# PAL Machine Learning Workshop Week 1: Linear Algebra Recap

02/22/2023 18:00 Future Technologies Lab

# Contents

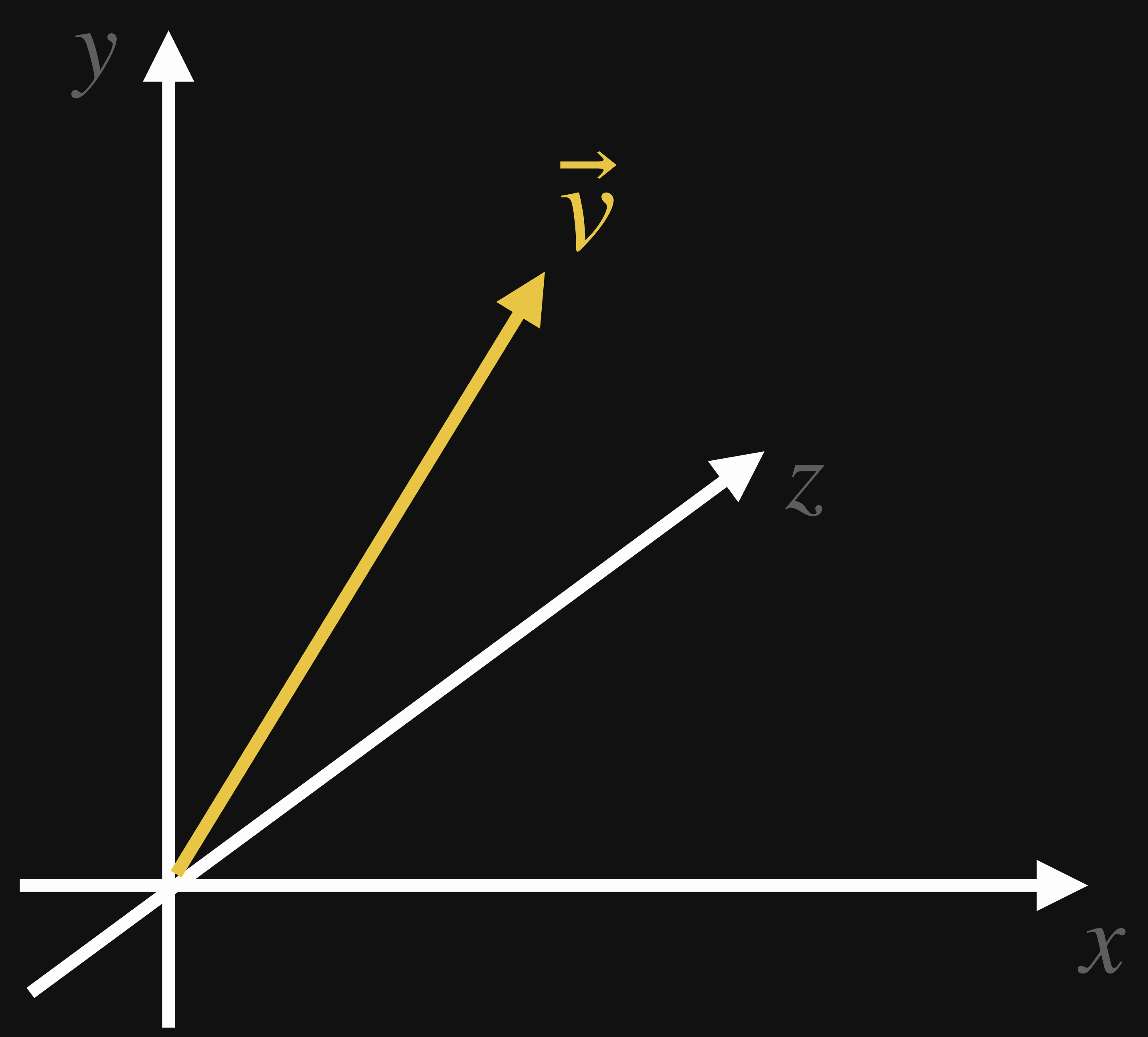
- Vectors
- Matrices
- Space Transformations
- Determinants and Singularity
- Inverses
- Dot Products

# Why study Linear Algebra?

- Provides a mathematical framework for understanding many of the concepts used in machine learning algorithms.
- Allows for the efficient representation and manipulation of large amounts of data.
- Plays a key role in optimization techniques used to train many machine learning models.

# Vectors

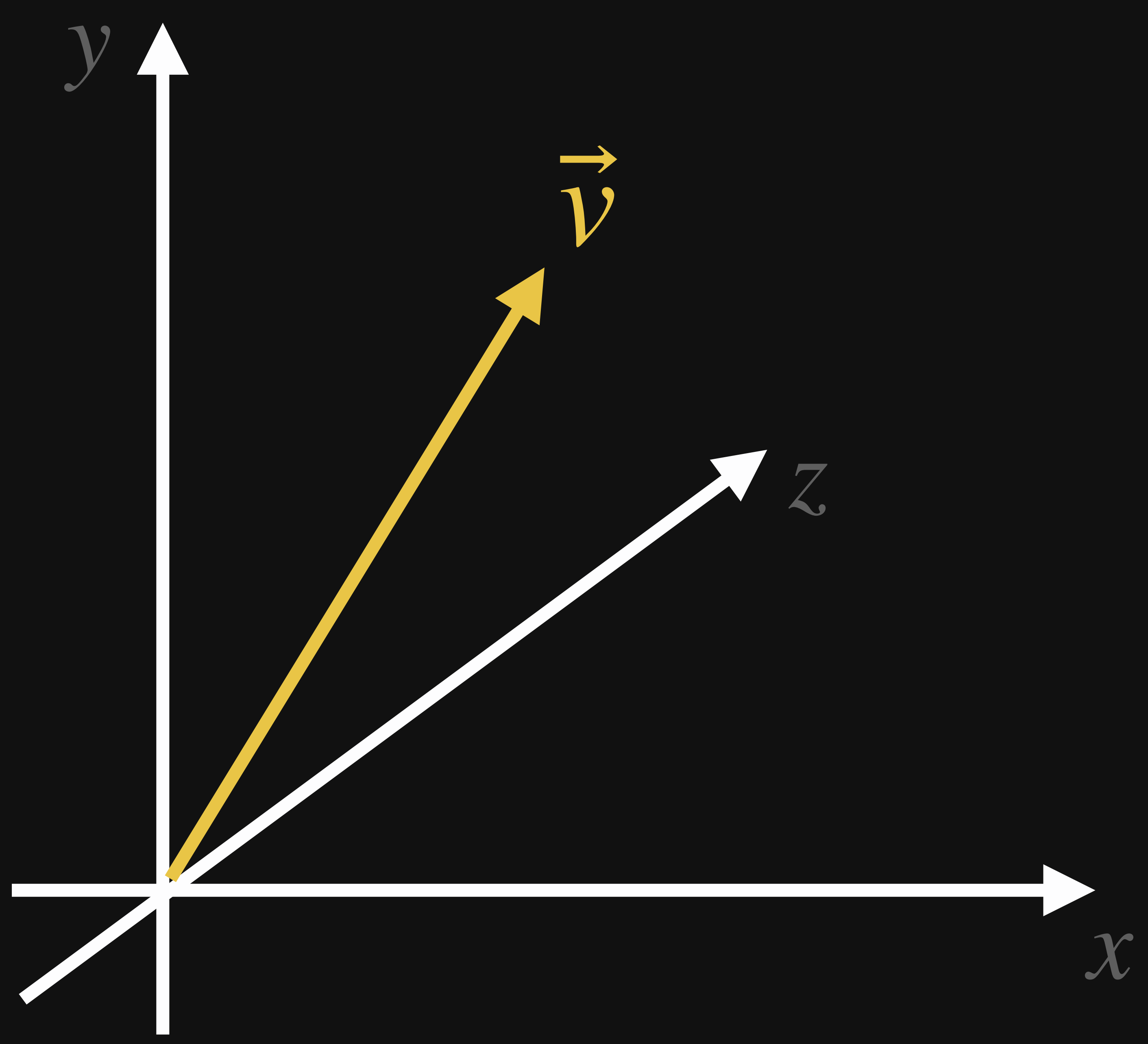
- Vector can be represented as a list of numbers indicating the coordinates of a point in space.



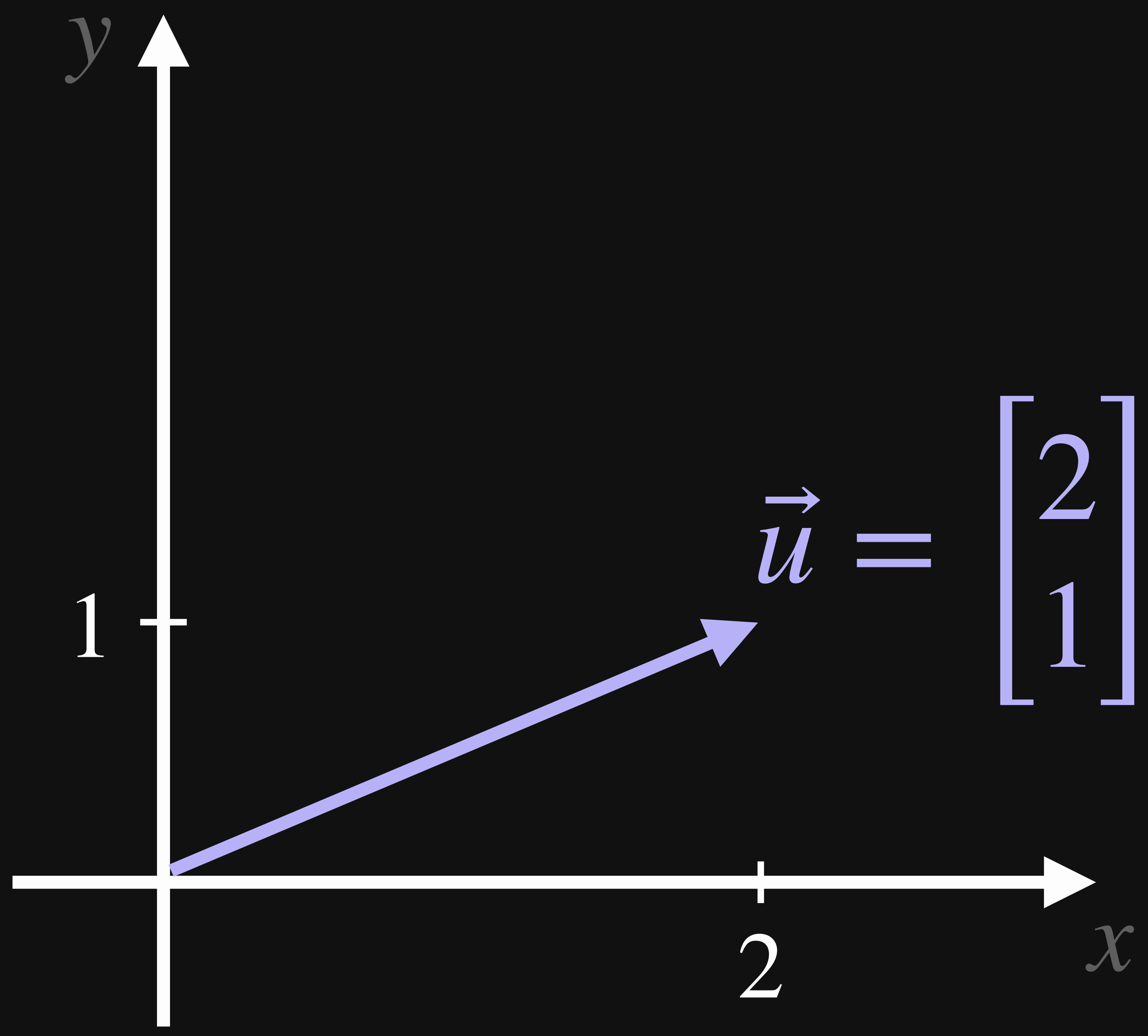

$$\vec{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

# Vectors

- Vector can be represented as a list of numbers indicating the coordinates of a point in space.



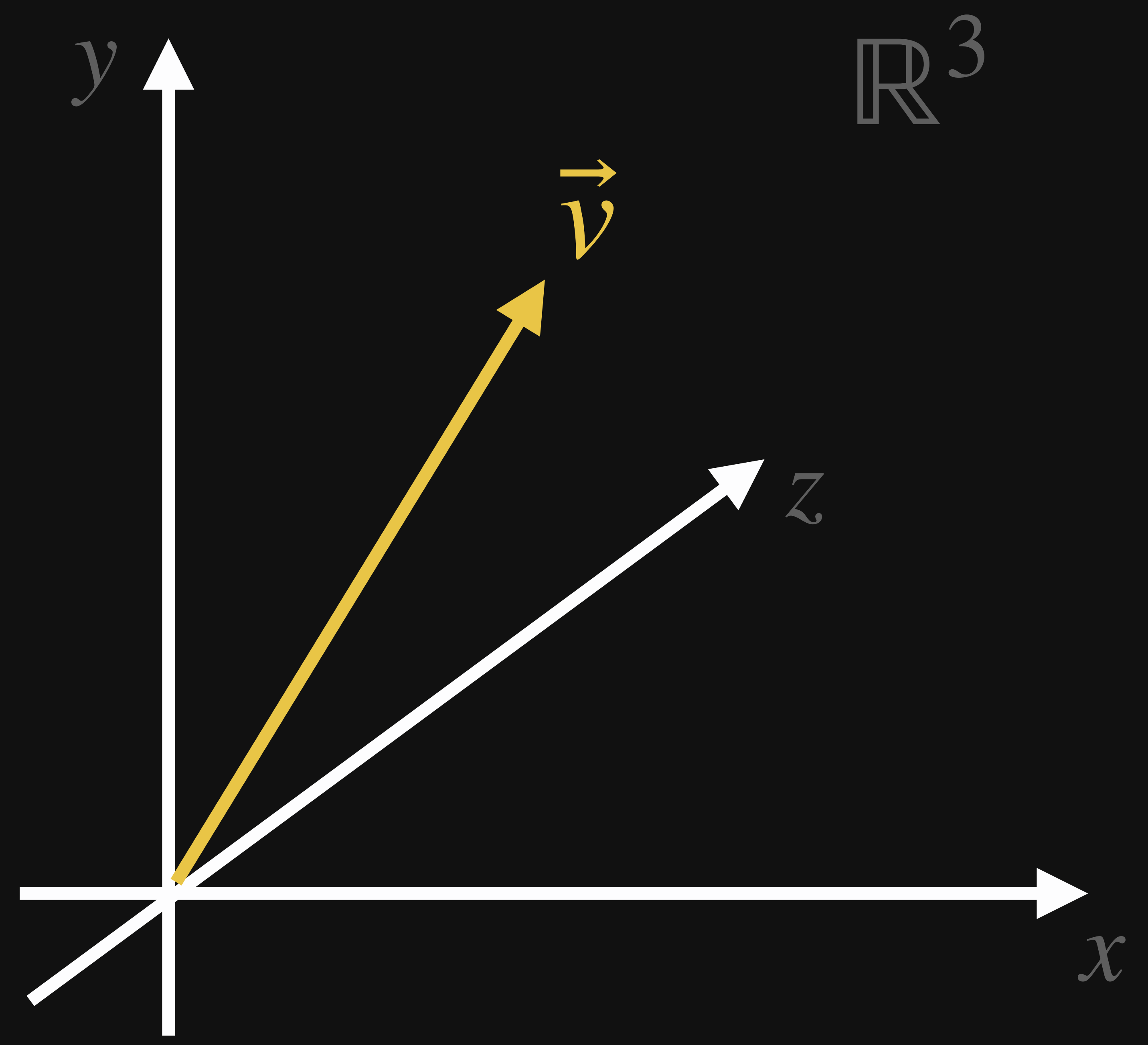

$$\vec{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$




$$\vec{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

# Vectors

- Mathematically, we define the dimensionality of a vector as $v \in \mathbb{R}^d$ , which means that $v$ is an element of a $d$-dimensional real space.

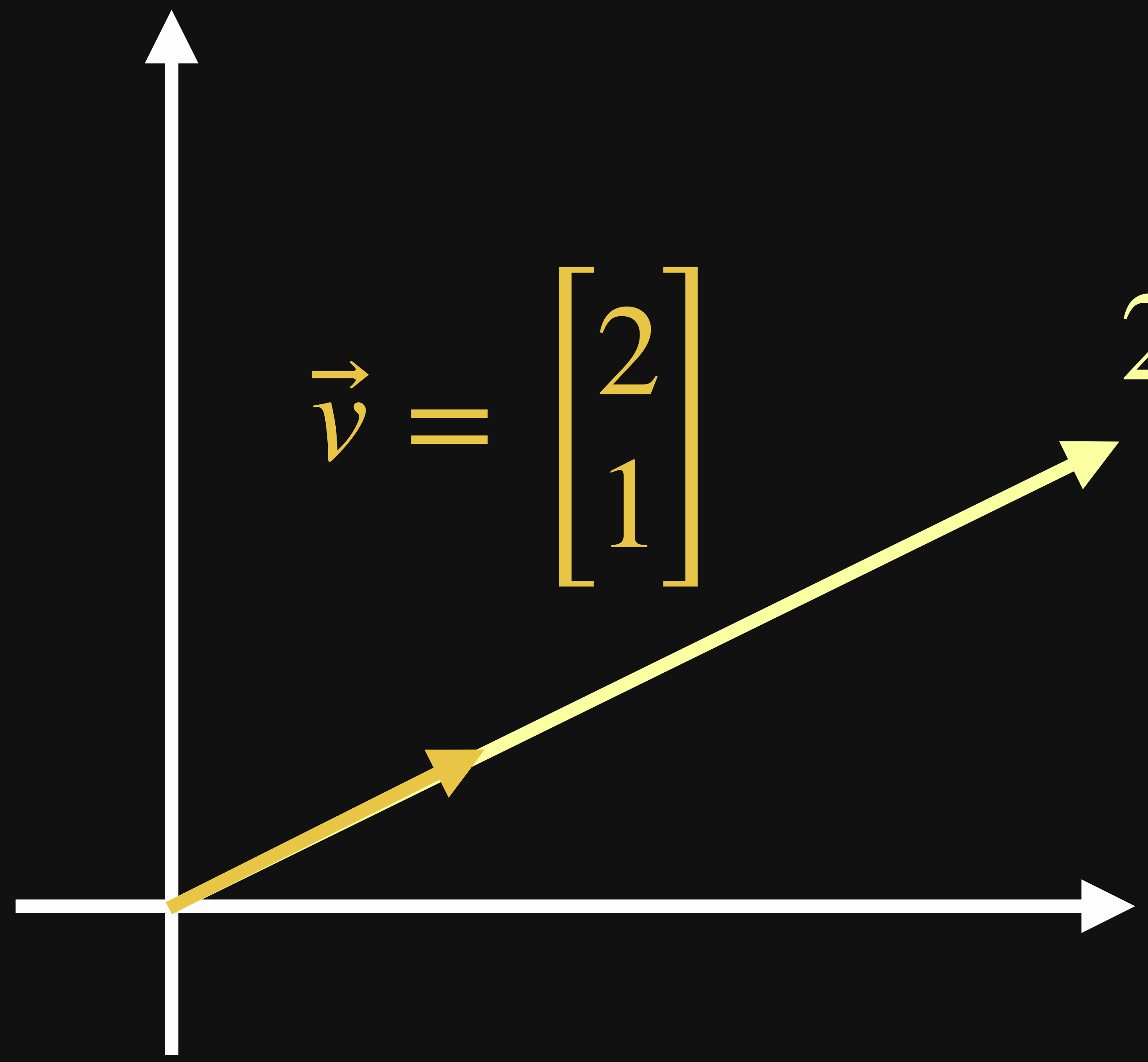



$$\vec{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

# Vectors

## Scalar Multiplication

- Multiplying vector by a scalar means multiplying each of its components by that scalar.



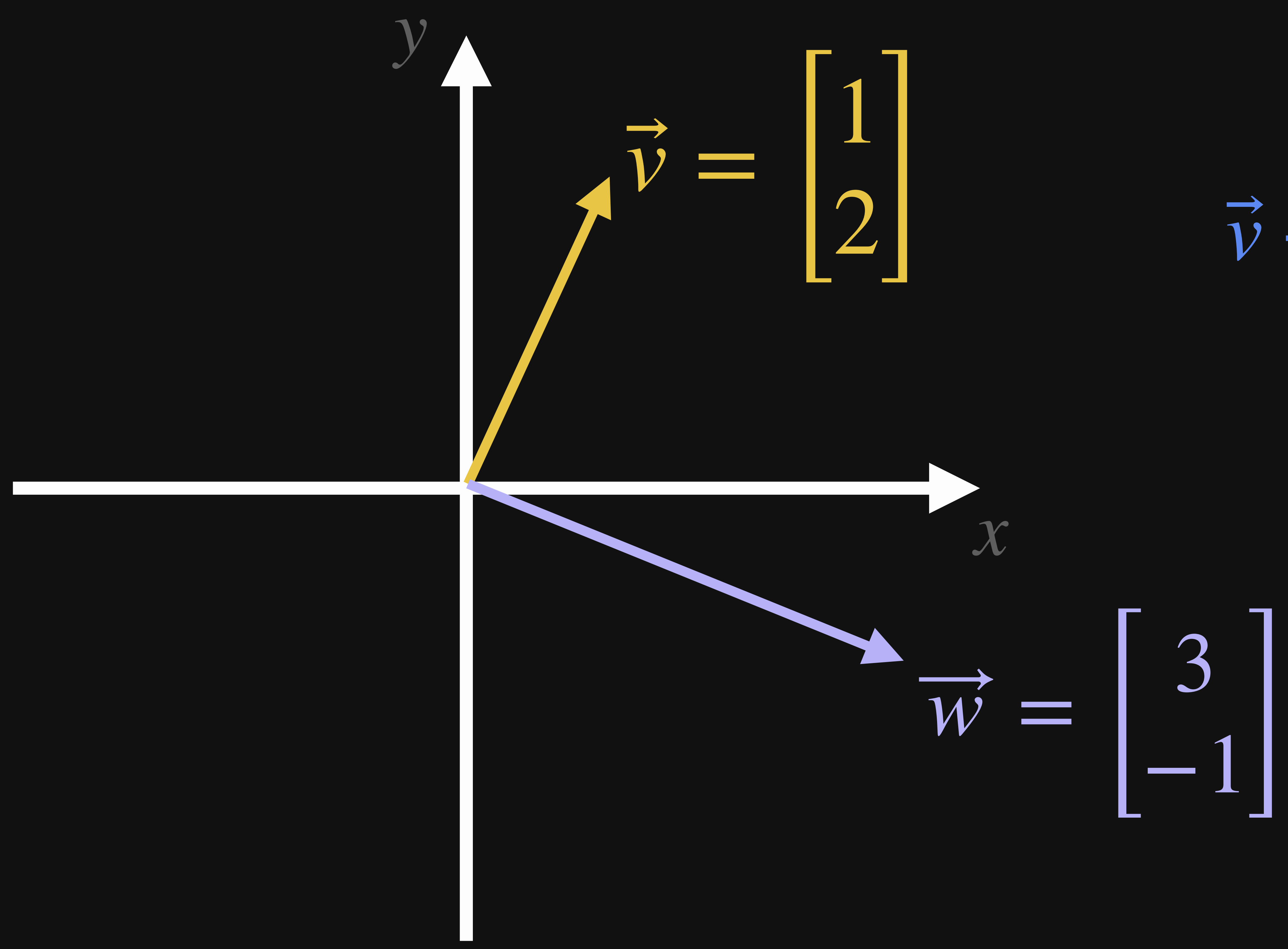
$$\vec{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$2\vec{v} = 2\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

# Vectors

## Addition

- Vectors are added component-wise.

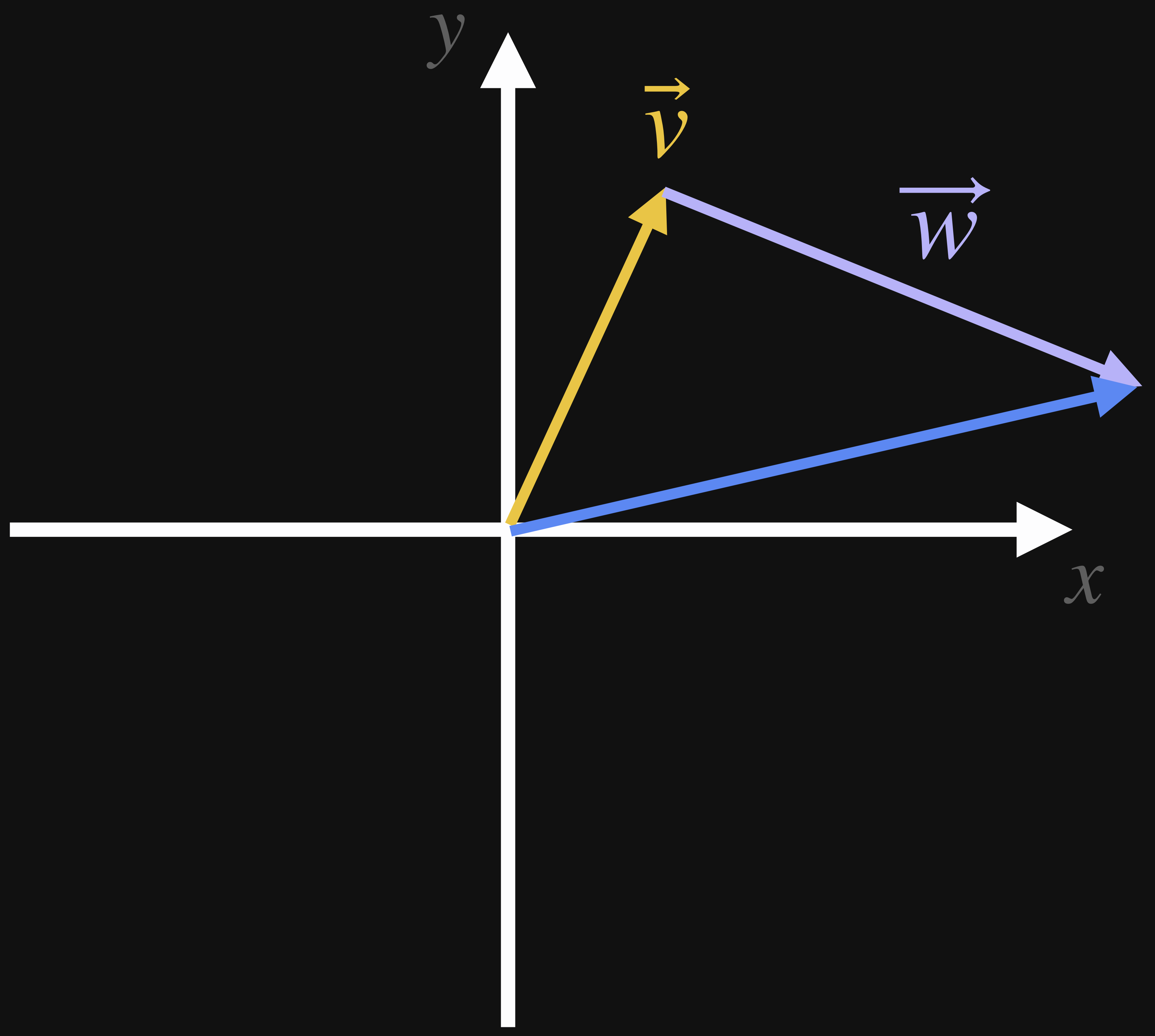



$$\vec{v} + \vec{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

# Vectors

## Addition

- Geometrically, this can be represented as placing the tail of the second vector at the tip of the first vector and drawing the third vector from the tail of the first vector to the tip of the second vector.




$$\vec{v} + \overrightarrow{w} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

# Vectors

## Span

- The *linear combination* of two vectors is defined as a scalar multiple of one vector added to a scalar multiple of the other vector.

$$a\vec{v} + b\vec{w}$$

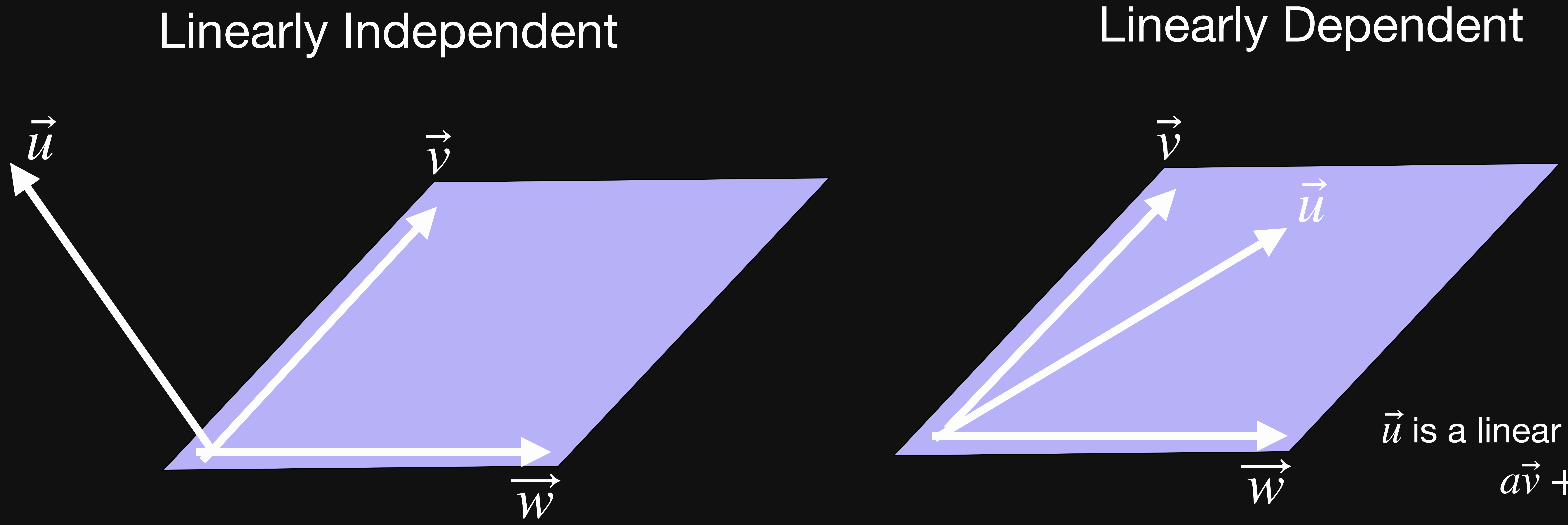
Scalars

- The *span of* $\vec{v}$ and $\vec{w}$ is the set of all their linear combinations.

- If two vectors are *linearly independent* then their span will cover the entire 2D plane.

# Vectors

## Linear Independence in 3D




$\vec{u}$ is a linear combination
$a\vec{v} + b\vec{w}$

# Matrices

- We denote the size of a matrix as: $A \in \mathbb{R}^{m \times n}$.
- You can think of it simply as a two dimensional array of numbers.

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \\ a_{m1} & & a_{mn} \end{bmatrix}$$

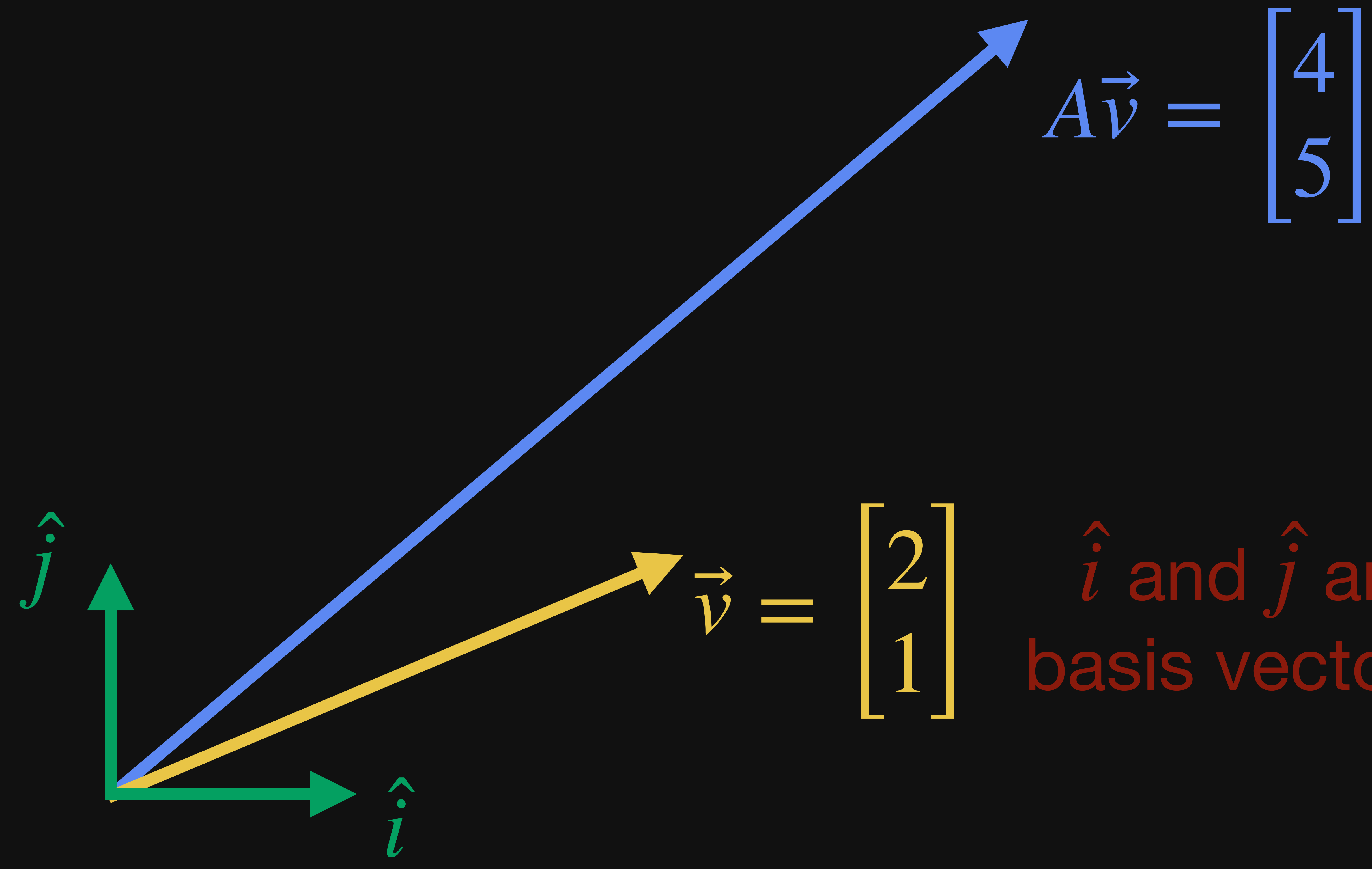
# Matrices

## Multiplication

- To multiply a matrix by a vector, the number of columns in the matrix must be equal to the number of rows in the vector. The result of the multiplication is a new vector.

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \qquad \vec{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$A\vec{v} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$




$\hat{i}$ and $\hat{j}$ are basis vectors!

# Matrices

## Transformations

- Transformation can be though of as a *function* that takes some input (e.g. coordinates of a vector) and returns some output (e.g. where that vector lands).

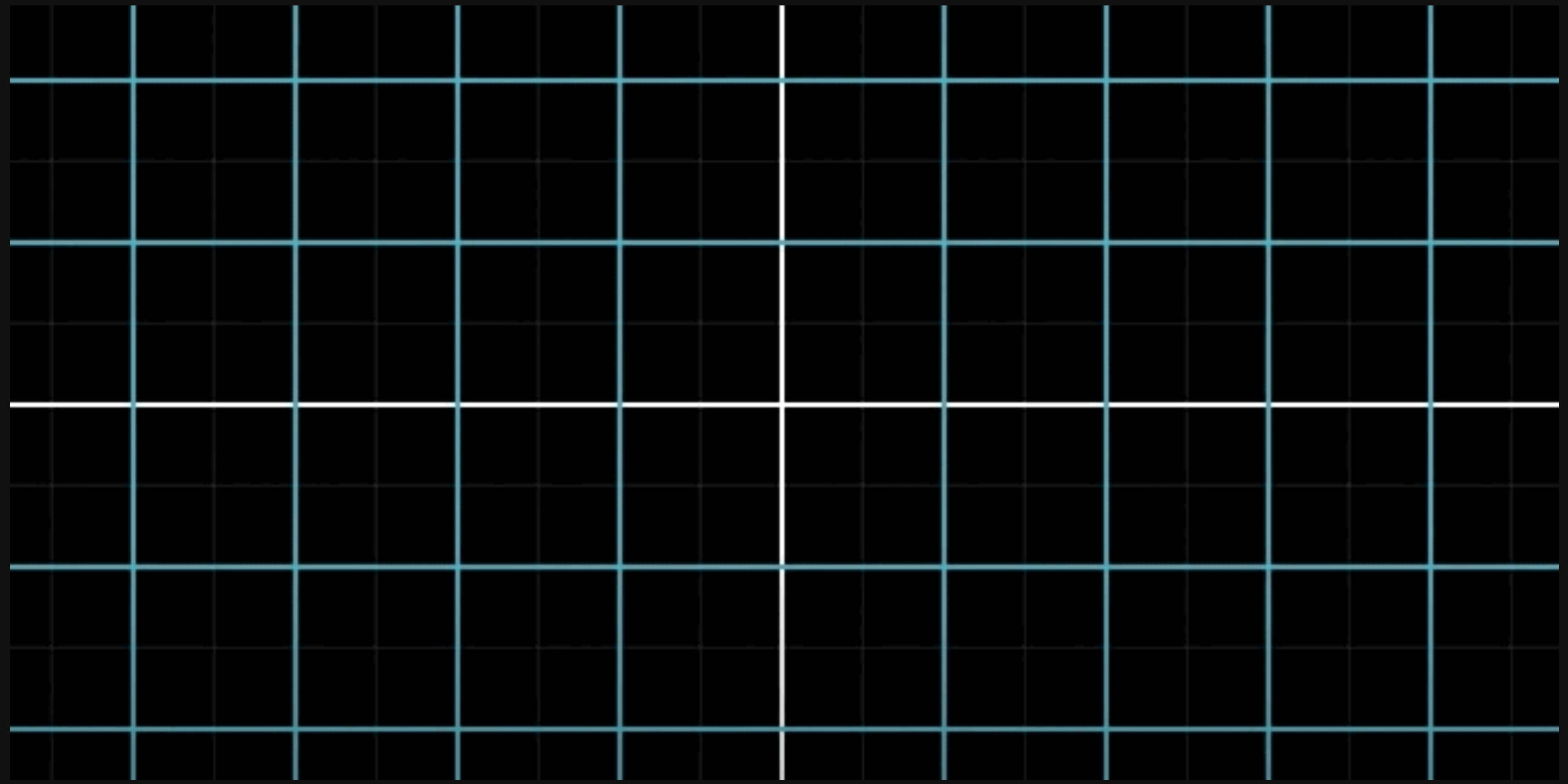
$$\begin{bmatrix} x_{in} \\ y_{in} \end{bmatrix} \rightarrow ??? \rightarrow \begin{bmatrix} x_{out} \\ y_{out} \end{bmatrix}$$

# Matrices

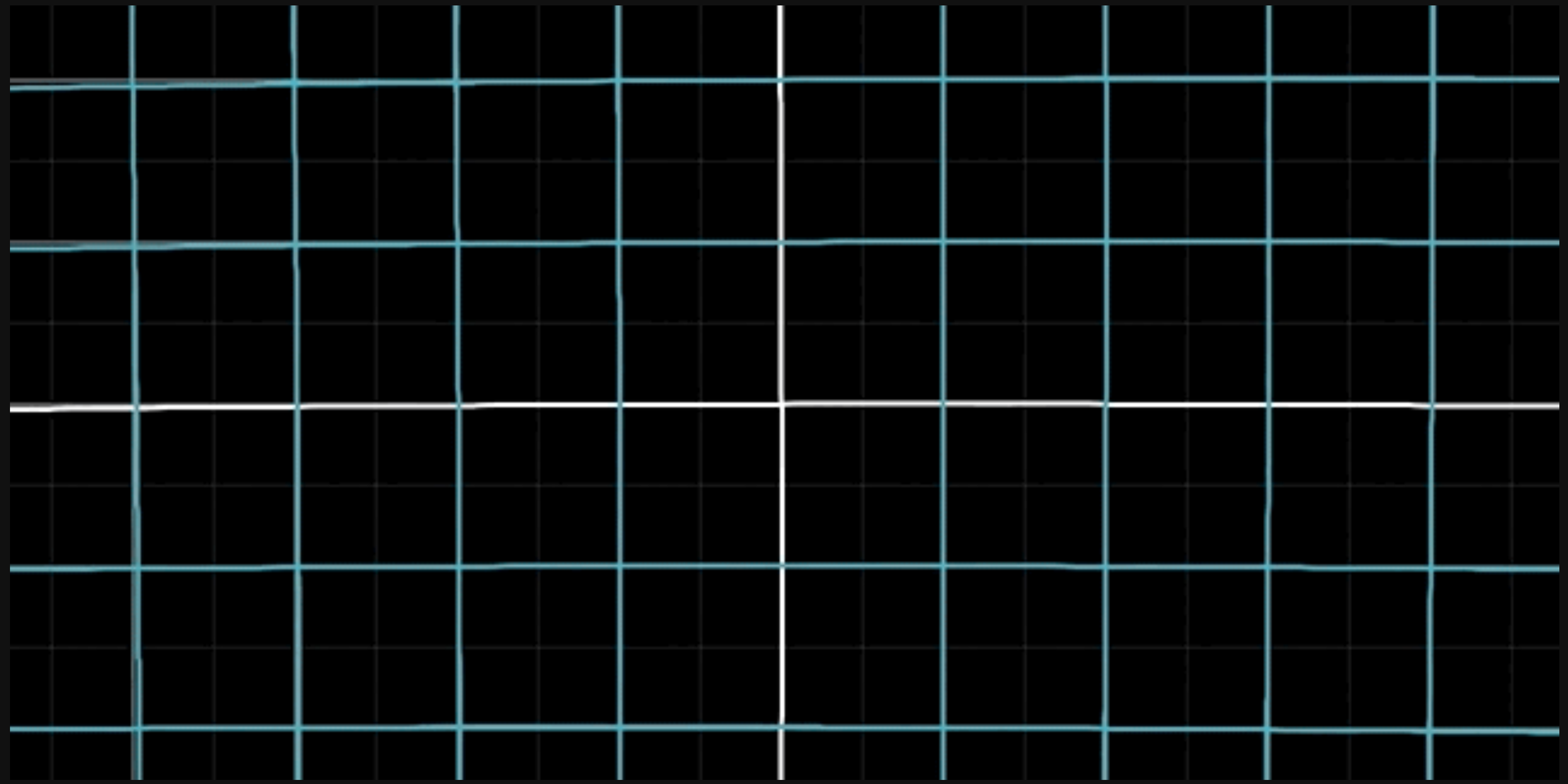
## Transformations

- A transformation is linear if:
    1. All lines remain parallel and evenly spaced
    2. Origin remains fixed

Linear Transformation



Non-Linear Transformation

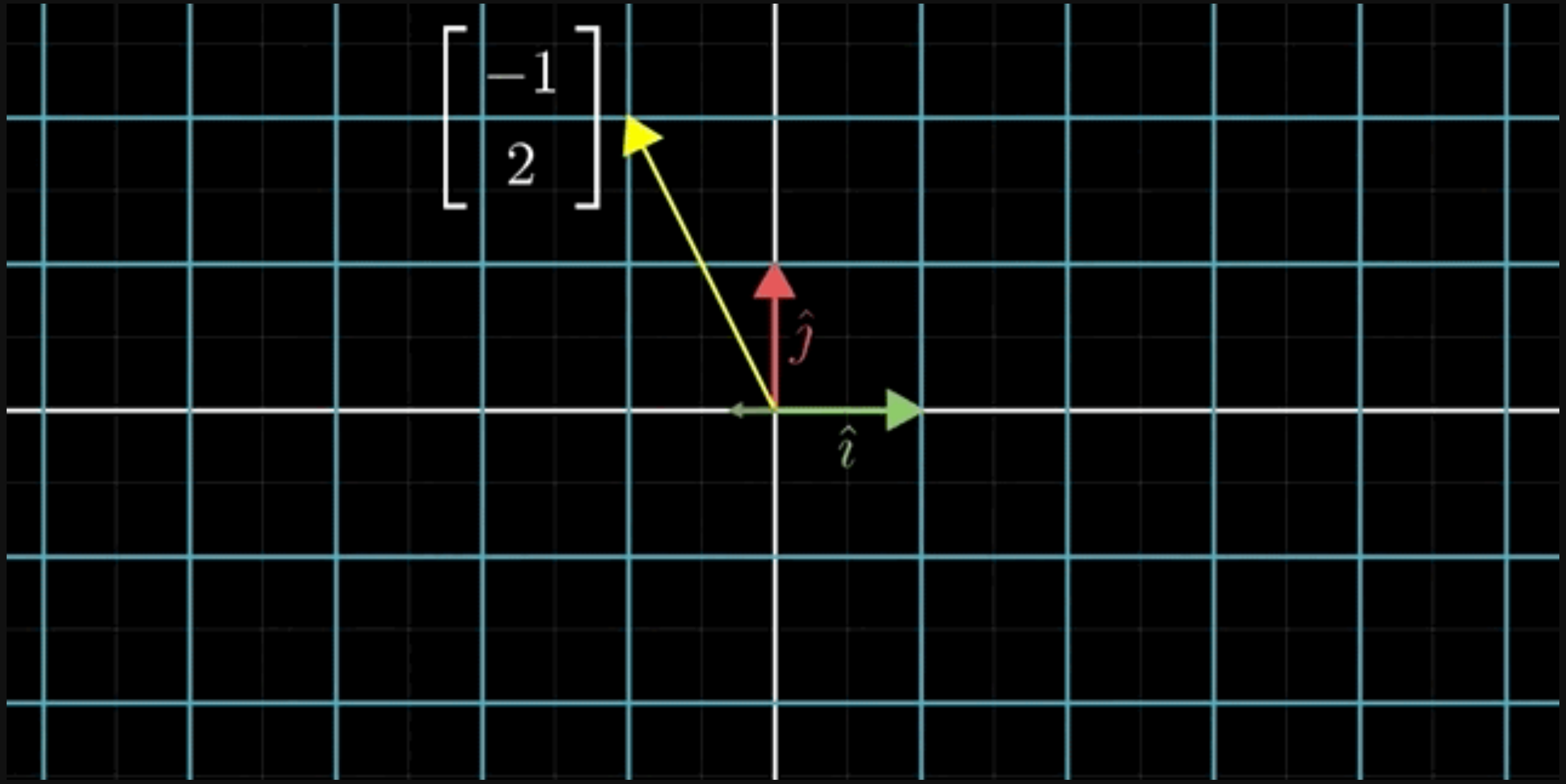


Source: YouTube 3Blue1Brown

# Matrices

## Transformations

- These properties allow us to determine where $\vec{v}$ lands if we know where $\hat{i}$ and $\hat{j}$ land.




Source: YouTube 3Blue1Brown

# Matrices

## Transformations

- These properties allow us to determine where $\vec{v}$ lands if we know where $\hat{i}$ and $\hat{j}$ land.

$$x\begin{bmatrix}3\\-2\end{bmatrix}+y\begin{bmatrix}2\\1\end{bmatrix}=\begin{bmatrix}3x+2y\\-2x+1y\end{bmatrix}$$

- This can also be described by a $2 \times 2$ matrix:

$$\begin{bmatrix}3 & 2\\-2 & 1\end{bmatrix}$$

Where $\hat{i}$ lands    Where $\hat{j}$ lands

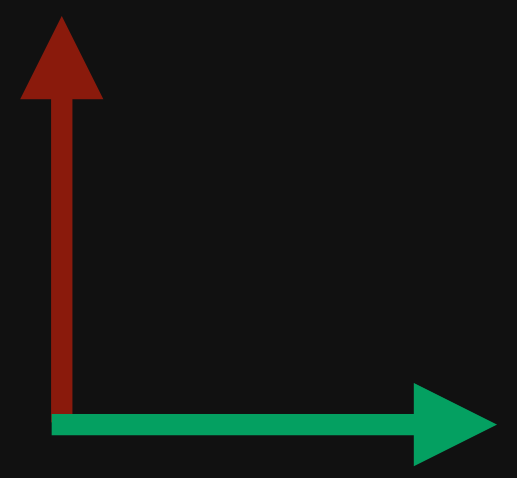
# Matrices

## Transformations

- What will this transformation do?

$$\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\hat{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



$$\hat{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

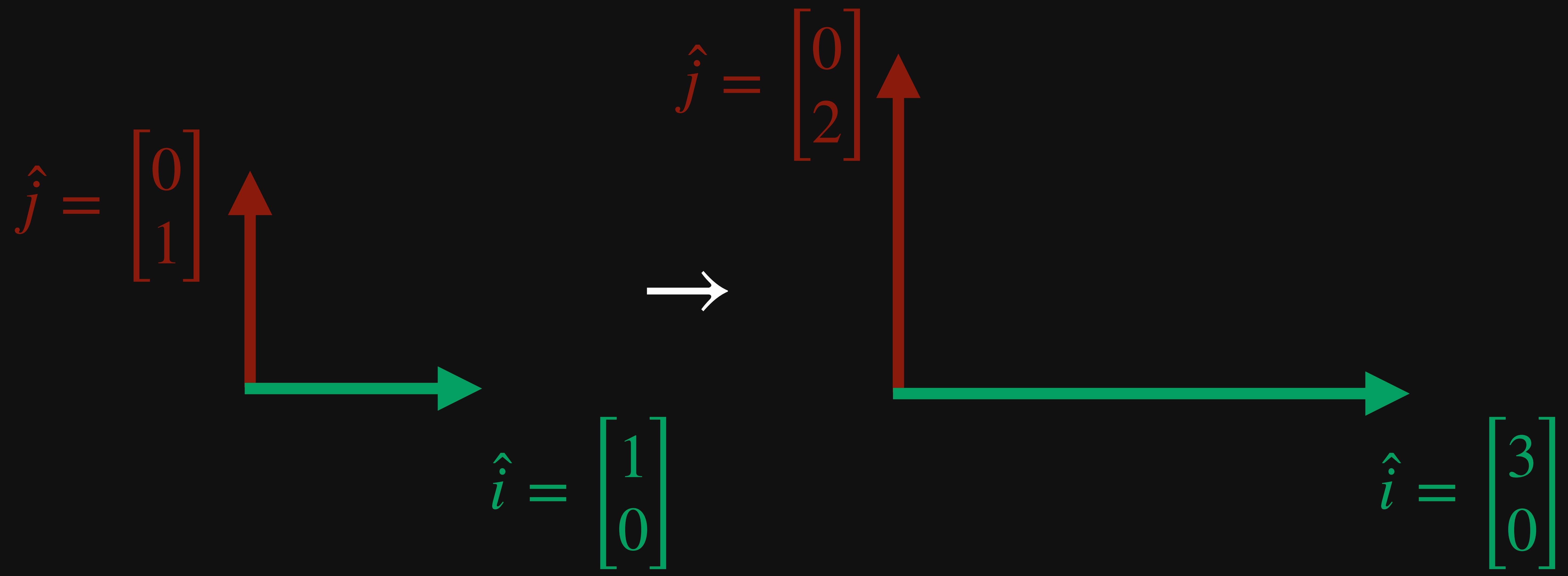
$$\rightarrow$$

# Matrices

## Transformations

- What will this transformation do?

$$\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

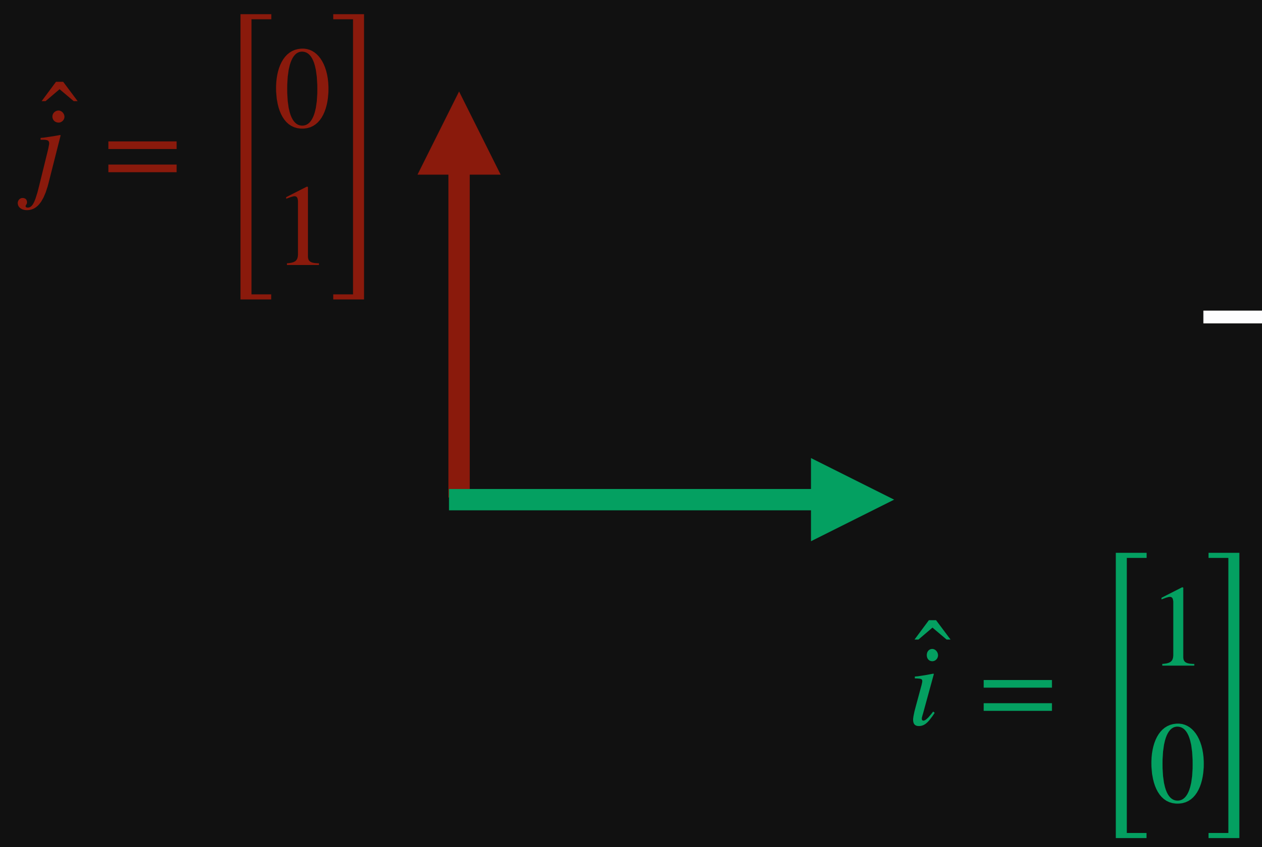


# Matrices

## Transformations

- What will this transformation do?

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$



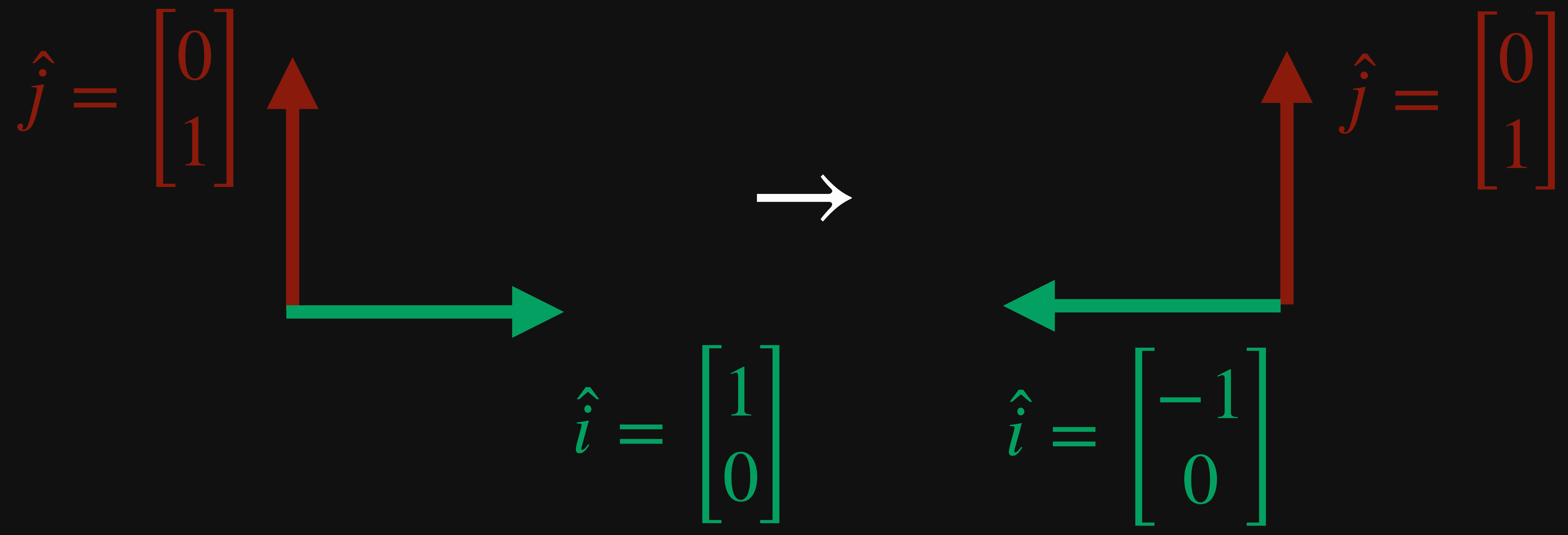

$\rightarrow$

# Matrices

## Transformations

- What will this transformation do?

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

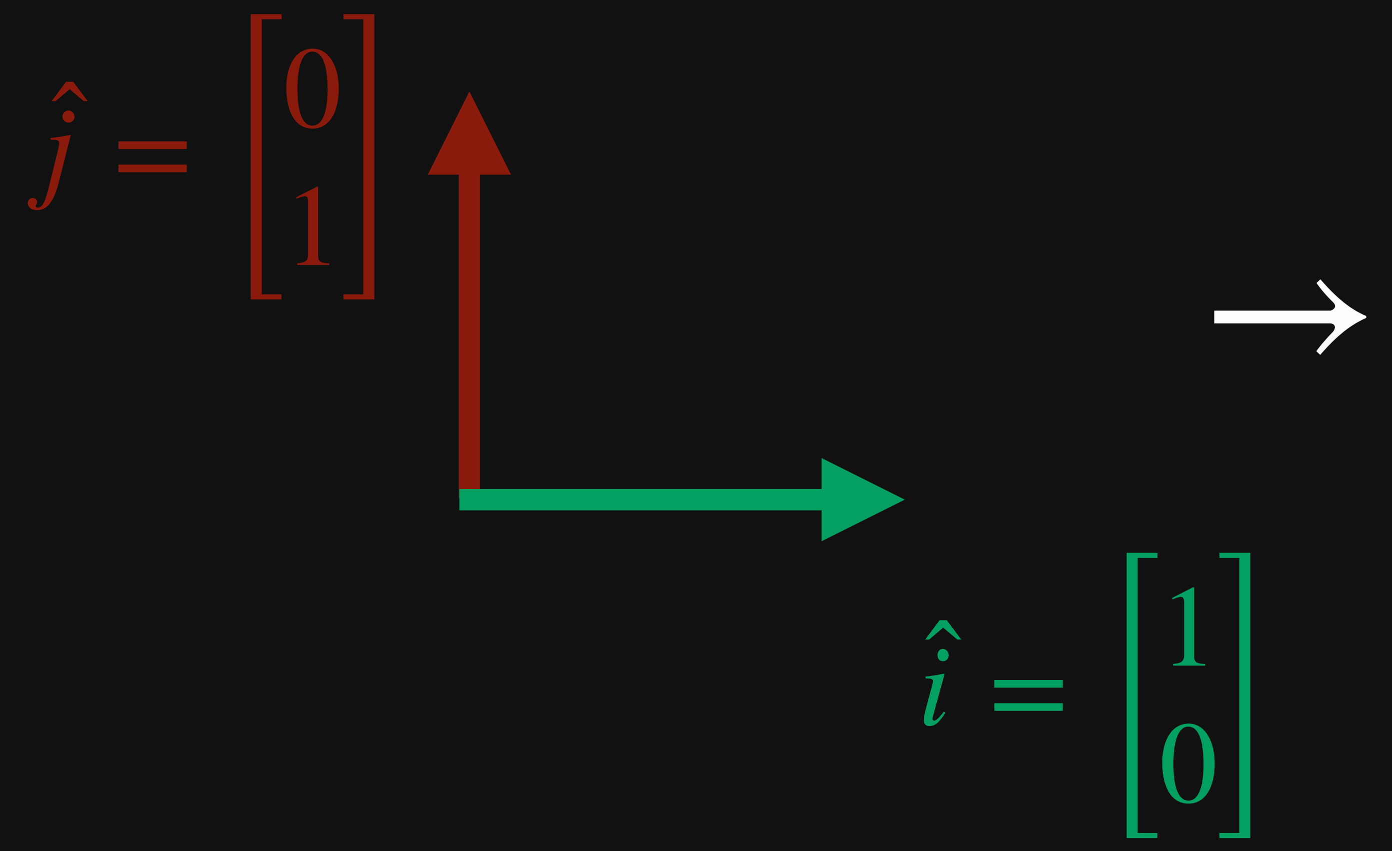


# Matrices

## Transformations

- What will this transformation do?

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

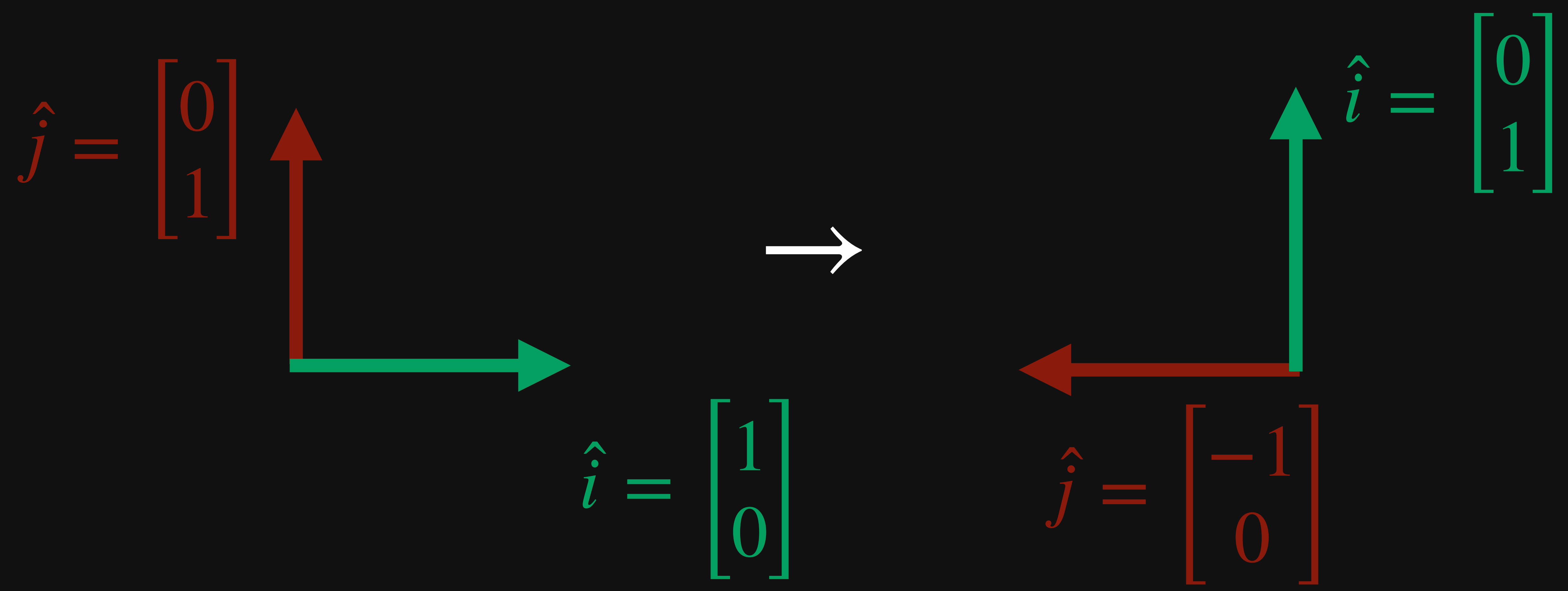


# Matrices

## Transformations

- What will this transformation do?

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$



# Matrices

## Composition

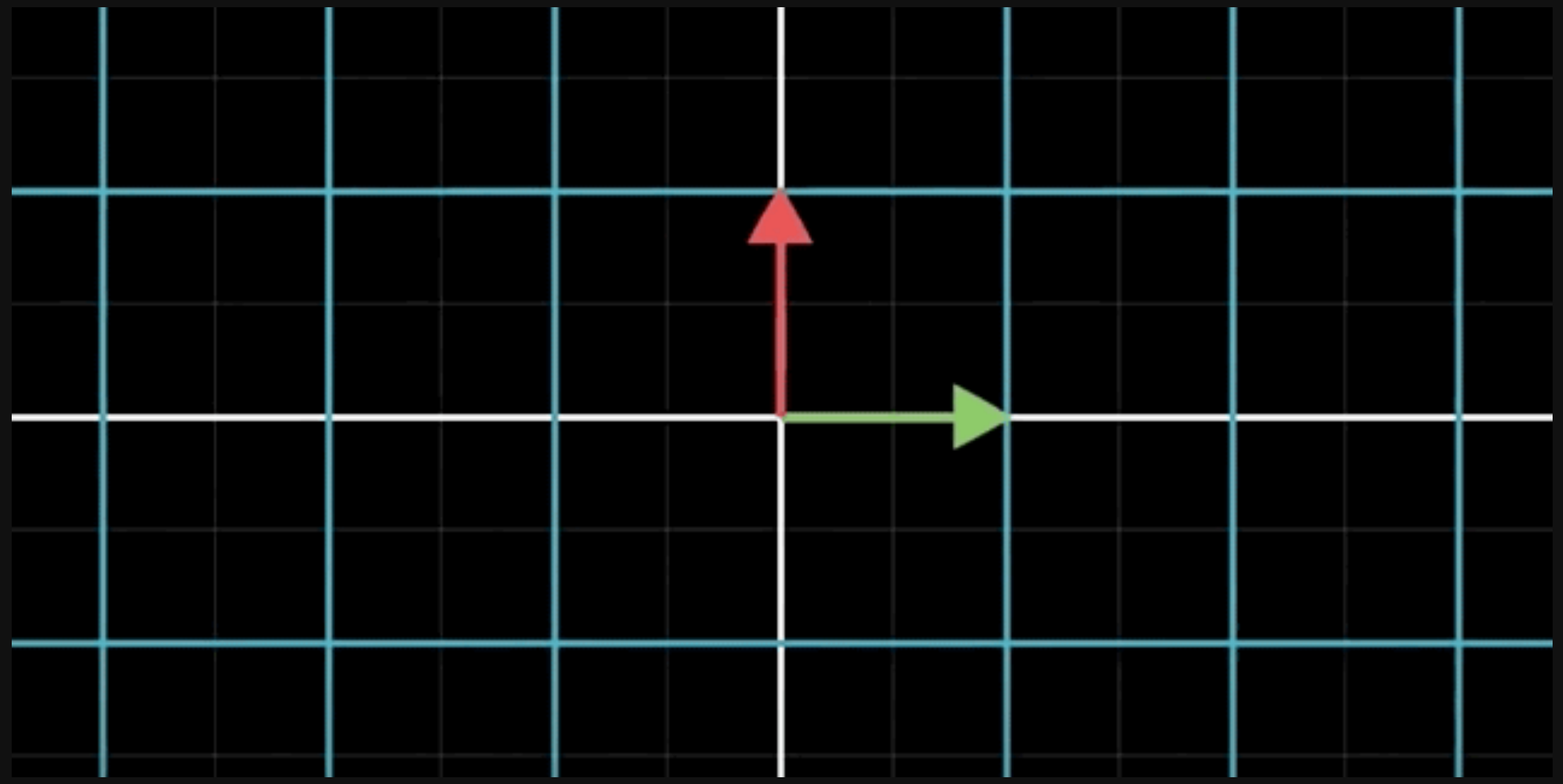
- Is the process of combining two or more matrices to form a new matrix.
- The result of the multiplication of two matrices A and B is a new matrix C, where each element of C is obtained by taking the dot product of a row of A and a column of B.
- Properties:

  1. Associative: $A_1(A_2A_3) = (A_1A_2)A_3$

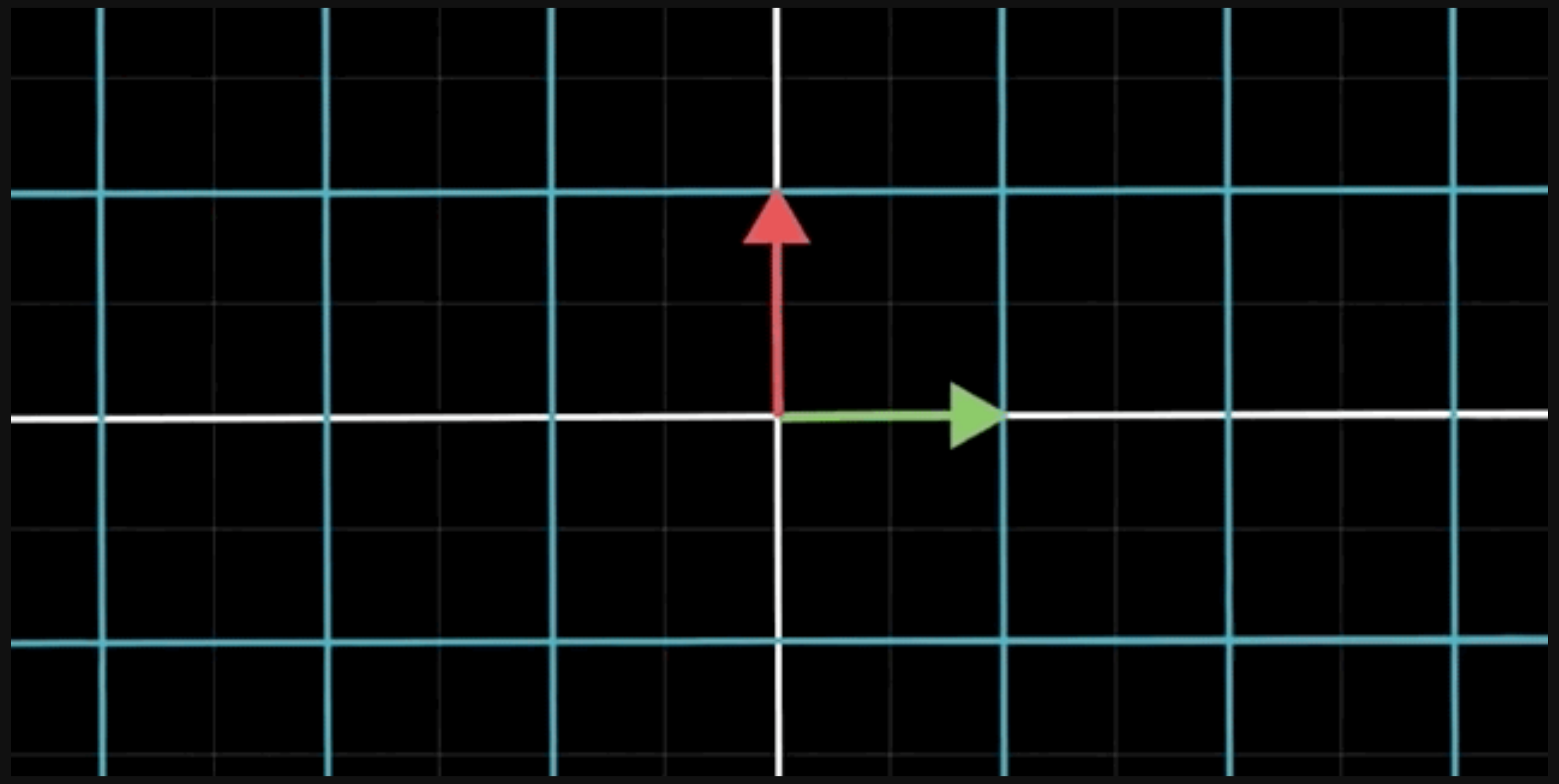
  2. Not Commutative: $A_1A_2 \neq A_2A_1$

# Matrices

## Composition



$$\underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}_{\text{Shear}} \left( \underbrace{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}_{\text{Rotation}} \begin{bmatrix} x \\ y \end{bmatrix} \right)$$



$$\underbrace{\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}}_{\text{Rotation + Shear}} \begin{bmatrix} x \\ y \end{bmatrix}$$

Source: YouTube 3Blue1Brown

# Matrices

## Transpose

- The transpose of a matrix is a new matrix that is obtained by reflecting the original matrix over its main diagonal.
- It switches the row and column indices of the matrix $A$ by producing another matrix, often denoted by $A^T$.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} \quad A_{ij} = A^T_{ji} \qquad M = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

Main diagonal

# Matrices

## Identity Matrix

- A special type of matrix that has 1's on the main diagonal and 0's everywhere else.
- If $A$ is any matrix and $I$ is it's identity matrix, then: $AI = IA = A$.
- Only exists for square matrices!

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad AI = A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

# Matrices

## Diagonal Matrix

- A matrix in which the entries outside of the main diagonal are all zero.
- Usually refers to square matrices.
- Element of the main diagonal can either be zero to non-zero.
- Any identity matrix of any size is a diagonal matrix.

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

# Matrices

## Trace

- The trace of a square matrix is the sum of the elements on its main diagonal.

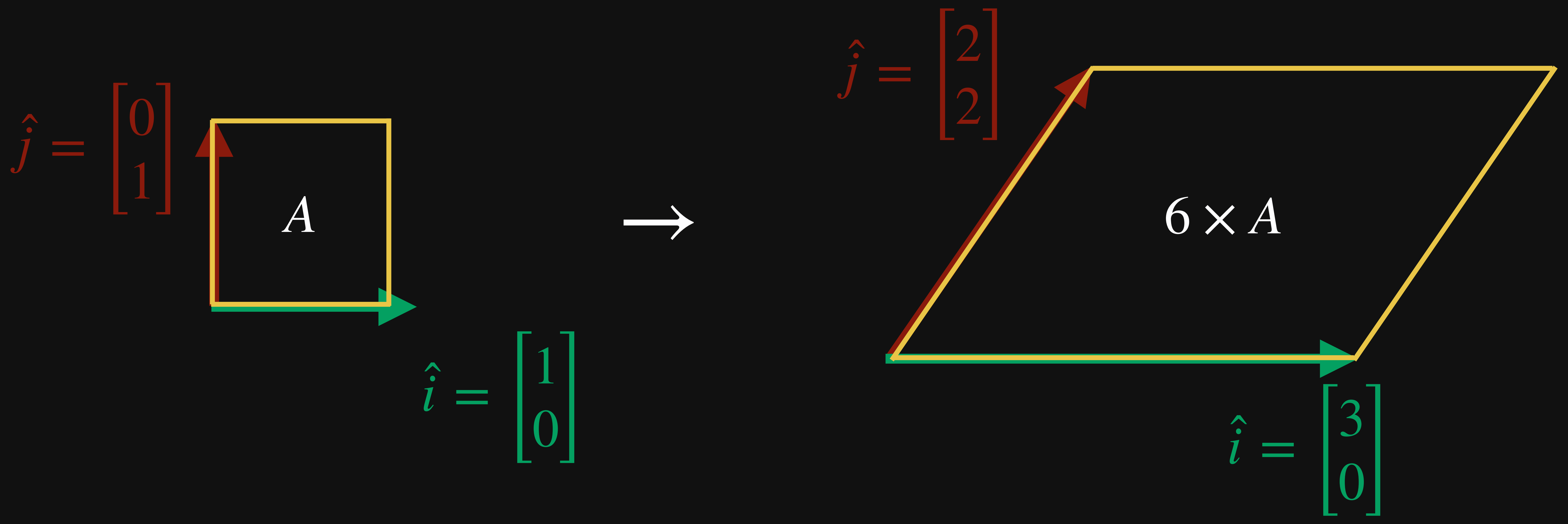
$$tr(A) = \sum_{i=1}^{n} a_{ii}$$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{bmatrix} \quad tr(A) = 1 + 4 + (-2) = 3$$

# Matrices

## Determinants

- The determinant of a square matrix is the scaling factor by which this matrix changes any area/volume inside our space.



# Matrices

## Determinants

- The determinant of a $2 \times 2$ matrix is defined as:

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

- What is the determinant of this matrix?

$$\det\begin{pmatrix} 3 & 2 \\ 0 & 2 \end{pmatrix} = ?$$

# Matrices

## Determinants

- The determinant of a $2 \times 2$ matrix is defined as:

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

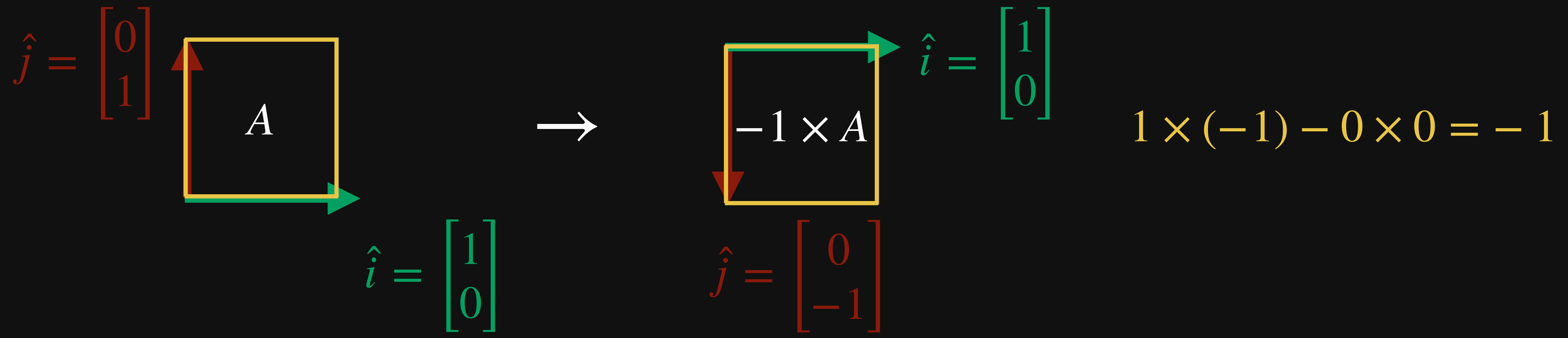
- What is the determinant of this matrix?

$$\det \begin{pmatrix} 3 & 2 \\ 0 & 2 \end{pmatrix} = 3 \times 2 - 0 \times 2 = 6$$

# Matrices

## Determinants

- The determinant of a matrix can be negative if the matrix represents a transformation that involves a reflection.

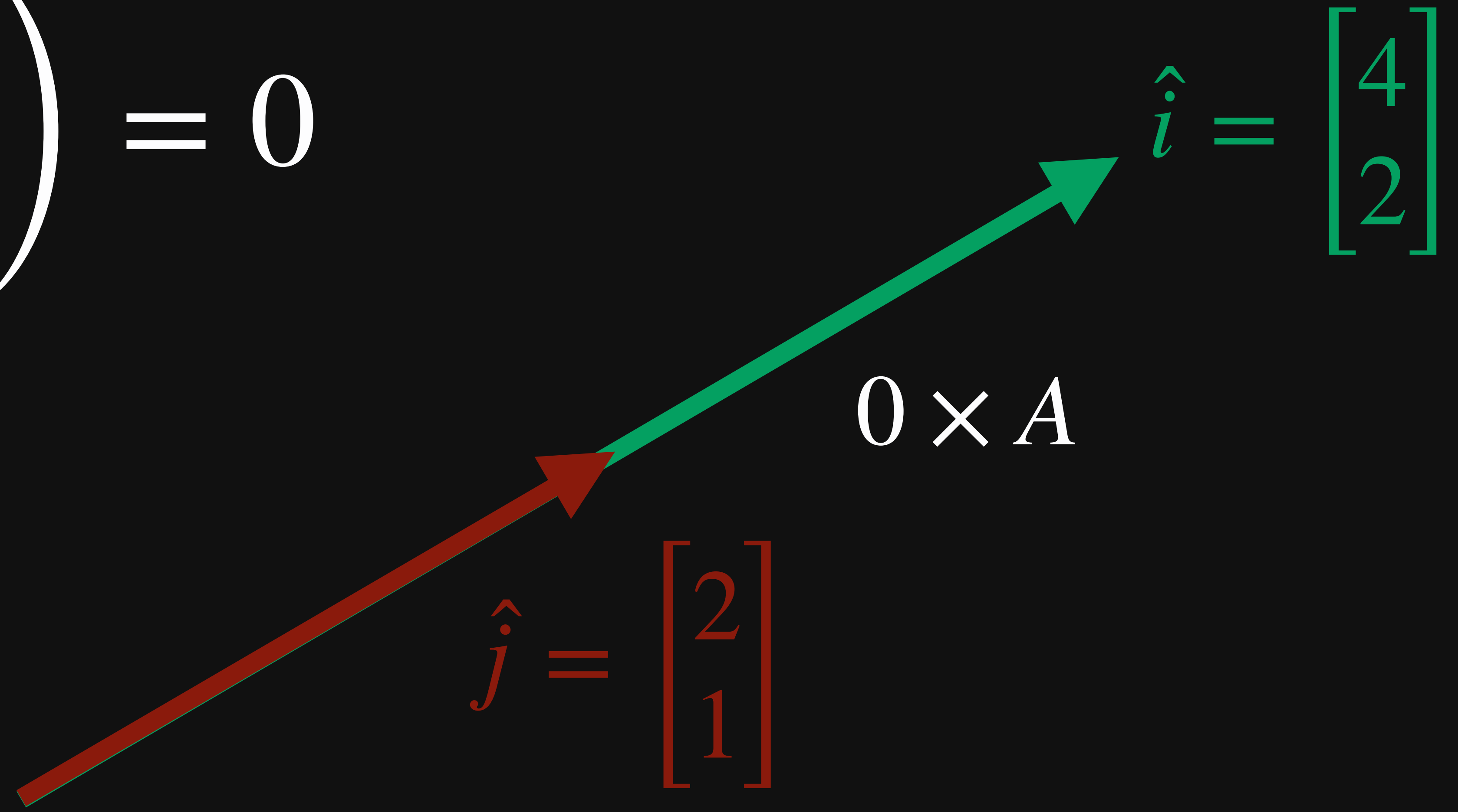


# Matrices

## Determinants

- If $\det(A)$ is $0$ it means that 2D space gets squished to a 1D line and $A$ is *not invertible.*

$$\det\begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix} = 0$$

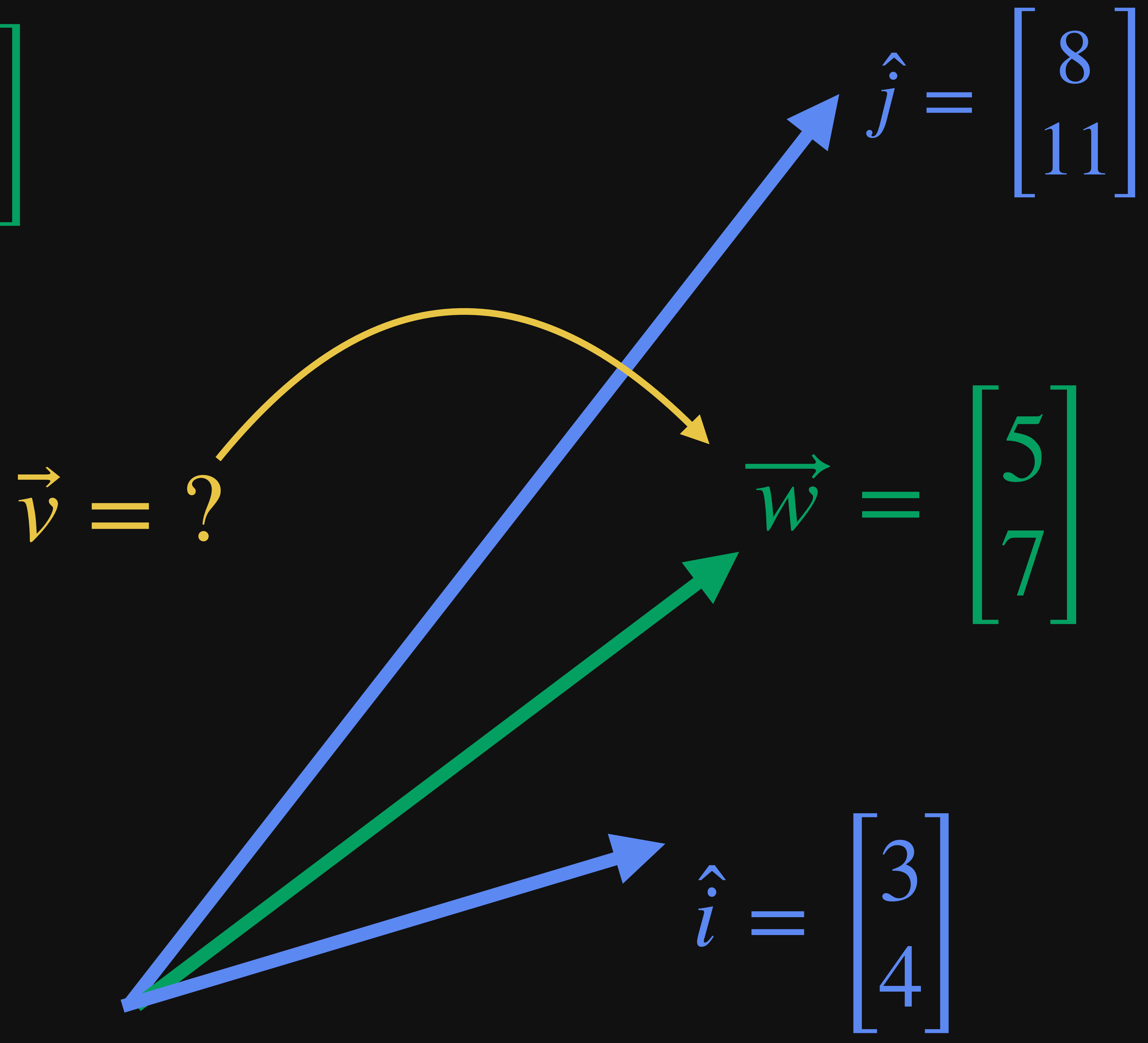


# Matrices

## Inverses

- Matrices can be used to solve systems of linear equations!

$$\begin{aligned} 3x + 8y &= 5 \\ 4x + 11y &= 7 \end{aligned} \qquad \begin{bmatrix} 3 & 8 \\ 4 & 11 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$



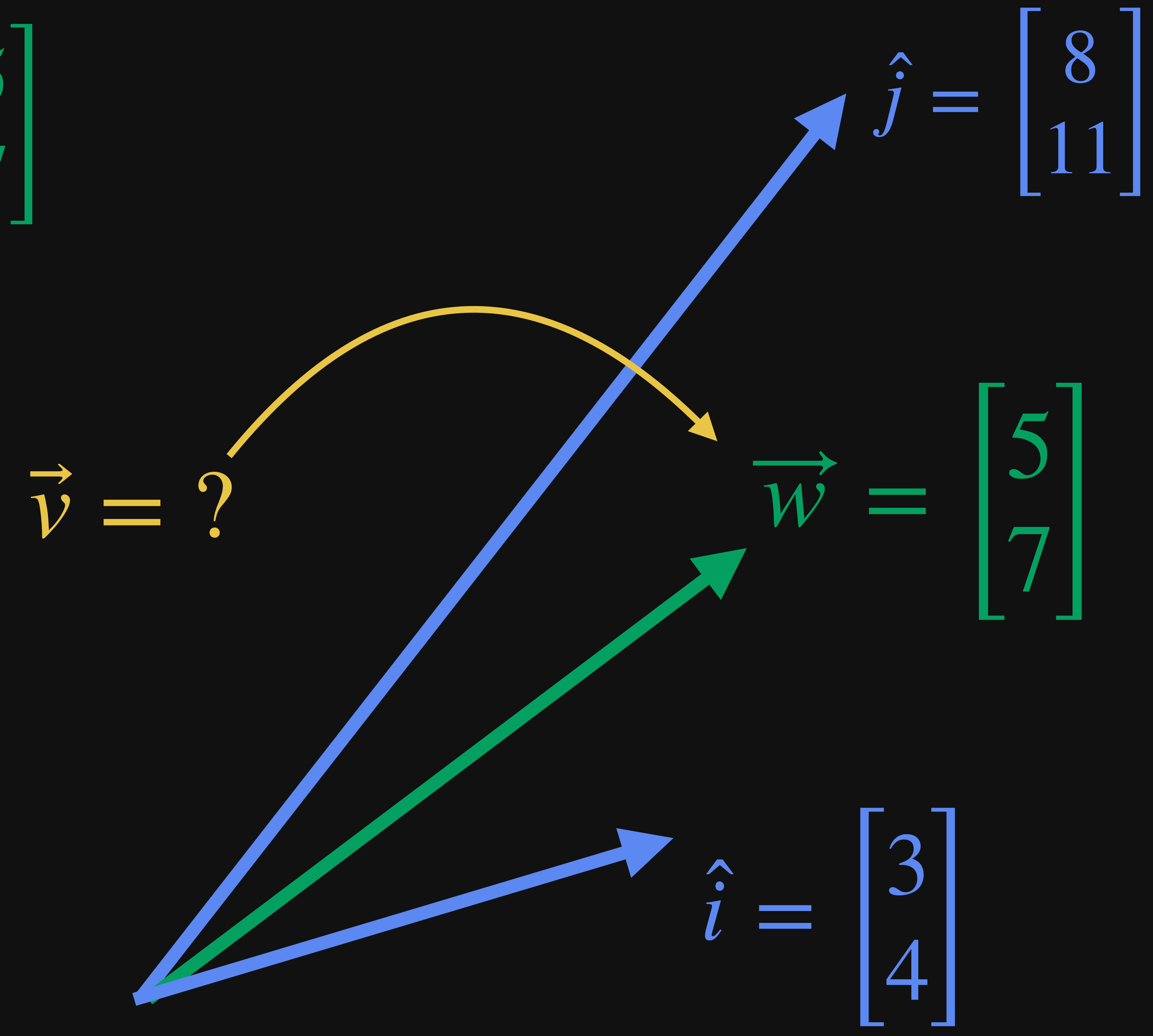
# Matrices

## Inverses

- Matrices can be used to solve systems of linear equations!

$$\begin{aligned} 3x + 8y &= 5 \\ 4x + 11y &= 7 \end{aligned} \qquad \begin{bmatrix} 3 & 8 \\ 4 & 11 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

- Inverse $A^{-1}$ "undoes" the effects of the original matrix transformation.



# Matrices
## Inverses

- Solving for $\vec{v}$:

We know: $A\vec{v} = \vec{w}$

Then: $A^{-1}A\vec{v} = A^{-1}\vec{w}$ $\quad A^{-1}A = I$

$I\vec{v} = A^{-1}\vec{w}$ $\quad I\vec{v} = \vec{v}$

$\vec{v} = A^{-1}\vec{w}$

# Matrices

## Inverses

- Properties of Inverses:
  - $A$ is *invertible* only if $AA^{-1} = I$
  - $A^{-1}$ exists only if $\det A \neq 0$

# Matrices

## Singular Matrix

- A square matrix is *singular* if the matrix has no *inverse*.
- To determine if the matrix is singular we need co compute its determinant.
- A square matrix is singular *iff* $\det = 0$.

# Dot Product

- Algebraic interpretation:

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = a_1 b_1 + a_2 b_2 + a_3 b_3$$

$$\begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = 2 \times 1 + 5 \times 3 + 1 \times 5 = 22$$

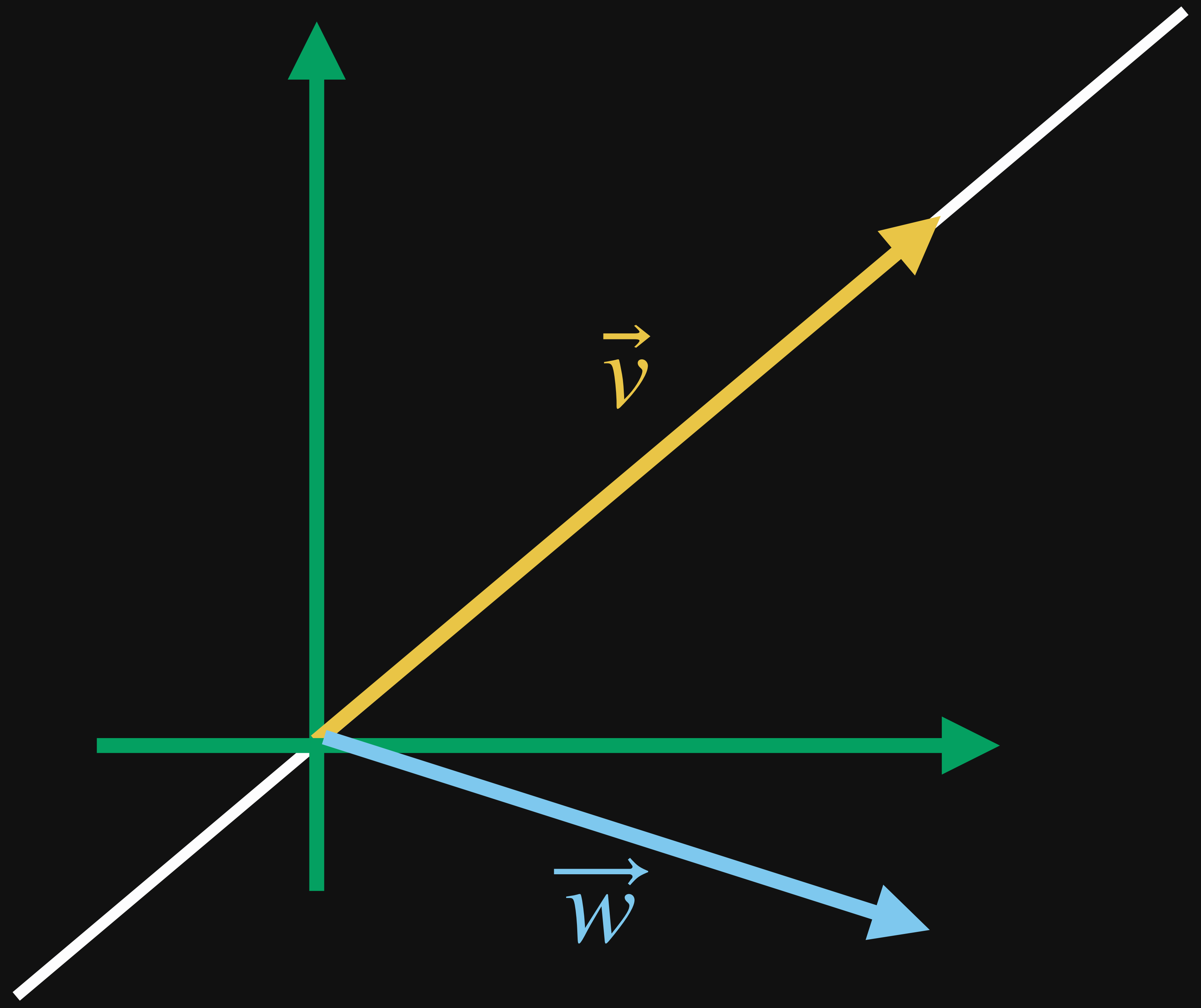
$$\begin{bmatrix} 1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 4 \end{bmatrix} = 1 \times 2 + 3 \times 4 = 14$$

# Dot Product

- Geometric interpretation:

$$a \cdot b = \|a\| \, \|b\| \cos \theta$$

$$\|a\| = \sqrt{\sum_{i}^{n} a_i^2}$$

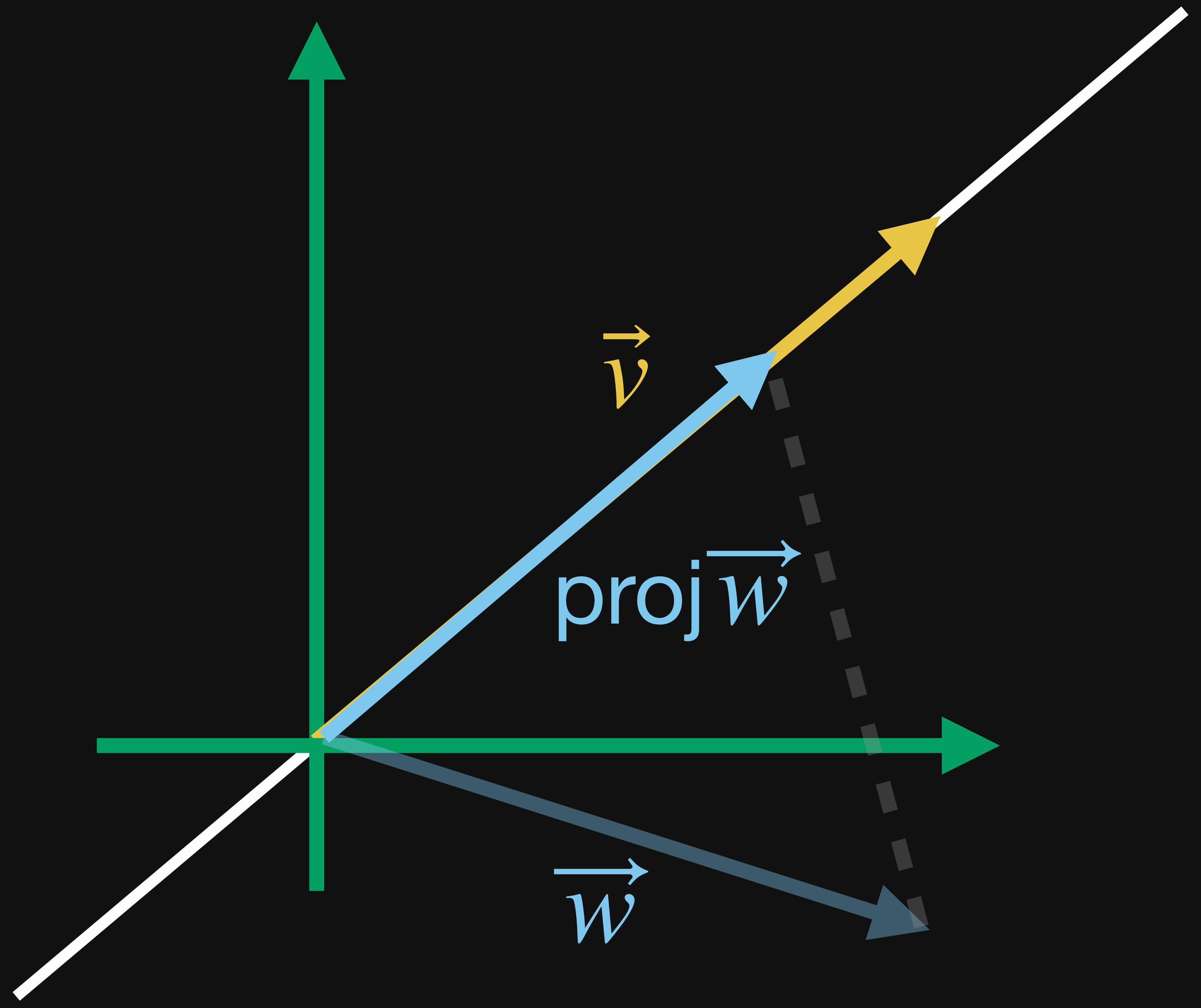


# Dot Product

- Geometric interpretation:

$$a \cdot b = \|a\| \, \|b\| \cos \theta$$

$$\|a\| = \sqrt{\sum_{i}^{n} a_i^2}$$



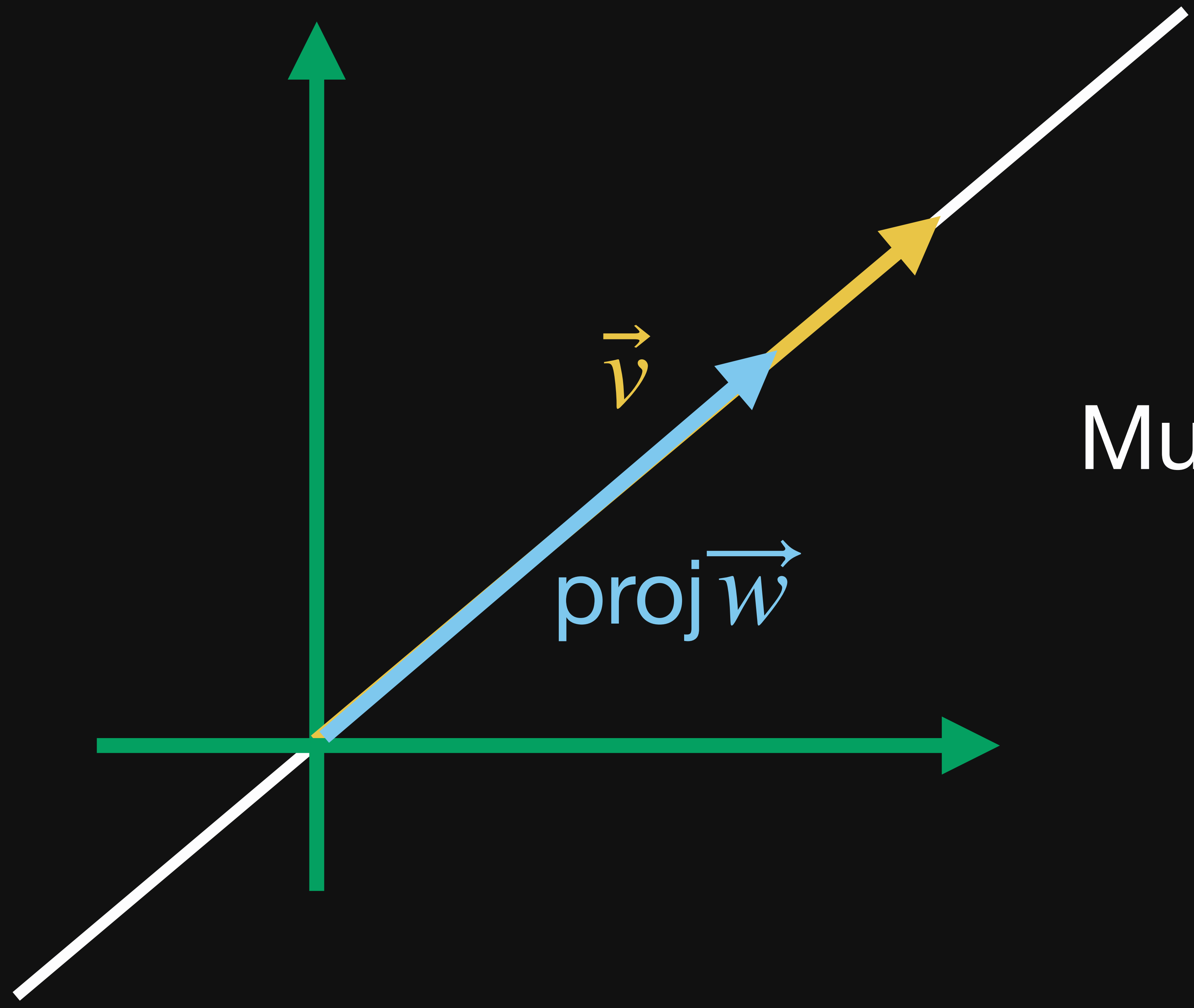
# Dot Product

- Geometric interpretation:

$$a \cdot b = \|a\| \, \|b\| \cos\theta$$

$$\|a\| = \sqrt{\sum_i^n a_i^2}$$

The result of the transformation lies on a number line, not in a 2d space




Multiply $\|\vec{v}\|$ proj$\|\vec{w}\|$

# Useful Links

- 3Blue1Brown Essence of Linear Algebra
- Mathematics from Machine Learning Specialisation
- Marc Peter Deisenroth *"Mathematics for Machine Learning"*, Chapters 2,3