 \\ 9 \end{bmatrix}$.



What is the orthogonal projection of $\vec{a}$ onto $\text{span}(\vec{b})$?

Your answer should be of the form $w^*\vec{b}$, where w^* is a scalar.

$$w^* = \frac{\vec{b} \cdot \vec{a}}{\boxed{\vec{b} \cdot \vec{b}}} = \frac{(-1) \cdot 5 + (9) \cdot 2}{(-1)^2 + (9)^2} = \frac{-5+18}{1+81} = \frac{13}{82}$$

orthogonal projection of $\vec{a}$ onto the span$\{\vec{b}\}$ is $\frac{13}{82}\vec{b}$

ortho. proj. of b onto span$\{\vec{a}\}$

$= \frac{13}{29}\vec{a}$

$$w^* = \frac{\vec{b} \cdot \vec{a}}{\boxed{\vec{a} \cdot \vec{a}}} = \frac{13}{29}$$

Moving to multiple dimensions

- Let's now consider three vectors, $\vec{y}$, $\vec{x}^{(1)}$, and $\vec{x}^{(2)}$, all in $\mathbb{R}^n$.
- **Question**: What vector in $\text{span}(\vec{x}^{(1)}, \vec{x}^{(2)})$ is closest to $\vec{y}$?
  - Vectors in $\text{span}(\vec{x}^{(1)}, \vec{x}^{(2)})$ are of the form $w_1\vec{x}^{(1)} + w_2\vec{x}^{(2)}$, where w_1, $w_2 \in \mathbb{R}$ are scalars.
- Before trying to answer, let's watch 🎥 this animation that Jack, a previous tutor, made.