

Orthogonal Sets

Geometric Algorithms

Lecture 22

Practice Problem

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2$$

$$[\vec{v}]_1 \rightarrow [\vec{v}]_{\mathcal{B}} \quad [\vec{b}_1 \dots \vec{b}_n]^{-1} \mathcal{B} = \left(\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}, \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} \right)$$

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

$$\left[\begin{array}{cc|c} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 2 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 3 \end{array} \right]$$

$\nwarrow B^{-1}$
change-of-basis matrix

$$\frac{1}{\det(B)} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2\frac{\sqrt{2}}{2} + 3\frac{\sqrt{2}}{2} \\ -2\frac{\sqrt{2}}{2} + 3\frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} 5\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$

$$\det(B) = \left(\frac{\sqrt{2}}{2} \right)^2 + \left(\frac{\sqrt{2}}{2} \right)^2 = \frac{2}{4} + \frac{2}{4} = 1$$

Answer

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

$$\mathcal{B} = \left(\underbrace{\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 2 \end{bmatrix}}_{\vec{v}_1}, \underbrace{\begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 2 \end{bmatrix}}_{\vec{v}_2} \right)$$

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

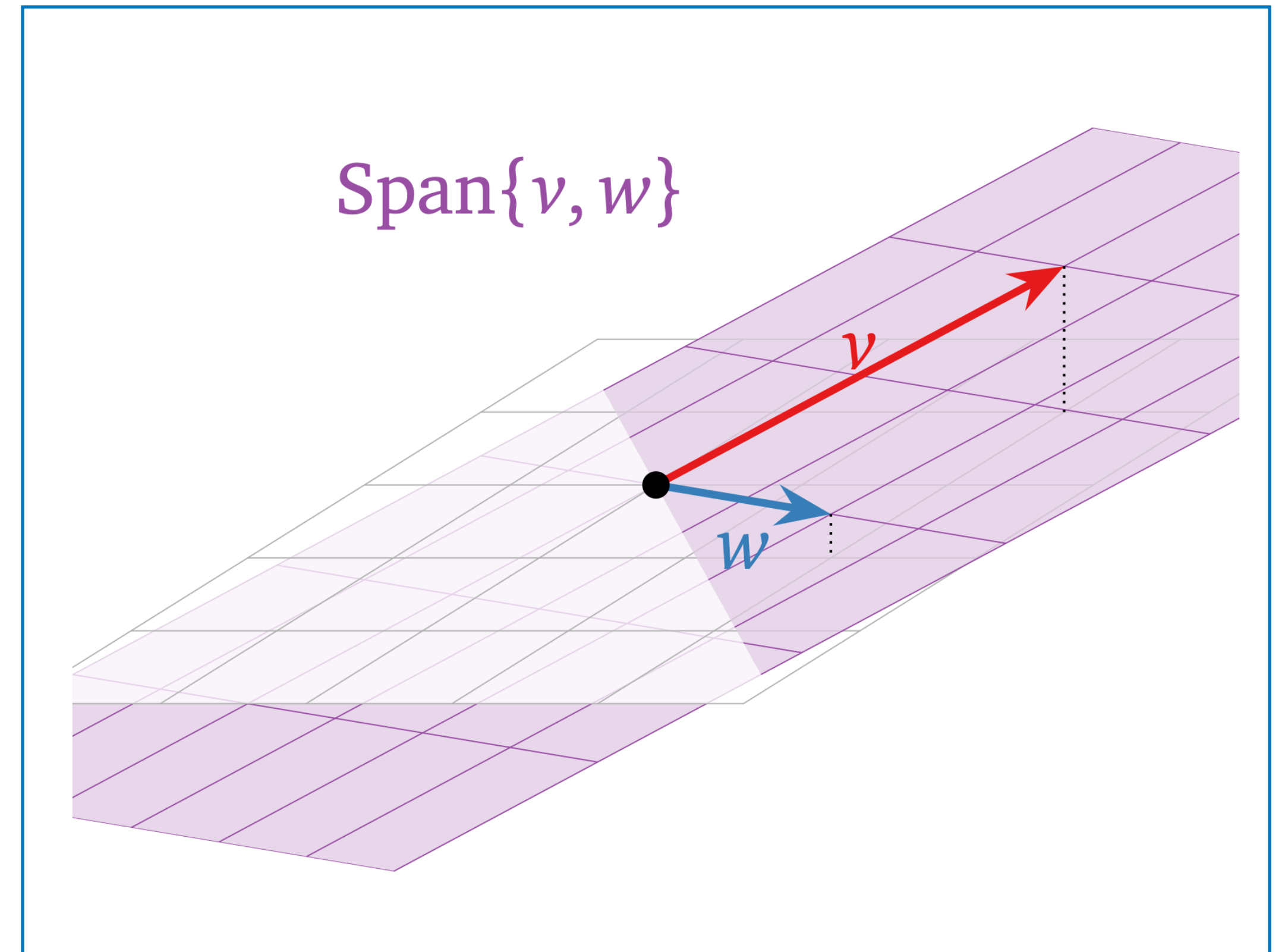
Recap: Analytic Geometry

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

$$[u_1 \quad u_2 \quad u_3 \quad u_4] \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = u_1 v_1 + u_2 v_2 + u_3 v_3 + u_4 v_4$$

Definition. The **inner product** of two vectors **u** and **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

$$[u_1 \ u_2 \ u_3 \ u_4] \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = u_1 v_1 + u_2 v_2 + u_3 v_3 + u_4 v_4$$

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

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

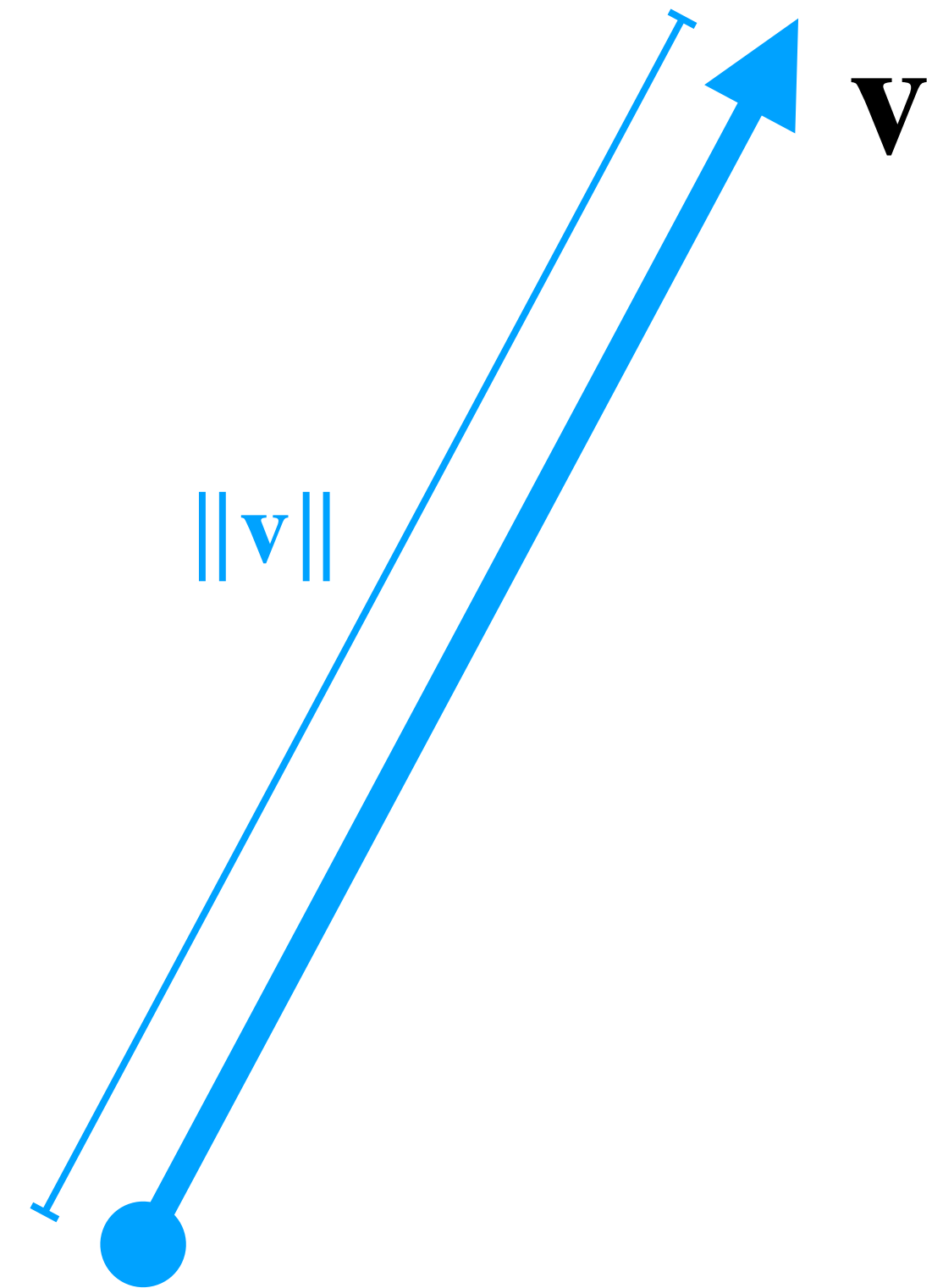
The norm of a vector is the square root of the inner product with itself.

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

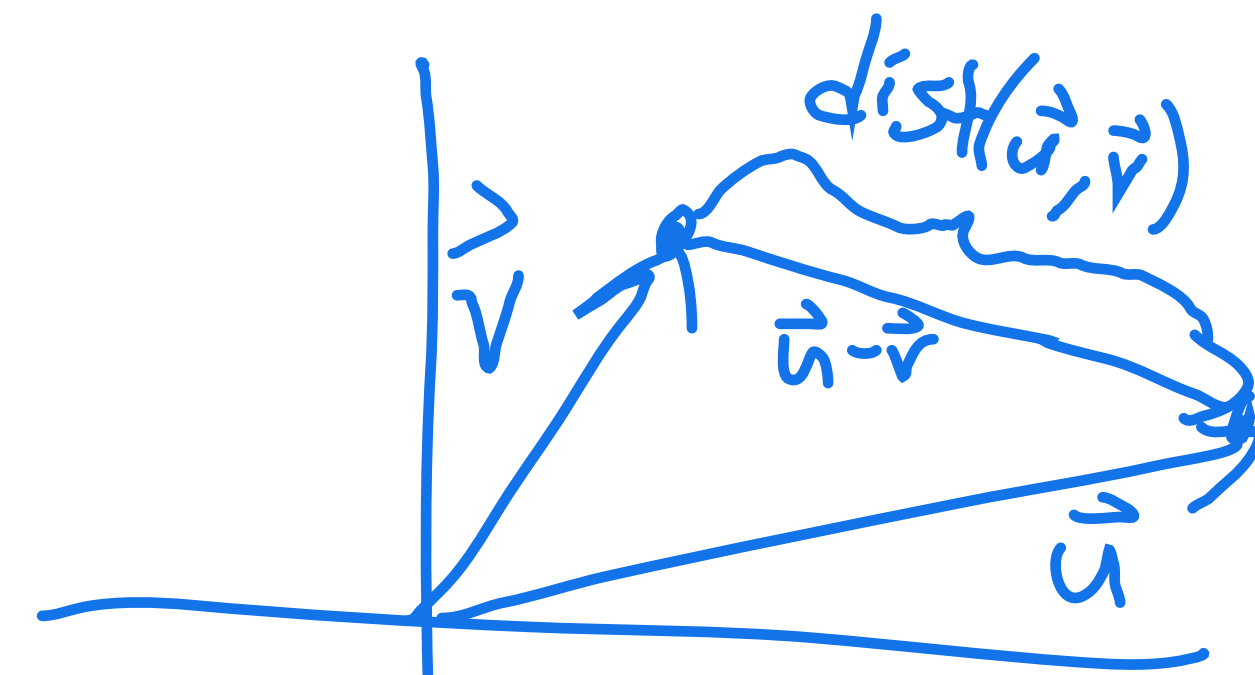
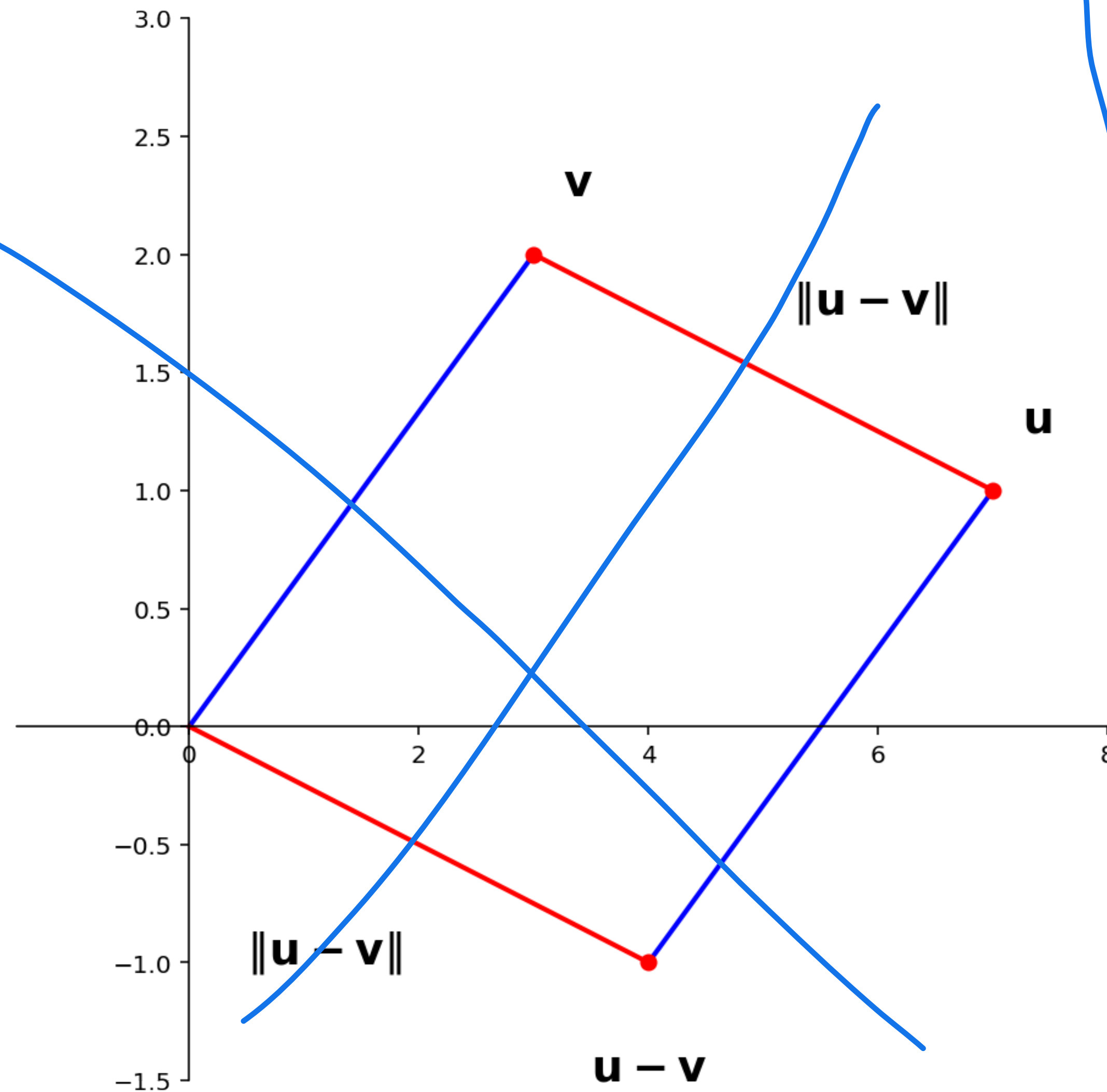
Recall: Norms and Length

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)



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

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

Recall: Distance (Algebraically)

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

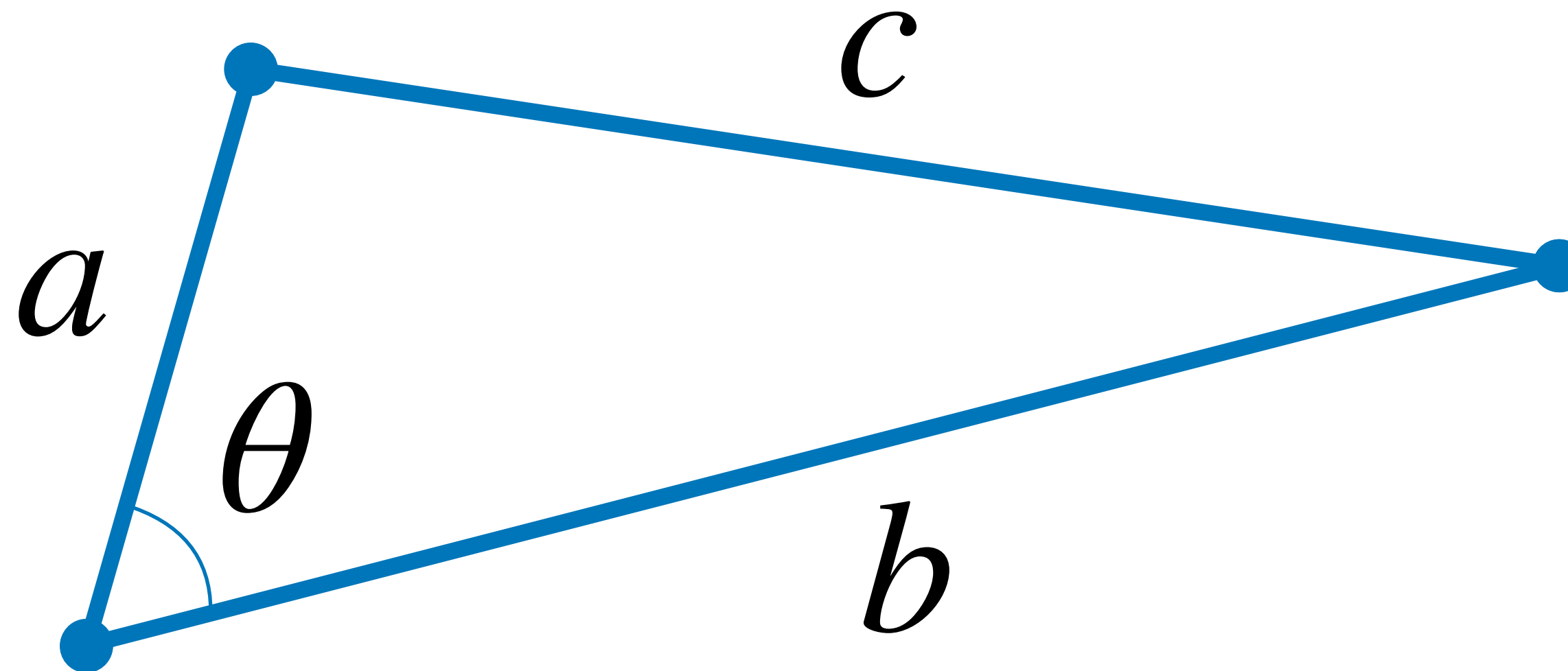
$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

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

$$\vec{u} - \vec{v} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$\|\vec{u} - \vec{v}\| = \sqrt{16+1} = \sqrt{17}$$

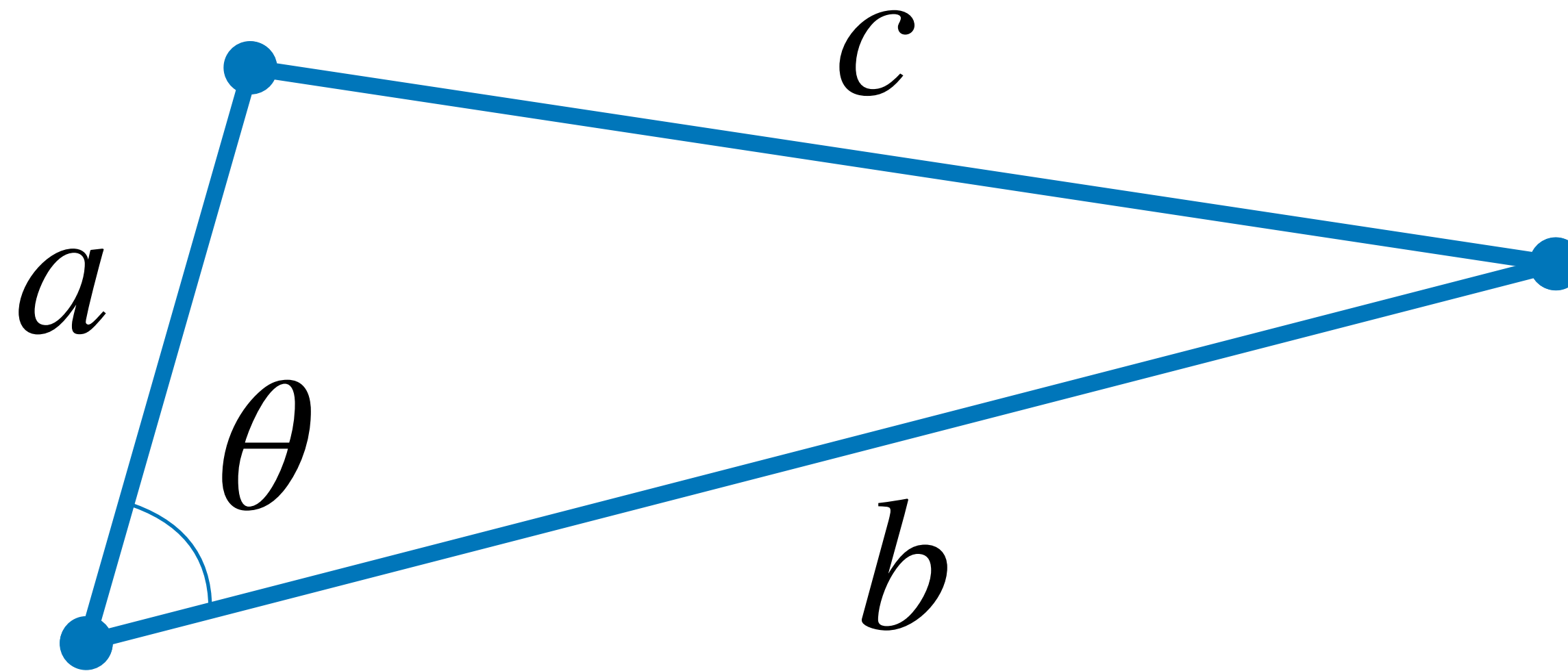
Recall: Law of Cosines



Theorem.

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

Recall: Law of Cosines

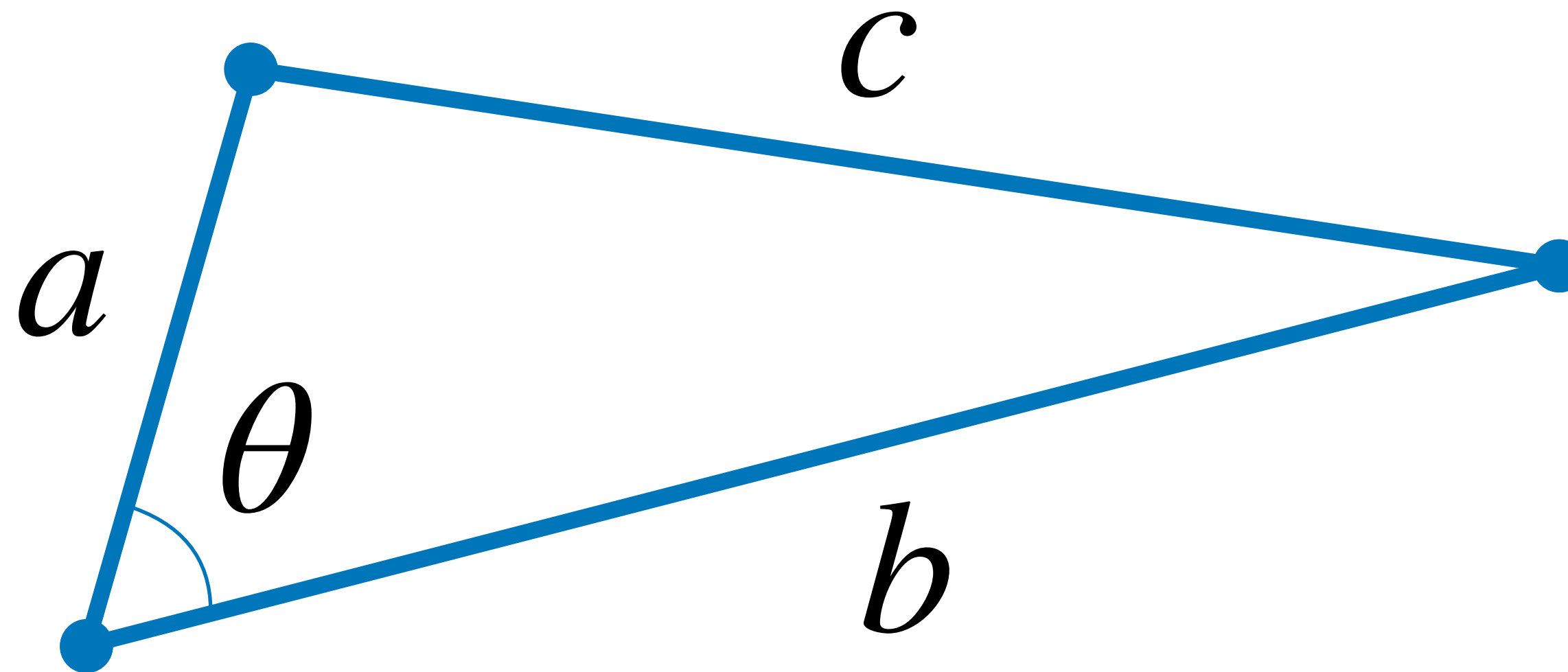


Theorem.

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

Generalized the Pythagorean Theorem

Recall: Law of Cosines



Theorem.

θ 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,

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

Handwritten notes in blue: "direction-only (scale/length-independent)" with arrows pointing to the unit vectors in the formula, and an arrow pointing to the inner product term.

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,

$$\cos 90^\circ = \cos \frac{\pi}{2} = 0$$

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

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

Orthogonality

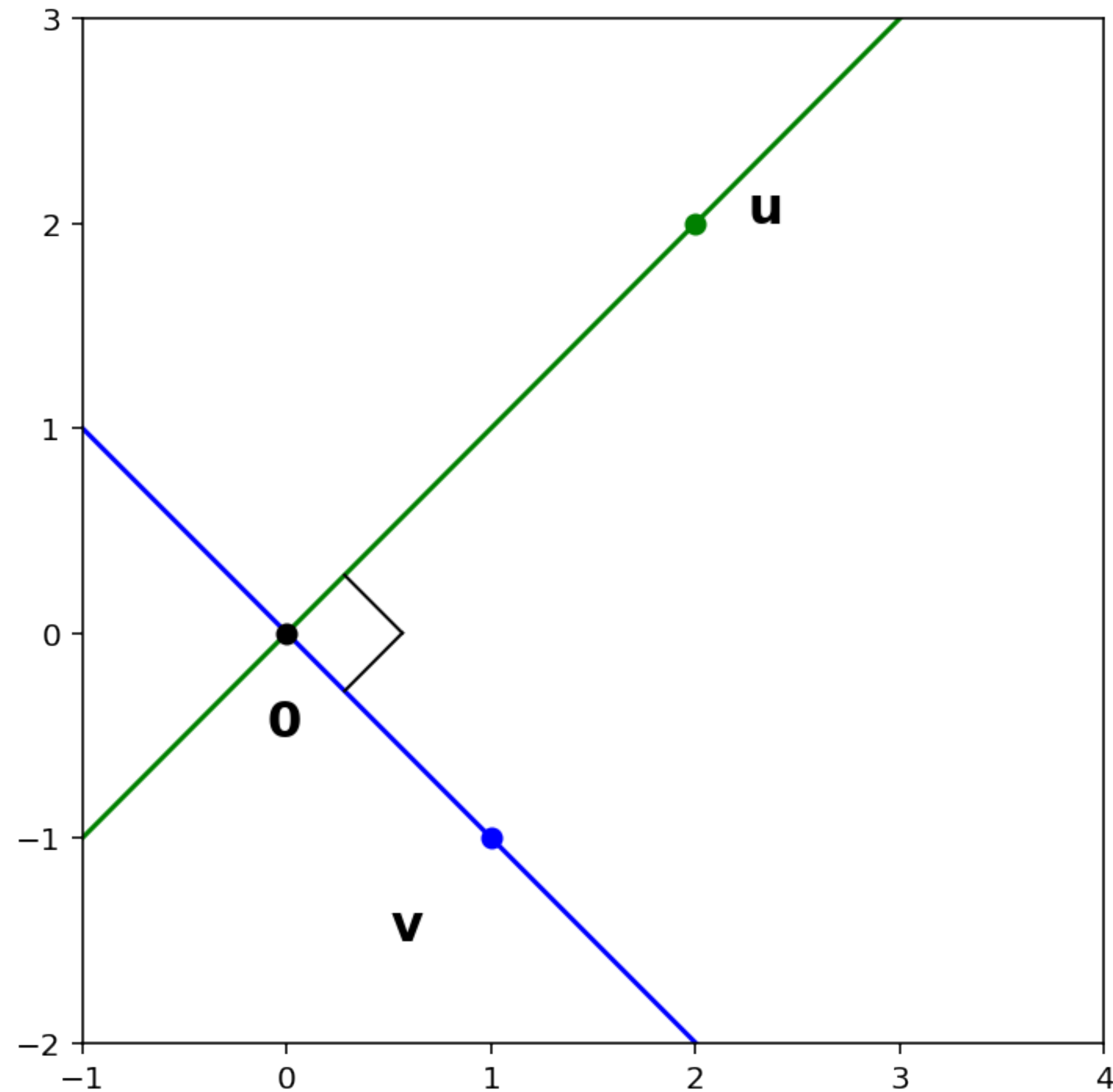
Definition. Vectors \mathbf{u} and \mathbf{v} are **orthogonal** if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

This definition gives an easy computational way to determine orthogonality.

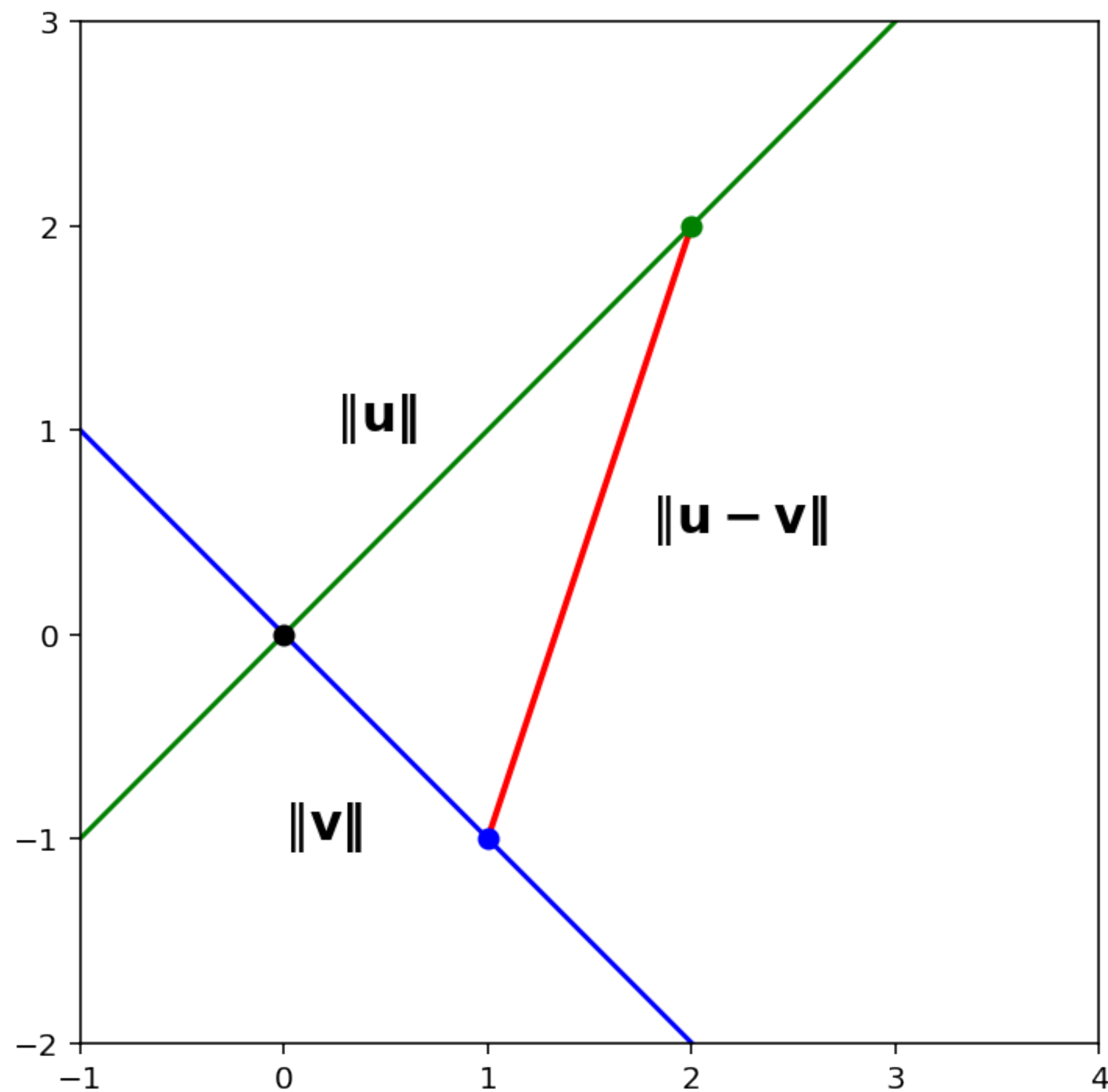
Example.

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

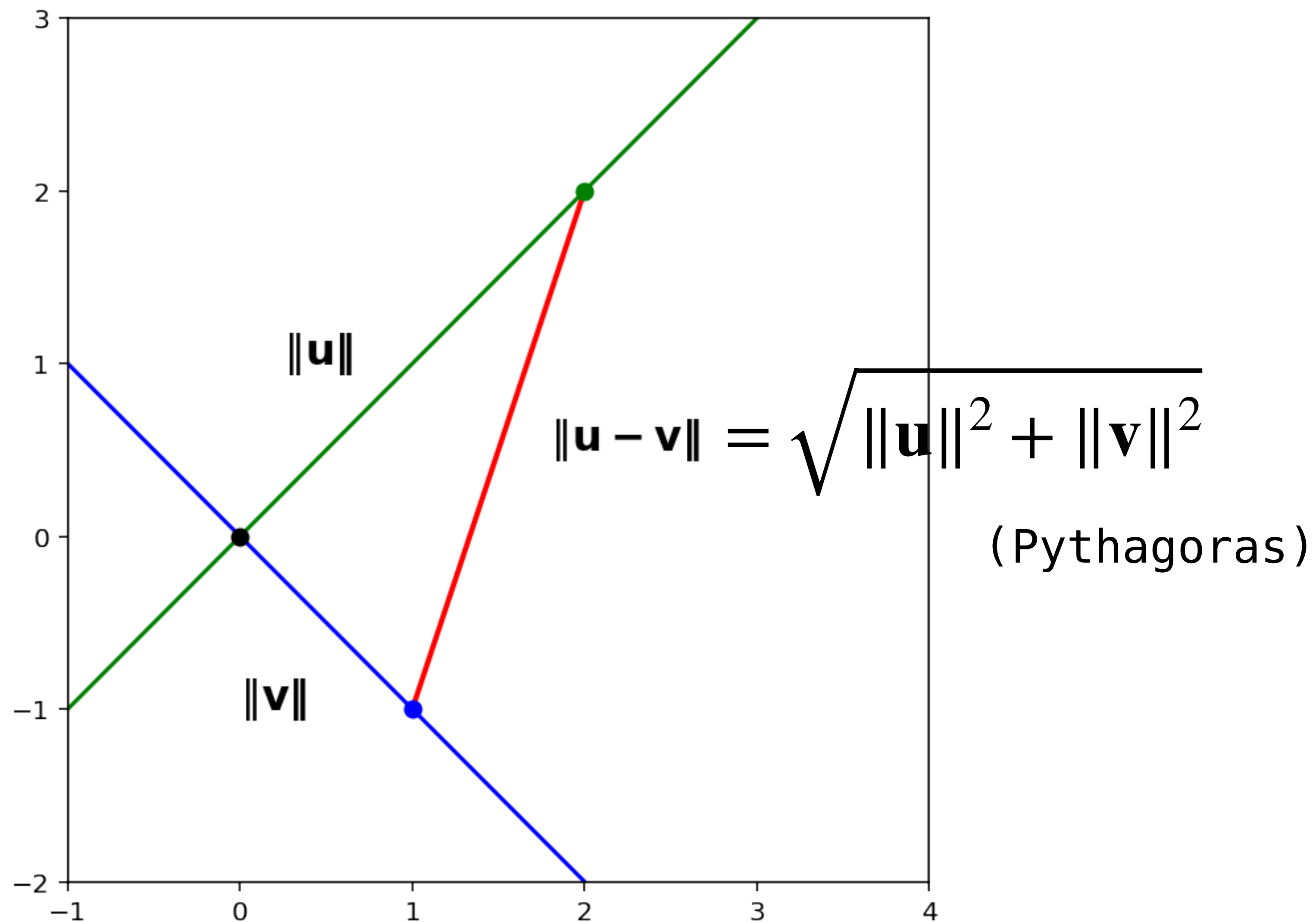
Derivation by Picture



Derivation by Picture



Derivation by Picture



Derivation by Algebra

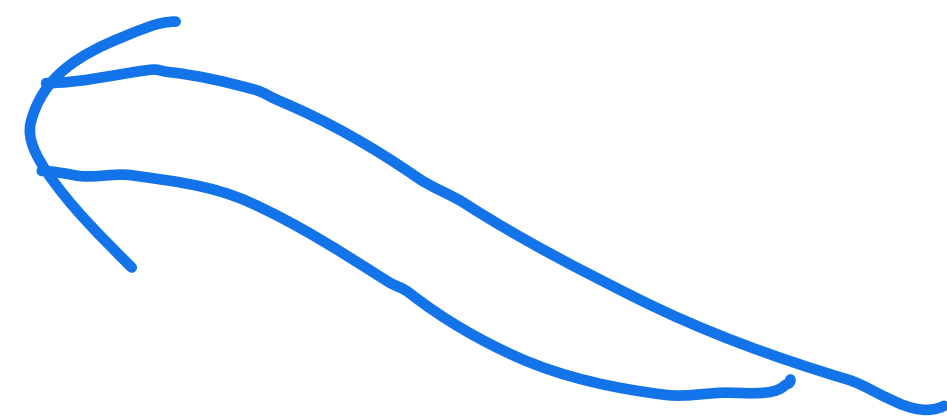
\mathbf{u} and \mathbf{v} are orthogonal exactly when

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

Let's simplify this a bit:

$$-2\langle \vec{u}, \vec{v} \rangle = 0$$

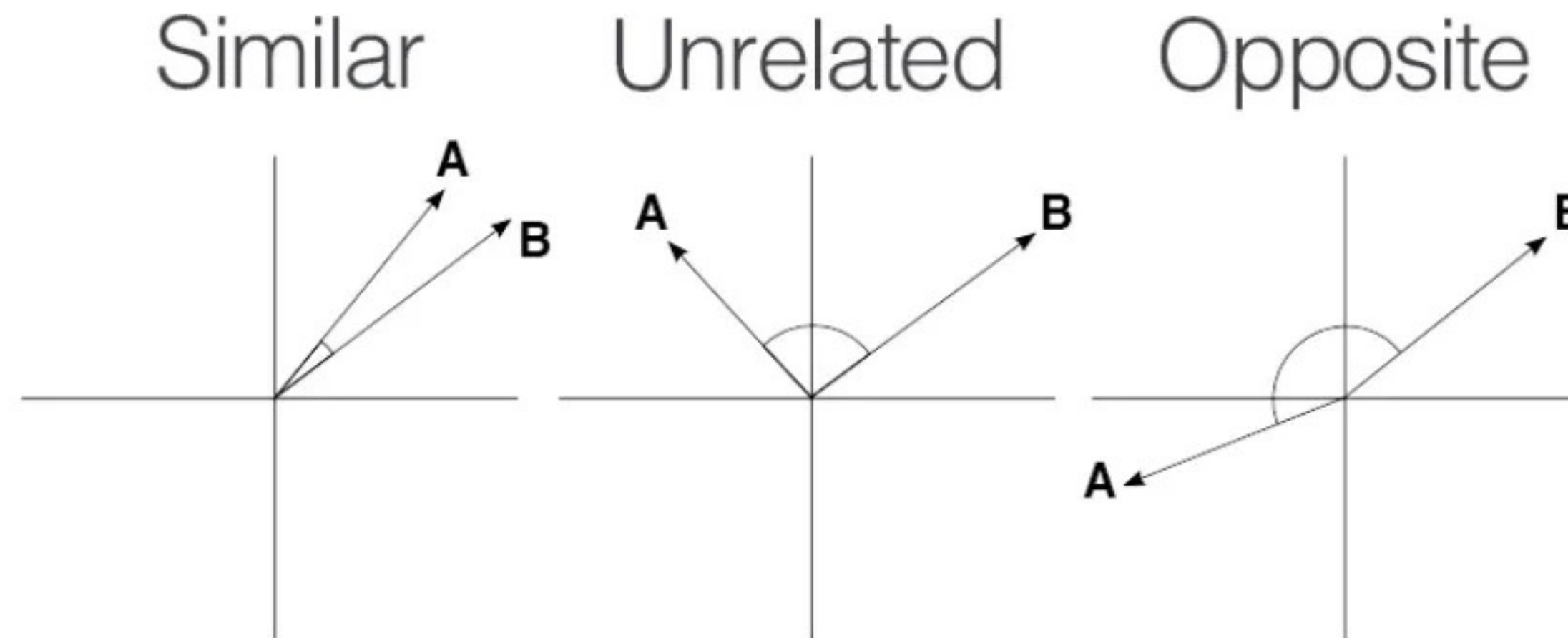
$$\downarrow$$
$$\langle \vec{u}, \vec{v} \rangle = 0$$



$$\begin{aligned} &= \langle \vec{u} - \vec{v}, \vec{u} - \vec{v} \rangle \\ &= \langle \vec{u}, \vec{u} - \vec{v} \rangle - \langle \vec{v}, \vec{u} - \vec{v} \rangle \\ &= \langle \vec{u}, \vec{u} \rangle - \langle \vec{u}, \vec{v} \rangle - \langle \vec{v}, \vec{u} \rangle + \langle \vec{v}, \vec{v} \rangle \\ &= \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\langle \vec{u}, \vec{v} \rangle \end{aligned}$$

Application: Cosine Similarity

High Level



Data points are very big vectors.

Similar vectors "point in nearly the same direction."

Example: Netflix Users

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

A Netflix user might be represented as a vectors whose i th 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.

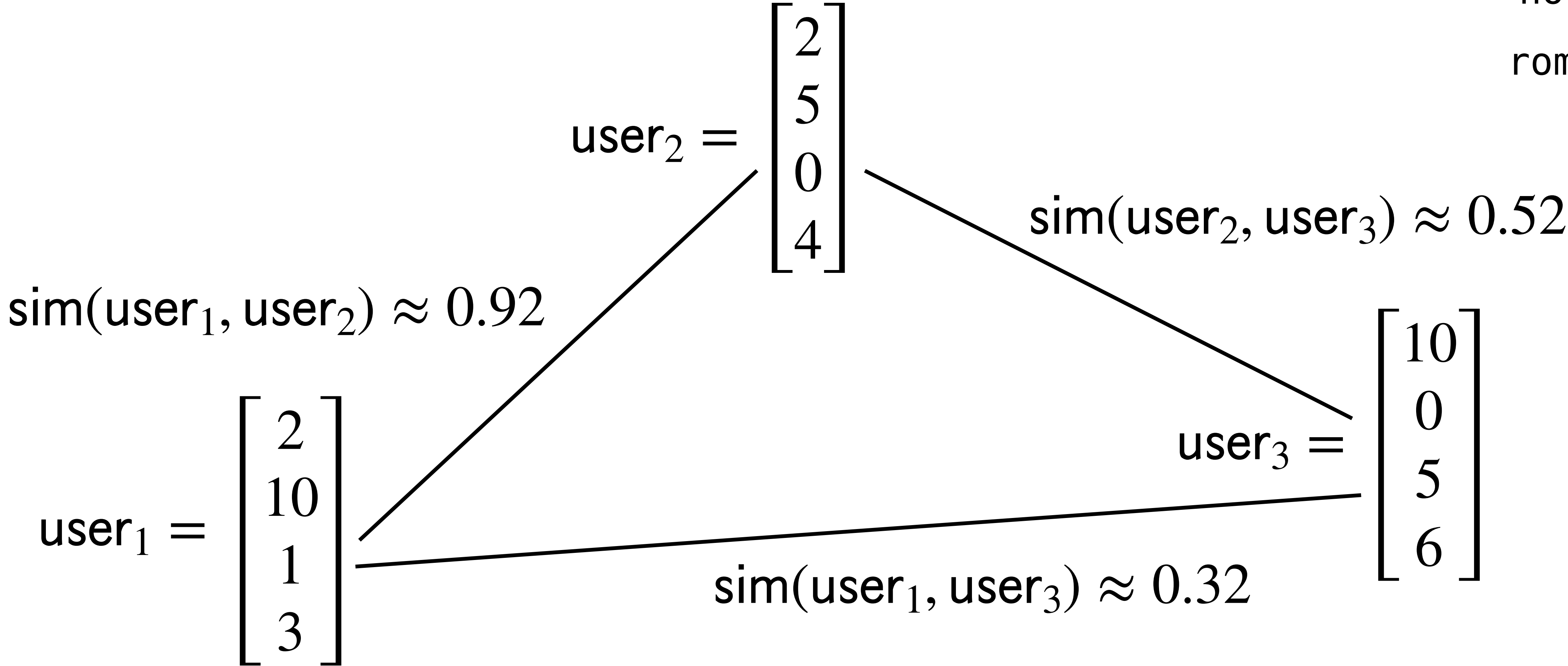
Example: Netflix Users

comedy

drama

horror

romance



Other Examples

- *Document similarity*
 - Documents \mapsto word count vectors
 - Similar documents should use similar words
- *Word2Vec*
 - Words \mapsto 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 $\{u_1, u_2, \dots, u_p\}$ 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

$$\binom{n}{k} \quad \binom{n}{2} = \frac{n!}{2!(n-2)!}$$

$$u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

$$u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$$

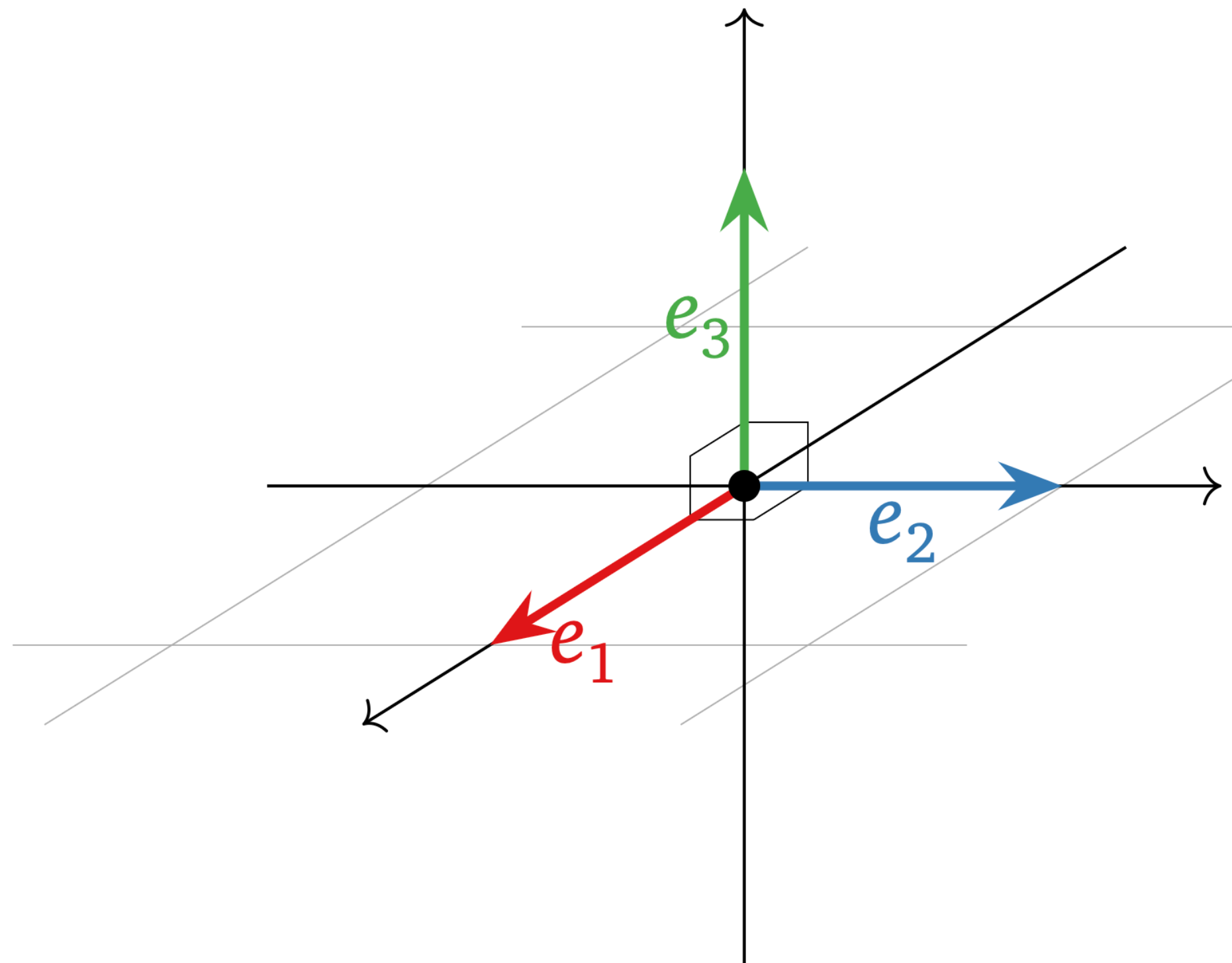
$$u_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

Verify:

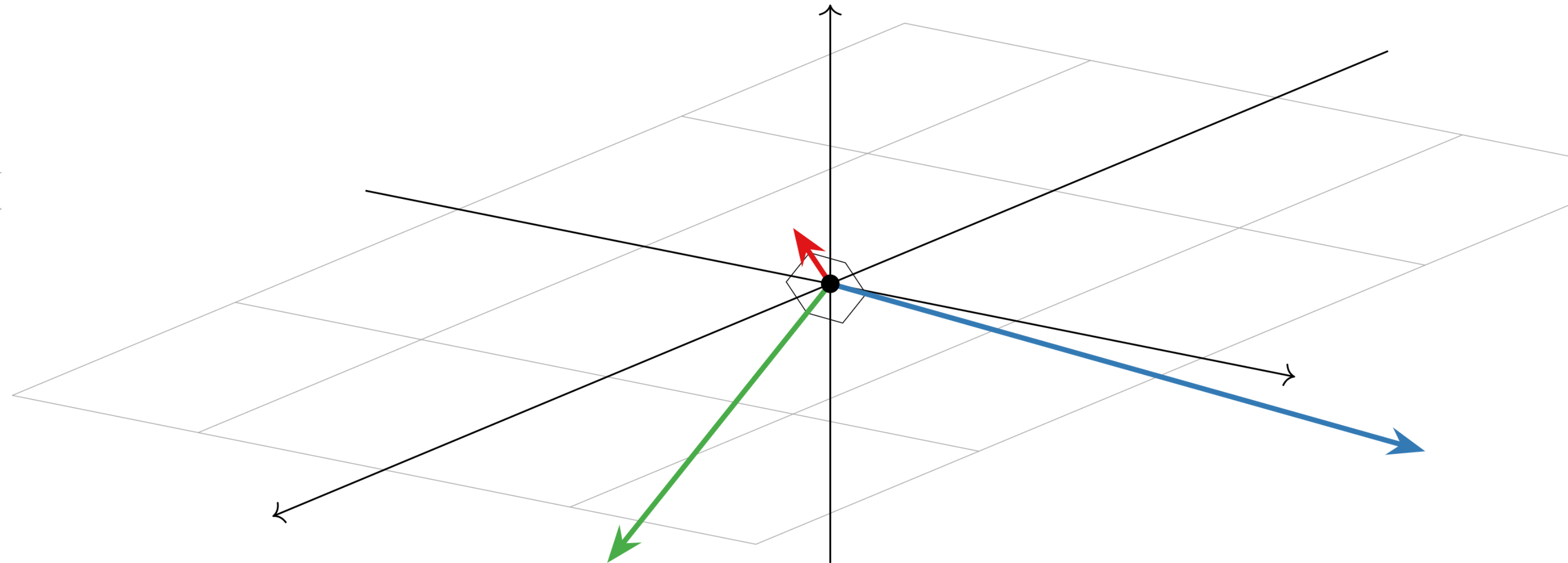
$$\vec{u}_1 \cdot \vec{u}_2 = -3 + 2 + 1 = 0$$
$$\vec{u}_1 \cdot \vec{u}_3 = -3/2 - 2 + 7/2 = -3/2 - 4/2 + 7/2 = 0$$
$$\vec{u}_2 \cdot \vec{u}_3 = 1/2 - 4 + 7/2 = 1/2 - 8/2 + 7/2 = 0$$

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
and reflections

Orthogonal Sets and Independence

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

Verify: $\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \dots + \alpha_k \vec{u}_k = \vec{0}$ (dot prod. with \vec{u}_1)
 $\alpha_1 (\vec{u}_1 \cdot \vec{u}_1) + \alpha_2 (\vec{u}_2 \cdot \vec{u}_1) + \dots + \alpha_k (\vec{u}_k \cdot \vec{u}_1) = \vec{0} \cdot \vec{u}_1 = 0$
 $\alpha_1 \|\vec{u}_1\|^2 = 0 \Rightarrow \alpha_1 = 0$

Analogous argument with $\vec{u}_j \Rightarrow \alpha_j = 0$

Get only the trivial sol'n

The Takeaway

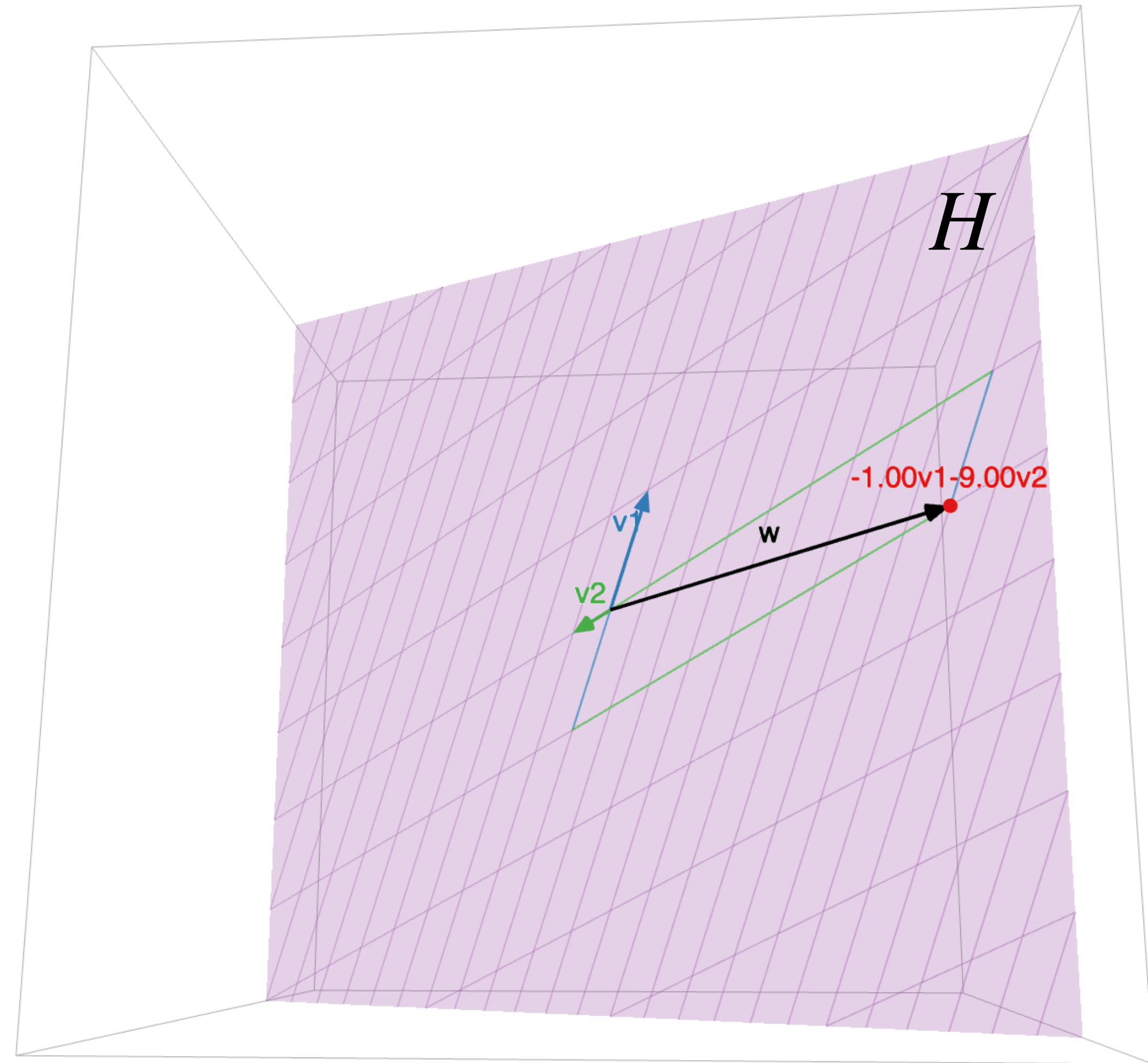
If $\{u_1, u_2, \dots, u_k\}$ is an orthogonal set,
then it is a **basis** for $\text{span}\{u_1, u_2, \dots, u_k\}$

Orthogonal Basis

Definition. An orthogonal basis for a subspace W of \mathbb{R}^n is a basis for W which is also an orthogonal set.

Orthogonal Basis

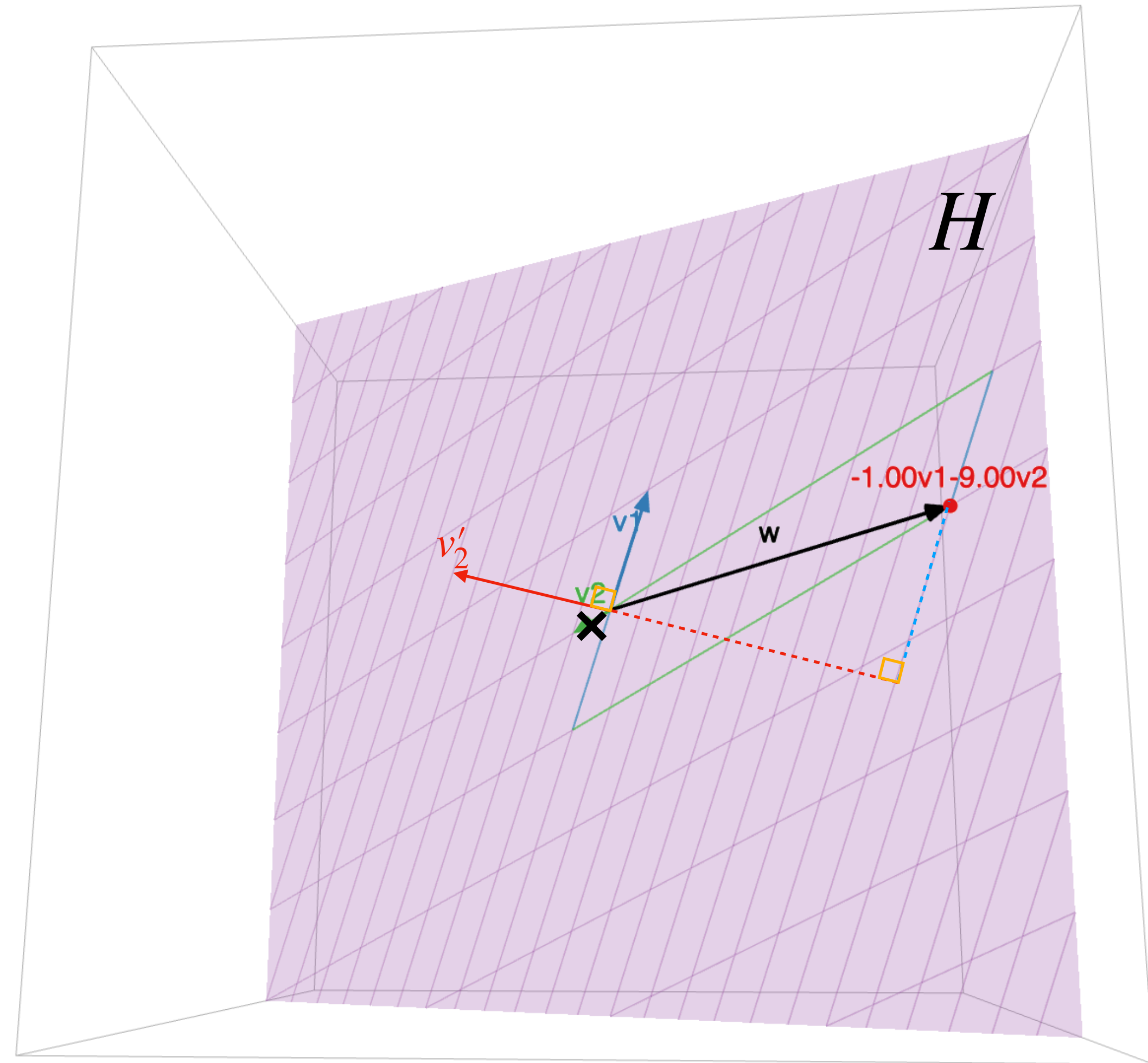
Definition. An orthogonal basis for a subspace W of R^n is a basis for W which is also an orthogonal set.



v_1 and v_2 form a basis of H

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

What's nice about an
orthogonal basis?

Recall: How To: Bases

Recall: How To: Bases

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

(unique)

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

Recall: How To: Bases

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

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

Solution. Solve the vector equation

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

by Gaussian elimination, matrix inversion, etc.

$$\begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_p \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = \begin{bmatrix} \vec{w} \end{bmatrix}$$

Recall: How To: Bases

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

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

Solution. Solve the vector equation

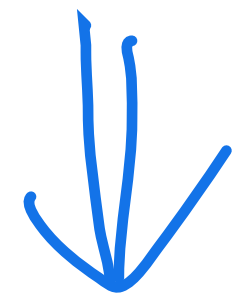
$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \dots x_p\mathbf{u}_p = \mathbf{w}$$

by Gaussian elimination, matrix inversion, etc.

This takes work

Orthogonal Bases and Linear Combinations

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



$$c_j = \frac{\mathbf{y}^T \mathbf{u}_j}{\mathbf{u}_j^T \mathbf{u}_j}$$

Verify: $\vec{y} \cdot \vec{u}_j = c_1(\vec{u}_1 \cdot \vec{u}_j) + c_2(\vec{u}_2 \cdot \vec{u}_j) + \dots + c_p(\vec{u}_p \cdot \vec{u}_j)$
 $= c_j \|\vec{u}_j\|^2 \Rightarrow c_j = \frac{\vec{y} \cdot \vec{u}_j}{\vec{u}_j \cdot \vec{u}_j}$

How To: Orthogonal Bases

How To: Orthogonal Bases

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

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

How To: Orthogonal Bases

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

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

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

How To: Orthogonal Bases

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

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

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

Much easier to compute.

Question

$$\vec{w} = \vec{x}_{11}$$

$$\begin{bmatrix} 3 & -1 & -1/2 & | & 6 \\ 1 & 2 & -2 & | & 1 \\ 1 & 1 & 7/2 & | & -8 \end{bmatrix}$$

~~→~~ slower

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

$$u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

$$u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$$

orthogonal set

$$u_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

$$C_1 = \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} = \frac{18 + 1 - 8}{9 + 1 + 1} = \frac{11}{11} = 1$$

$$C_2 = \frac{\vec{y} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} = \frac{-6 + 2 - 8}{1 + 4 + 1} = \frac{-12}{6} = -2$$

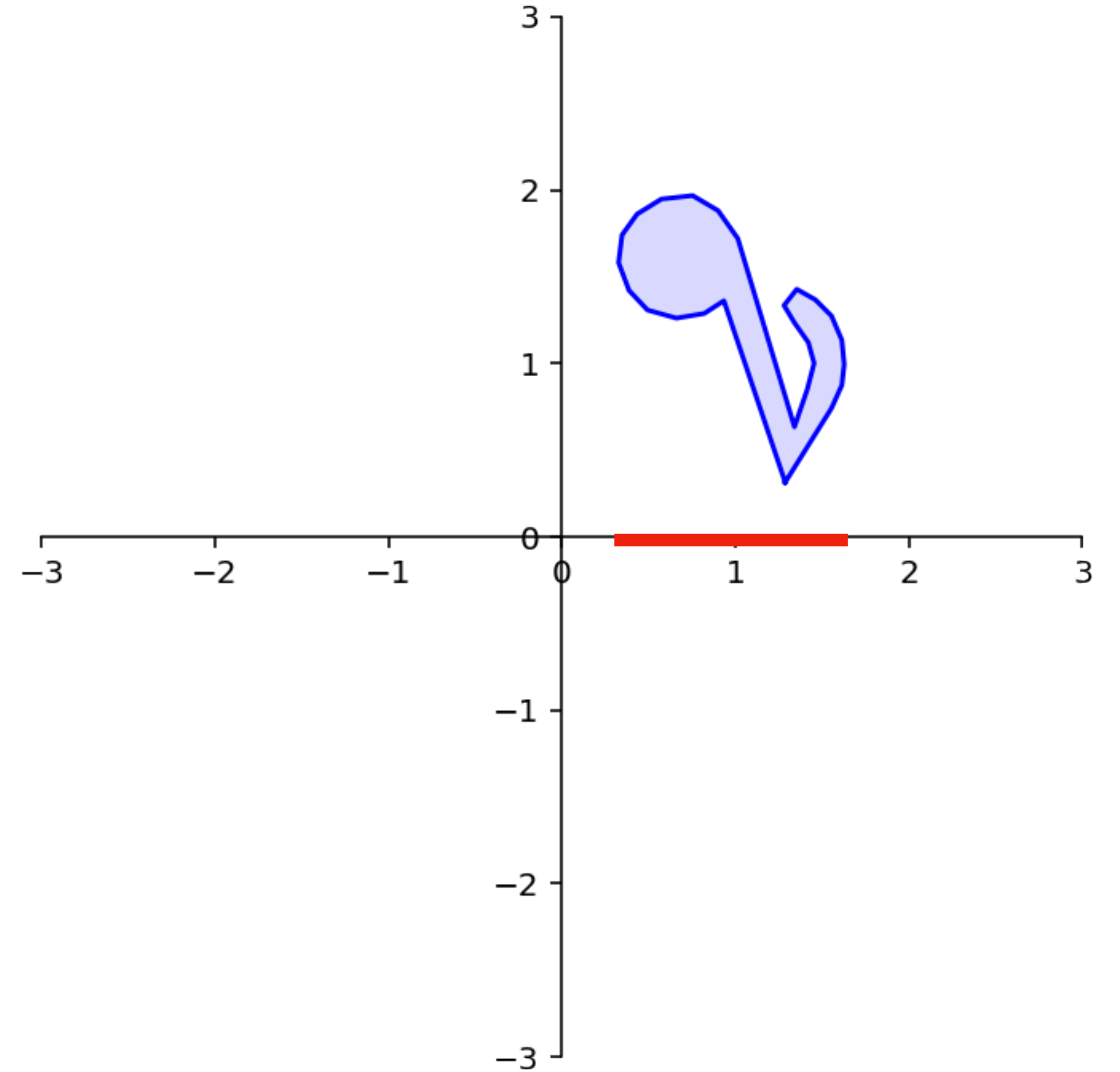
$$C_3 = \frac{\vec{y} \cdot \vec{u}_3}{\vec{u}_3 \cdot \vec{u}_3} = \frac{-6/2 - 2 - 56/2}{1/4 + 4 + 49/4} = \frac{(-6 - 4 - 56)/2}{(1 + 16 + 49)/4} = \frac{-66/2}{66/4} = \frac{-33}{33/2} = -2$$

Answer: $u_1 - 2u_2 - 2u_3$

Orthogonal Projection

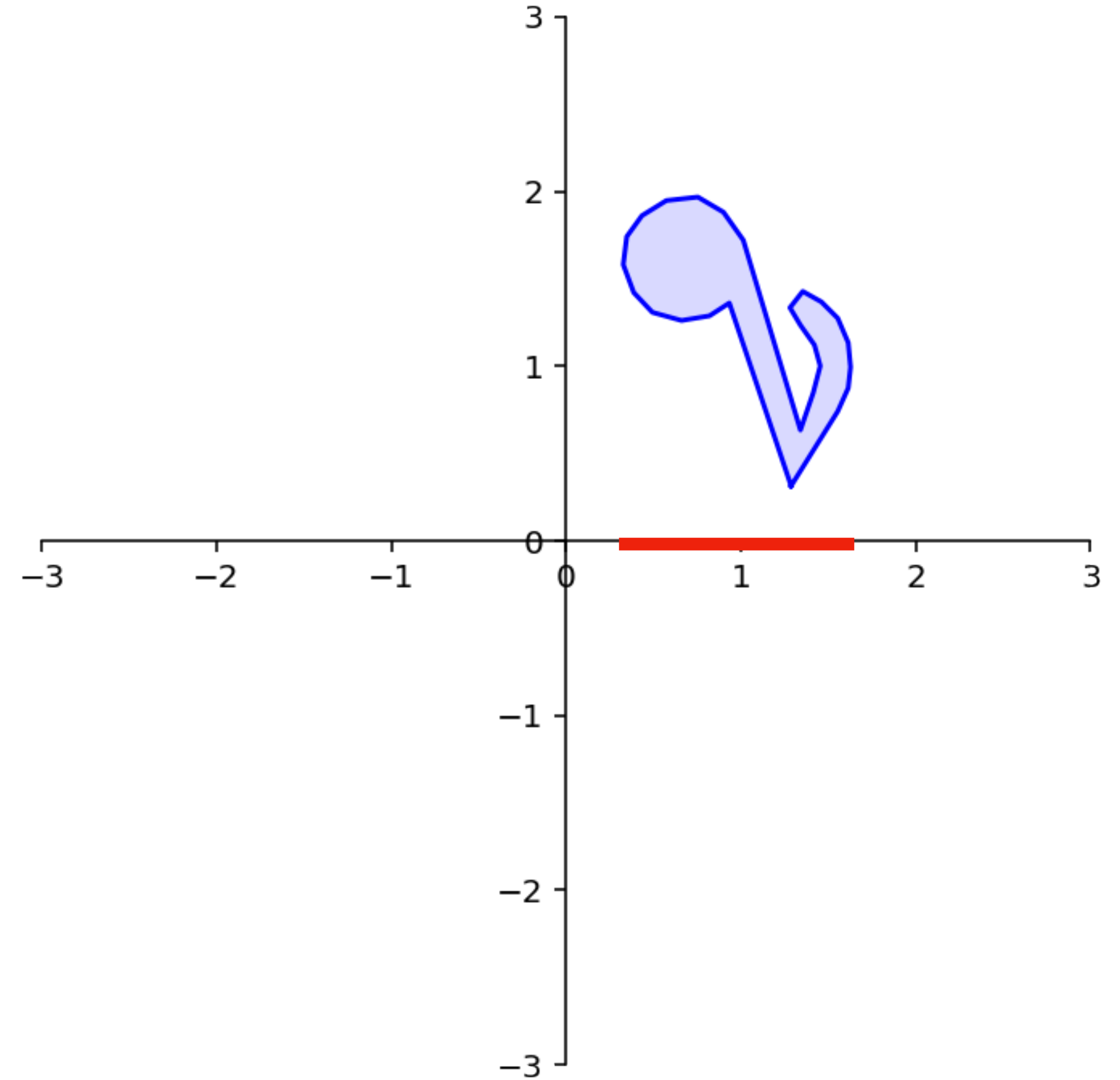
Why does that formula in
the last example work?

Recall: Projection onto the x -axis



Recall: Projection onto the x -axis

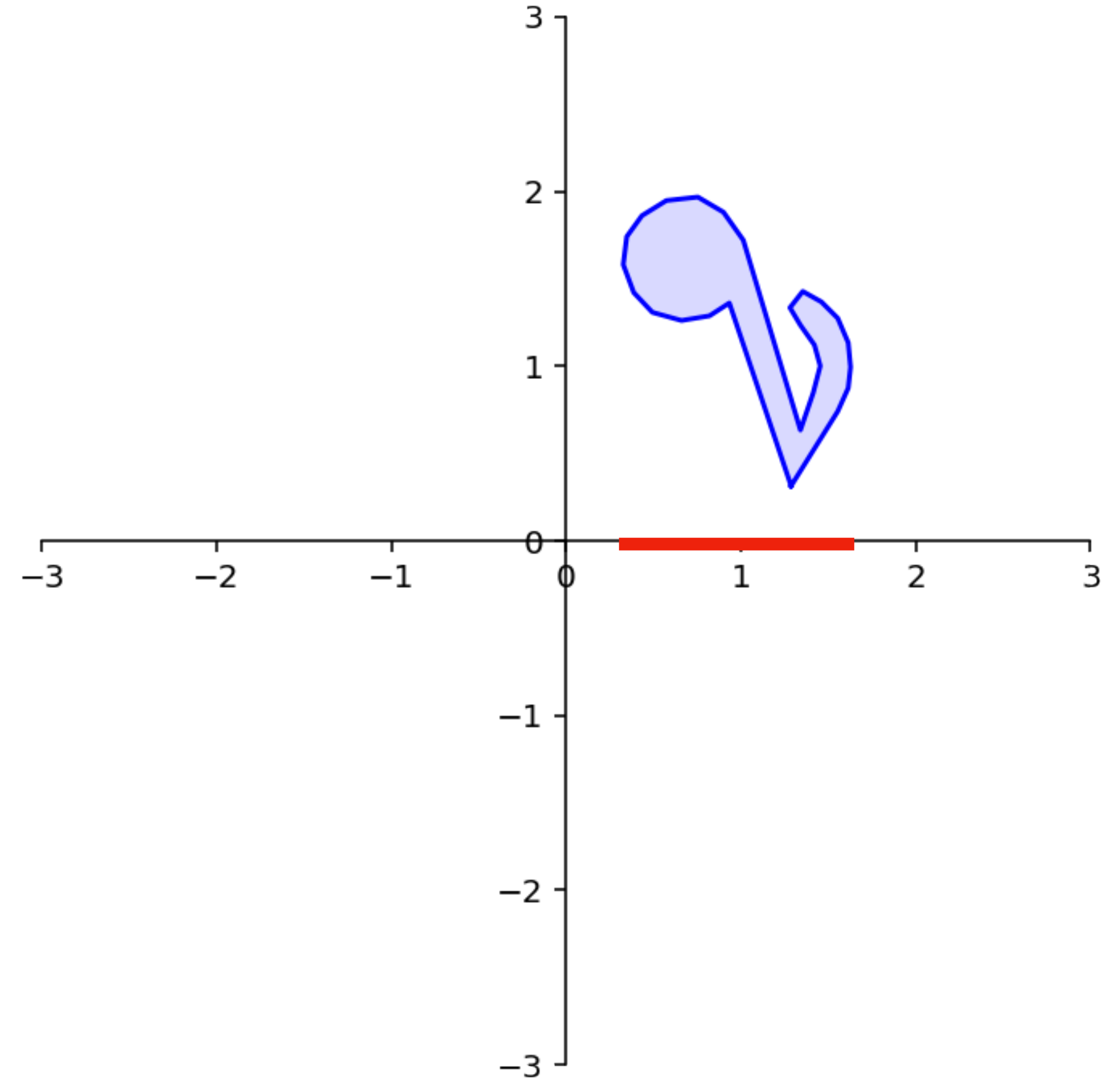
We've seen simple
projections in R^2



Recall: Projection onto the x -axis

We've seen simple
projections in R^2

We're going to
generalize this idea

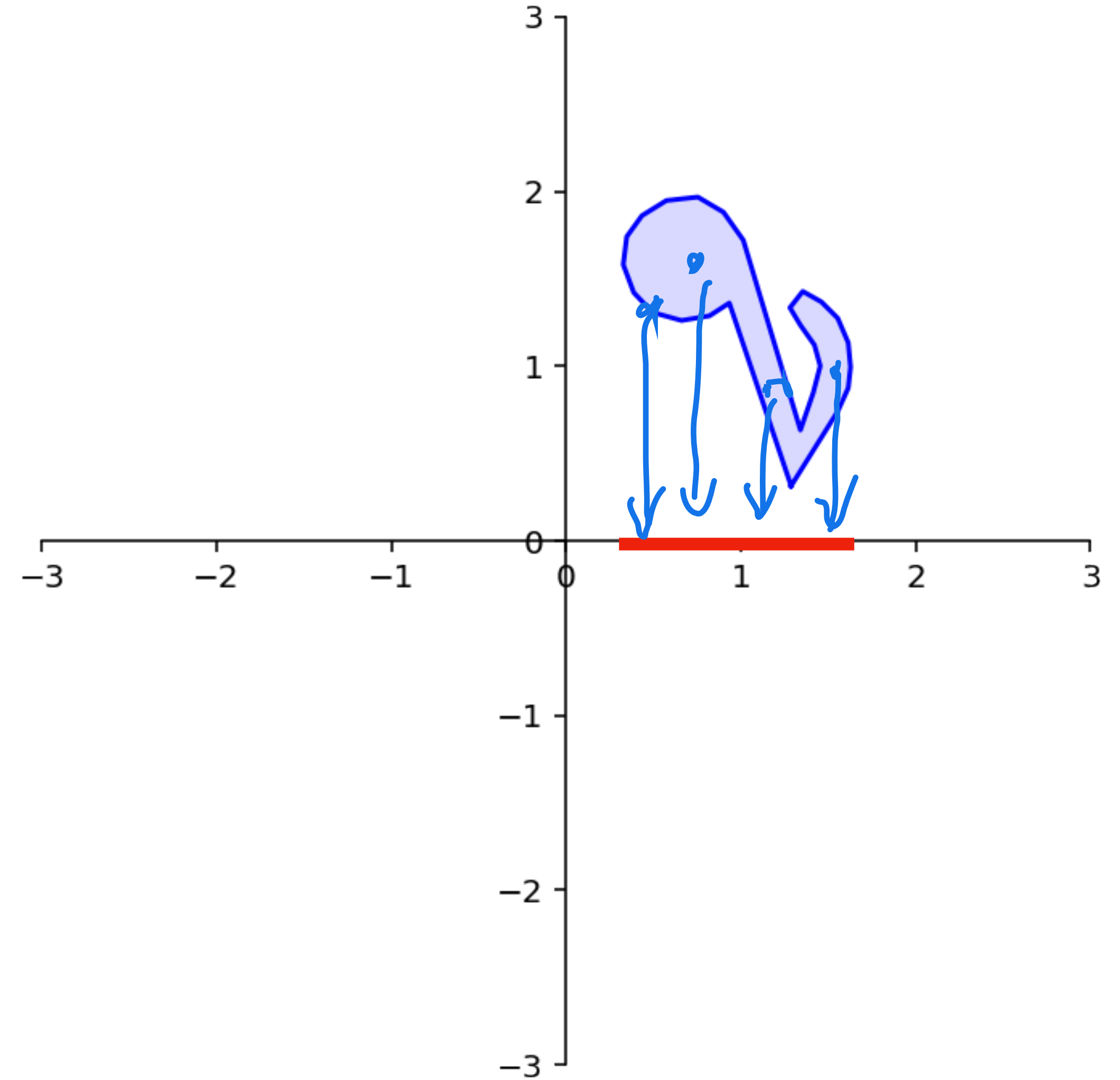


Recall: Projection onto the x -axis

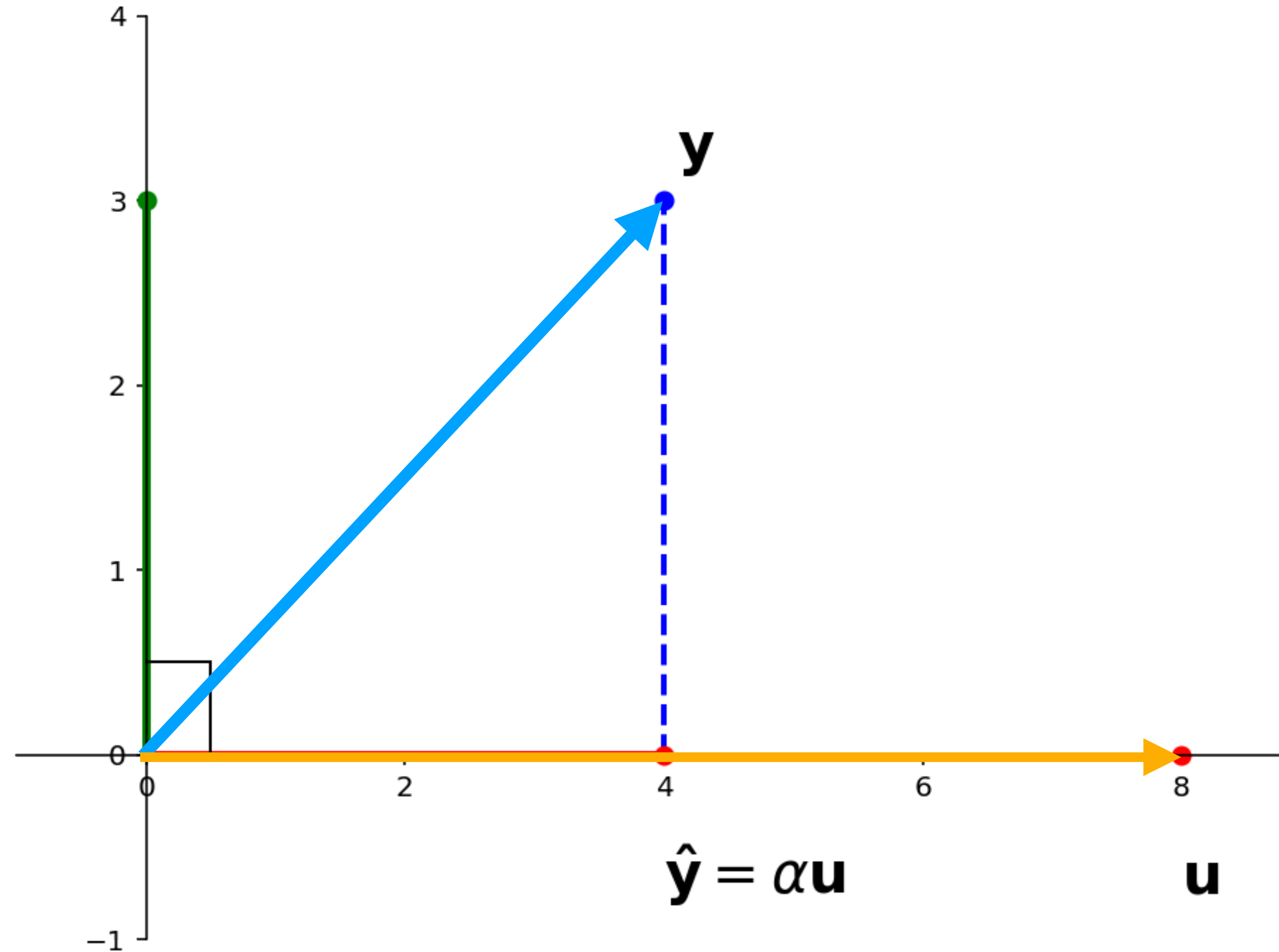
We've seen simple projections in R^2

We're going to generalize this idea

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

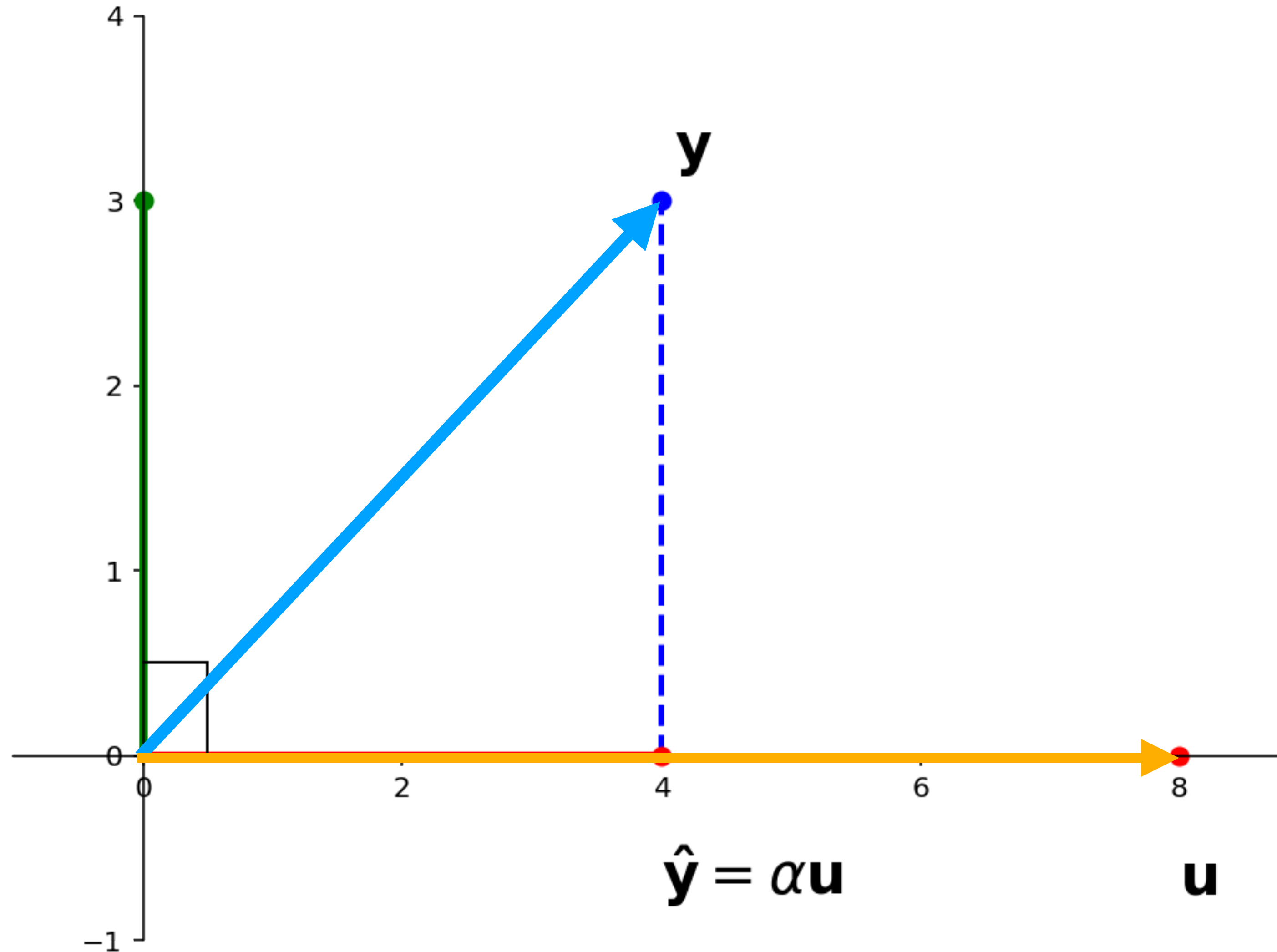


Orthogonal Projection



Orthogonal Projection

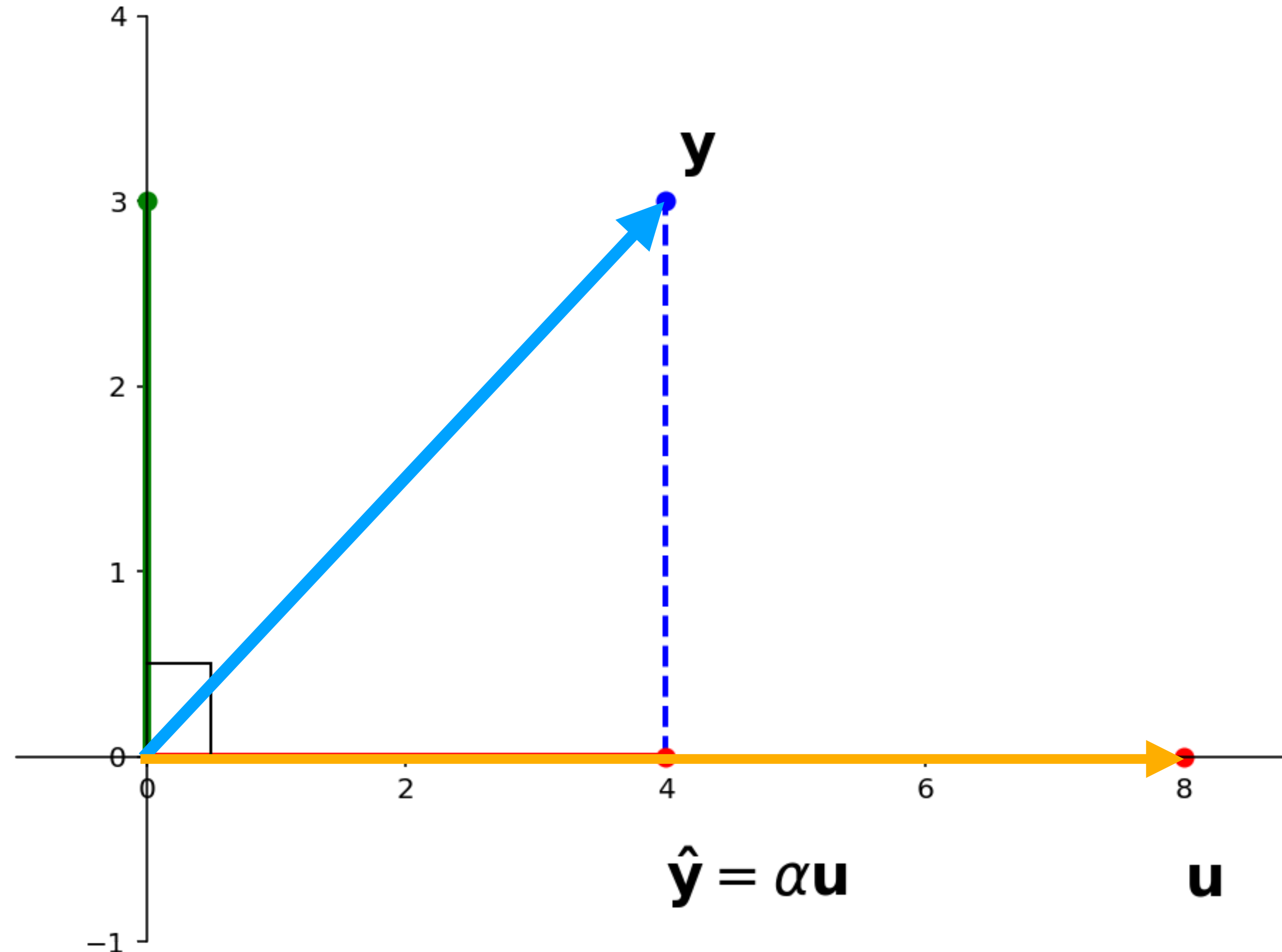
Question. Given vectors \mathbf{y} and \mathbf{u} in R^n , find vectors $\hat{\mathbf{y}}$ and \mathbf{z} such that



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., $\mathbf{z} \cdot \mathbf{u} = 0$)

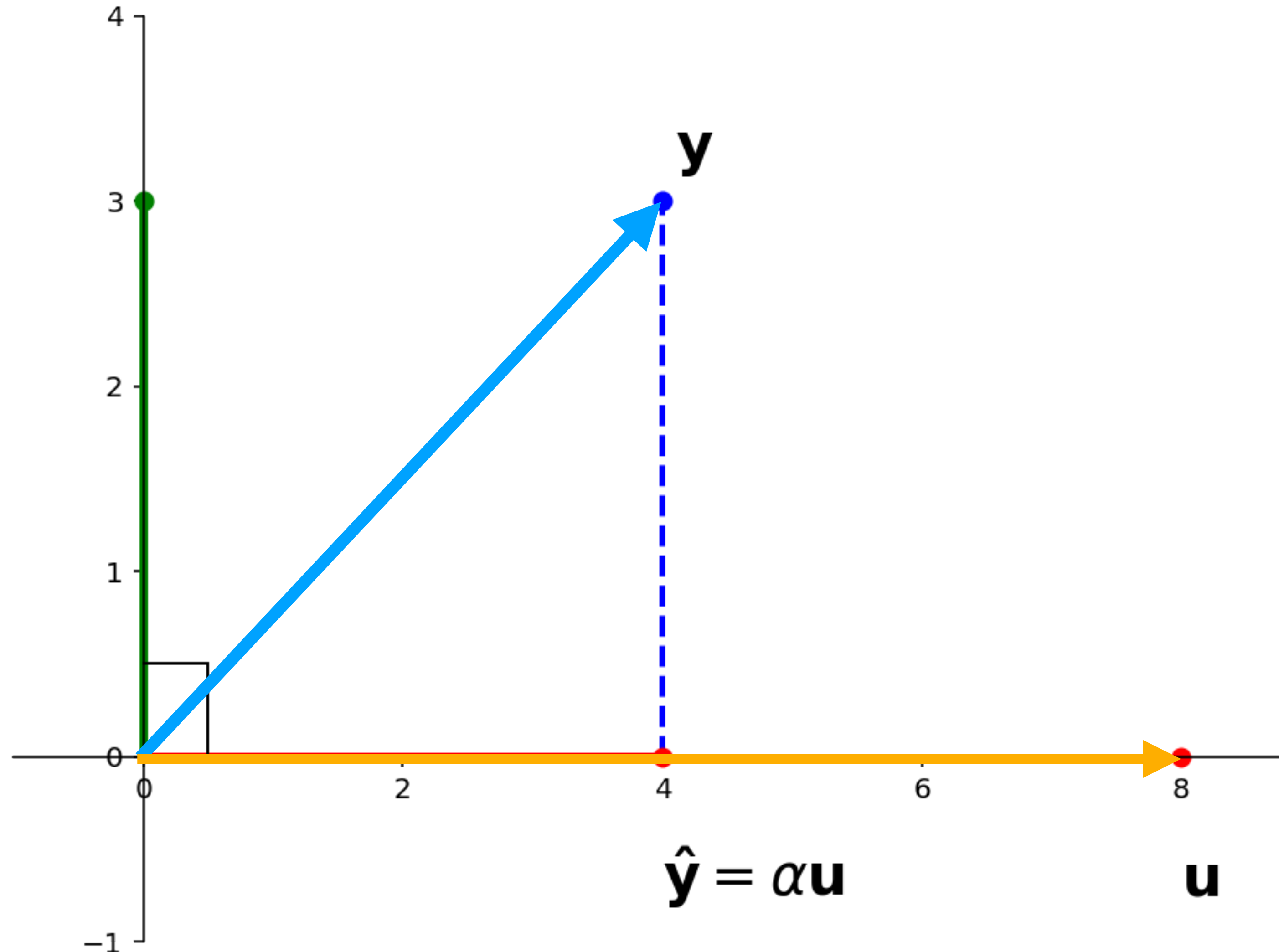


Orthogonal Projection

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., $\mathbf{z} \cdot \mathbf{u} = 0$)

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



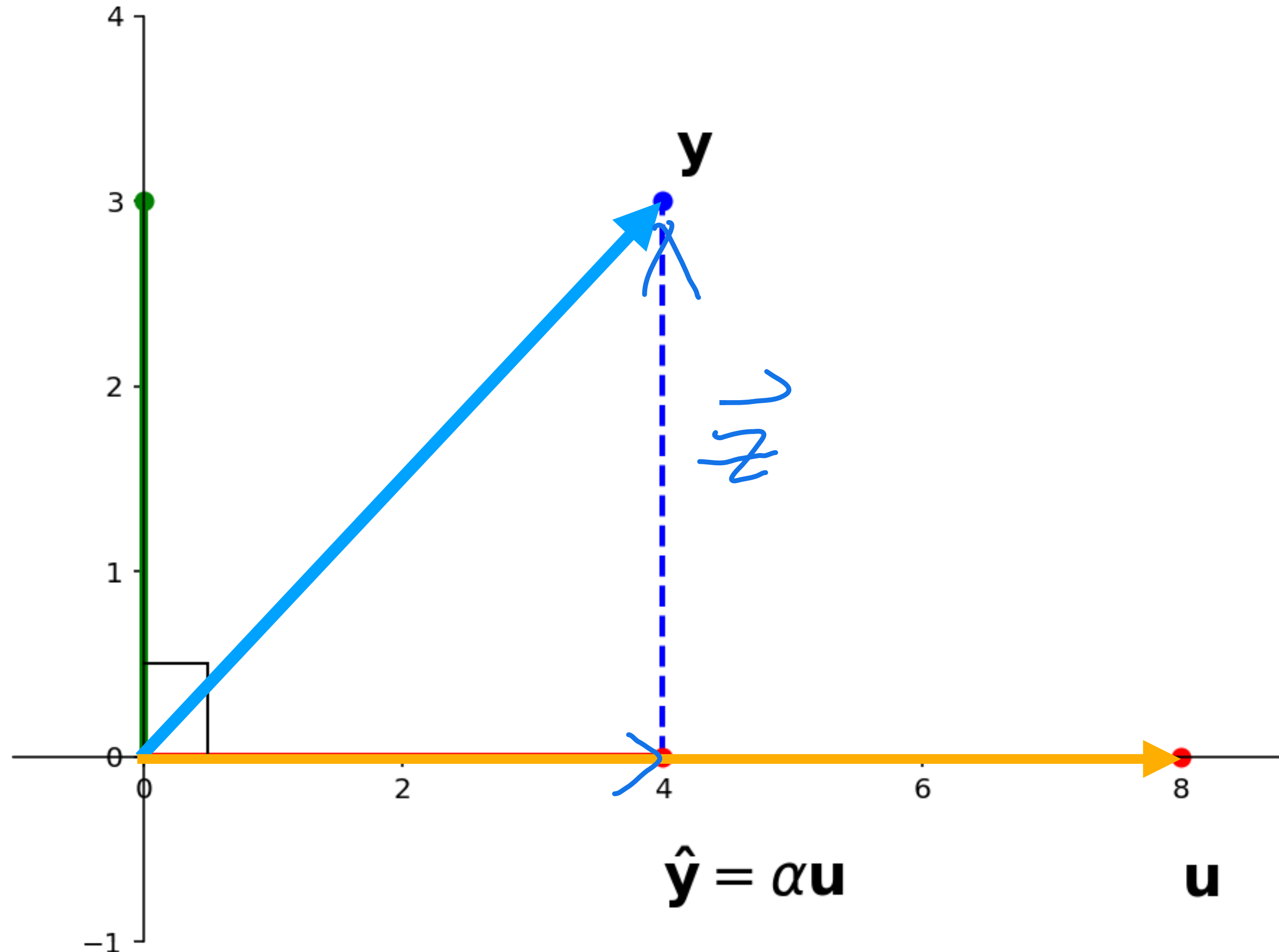
Orthogonal Projection

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

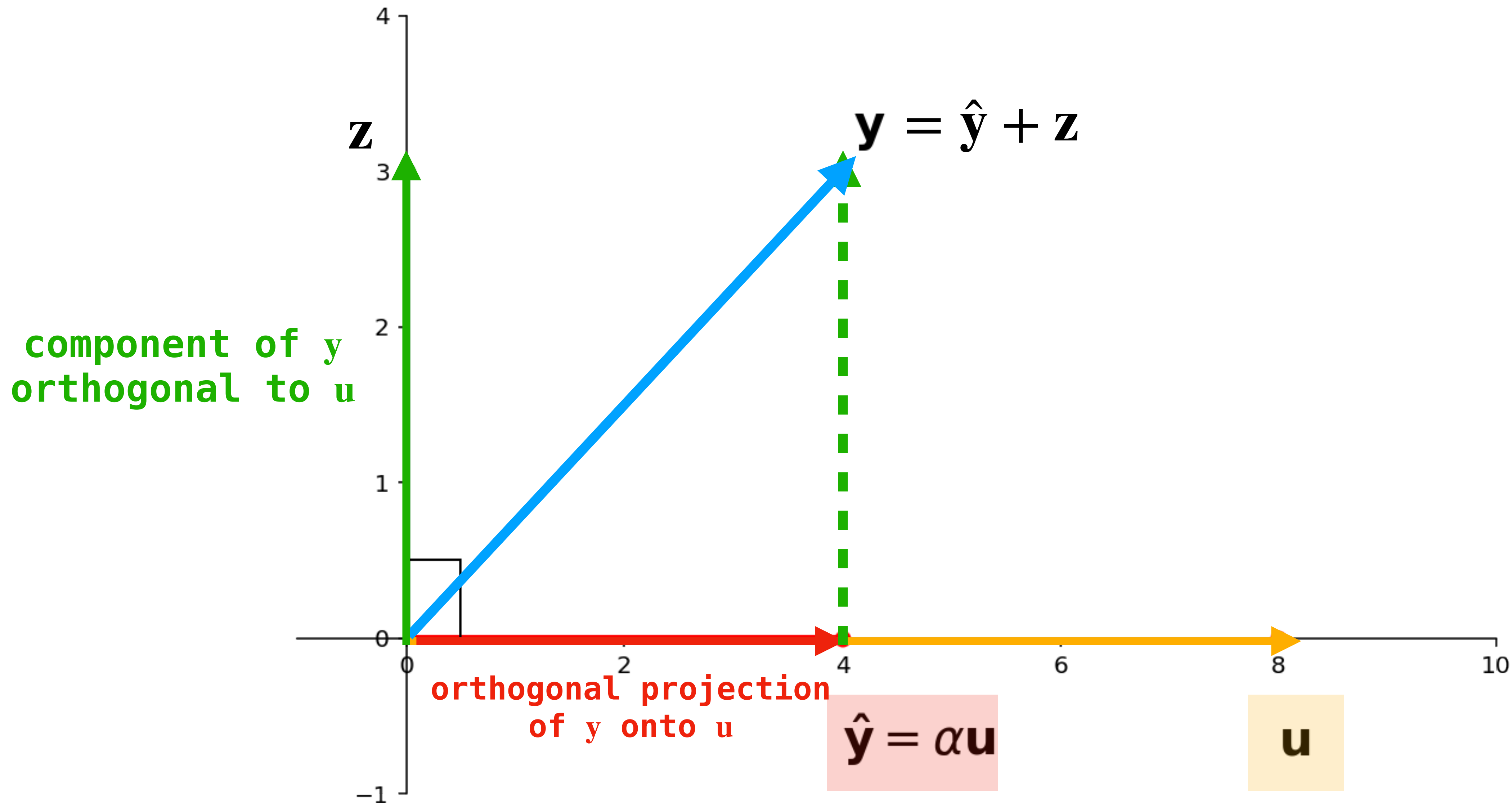
» \mathbf{z} is orthogonal to \mathbf{u}
(i.e., $\mathbf{z} \cdot \mathbf{u} = 0$)

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

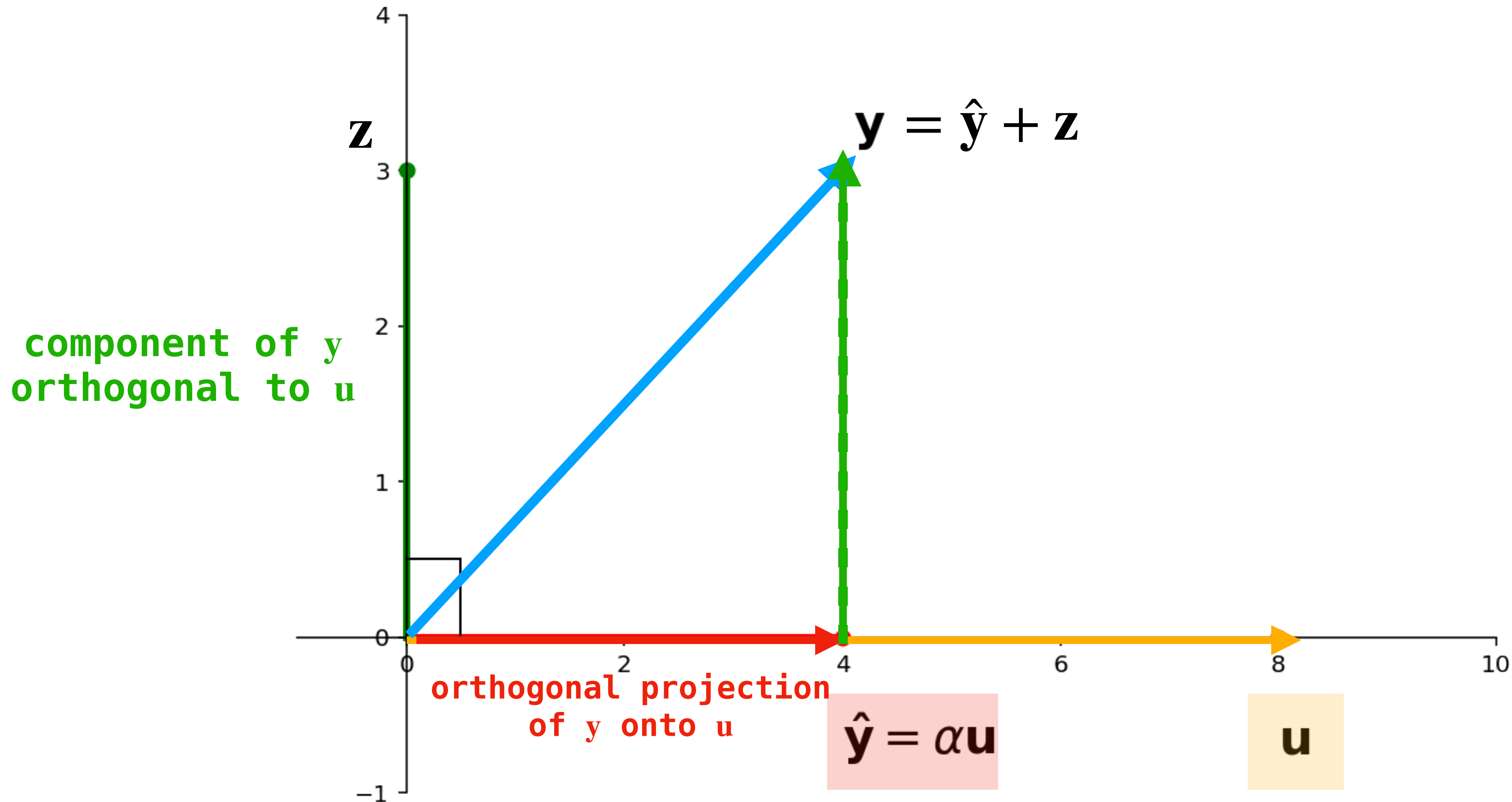
» $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$



Orthogonal Projection

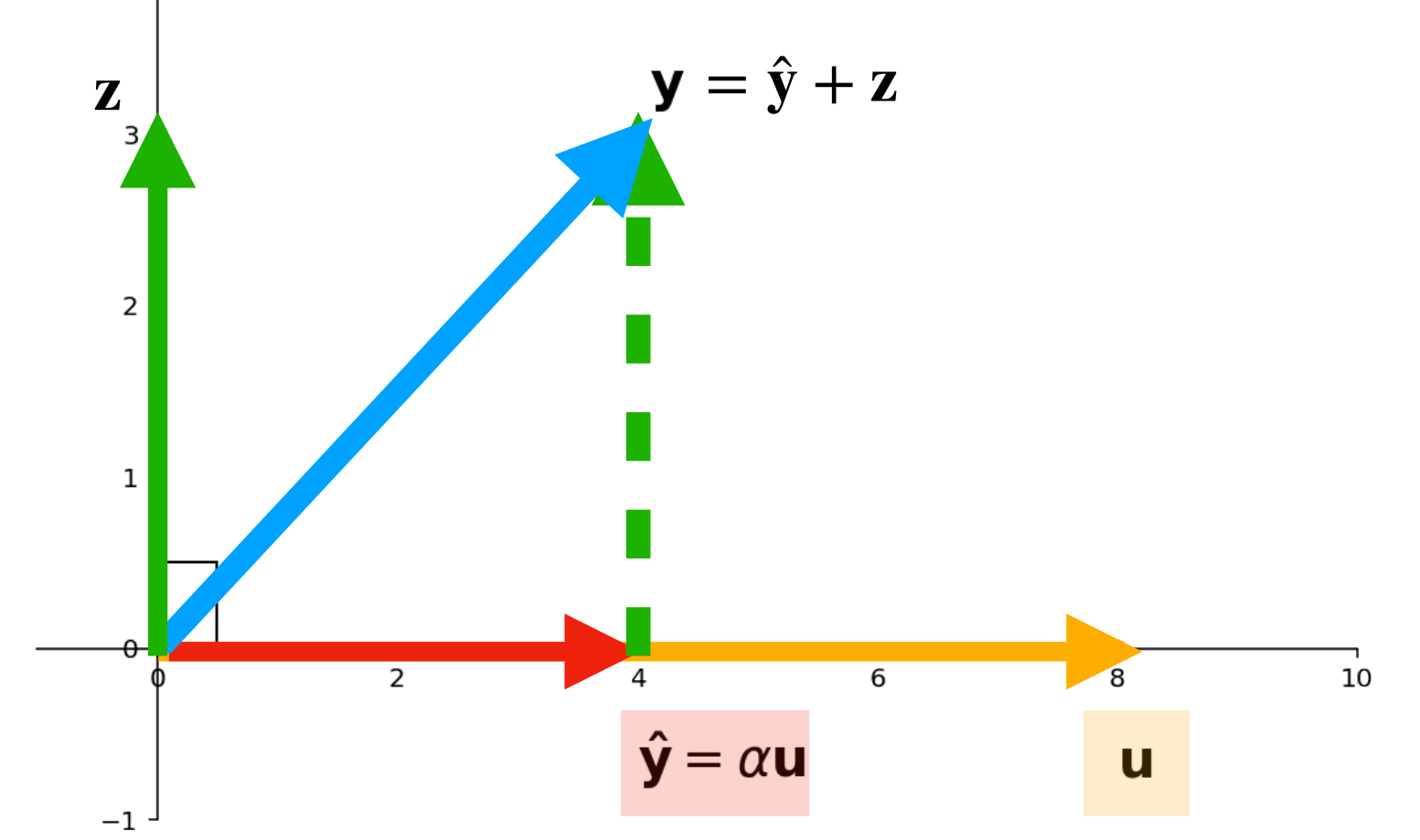


Orthogonal Projection

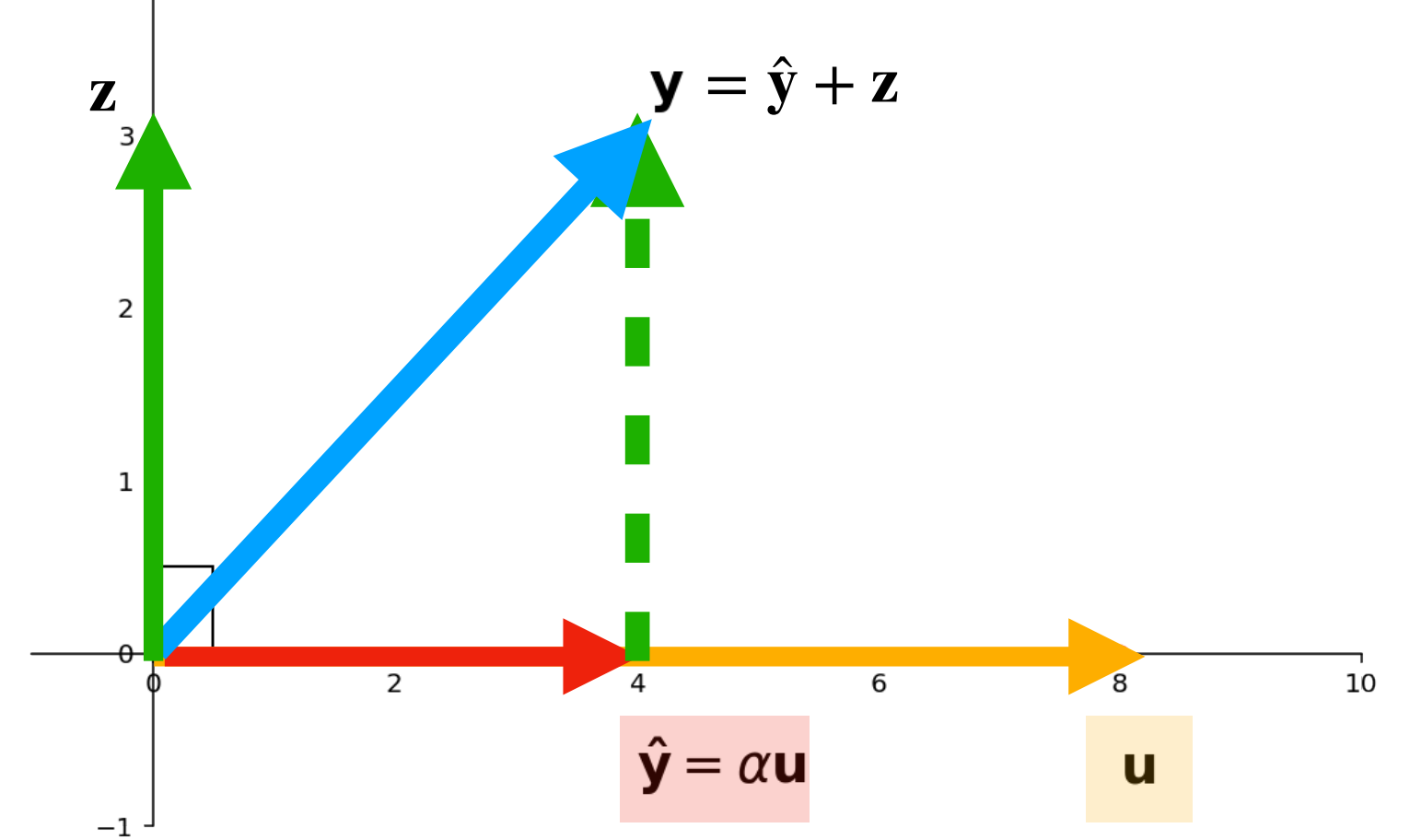


How do we find the orthogonal
projection and orthogonal component?

What we know

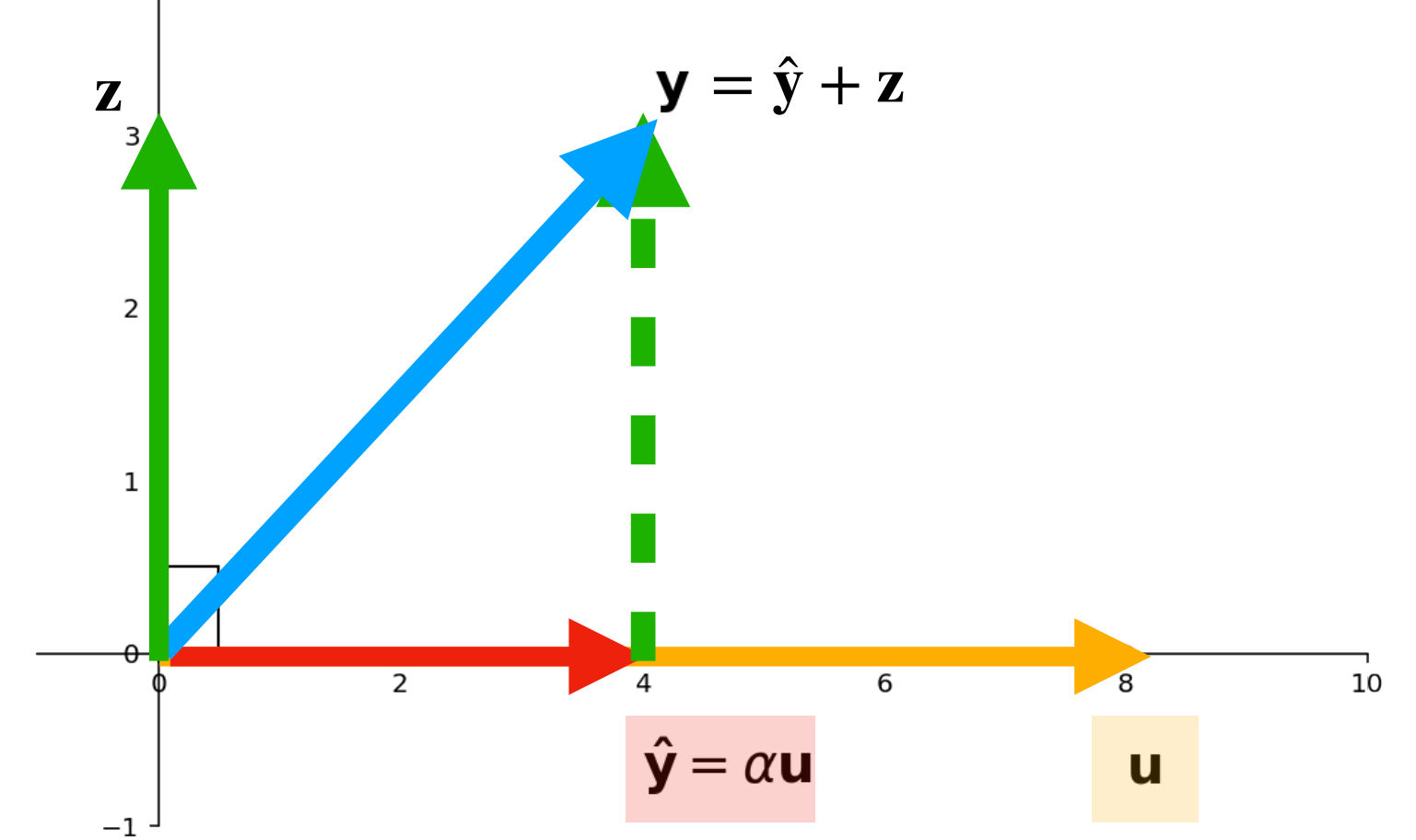


What we know



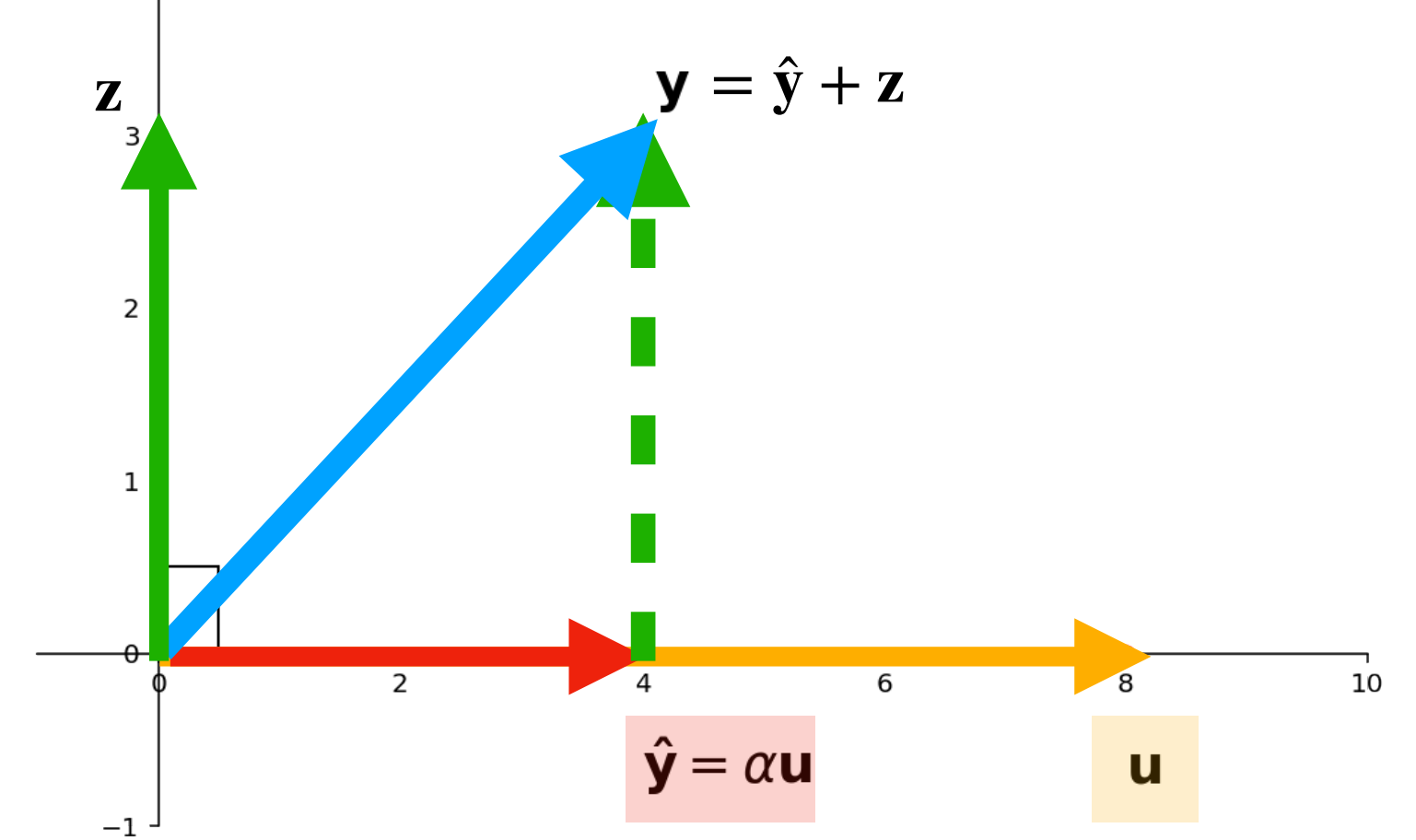
- $\hat{\mathbf{y}} = \alpha \mathbf{u}$ for some scalar α (since $\hat{\mathbf{y}} \in \text{span}\{\mathbf{u}\}$)

What we know



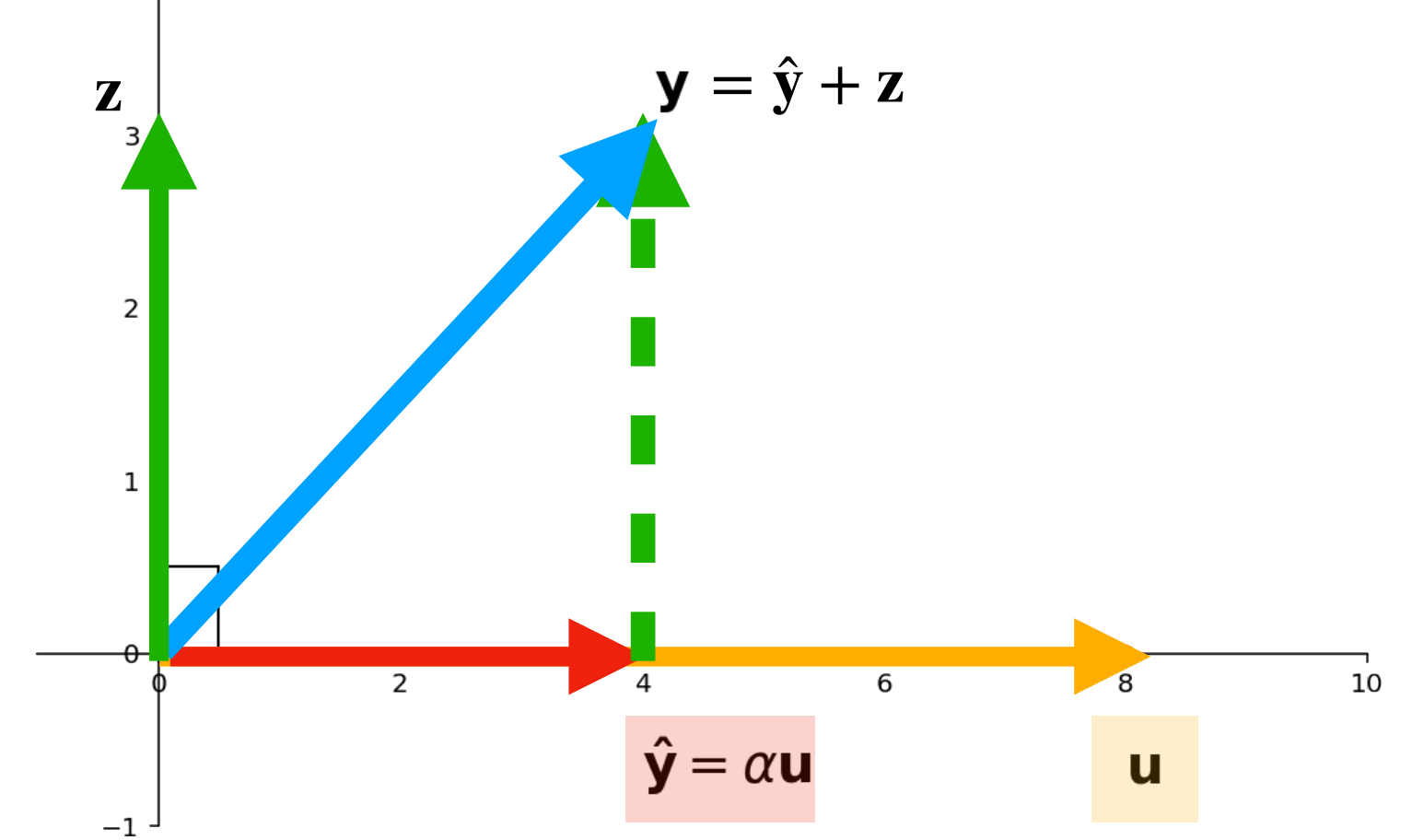
- $\hat{\mathbf{y}} = \alpha \mathbf{u}$ for some scalar α (since $\hat{\mathbf{y}} \in \text{span}\{\mathbf{u}\}$)
- $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \alpha \mathbf{u}$ (since $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$)

What we know



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

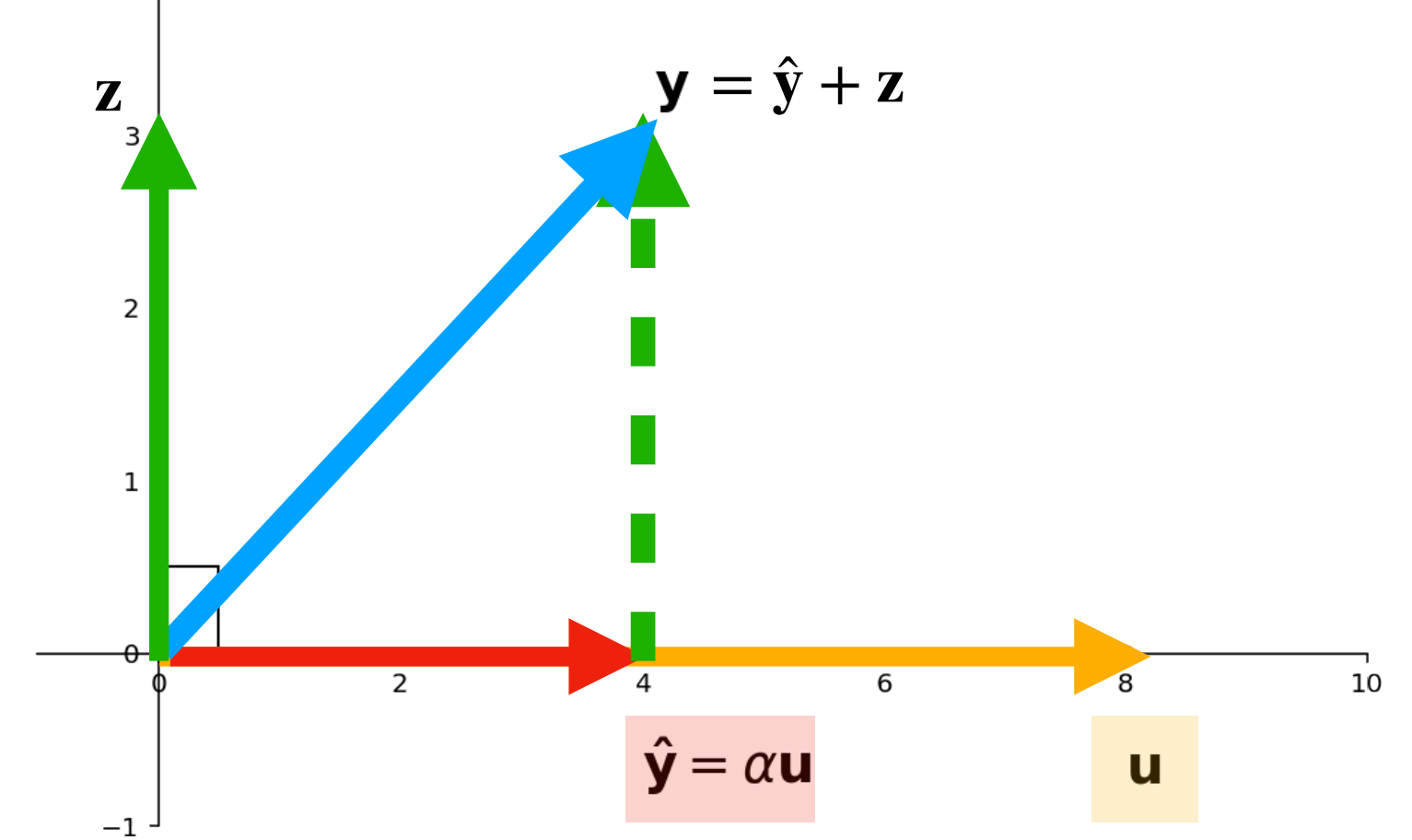


- $\hat{\mathbf{y}} = \alpha \mathbf{u}$ for some scalar α (since $\hat{\mathbf{y}} \in \text{span}\{\mathbf{u}\}$)
- $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \alpha \mathbf{u}$ (since $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$)
- $\langle \mathbf{z}, \mathbf{u} \rangle = 0$ (since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

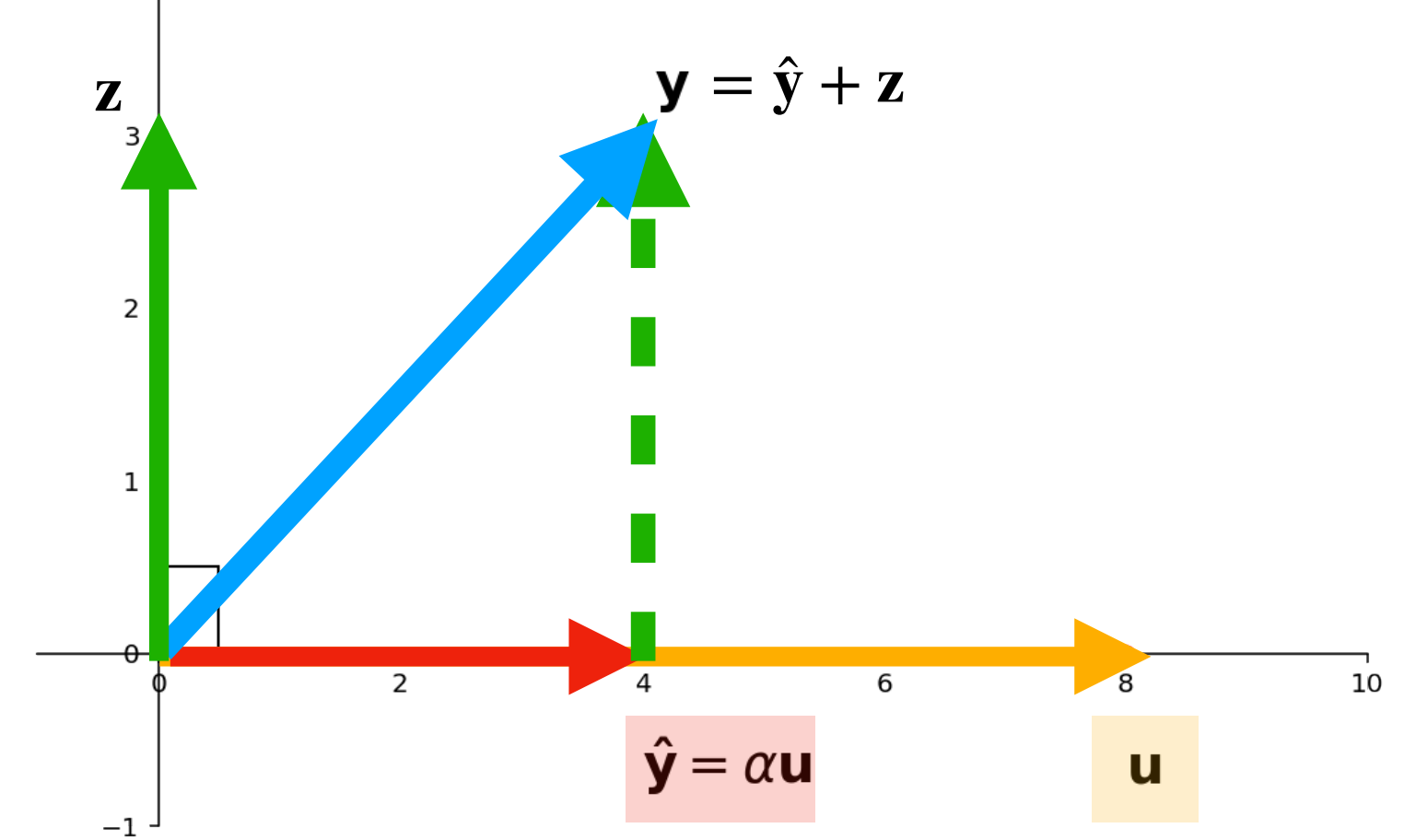
Therefore:

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

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} :

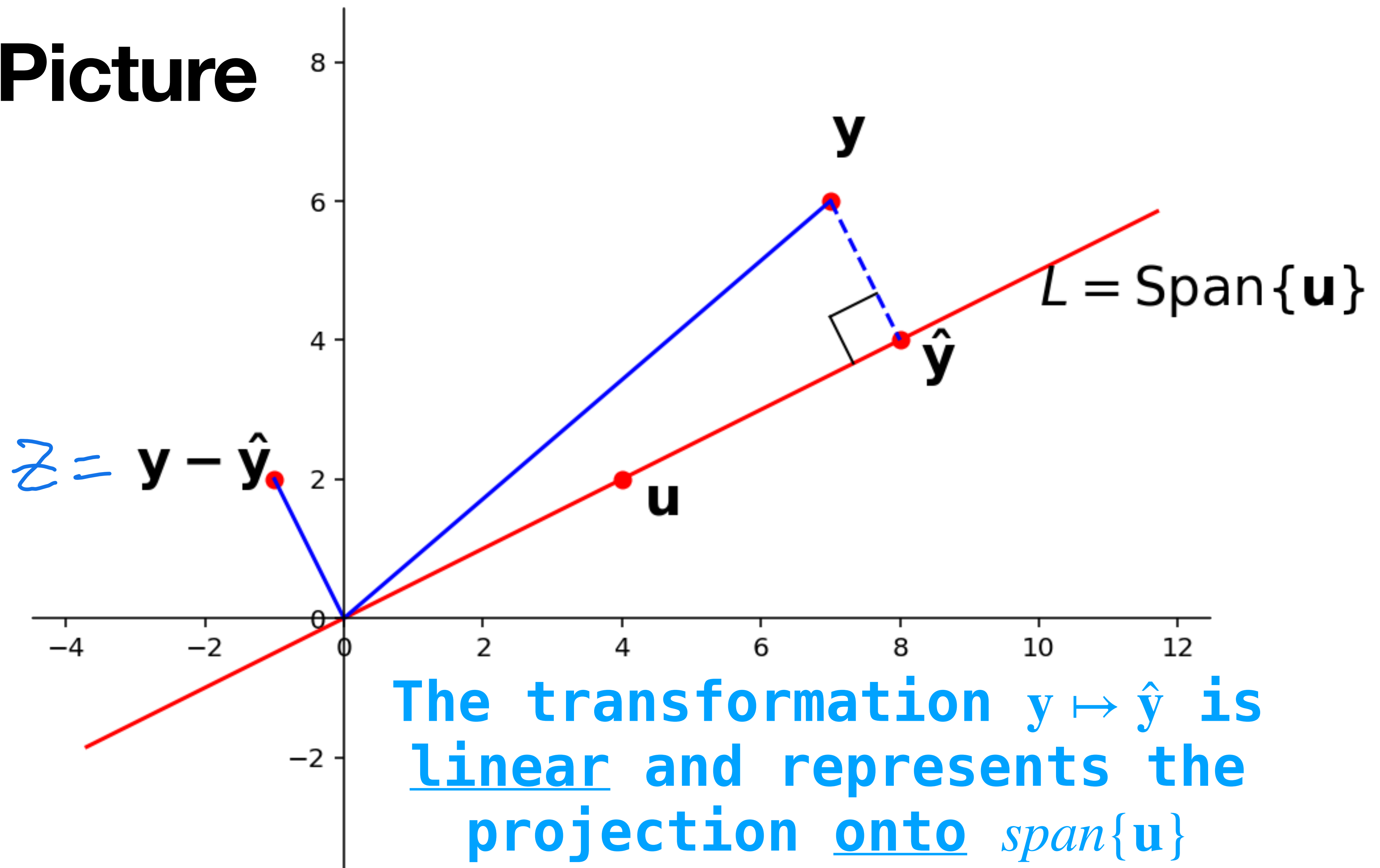
$$\langle \vec{y}, \vec{u} \rangle - \alpha \langle \vec{u}, \vec{u} \rangle = 0$$

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

$$\hat{\mathbf{y}} = \alpha \mathbf{u} = \frac{\vec{y}^T \vec{u}}{\vec{u}^T \vec{u}} \vec{u}$$

$$\vec{z} = \vec{y} - \alpha \vec{u}$$

The Picture

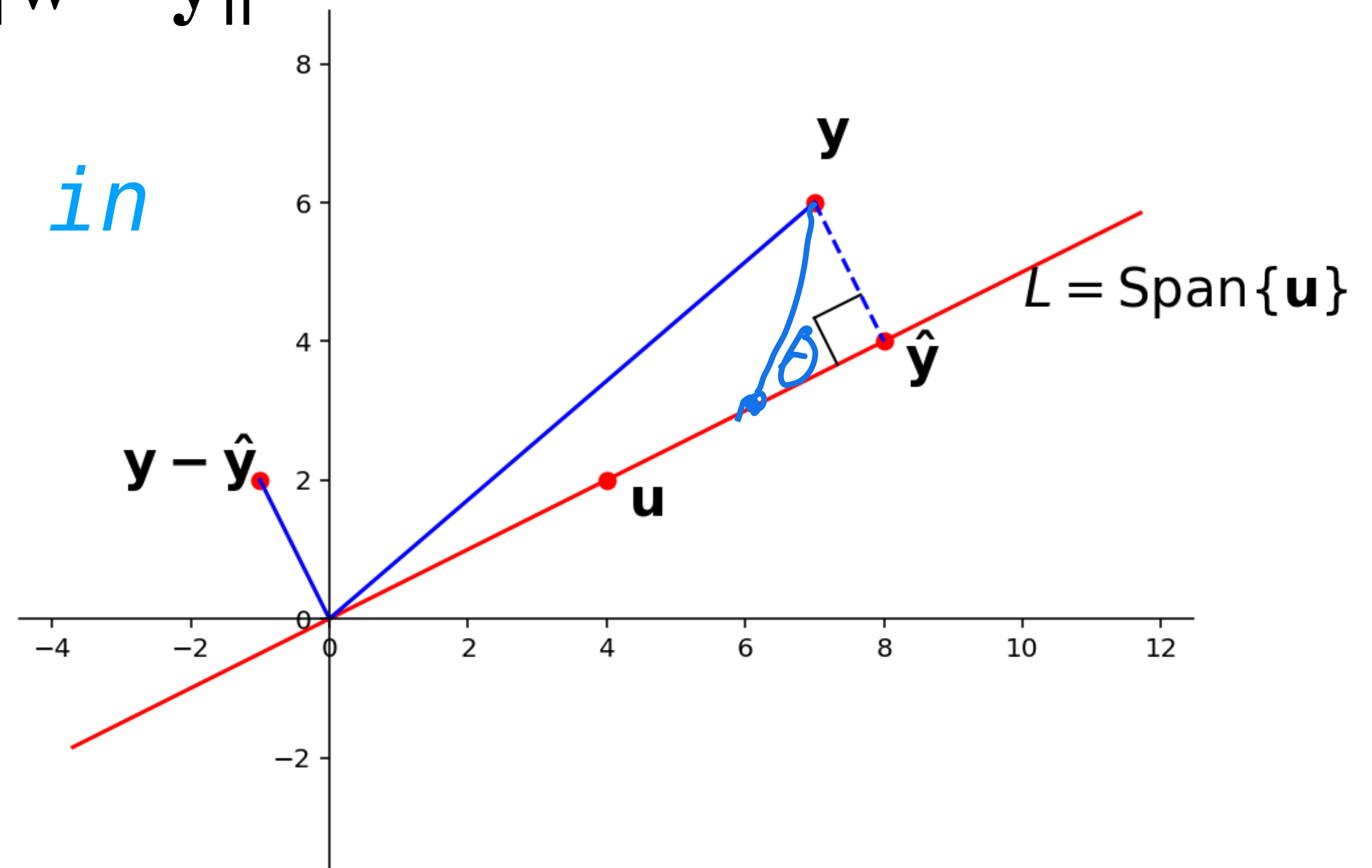


$\hat{\mathbf{y}}$ and Distance

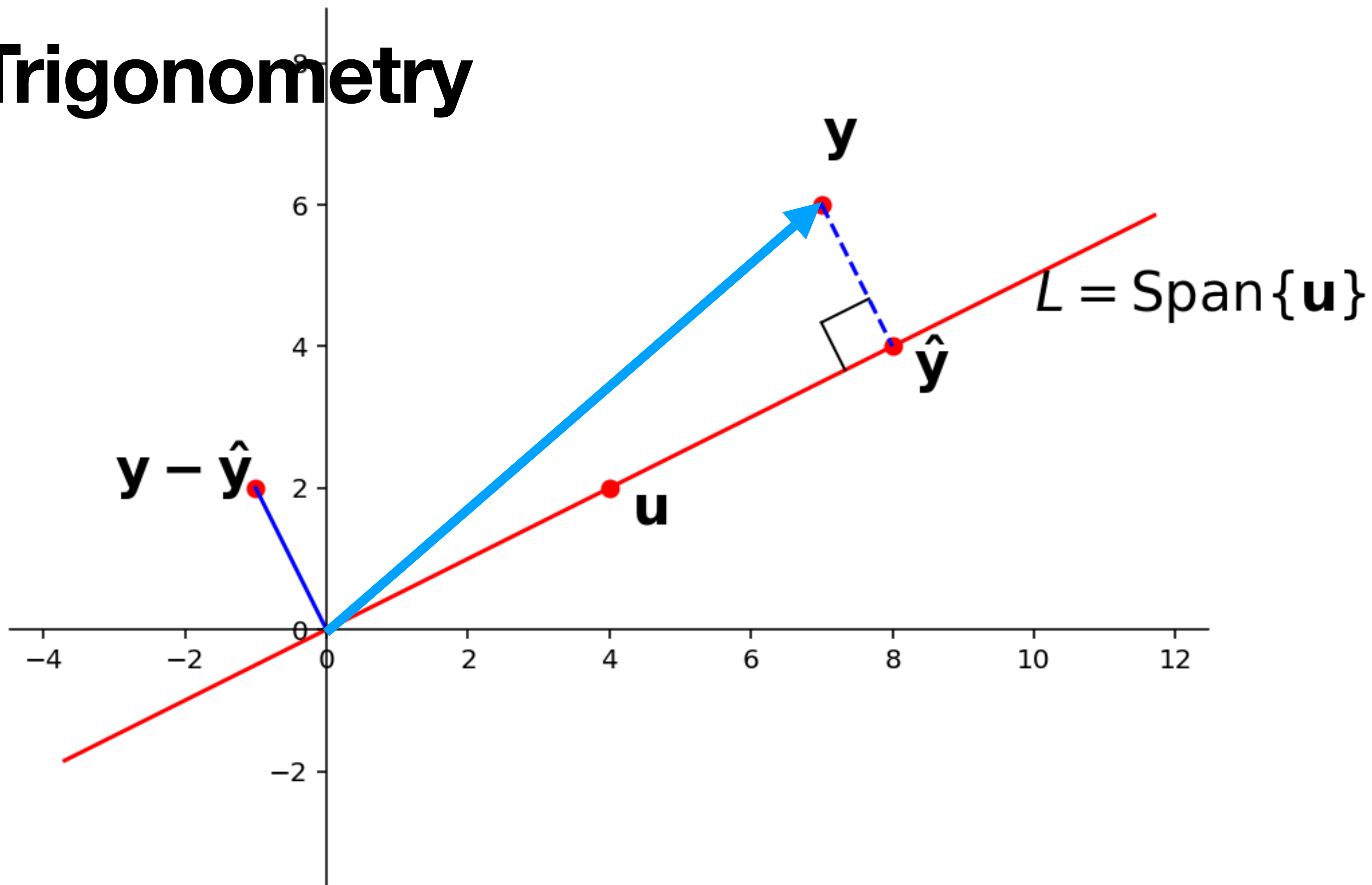
Theorem. $\|\hat{\mathbf{y}} - \mathbf{y}\| = \min_{\mathbf{w} \in \text{span}\{\mathbf{u}\}} \|\mathbf{w} - \mathbf{y}\|$

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

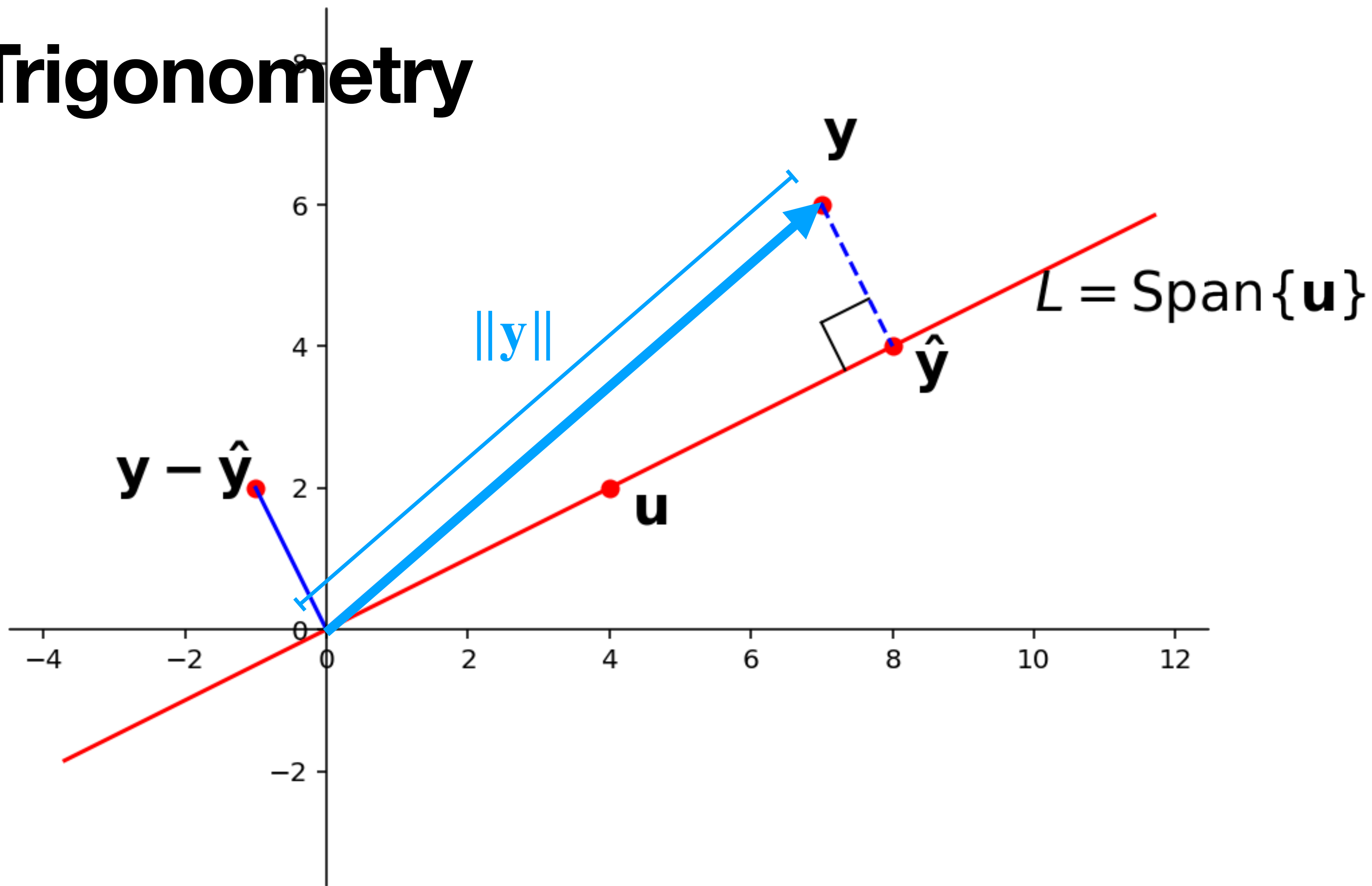
"Proof" by inspection:



The Trigonometry

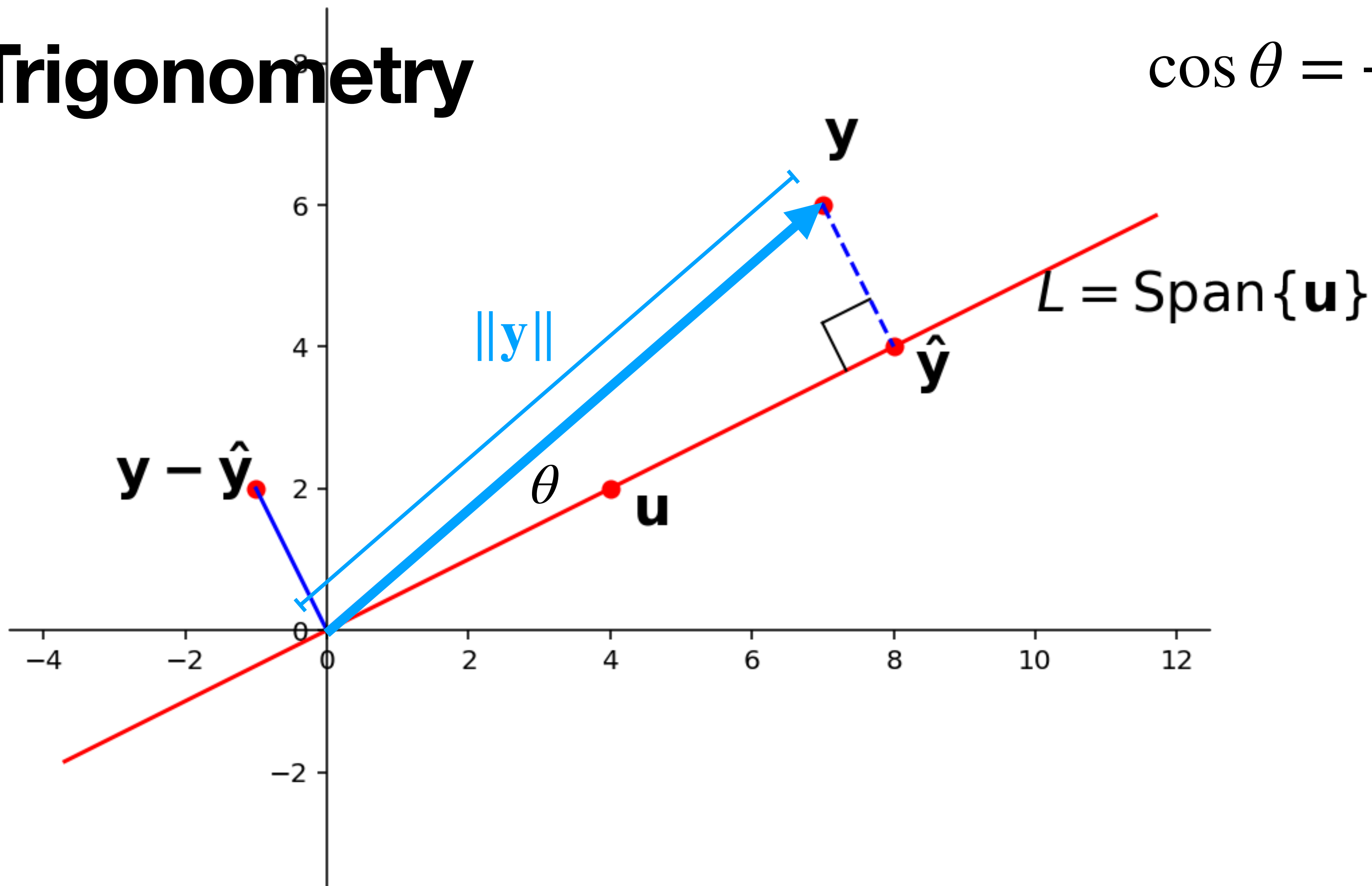


The Trigonometry



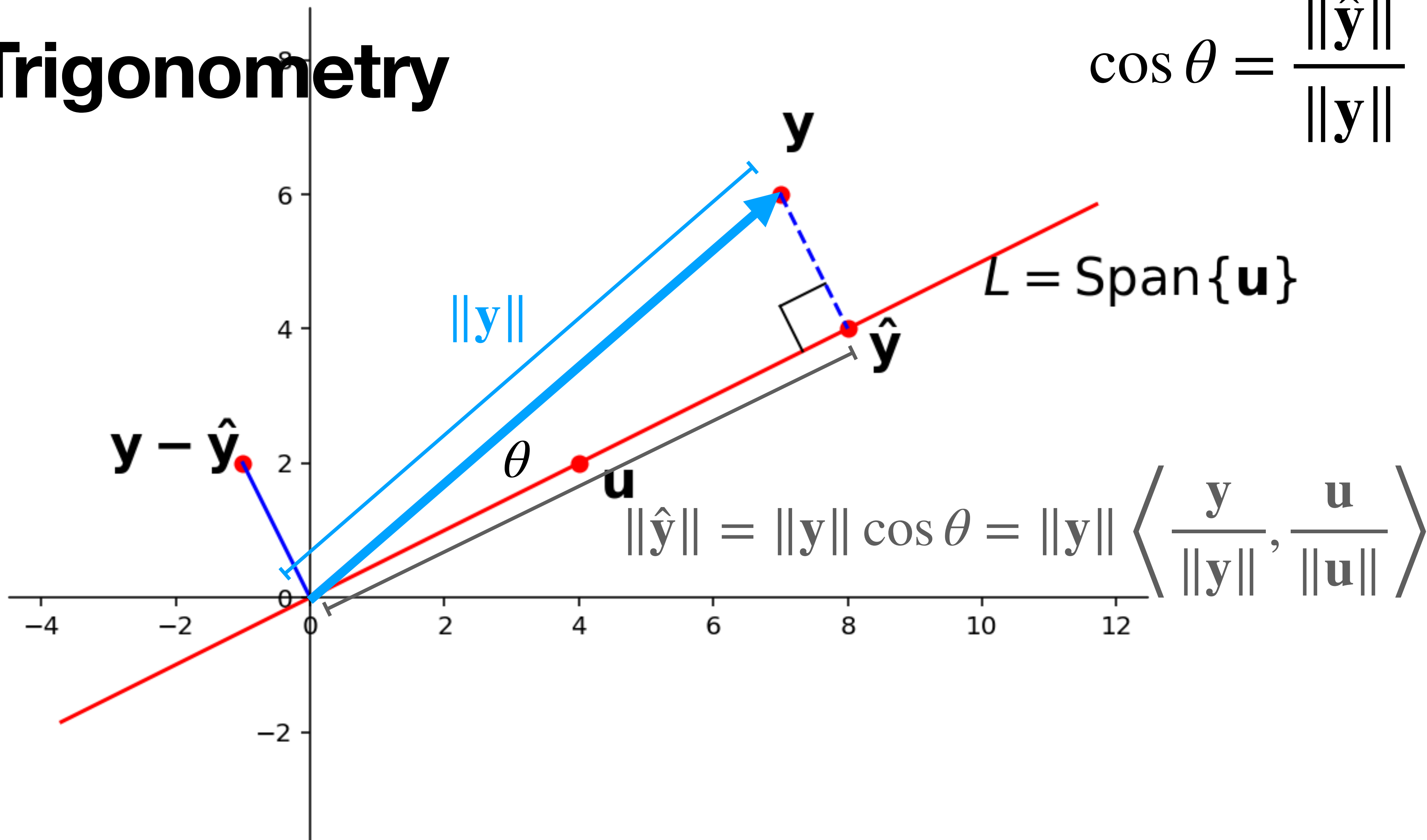
The Trigonometry

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



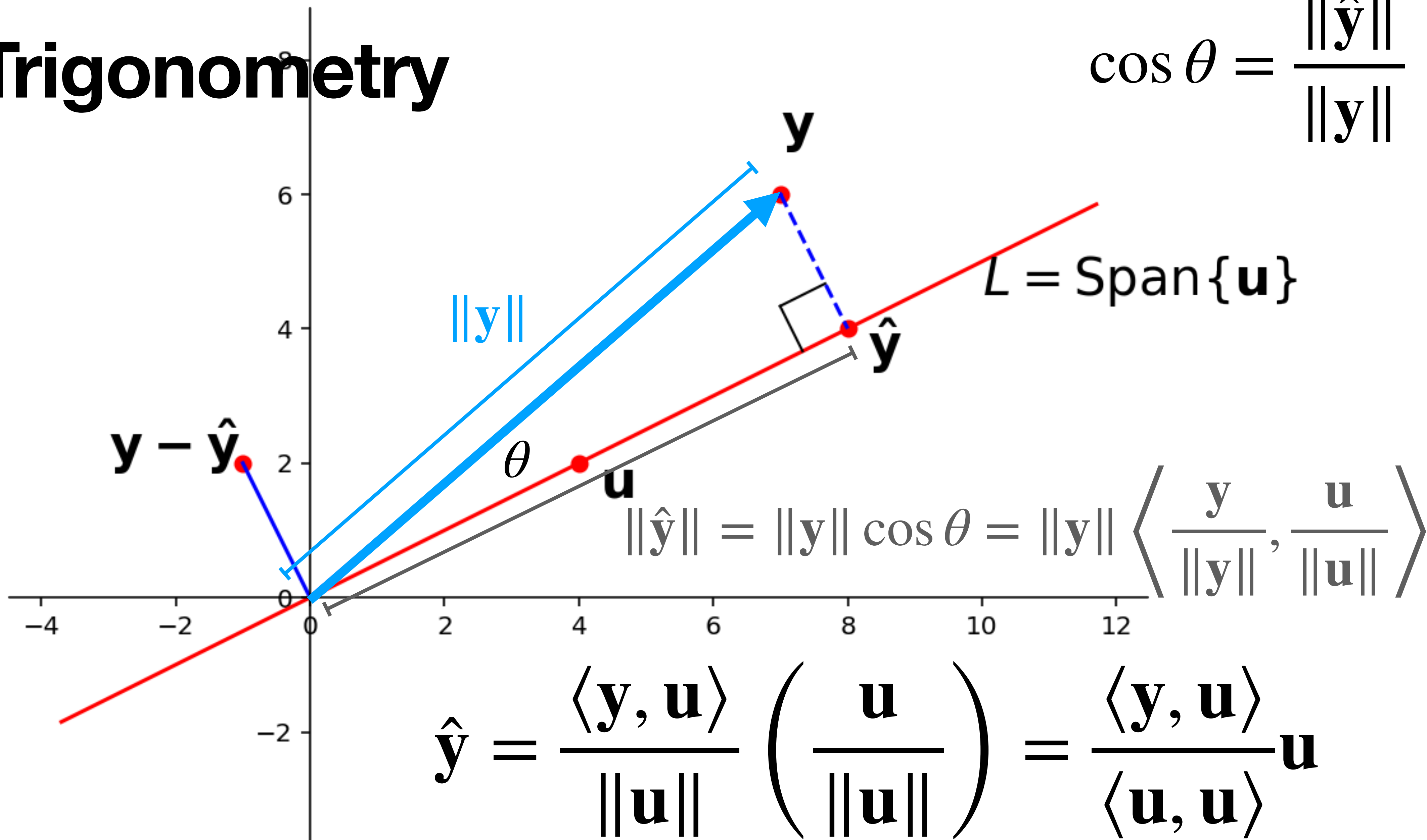
The Trigonometry

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

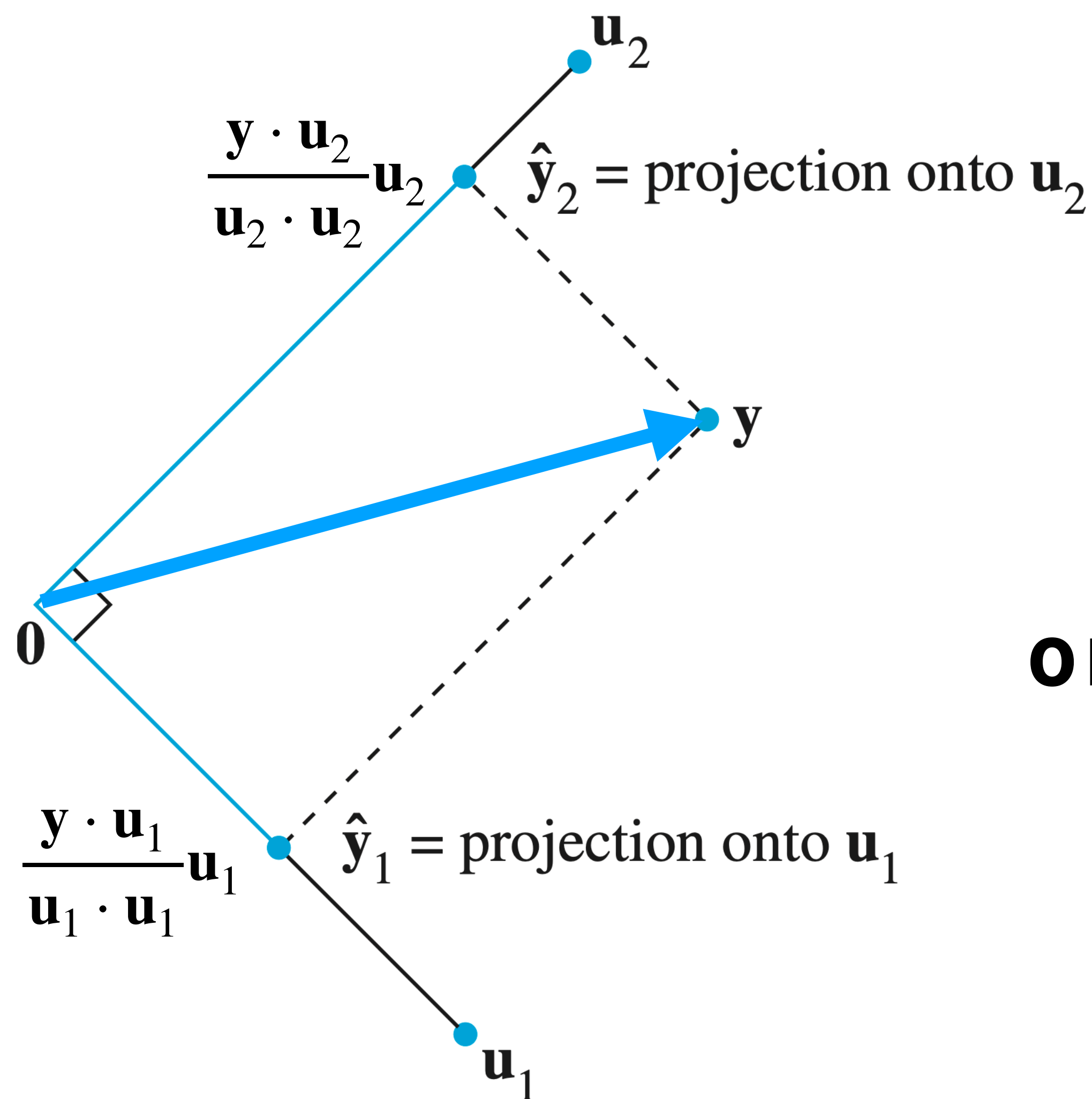


The Trigonometry

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



Orthogonal Projections and Orthogonal Bases



Each component of y written in terms of an *orthogonal* basis is an **orthogonal projection onto to a basis vector**

$$y = \hat{y}_1 + \hat{y}_2$$

How To:

Question. Find the projection of y onto the span of u

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

Question

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

output dim'n $\rightarrow m \times n$

input dim'n

Goal: 3-by-3

$$\vec{y} \mapsto \alpha \vec{u} = \left(\frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \right) \vec{u}$$

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

$$= \vec{u} \left(\frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \right) = \vec{u} \left(\frac{\vec{u} \cdot \vec{y}}{\vec{u} \cdot \vec{u}} \right) = \vec{u} \left(\frac{\vec{u}^T \vec{y}}{\vec{u} \cdot \vec{u}} \right)$$

$$\frac{\vec{u} \vec{u}^T}{\vec{u} \cdot \vec{u}} \vec{y} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \left(\frac{1}{1+1+4} \right) \begin{bmatrix} 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

Answer

$$\frac{1}{6\sqrt{5}} \begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{bmatrix}$$

Orthonormal Sets

Orthogonal sets would be easier to
work with if every vector was a
unit vector

Orthonormality

Orthonormality

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

Orthonormality

Definition. A set $\{u_1, u_2, \dots, u_p\}$ 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 $\{u_1, u_2, \dots, u_p\}$ 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 $\{u_1, u_2, \dots, u_p\}$ is an **orthonormal set** if it is 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 $\{u_1, u_2, \dots, u_p\}$ is an **orthonormal set** if it is 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

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

$$\overset{(n \times m)}{U^T} \overset{(m \times n)}{U} = I_n$$

Verify:

$$\begin{bmatrix} \vec{u}_1^T \\ \vdots \\ \vec{u}_n^T \end{bmatrix} \begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_n \end{bmatrix} = \begin{bmatrix} \vec{u}_i^T \vec{u}_j \end{bmatrix}$$

Handwritten notes:

- i, j th entry (pointing to the matrix)
- for $i = j$, $\vec{u}_i^T \vec{u}_i = 1$
- 0 for $i \neq j$

Inverses of Orthogonal Matrices

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

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

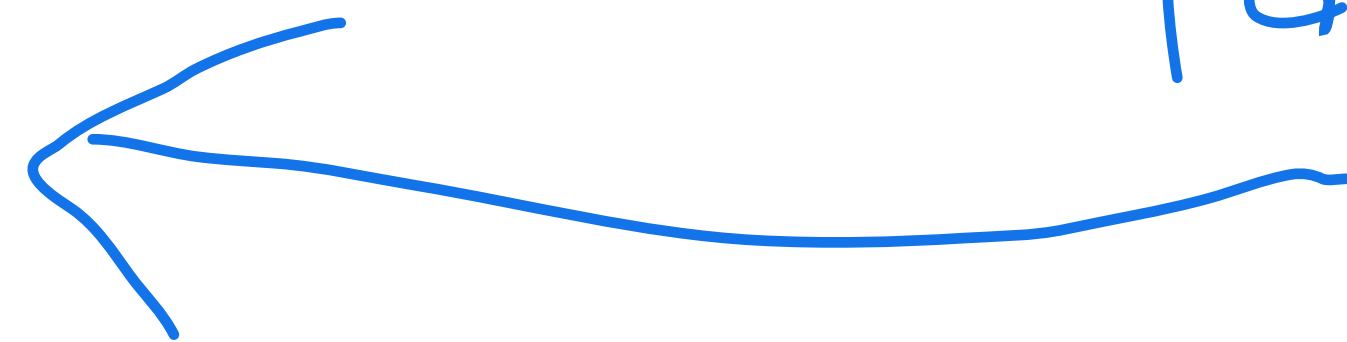
$$\det(U^{-1}) = \frac{1}{\det U}$$

Verify: By prior argument for $m=n$

$$\det(U^T) = \det(U)$$

$$\det(U) = \pm 1$$

$$(\det(U))^2 = 1$$



Orthonormal Matrices and Inner Products

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

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

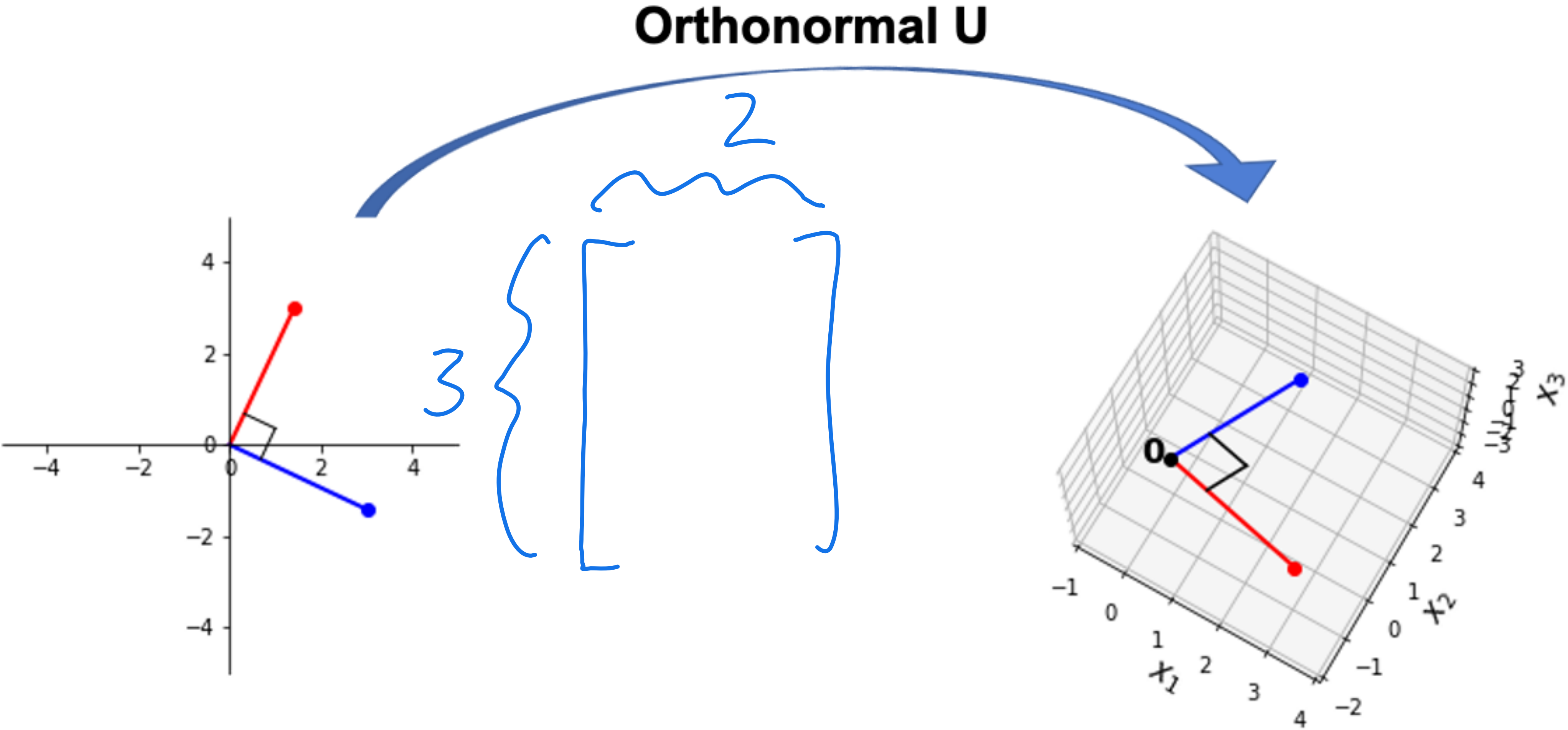
Orthonormal matrices preserve inner products

Verify: $(U\vec{x})^T U\vec{y} = \vec{x}^T U^T U \vec{y} = \vec{x}^T \vec{y}$

Length, Angle, Orthogonality Preservation

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

The Picture



Example

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

Show $\|Ux\|^2 = \|x\|^2$

$$x = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$$

Question (Conceptual)

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

Answer

If $A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n]$ 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