Orthogonal Sets

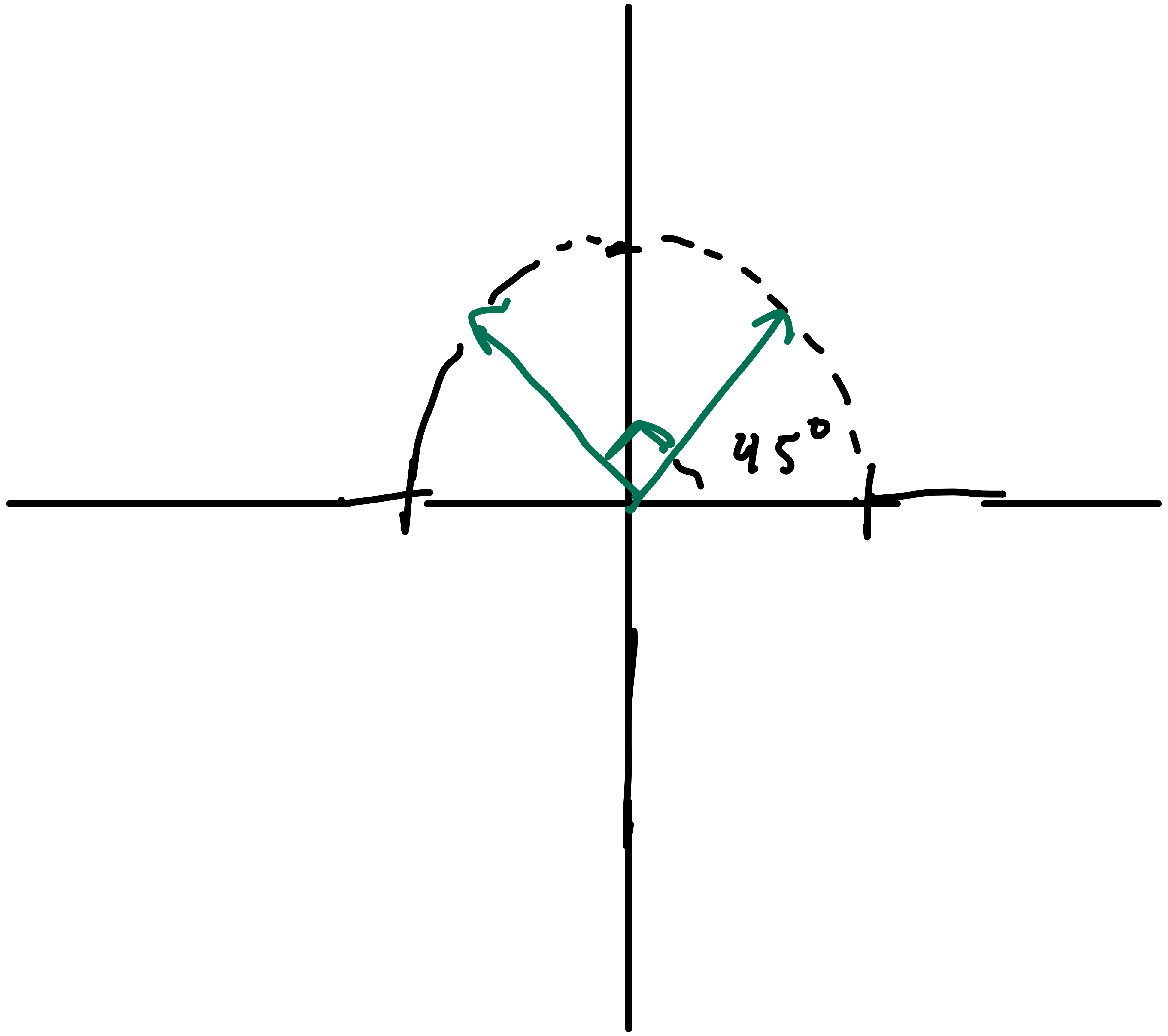
Geometric Algorithms
Lecture 22

CAS CS 132

Practice Problem

$$\mathscr{B} = \left(\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}, \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} \right)$$

Determine $\begin{bmatrix} 2 \\ 3 \end{bmatrix}_{\mathscr{B}}$



Answer

$$x_1 \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} + x_2 \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix}_{\mathscr{B}}$$

$$\mathscr{B} = \left(\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}, \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} \right)$$

$$\vec{x} \mapsto [\vec{x}]_{\mathcal{B}}$$

$$\vec{x} \mapsto A^{-1} \vec{x}$$

$$A = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$$

$$A^{-1} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2 \begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix} + 3 \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$$

$$\det A = \left(\frac{\sqrt{2}}{2}\right)^2 + \left(\frac{\sqrt{2}}{2}\right)^2 = \frac{1}{2} + \frac{1}{2} = 1$$

$$= \begin{bmatrix} \frac{5\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$$

Objectives

1. Recap analytic geometry in R^n
2. Try to understand why it is useful to work with orthogonal vectors
3. Get a sense of how to compute orthogonal vectors
4. Start to connect orthogonality to matrices and linear transformations

Keywords

orthogonal

orthogonal set

orthogonal basis

orthogonal projection

orthogonal component

orthonormal

orthonormal set

orthonormal basis

orthonormal matrix

orthogonal matrix

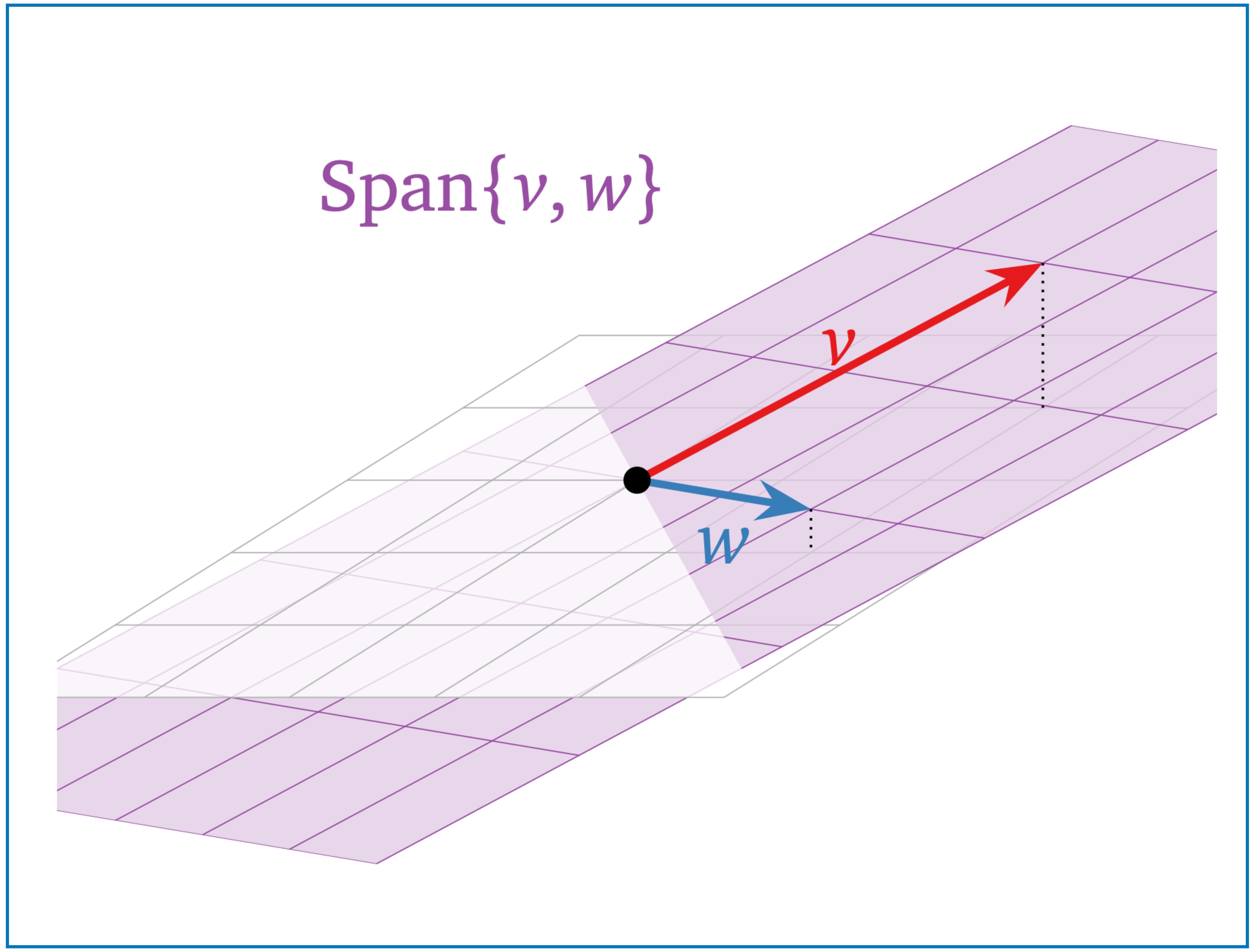
Recap: Analytic Geometry

Recall: The First Key Idea

Angles make sense in *any* dimension

Any pair of vectors in $\mathbb{R}^n$ span a (2D) plane

(We could formalize this via change of bases)



Recall: The Second Key Idea

All of the basic concepts of analytic geometry can be defined *in terms of inner products*

Spaces with inner products (like $\mathbb{R}^n$) are places where you can do analytic geometry

Recall: Inner Products

$$\begin{bmatrix} u_1 & u_2 & u_3 & u_4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = u_1v_1 + u_2v_2 + u_3v_3 + u_4v_4$$

Definition. The **inner product** of two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ is

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}$$

Recall: Inner Products

$$\begin{bmatrix} u_1 & u_2 & u_3 & u_4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = u_1v_1 + u_2v_2 + u_3v_3 + u_4v_4$$

Definition. The **inner product** of two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ is

a.k.a. dot product

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}$$

Recall: Norms and Inner Products

Definition. The ℓ^2 **norm** of a vector $\mathbf{v}$ in $\mathbb{R}^n$ is

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$$

The norm of a vector is the square root of the inner product with itself.

Recall: Norms and Inner Products

Definition. The ℓ^2 **norm** of a vector $\mathbf{v}$ in $\mathbb{R}^n$ is

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$$

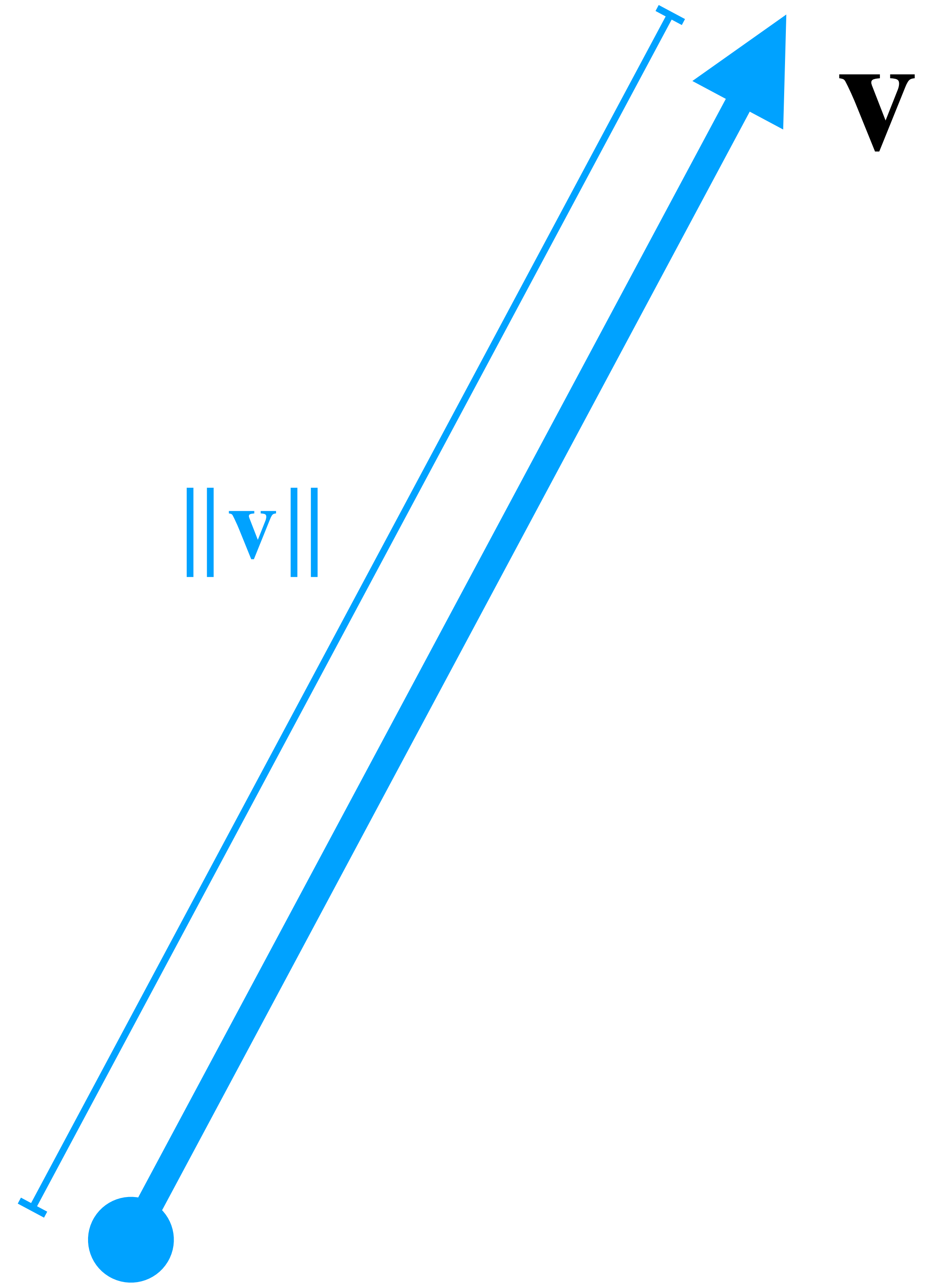
The norm of a vector is the square root of the inner product with itself.

It's important that $\mathbf{v}^T\mathbf{v}$ is nonnegative.

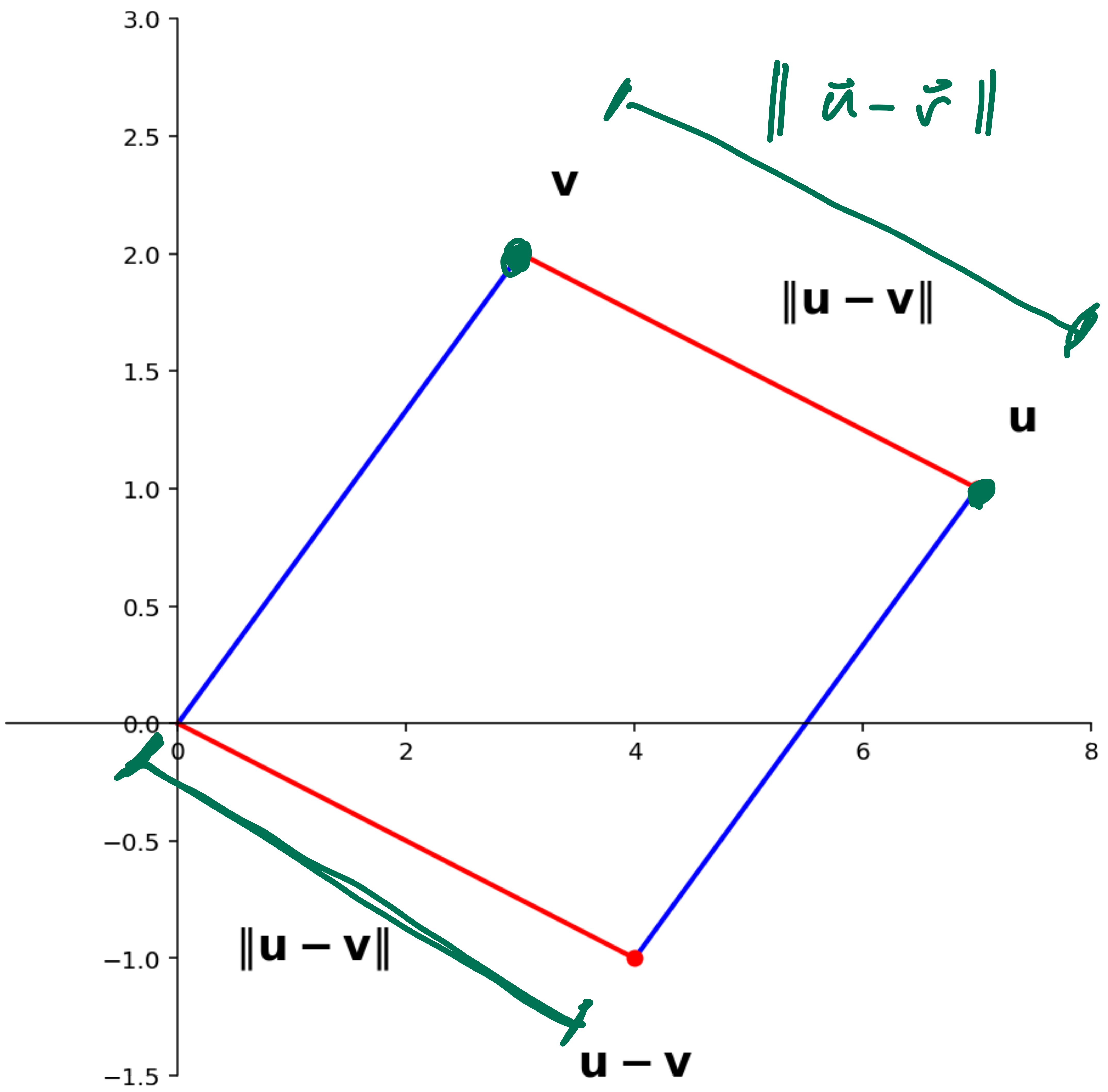
Recall: Norms and Length

Norms give us a notion of length.

In $\mathbb{R}^2$ and $\mathbb{R}^3$ this is our existing notion of length.



Recall: Distance (Pictorially)



$\vec{w}$ s.t. $\vec{v} + \vec{w} = \vec{u}$

so $\vec{w} = \vec{u} - \vec{v}$

$\|\vec{u} - \vec{v}\|$

$\vec{v} + (\vec{u} - \vec{v}) = \vec{u}$

Recall: Distance (Algebraically)

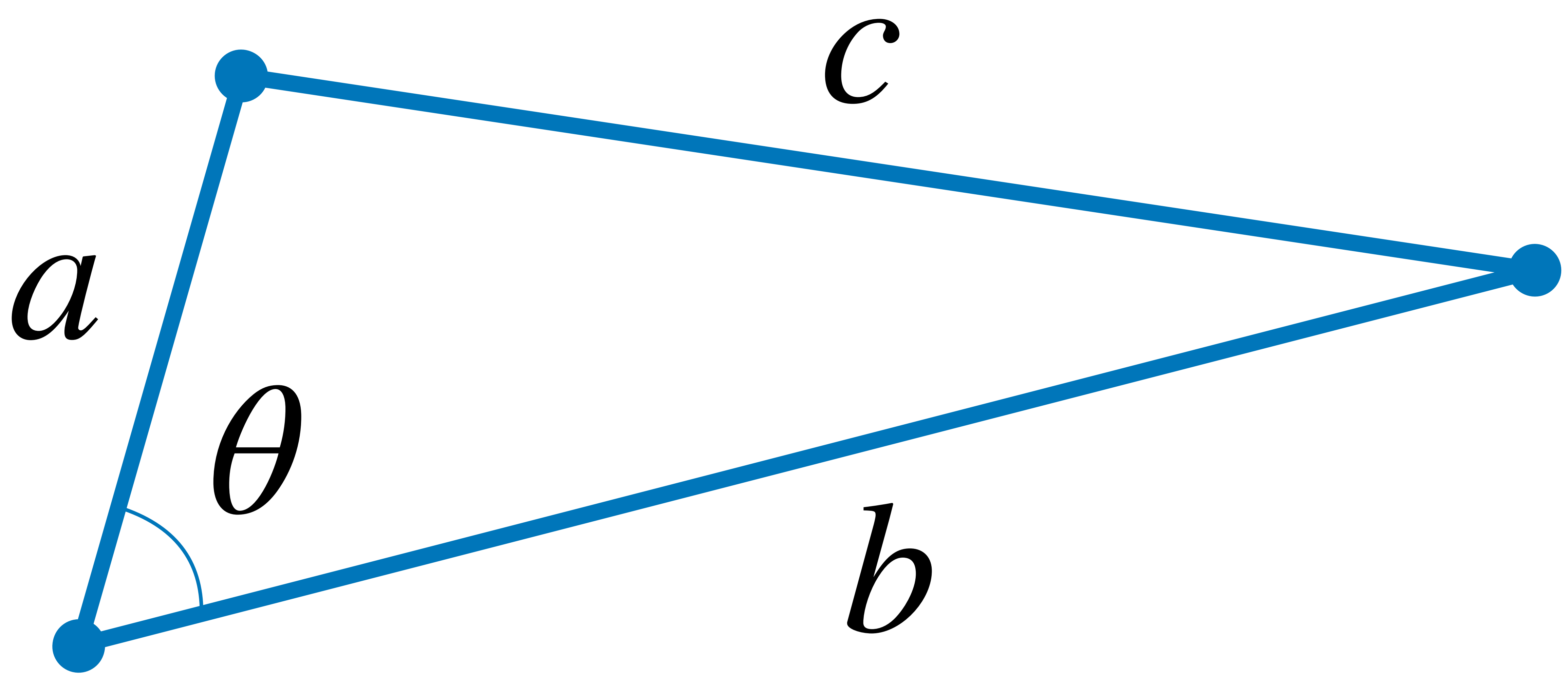
Definition. The distance between two points $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ is given by

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \| \vec{v} - \vec{u} \|$$

e.g., $\mathbf{u} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

$$\| v \| = \| - \vec{v} \|$$

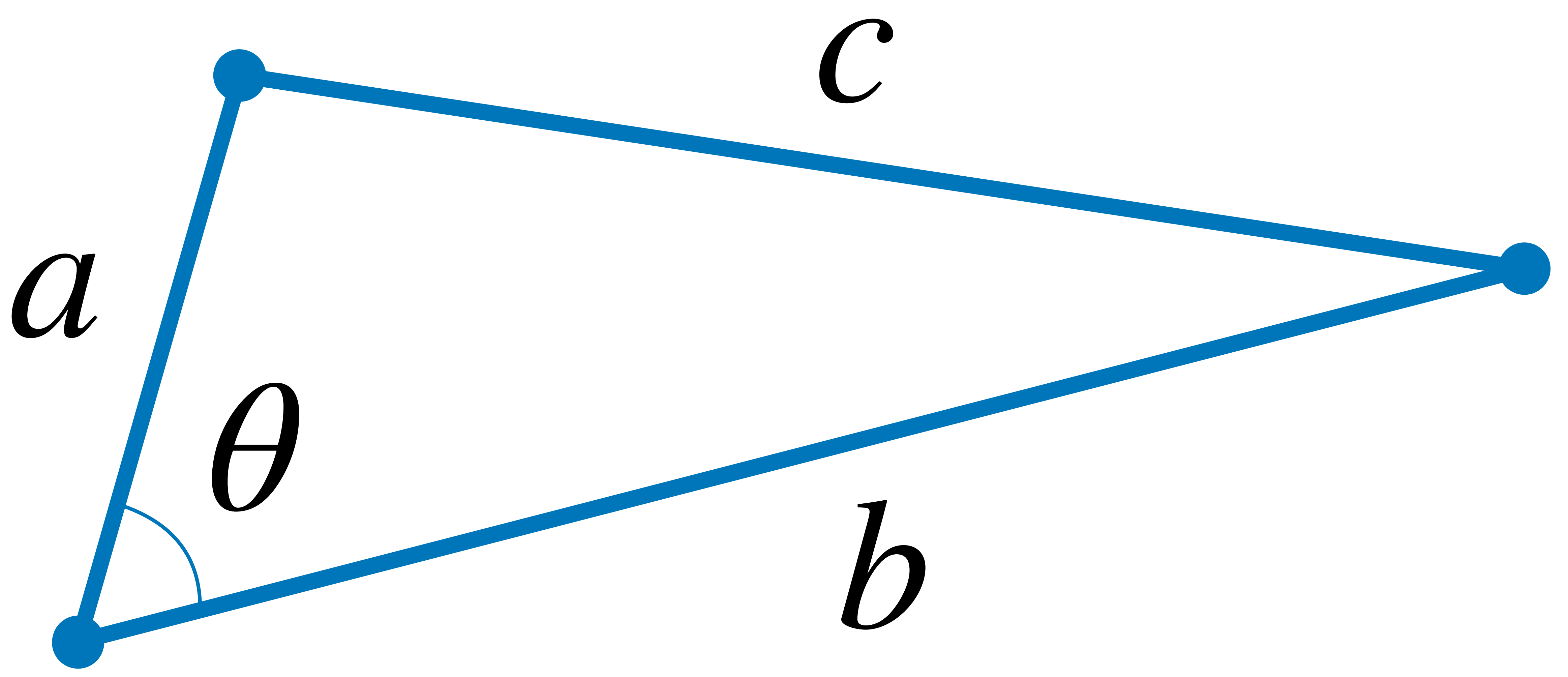
Recall: Law of Cosines



Theorem.

$$c^2 = a^2 + b^2 - 2ab\cos\theta$$

Recall: Law of Cosines

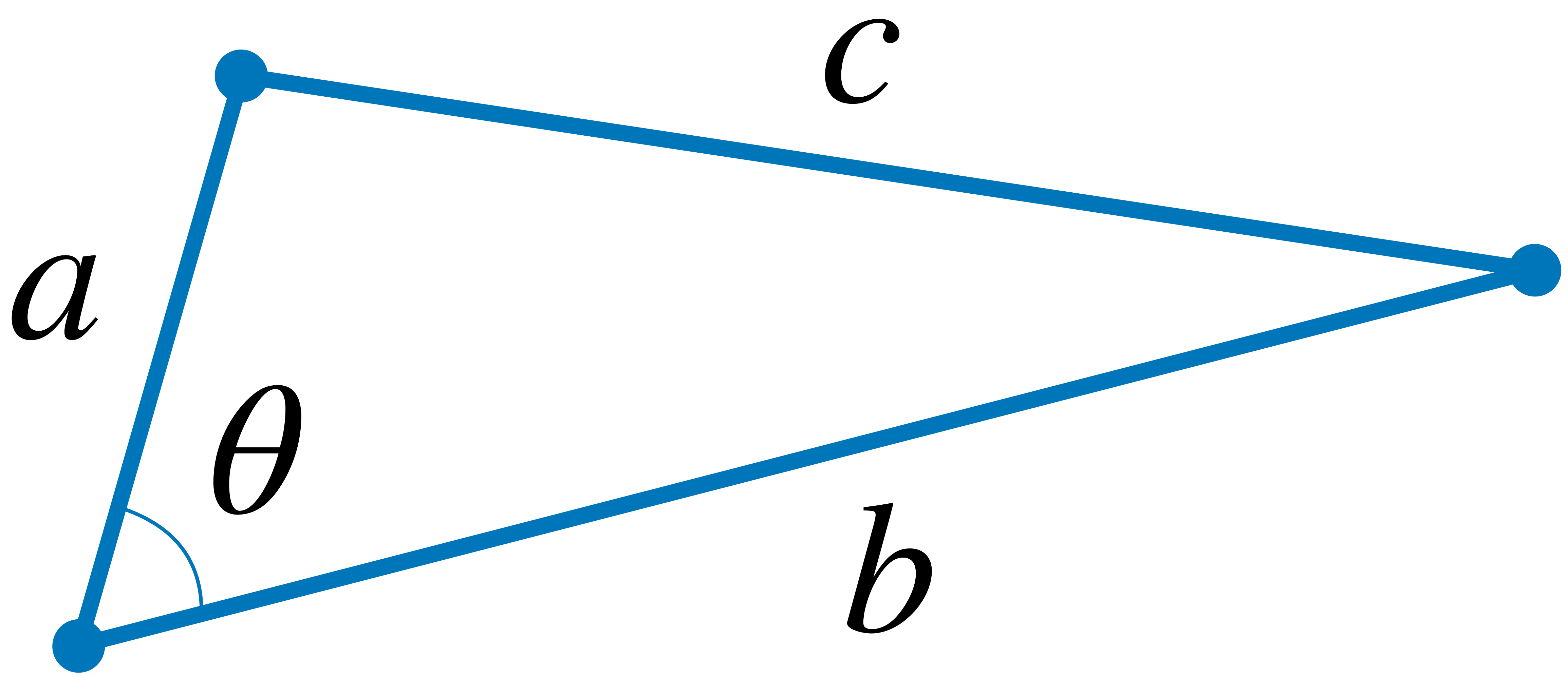


Theorem.

$$c^2 = a^2 + b^2 - 2ab\cos\theta$$

Generalized the Pythagorean Theorem

Recall: Law of Cosines



Theorem.

0 exactly when $\theta = 90^\circ$

$$c^2 = a^2 + b^2 - 2ab\cos\theta$$

Generalized the Pythagorean Theorem

Recall: Cosines and Unit Vectors

Theorem. For vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ with an angle θ between them,

$$\cos\theta = \left\langle \frac{\mathbf{u}}{\|\mathbf{u}\|}, \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\rangle$$

The cosine of the angle between two vectors is the inner product of their ℓ^2 normalizations

Orthogonality (Perpendicularity)

A Simpler Fundamental Question

How do we determine if angle between any two vectors is 90°?

Recall: Cosines and Unit Vectors

Theorem. For vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ with an angle θ between them,

$$\cos\theta = \left\langle \frac{\mathbf{u}}{\|\mathbf{u}\|}, \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\rangle$$

The cosine of the angle between two vectors is the inner product of their ℓ^2 normalizations.

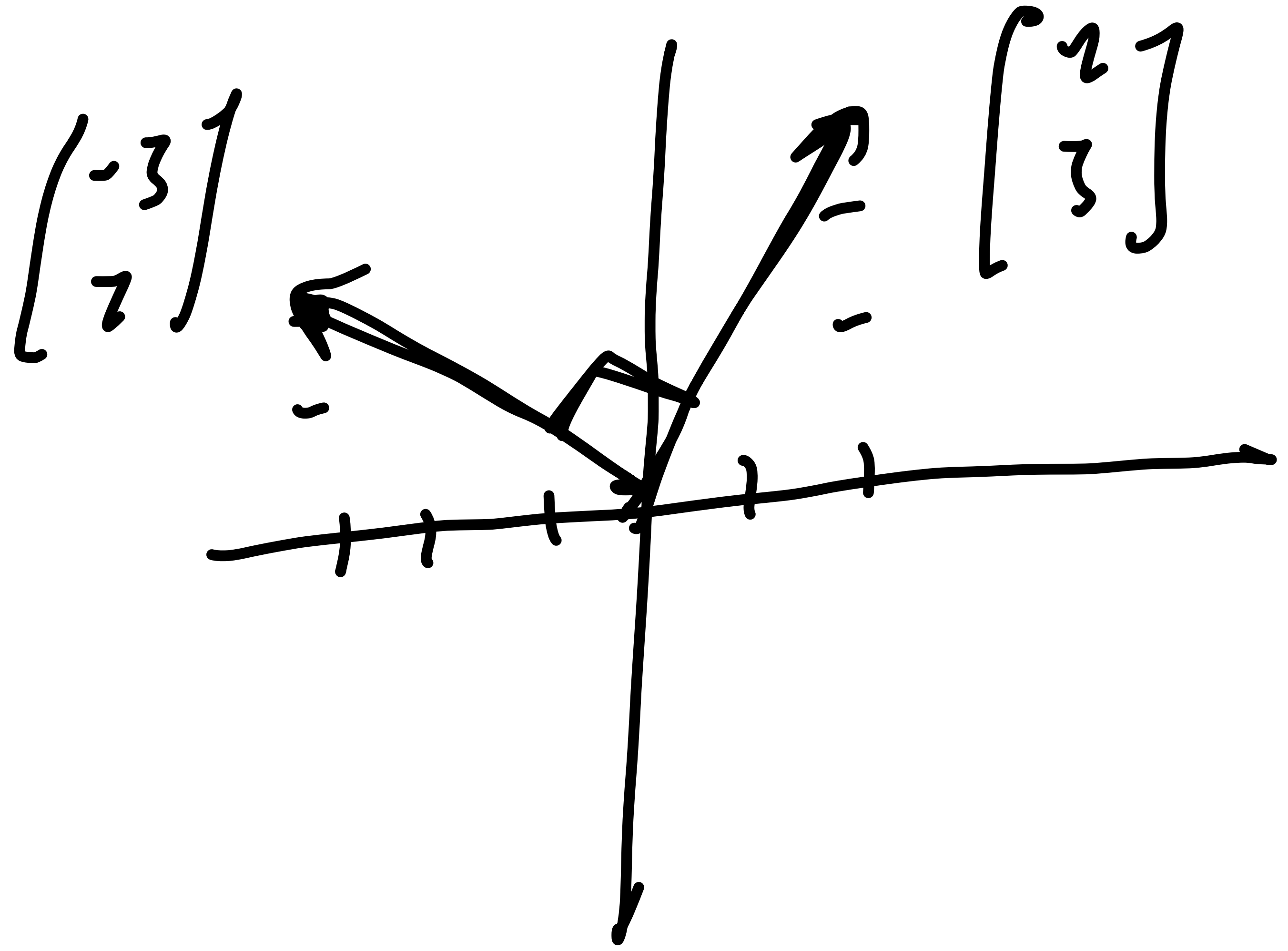
Orthogonality

Definition. Vectors $\mathbf{u}$ and $\mathbf{v}$ are **orthogonal** if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

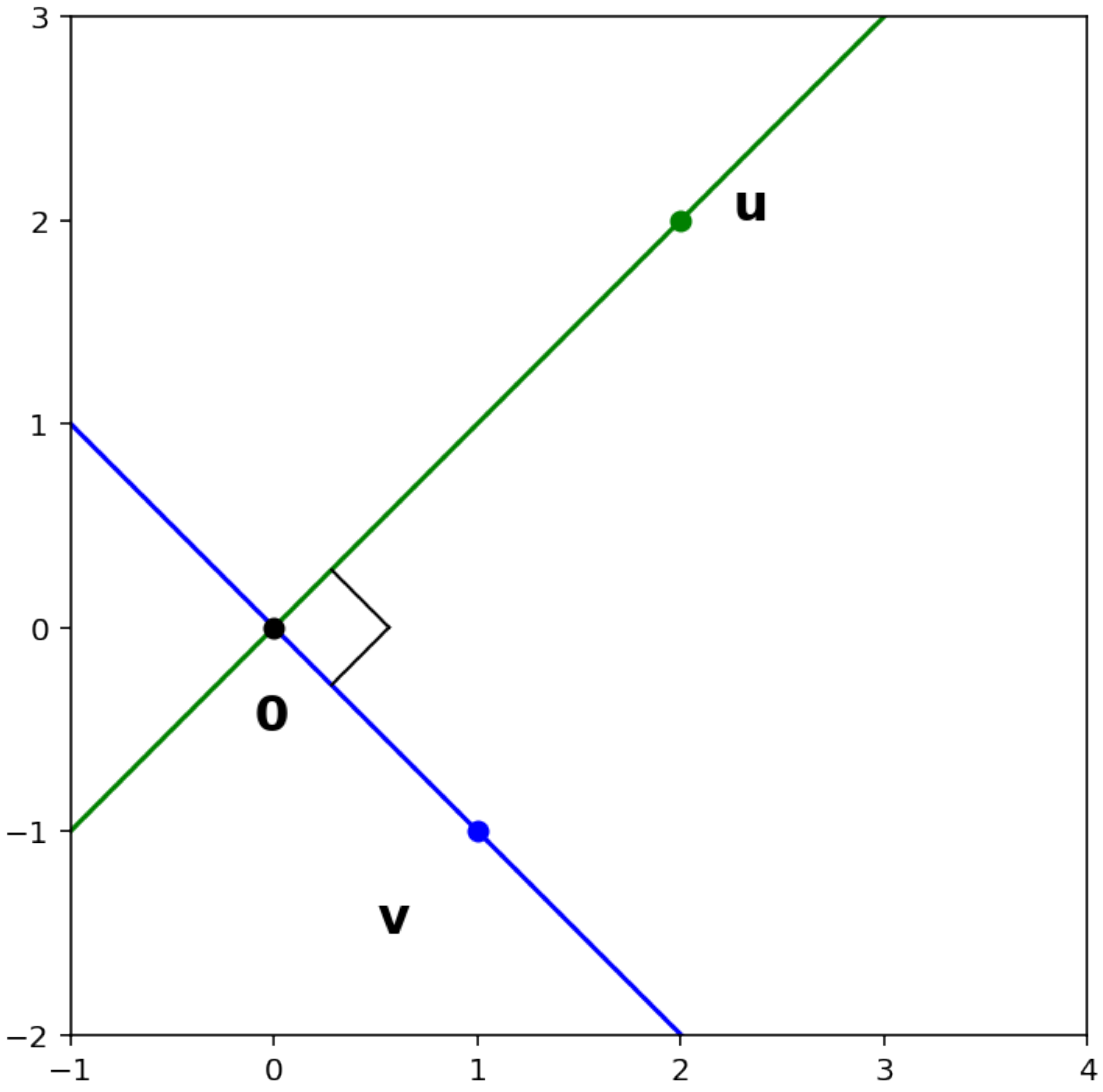
This definition gives an easy computational way to determine orthogonality.

<u>Example.</u>

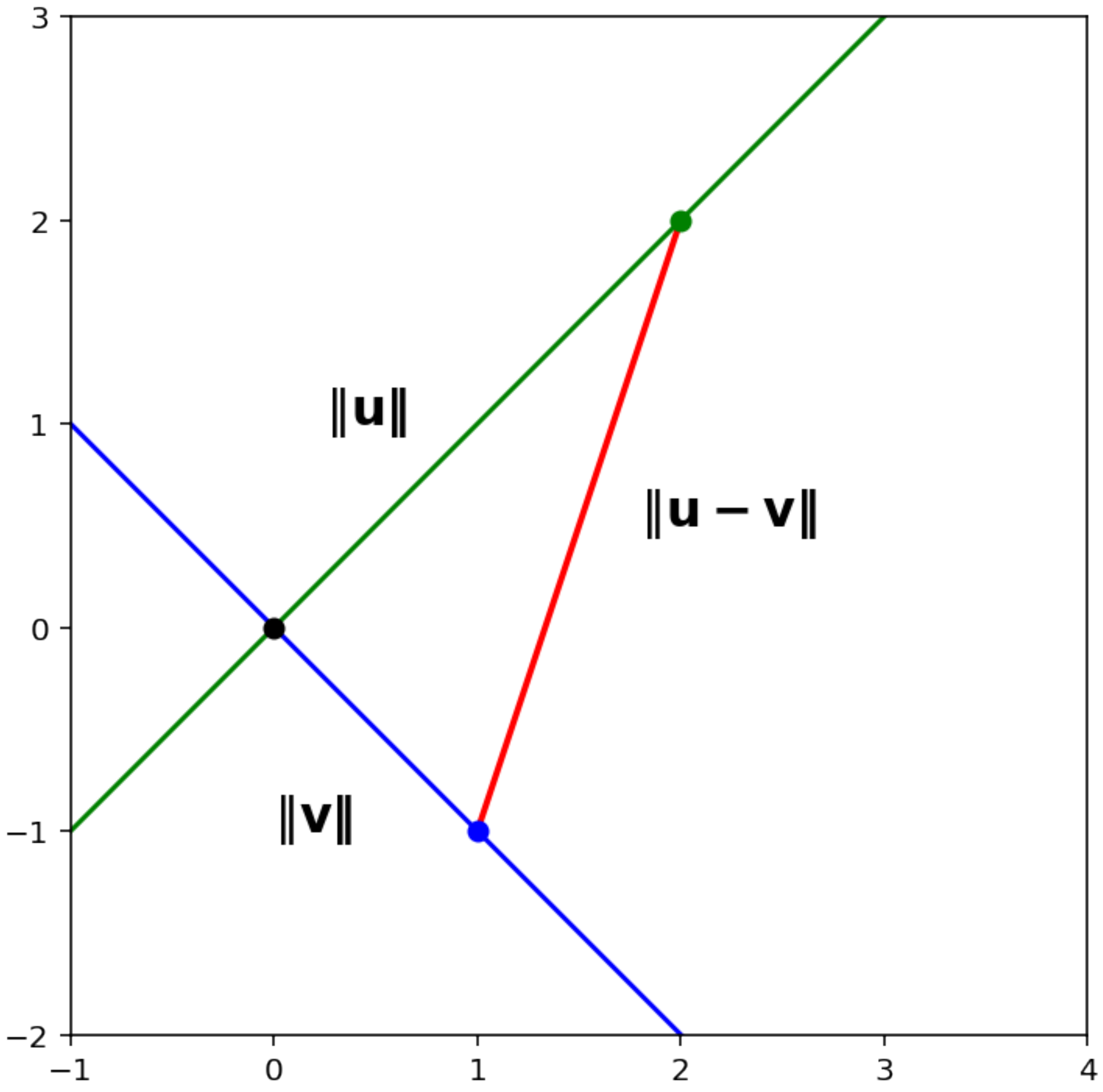
$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad \begin{bmatrix} -3 \\ 2 \end{bmatrix}$$



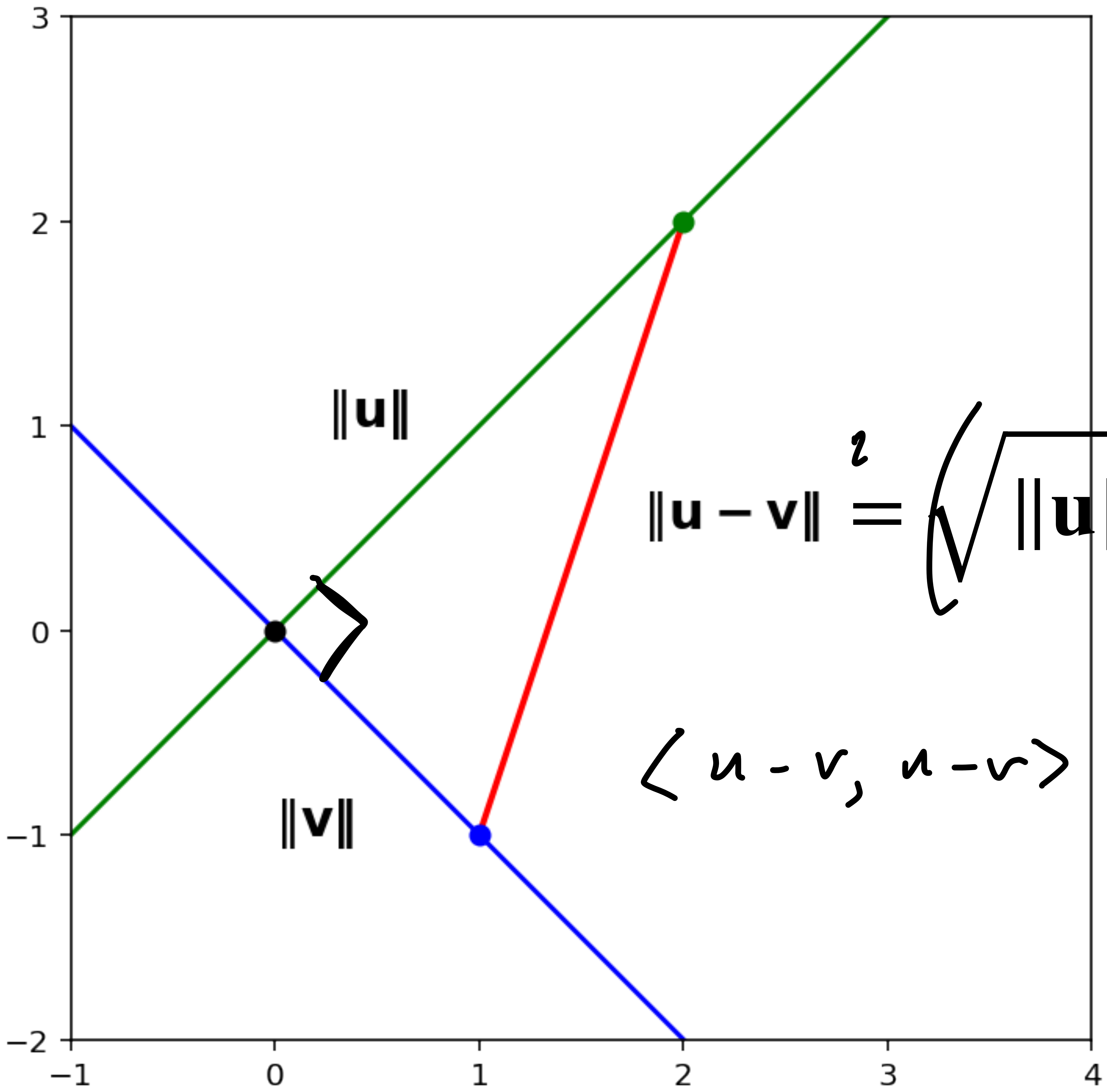
Derivation by Picture



Derivation by Picture



Derivation by Picture



$$\|\mathbf{u}-\mathbf{v}\|^2 = \left(\sqrt{\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2}\right)^2$$

(Pythagoras)

$$\langle u-v, u-v\rangle = \langle u,u\rangle + \langle v,v\rangle$$

Derivation by Algebra

$\mathbf{u}$ and $\mathbf{v}$ are orthogonal exactly when

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

Let's simplify this a bit:

$$\langle u, u \rangle + \langle v, v \rangle = \langle u-v, u-v \rangle = \langle u, u-v \rangle - \langle v, u-v \rangle$$

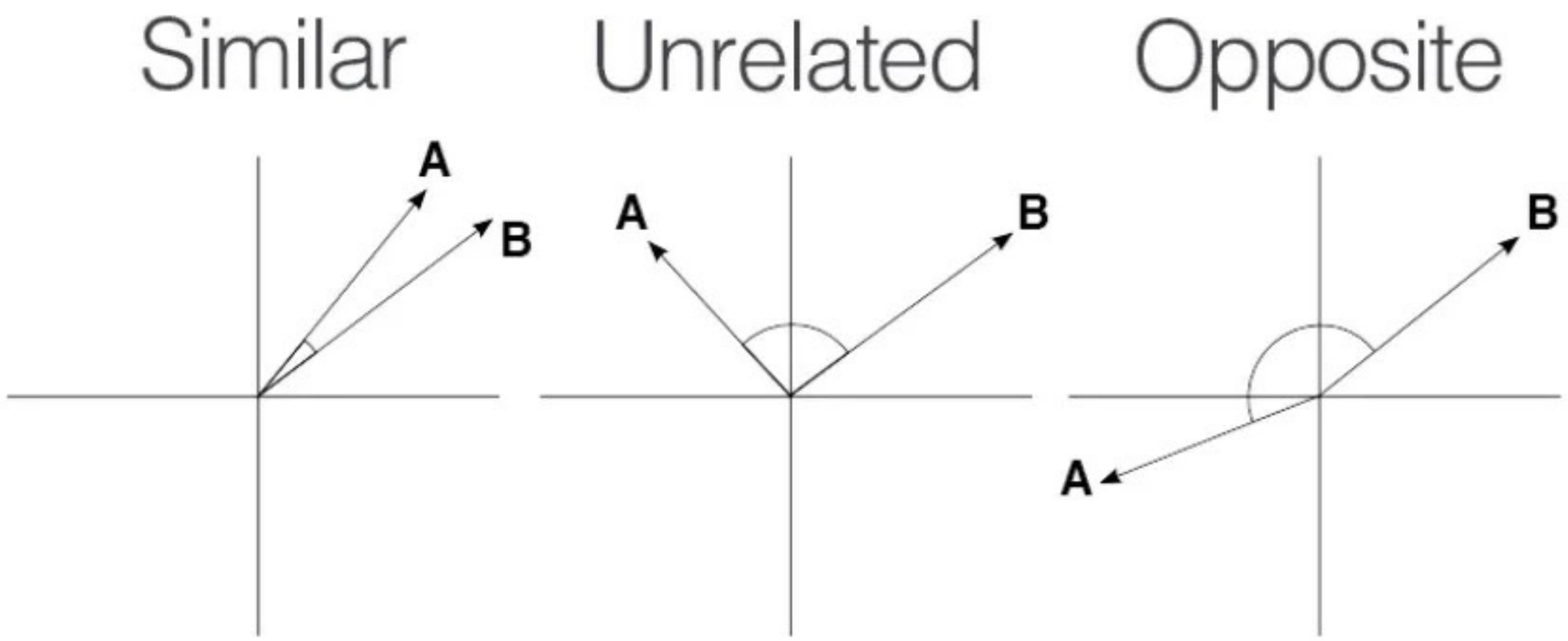
$$= \langle u, u \rangle - \langle u, v \rangle - \langle v, u \rangle + \langle v, v \rangle$$

$$2\langle u, v \rangle = 0$$

$$\langle u, v \rangle = 0$$

Application: Cosine Similarity

High Level



Data points are very big vectors.

Similar vectors "point in nearly the same direction."

https://medium.com/@milana.shxanukova15/cosine-distance-and-cosine-similarity-a5da0e4d9ded

Example: Netflix Users

$$\text{user}_1 = \begin{bmatrix} 2 \\ 10 \\ 1 \\ 3 \end{bmatrix} \quad \text{user}_2 = \begin{bmatrix} 2 \\ 5 \\ 0 \\ 4 \end{bmatrix} \quad \text{user}_3 = \begin{bmatrix} 10 \\ 0 \\ 5 \\ 6 \end{bmatrix} \quad \begin{matrix} \texttt{comedy} \\ \texttt{drama} \\ \texttt{horror} \\ \texttt{romance} \end{matrix}$$

A Netflix user might be represented as a vectors whose ith entry is the number of movies they've watched in a particular genre.

Who are more likely to share similar interests in movies?

Cosine Similarity

Definition. The **cosine similarity** of two vectors is the cosine of the angle between them.

If its close to 0, then two Netflix users watch very different movies.

If its close to 1, then two Netflix users watch very similar movies.

Example: Netflix Users

comedy
drama
horror
romance

$$\text{user}_2 = \begin{bmatrix} 2 \\ 5 \\ 0 \\ 4 \end{bmatrix}$$

$$\text{sim}(\text{user}_1, \text{user}_2) \approx 0.92$$

$$\text{sim}(\text{user}_2, \text{user}_3) \approx 0.52$$

$$\text{user}_3 = \begin{bmatrix} 10 \\ 0 \\ 5 \\ 6 \end{bmatrix}$$

$$\text{user}_1 = \begin{bmatrix} 2 \\ 10 \\ 1 \\ 3 \end{bmatrix}$$

$$\text{sim}(\text{user}_1, \text{user}_3) \approx 0.32$$

Other Examples

- ***Document similarity***
  - Documents ↦ word count vectors
  - Similar documents should use similar words
- ***Word2Vec***
  - Words ↦ vector *somehow*
  - This underlies modern natural language processing (NLP)

Recall: Orthogonality

Definition. Vectors $\mathbf{u}$ and $\mathbf{v}$ are **orthogonal** if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$

Orthogonal and perpendicular are the same thing.

With inner product we can...

- Given a vector we can determine its length
- Given two points (vectors) we can determine the distance between them
- Given two vectors we can determine the angle between them

Orthogonal Sets

Orthogonal Sets

Definition. A set $\{u_1, u_2, \ldots, u_p\}$ of vectors from R^n is an **orthogonal set** if every pair of distinct vectors is orthogonal: if $i \neq j$ then

$$\langle u_i, u_j \rangle = 0$$

Each vector is pairwise/mutually perpendicular

Example

$$u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \qquad u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \qquad u_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

Verify:

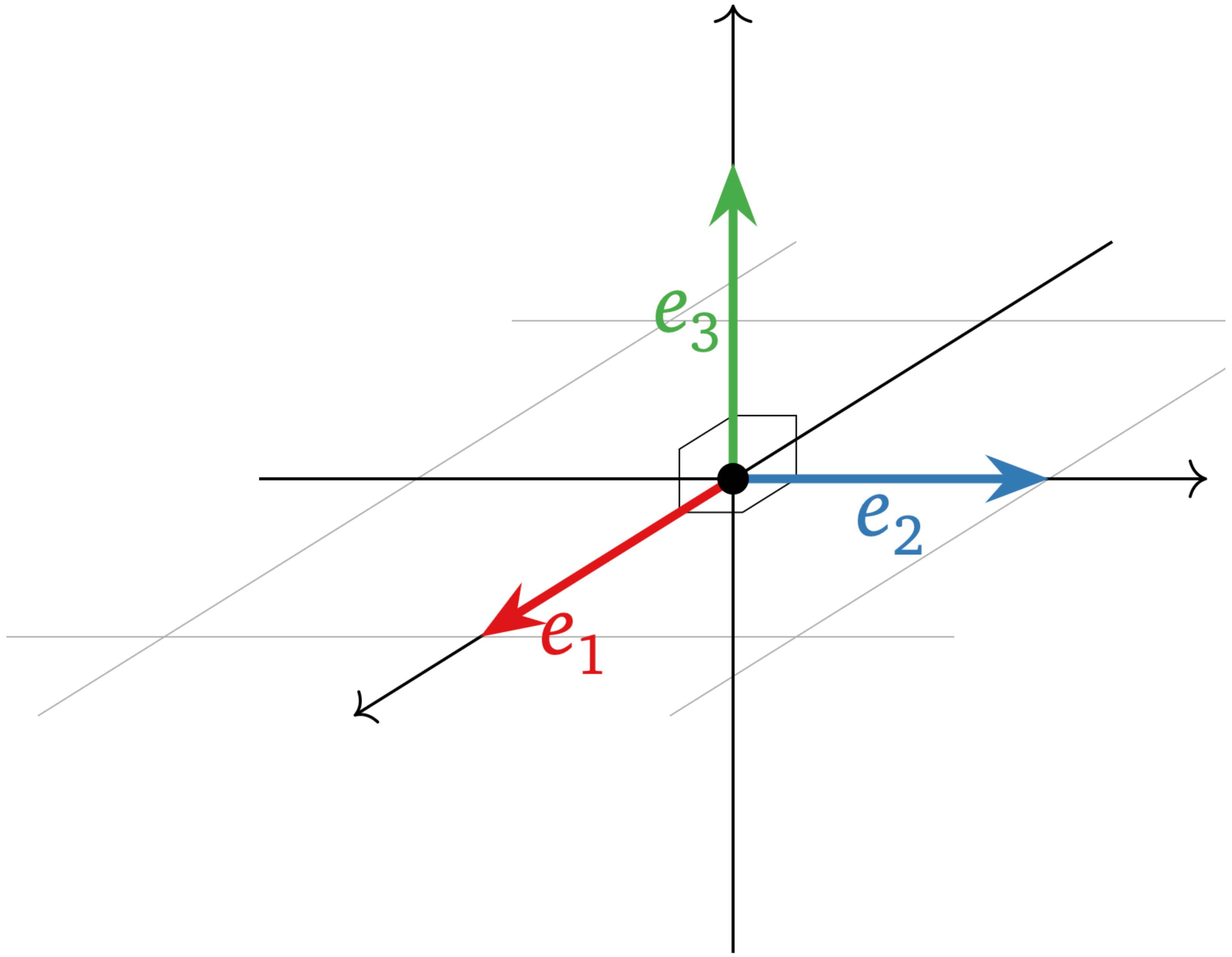
$$\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = -3 + 2 + 1 = 0 \checkmark$$

$$\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix} = -3/2 + -4/2 + 7/2 = 0 \checkmark$$

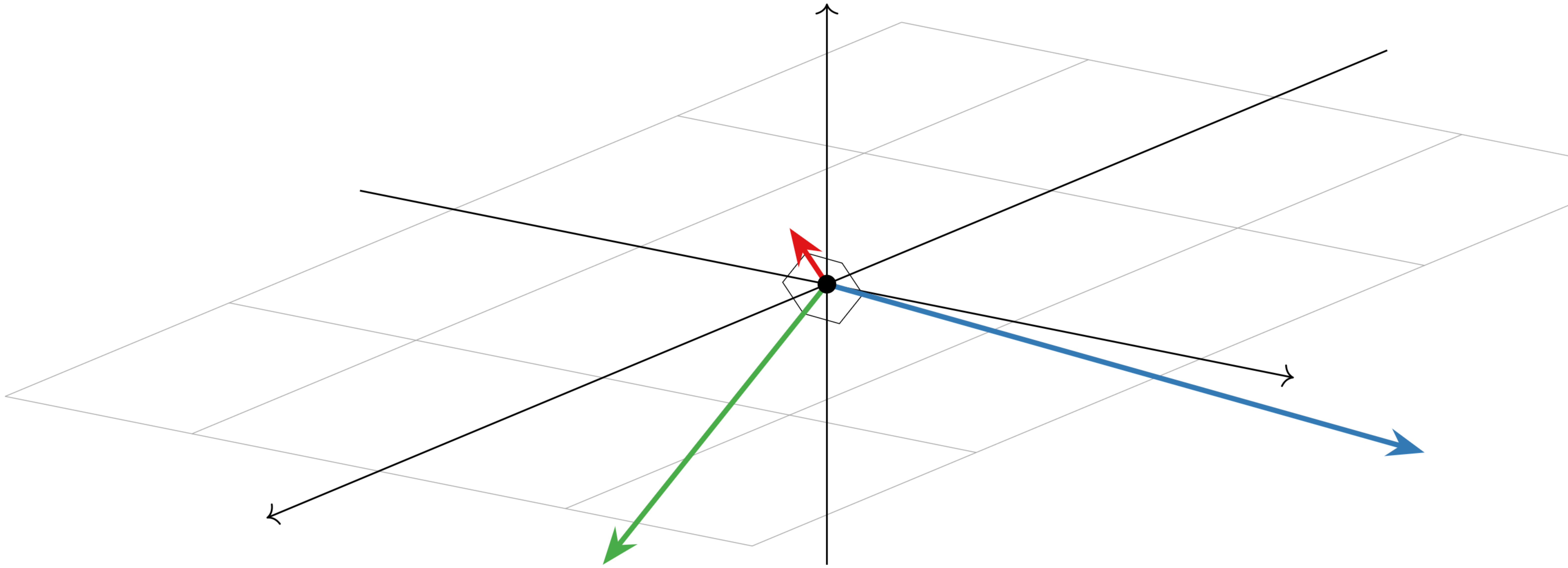
$$\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix} = \frac{1}{2} + \frac{-8}{2} + \frac{7}{2} = 0 \checkmark$$

What do orthogonal sets look like?

The Picture



the standard basis forms a "centered" orthogonal set



an orthogonal set is like the standard basis *after some rotations and scalings*

https://textbooks.math.gatech.edu/ila/orthogonal-sets.html

Orthogonal Sets and Independence

Theorem. If $\{u_1, u_2, \ldots, u_k\}$ is an orthogonal set of *nonzero* vectors from R^n, then it is linearly independent

Verify: $\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \ldots + \alpha_k \vec{u}_k = \vec{0}$

$$\left\langle \sum_{i=1}^{k} \alpha_i \vec{u}_i , \vec{u}_1 \right\rangle = \sum_{i=1}^{k} \alpha_i \langle \vec{u}_i , \vec{u}_1 \rangle = \alpha_1 \langle \vec{u}_1 , \vec{u}_1 \rangle$$

$\|$

0

$\langle u_i , u_j \rangle = 0$ if $i \neq j$

$\langle \vec{u}_1 , \vec{u}_1 \rangle > 0$

so $\alpha_1 = 0$, generalizes to all α_j

The Takeaway

If $\{u_1, u_2, \ldots, u_k\}$ is an orthogonal set, then it is a **basis** for $span\{u_1, u_2, \ldots, u_k\}$

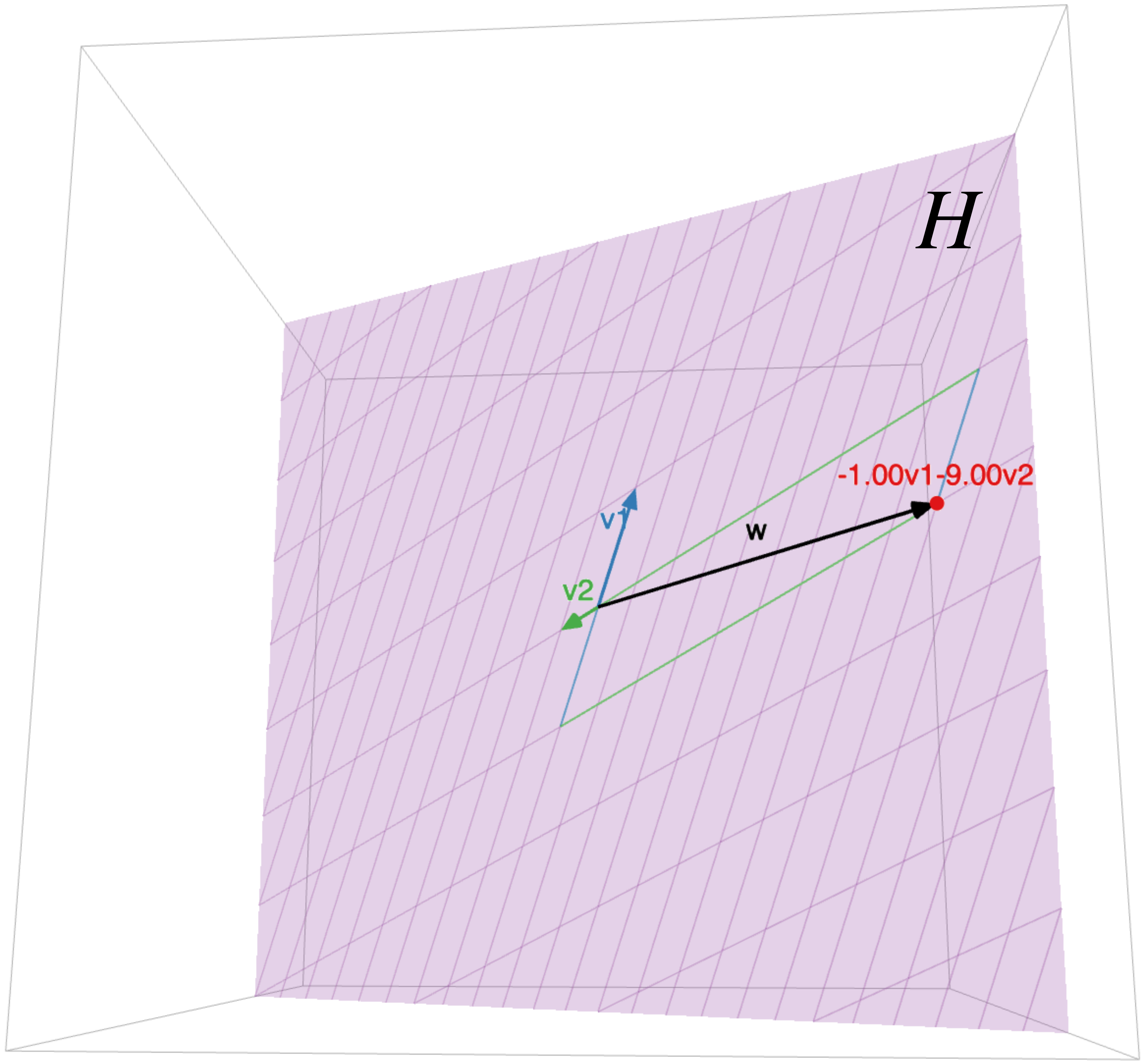
Orthogonal Basis

Definition. An **orthogonal basis** for a subspace W of R^n is a basis for W which is also an orthogonal set.

https://textbooks.math.gatech.edu/ila/spans.html

Orthogonal Basis

Definition. An **orthogonal basis** for a subspace W of R^n is a basis for W which is also an orthogonal set.

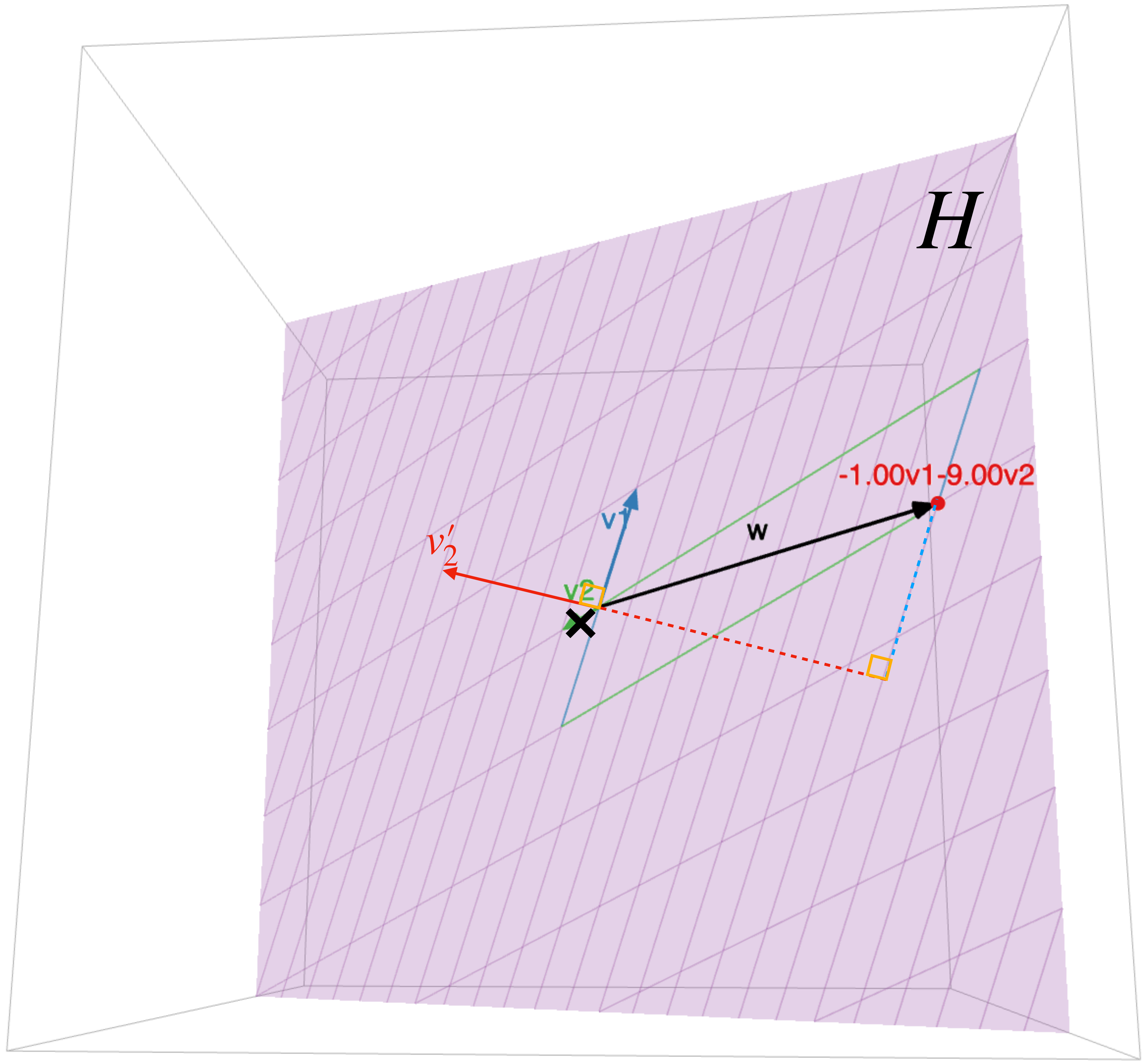


v_1 and v_2 form a basis of H

https://textbooks.math.gatech.edu/ila/spans.html

Orthogonal Basis

Definition. An **orthogonal basis** for a subspace W of R^n is a basis for W which is also an orthogonal set.



v_1 and v_2 form a basis of H

v_1 and v_2' form an **orthogonal** basis of H

https://textbooks.math.gatech.edu/ila/spans.html

What's nice about an orthogonal basis?

Recall: How To: Bases

Recall: How To: Bases

Question. Given a basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ for a subspace W of R^n and a vector $\mathbf{w}$ in W, weights $c_1, c_2, \ldots, c_p$ such that

$$\mathbf{w} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots c_p\mathbf{u}_p$$

Recall: How To: Bases

Question. Given a basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ for a subspace W of R^n and a vector $\mathbf{w}$ in W, weights $c_1, c_2, \ldots, c_p$ such that

$$\mathbf{w} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots c_p\mathbf{u}_p$$

Solution. Solve the vector equation

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \ldots x_p\mathbf{u}_p = \mathbf{w}$$

by Gaussian elimination, matrix inversion, etc.

Recall: How To: Bases

Question. Given a basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ for a subspace W of R^n and a vector $\mathbf{w}$ in W, weights $c_1, c_2, \ldots, c_p$ such that

$$\mathbf{w} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots c_p\mathbf{u}_p$$

Solution. Solve the vector equation

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \ldots x_p\mathbf{u}_p = \mathbf{w}$$

by Gaussian elimination, matrix inversion, etc.

This takes work

Orthogonal Bases and Linear Combinations

Theorem. For an orthogonal set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$, if $\mathbf{y} = c_1\mathbf{u}_1 + \ldots + c_p\mathbf{u}_p$ then for $j = 1, \ldots, p$

$$c_j = \frac{\mathbf{y}^T\mathbf{u}_j}{\mathbf{u}_j^T\mathbf{u}_j} = \frac{\langle y, u_j \rangle}{\langle u_j, u_j \rangle}$$

Verify: $\langle y, u_j \rangle = \left\langle \sum_{i=1}^{p} c_i \vec{u}_i, u_j \right\rangle = \sum_{i=1}^{p} c_i \langle \vec{u}_i, \vec{u}_j \rangle$

$= c_j \langle \vec{u}_j, \vec{u}_j \rangle$

How To: Orthogonal Bases

How To: Orthogonal Bases

Question. Given an **orthogonal** basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ for a subspace W of R^n and a vector $\mathbf{w}$ in W, weights $c_1, c_2, \ldots, c_p$ such that

$$\mathbf{w} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots c_p\mathbf{u}_p$$

How To: Orthogonal Bases

Question. Given an **orthogonal** basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ for a subspace W of R^n and a vector $\mathbf{w}$ in W, weights $c_1, c_2, \ldots, c_p$ such that

$$\mathbf{w} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots c_p\mathbf{u}_p$$

Solution. $c_j = \dfrac{\mathbf{w} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j}$

How To: Orthogonal Bases

Question. Given an **orthogonal** basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ for a subspace W of R^n and a vector $\mathbf{w}$ in W, weights $c_1, c_2, \ldots, c_p$ such that

$$\mathbf{w} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots c_p\mathbf{u}_p$$

Solution. $c_j = \dfrac{\mathbf{w} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j}$

Much easier to compute.

Question

Express $[6 \;\; 1 \;\; (-8)]^T$ *as a linear combination of vectors in* $\{u_1, u_2, u_3\}$ *where*

$$u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \qquad u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \qquad u_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

Answer: $u_1 - 2u_2 - 2u_3$

$\vec{v} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$

$\langle \vec{v}, \vec{u}_1 \rangle = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = 18 + 1 - 8 = 11$

$c_1 = 1$

$\langle \vec{u}_1, \vec{u}_1 \rangle = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = 9 + 1 + 1 = 11$

$\langle \vec{v}, \vec{u}_2 \rangle = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = -6 + 2 - 8 = -12$

$c_2 = -2$

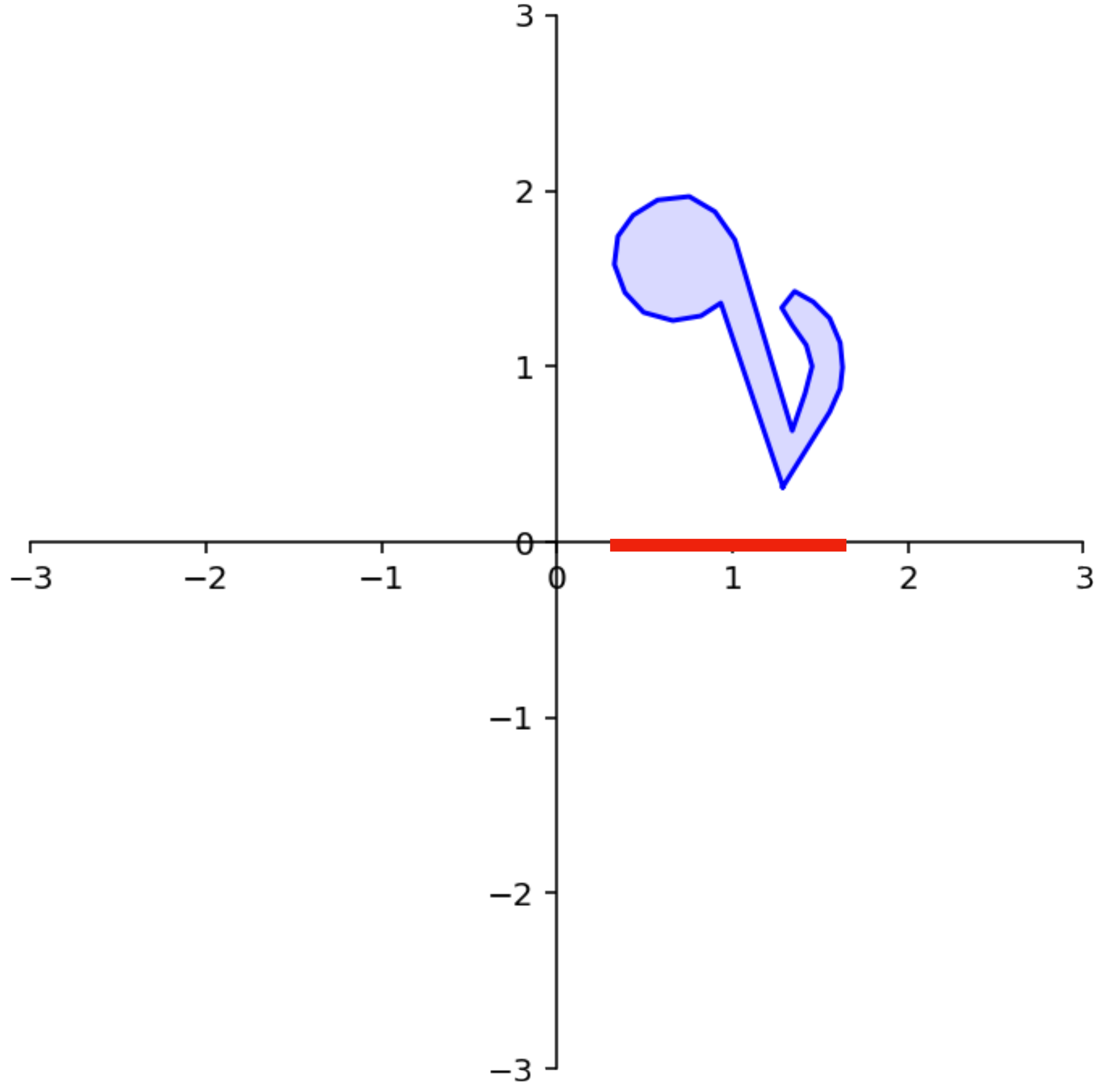
$\langle \vec{u}_2, \vec{u}_2 \rangle = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = 1 + 4 + 1 = 6$

and so on...

Orthogonal Projection

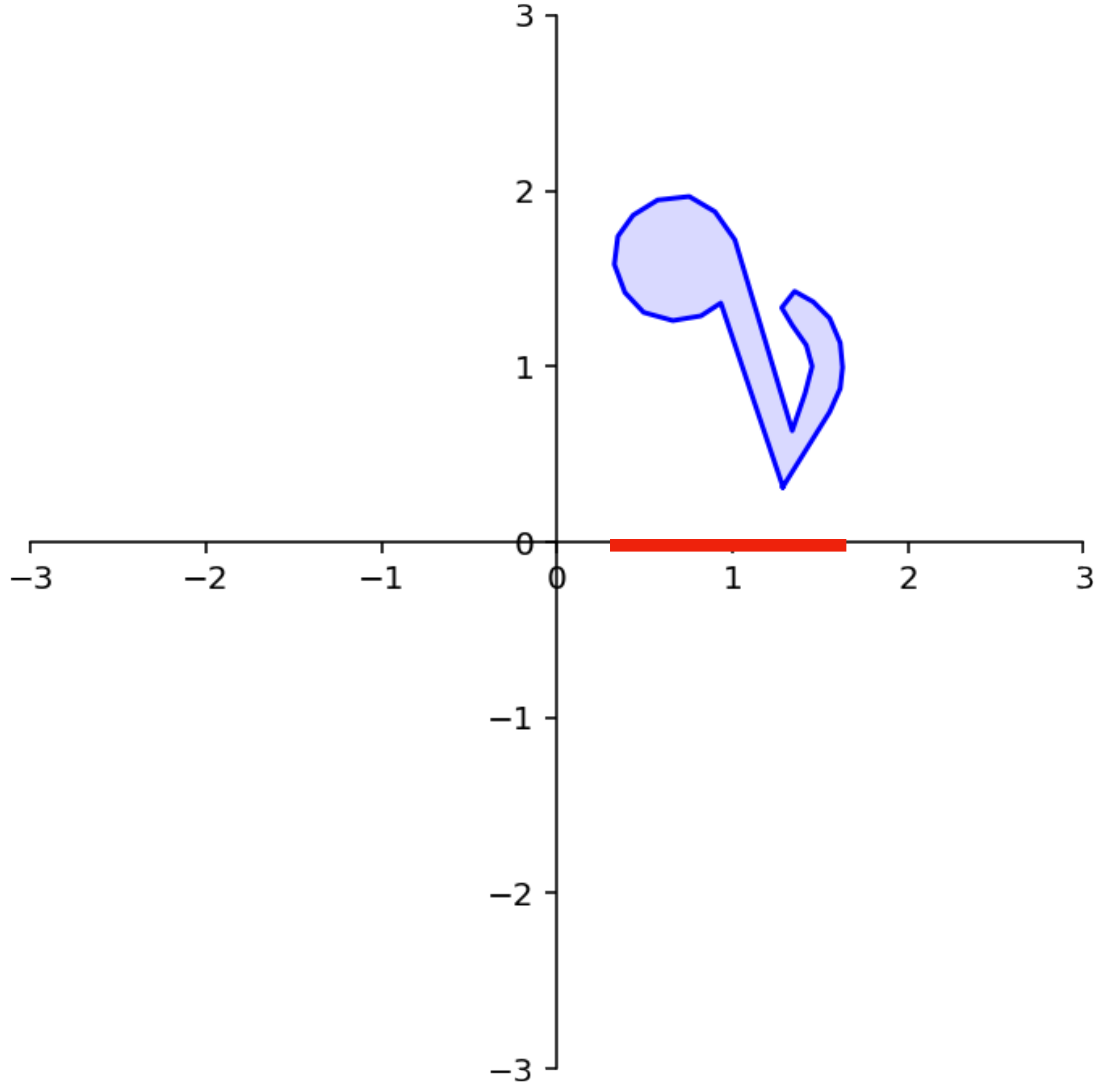
Why does that formula in the last example work?

Recall: Projection onto the x-axis



Recall: Projection onto the x-axis

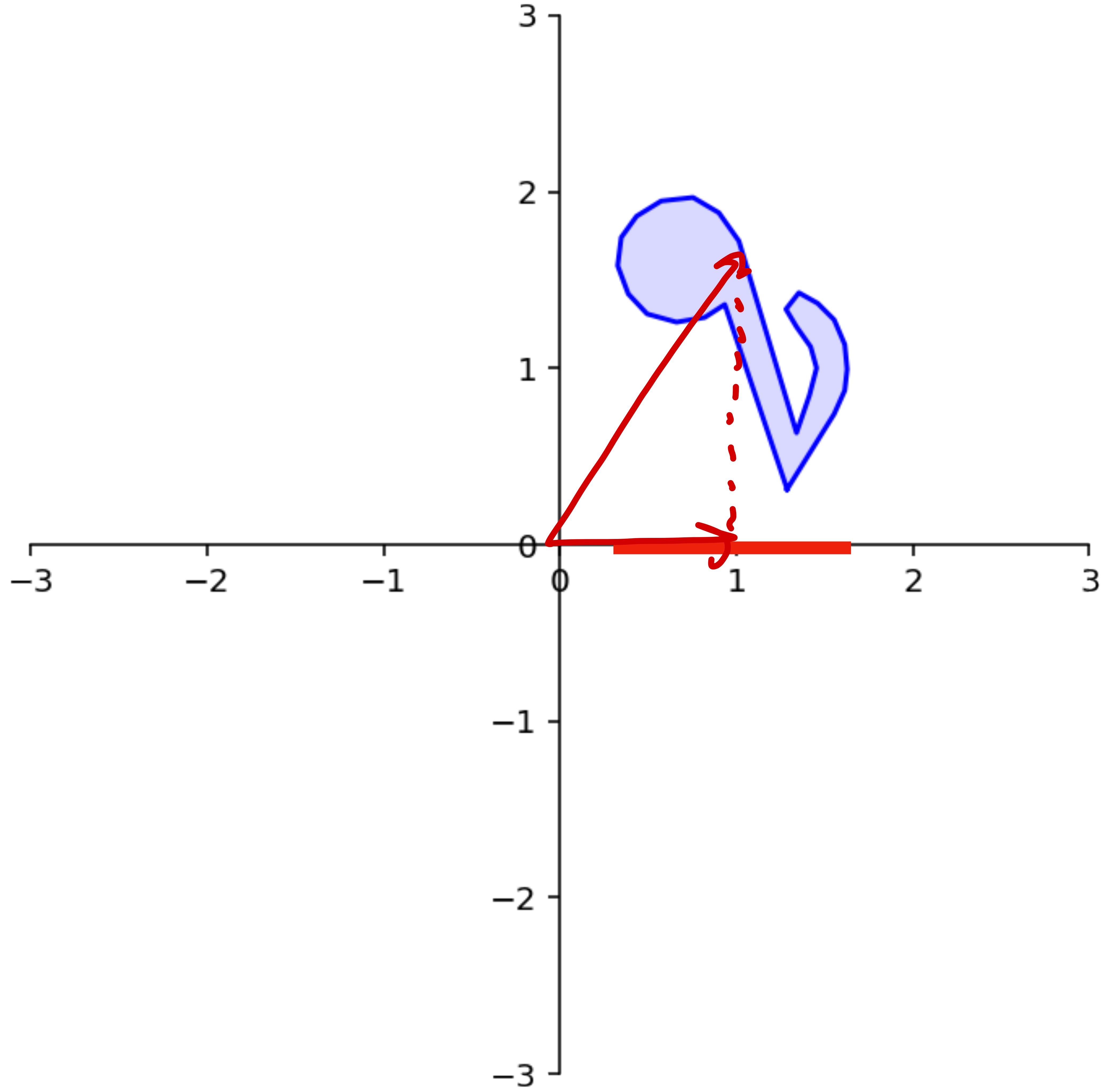
We've seen simple projections in R^2



Recall: Projection onto the x-axis

We've seen simple projections in R^2

We're going to generalize this idea

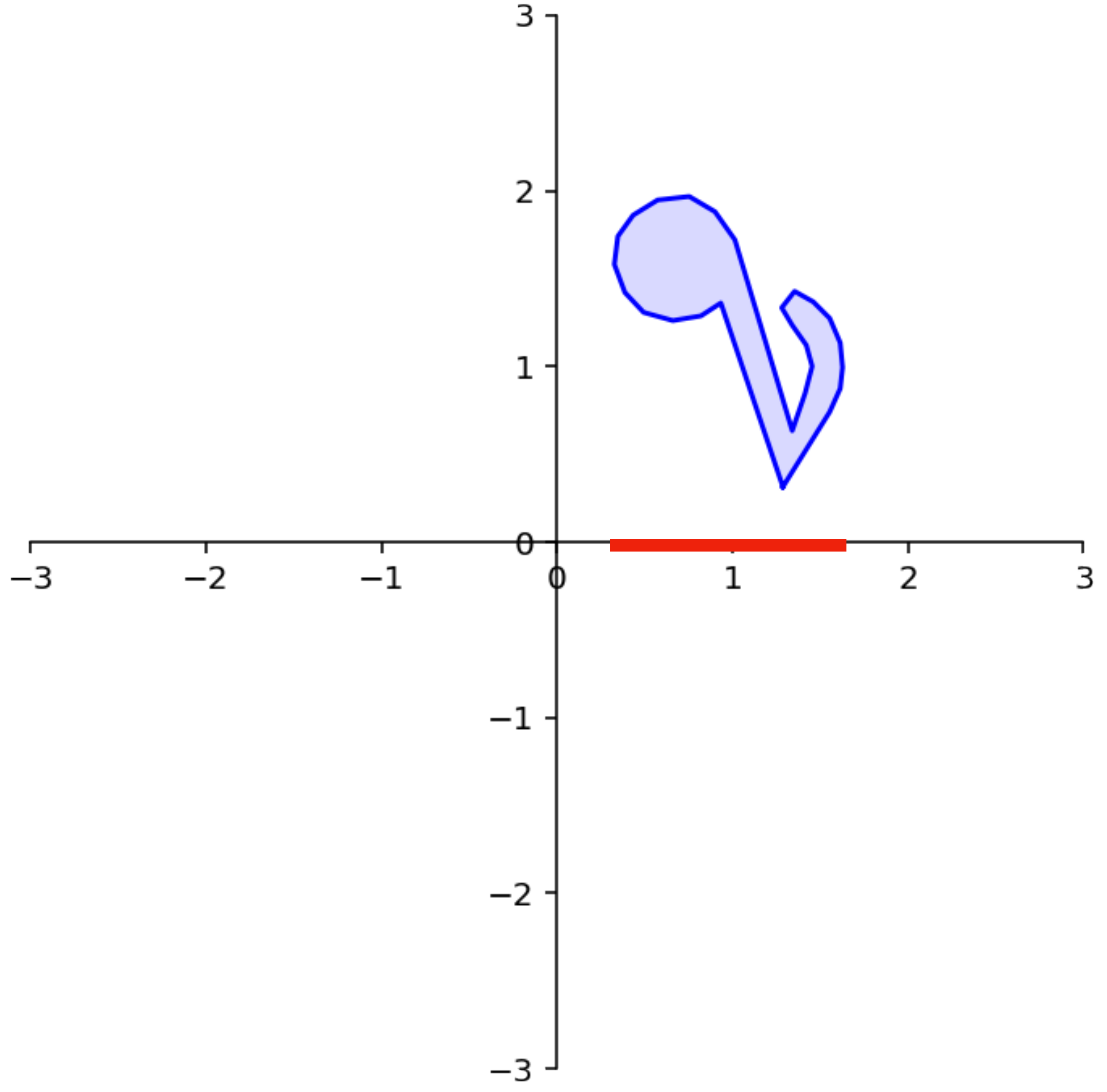


Recall: Projection onto the x-axis

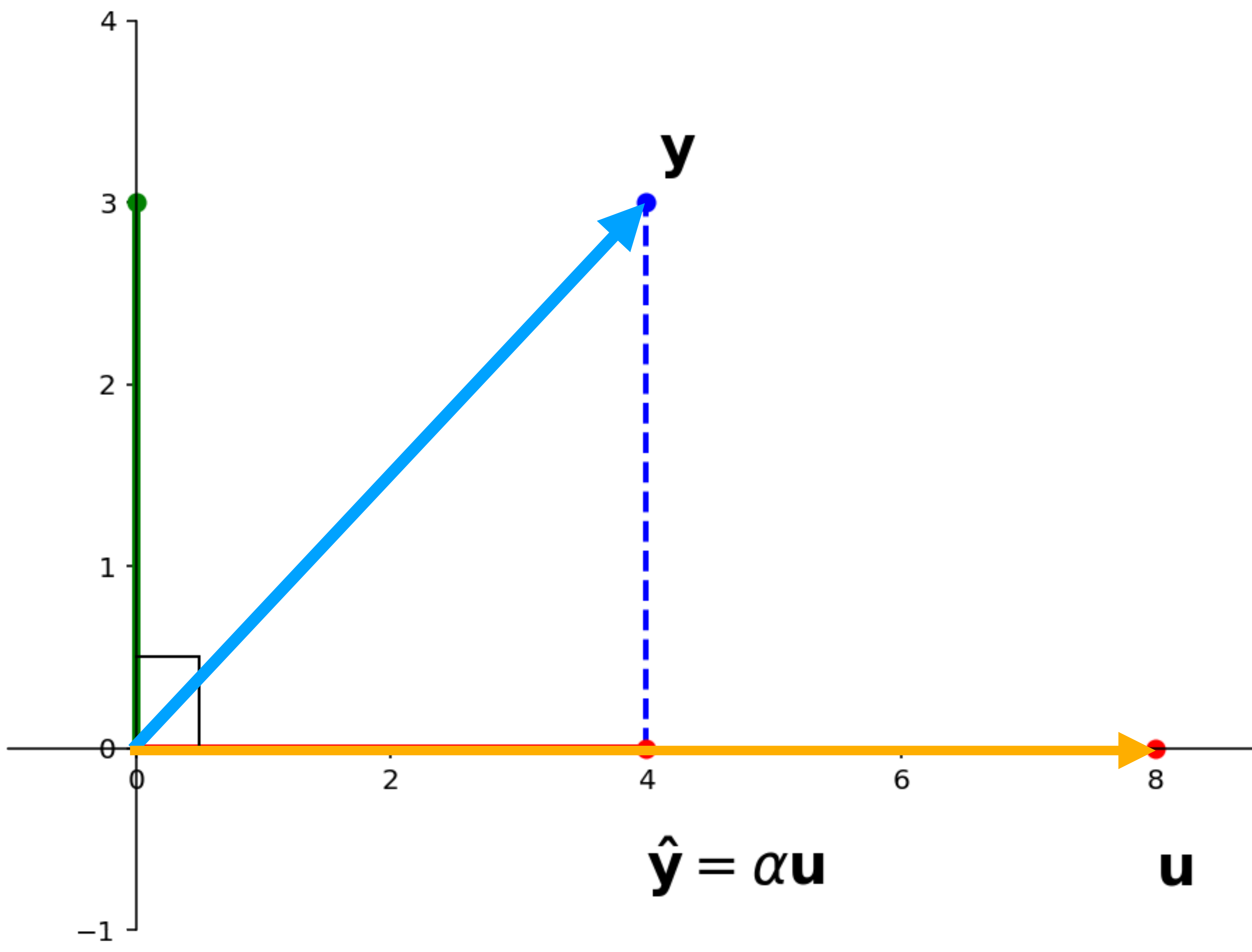
We've seen simple projections in R^2

We're going to generalize this idea

What we really did was a kind of projection onto the basis vectors

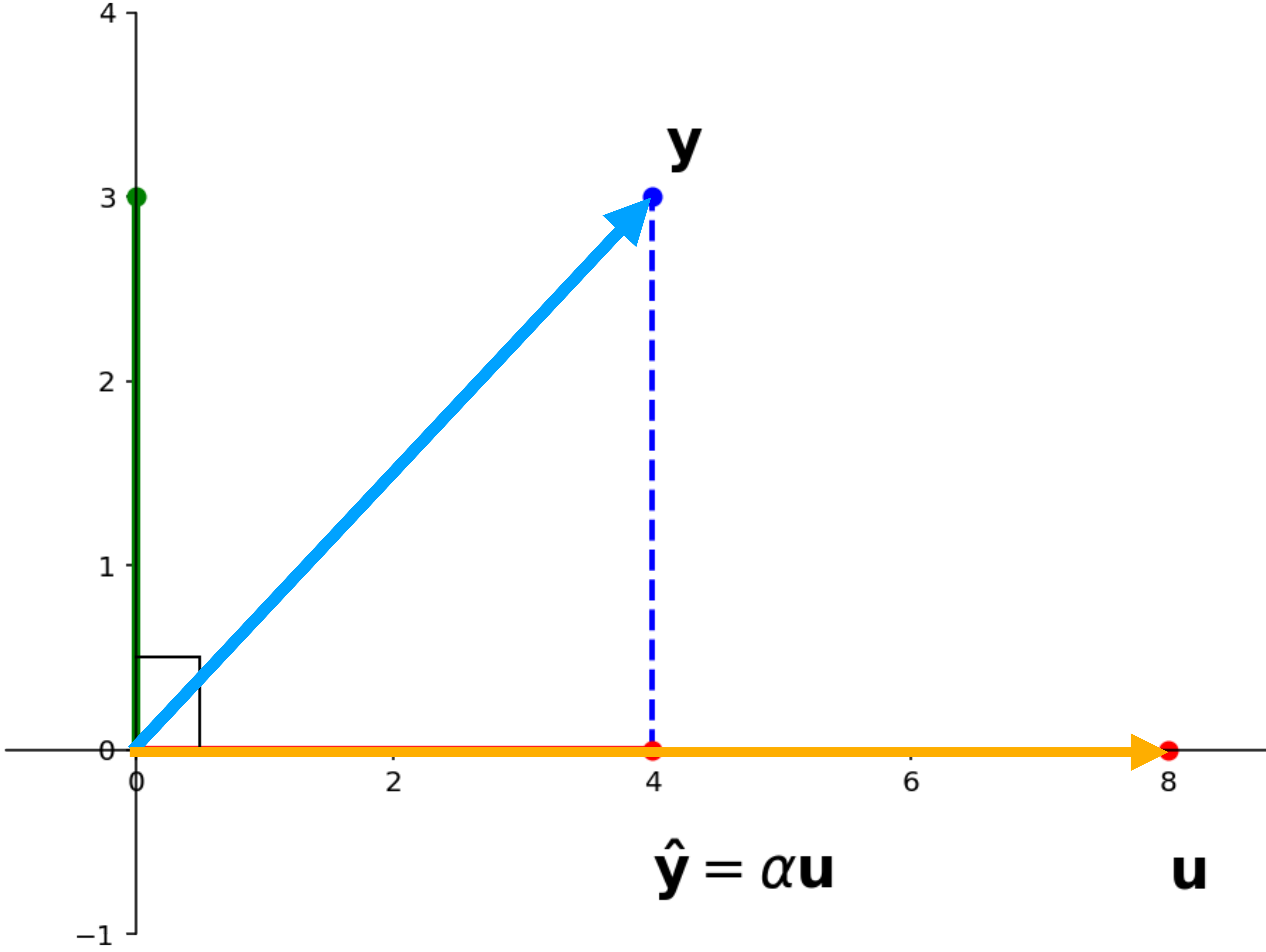


Orthogonal Projection



Orthogonal Projection

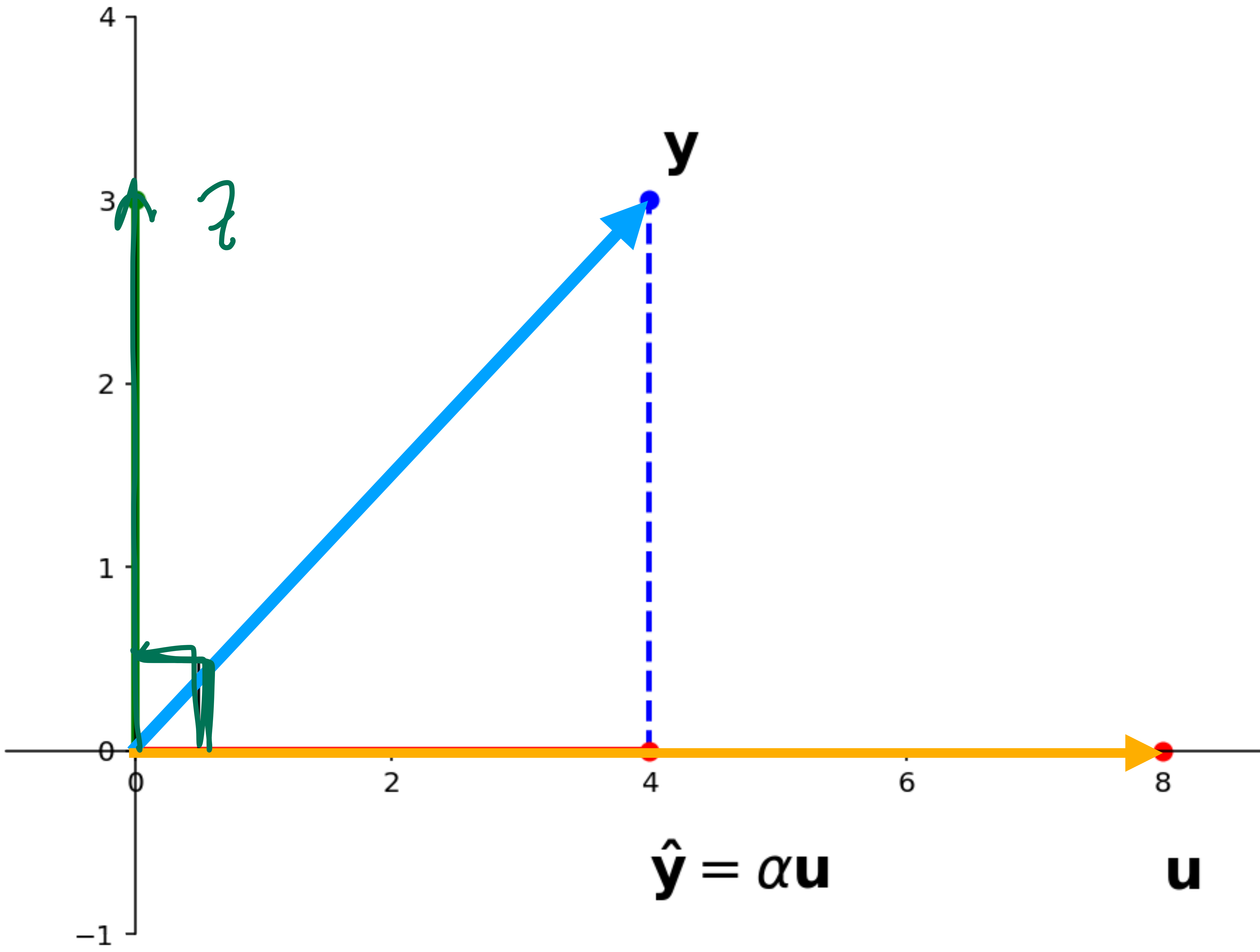
Question. Given vectors $\mathbf{y}$ and $\mathbf{u}$ in R^n, find vectors $\hat{\mathbf{y}}$ and $\mathbf{z}$ such that



Orthogonal Projection

Question. Given vectors $\mathbf{y}$ and $\mathbf{u}$ in R^n, find vectors $\hat{\mathbf{y}}$ and $\mathbf{z}$ such that

» $\mathbf{z}$ is orthogonal to $\mathbf{u}$ (i.e., $\mathbf{z} \cdot \mathbf{u} = 0$)

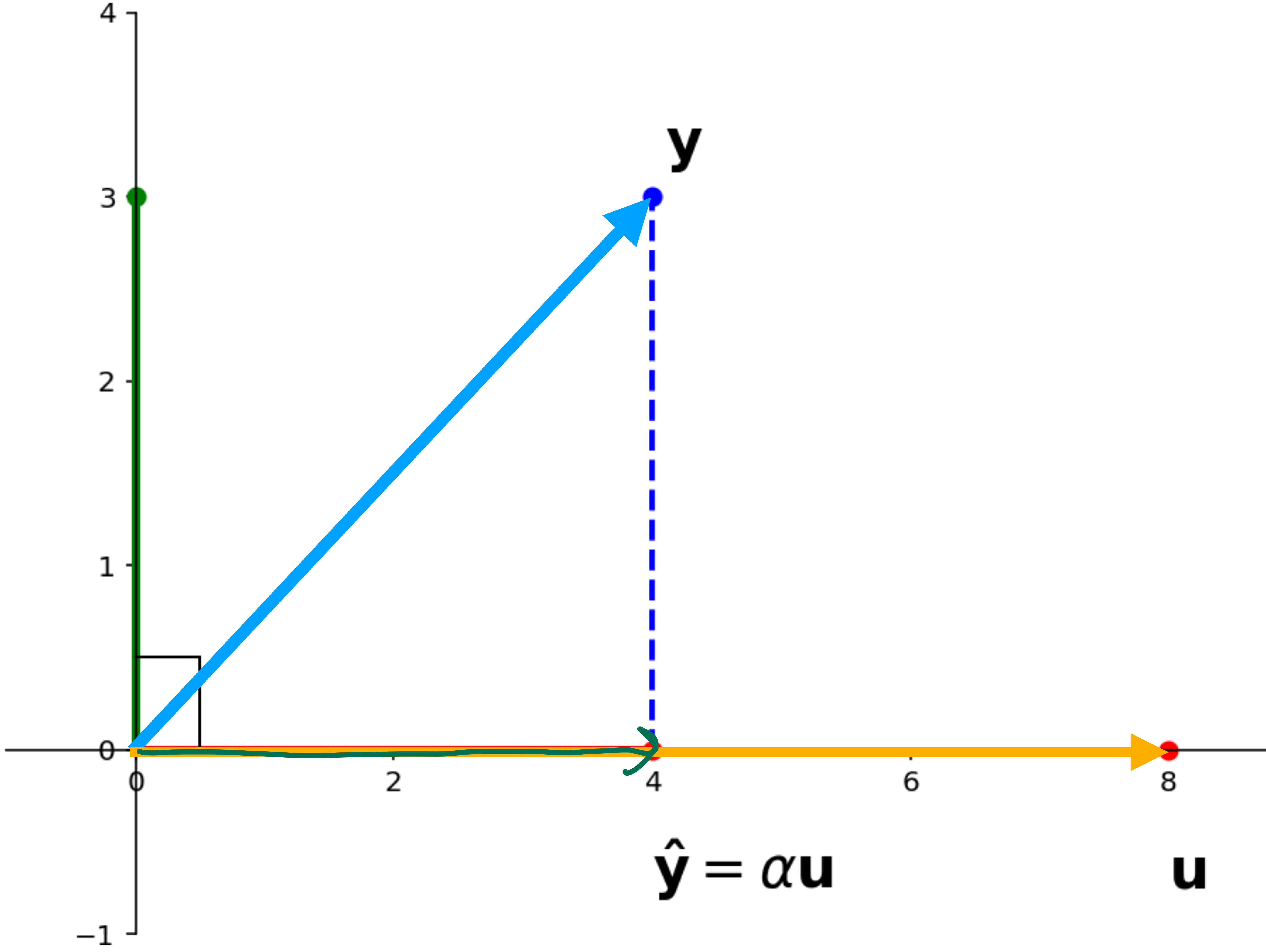


Orthogonal Projection

Question. Given vectors $\mathbf{y}$ and $\mathbf{u}$ in R^n, find vectors $\hat{\mathbf{y}}$ and $\mathbf{z}$ such that

» $\mathbf{z}$ is orthogonal to $\mathbf{u}$ (i.e., $\mathbf{z} \cdot \mathbf{u} = 0$)

» $\hat{\mathbf{y}} \in span\{\mathbf{u}\}$



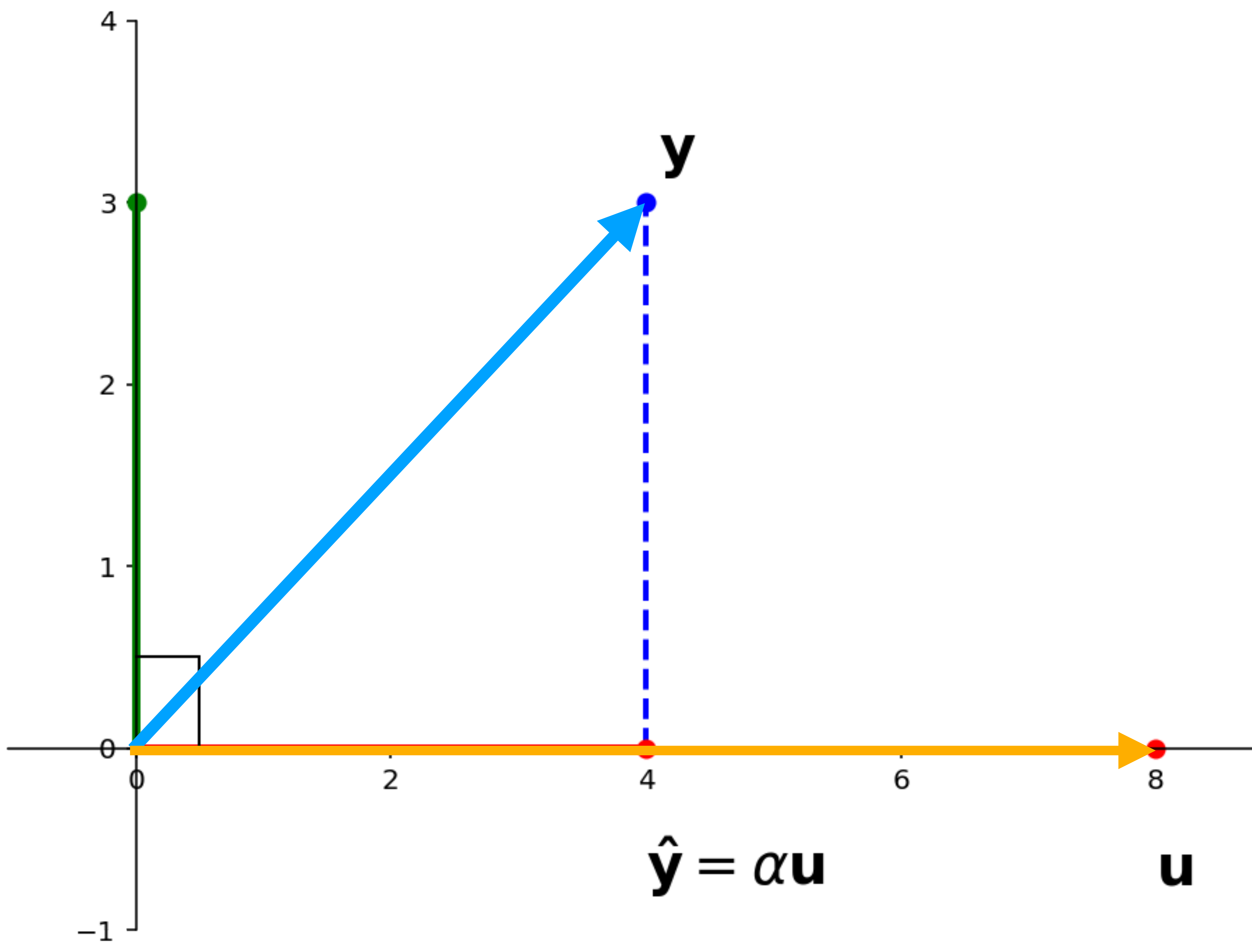
Orthogonal Projection

Question. Given vectors $\mathbf{y}$ and $\mathbf{u}$ in R^n, find vectors $\hat{\mathbf{y}}$ and $\mathbf{z}$ such that

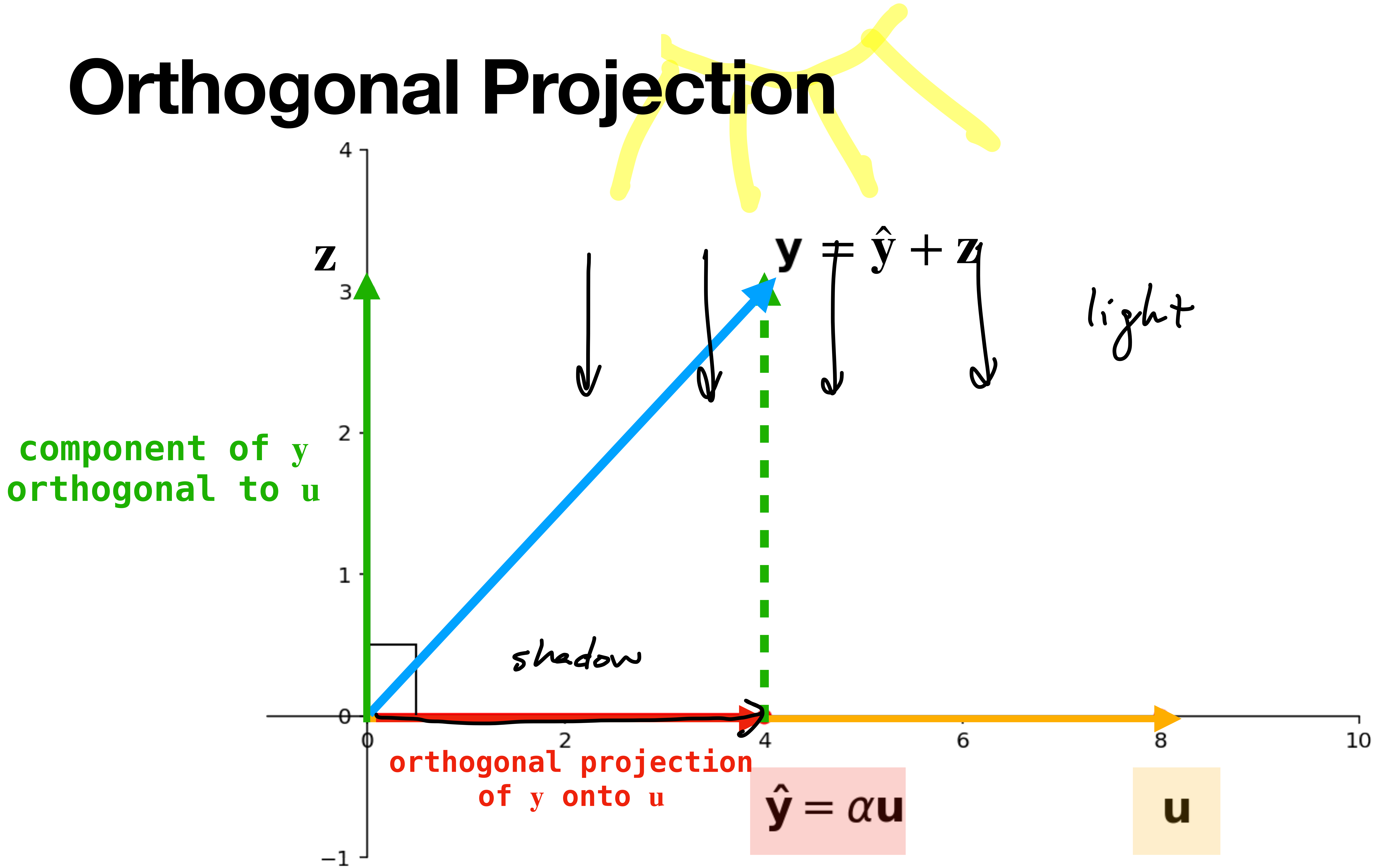
» $\mathbf{z}$ is orthogonal to $\mathbf{u}$ (i.e., $\mathbf{z} \cdot \mathbf{u} = 0$)

» $\hat{\mathbf{y}} \in span\{\mathbf{u}\}$

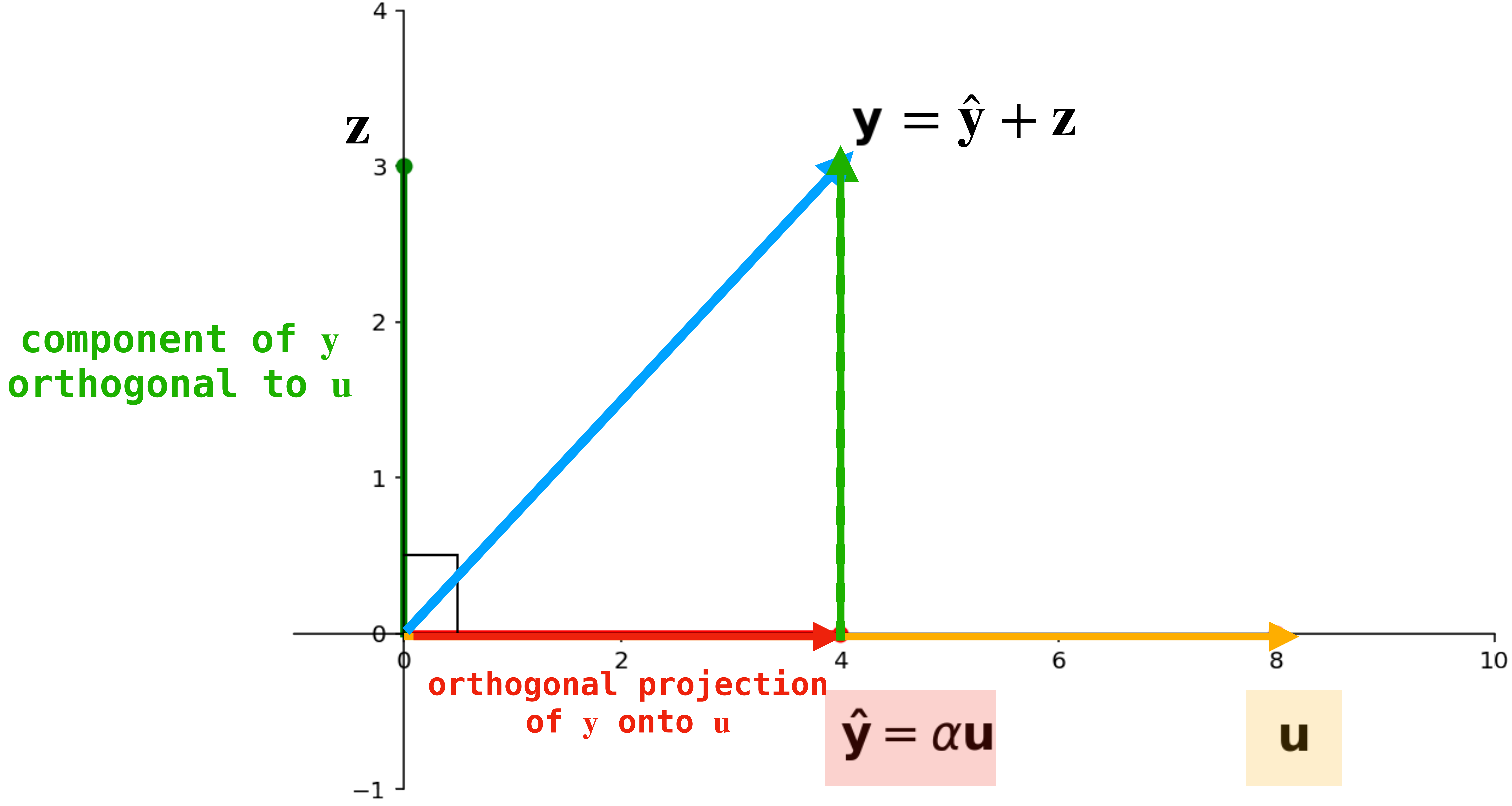
» $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$



Orthogonal Projection

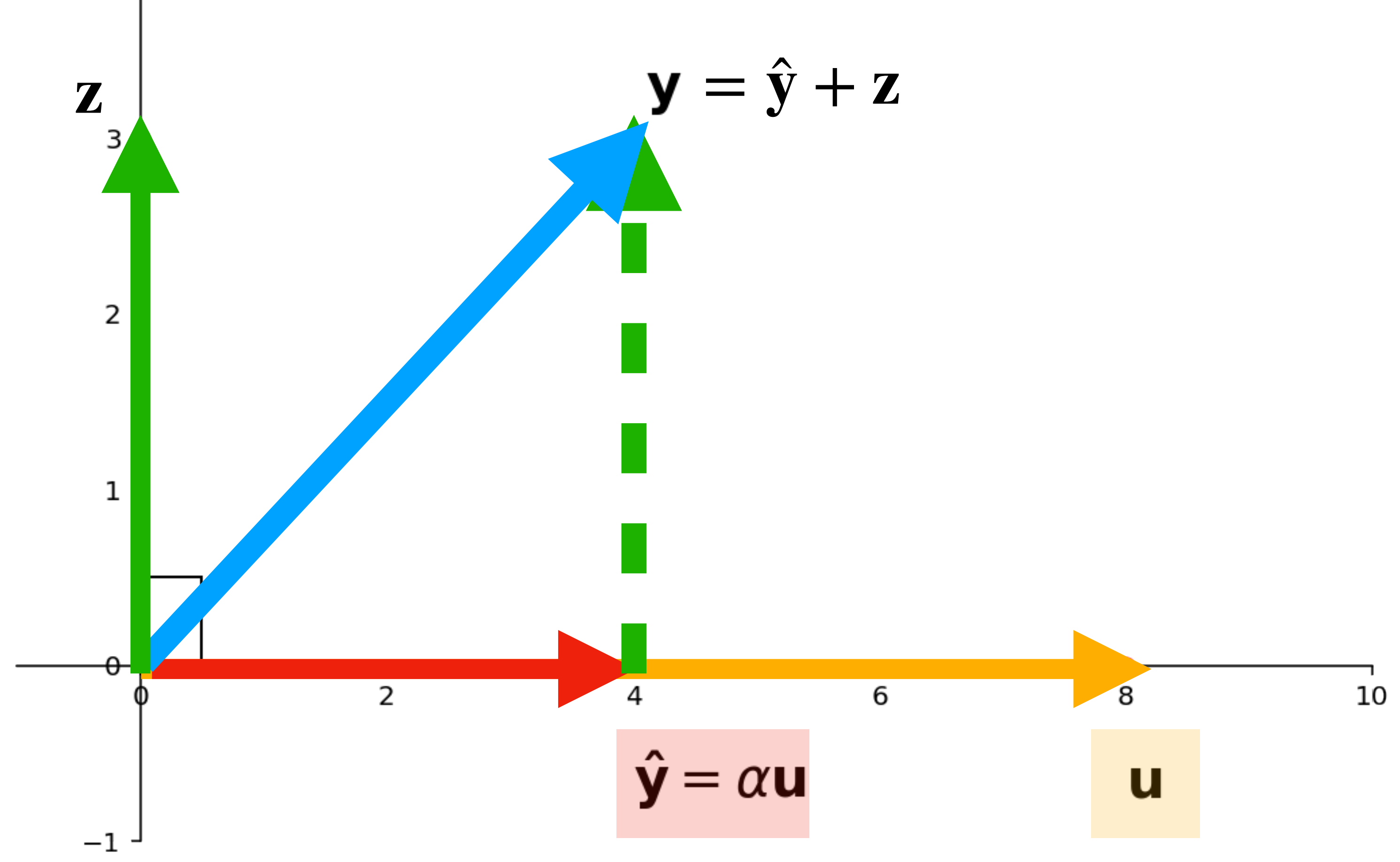


Orthogonal Projection



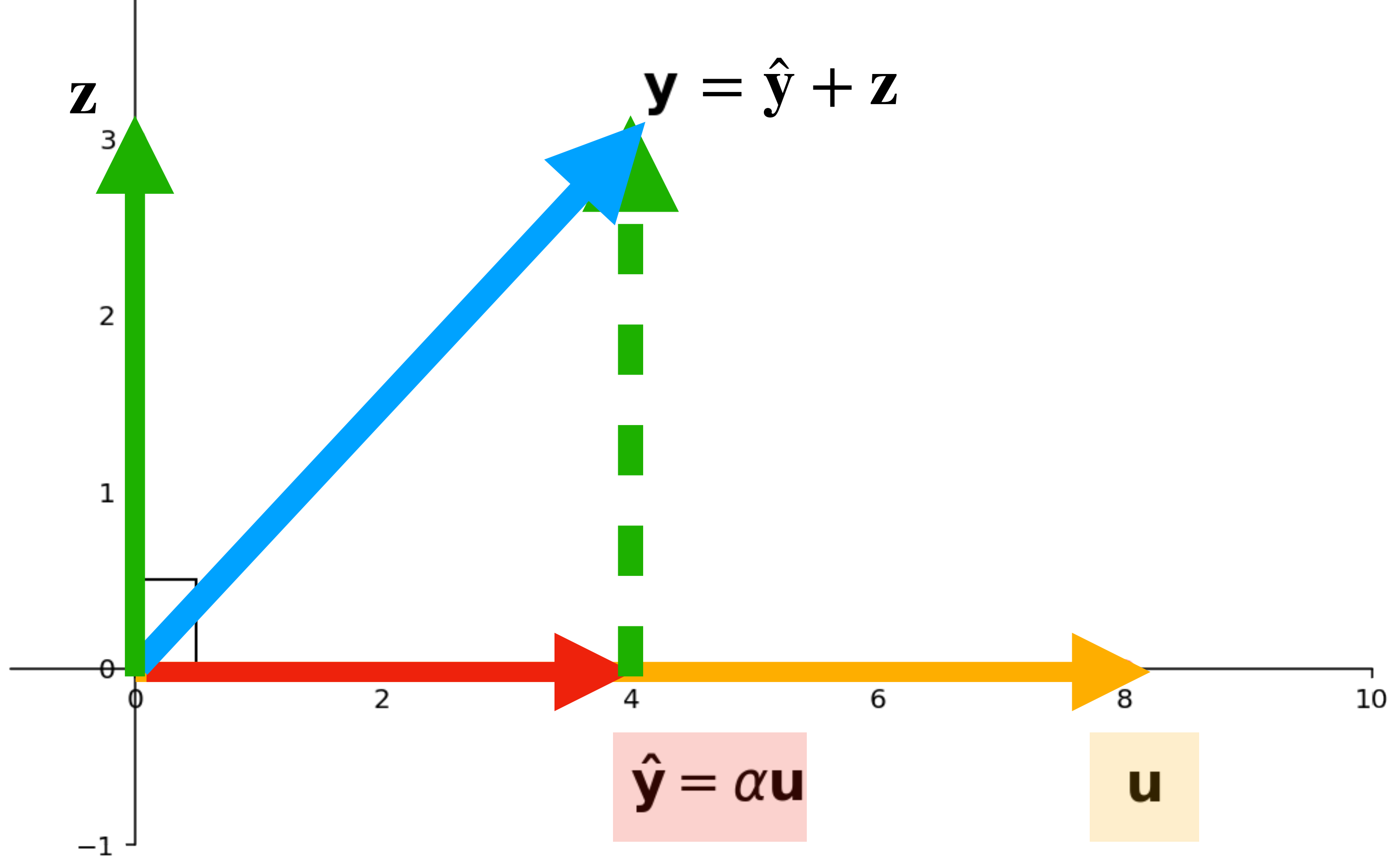
How do we find the orthogonal projection and orthogonal component?

What we know



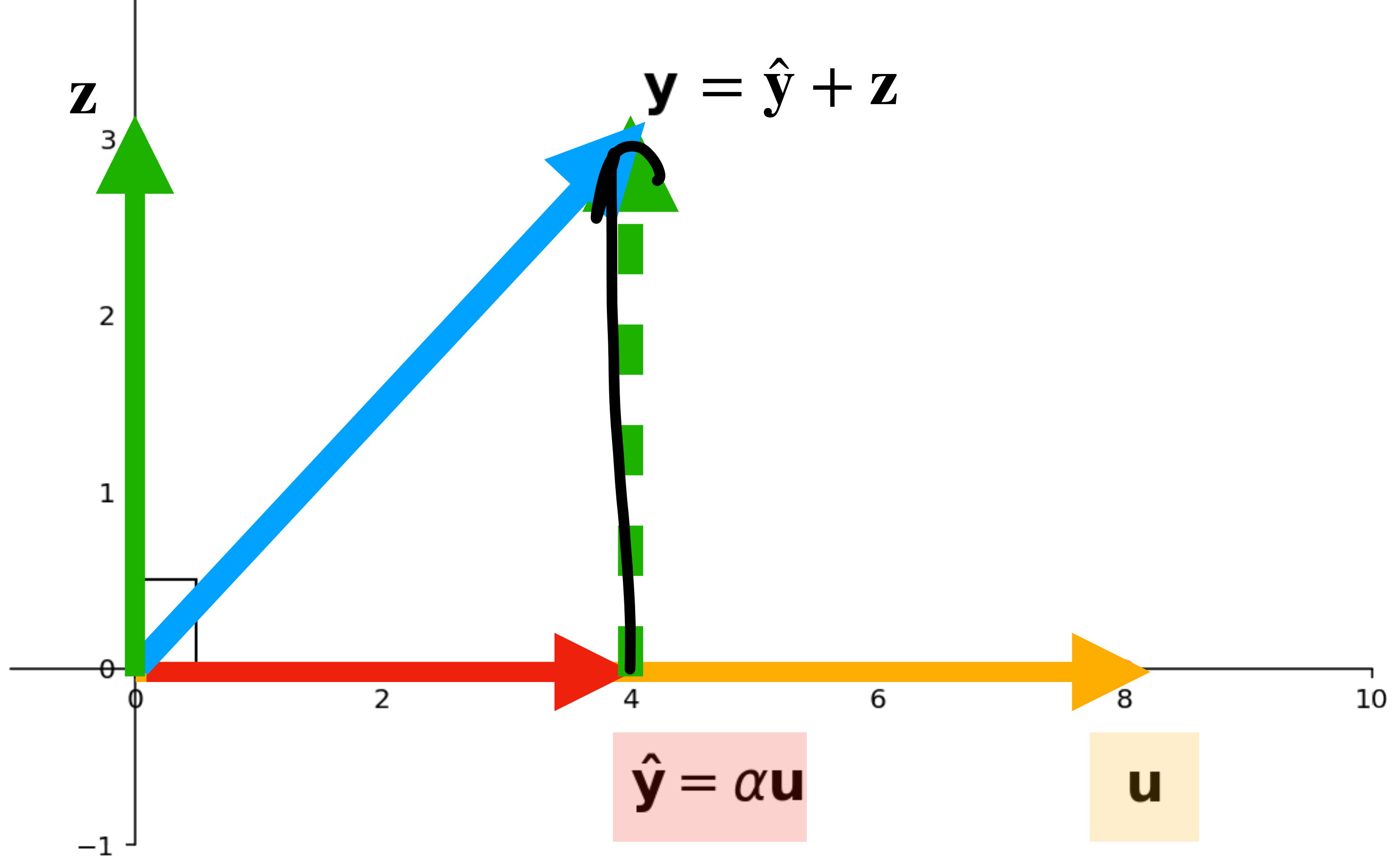
$\mathbf{z}$
$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$
3
2
1
0
−1
0
2
4
6
8
10
$\hat{\mathbf{y}} = \alpha\mathbf{u}$
$\mathbf{u}$

What we know



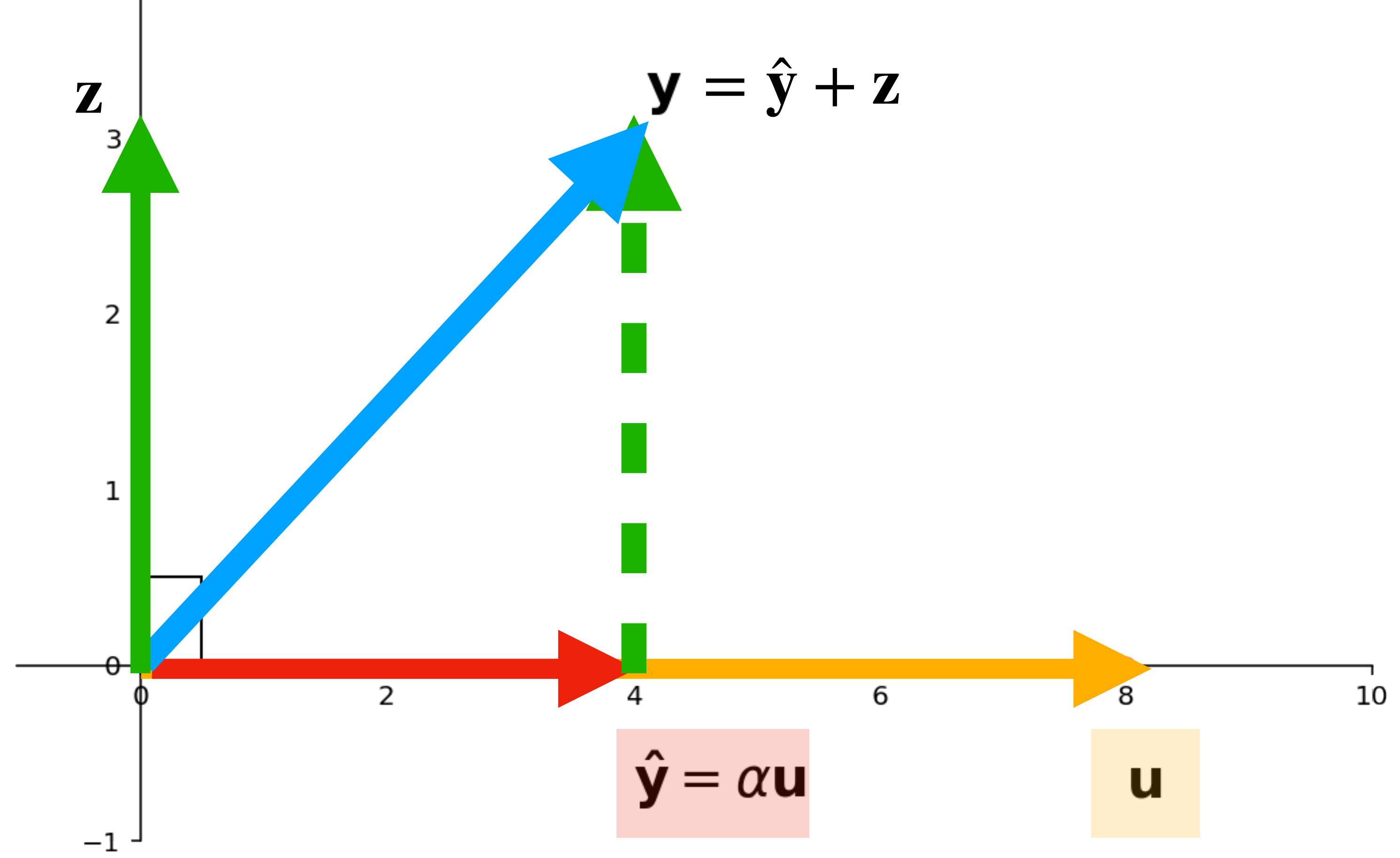
- $\hat{\mathbf{y}} = \alpha\mathbf{u}$ for some scalar α (since $\hat{\mathbf{y}} \in span\{\mathbf{u}\}$)

What we know



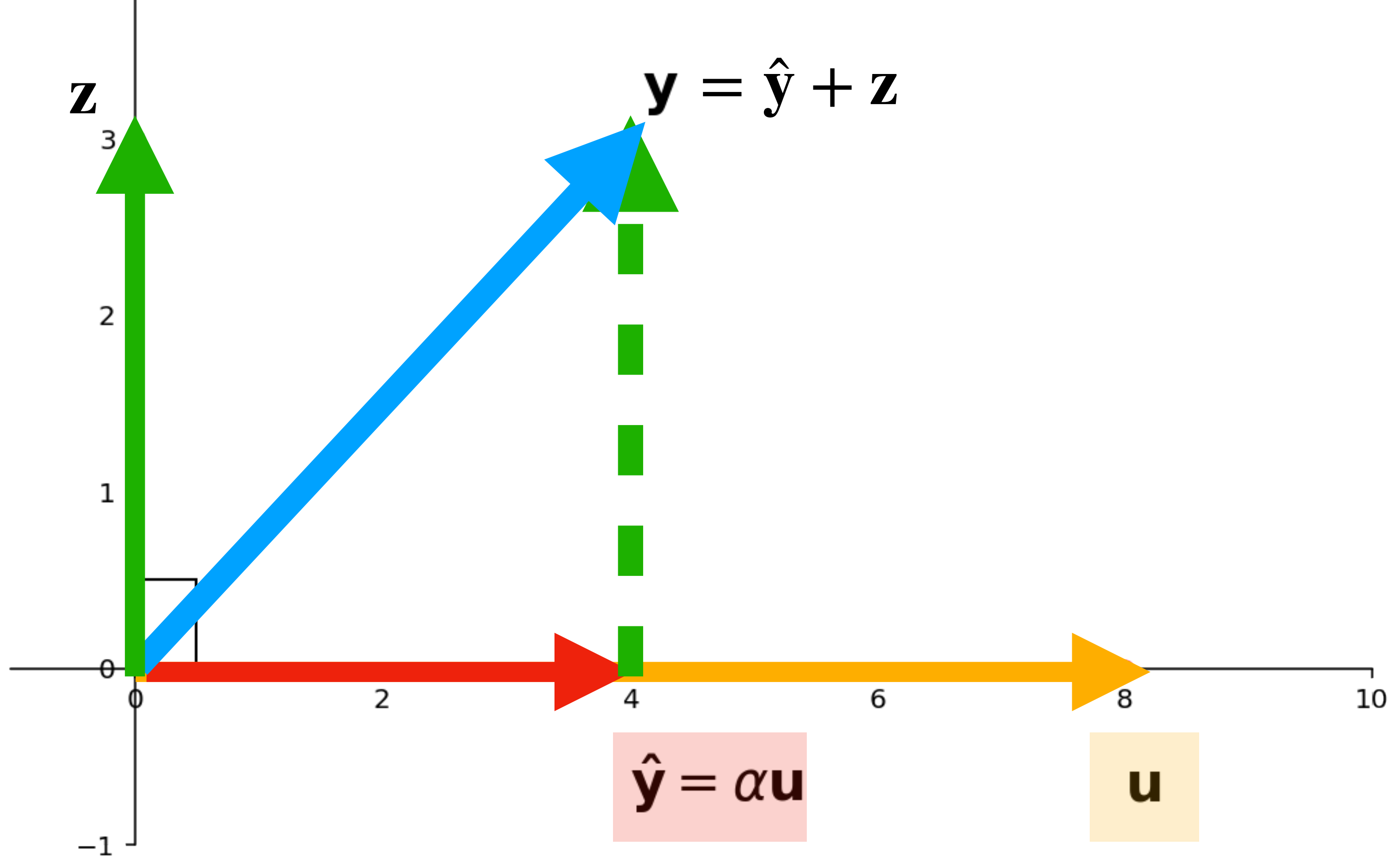
- $\hat{\mathbf{y}} = \alpha\mathbf{u}$ for some scalar α (since $\hat{\mathbf{y}} \in span\{\mathbf{u}\}$)
- $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \alpha\mathbf{u}$ (since $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$) $\hat{y} + \vec{y} - \hat{y} = \vec{y}$

What we know



- $\hat{\mathbf{y}} = \alpha\mathbf{u}$ for some scalar α (since $\hat{\mathbf{y}} \in span\{\mathbf{u}\}$)
- $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \alpha\mathbf{u}$ (since $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$)
- $\langle \mathbf{z}, \mathbf{u} \rangle = 0$ (since $\mathbf{z}$ is orthogonal with $\mathbf{u}$)

What we know

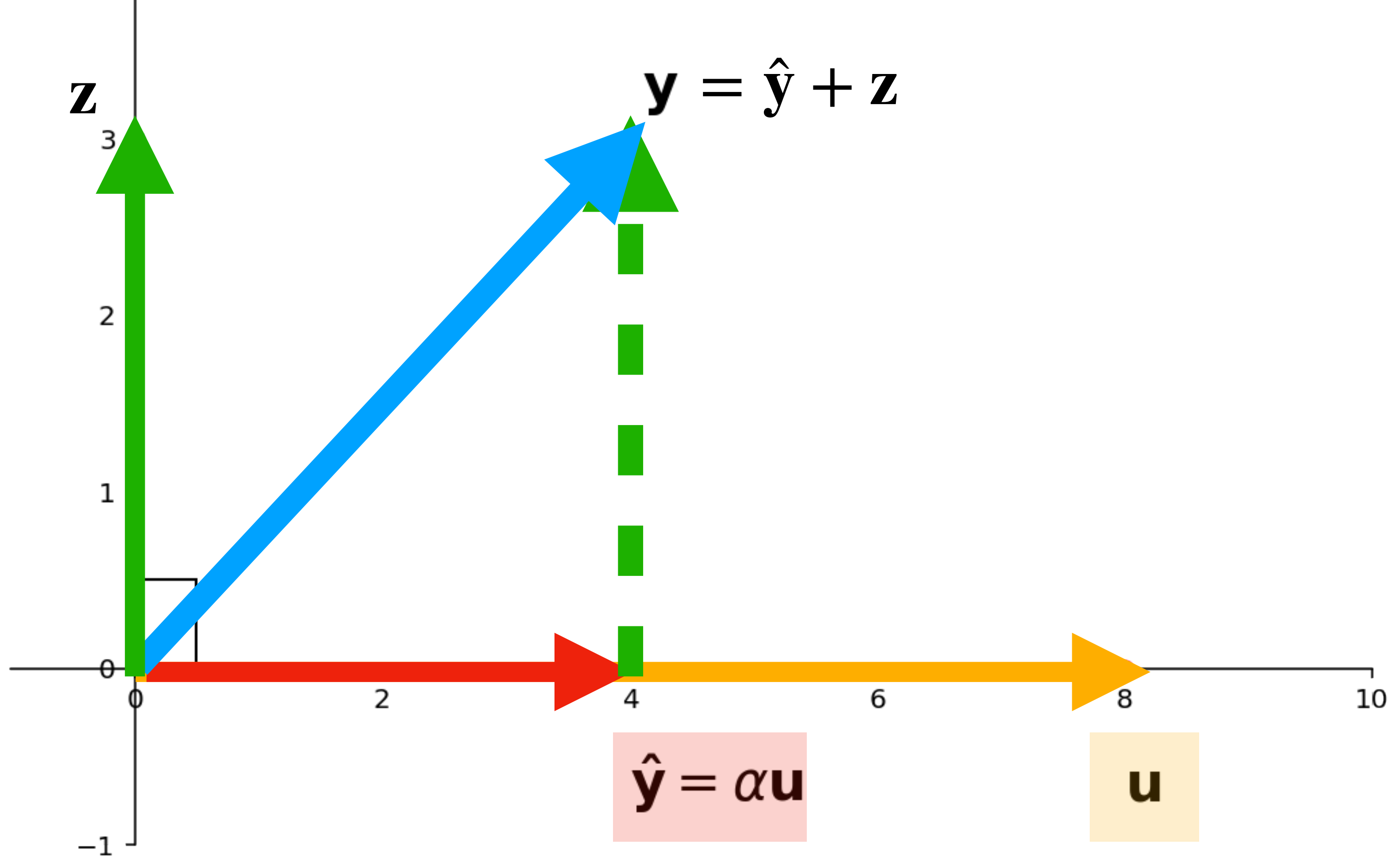


- $\hat{\mathbf{y}} = \alpha\mathbf{u}$ for some scalar α (since $\hat{\mathbf{y}} \in span\{\mathbf{u}\}$)
- $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \alpha\mathbf{u}$ (since $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$)
- $\langle \mathbf{z}, \mathbf{u} \rangle = 0$ (since $\mathbf{z}$ is orthogonal with $\mathbf{u}$)

Therefore:

$$\langle \mathbf{y} - \alpha\mathbf{u}, \mathbf{u} \rangle = 0$$

What we know



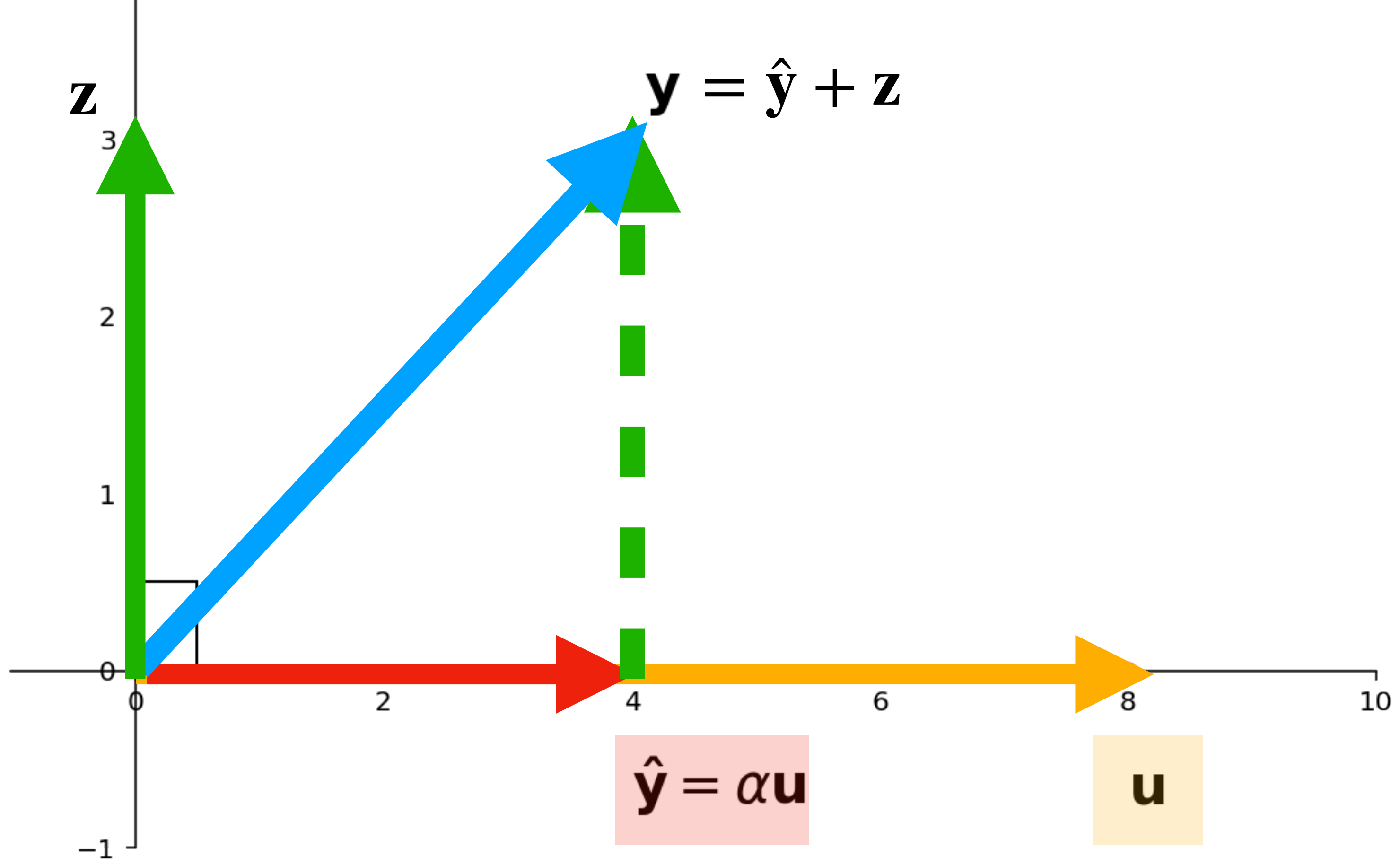
- $\hat{\mathbf{y}} = \alpha\mathbf{u}$ for some scalar α (since $\hat{\mathbf{y}} \in span\{\mathbf{u}\}$)
- $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \alpha\mathbf{u}$ (since $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$)
- $\langle \mathbf{z}, \mathbf{u} \rangle = 0$ (since $\mathbf{z}$ is orthogonal with $\mathbf{u}$)

Therefore:

$$\langle \mathbf{y} - \alpha\mathbf{u}, \mathbf{u} \rangle = 0$$

Once we have α, we can compute both $\hat{\mathbf{y}}$ and $\mathbf{z}$

Step 1: Finding α



$$\langle \mathbf{y} - \alpha\mathbf{u}, \mathbf{u} \rangle = 0$$

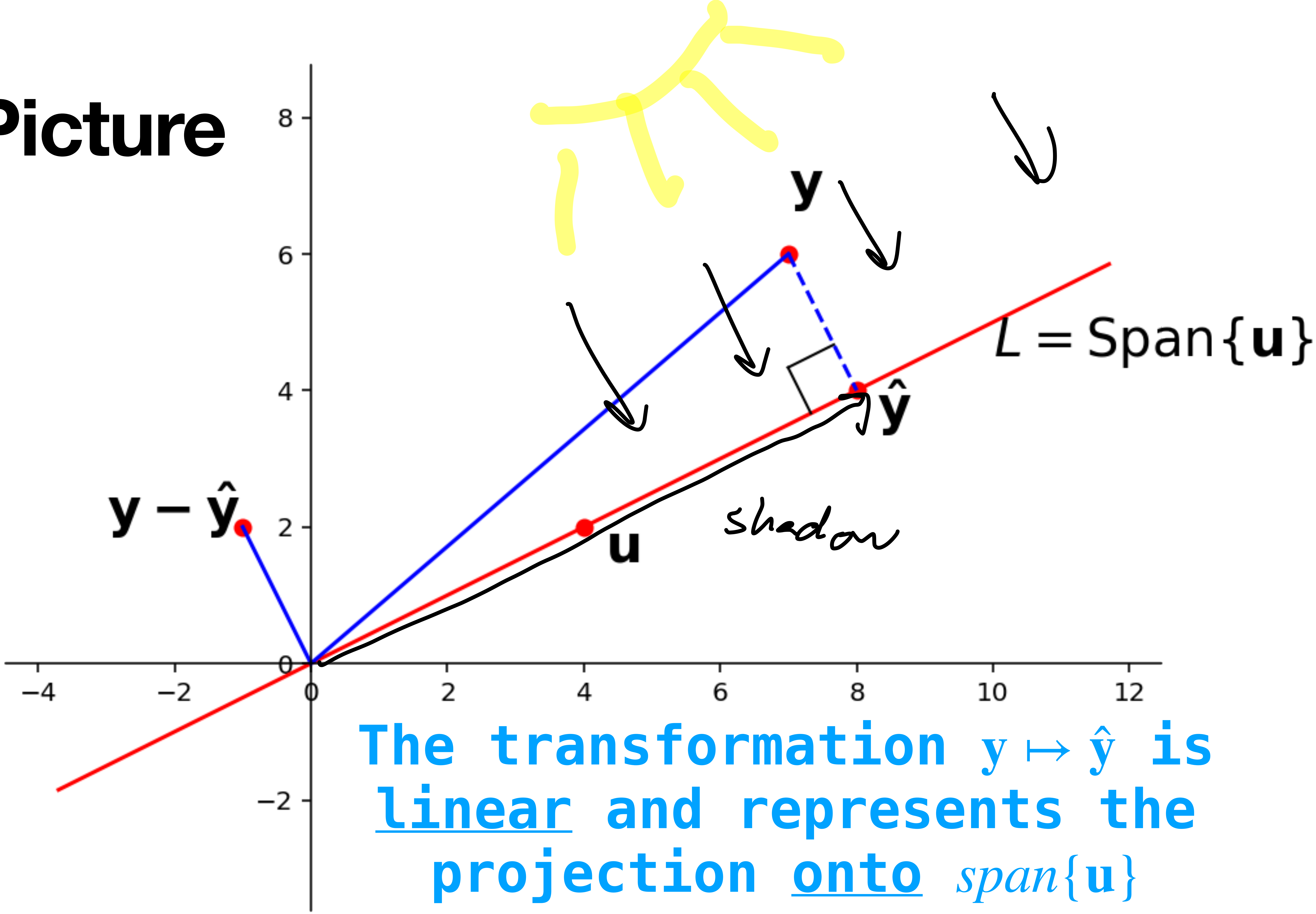
Let's solve for α, $\hat{\mathbf{y}}$ and $\mathbf{z}$:

$$\langle \vec{y} - \alpha\vec{u}, \vec{u} \rangle = \langle \vec{y}, u \rangle - \alpha \langle \vec{u}, \vec{u} \rangle = 0$$

$$\alpha = \frac{\langle y, u \rangle}{\langle u, u \rangle}$$

exactly equation from previous slide

The Picture



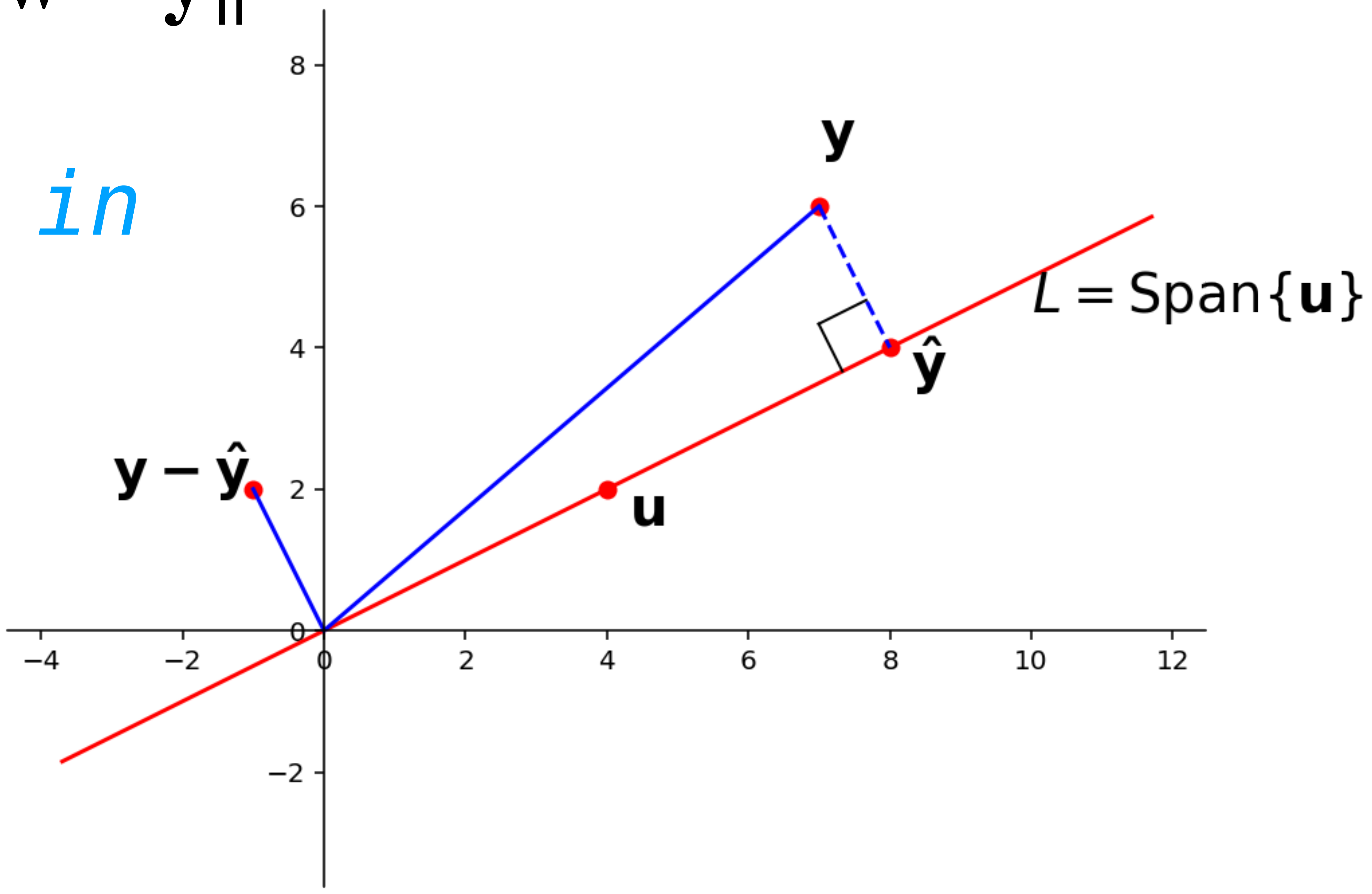
The transformation $y \mapsto \hat{y}$ is linear and represents the projection onto $span\{\mathbf{u}\}$

$\hat{\mathbf{y}}$ and Distance

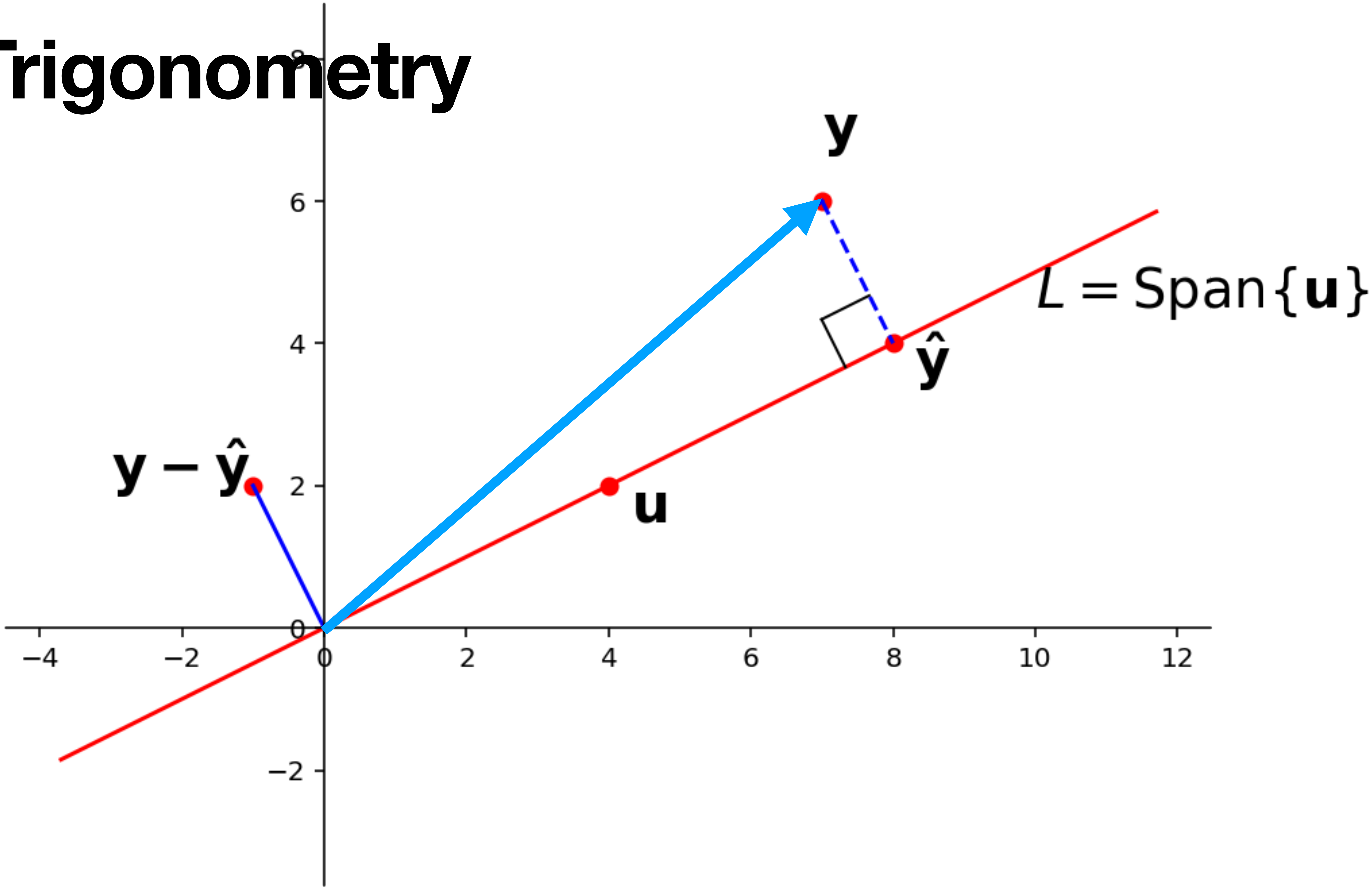
Theorem. $\|\hat{\mathbf{y}} - \mathbf{y}\| = \min_{\mathbf{w} \in span\{\mathbf{u}\}} \|\mathbf{w} - \mathbf{y}\|$

$\hat{\mathbf{y}}$ is the closest vector in $span\{\mathbf{u}\}$ to $\mathbf{y}$.

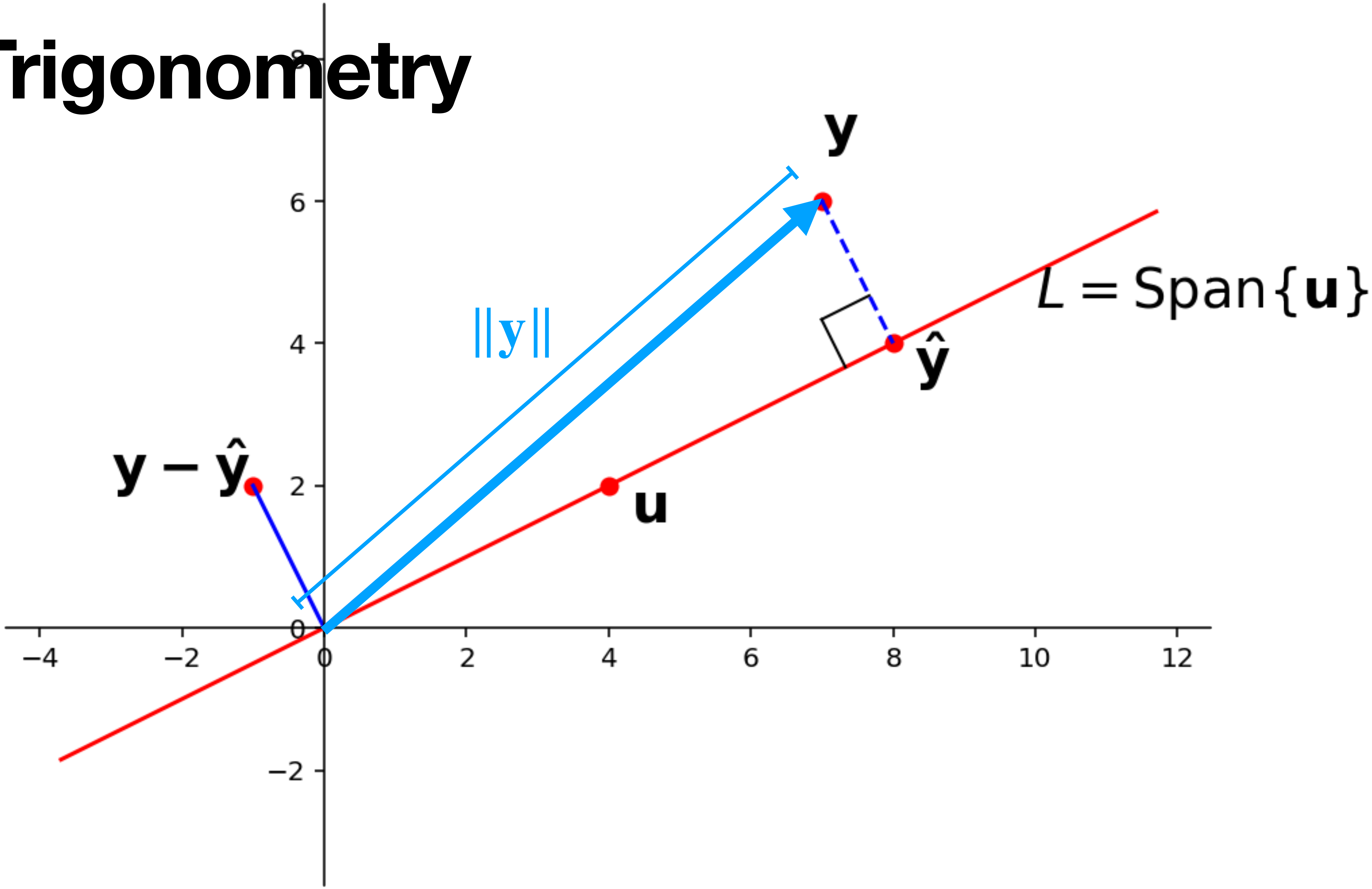
"Proof" by inspection:



The Trigonometry

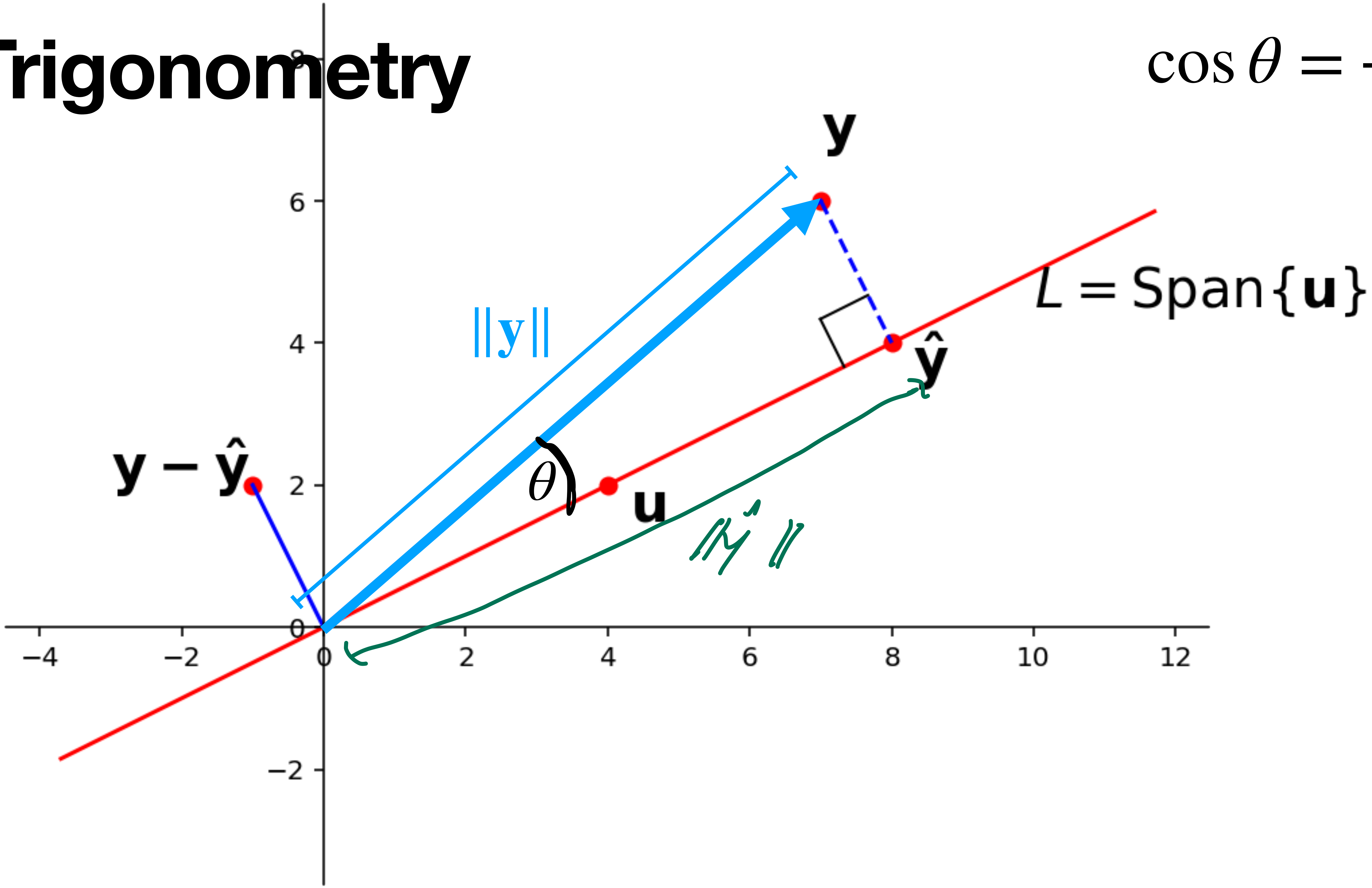


The Trigonometry



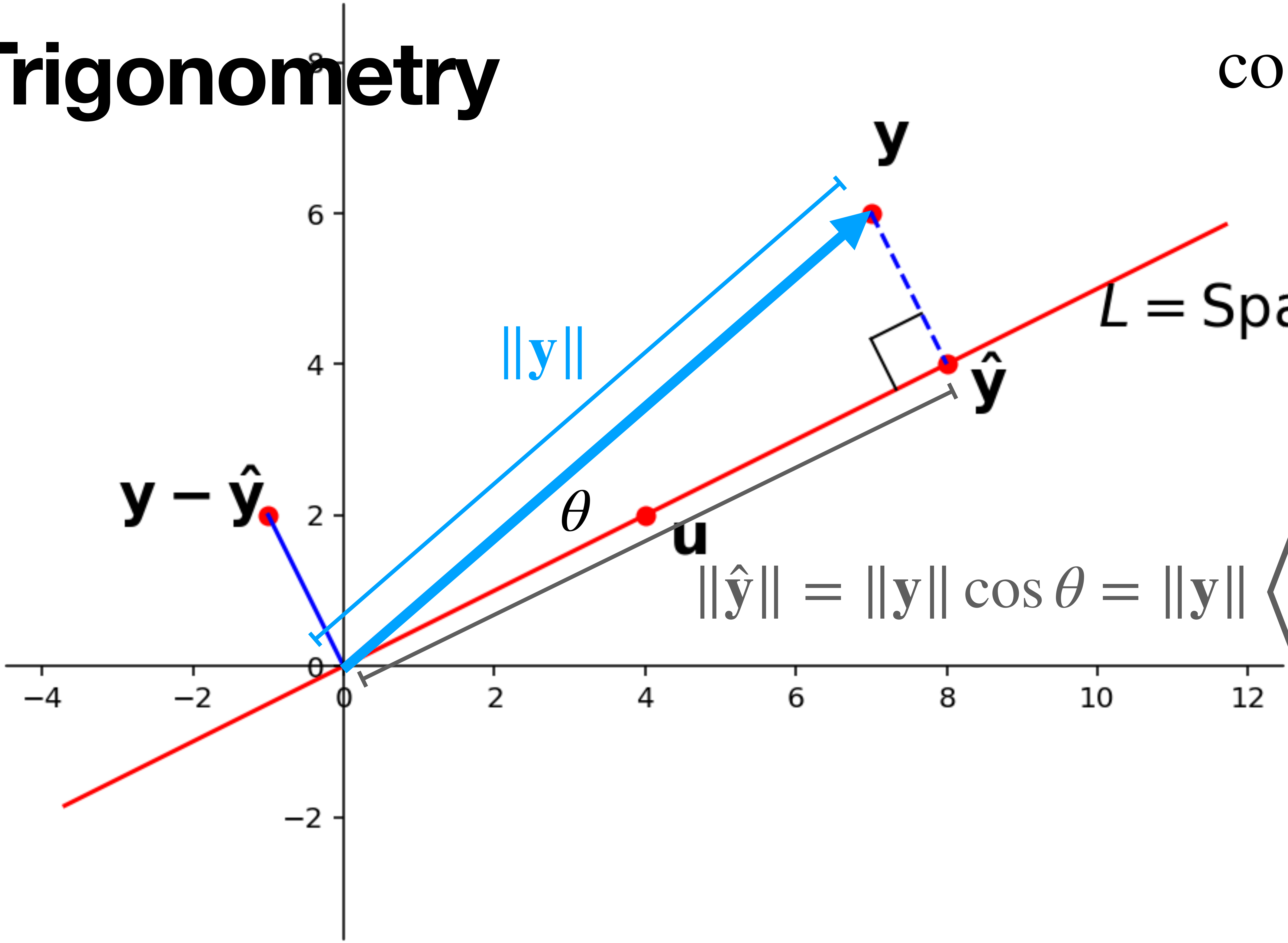
The Trigonometry

$$\cos\theta = \frac{\|\hat{\mathbf{y}}\|}{\|\mathbf{y}\|}$$



The Trigonometry

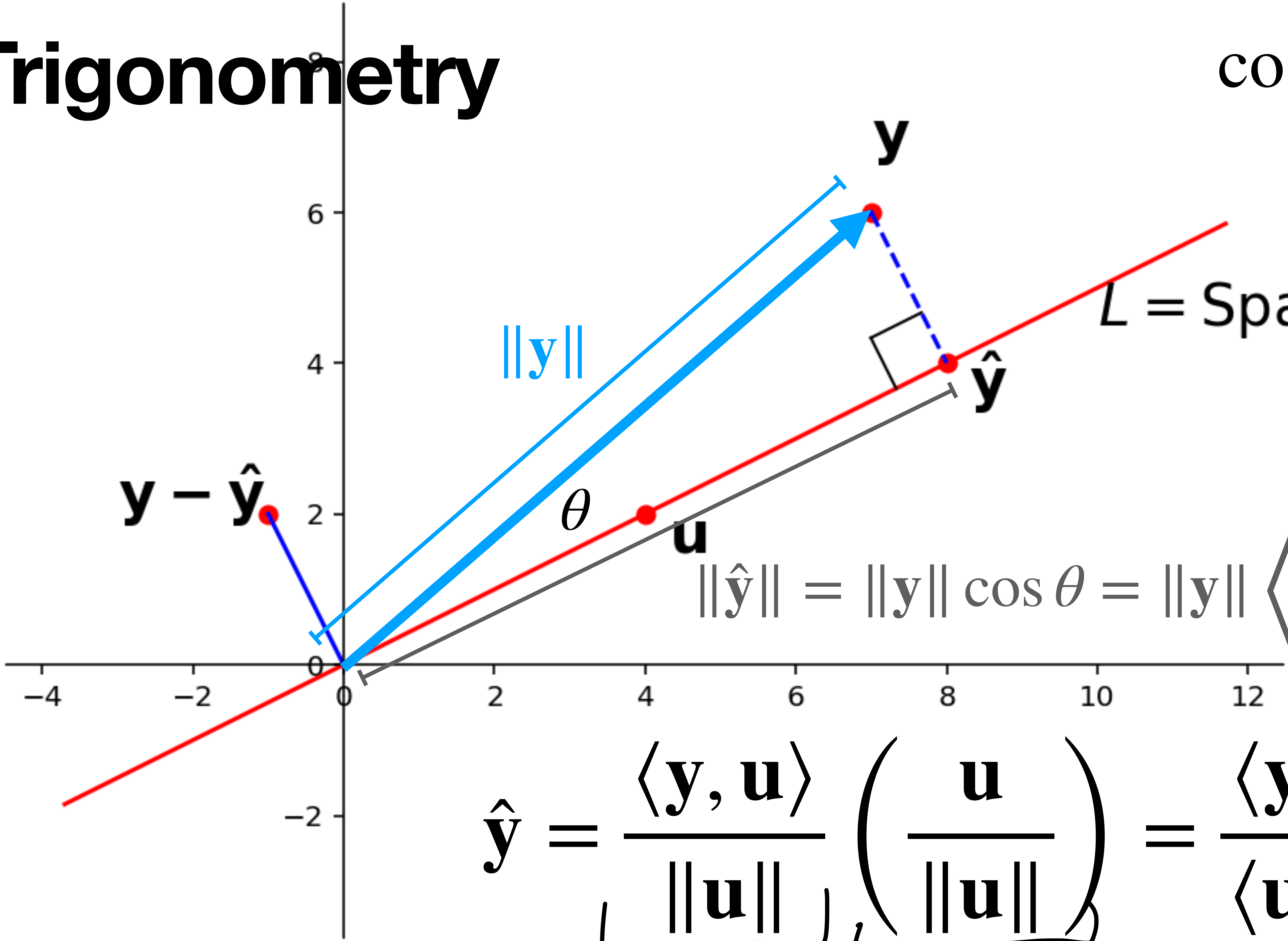
$$\cos\theta = \frac{\|\hat{\mathbf{y}}\|}{\|\mathbf{y}\|}$$



$$\|\hat{\mathbf{y}}\| = \|\mathbf{y}\|\cos\theta = \|\mathbf{y}\|\left\langle \frac{\mathbf{y}}{\|\mathbf{y}\|}, \frac{\mathbf{u}}{\|\mathbf{u}\|} \right\rangle$$

The Trigonometry

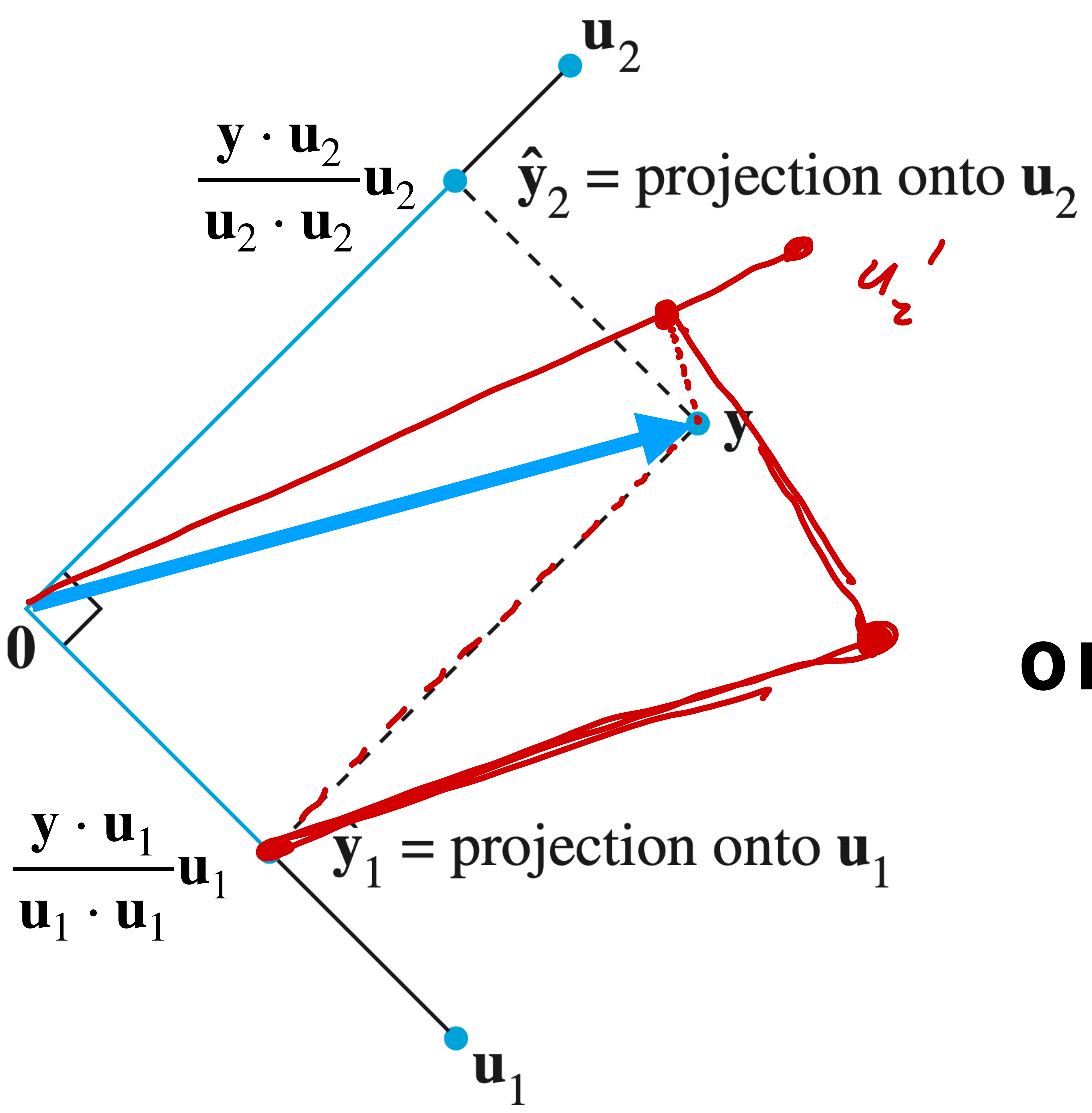
$$\cos\theta = \frac{\|\hat{\mathbf{y}}\|}{\|\mathbf{y}\|}$$



$$\|\hat{\mathbf{y}}\| = \|\mathbf{y}\|\cos\theta = \|\mathbf{y}\|\left\langle \frac{\mathbf{y}}{\|\mathbf{y}\|}, \frac{\mathbf{u}}{\|\mathbf{u}\|} \right\rangle$$

$$\hat{\mathbf{y}} = \underbrace{\frac{\langle \mathbf{y}, \mathbf{u}\rangle}{\|\mathbf{u}\|}}_{\|\hat{y}\|} \underbrace{\left(\frac{\mathbf{u}}{\|\mathbf{u}\|}\right)}_{\text{unit vector}} = \frac{\langle \mathbf{y}, \mathbf{u}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u}$$

Orthogonal Projections and Orthogonal Bases



Each <u>component</u> of $\mathbf{y}$ written in terms of an *orthogonal* basis is an **orthogonal projection onto to a basis vector**

How To:

Question. Find the projection of $\mathbf{y}$ onto the span of $\mathbf{u}$

Solution. Calculate $\alpha = \dfrac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}$, then the solution is $\alpha\mathbf{u}$

Question

$$\alpha = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}$$

Find the matrix which implements orthogonal projection onto the span of $\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$

Answer

$$\frac{1}{\sqrt{5}}\begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{bmatrix}$$

Orthonormal Sets

Orthogonal sets would be easier to work with if every vector was a <u>unit</u> vector

Orthonormality

Orthonormality

Definition. A set $\{u_1, u_2, \ldots, u_p\}$ is an **orthonormal set** if of it an orthogonal set of <u>unit</u> vectors

Orthonormality

Definition. A set $\{u_1, u_2, \ldots, u_p\}$ is an **orthonormal set** if of it an orthogonal set of unit vectors

Definition. An **orthonormal basis** of the subspace W is a basis of W which is an orthonormal set

Orthonormality

Definition. A set $\{u_1, u_2, \ldots, u_p\}$ is an **orthonormal set** if of it an orthogonal set of unit vectors

Definition. An **orthonormal basis** of the subspace W is a basis of W which is an orthonormal set

ortho·normal

Orthonormality

Definition. A set $\{u_1, u_2, \ldots, u_p\}$ is an **orthonormal set** if of it an orthogonal set of unit vectors

Definition. An **orthonormal basis** of the subspace W is a basis of W which is an orthonormal set

ortho·normal

orthogonal/perpendicular

Orthonormality

Definition. A set $\{u_1, u_2, \ldots, u_p\}$ is an **orthonormal set** if of it an orthogonal set of <u>unit</u> vectors

Definition. An **orthonormal basis** of the subspace W is a basis of W which is an orthonormal set

ortho·normal

orthogonal/perpendicular normalized/made unit vectors

Orthonormal Matrices

Definition. A matrix is **orthonormal** if its columns form an orthonormal set

The notes call a square orthonormal matrix an **orthogonal** matrix.

Orthonormal Matrices

Definition. A matrix is **orthonormal** if its columns form an orthonormal set

The notes call a square orthonormal matrix an **orthogonal** matrix.

This is incredibly confusing, but we'll try to be consistent and clear

Orthonormal Matrices and Transposition

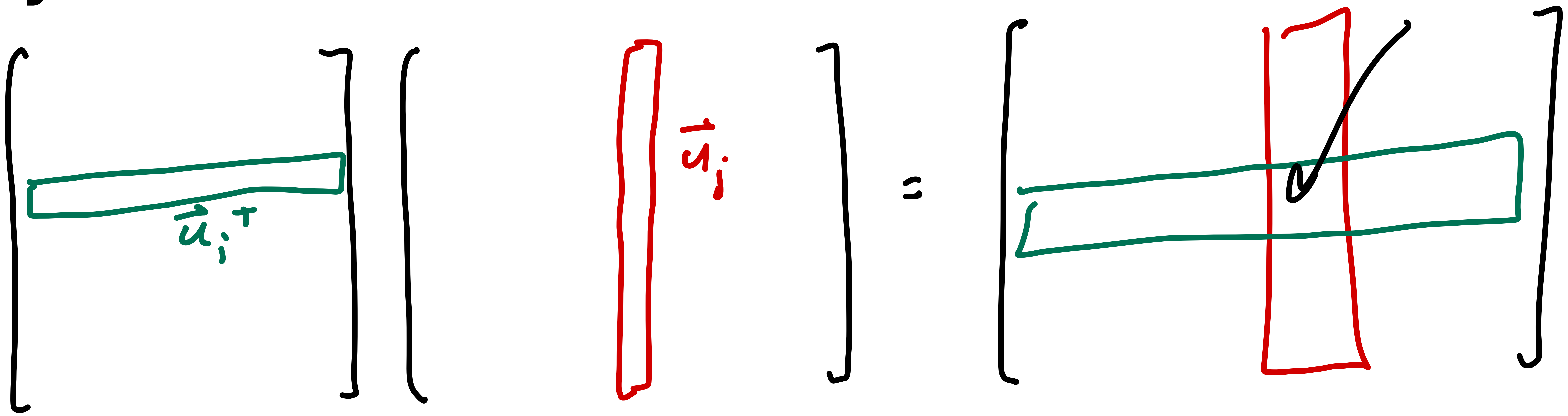
Theorem. For an $m \times n$ orthonormal matrix U

$$(U^TU)_{ij} = \langle u_i, u_j \rangle = \begin{cases} 1 & i=j \\ 0 & \text{otherwise} \end{cases}$$

$$U^TU = I_n$$

$$\vec{u}_i^T \vec{u}_j = \langle u_i, u_j \rangle$$

Verify:



Inverses of Orthogonal Matrices

Theorem. If an $n \times n$ matrix U is orthogonal (square orthonormal) then it is invertible and

$$U^{-1} = U^T$$

Verify:

$$U U^T = I$$

Orthonormal Matrices and Inner Products

Theorem. For a $m \times n$ orthonormal matrix U, and any vectors x and y in R^n

$$\langle Ux, Uy \rangle = \langle x, y \rangle$$

Orthonormal matrices preserve inner products

Verify: $\langle Ux, Uy \rangle = (Ux)^T (Uy)$

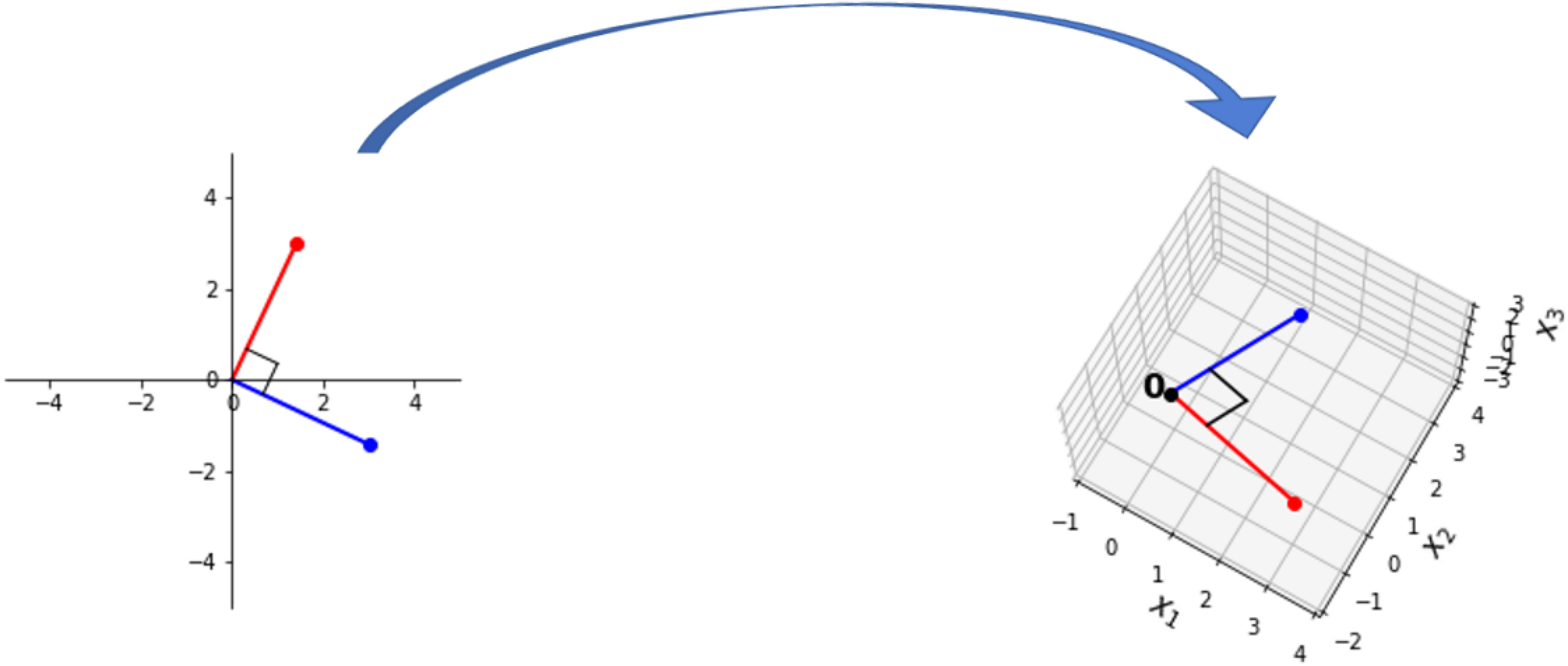
$= x^T U^T U y = x^T y$

Length, Angle, Orthogonality Preservation

Since lengths and angles are defined in terms of inner products, they are also preserved by orthonormal matrices:

The Picture

Orthonormal U



Example

$$U = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \qquad x = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$$

Question (Conceptual)

Suppose A is an $m \times n$ matrix with orthogonal but ***not*** *orthonormal columns. What is $A^T A$?*

Answer

If $A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_n]$ then $A^T A$ is a diagonal matrix D where

$$D_{ii} = \|\mathbf{a}_i\|^2$$

Summary

Orthogonal sets allow for simpler calculations of coordinates

Finding these coordinates is a really about find the orthogonal projections onto each vector in the orthogonal set

We can apply these ideas to matrices and describe a class of very well behaved transformations via orthonormal matrices