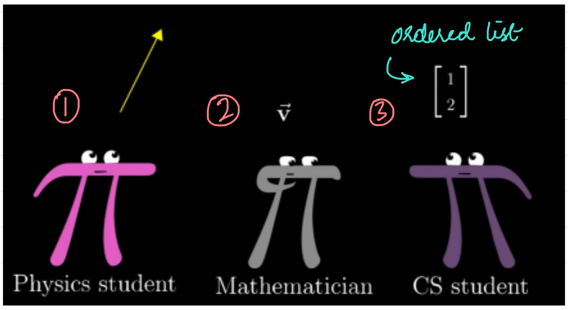
Vectors | Chapter 1, Essence of linear algebra + Math for ML book

↳ what are scalars ? real numbers or elements in the field.

↳ what are vectors ? 3 perspectives



↳ coordinate system & vectors rooted in origin.

↳ nD vector as a point in nD space

↳ vector addition
Δ law (11th law)

↳ multiplying vectors by scalar → scaling

↳ components of a vector
$a\hat{i} + b\hat{j}$

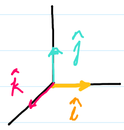
Linear combinations, span, and basis vectors | Chapter 2, Essence of linear algebra

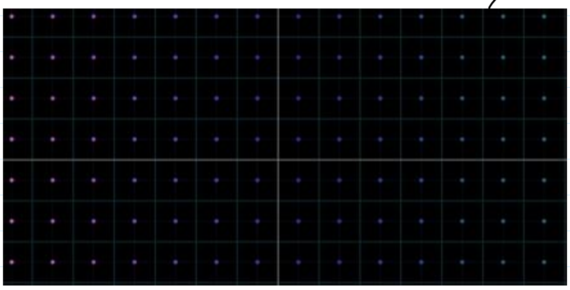
↳ think about the coordinates as scalars. & then the addition/subtraction/multiplication

(Basis aka unit vectors)

↳ special vectors $\hat{i}$, $\hat{j}$ are basis in 2D vector space

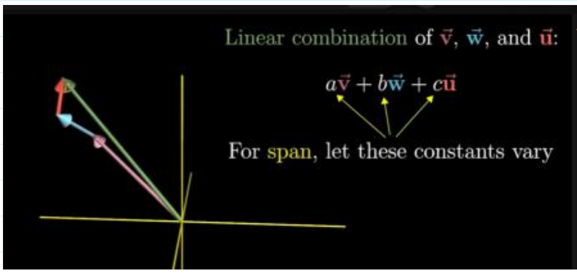
↳ why linear combination ? Reaching all possible vectors with pairs of 2D vectors is called span of those vectors.





think of collection of vectors as points in the vector space

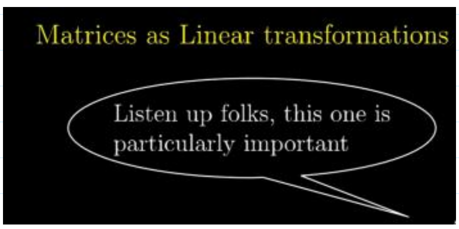
↳ span of
1 vector → line
2 LI vectors → plane
3 LI vectors → hyperplane



→ 3 linearly independent (LI) vectors span 3D space

→ we need n LI vectors to span nD space

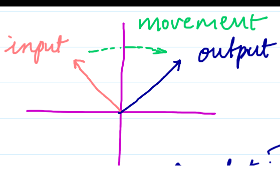
Linear transformations and matrices | Chapter 3



↳ matrix is best understood by vizualization.

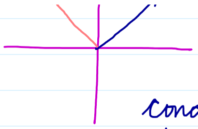
↳ matrix vector multiplication or linear transformation

vector input → fuction → vector output.

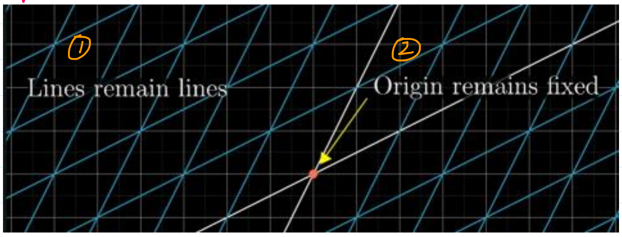


↳ imagine movement of all vectors in the space

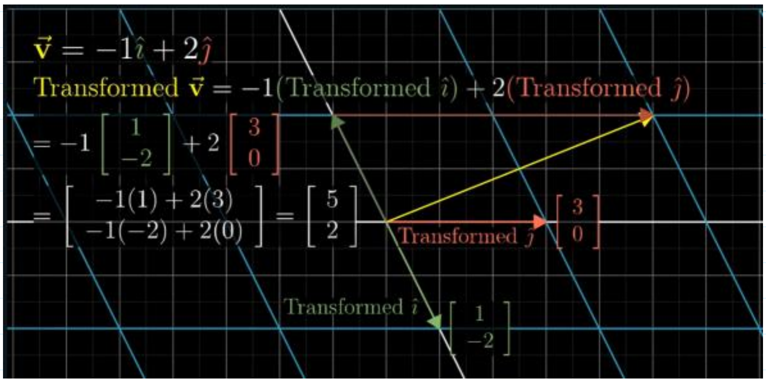
space



conditions for linear transformation

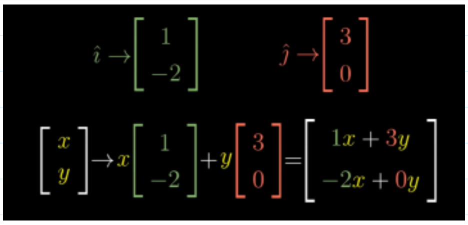


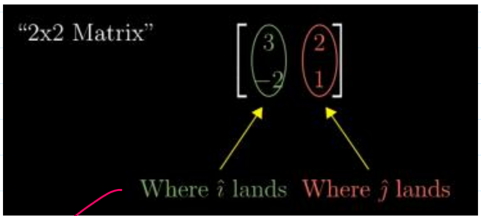
③ ↳ grid lines must be parallel and equally spaced.



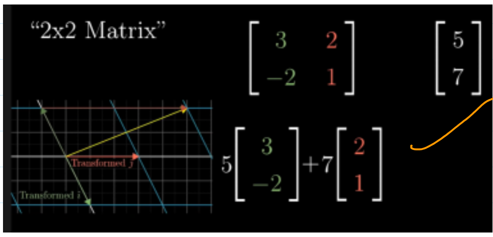
← to understand the transformation, follow the basis vectors. How their positions change? where do they land?

→ pluging these transformed values of $\hat{\imath}$ & $\hat{\jmath}$ in the original vector equation will give us the output vector.



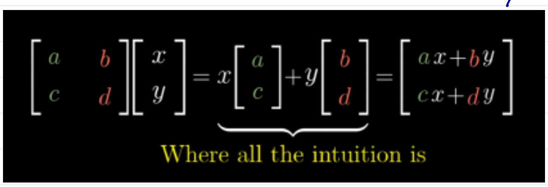


↳ transformed unit vectors.



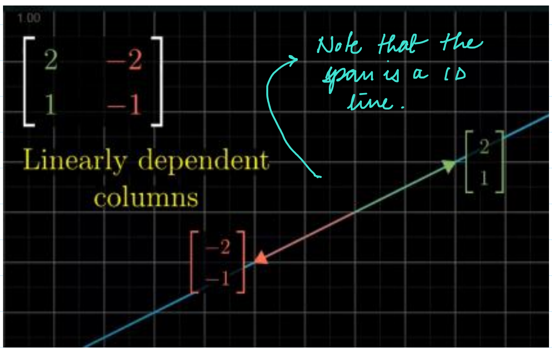
→ scaled version of basis vectors } linear combination ⇒ output vector.

→ generalization



eg: rotation by 90°
shear

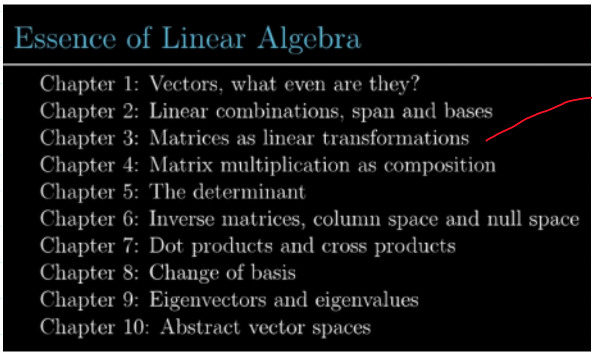
↳ we can also start from a matrix & deduce its transformation.



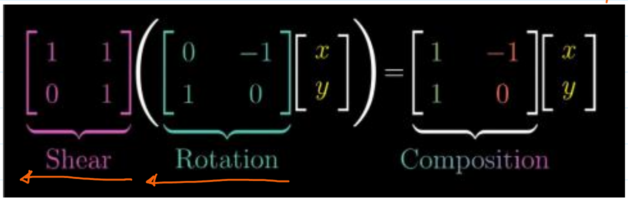
Note that the span is a 1D line.

→ Linearly dependent vectors can not span the nD vector space

Matrix multiplication as composition | Chapter 4



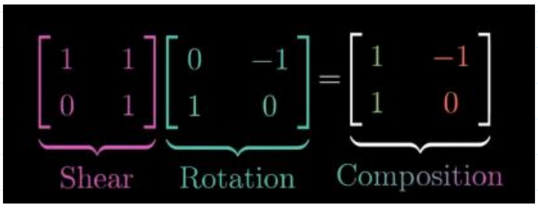
All we learnt in video 4 is the underlying concept of following videos.



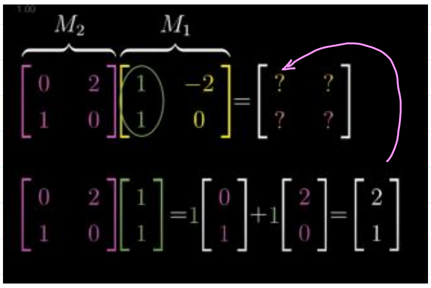
applying rotation & then shear

(read from right to left)

similar to functions $\underset{②}{f}(\underset{①}{g}(x))$



where $\hat{i}$ goes →



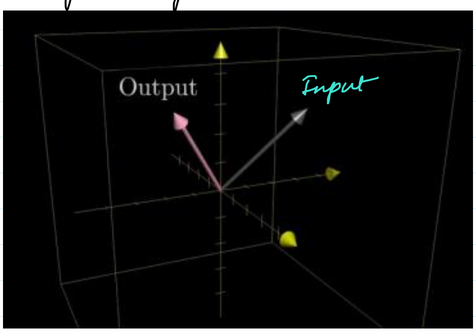
mat mult^n is just composition of transformations and get a new transformation.

Note, matrix mult^n is associative

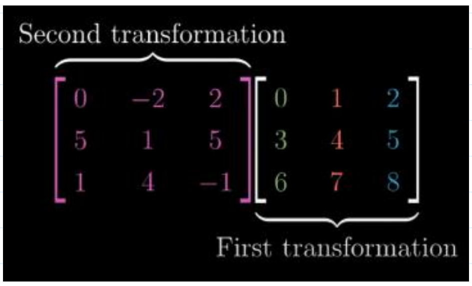
$M_1(M_2 M_3) = (M_1 M_2)M_3$ → consider as applying transf^n one after the other.

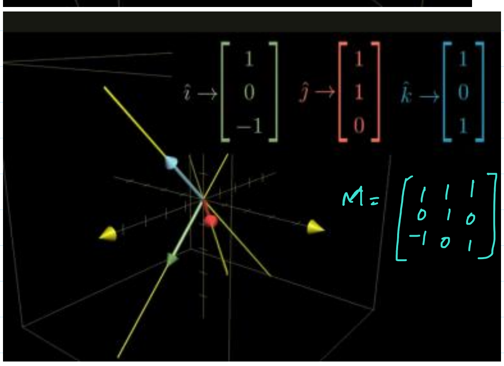
Three-dimensional linear transformations | Chapter 5

↳ in previous two videos we saw linear transformations in 2D they easily translate to higher Ds

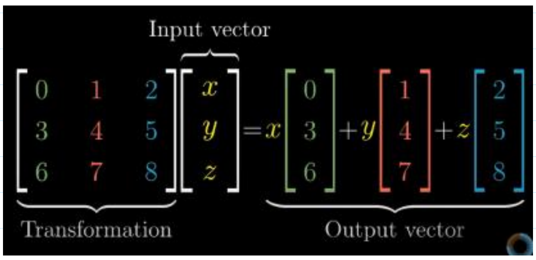


→ as discussed earlier, we see the transf^n of unit vectors now there are 3 ($\hat{i}, \hat{j}, \hat{k}$)





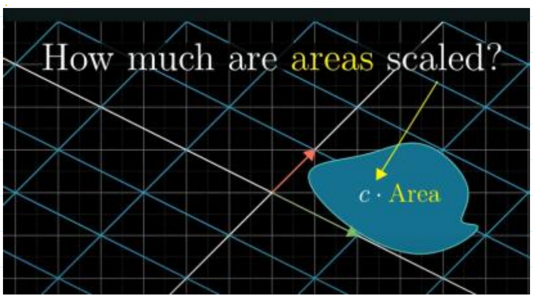
$$M = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$



↳ composition of transf^n is also the same as seen.

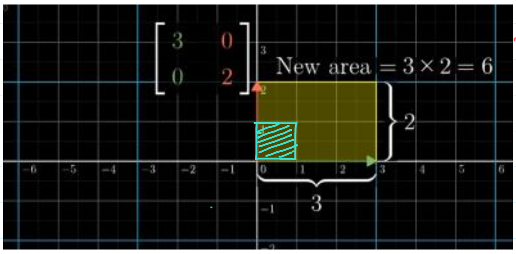
The determinant | Chapter 6

↳ transformations generally stretch or squeeze the space.



↳ Note: since the grid lines are parallel and equispaced, the transf^n applied to one square is translated to all the squares in the space

↳ the scaling factor of the area is determinant. (linear transf^n) → matrix



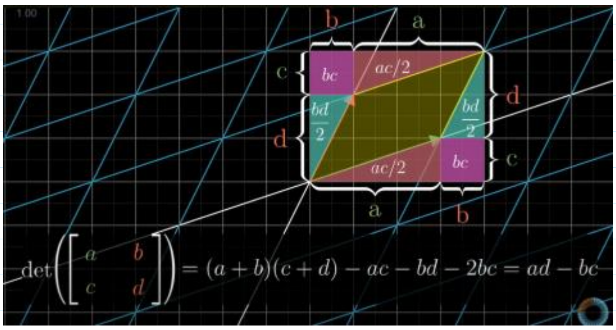
→ unit square scaled by a factor of 6.

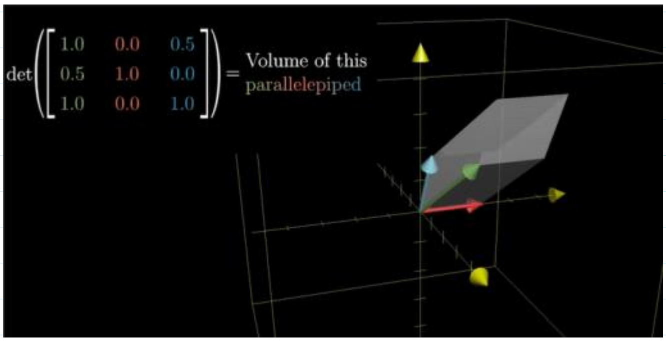
→ if scaling factor is -ve it rotates / changes orientation of the space.

↳ In 3D we get volume instead of area (how?)

↳ right hand rule can tell the orientation.

Note: when we rotate the space $\hat{i}$ & $\hat{j}$ come closer and at some point overlap. Here the det is 0. And if we go further, it becomes -ve.





Inverse matrices, column space and null space | Chapter 7

↳ Linear algebra has its use mainly because it helps to solve linear system of equations

$$A\vec{x} = \vec{b}$$

we are looking for the vector x which when transformed by A, it lands on $\vec{b}$.

↳ any transf^n that undos the effect of a transf^n is called inverse of that transf^n.

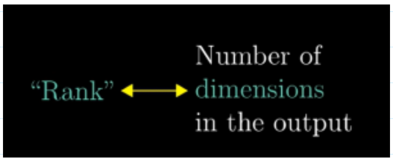
↳ the solution depends on the det(A) does it ↓ dims or leave as is.

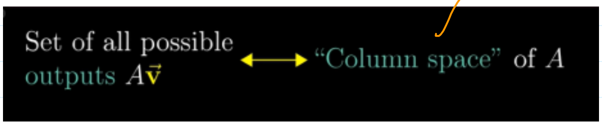
↳ when # variables = # equations then probably we have unique solution

↳ only when det(A) ≠ 0 we have A^{-1} as det(A) = 0 squeezes the space to a

↳ only when det(A) ≠ 0 we have A^{-1}
as det(A) = 0 squeezes the space to a
lower dim. → we can not undo
to go to higher dim. At least not by
using a function. as it can take single i/p
and single o/p at a time

→ zero area (2D) or zero volume (3D)

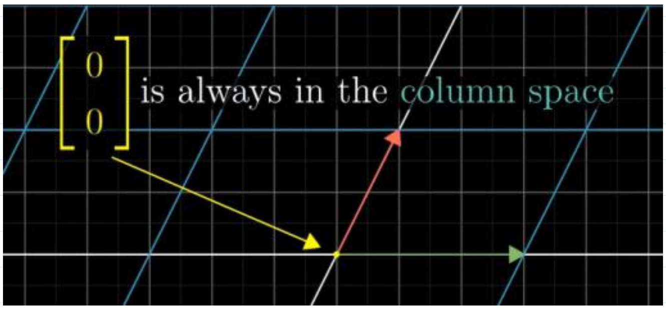




Basis vectors (transformed)

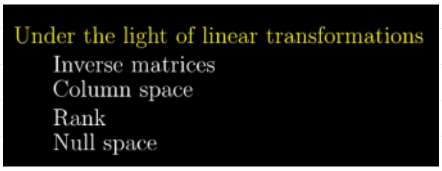
→ span of columns of matrix

dims in column space
= rank



↳ in case of full rank,
only 0 vector itself can
land on origin.

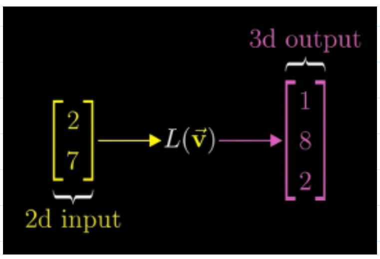
↳ but when rank is not full,
dim is reduced set of
many vectors can pass the origin
This set is called null space
or kernel of the matrix.



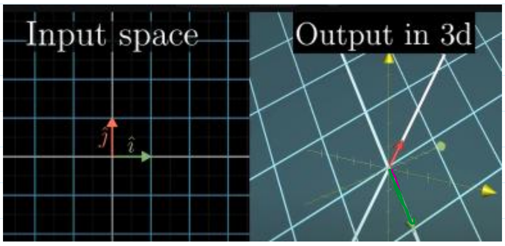
Non-square matrices as transformations between dimensions | Chapter 8

↳ so far we have discussed linear transf^n
in 2D, 3D.

→ transf^n b/w dims



↳ for non square transf^n,
again think about
the unit vectors,
where do they land after
the transf^n is applied?

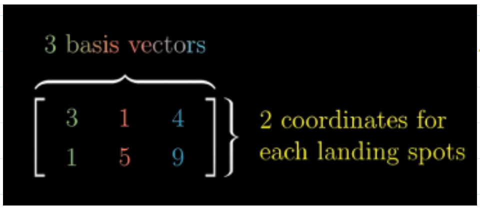


$$\begin{bmatrix} 2 & 0 \\ -1 & 1 \\ -2 & 1 \end{bmatrix}$$

span of columns
$\Updownarrow$
column space

modif $\hat{\imath}$ modif $\hat{j}$

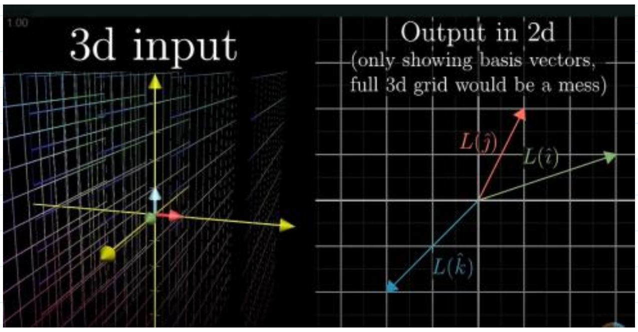
matrix is full rank



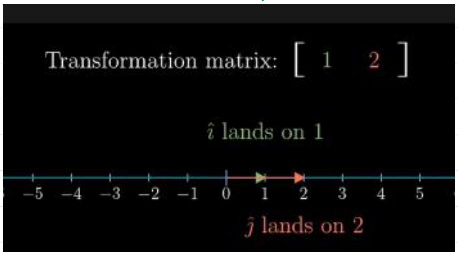
→ transf^n from 3D to 2D

similarly, we can think of 2D to 1D transf^n

a number on number line



ties to dot product



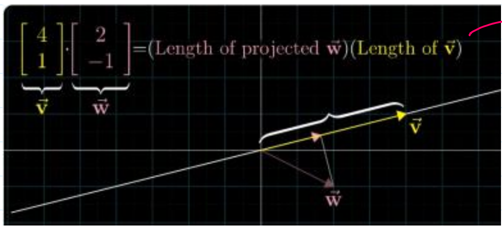
Continued in page 2 ...

2. Video 9- 16

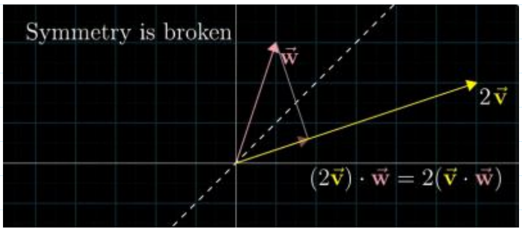
28 October 2024 10:24

Dot products and duality | Chapter 9

↳ must understand transf^n first

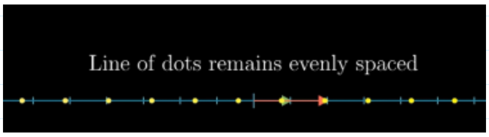


→ either of the vectors can be projected & multiplied in any order

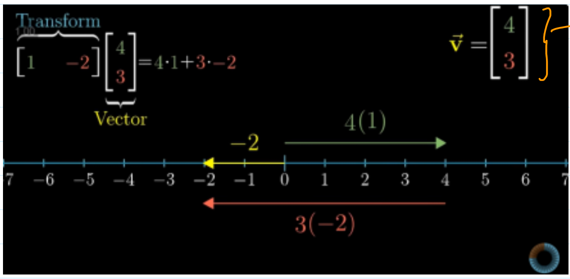


→ think about the case when $\vec{v}$ & $\vec{w}$ are of same length here symmetry holds but on scaling either vector the respective projection is also scaled by the same factor.

linear transf^n from nD to 1D



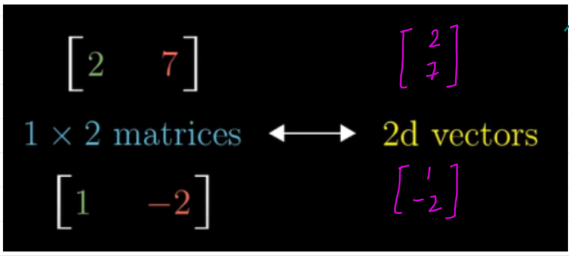
→ if transf^n is not linear, then the dots will be unevenly spaced



→ lets' track basis vectors the $\hat{i}$ & $\hat{j}$ where they land when $\begin{bmatrix}4\\3\end{bmatrix}$ is transformed from 2D to 1D using $[1,-2]$

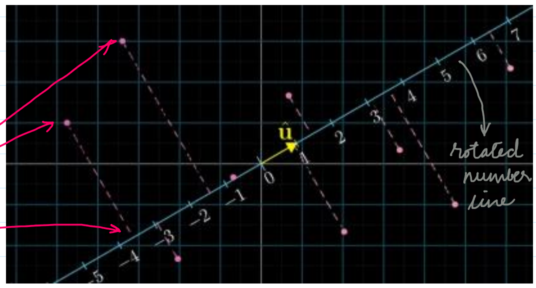
$\hat{i}$ lands on 1 and → scales by 4

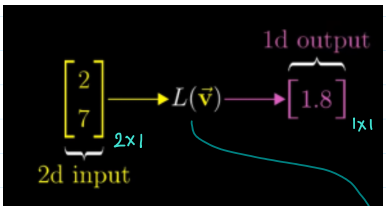
$\hat{j}$ lands on −2 ⟶ scales by 3



→ what does this association mean geometrically?

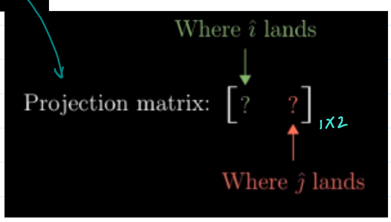
↳ linear transf^n that takes numbers to vectors & vise versa



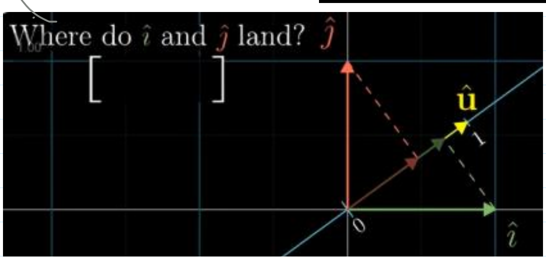


2D vectors

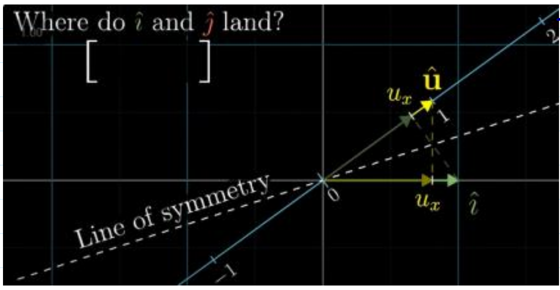
projected on the number line



→ these will be the columns of transf^n matrix (projection)



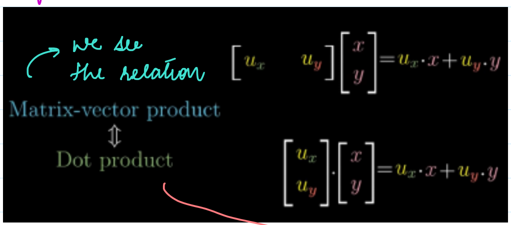
→ Because of symmetry, note that $\hat{i}$ projection on $\hat{u}$ ie u_x is same as

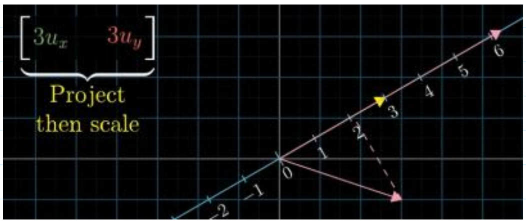


note that î projection on û ie u_x is same as û projection on î ie u_x

similarly for ĵ we get u_y

projection matrix = $[u_x \ u_y]$



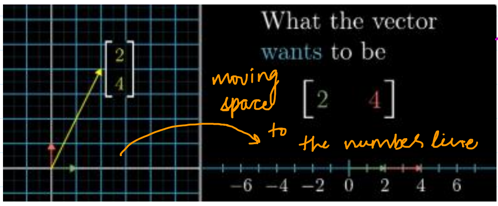


→ for non unit vectors.

this is concept of 'duality'. Anytime we have a 2d to 1d linear transf^n there exists a unique vector such that applying the transf^n is same as taking dot product with that vector.

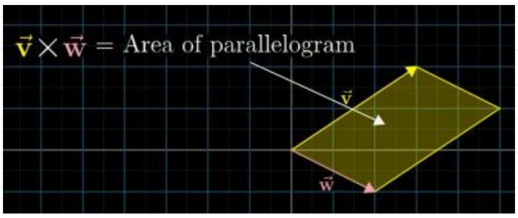
duality ⟷ natural but surprising correspondence.

To sum up, dot product is tool to know
- direction of vectors wrt one another
- how the vector transforms (how it is & what it wants to be)



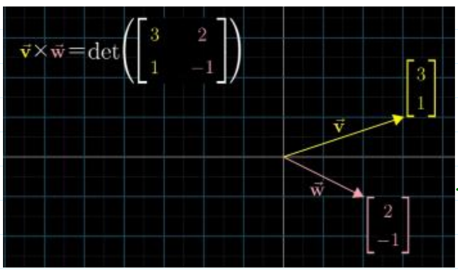
→ transf^n on arrows in space
vs
moving the space

Cross products | Chapter 10



→ order of basis vector defines orientation

→ when $\vec{v}$ & $\vec{w}$ is swapped, the area is -ve.



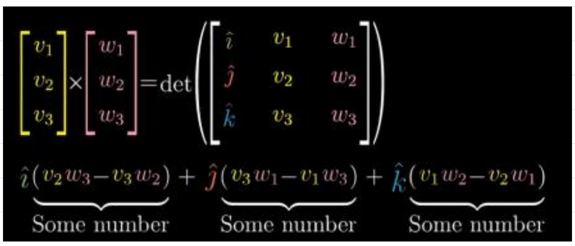
→ recall that determinant (linear transf) measures how the area of unit square changes.

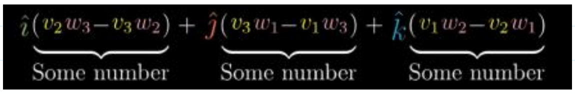
→ thus when columns are interchanged, the sign of area (det) also changes

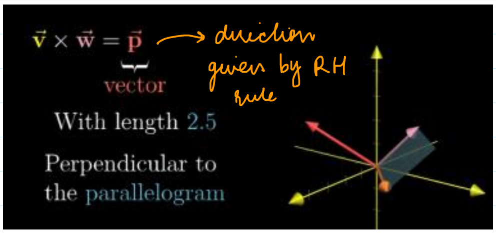
→ $(3\vec{v}) \times \vec{w} = 3(\vec{v} \times \vec{w})$

→ when vectors are ⊥ area is max^m

↪ we use right hand rule to find direction of the result (cross product)



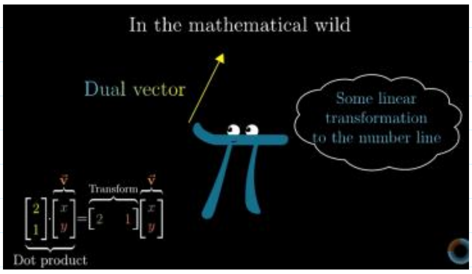


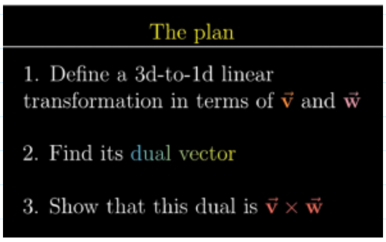


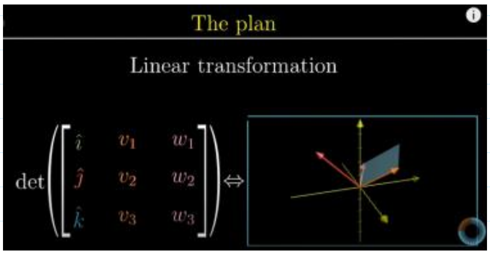
→ length of $\vec{p}$ is area of ||gm

Cross products in the light of linear transformations | Chapter 11

Recall

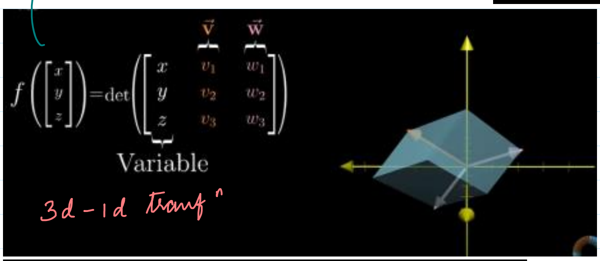






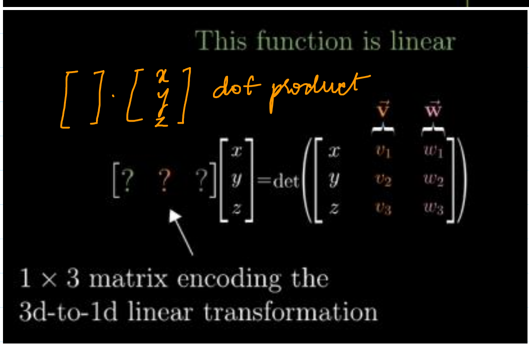
Recall the 2d cross product & compare with 3d cross prod.

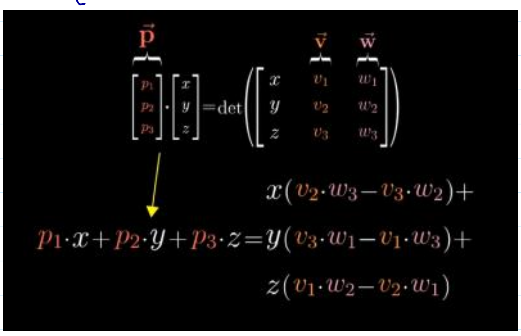
consider this function

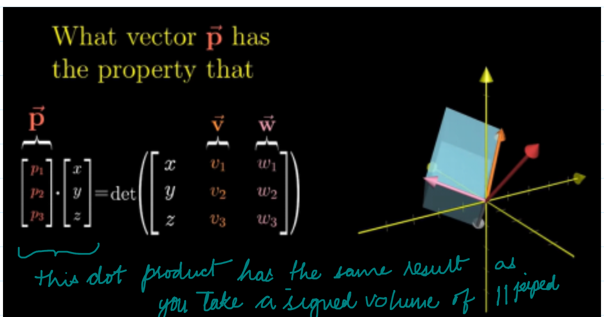


→ compare the coefficients to get p_1, p_2, p_3

the x, y, z are comparable to $\hat{\imath}, \hat{\jmath}, \hat{k}$ as we saw earlier.





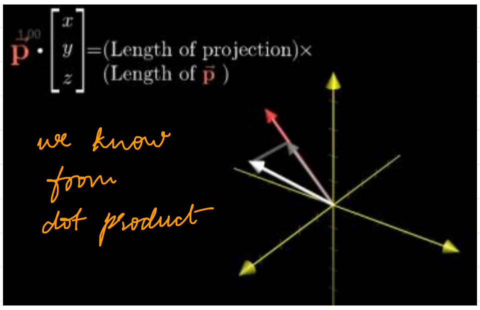


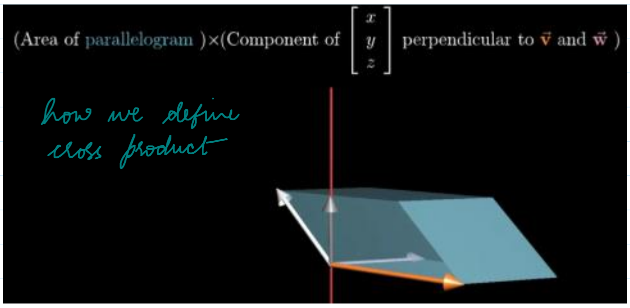
defined by this vector $\begin{matrix} x \\ y \\ z \end{matrix}$ along with $\vec{v}$ & $\vec{w}$

→ if we choose the appropriate direction of $\vec{p}$ the cases will line up where the area of ||gm is -ve. (the orientation is given by RH rule)

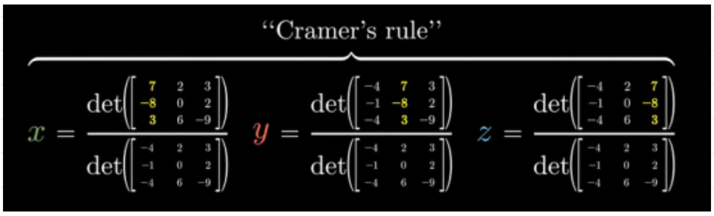
→ this is the same thing as taking a dot product of $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and

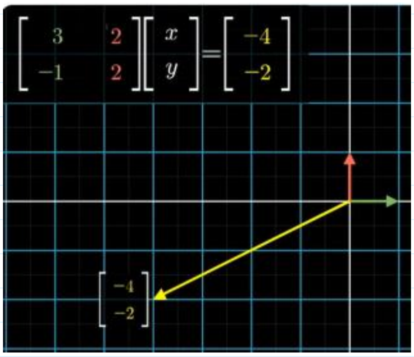
this is the same thing as taking a dot product of $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and a vector $(\vec{p})$ that is $\perp$ to $\vec{v}$ & $\vec{w}$ and length is equal to the area of ||gm.

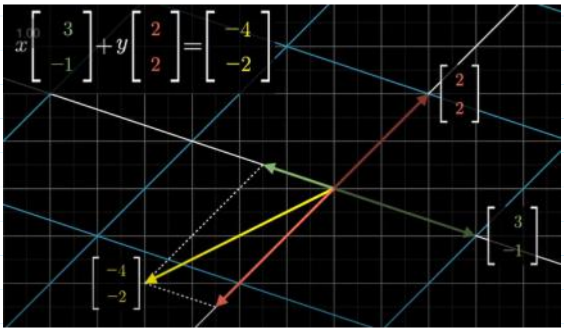


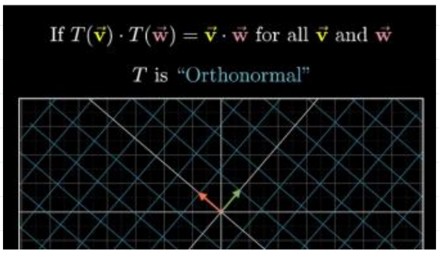


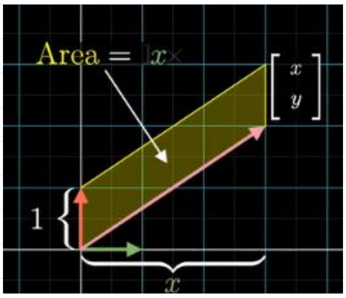
Cramer's rule, explained geometrically | Chapter 12



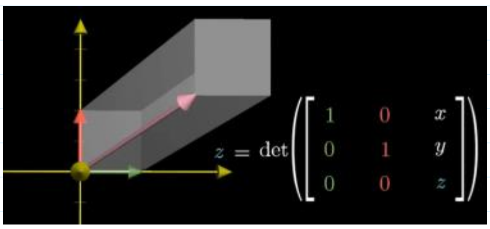


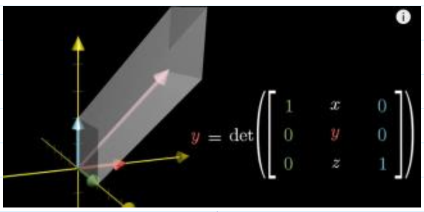


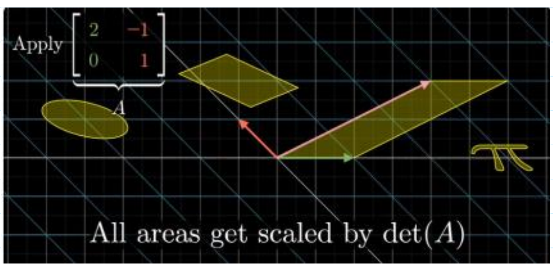




To preserve the properties of resultant of dot product of two vectors, the transf^n must be Orthonormal.

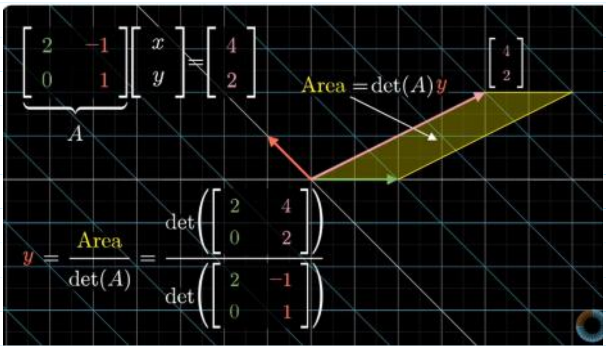


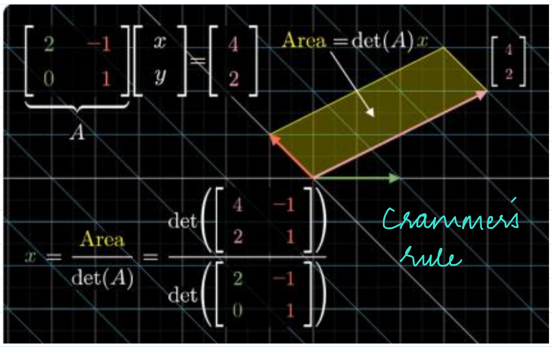




similarly,

$$x = \det\left(\begin{bmatrix} x & 0 & 0 \\ y & 1 & 0 \\ z & 0 & 1 \end{bmatrix}\right)$$

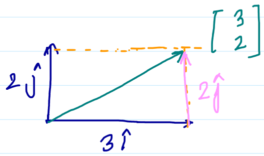


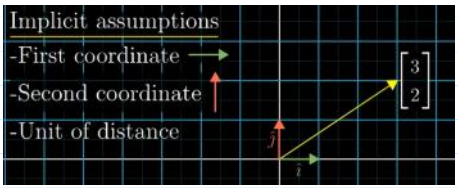


$$3x + 2y - 7z = 4$$
$$1x + 2y - 4z = 2$$
$$4x + 0y + 1z = 5$$
$$\begin{bmatrix} 3 & 2 & -7 \\ 1 & 2 & -4 \\ 4 & 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} x \\ y \\ z \end{bmatrix}}_{\text{Mystery input}} = \begin{bmatrix} 4 \\ 2 \\ 5 \end{bmatrix}$$

→ take for example the z coordinate of input as the volume spanned by $\hat{\imath}$, $\hat{\jmath}$ and the input vector
what happens to this volume after the transf^n. What ways can you find the volume?

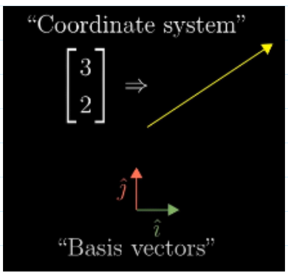
Change of basis | Chapter 13



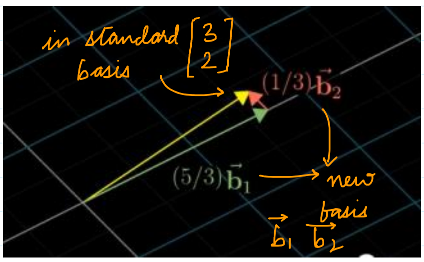


why do we care about change of basis?
↳ we see this in eigen values & eigen vectors

↳ any way to translate b/w set of numbers & vectors is called coordinate system



→ we now introduce the idea of using different standard vectors.

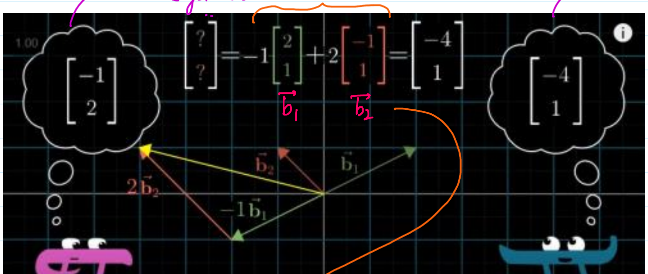


$\hat{\imath} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $\hat{\jmath} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

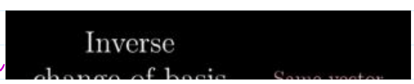
and $\vec{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$\vec{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ in new system $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ & $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

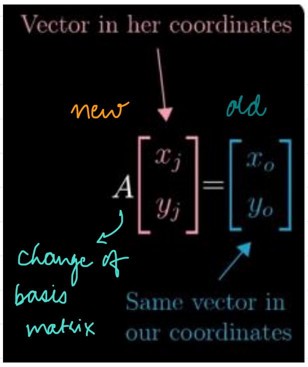
new system matrix multip^n standard system

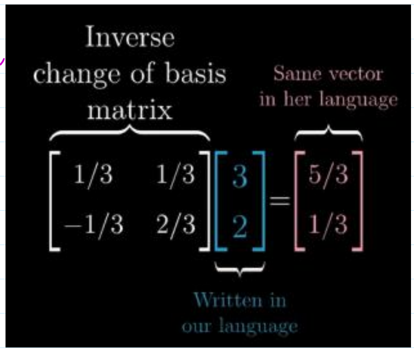


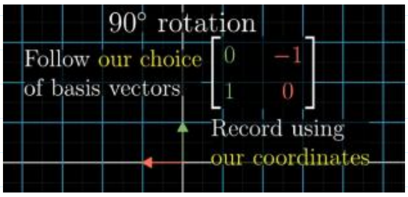
$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} \\ \end{bmatrix}$ vector in $\hat{\imath}\,\hat{\jmath}$ system

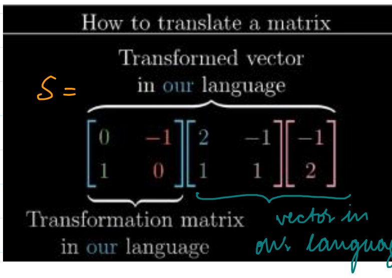


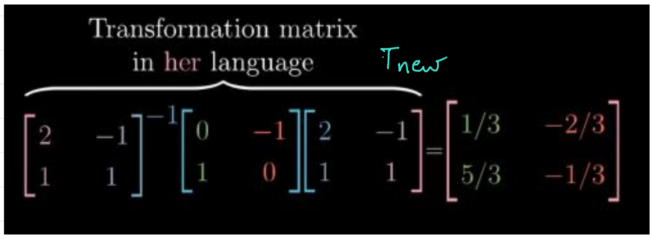
$$\underbrace{\begin{bmatrix}2 & -1\\ 1 & 1\end{bmatrix}}_{\text{transf}^n}\underbrace{\begin{bmatrix}-1\\ 2\end{bmatrix}}_{\text{vector in new system}} = \begin{bmatrix} \ \\ \ \end{bmatrix}$$ vector in $\hat{i}\,\hat{j}$ system









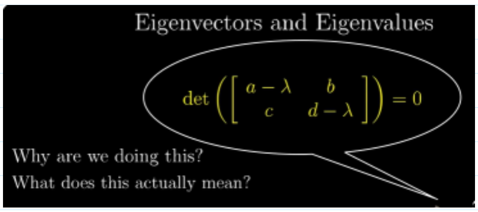


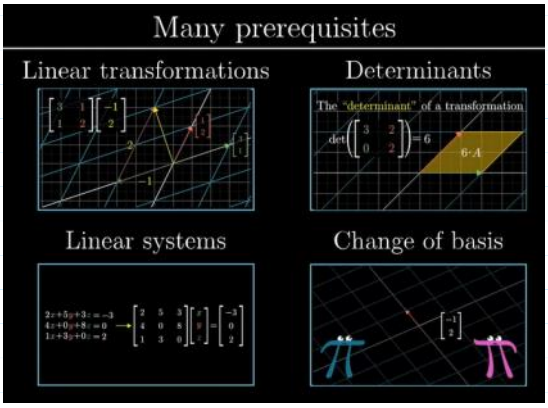
transformed vector in new language $= \begin{bmatrix}2 & -1\\ 1 & 1\end{bmatrix}^{-1} S$

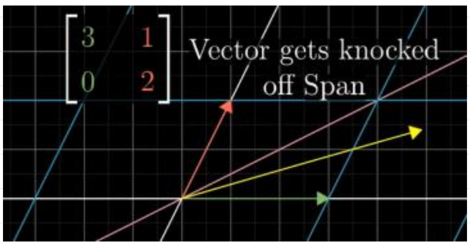
$\vec{V}_{new} = T_{new}\,\vec{V}$ — a vector in new language

transformed vector in new language

Eigenvectors and eigenvalues | Chapter 14

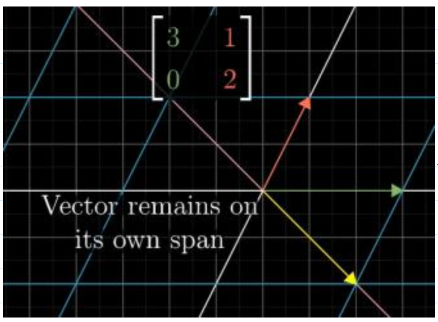




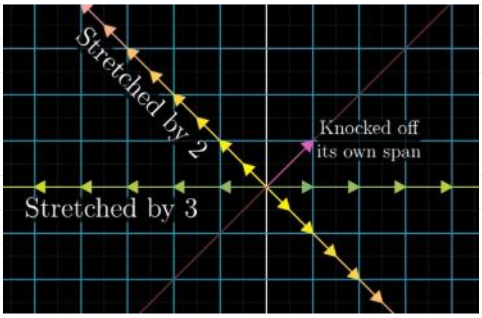


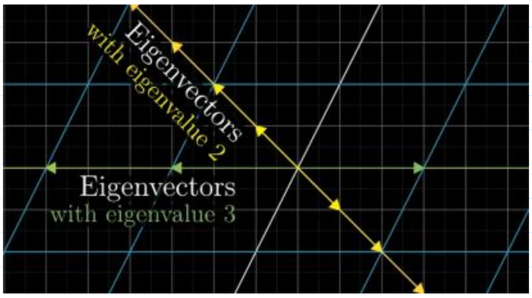
The concept itself is easy.

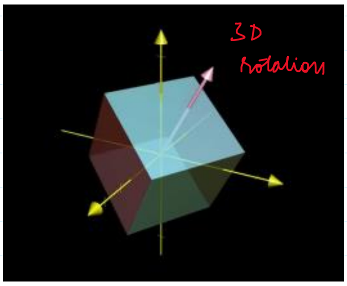
the concept of eigen value/vectors is hard to understand due to gap in prerequisites



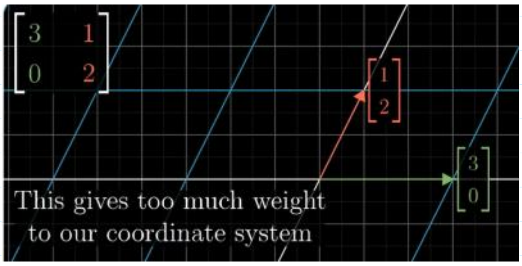
→ a transformation such that the span of vectors is preserved



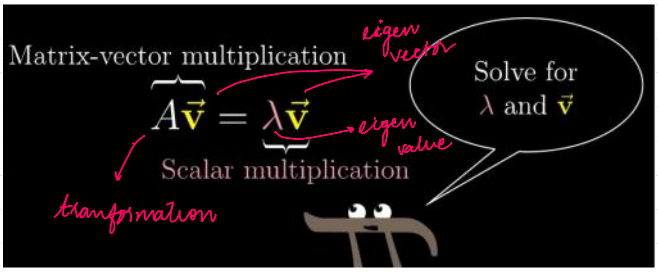




for 3D rotation, eigen vector is the axis of rotation & eigen value is 1 since there is no stretch or squeeze in case of rotations



→ what could be the alternate way, where we don't rely on coordinate system much?



LHS is matrix vector multipⁿ and RHS is scalar vector multipⁿ. How are they equal?

multiplⁿ by matrix

$\lambda \longleftrightarrow A$

scaling by λ

$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$ columns of matrix is λ times unit vectors $= \lambda(I)$

so,

$A\vec{v} = (\lambda I)\vec{v}$ now both LHS & RHS are matrix vector multipⁿ

$(A - \lambda I)\vec{v} = \vec{0}$

↑ we want non zero $\vec{v}$ (non trivial solution)

that means det $(A - \lambda I)$ must be zero

↳ (only then product of a matrix & non zero vector is zero)

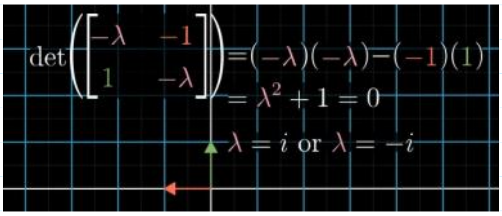
this means the transfⁿ squishes the space to lower dim.

→ to a line ie area = 0

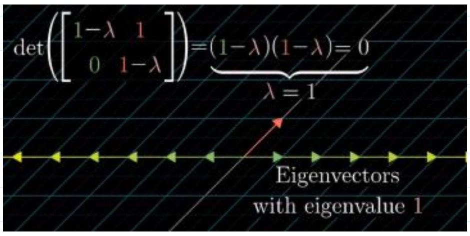
that corresponds to zero det.

so, we need to find such a λ.

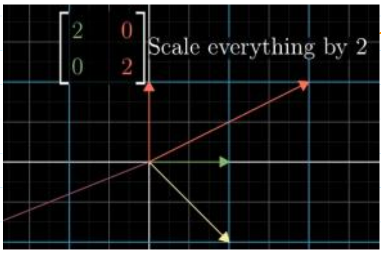
this vector v is eigen vector staying in its own span after transformation A



→ reason why rotation has no eigen values



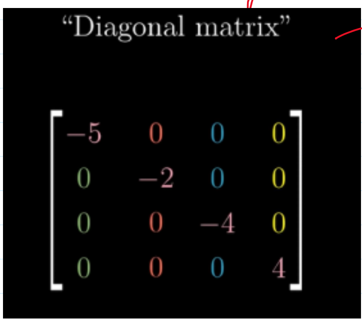
→ consider shear
it has only one eigen value
all eigen vectors have this eigen value



→ a single eigen value can have more than a line full of eigenvectors

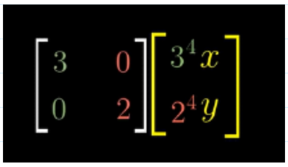
→ the eigen value is 2 but every vector in the plane is eigenvector

Eigen basis, imagine basis vectors to be the eigen vectors
Think of a diagonal matrix

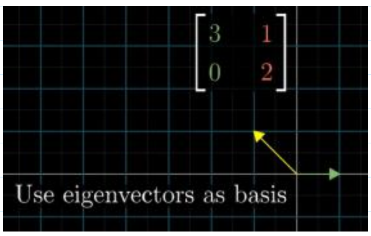


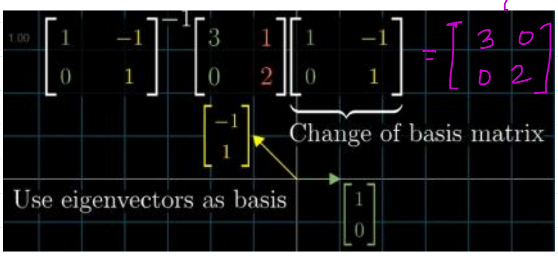
→ each column is eigen basis
and the diagonal entries are eigen values

↳ a set of basis vectors which are also eigen vectors are called eigen basis



→ $\begin{bmatrix}3^5 & x\\ 2^5 & y\end{bmatrix}$ can go to any number of repeated multiplication



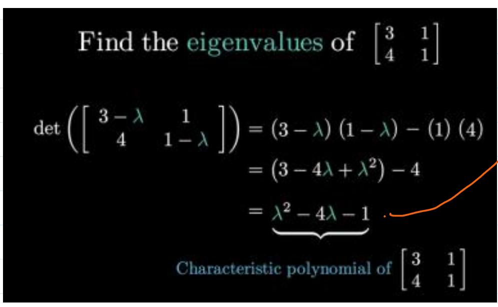


$= \begin{bmatrix}3 & 0\\ 0 & 2\end{bmatrix}$

guaranteed to be diagonal as the basis vectors get scaled during transf^n

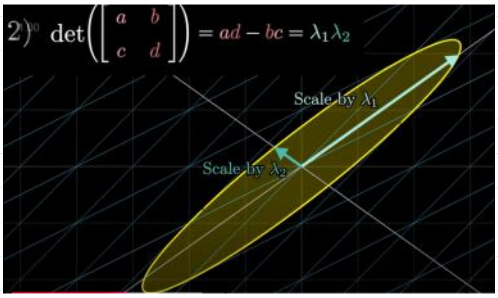
---

A quick trick for computing eigenvalues | Chapter 15

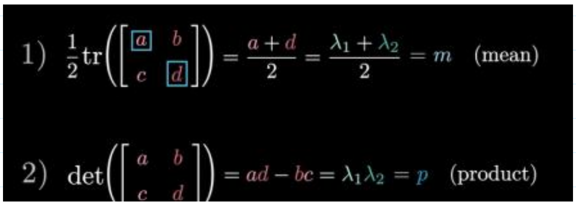


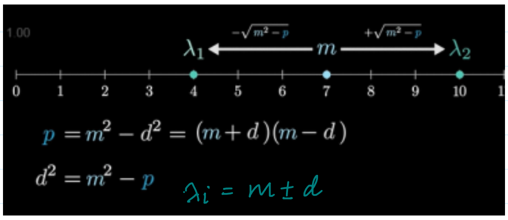
→ Roots of this eq^n are the eigen values

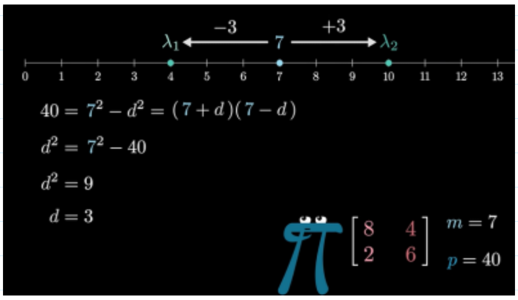
→ area scaled by



→ area scaled by $\lambda_1 \lambda_2$ so
product of eigen values = det (A)







Abstract vector spaces | Chapter 16