

Linear Algebra

An Intuitive and « Gentle » Introduction

Screenshots from 3blue1brown's animations:

https://www.youtube.com/playlist?list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab

Algebra

« Algebra is the intellectual instrument which has been created for rendering clear the quantitative aspect of the world » — Alfred North Whitehead

Math tends to be characterized by an extensive use of symbolism :

Generally, the use of these symbols mark the transition between arithmetic and algebra.

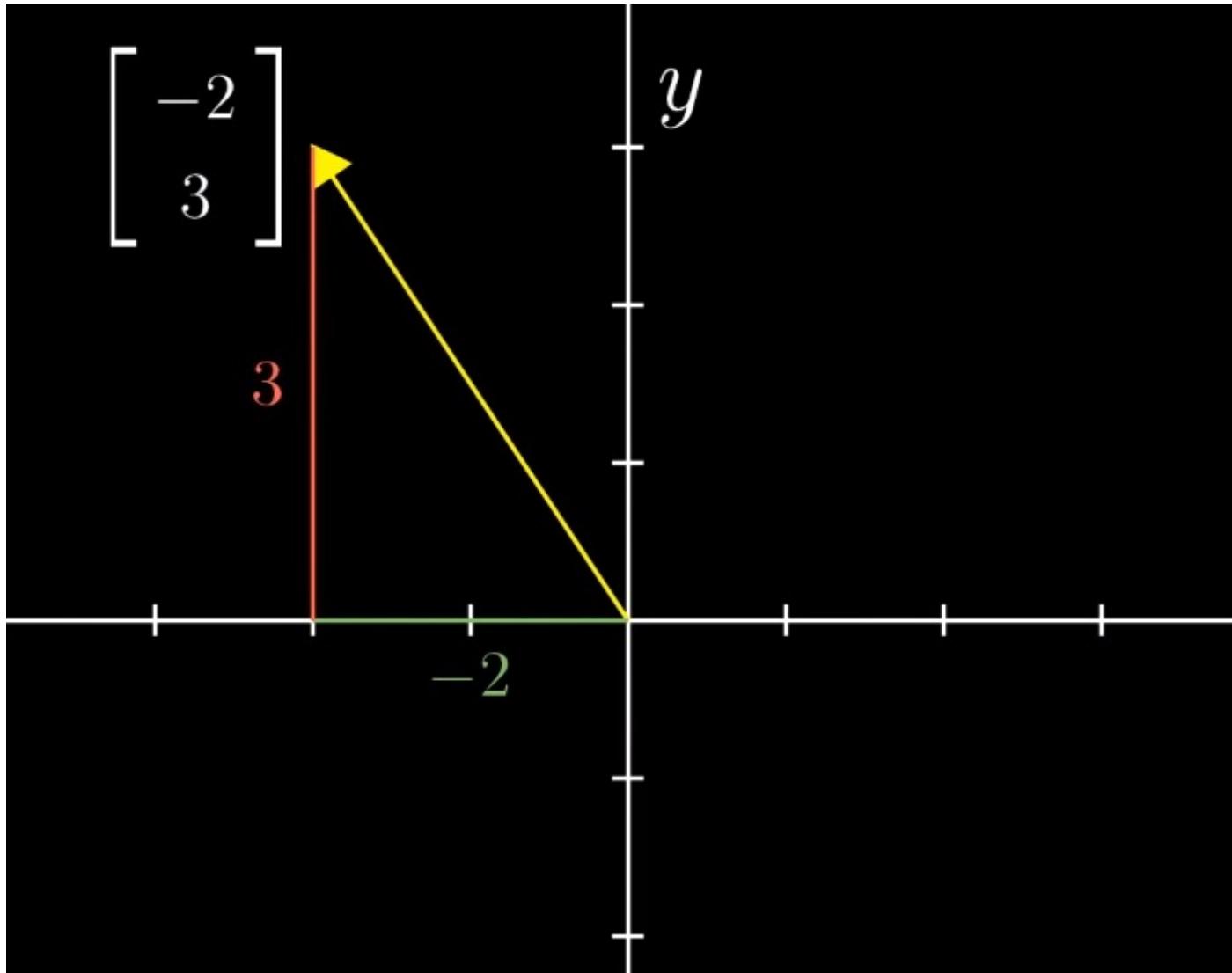
⇒ Linear algebra can be understood as a language applied to *lines* and objects which go beyond a single dimension (numbers).

Vectors

- Fundamental building block for linear algebra
- 3 perspectives :
 - Physics student
 - Vectors are arrows that we can move around
 - Computer Science student
 - Vectors are ordered lists of numbers => think of feature vectors
 - [200 sqm, 1.5M] vs. [500K, 150 sqm] => order matters
 - Math student
 - Generalizes both of these views

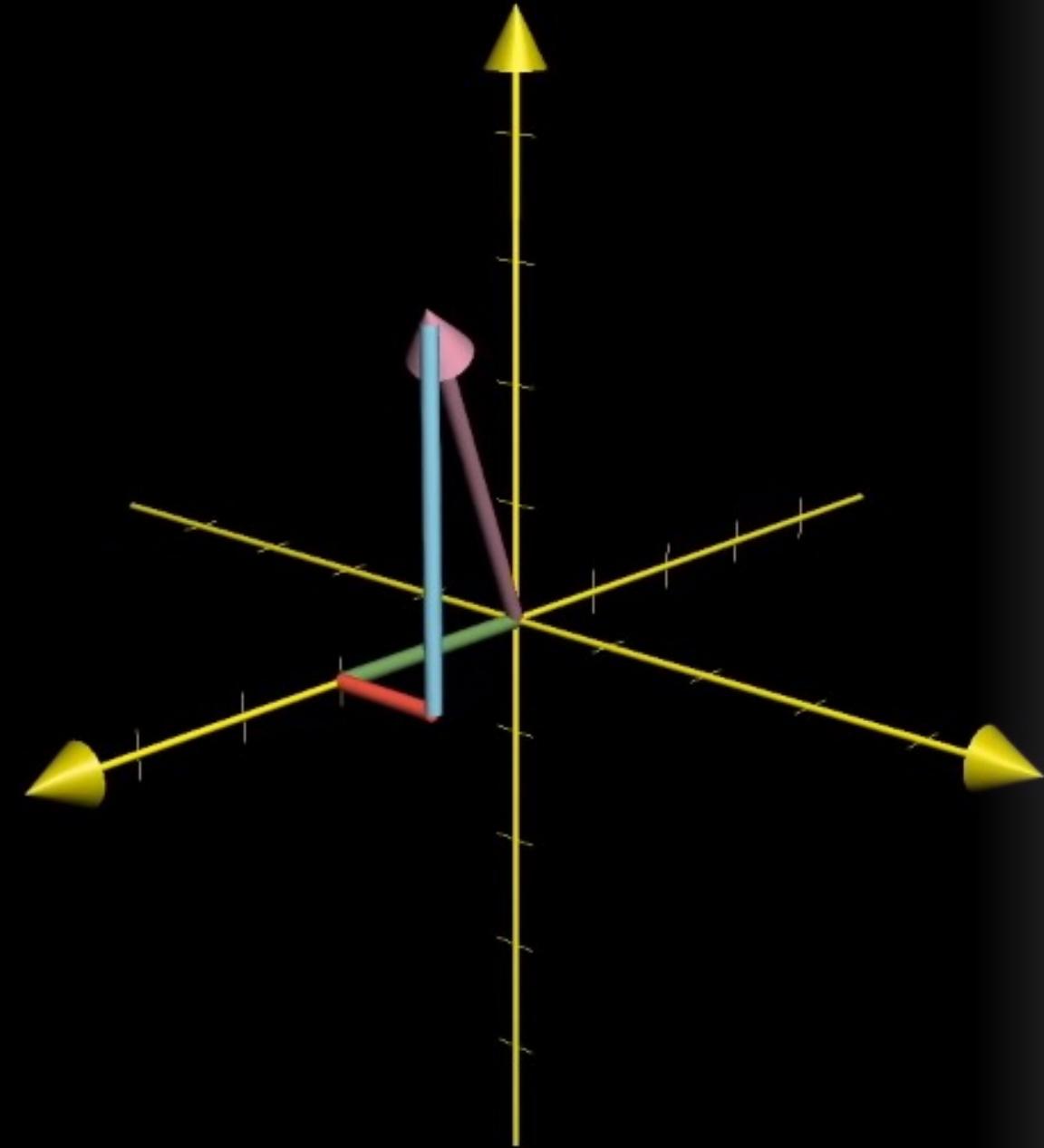
Vectors (2D for now)

- What do you think about when someone talks about vectors:
 - An arrow ?
 - A set of coordinates/list of numbers ?
- An arrow that sits inside a coordinate system, with its tail on the origin => where is the origin ?
- What are coordinates ?
 - Can be seen as a set of instructions to go from the base to the tip of the vector
 - Go along the x axis a certain amount and then along the y axis a certain amount



- A point's coordinates are traditionally written horizontally $(-2, 3)$
- Unlike vector coordinates which are vertical with square brackets $\begin{bmatrix} -2 \\ 3 \end{bmatrix}$

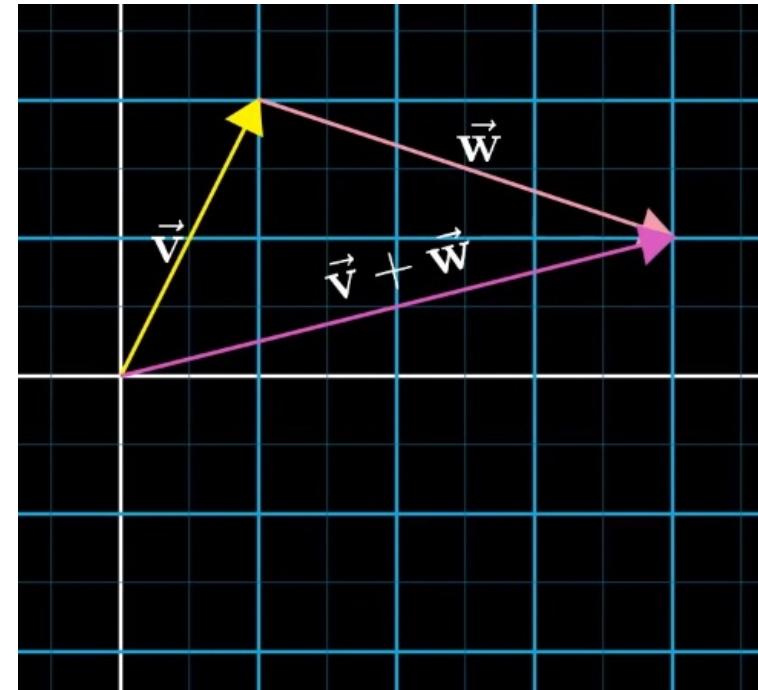
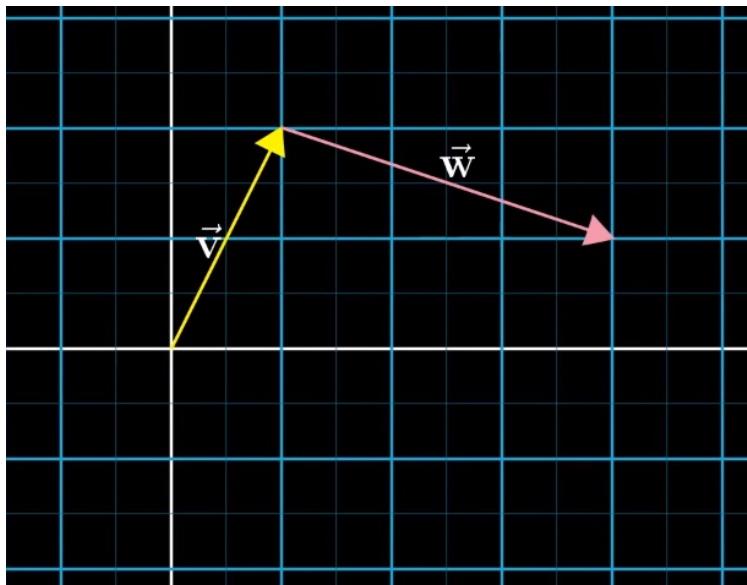
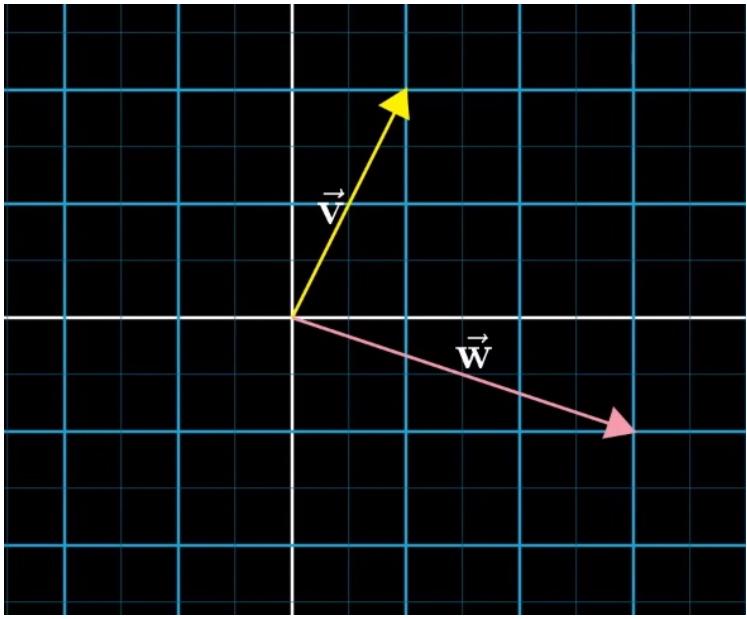
3rd Dimension ?



- Ordered Triplet of numbers
- Same « instruction » principle
 - move along the x axis, then parallel to the y axis then parallel to the z axis
- Gives you one unique vector in space.

2 Central Operations

- Addition
- Multiplication
- This you know how to apply to numbers
- What happens when we enter 2D space ?

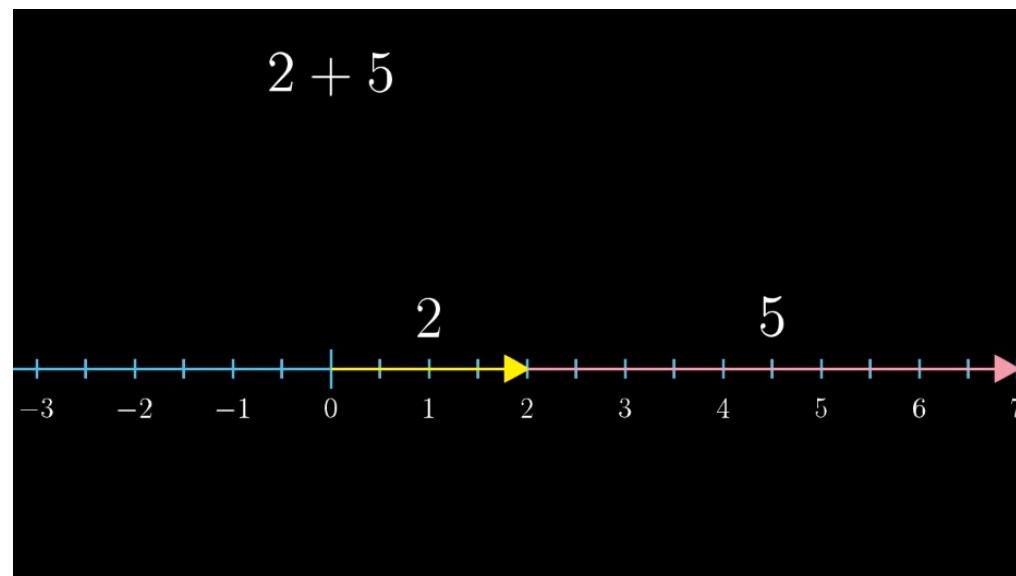


Vector Addition

$$v + w$$

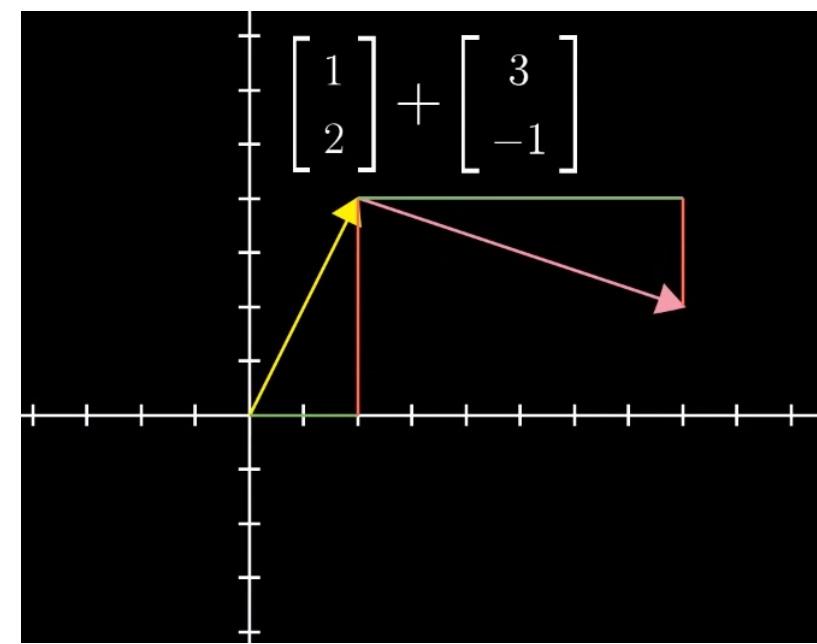
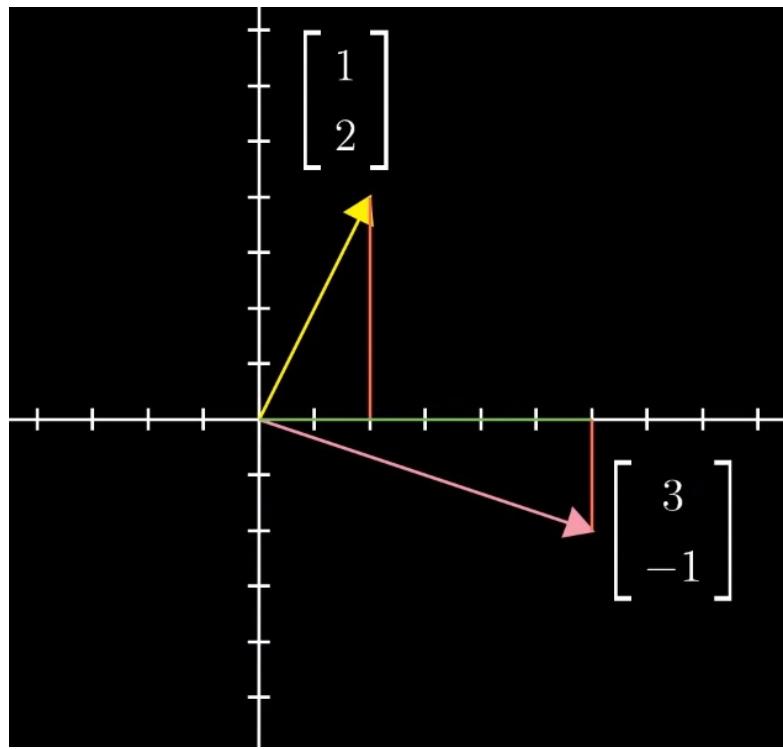
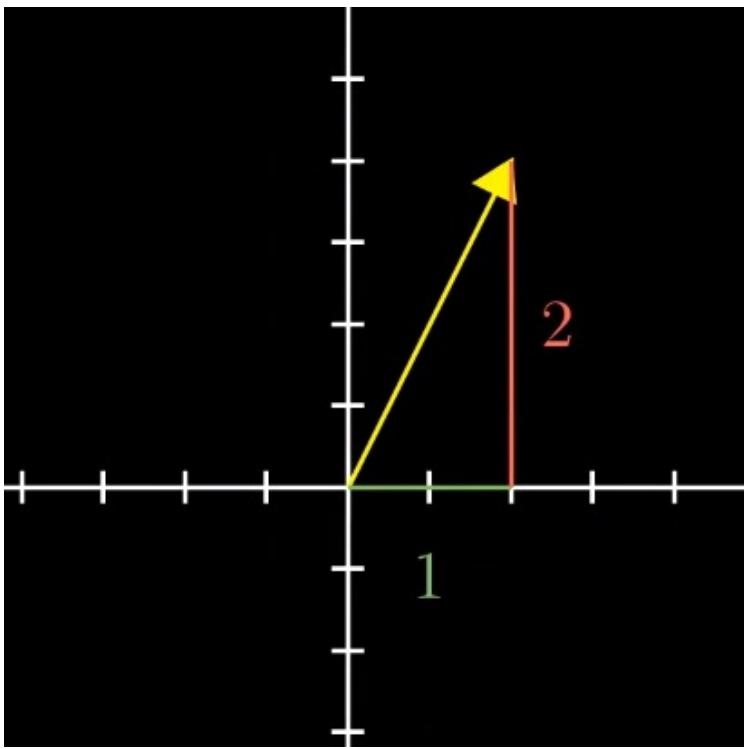
Why this definition of addition ?

- Each vector represents a certain movement, a step in a direction
- So going along the sum of vectors which starts from the origin takes you exactly to the place where you end up by going along one vector and then the other
- Same thing in 1-dimension when adding $2 + 5$ for example.

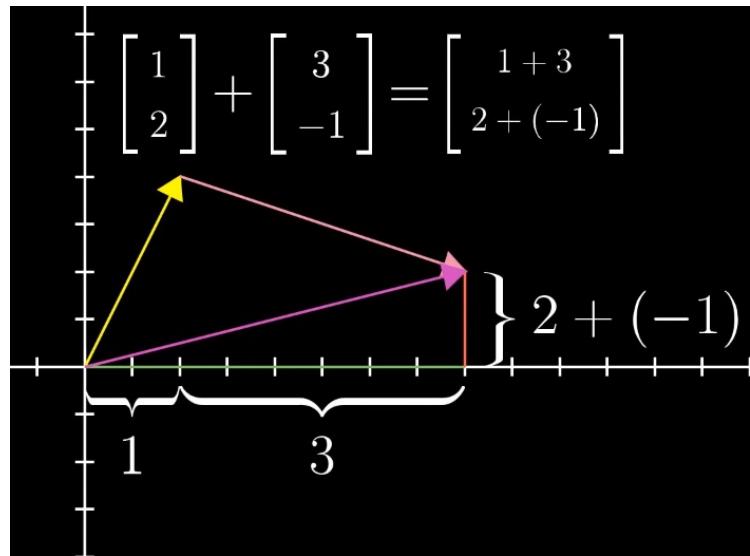


Example

Can think of vector addition as a *4 step path* from the origin to the tip of the second vector.



Example and General Formula



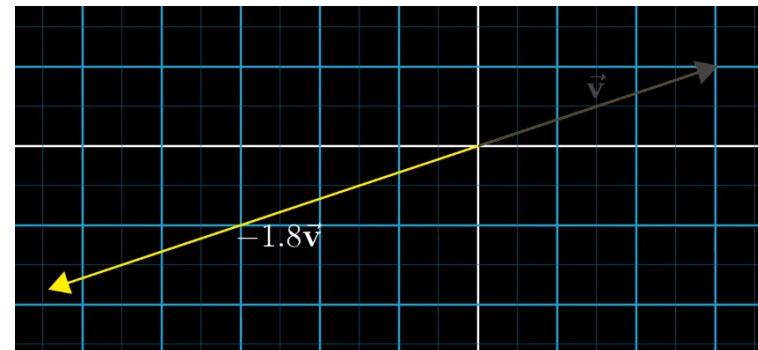
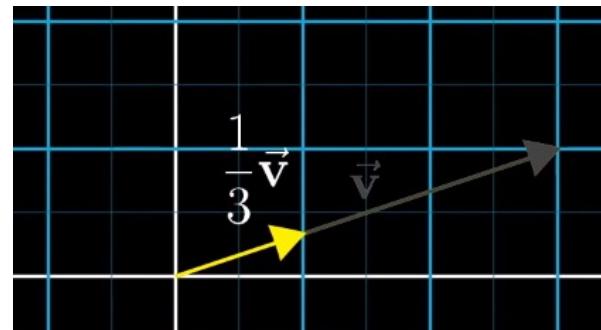
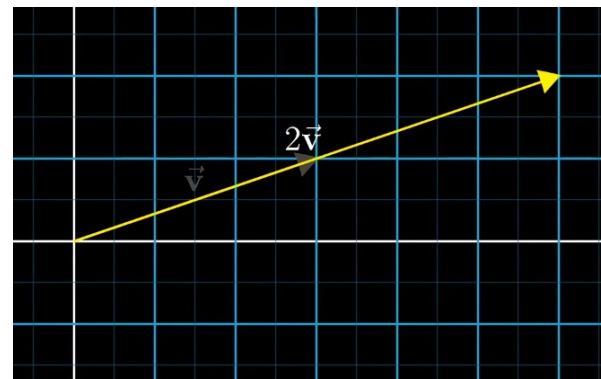
$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}$$

- This is the same as
 - moving along the x axis a certain number of steps
 - and then moving parallel to the y axis for a certain number of steps.

- What about subtracting one vector from another ?

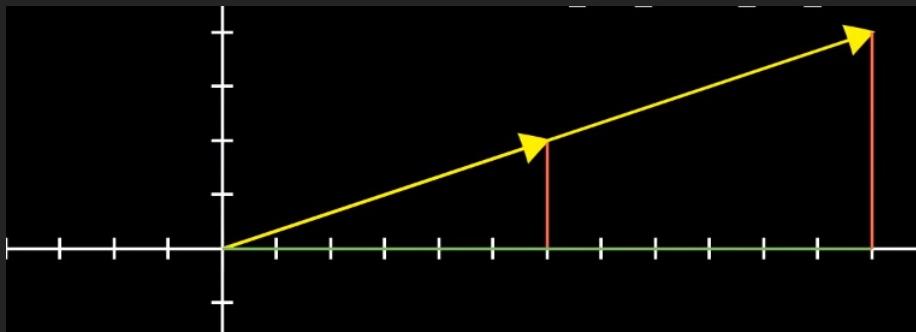
Vector Multiplication

- A few examples :
 - $2\mathbf{v} \Rightarrow$ stretches the vector \mathbf{v}
 - $\frac{1}{3}\mathbf{v} \Rightarrow$ squishes vector \mathbf{v}
 - $-1.8\mathbf{v}$ flips it around and stretches it



- This is called « scaling » a vector
- 2, -1.8 etc... are also known as « scalars »
- General Idea:

$$2 \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$



Exercises

$$\mathbf{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \text{ and } \mathbf{v} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

Compute $\mathbf{u} + 2\mathbf{v}$ and $\mathbf{u} - \mathbf{v}$

Exercises

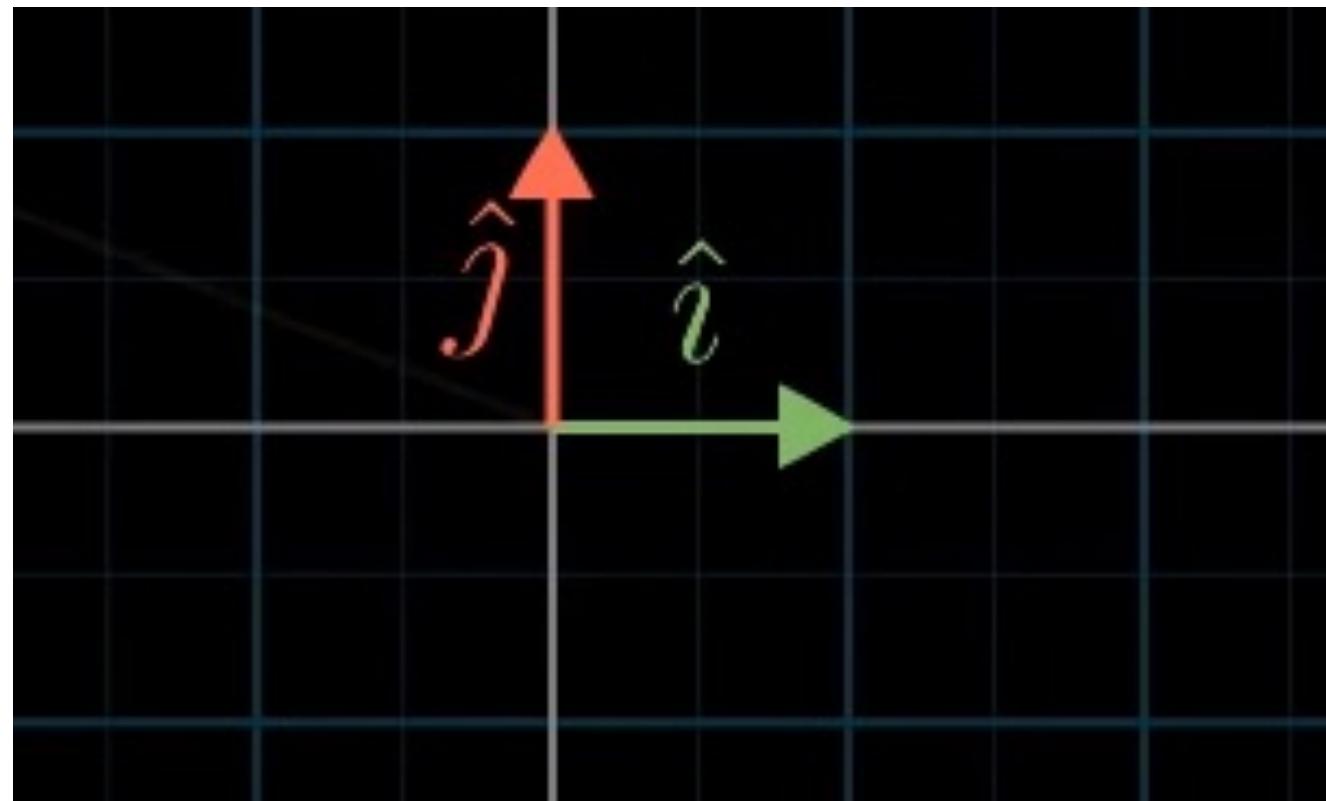
- Display the following vectors using arrows on an xy-graph:
- u ,
- v ,
- $-v$,
- $2v$,
- $u + 2v$
- $u - v$

How to reconcile the different views mentioned ?

- Linear Algebra is in fact about the ability to go back and forth between the geometric / spatial view and the ordered list of numbers view
- CS student can now visualize his many lists of numbers which can be very helpful to find patterns in data, see how it is distributed
- Physics students and computer graphics programmers can now describe space mathematically and manipulate it.

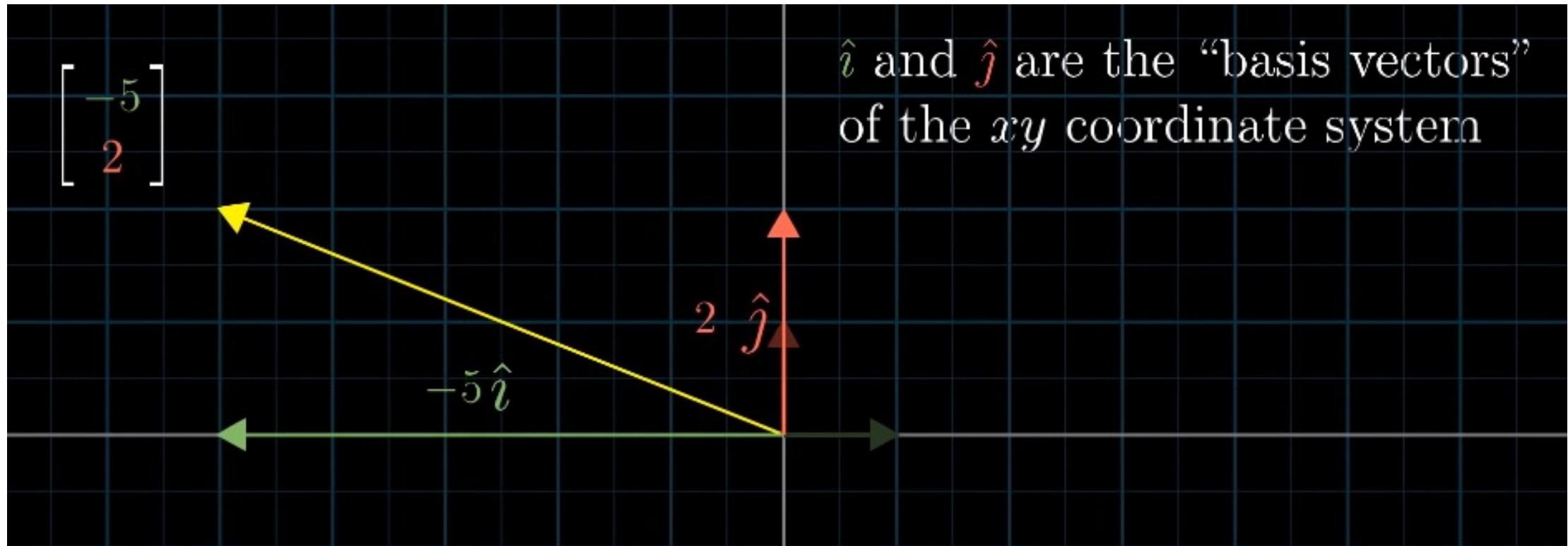
Basis Vectors

- Basis vectors : $i \hat{}$ and $j \hat{}$
- Vectors which determine the basic unit for the space
- $i \hat{}$ = $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$
- $j \hat{}$ = $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$



Linear Combinations

- Any vector in 2D space can be expressed using these 2 vectors :



Linear Combinations

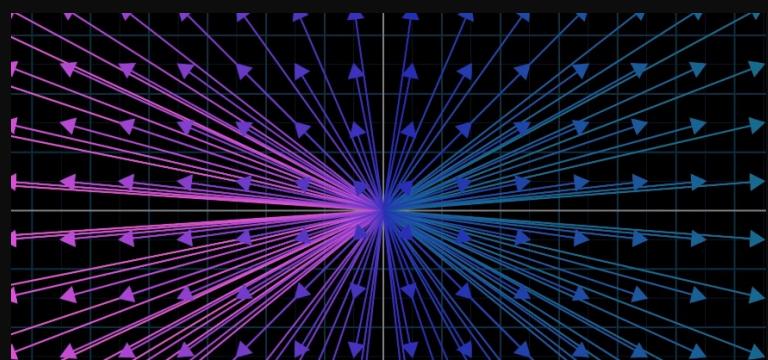
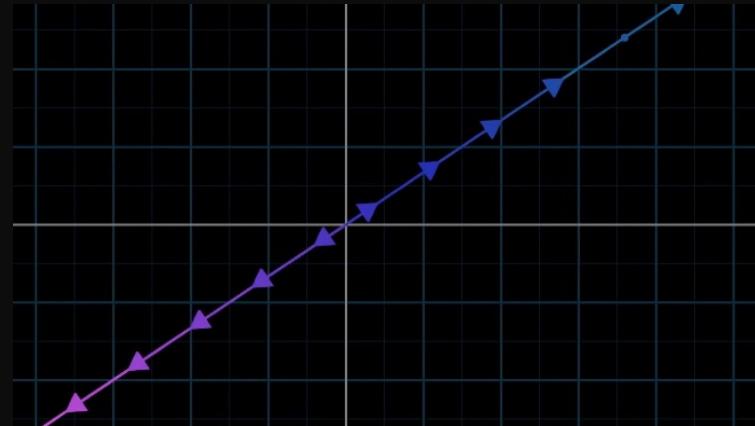
- So any time we describe vectors, it depends on an implicit choice of basis vectors.
- We can actually choose many other vectors as « basis » vectors, we are not limited to i_{hat} and j_{hat}

$$\begin{bmatrix} 5 \\ -2 \end{bmatrix} = 5i_{hat} + (-2)j_{hat}$$

- Any time we scale and add vectors => we call this a linear combination

Span

- Can reach any point in space by using linear combinations (barring certain specific cases where vectors are colinear or one of the vectors is the 0 vector)
- This is called the *span* of 2 vectors, the amount of space they can cover basically.
- In 2D : this can be
 - a line if the vectors are colinear
 - or a plane, ie all of the 2D space



Mini exercise

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 8 \end{bmatrix}$$

Is the vector \vec{b} a linear combination of the above vectors ?

$$\vec{b} = \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}$$

Solution

$$\vec{b} = 3\vec{v}_1 + 0\vec{v}_2 + 0\vec{v}_3$$

Quick Recap

- Vector addition
- Vector multiplication
- Basis vectors can be any vectors really, not just \hat{i} and \hat{j}
- We can describe any vector in a particular space as a linear combination of basis vectors.

Linear Transformations and Matrices

More intuitive to figure out what they do *visually*, so let's try and do this

Unfortunately, no one can be told what the Matrix is. You have to see it for yourself.

-Morpheus

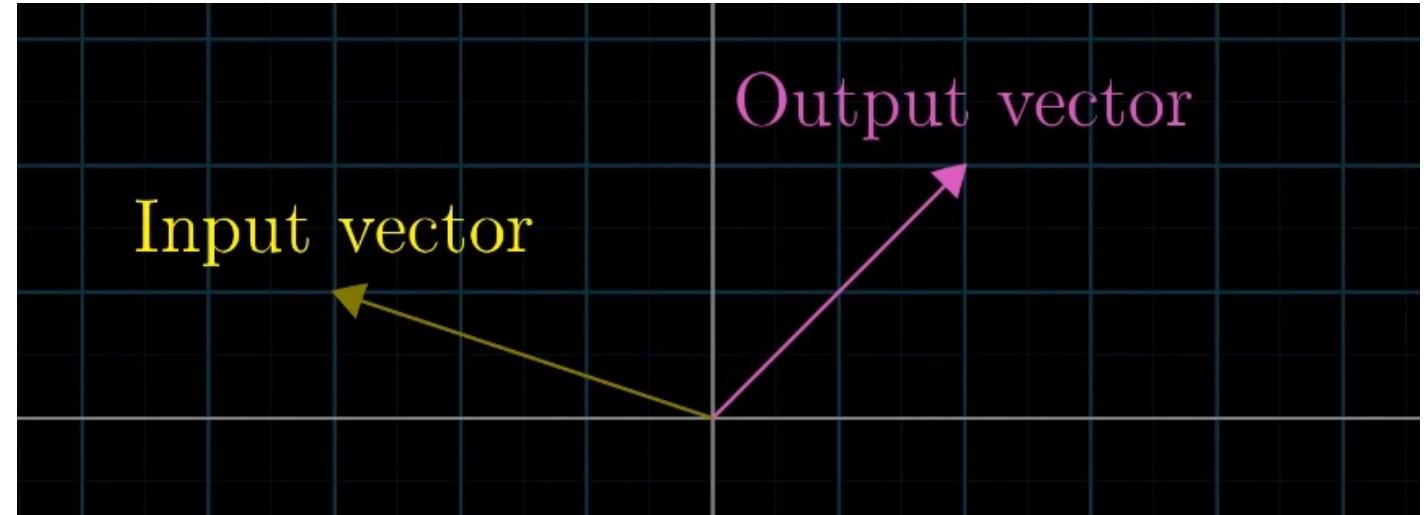
- Get an intuition for
 - matrix / vector multiplication
 - and later on vector / vector multiplication or dot product
 - (so far we've only seen what happens when we multiply vectors by scalars)
- Goal is to try not to rely too much on memorization for these operations which neural nets rely on and which are generally used in machine learning.

Transformation

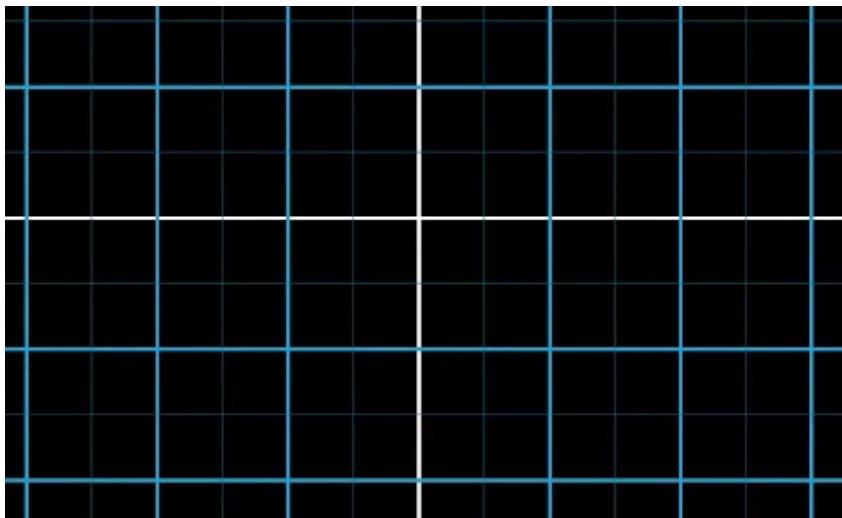
- Fancy word for function:
 - Takes an input and spits out an output
 - Vector 2 Vector in Linear Algebra
- Transformation has maybe more of a visual connotation than function however...
- If a transformation maps an input vector to an output vector, we can imagine that input vector *moving over* to the output vector.

Illustration

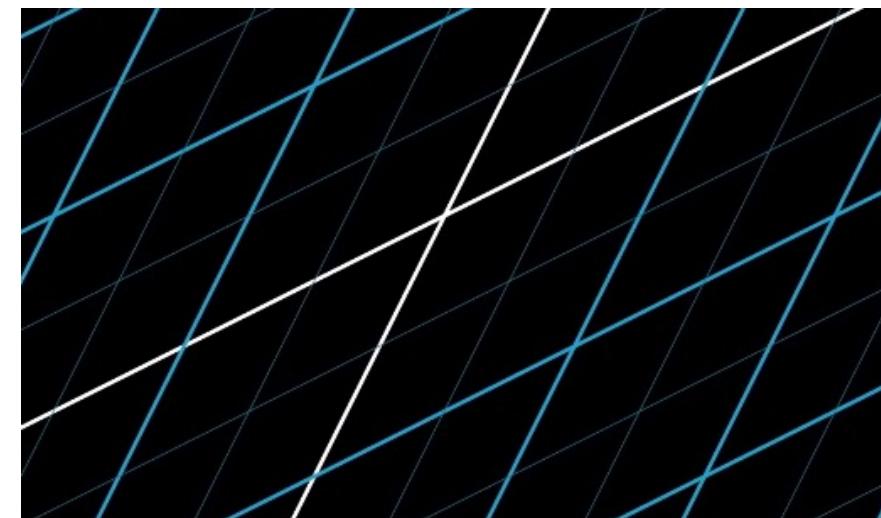
- For a single vector :



- For every vector in the space (represented by an infinite grid):



=>

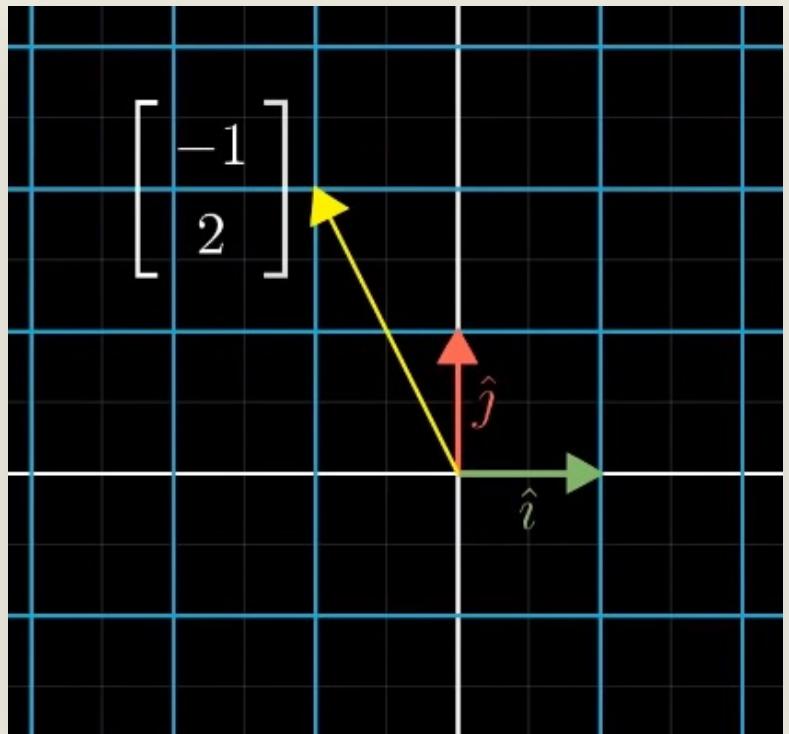


Linear (Transformation)

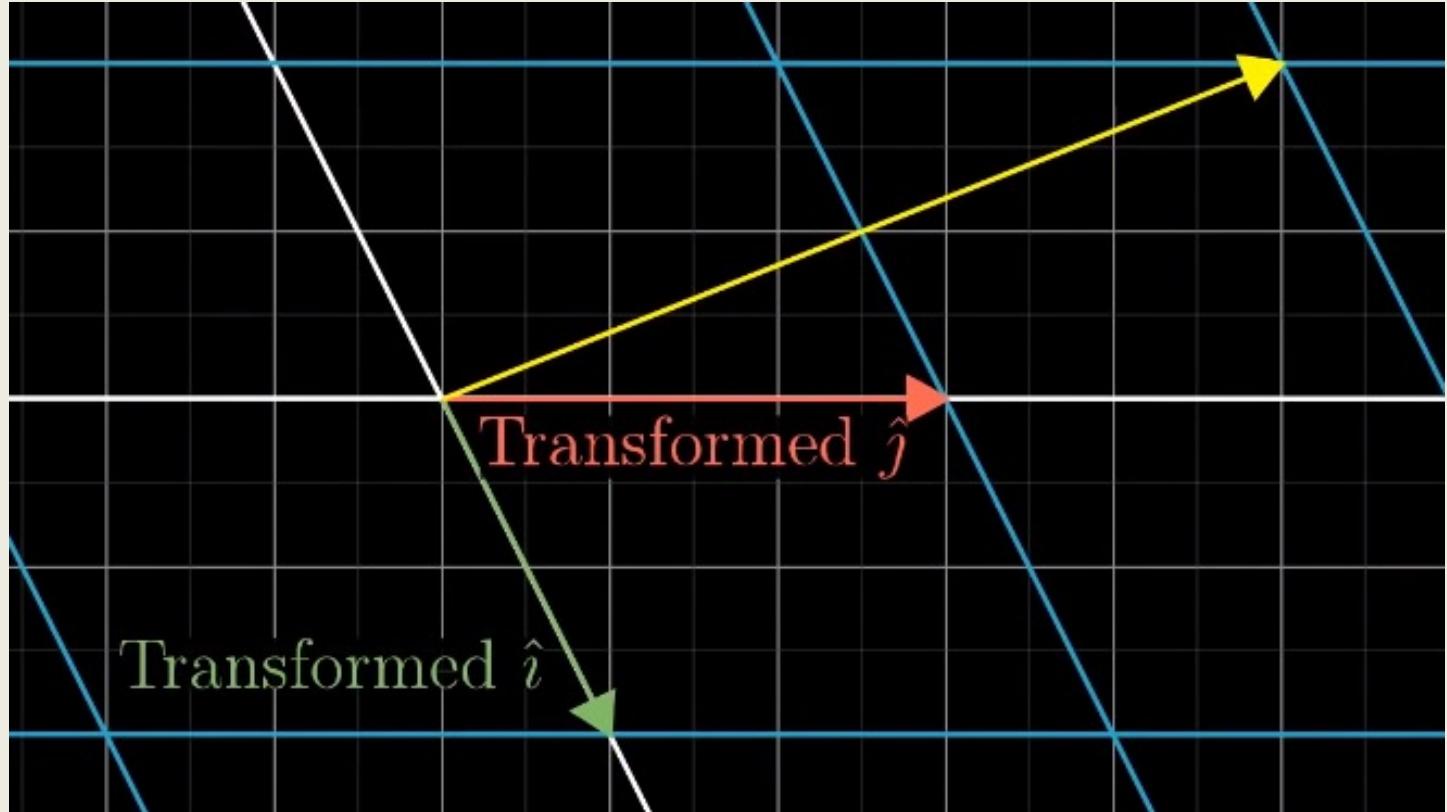
- All lines must remain lines without getting curved
- The origin must remain fixed in place
- Everything remains parallel and evenly spaced, which has some nice properties.

- To know where a vector is going to land after going through a linear function/linear transformation, all we need to do is look at where \hat{i} and \hat{j} land.
- Remember :
 - if a vector in the input space is a linear combination \hat{i} and \hat{j}
 - Then it will be equal to a similar linear combination in the transformed space, but with \hat{i} and \hat{j} transformed (ie. where these basis vectors landed).

Example



$$\vec{v} = -1\hat{i} + 2\hat{j}$$



$$\text{Transformed } \vec{v} = -1(\text{Transformed } \hat{i}) + 2(\text{Transformed } \hat{j})$$

- In the example, we see where the vector landed...
- But by **only looking** at where the **basis vectors** land, we can deduce any transformed vector's new position.
- So we basically only need these 2 new « basis » vectors :

$$\hat{i} \rightarrow \begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad \hat{j} \rightarrow \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow x \begin{bmatrix} 1 \\ -2 \end{bmatrix} + y \begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 1x + 3y \\ -2x + 0y \end{bmatrix}$$

Matrices

- We commonly package these coordinates into a 2x2 grid of numbers
 - => a 2x2 Matrix

“2x2 Matrix”

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



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

Generalized Formula of a Matrix / Vector Product

- A matrix / vector product is like applying a function to a vector where:
 - The columns of the matrix are the transformed basis vectors
 - The result of the operation is a linear combination of those column vectors.
- When learning about matrix vector products, seeing it as a linear combination is usually not made explicit... and you're left with having to remember by heart $ax + by$ etc...

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \underbrace{x \begin{bmatrix} a \\ c \end{bmatrix} + y \begin{bmatrix} b \\ d \end{bmatrix}}_{=} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$$

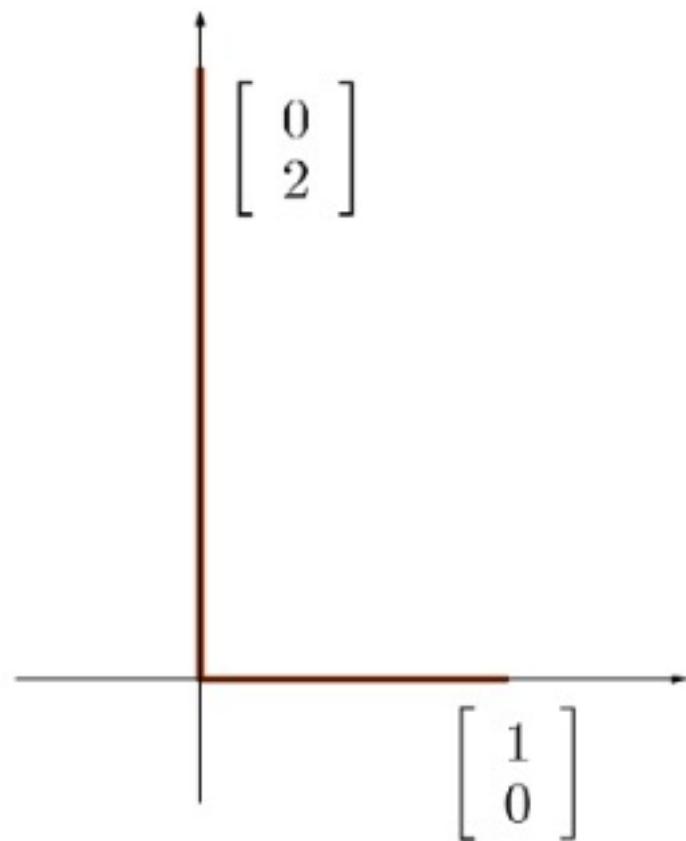
Where all the intuition is

Transformation examples

- Rotation
 - 90° for example
- Shear (lateral shift)
- => so everytime you see a matrix, you can interpret it as a certain transformation of space (where everything stays linear, no curves!)

Exercise

- Show the effect of the linear transformation T on a letter L made up of the vectors



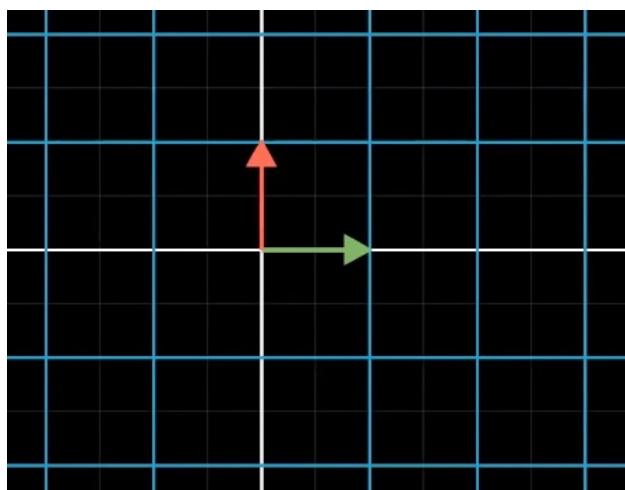
$$T(\mathbf{x}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \mathbf{x}$$

Matrix Multiplication

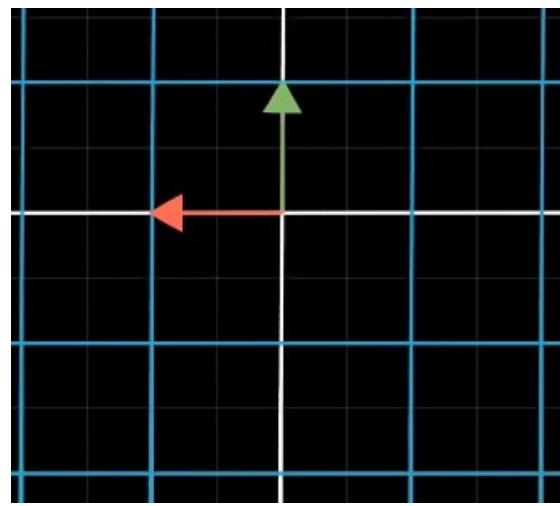
- What happens when we multiply 1 matrix by another ?
- What does it mean intuitively, what happens ?
- Always a good idea to think of the geometric meaning and not get lost in the numbers at first

Matrix Multiplication

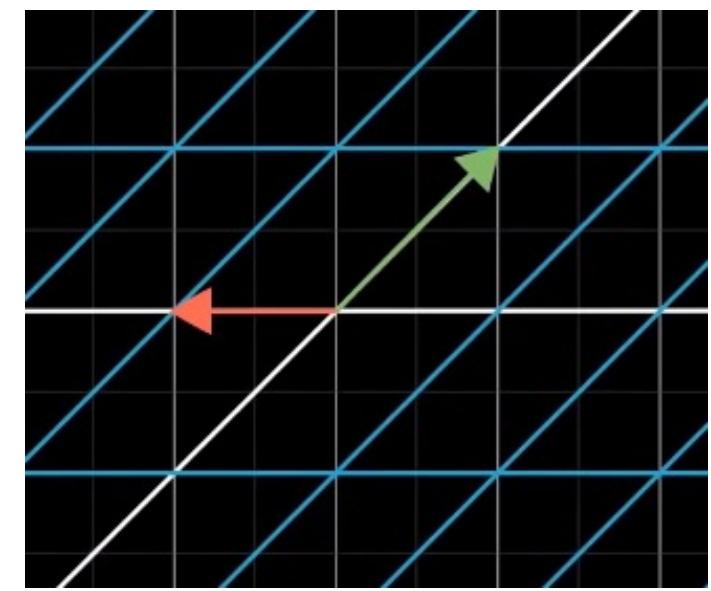
- Apply one transformation to another: move space around twice !
- The combined transformation is called a « composition »
- Rotation then shear for example : new basis vectors are the column of the composition matrix



=>



=>



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

Matrix Multiplication

- The composition matrix can be understood as the result of a product of matrices.
- That product should be read right to left, as with functions:

$$f(g(x))$$

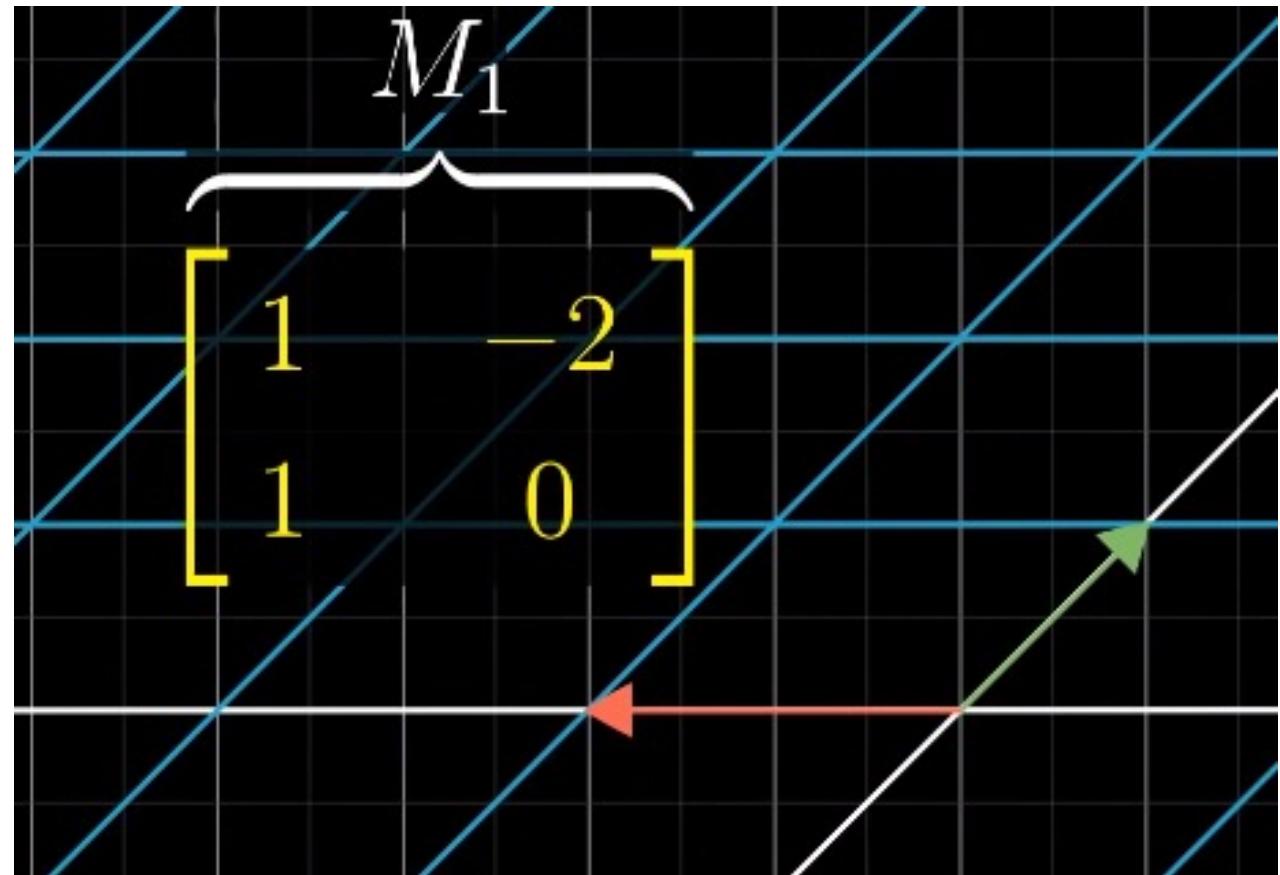
- example :

$$\begin{aligned}g(x) &= 3x \\f(x) &= 2x \\f(g(x)) &=?\end{aligned}$$

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}_{\text{Shear}} \underbrace{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}_{\text{Rotation}} = \underbrace{\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}}_{\text{Composition}}$$

Example

- We are going to apply 2 consecutive transformations, and follow the basis vectors:
- Matrix M_1 and its transformation:



Example

- Matrix M_2 and its transformation :

$$M_2 = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}$$



Example

- Total effect gives us a new transformation !
- Although we could figure this out geometrically/visually, this isn't necessarily immediately intuitive like in the previous example...
- So let's compute this composition numerically without looking at the vectors in space.

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

Example

- First column of M_1 tells us where \hat{i} lands. Let's follow this vector to begin with.

$$\underbrace{\begin{bmatrix} M_2 & M_1 \end{bmatrix}}_{\left[\begin{array}{cc|c} 0 & 2 & 1 \\ 1 & 0 & 1 \end{array} \right]} \left[\begin{array}{cc} -2 \\ 0 \end{array} \right] = \left[\begin{array}{cc} ? & ? \\ ? & ? \end{array} \right]$$

Example

- To figure out what happens to this vector when we then apply the transformation defined by M_2 ,
- we can use a simple matrix vector product, just like in the previous slides !

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

Example

- And if you remember the intuition behind it,
- This is nothing but a linear combination :

$$\begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Mini Exercise

- Now try and figure out the values in the second column of the matrix!

$$\overbrace{\begin{bmatrix} M_2 & M_1 \end{bmatrix}}^{} \left[\begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & 0 \end{bmatrix} \right] = \begin{bmatrix} 2 & ? \\ 1 & ? \end{bmatrix}$$

Example

- Same thing as for the first column :

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

General formula

- As previously, let's first take care of \hat{t} and its resulting transformation

$$\overbrace{\begin{bmatrix} a & b \\ c & d \end{bmatrix}}^{M_2} \overbrace{\begin{bmatrix} e & f \\ g & h \end{bmatrix}}^{M_1} = \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e \\ g \end{bmatrix} = e \begin{bmatrix} a \\ c \end{bmatrix} + g \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} ae + bg \\ ce + dg \end{bmatrix}$$

General Formula

- And then follow the same process for \hat{j} !

$$\overbrace{\begin{bmatrix} a & b \\ c & d \end{bmatrix}}^{M_2} \overbrace{\begin{bmatrix} e & f \\ g & h \end{bmatrix}}^{M_1} = \begin{bmatrix} ae + bg & ? \\ ce + dg & ? \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} f \\ h \end{bmatrix} = f \begin{bmatrix} a \\ c \end{bmatrix} + h \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} af + bh \\ cf + dh \end{bmatrix}$$

General Formula

- The final result :

$$\begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$$

- It is very common to be taught this formula as something to memorize as a process...
- Without anyone really ever going through where it comes from and what it represents.
- Can be very useful to understand that we are actually applying 2 consecutive linear transformations.

Practice

- $E = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$
- $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
- Let $H = EA$.
- Find H.
- Now compute AE

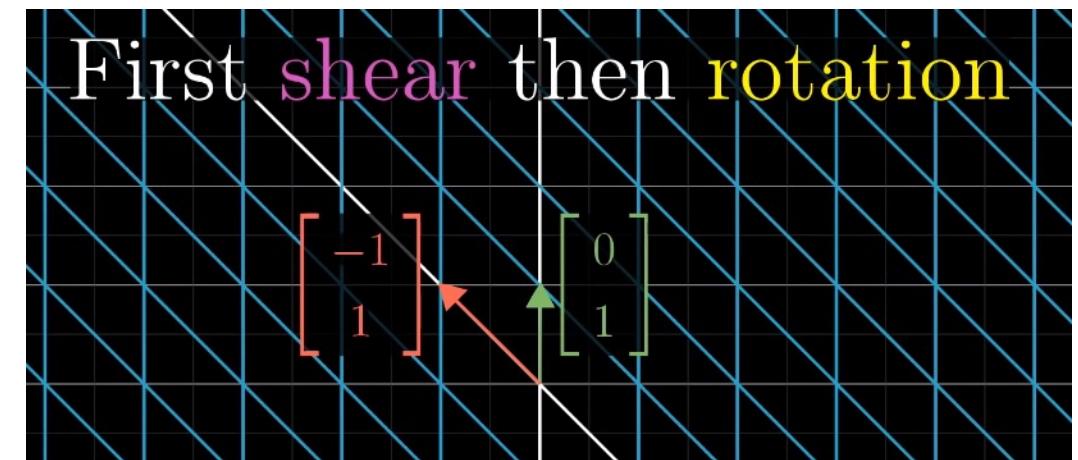
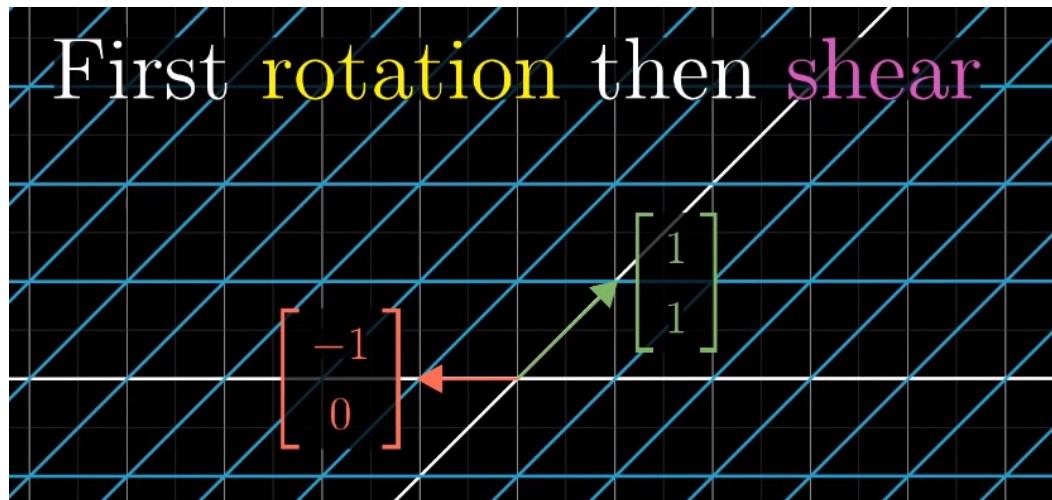
Matrix Multiplication is not commutative

- Now you know what this operation represents, here's a question :
- Does order matter ?

$$M_1 M_2 \stackrel{???}{=} M_2 M_1$$

Extra example

- Because you now understand what these matrices represent, it seems intuitive that those 2 transformations are not equal if you go through them in your head.
- No need to go do calculations and compare the results !
- $M_1M_2 \neq M_2M_1$



Non-square transformations

- Can go from 2d to 3d with a linear transformation, but this is slightly less intuitive as the spaces are unconnected compared to 2d \rightarrow 2
- # Columns : # of basis vectors in the input space
- # Rows = landing spot for the basis vectors has 3 separate coordinates
- These matrices map from (#columns) dimensions to (#rows) dimensions

$$\mathbb{R}^{3 \times 2}$$

3 × 2 matrix

2 columns

3 rows

$$\left\{ \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ -2 & 1 \end{bmatrix} \right.$$

Practice

$$\cdot \begin{bmatrix} 1 & 3 \\ 2 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \Rightarrow ?$$

$$\cdot \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \Rightarrow ?$$

Recap

- Matrix matrix multiplication :
 - Consists in applying 1 transformation after another
 - The « product » of 2 matrices is called a « composition » and describes the combined, overall transformation.
 - Therefore the « product » of 2 matrices yields a new matrix.
 - Lastly, matrices can be « rectangular » and map 1 set of dimensions to another higher

Dot Product

- Now let's look at the dot product, ie. « Vector / vector multiplication »
- Keep linear transformations in mind !

Standard view first

- If you have 2 vectors with the same dimensions, then the dot product is equal to :
- Summing the products of all the paired up coordinates.

$$\begin{bmatrix} 2 \\ 7 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 8 \\ 2 \\ 8 \end{bmatrix} = 2 \cdot 8 + 7 \cdot 2 + 1 \cdot 8$$

Dot product

Okay so what about...

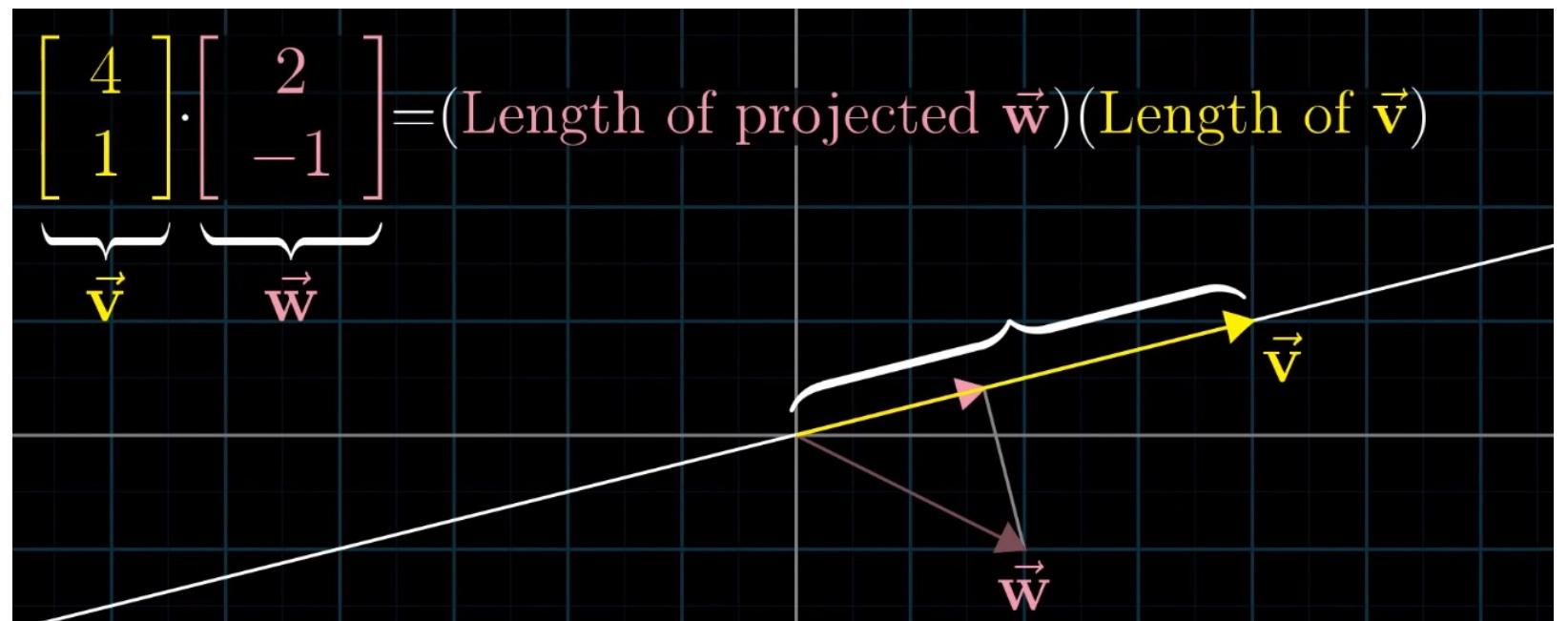
- $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} \dots ?$

- $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \dots ?$

- What do you notice about the result ?

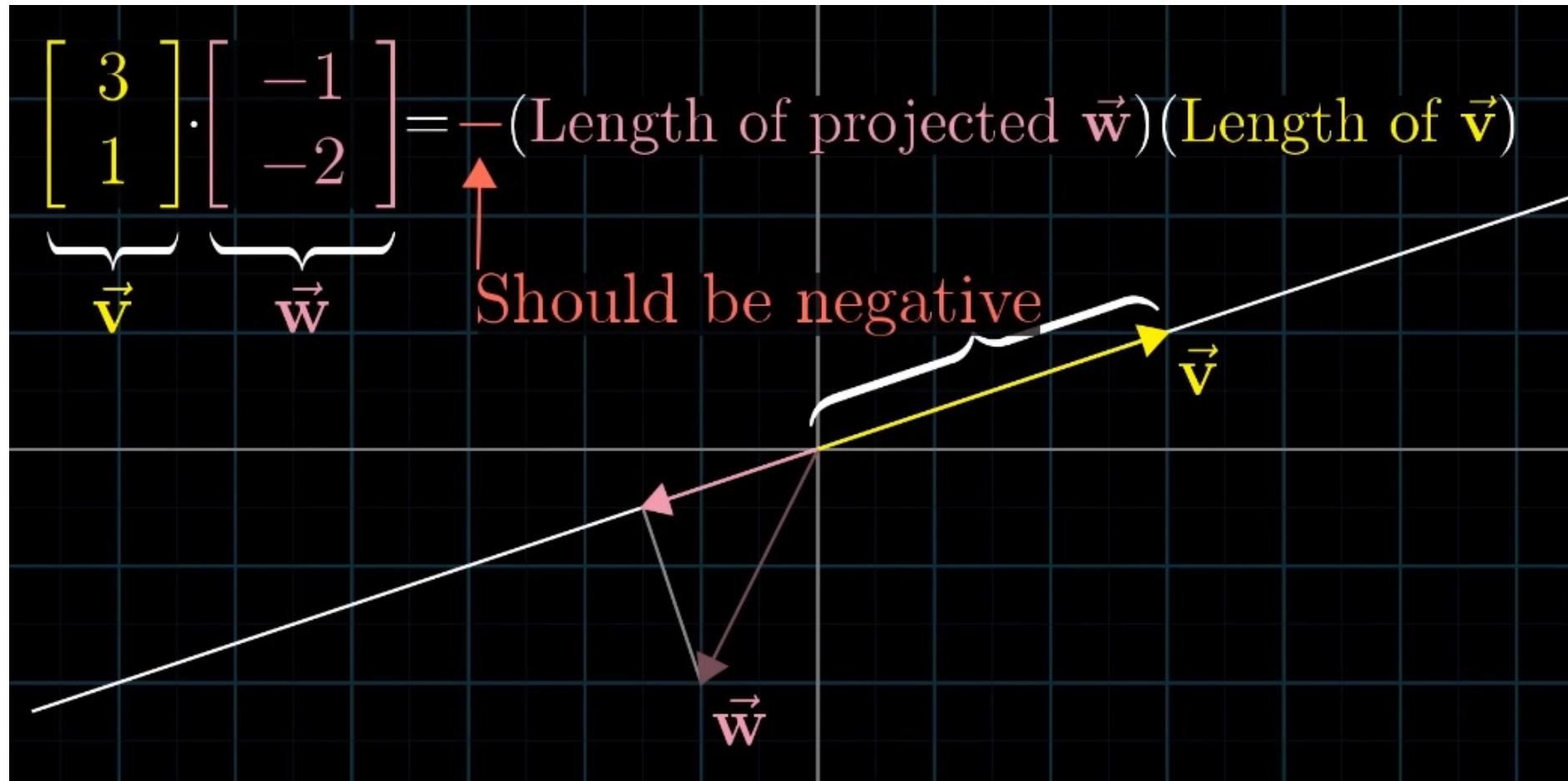
Geometric Interpretation

- This often accompanies the dot product's introduction.
- Projecting \vec{w} onto the line which passes through the origin and \vec{v} 's tip, (imagine shining a torch, where the projection is \vec{w} 's shadow)
- the dot product is equal to the length of that projection multiplied by the length of \vec{v} .



Negative dot product

When \mathbf{w} points in the opposite direction, the dot product will actually be negative however.



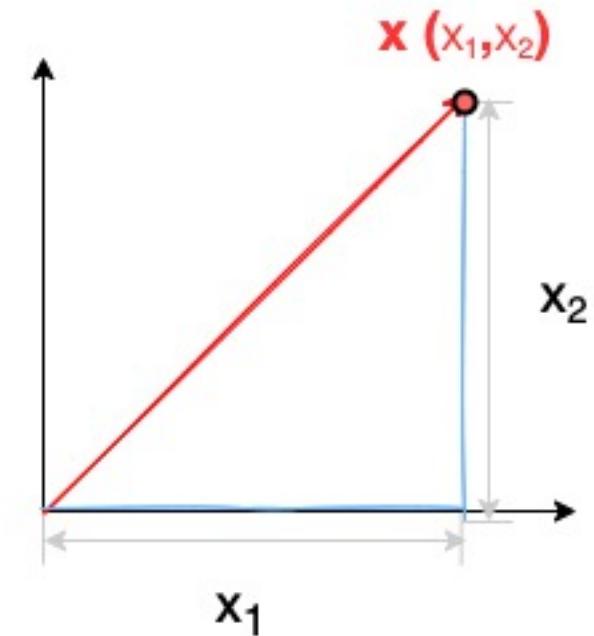
Vector Norm

- How can we calculate the length of a vector ?

- $\| \mathbf{a} \| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$

- Euclidian norm or L² norm

- Measures the shortest distance from the origin

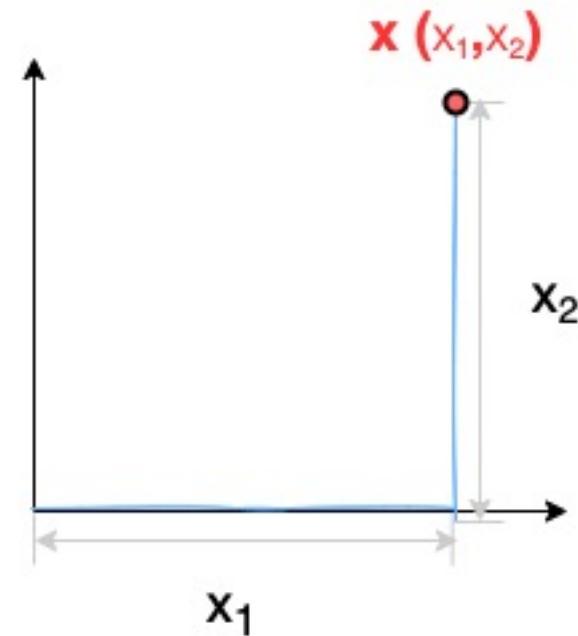


Vector Norm

- $\| a \| = |a_1| + |a_2| + \cdots + |a_n|$

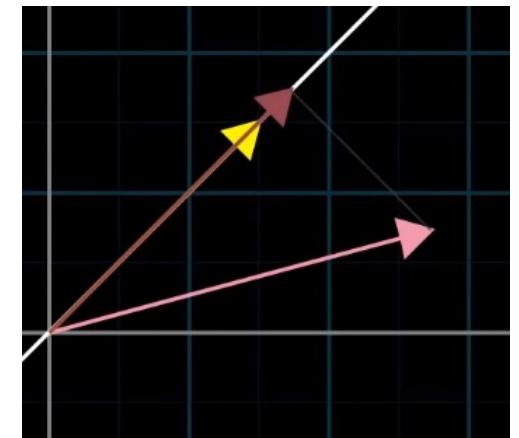
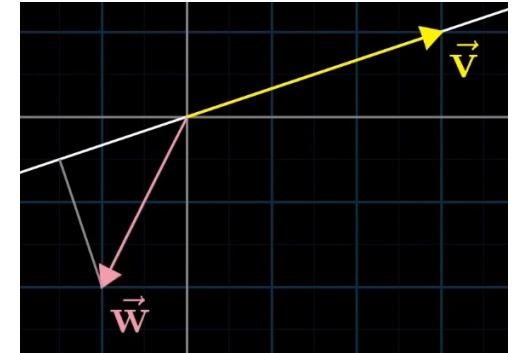
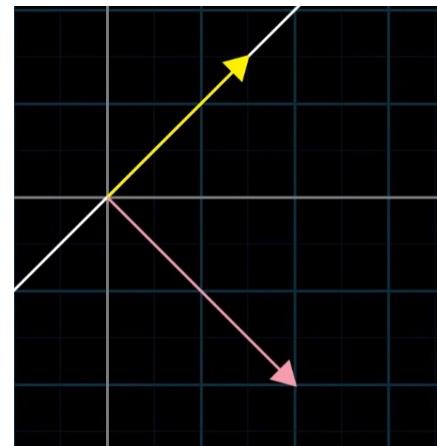
- L_1 norm or Manhattan norm

- Sum of the absolute values of the components of the vector



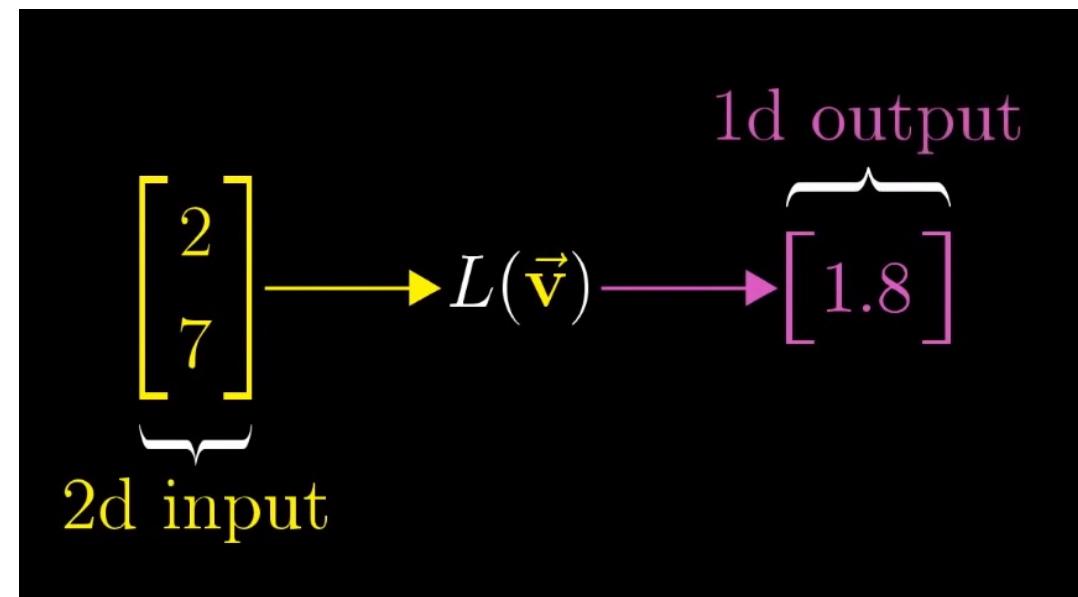
A couple important properties

- The dot product between 2 vectors is :
- < 0 , when the projection faces the opposite direction
- > 0 , when vectors point in the same general direction
- $= 0$ when perpendicular



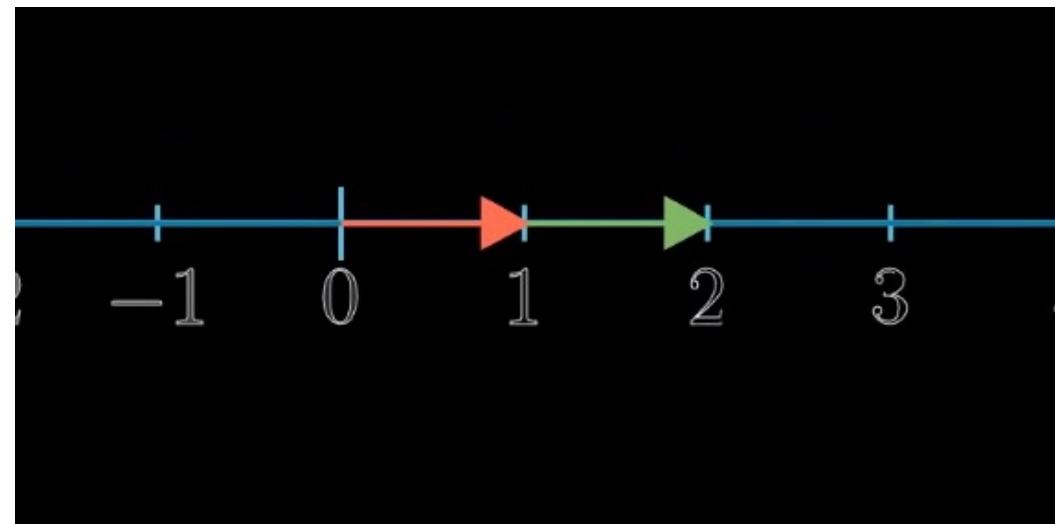
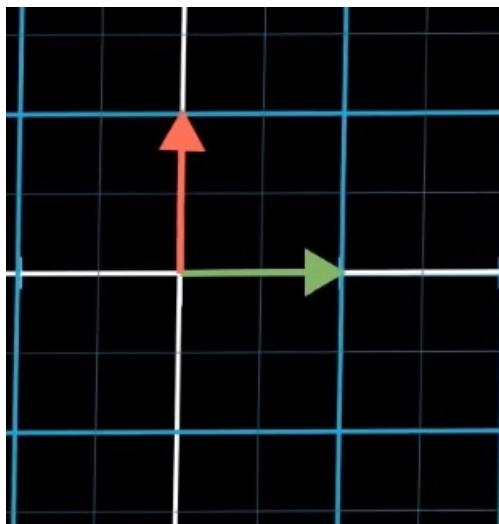
Back to linear transformations : Multi-D to 1D

- This type of transformation is a function which takes in a **vector** and spits out a **number**



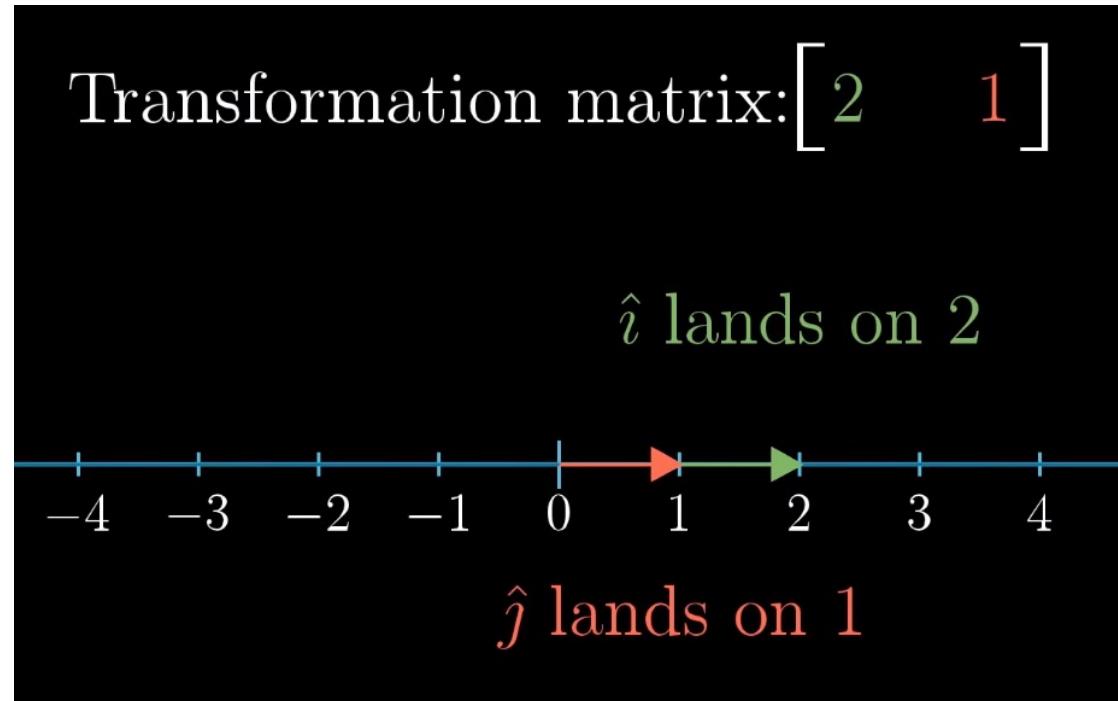
2D to 1D transformation

- Remember a transformation can be determined by looking at where \hat{i} and \hat{j} land...
- In this case, each basis vector just lands on a number vs. A point in 2D space.



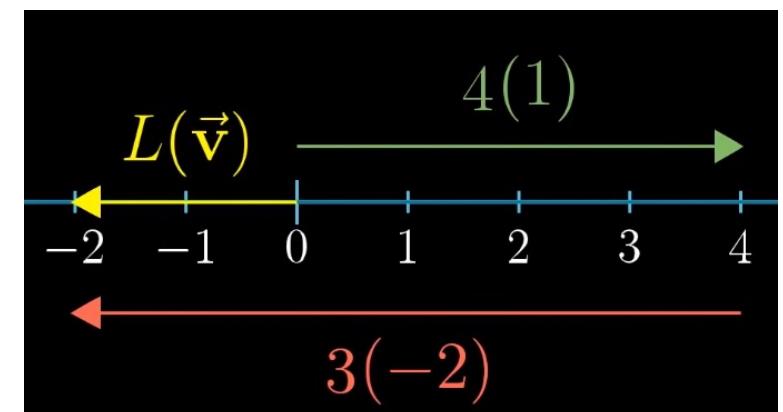
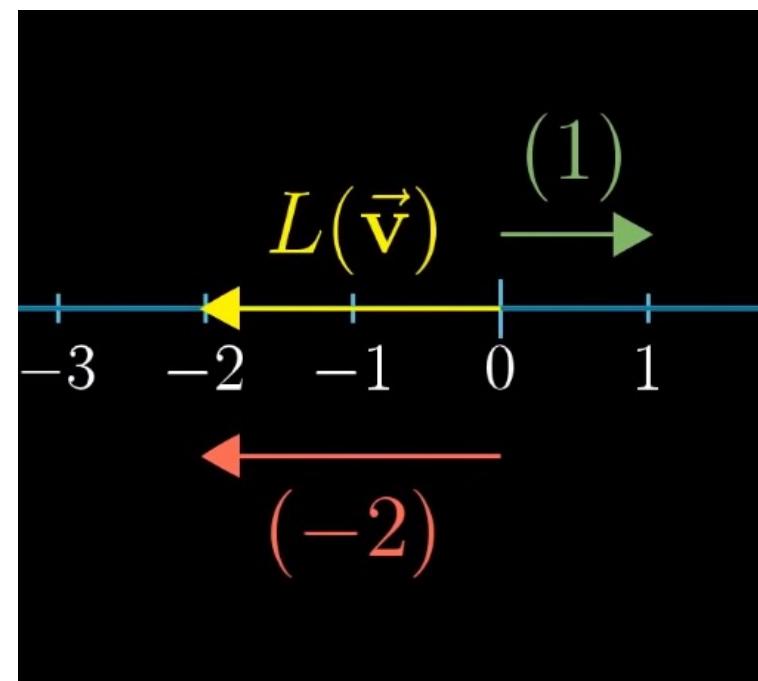
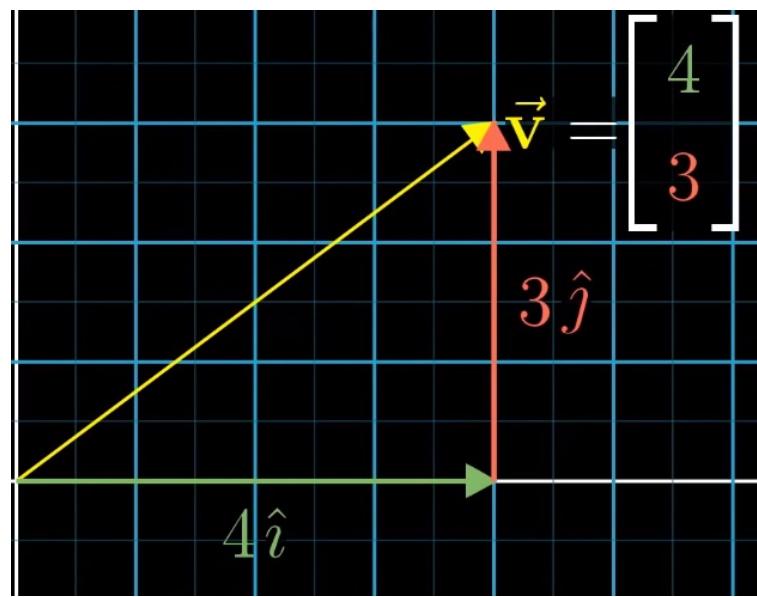
Matrix

- So when we record where these vectors land as the columns of the matrix
- Each column just has a single number !
- This is a 1×2 matrix ($\mathbb{R}^{1 \times 2}$)



Applying the transformation to a vector $\vec{v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$

- The vector is defined as a linear combination of the basis vectors,
- and a consequence of linearity is that
- The same linear combination applies
- Using the « landing spots » of \hat{i} and \hat{j} .



Matrix / vector operation vs. Dot Product

- Looking at things numerically, the matrix vector product in this case feels just like taking the dot product between 2 vectors !
- A 1×2 matrix just looks like a vector, but tipped on its side

Transform

$$\begin{bmatrix} 1 & -2 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \end{bmatrix} = 4 \cdot 1 + 3 \cdot -2$$

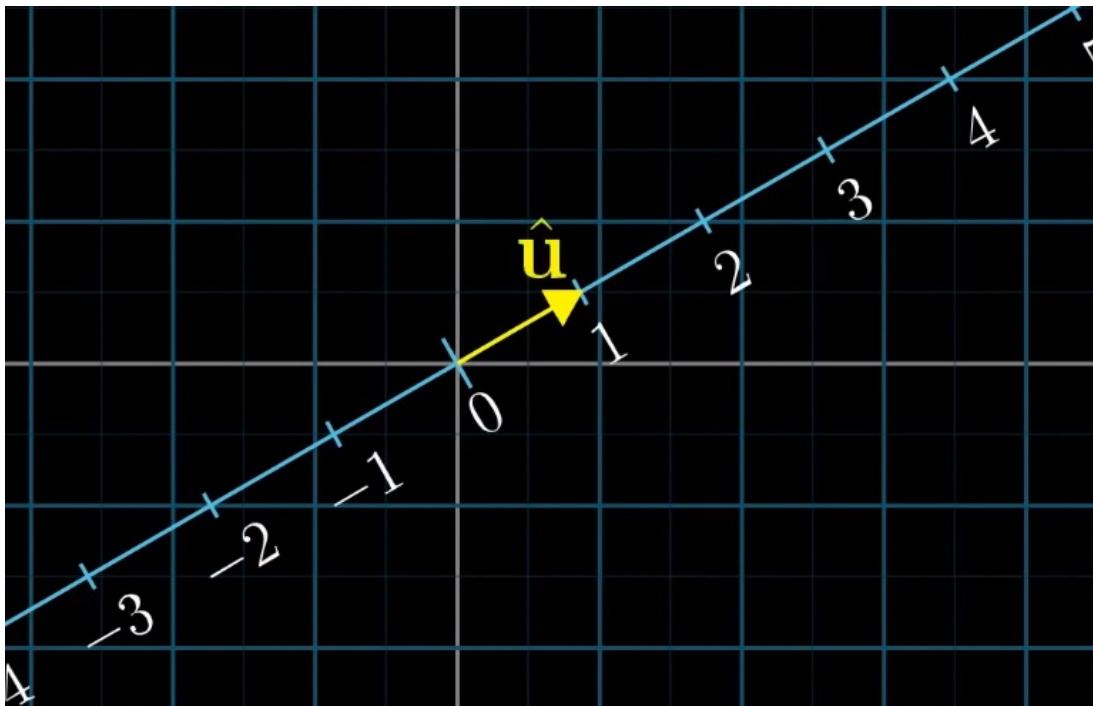
Vector

Dot product

$$\begin{bmatrix} 2 \\ 7 \\ 1 \end{bmatrix} \dot \begin{bmatrix} 8 \\ 2 \\ 8 \end{bmatrix} = 2 \cdot 8 + 7 \cdot 2 + 1 \cdot 8$$

What about the projection of one vector onto another ? (optional)

- One clever way to determine where \hat{i} and \hat{j} land is to take a copy of the number line and embed it diagonally in 2D space.
- Then add a 2D unit (length = 1) vector \hat{u} whose tip sits on the number 1 on the number line.

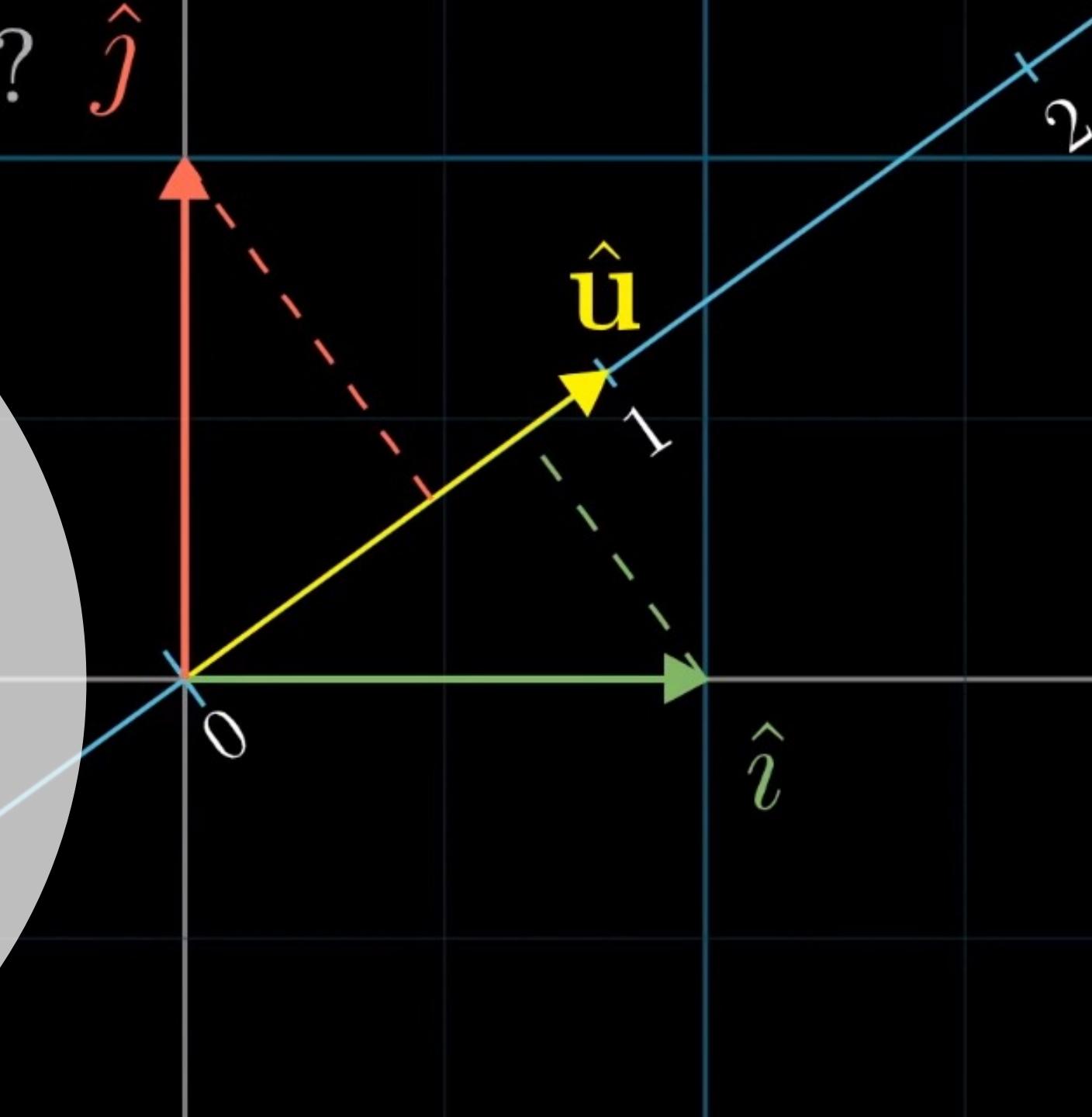


here do \hat{i} and \hat{j} land? \hat{j}

[]

Geometric Interpretation

- \hat{u} lives in the **input** space, it is just situated in such a way that it overlaps with the number line...
- So now how do we find a 1×2 matrix which defines a transformation such that a vector is « translated » onto the number line (into 1D).
- Need to find where \hat{i} and \hat{j} each land...

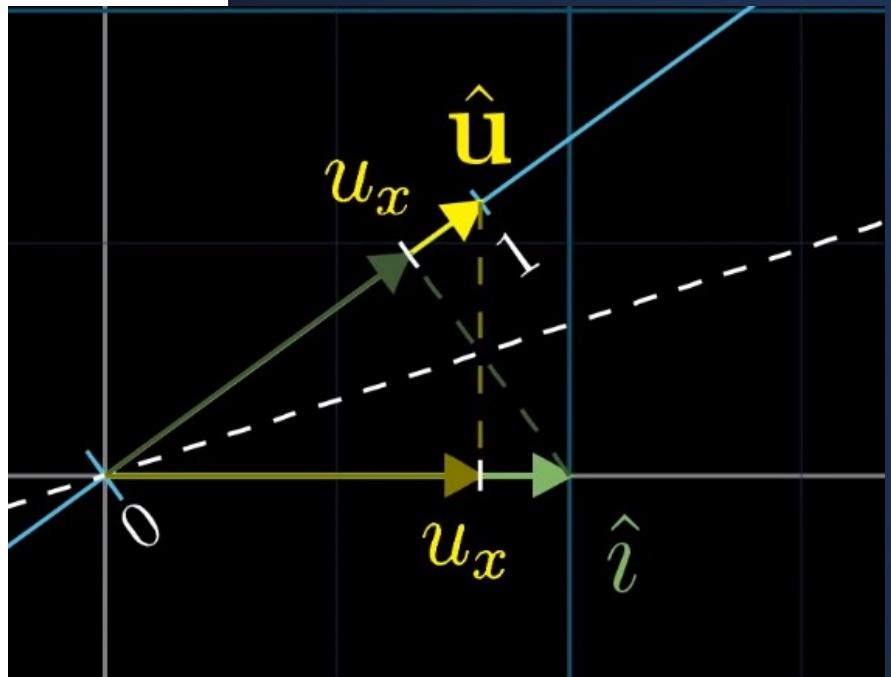


Geometric Interpretation

- This is where projections come in.
- Can reason by symmetry because \hat{i} and \hat{u} are unit vectors (both have length 1):
 - Projecting \hat{i} onto the line of numbers looks completely symmetric to projecting \hat{u} onto the x axis...
 - So where does \hat{i} land ? => Same place as \hat{u} on the x axis !

$$T(\hat{i}) = u_x$$

- The same can be done for \hat{j}



Geometric Interpretation

- So the entries of the 1×2 matrix describing the transformation are going to be the coordinates of \hat{u} ...

$$\begin{bmatrix} u_x & u_y \end{bmatrix}$$

- Therefore, projecting any vector $\begin{bmatrix} x \\ y \end{bmatrix}$ onto the unit vector \hat{u} is just the sum of the projections of \hat{i} and \hat{j} scaled by x and y :

Input space : $x\hat{i} + y\hat{j}$

Output space (number line) : $x T(\hat{i}) + y T(\hat{j})$

Geometric Interpretation

- What about when we project onto a non-unit vector ?
- If the vector we project onto has a magnitude $\neq 1$, then we just scale up that projection. This brings us exactly to our original geometric definition !
- This shows the close relationship between the dot product, projections and linear transformations.

