# Orthogonal Sets and Projection

Geometric Algorithms Lecture 22

## Introduction

## Recap Problem

(**Final Review**) Find a set of vectors which forms a basis for the hyperplane given by the equation

$$x_1 + 3x_2 - 4x_3 + 6x_4 = 0$$

## Answer

$\lceil -3 \rceil$	$\lceil 4 \rceil$	-6
1	0	0
0	1	0
		_ 1 _

## Objectives

- 1. Recap analytic geometry in  $\mathbb{R}^n$ .
- 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

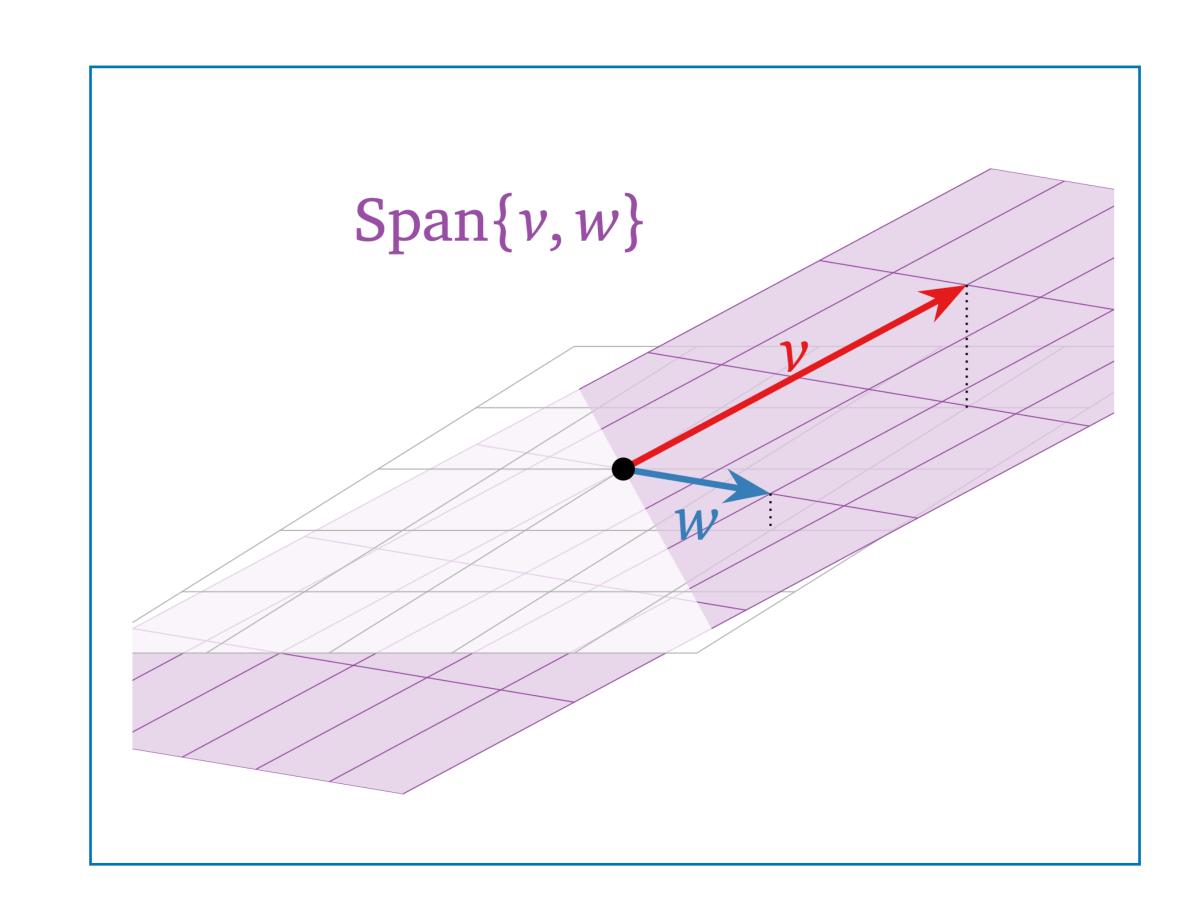
## Recap: Analytic Geometry

## 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

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

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

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

#### Recall: Inner Products

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

**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  $\ell^2$  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  $\ell^2$  norm of a vector  $\mathbf{v}$  in  $\mathbb{R}^n$  is

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

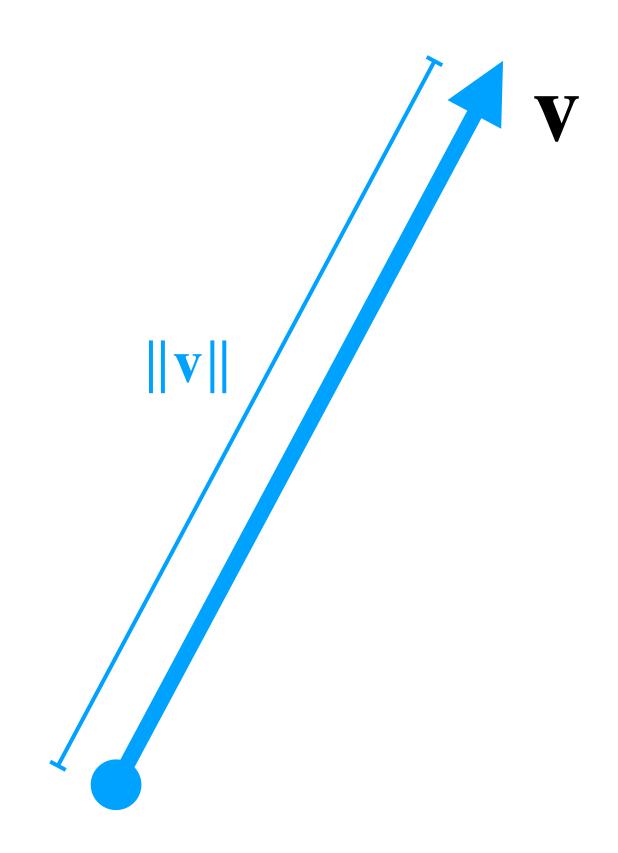
The norm of a vector is the square root of the inner product with itself.

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

## Recall: Norms and Length

Norms give us a notion of length.

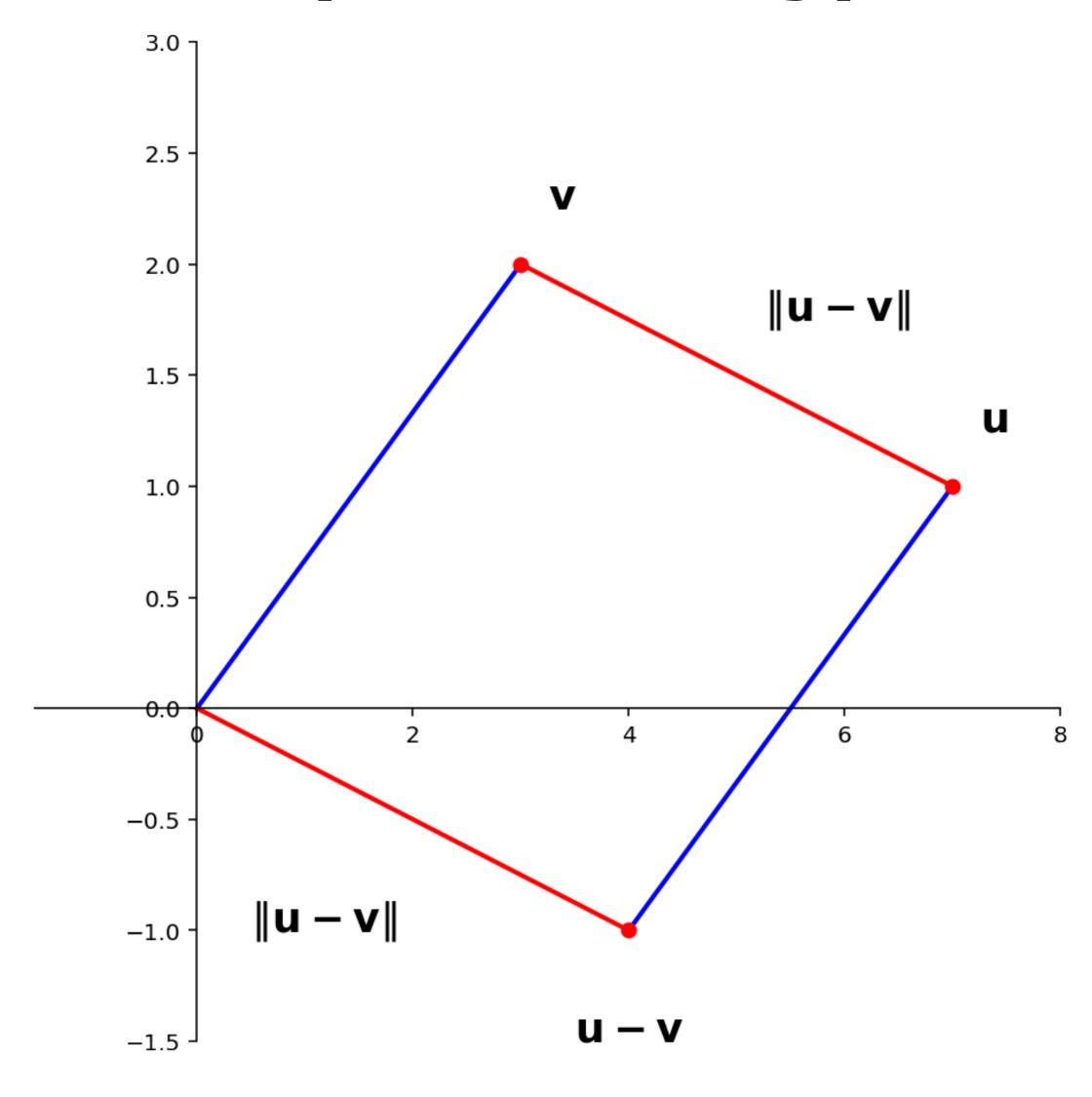
In  $\mathbb{R}^2$  and  $\mathbb{R}^3$  this is our existing notion of length.



#### Recall: Distance

If we know how to calculate lengths of vectors, we know how to calculate distances.

## Recall: Distance (Pictorially)



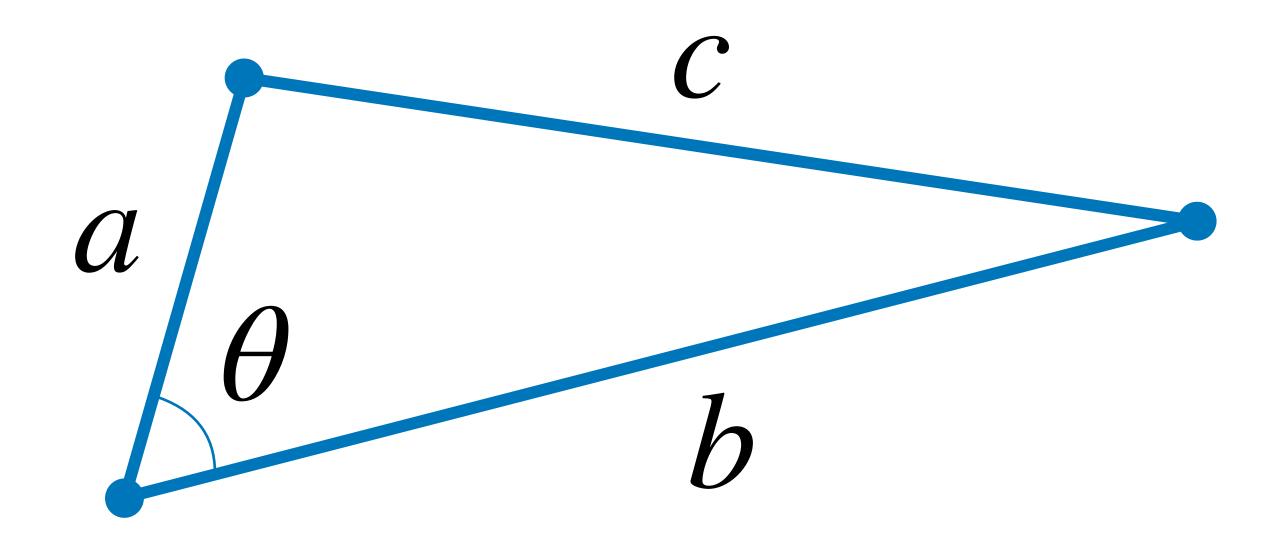
## Recall: Distance (Algebraically)

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

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

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

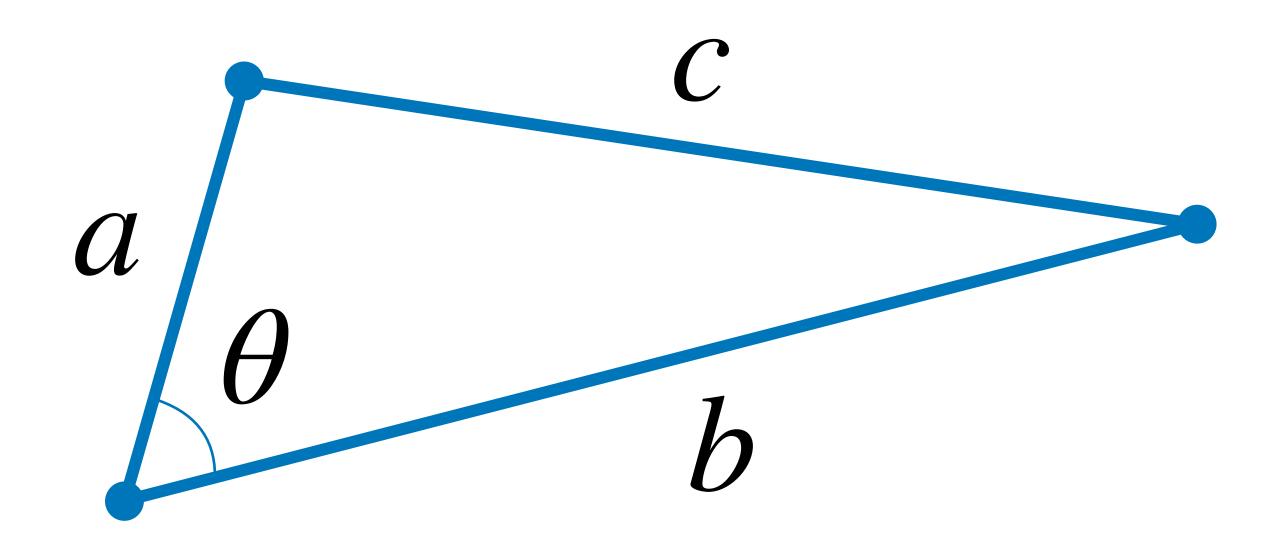
#### Recall: Law of Cosines



#### Theorem.

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

#### Recall: Law of Cosines

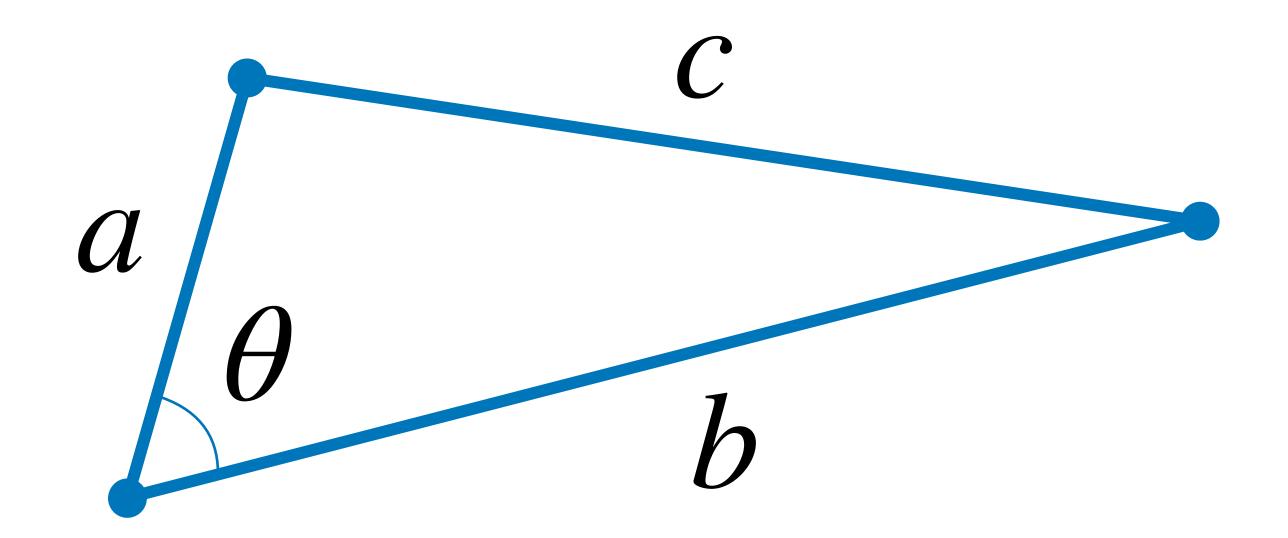


Theorem.

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

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  $\theta$  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  $\ell^2$  normalizations.

**Definition (Informal).** Two nonzero vectors  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathbb{R}^n$  are **orthogonal** if the angle between them is  $90^{\circ}$ .

**Definition (Informal).** Two nonzero vectors  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathbb{R}^n$  are **orthogonal** if the angle between them is  $90^{\circ}$ .

Orthogonal and perpendicular are the same thing.

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

Verify:

Example:

#### In All

#### With inner products:

- 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  $\{u_1, u_2, ..., u_p\}$  of vectors from  $\mathbb{R}^n$  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.

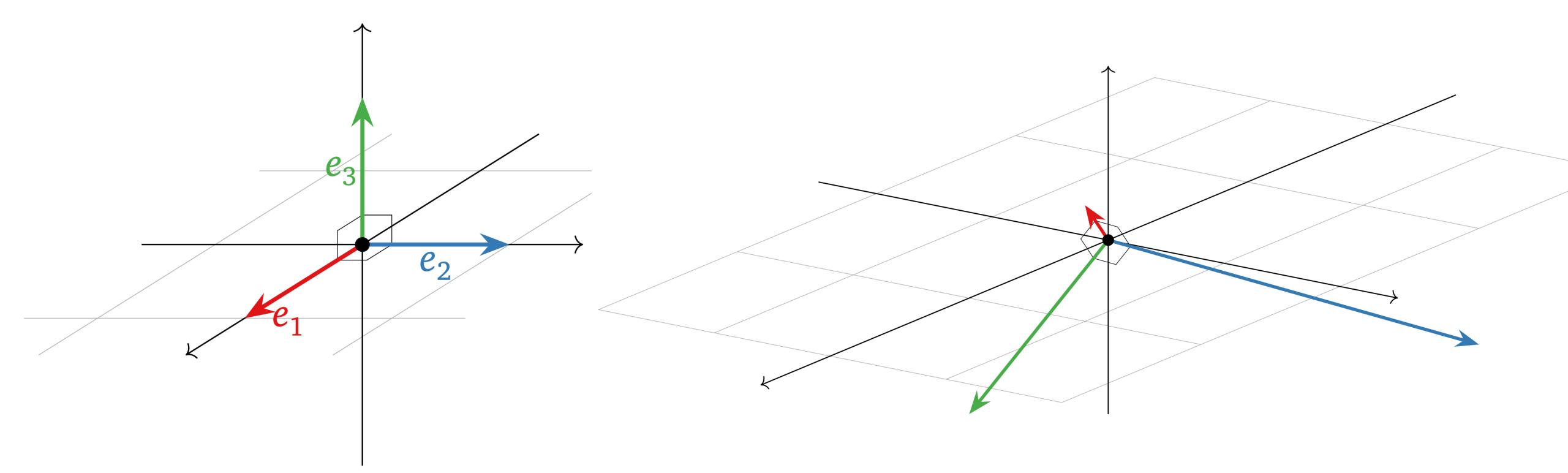
### Example

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

Verify:

# 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

## Orthogonal Sets and Independence

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

Verify:

## The Takeaway

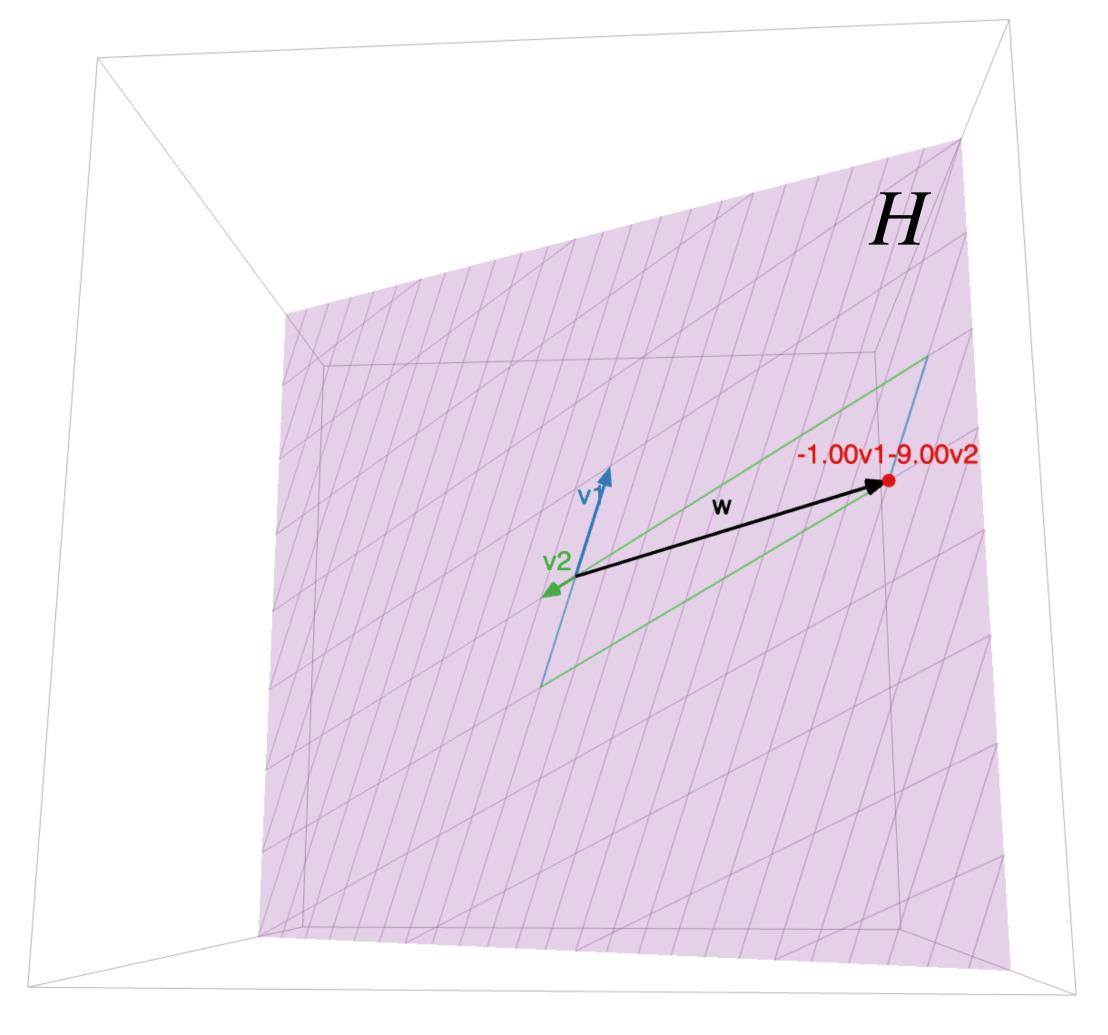
```
If \{u_1, u_2, ..., u_k\} is an orthogonal set, then it is a basis for span\{u_1, u_2, ..., u_k\}.
```

## 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

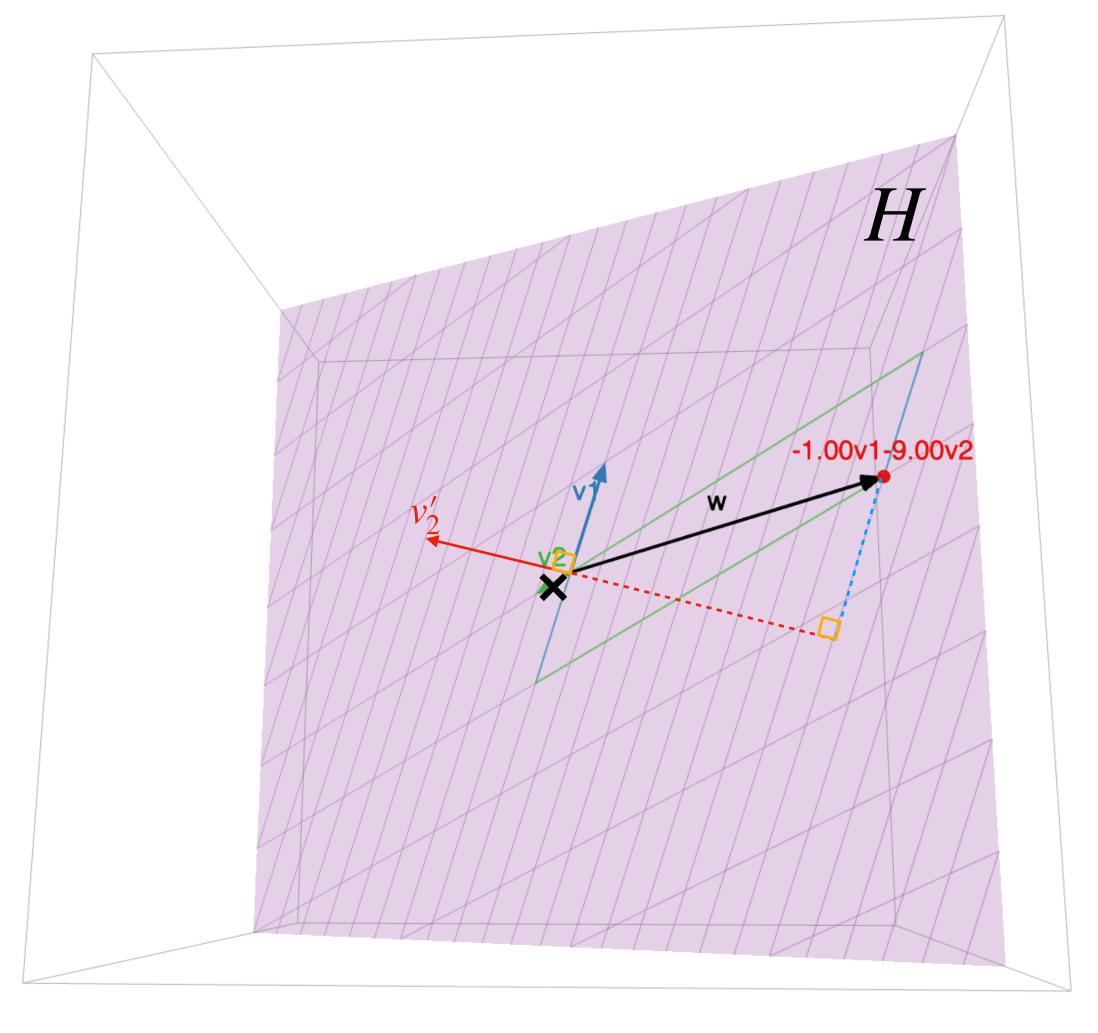
**Definition.** An **orthogonal basis** for a subspace W of  $\mathbb{R}^n$  is a basis for W which is also an orthogonal set.



 $v_1$  and  $v_2$  form a basis of H

## Orthogonal Basis

**Definition.** An **orthogonal basis** for a subspace W of  $\mathbb{R}^n$  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?

**Question.** Given a basis  $\{\mathbf u_1, \mathbf u_2, ..., \mathbf u_p\}$  for a subspace W of  $R^n$  and a vector  $\mathbf w$  in W, weights  $c_1, c_2, ..., c_p$  such that

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

**Question.** Given a basis  $\{\mathbf u_1, \mathbf u_2, ..., \mathbf u_p\}$  for a subspace W of  $R^n$  and a vector  $\mathbf w$  in W, weights  $c_1, c_2, ..., c_p$  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.

**Question.** Given a basis  $\{\mathbf u_1, \mathbf u_2, ..., \mathbf u_p\}$  for a subspace W of  $R^n$  and a vector  $\mathbf w$  in W, weights  $c_1, c_2, ..., c_p$  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  $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_p\}$ , if  $\mathbf{y} = c_1 \mathbf{u}_1 + ... + c_p \mathbf{u}_p$  then for j = 1,...,p

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

Verify:

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

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

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

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

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

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

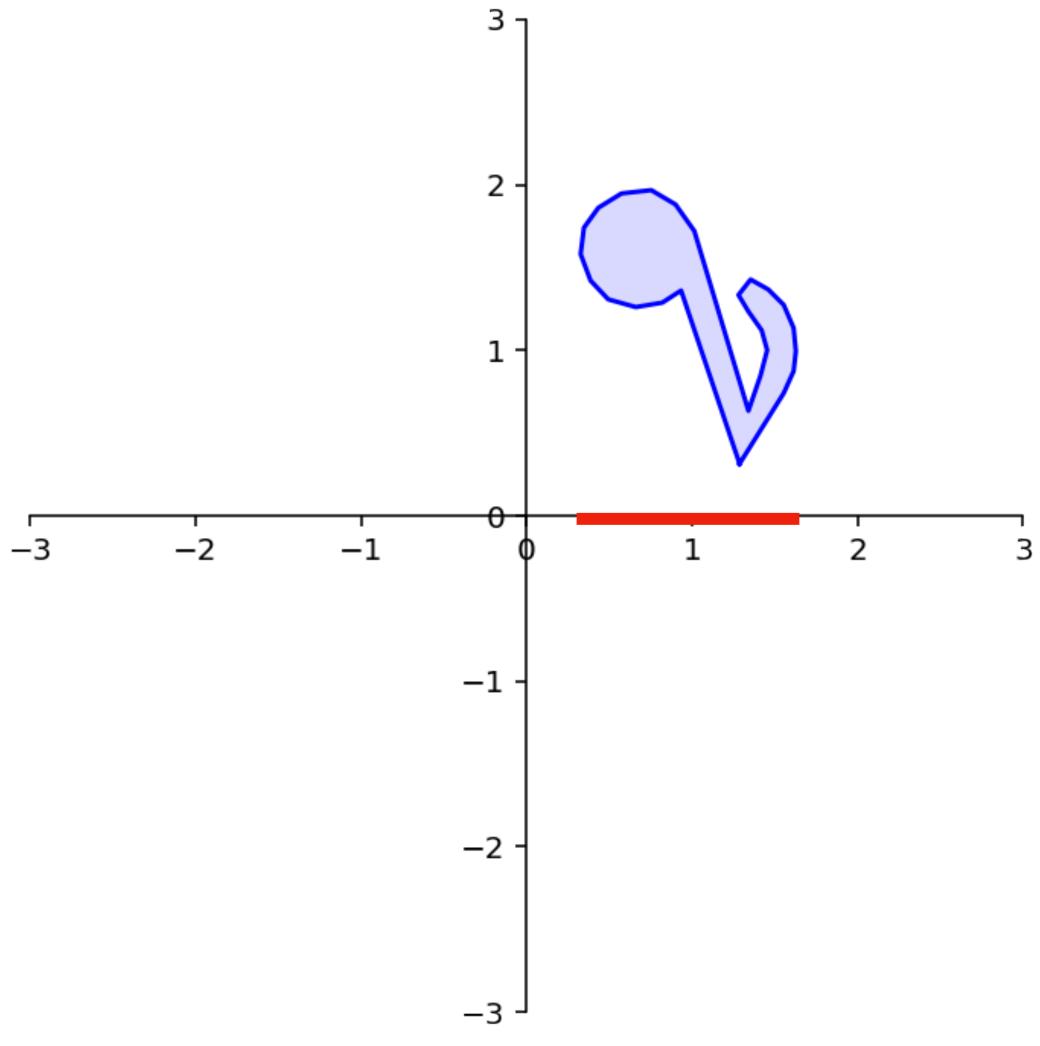
#### Question

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

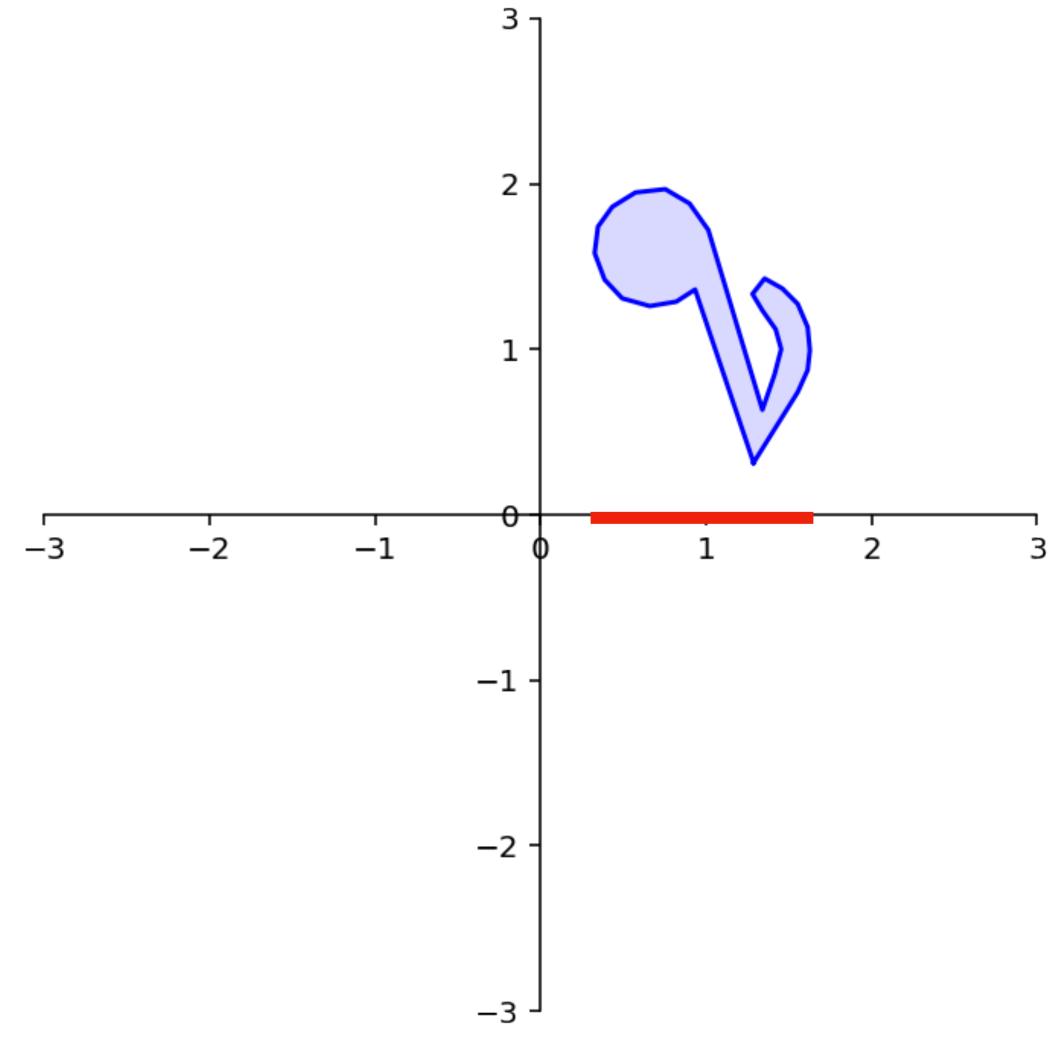
$$u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \qquad u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \qquad u_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

Answer:  $u_1 - 2u_2 - 2u_3$ 

# Why does that formula in the last example work?

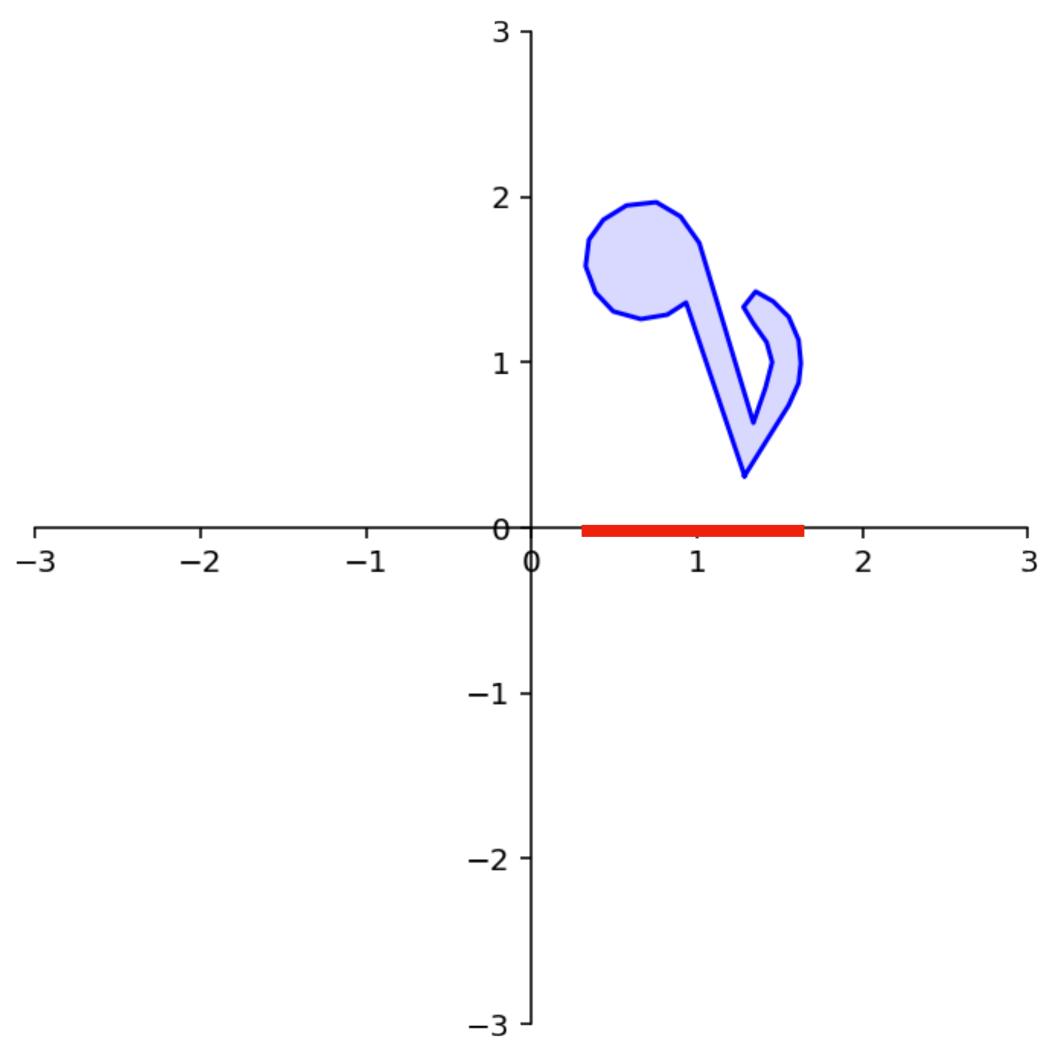


We've seen simple projections in  $\mathbb{R}^2$ .



We've seen simple projections in  $\mathbb{R}^2$ .

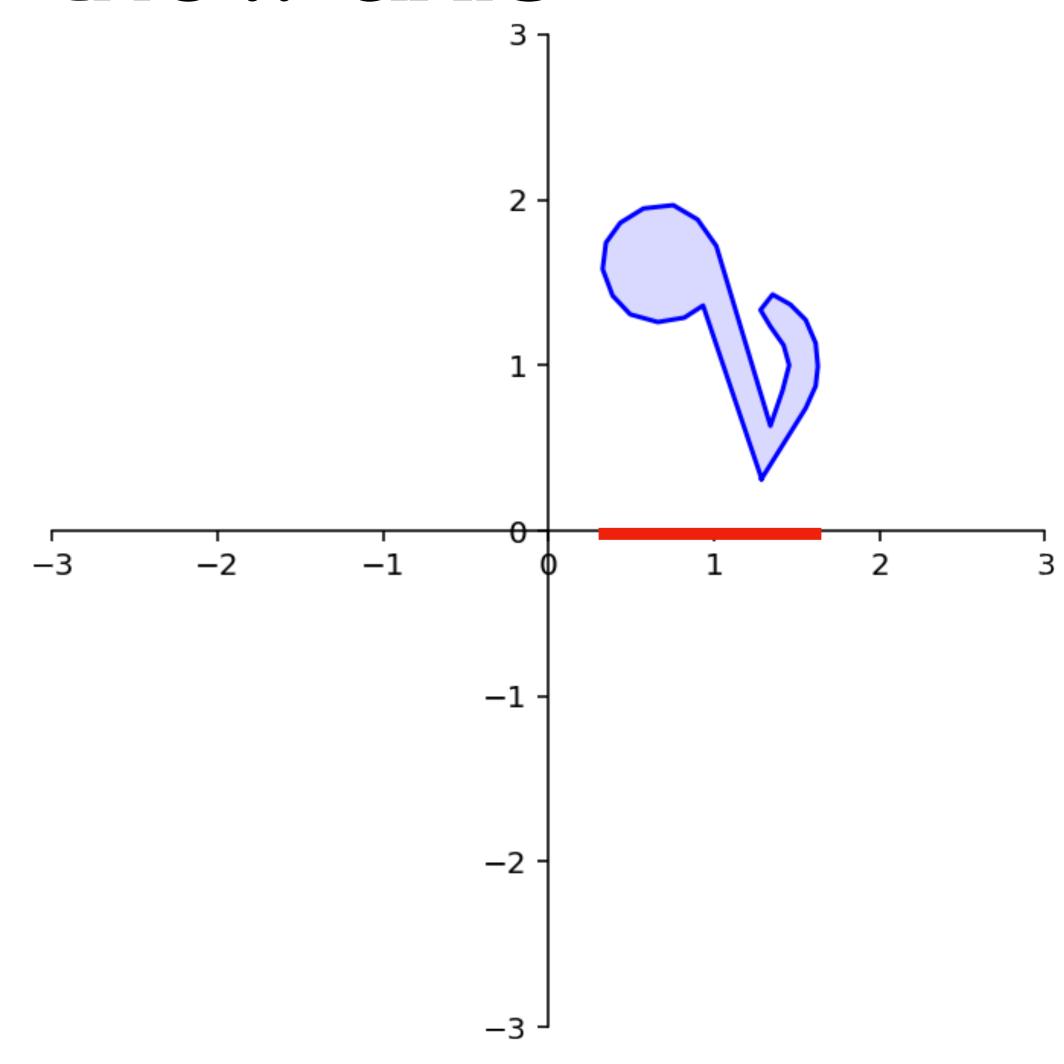
We're going to generalize this idea.

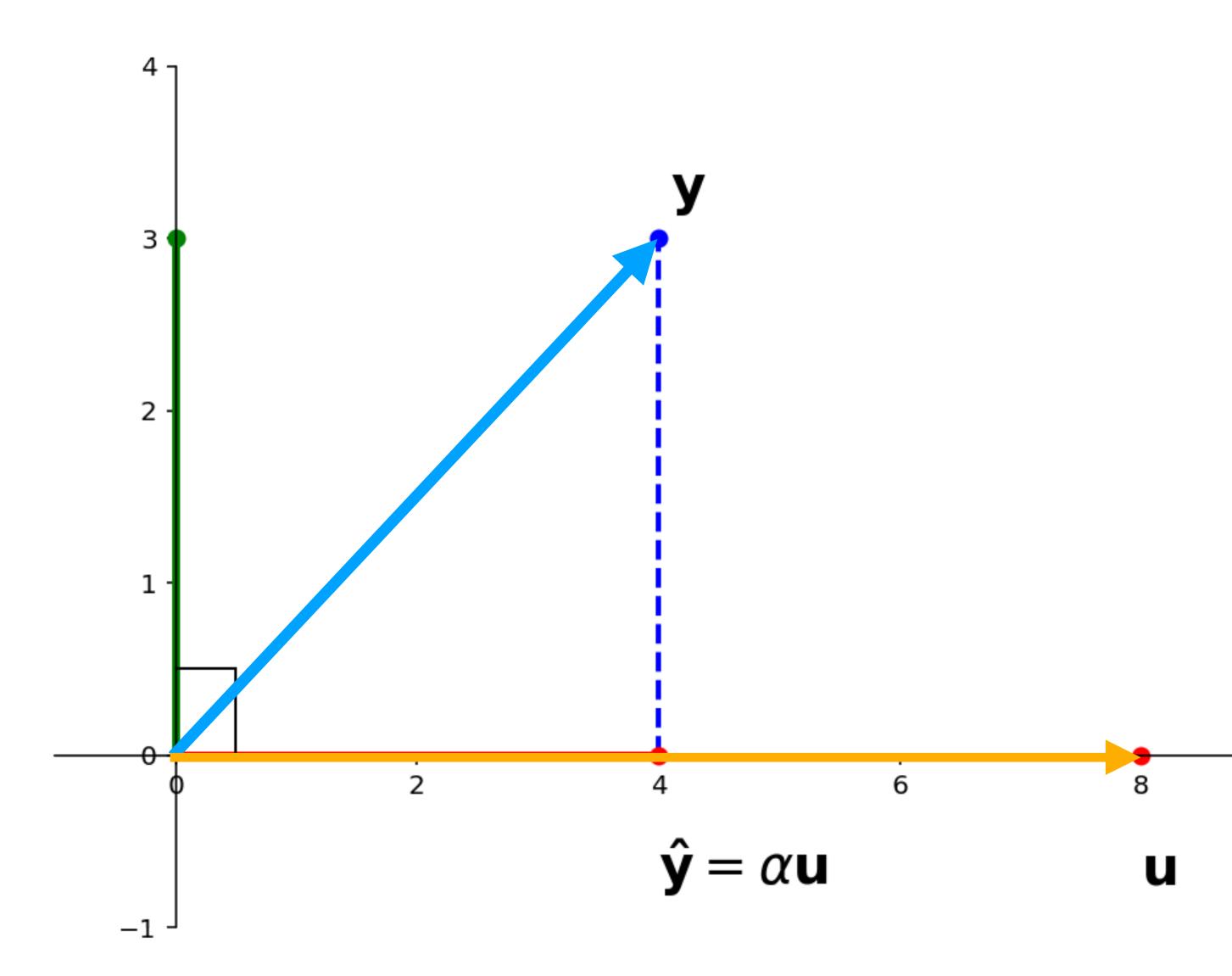


We've seen simple projections in  $\mathbb{R}^2$ .

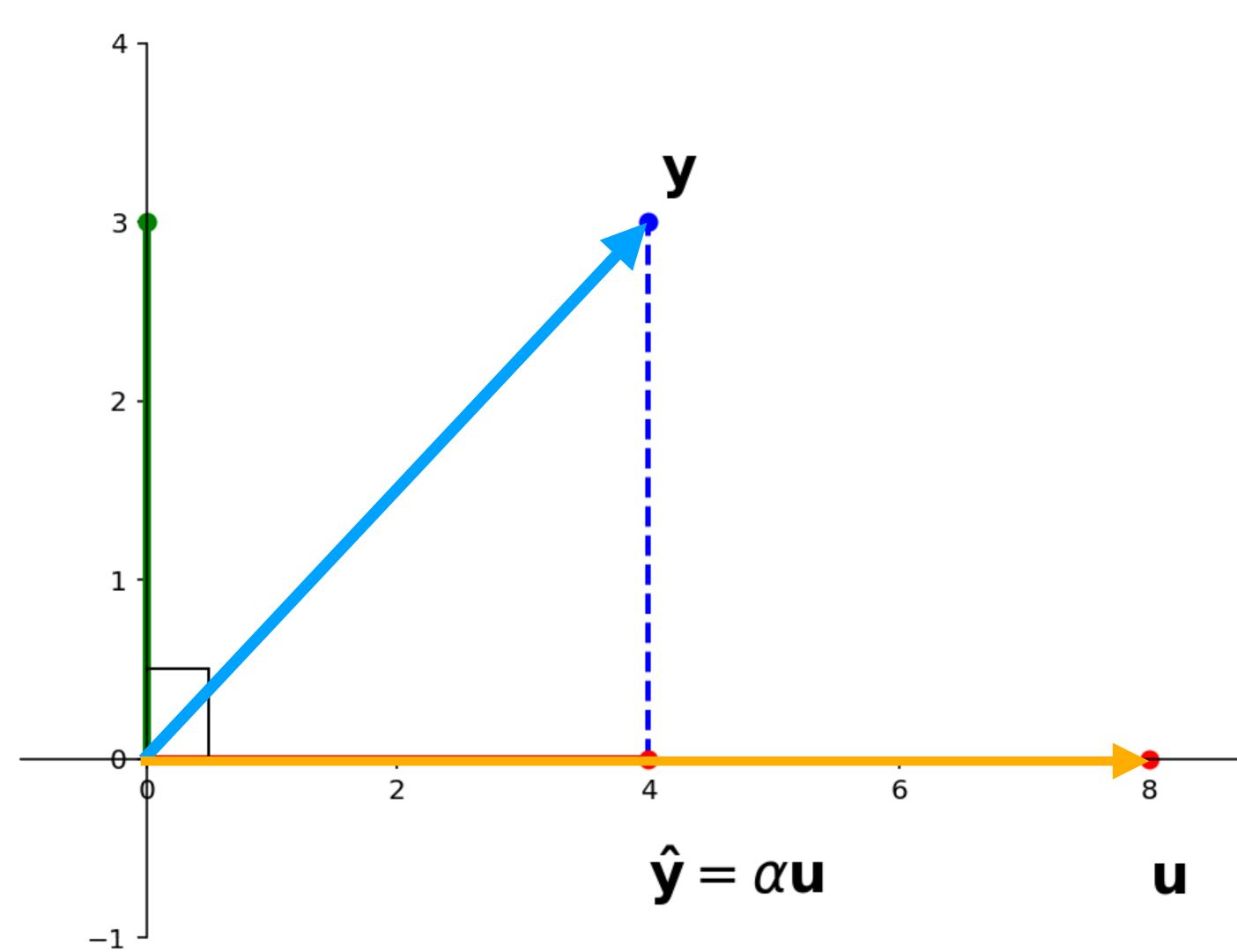
We're going to generalize this idea.

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



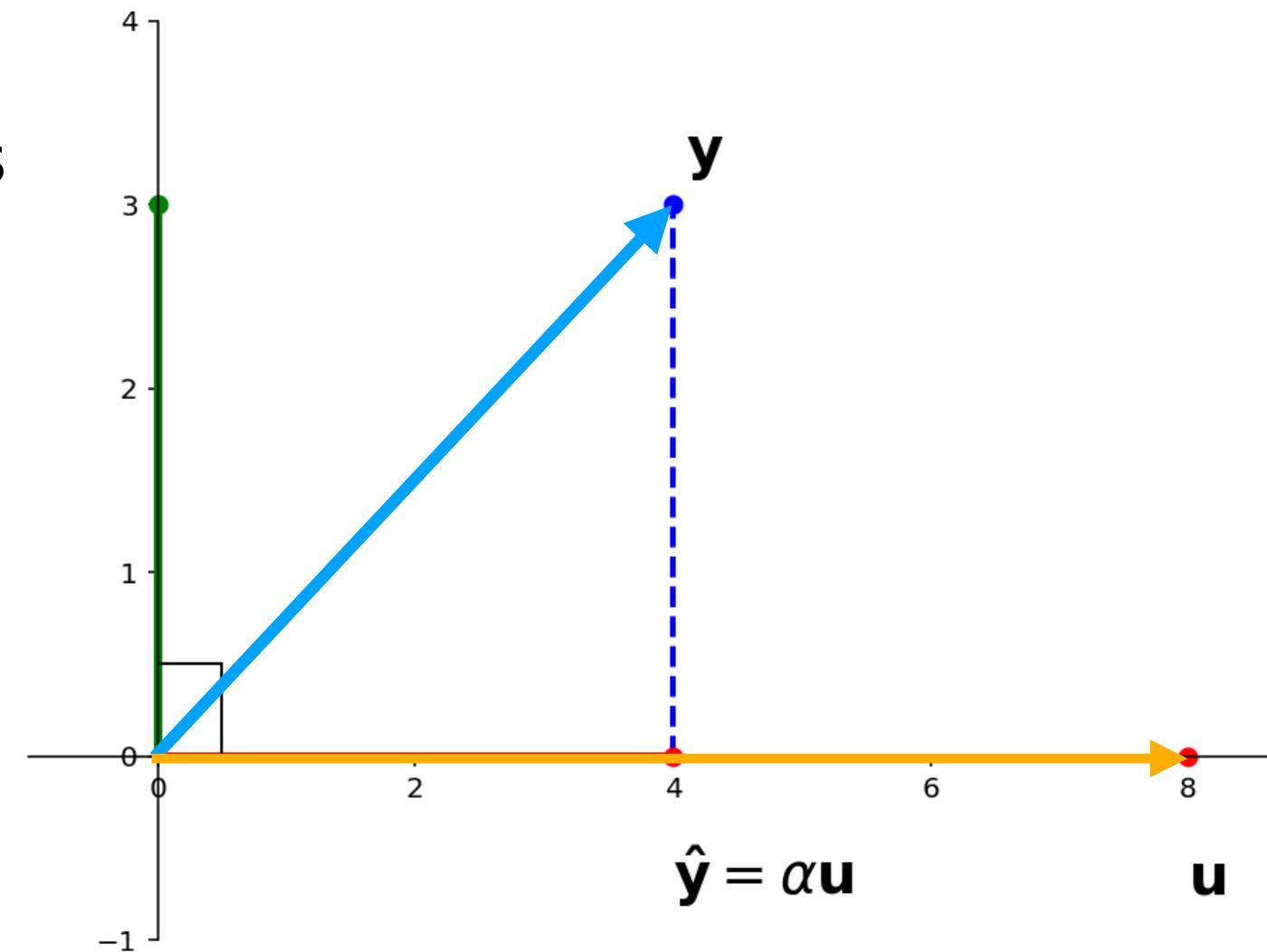


Question. Given vectors y and u in  $R^n$ , find vectors  $\hat{y}$  and z such that



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

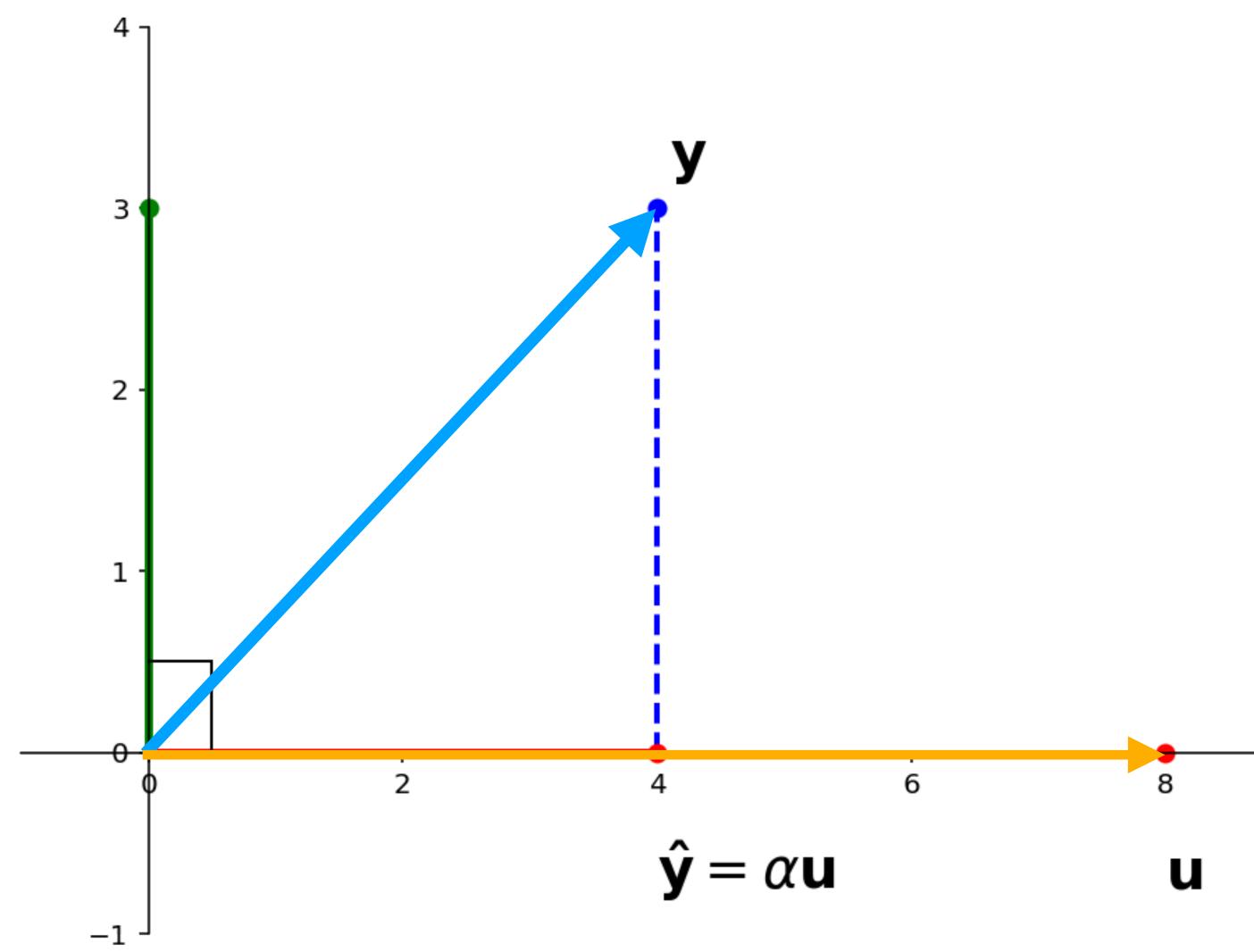
 $\Rightarrow$  z is orthogonal to u (i.e.,  $z \cdot u = 0$ )



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

 $\Rightarrow$  z is orthogonal to u (i.e.,  $z \cdot u = 0$ )

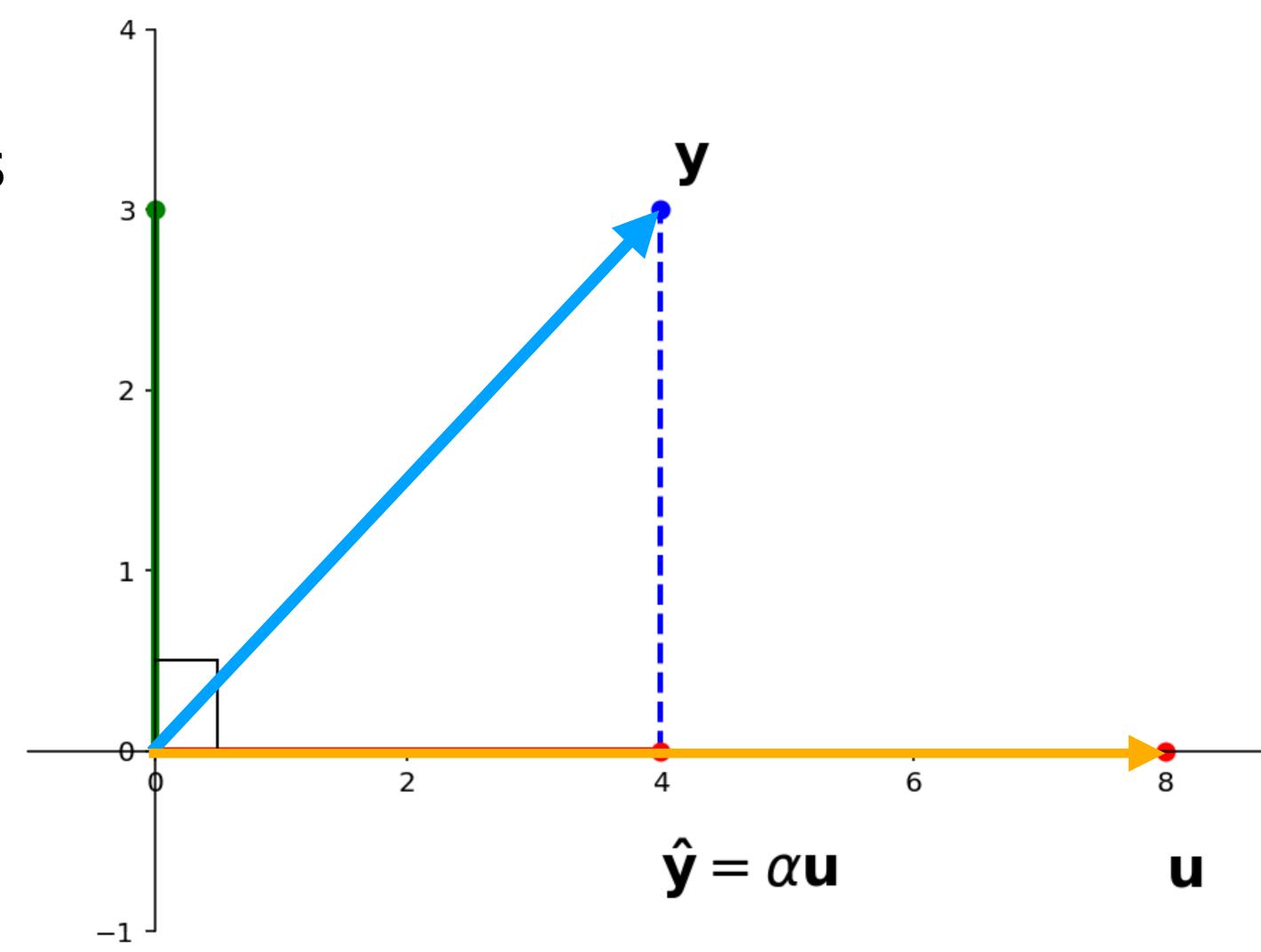
 $\Rightarrow \hat{\mathbf{y}} \in span\{\mathbf{u}\}$ 

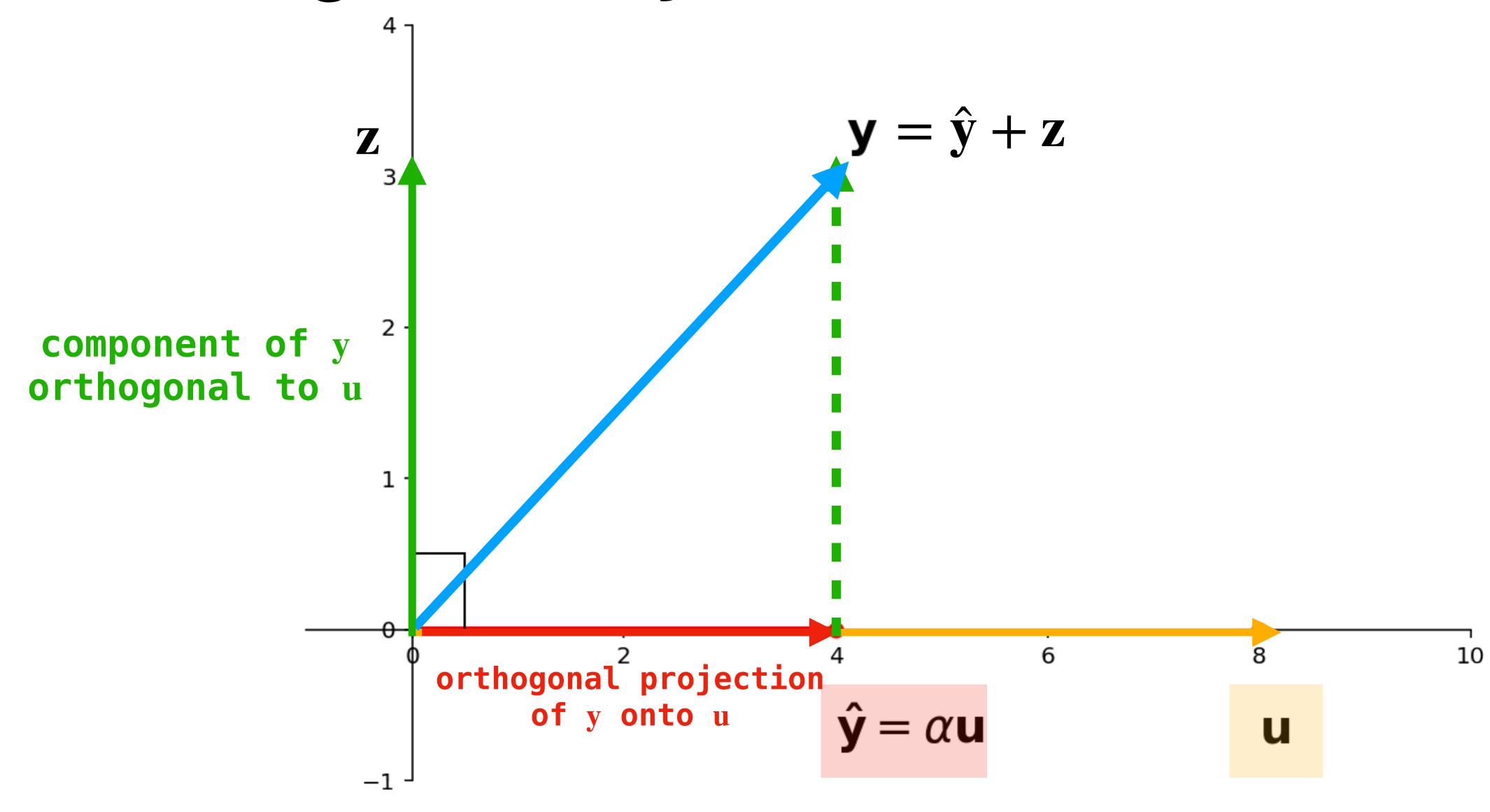


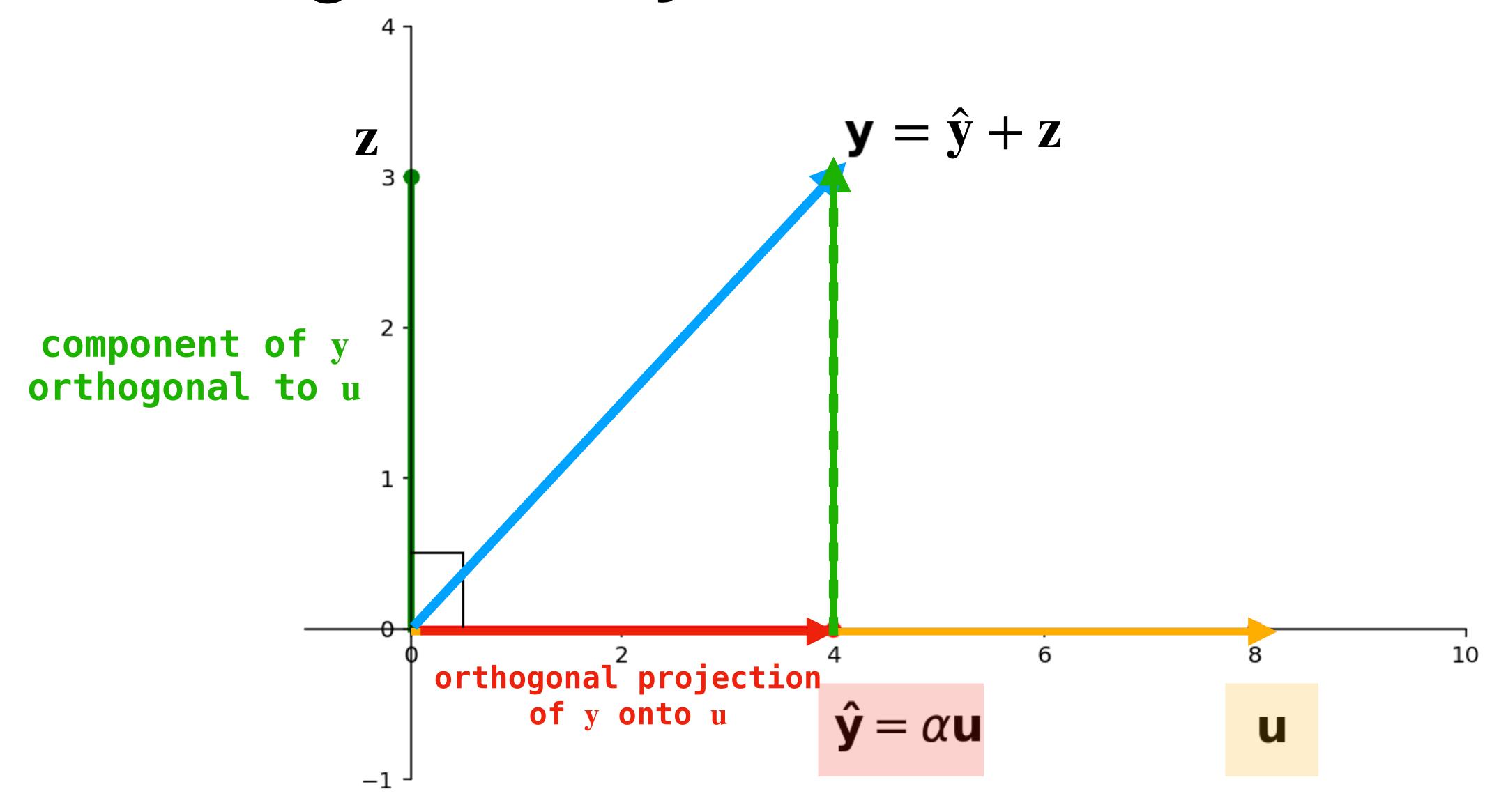
Question. Given vectors y and u in  $R^n$ , find vectors  $\hat{y}$  and z such that

 $\Rightarrow$  z is orthogonal to u (i.e.,  $z \cdot u = 0$ )

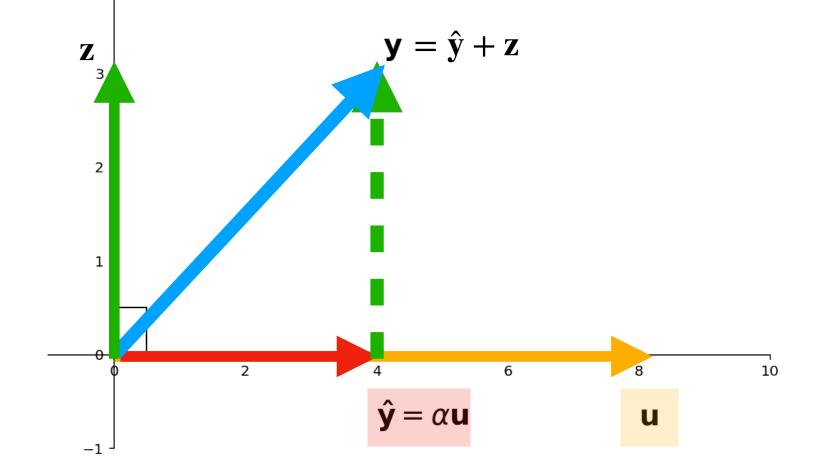
- $\Rightarrow \hat{\mathbf{y}} \in span\{\mathbf{u}\}$
- $y = \hat{y} + z$

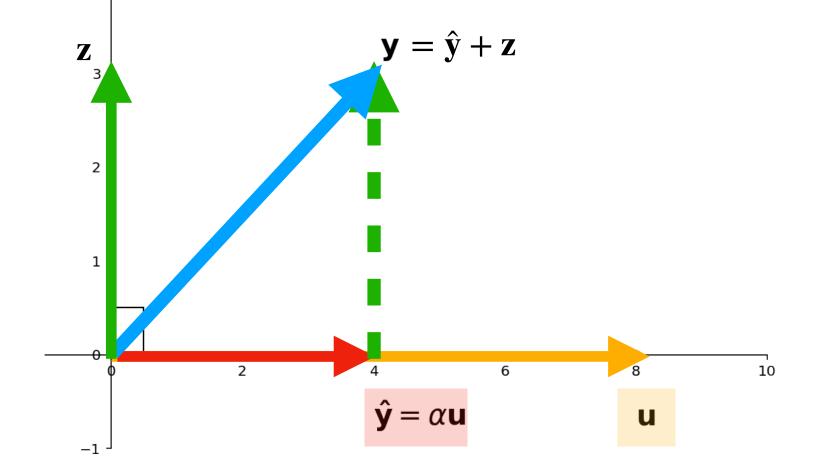




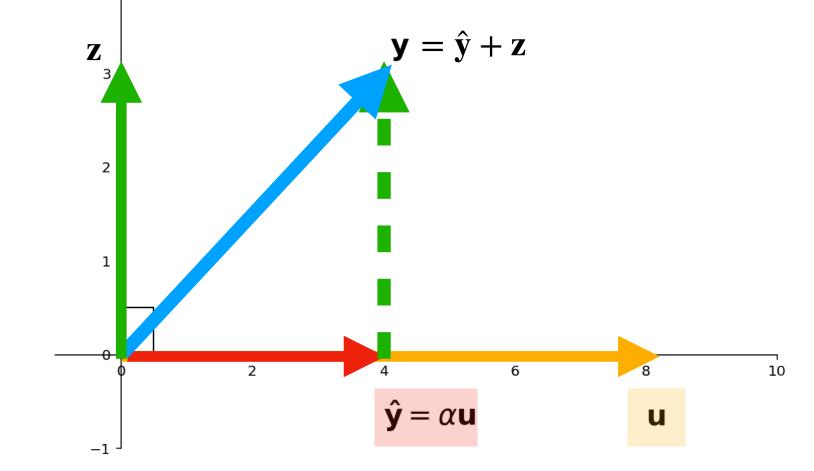


# How do we find the orthogonal projection and orthogonal component?

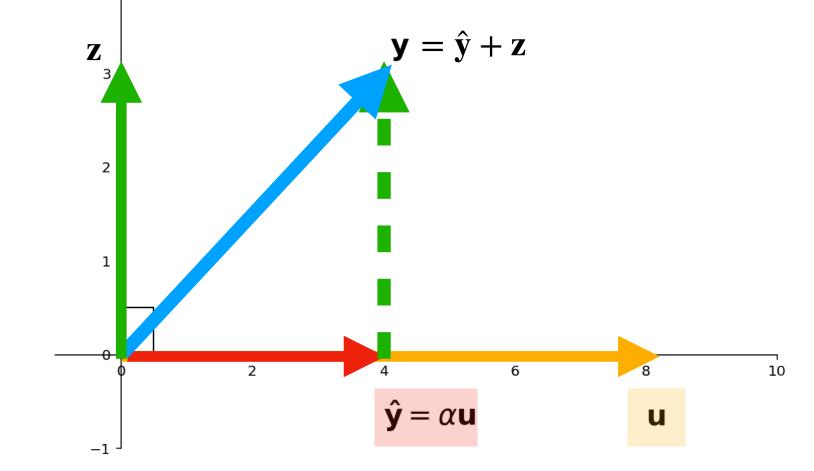




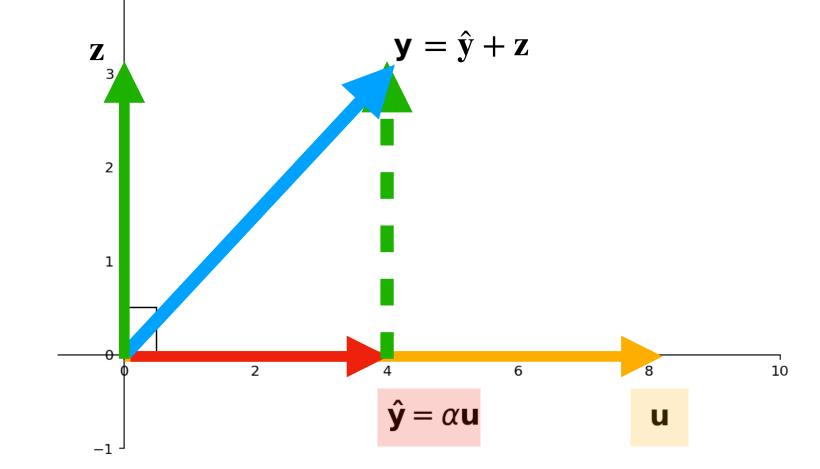
•  $\hat{\mathbf{y}} = \alpha \mathbf{u}$  for some scalar  $\alpha$  (since  $\hat{\mathbf{y}} \in span\{\mathbf{u}\}$ )



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



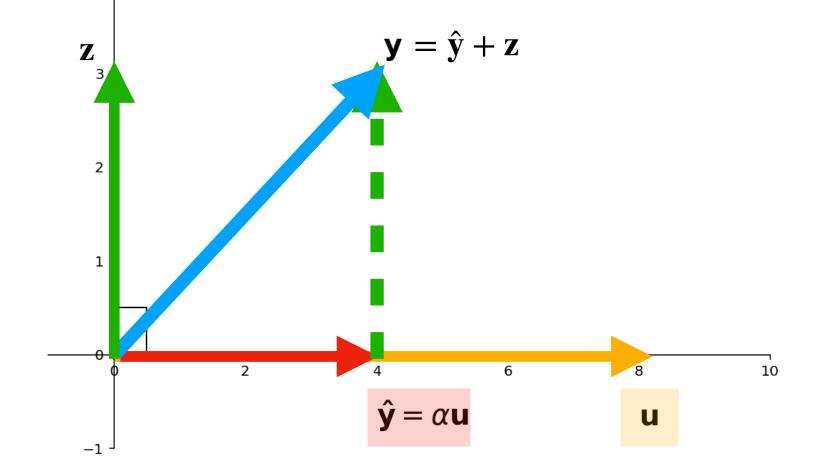
- $\hat{\mathbf{y}} = \alpha \mathbf{u}$  for some scalar  $\alpha$  (since  $\hat{\mathbf{y}} \in span\{\mathbf{u}\}$ )
- $\mathbf{z} = \mathbf{y} \hat{\mathbf{y}} = \mathbf{y} \alpha \mathbf{u}$  (since  $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ )
- $\langle \mathbf{z}, \mathbf{u} \rangle = 0$ 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



- $\hat{\mathbf{y}} = \alpha \mathbf{u}$  for some scalar  $\alpha$  (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:

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



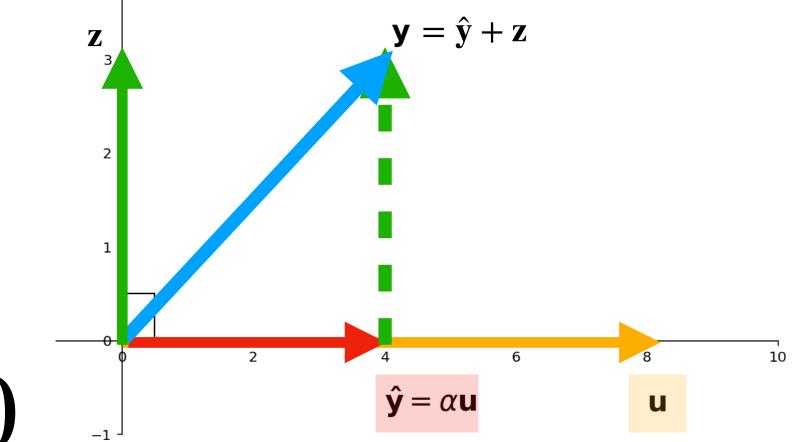
- $\hat{y} = \alpha u$  for some scalar  $\alpha$  (since  $\hat{y} \in span\{u\}$ )
- $\mathbf{z} = \mathbf{y} \hat{\mathbf{y}} = \mathbf{y} \alpha \mathbf{u}$  (since  $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ )
- $\langle \mathbf{z}, \mathbf{u} \rangle = 0$  (since  $\mathbf{z}$  is orthogonal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

#### Therefore:

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

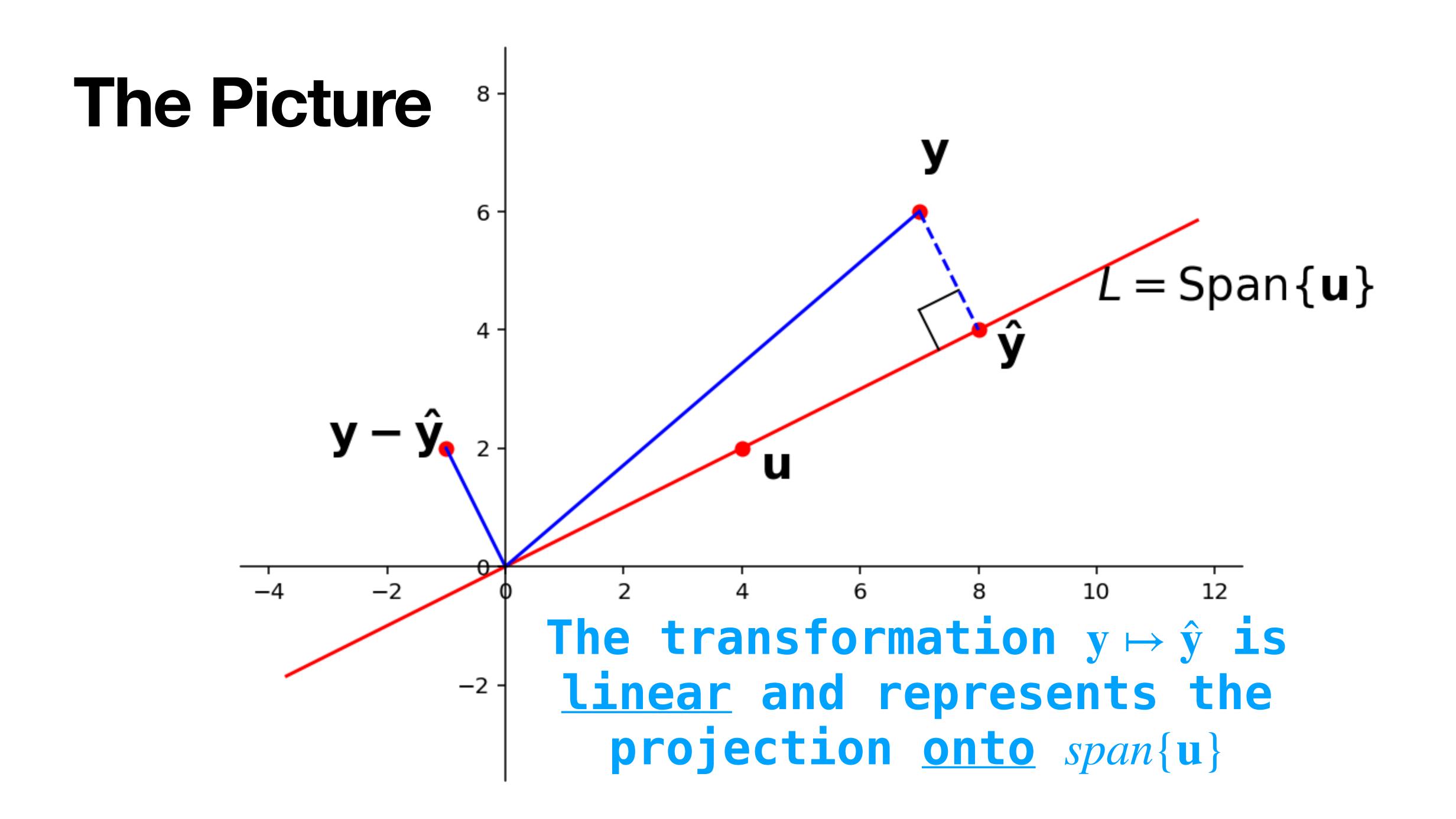
Once we have  $\alpha$ , we can compute both  $\hat{\mathbf{y}}$  and  $\mathbf{z}$ 

# Step 1: Finding $\alpha$



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

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

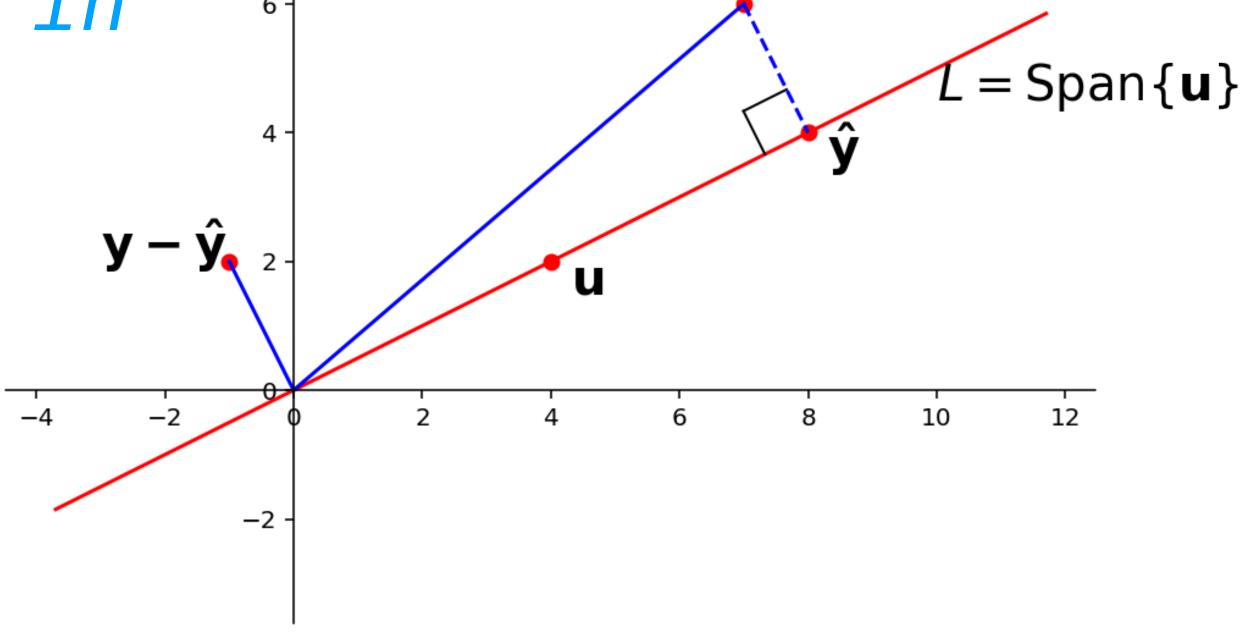


# ŷ and Distance

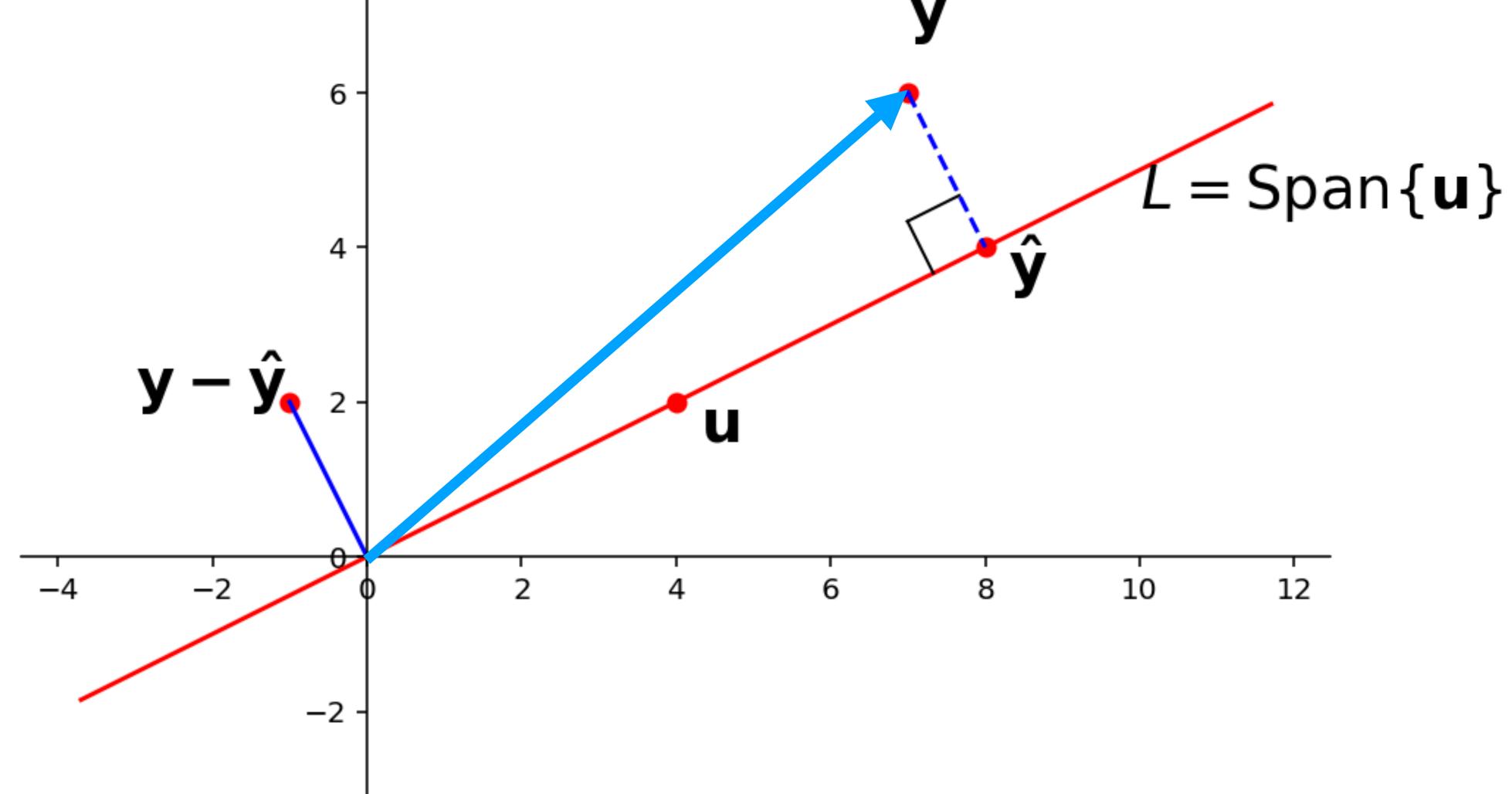
Theorem. 
$$\|\hat{\mathbf{y}} - \mathbf{y}\| = \min_{\mathbf{w} \in span\{\mathbf{u}\}} \|\mathbf{w} - \mathbf{y}\|$$

ŷ is the <u>closest</u> vector in span{u} to y.

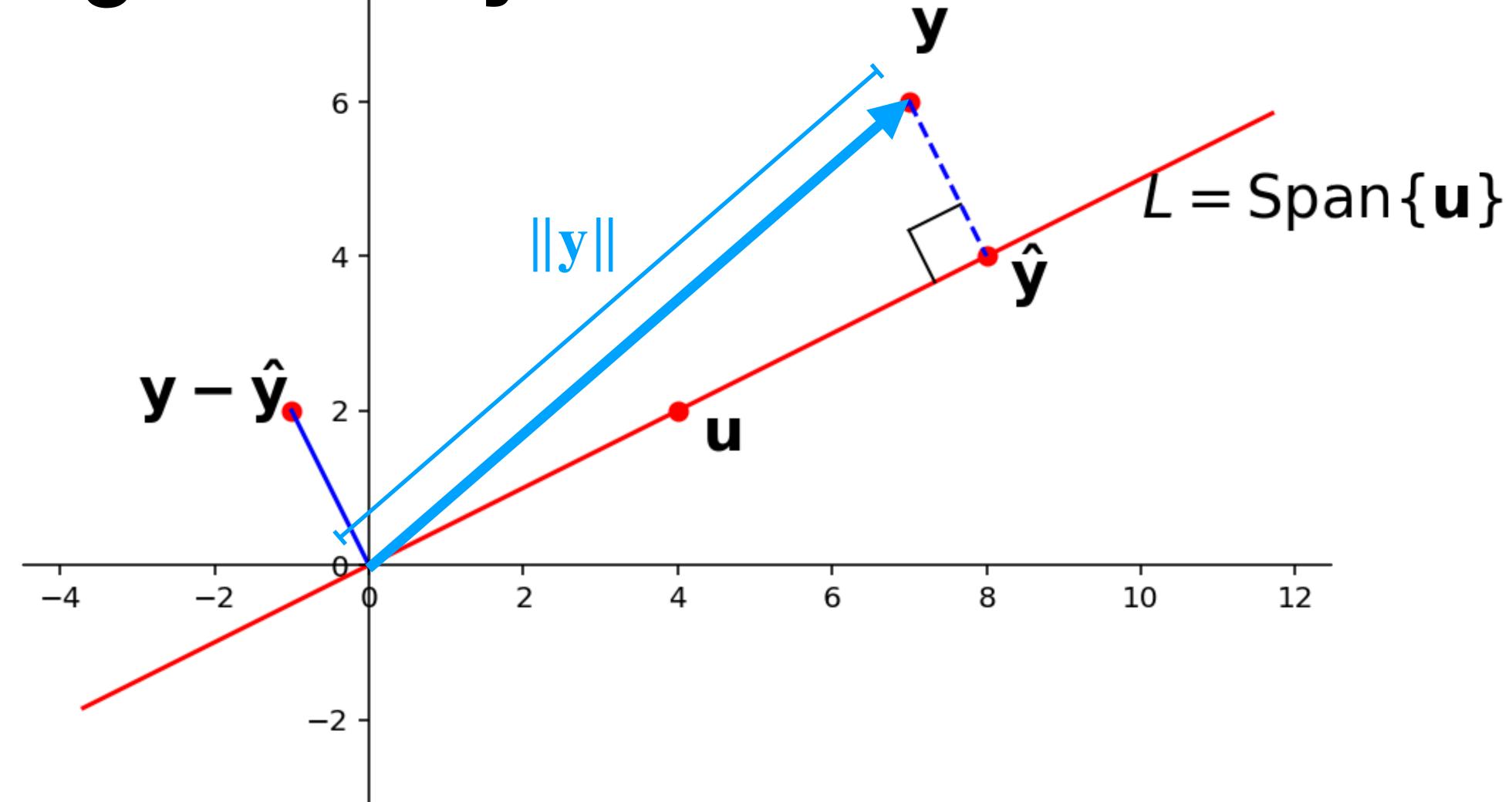
"Proof" by inspection:

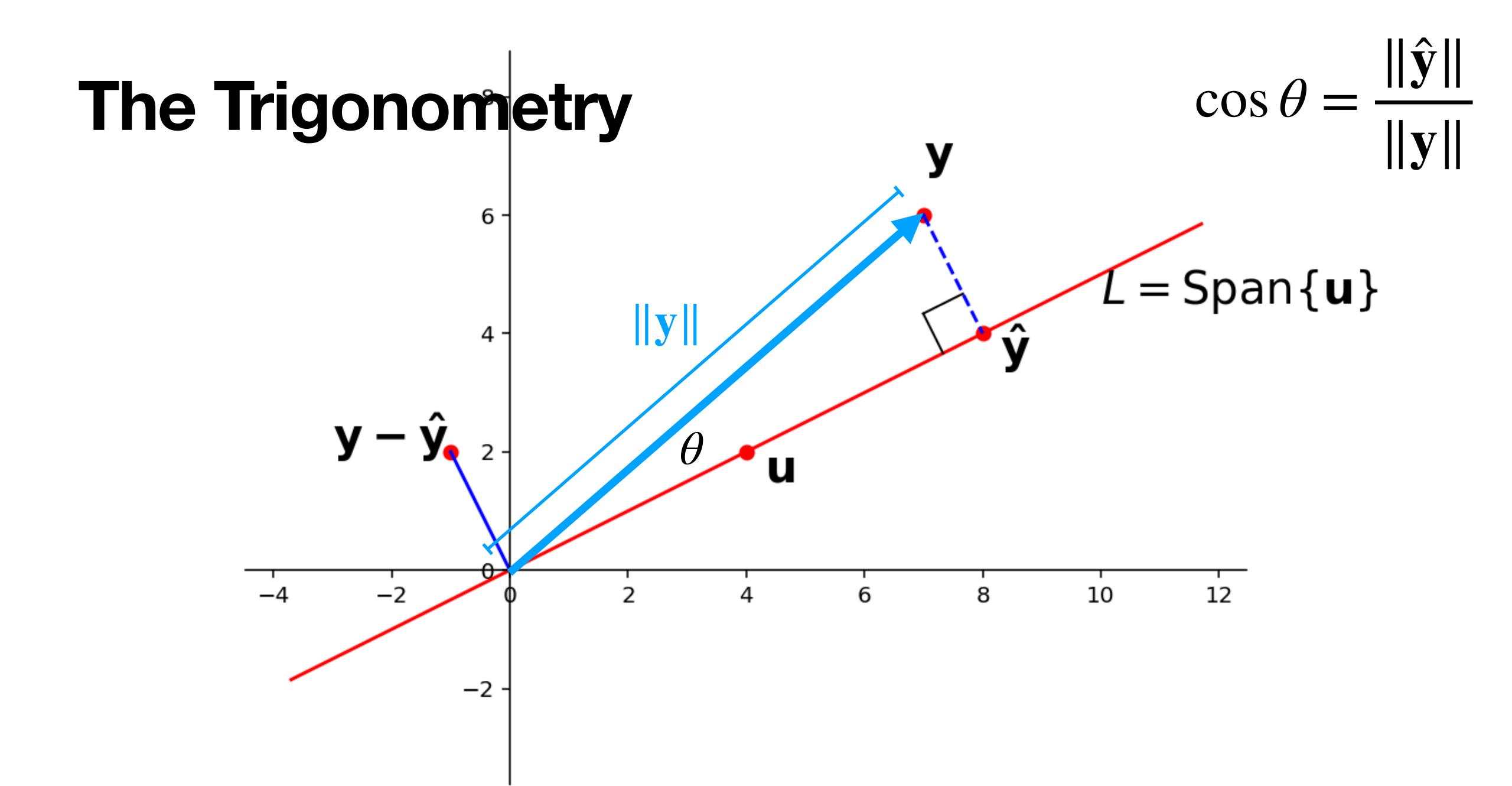


# The Trigonometry



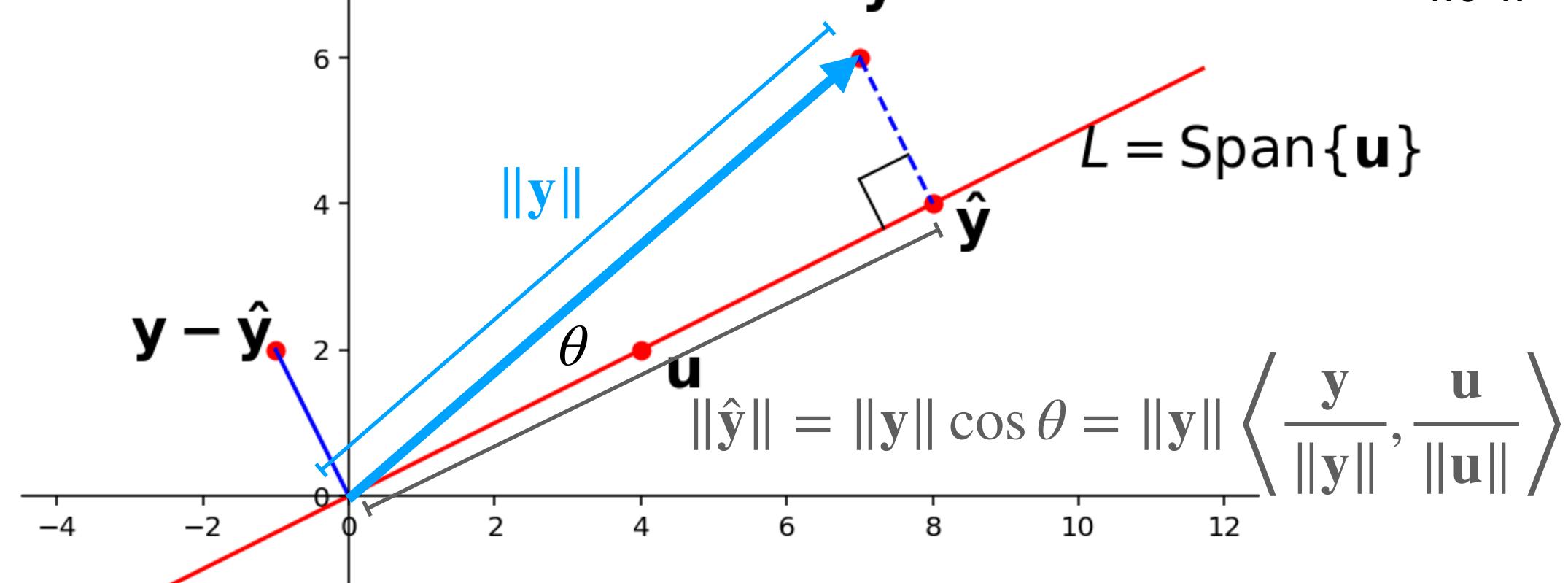
### The Trigonometry





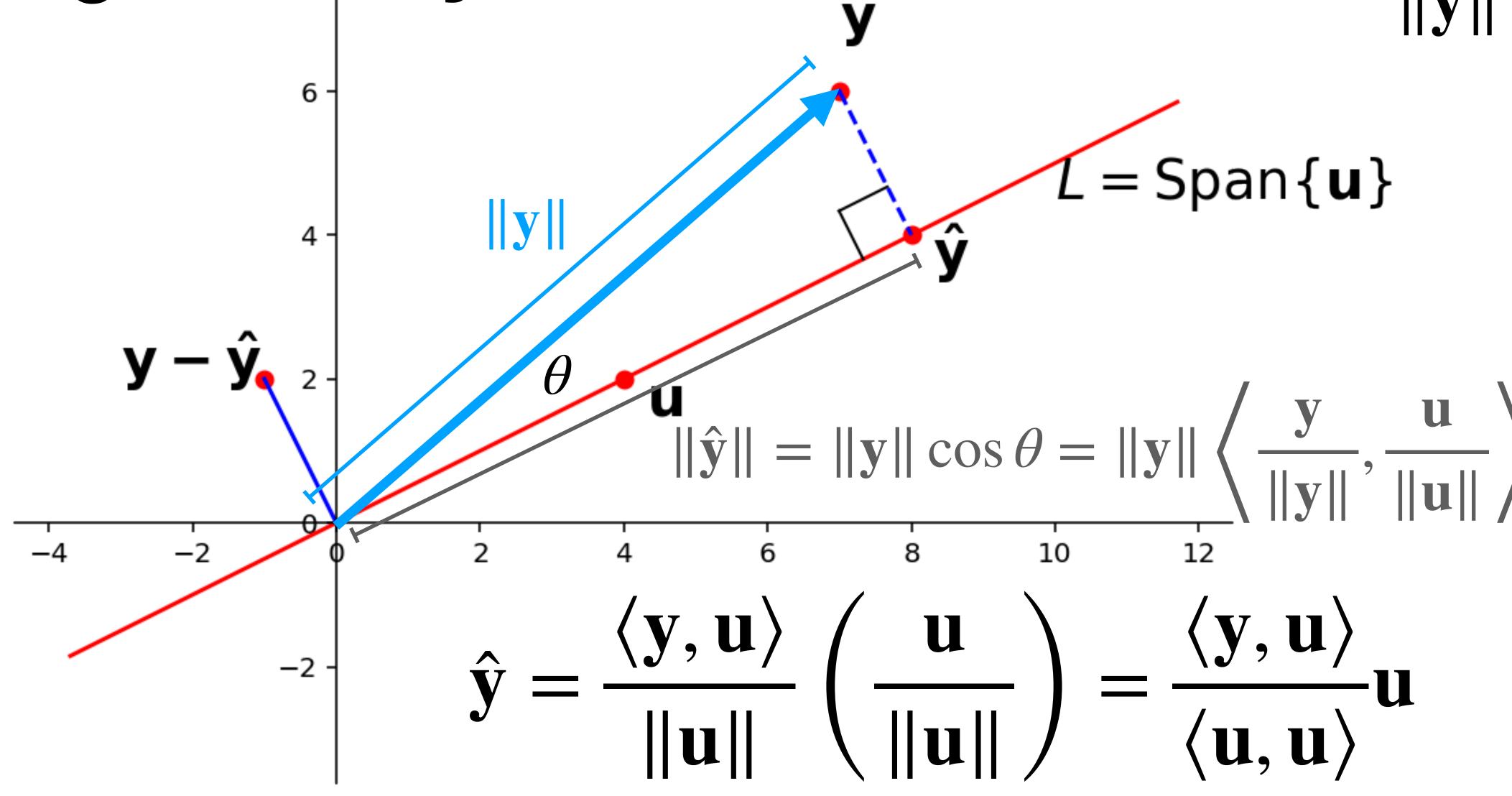
## The Trigonometry

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

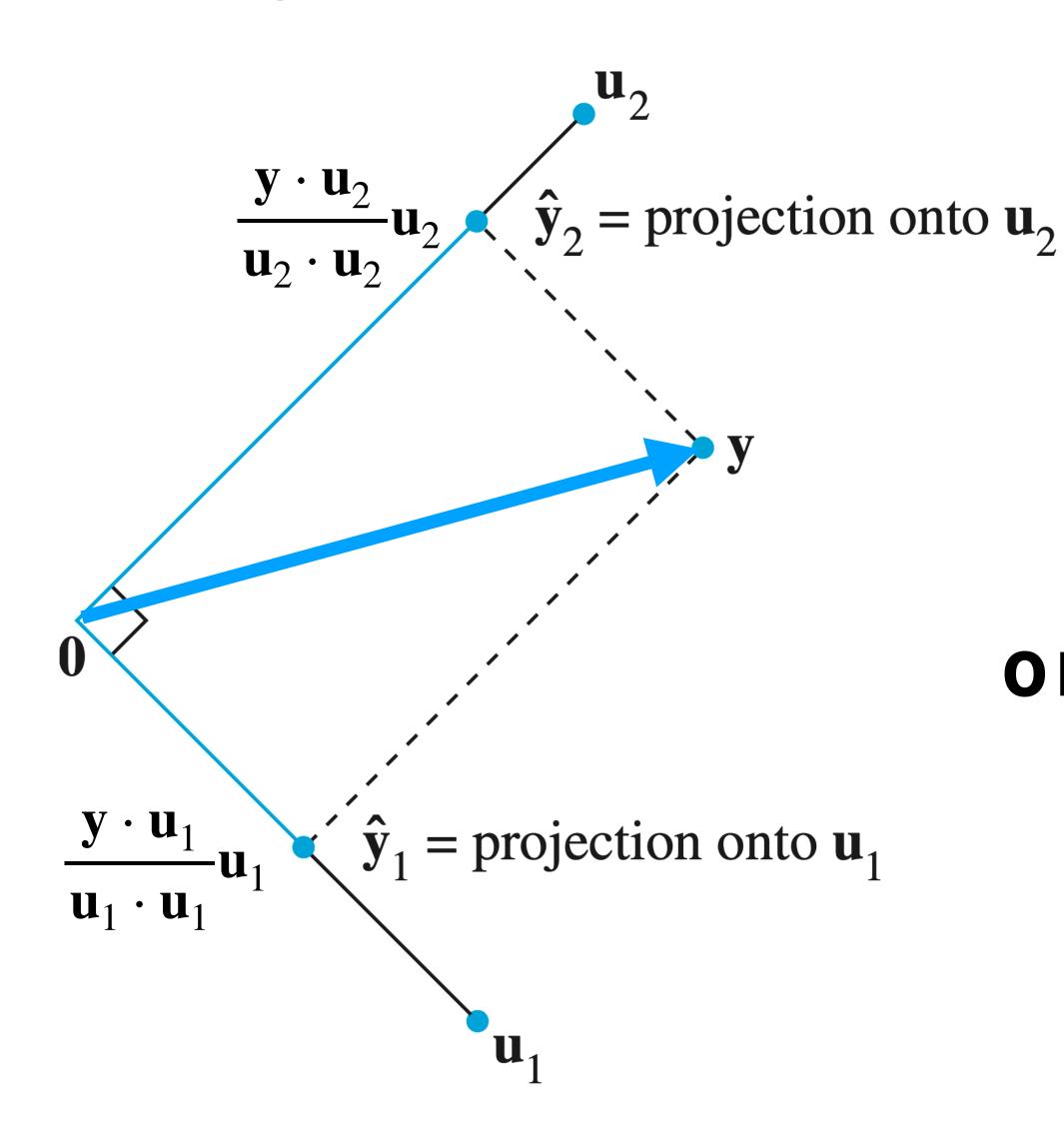


### The Trigonometry

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



#### Orthogonal Projections and Orthogonal Bases



Each <u>component</u> of y written in terms of an orthogonal basis is an orthogonal projection onto to a basis vector

#### How To:

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

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

#### Question

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 <a href="mailto:unit">unit</a> vector

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

**Definition.** A set  $\{u_1, u_2, ..., u_p\}$  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.

**Definition.** A set  $\{u_1, u_2, ..., u_p\}$  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

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

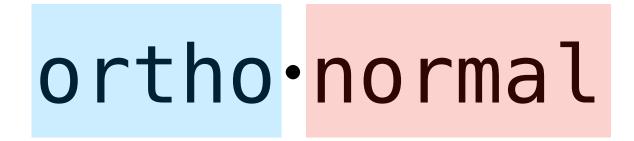
**Definition.** An **orthonormal basis** of the subspace W is a basis of W which is an orthonormal set.

ortho•normal

orthogonal/perpendicular

**Definition.** A set  $\{u_1, u_2, ..., u_p\}$  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.



#### 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

$$U^TU=I_m$$

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:

#### 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.

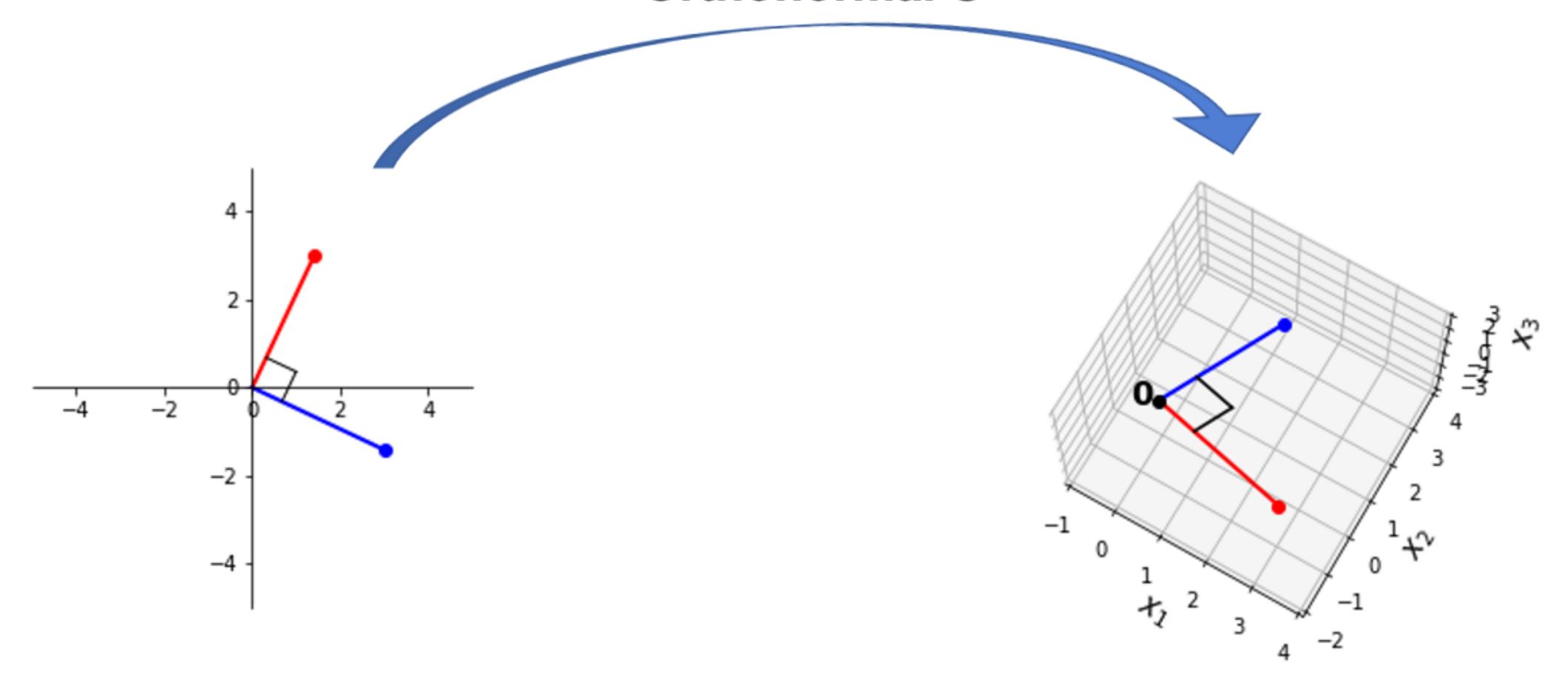
Verify:

#### Length, Angle, Orthogonality Preservation

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

#### The Picture

#### **Orthonormal U**



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

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

#### Question (Conceptual)

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

#### Answer

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

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

#### Summary

Orthogonal sets allow for <u>simpler calculations</u> of coordinates.

Finding these coordinates is a really about find the <u>orthogonal projections</u> onto each vector in the orthogonal set.

We can apply these ideas to matrices and describe a class of very well behaved transformations via <u>orthonormal matrices</u>.