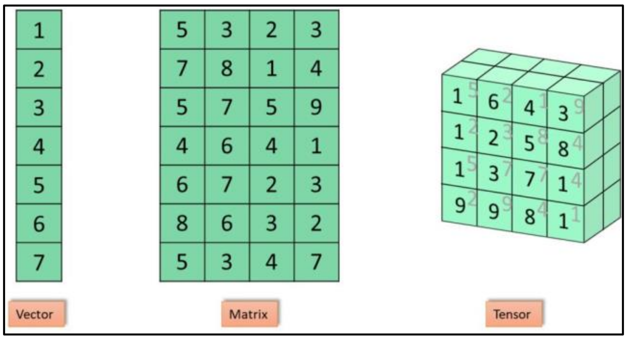
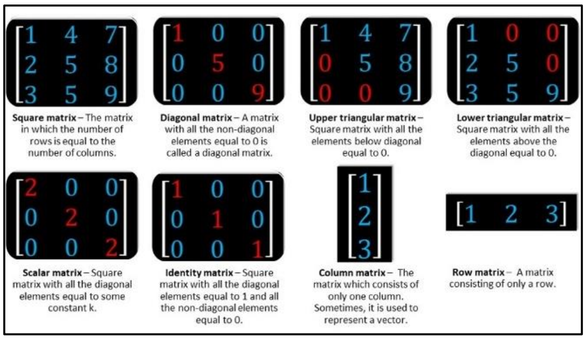
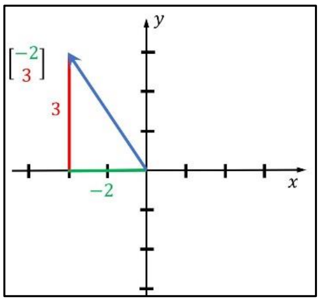
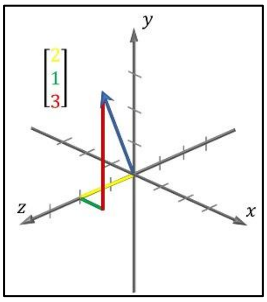
- **Vectors, Matrices, and Tensors**
    - In any ML/DL problem, most of data is presented as Vectors, Matrices, and Tensors.
    - A vector is a 1D array. a point in space is a vector of 3 coordinates (x, y, z). it is defined such that it has magnitude and direction.
    - A matrix is a 2D array of numbers, that has a fixed number of rows and columns. It contains a number at the intersection of $n^{th}\ row\ and\ d^{th} column$. A matrix is usually denoted by square brackets [].
    - A tensor is a generalization of vectors and matrices. E.g., a 1D tensor is a vector,2D tensor is a matrix.
    - We can have a 3D tensor like image with RGB colours. This continues to expand to n-dimensional tensors.
    - Following are some basic matrix terms to represent a matrix





- **Vectors**
    - In machine learning context, vectors are list of numbers. E.g. If we are predicting student final test marks based on his previous test marks, age and gender. $Student_1 = [65,15,0]$
    - In real word projects, this dimension can go up to thousands.
    - In linear algebra, vector coordinates tell distance & direction to go from the origin on that axis.





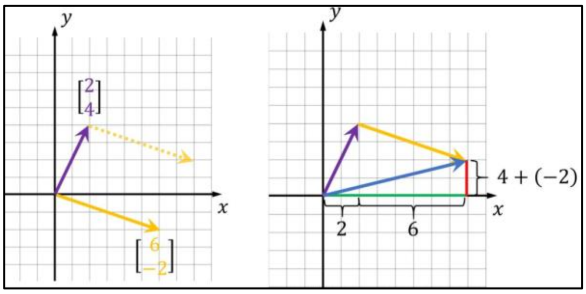
    - Row Vector- If we write all values in vector as a row (1 x n)

$$A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}_{(1\times n)}$$

    - Column Vector- If we write all values in vector as a column (n x 1)

$$B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}_{(n\times 1)}$$

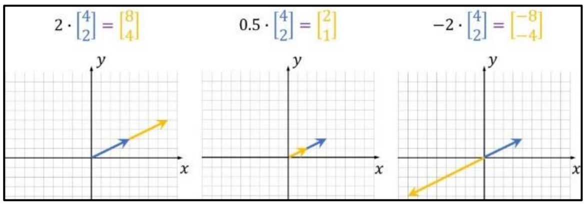
- **Vector Addition**



    - If we have 2 vectors $\vec{v}$, $\vec{w}$.
    - To add these $\vec{v}$, $\vec{w}$, we move the $\vec{w}$ so that its tail sits at the tip of the $\vec{v}$. The new vector is drawn from the tail of the $\vec{v}$ to the tip of the $\vec{w}$ after moving. It's because If you take a step along $\vec{v}$, then take a steps along $\vec{w}$, the overall effect is same as if you moved along $\vec{v} + \vec{w}$.

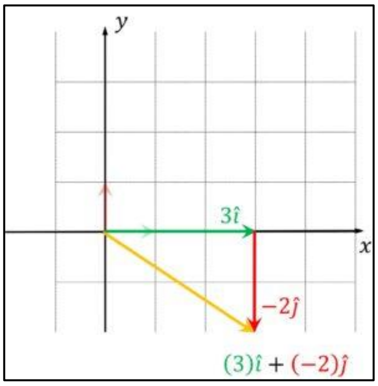
$$\vec{v} + \vec{w} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 6 \\ -2 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$$

- **Scalar-vector multiplication**

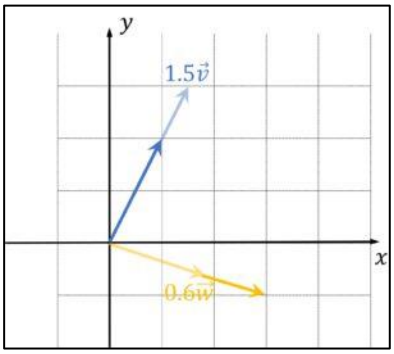


  - A vector is multiplied by a scalar, which is done by multiplying every element of the vector by the scalar.
  - This process of stretching the vector direction is called scaling. A scalar is a number that scales some vector.

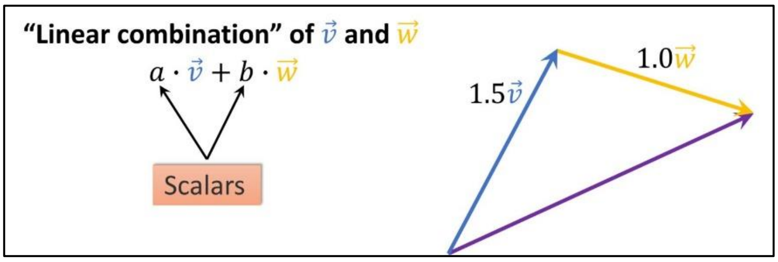
- **Unit Vectors (basis of a coordinate system)**
  - In the 2D coordinate system, there are two-unit vectors one on x direction & one in y $\hat{\imath} = \begin{bmatrix}1\\0\end{bmatrix}$ and $\hat{\jmath} = \begin{bmatrix}0\\1\end{bmatrix}$
  - A vector $[3, -2]$ scales $\hat{\imath}$ by a factor of 3 and $\hat{\jmath}$ by factor of -2.
  - These coordinates describe the sum of two scaled vectors $3\hat{\imath} + (-2\hat{\jmath})$
  - Coordinates as scalars are scaled unit vectors.



  - It is possible to have different basis or unit vectors. With these different basis vectors, we get a completely new coordinate system.

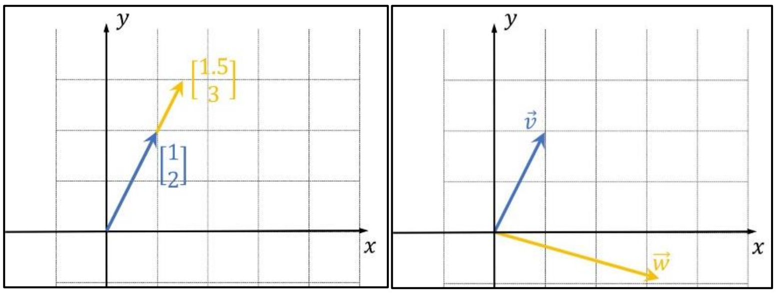


- **Linear combination & Span of vectors**
  - If two vectors are scaled with different scalars with any basis vectors, the addition of this 2 scaled vectors is linear combination.



  - The span of a vector is $\vec{v}$ and $\vec{w}$ is the set of all of their linear combinations i.e. the points they can reach. We will have the whole plane as a span for most of the vectors.

- **[Linearly independent and dependent vectors]**
  - $\vec{v}$ and $\vec{w}$ vectors are linearly dependent if they lie on the same direction.



  - If their addition adds a new span of dimensions then they are Linear independent. This vector will give a contribution to the span. For all values of a , $\vec{v} \neq a.\vec{w}$
  - Set of vectors is linearly dependant if $\beta_1 * a_1 + \cdots + \beta_k * a_k = 0$ where $\beta_1, \cdots, \beta_k \neq 0$. We can form the zero vector as a linear combination of the vectors.
  - When a collection of vectors is linearly dependent, at least one of the vectors can be expressed as a linear combination of the other vectors

$$a_1 = \begin{bmatrix} 0.2 \\ -7.0 \\ 8.6 \end{bmatrix} a_2 = \begin{bmatrix} -0.1 \\ 2.0 \\ -1.0 \end{bmatrix} a_1 = \begin{bmatrix} 0.0 \\ -1.0 \\ 2.2 \end{bmatrix} \text{ are linearly dependent, since } a_1 + 2a_2 - 3a_3 = 0$$

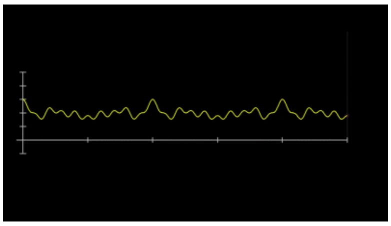
We can express any of these vectors as a linear combination of the other two $a_2 = \left(-\frac{1}{2}\right) a_1 + \left(\frac{3}{2}\right) a_3$

  - Set of vectors is linearly independent if $\beta_1 * a_1 + \cdots + \beta_k * a_k = 0$ where $\beta_1 = \cdots = \beta_k = 0$. We can form the zero vector as a linear combination of the vectors. We can't form the zero vector as a linear combination of the vectors unless all coefficients are zero

$$a_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} a_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} a_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \text{ are linearly independent}$$

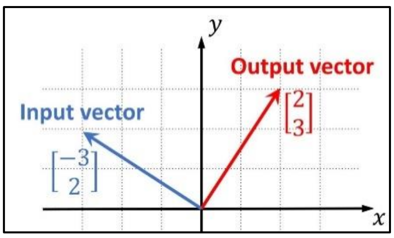
Suppose $\beta_1 * a_1 + \beta_2 * a_2 + \beta_3 * a_3 = 0$ However, this is only possible if all of the values are equal to zero.

  - A linearly independent collection of $n\text{-}vectors$ can have at most n elements.
  - A collection of n linearly independent $n\text{-}vectors$ (i.e., a collection of linearly independent vectors of the maximum possible size) is called a basis. i.e., the $n\text{-}vectors$ $a_1, \ldots, a_n$ are a basis, then any $n\text{-}vector\ b$ can be written as a linear combination of them or any $n\text{-}vector\ b$ can be written in a unique way as a linear combination of them.
  - In many areas of ML, we search for features and we examine their independence. a good understanding of basis helps transform data from one domain to another.
- **[Linear transformations and matrices]**
  - A linear transformation is a function that maps an input vector into an output vector of same length as input. e.g., Fourier Transform maps input sequence of N signal samples to a sequence of another N samples. New samples are complex numbers which capture the amplitude and phase (time-shift) of input signal.

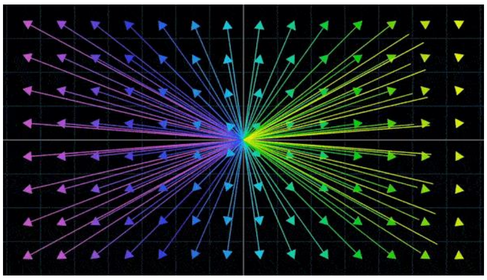
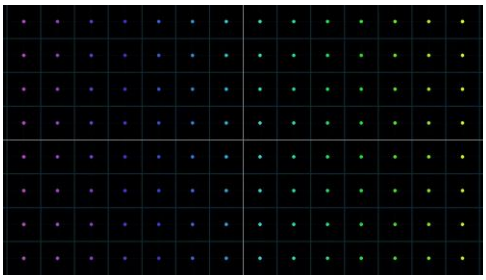


[Gif File]

  - With linear transformation vector is moved again in some particular way in plane to get output vector. E.g. input vector is rotated at 90° clockwise to form output vector.

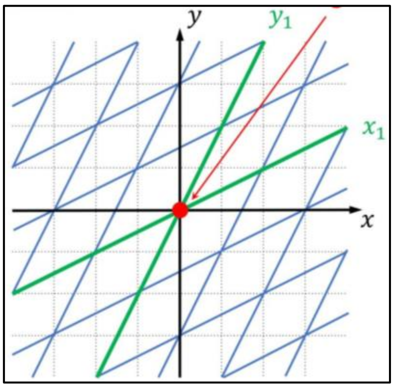


- This transform can also be applied on the whole set of vectors i.e. transform the whole plane to see where majority of the vectors will be mapped. We can visualize this is to represent vectors as points (Tip of vector).

[Gif File](#) [Gif File](#)

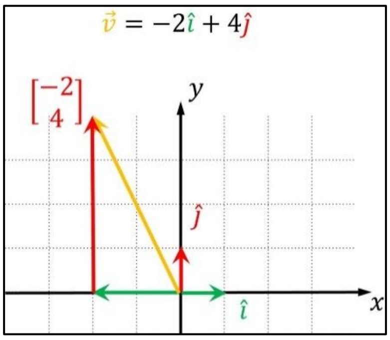
- linear transformation must obey 3 properties
  - Line should remain a line after transformation.
  - Origin should remain at the fixed place.
  - Distance between the grid lines should remain equidistant and parallel

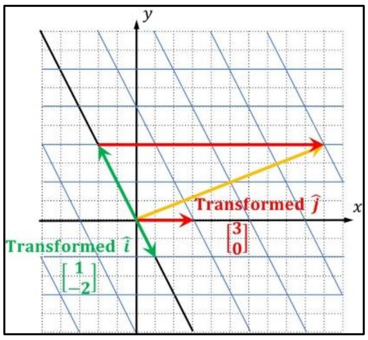


- Several important geometric transformations can be expressed as a matrix-vector product $w = Av$. A is 2×2 matrix. E.g., Scaling, Rotation, Reflection
- To transform input vector x_{in} and y_{in} to x_{out} and y_{out} basis vectors $\hat{\imath} = \begin{bmatrix}1\\0\end{bmatrix}$ and $\hat{\jmath} = \begin{bmatrix}0\\1\end{bmatrix}$ can be observed.

  E.g., To obtain vector $\vec{v} = \begin{bmatrix}-2\\4\end{bmatrix}$ and we have to multiply $\hat{\imath}$ with −2 and $\hat{\jmath}$ with 4.
- With linear transformation, we will obtain transformed $\hat{\imath}$ and transformed $\hat{\jmath}$. E.g. $\hat{\imath}_{trans} = \begin{bmatrix}1\\-2\end{bmatrix}$ and $\hat{\jmath}_{trans} = \begin{bmatrix}3\\0\end{bmatrix}$

$$\vec{v}_{trans} = -2\begin{bmatrix}1\\-2\end{bmatrix} + 4\begin{bmatrix}3\\0\end{bmatrix} = \begin{bmatrix}ax+by\\cx+dy\end{bmatrix} = \begin{bmatrix}1*-2+3*4\\-2*-2+0*4\end{bmatrix} = \begin{bmatrix}10\\4\end{bmatrix}$$





- To map any input vector $\begin{bmatrix}x\\y\end{bmatrix}$ with a linear transformation, the output vector will multiply them with transformed basis vectors.

$$\vec{v}_{trans} = \begin{bmatrix}1*x+3*y\\-2*x+0*y\end{bmatrix}$$

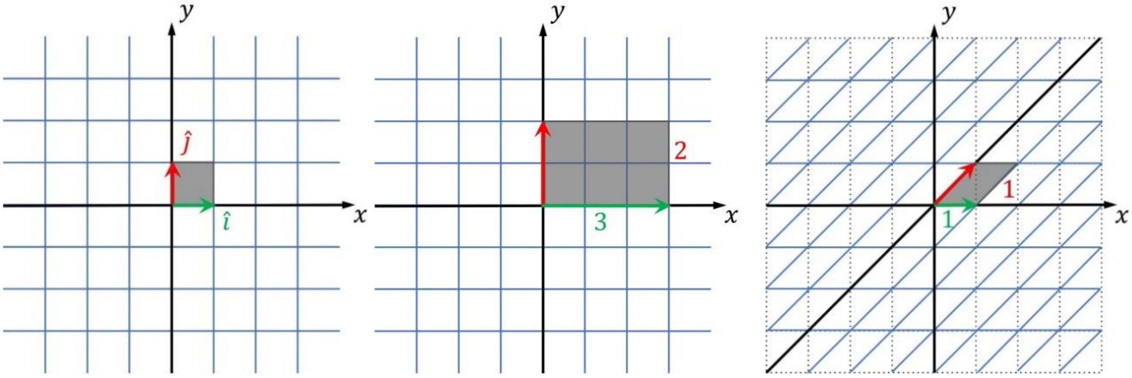
- We can map the whole 2-D plane if we know the transformed basis vectors.
- We can create 2×2 matrix of transformed basis vectors which can be used for vector processing. E.g. $\begin{bmatrix}1 & 3\\-2 & 0\end{bmatrix}$. Intuitively, it's summing two scaled column vectors. i.e. matrix-vector multiplication.

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

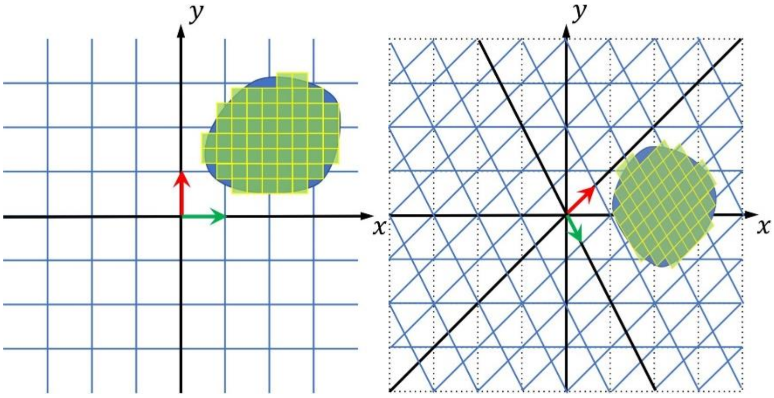
- If a matrix whose vectors are linearly dependent, then, a 2-D plane will be mapped to a single line.

- **The determinant**
  - A descriptive and more intuition of linear transformations will be the area of one object changed under a linear transformation. We will keep track of a unit square, defined by basis vectors, and how it is transformed under an operation. E.g. Dilation & Shear Transformation.

$$\begin{bmatrix}3 & 0\\0 & 2\end{bmatrix} \text{ New area } = 3 \times 2 = 6 \qquad \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix} \text{ New area } = 1 \times 1 = 1$$



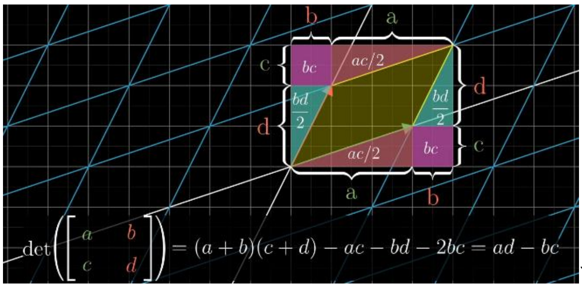
  - For any object in general which can be approximated by squares, every unit square will transform. So, every grid square of size 1 will be mapped into a parallelogram under shear operation.



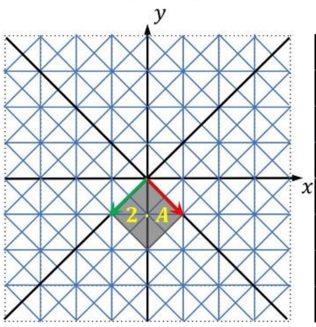
  - This new area is called as determinant of a transformation. The following operation doubles the area.

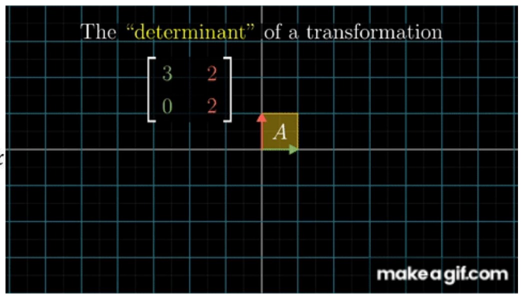
$$\det\left(\begin{bmatrix}a & b\\c & d\end{bmatrix}\right) = (a * d) - (b * c)$$

The $a * d$ term tells how much it has stretched in x & y direction while as $b * c$ how much it has been squished in diagonal direction



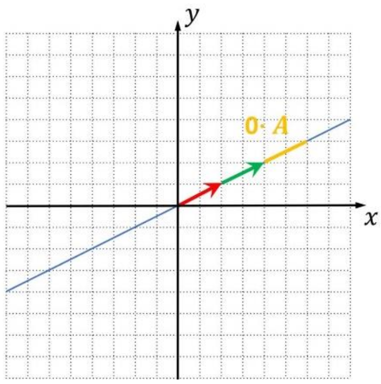
$$\det\left(\begin{bmatrix}-1 & 1\\-1 & -1\end{bmatrix}\right) = (-1 * -1) - (1 * -1) = 1 - (-1) = 1 + 1 = 2$$





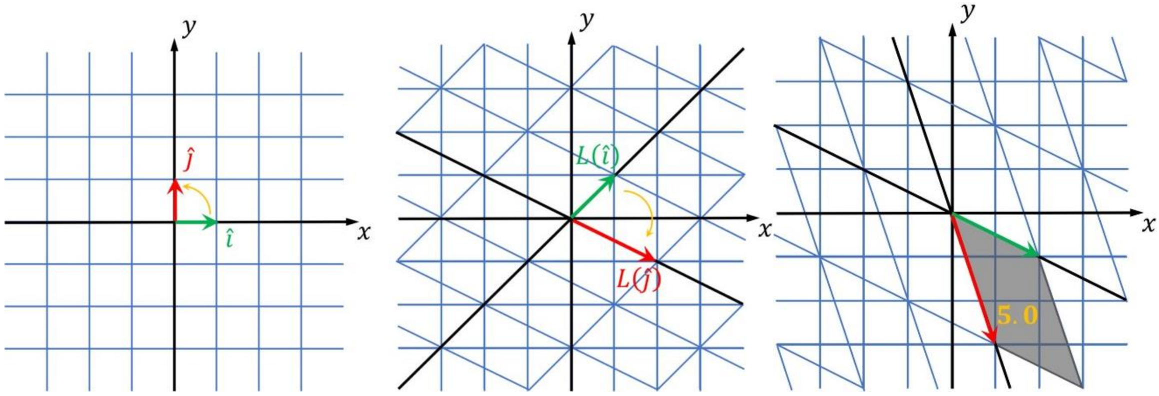
  - If a linear transformation map a 2-D space into a single line, the determinant value = 0

$$\det\left(\begin{bmatrix}4 & 2\\2 & 1\end{bmatrix}\right) = (4 * 1) - (2 * 2) = 4 - 4 = 0$$

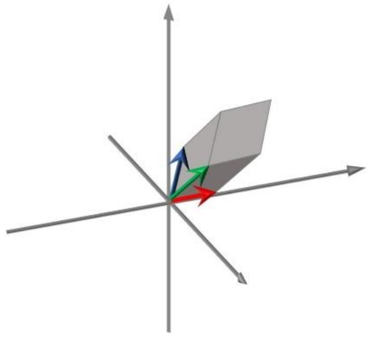


- If determinant is negative, it implies order of $\hat{\imath}$ and $\hat{\jmath}$ is reversed i.e., after a transformation $L(\hat{\jmath})$ is to the right of $L(\hat{\imath})$ or the orientation of the space has been reversed.

$$\det\left(\begin{bmatrix} 2 & 1 \\ -1 & -3 \end{bmatrix}\right) = (2 * -3) - (1 * -1) = -6 - (-1) = -6 + 1 = -5$$



- In 3D Space, we use determinant to calculate a volume transformation of unit cube to parallelepiped.



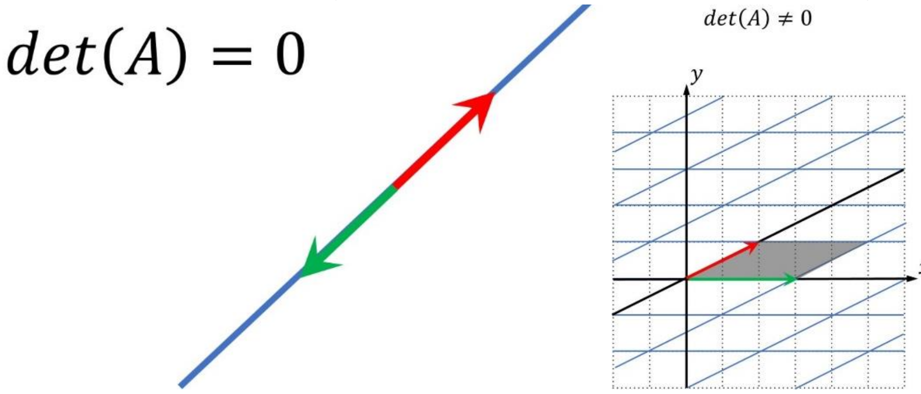
$$\det\left(\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}\right) = a * \det\left(\begin{bmatrix} e & f \\ h & i \end{bmatrix}\right) - b * \det\left(\begin{bmatrix} d & f \\ g & i \end{bmatrix}\right) + c * \det\left(\begin{bmatrix} d & e \\ g & h \end{bmatrix}\right)$$
$$= a(ei - fh) - b(di - fg) + c(dh - eg)$$
$$= aei - afh - bdi + bfg + cdh - ceg$$

- **Inverse matrices**
  - For Linear equation, variables x, y are only scaled by a coefficient and are summed. To solve them, we align variables on the left and constants on the right into another matrix.

$$2x + 2y = -4$$
$$1x + 3y = -1$$
$$\underbrace{\begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} -4 \\ -1 \end{bmatrix}}_{\vec{b}}$$
$$A\vec{x} = \vec{b}$$

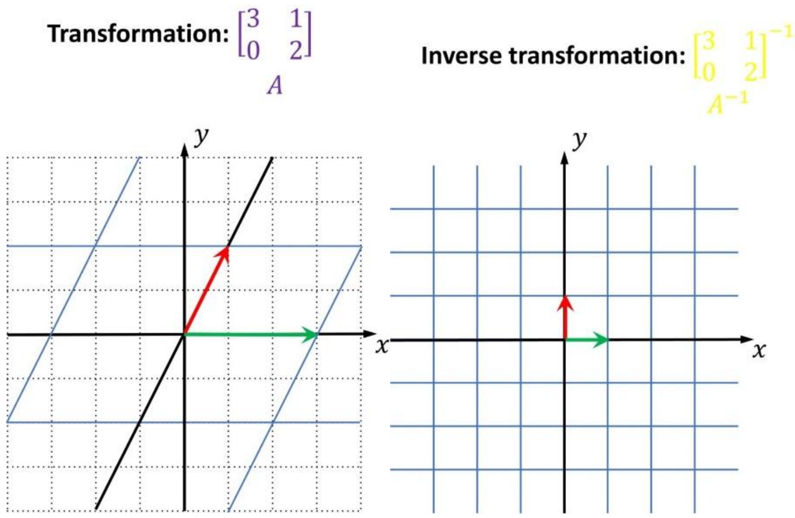
  - $\vec{x}$ is input vector and it is transformed to $\vec{b}$ using a matrix A. So, it moved somewhere in the 2-D space to obtain resulting vector $\vec{b}$.
  - So, we are looking for a vector $\vec{x}$ which lands on $\vec{b}$

- Case1: matrix A transforms 2D plane into a line i.e. columns of A are linearly dependent. i.e. det = 0
  Case2: matrix A transforms 2D plane into a 2D Plane i.e. columns of A are linearly independent. i.e. det ≠0



- If determinant is not 0, we will get the unique solution which will transform $\vec{x}$ into $\vec{b}$. As the $\vec{x}$ is input vector and it is transformed to $\vec{b}$ Vice versa we can go back from $\vec{b}$ in search of $\vec{x}$. This backward search is inverse transformation.

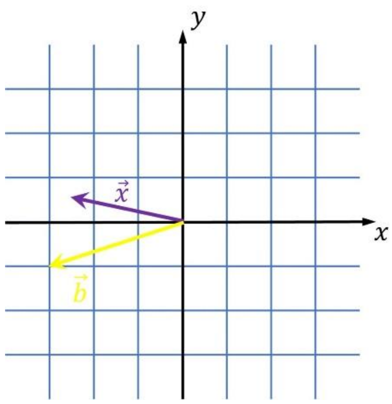
  E.g., $A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$ and Its inverse transformation will be $A^{-1} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}^{-1}$



- If we consecutively apply A and then A^{-1} we will go back to the basis vectors $\hat{\imath} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\hat{\jmath} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ .This consecutive transformation is called as matrix multiplication. This is also called as identity transformation as this transformation does nothing but gives us the identity matrix.
- To solve the equation $A\vec{x} = \vec{b}$ we will multiply both sides by A^{-1}

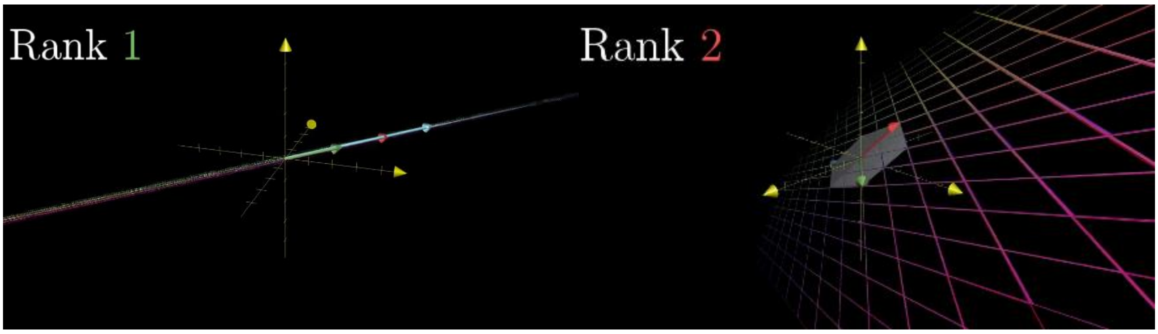
$$A^{-1}A\vec{x} = A^{-1}\vec{b}$$

- Intuitively we are transformation in reverse and following $\vec{b}$ thus wherever $\vec{x}$ will land is solution



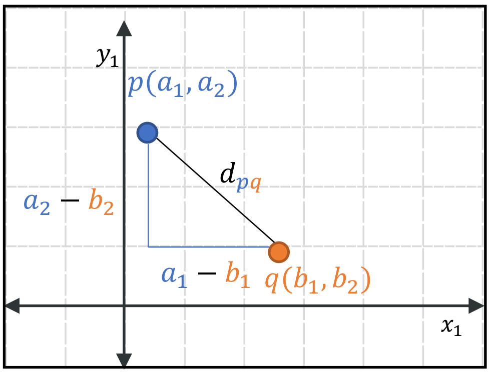
- when $det(A) = 0$ then any vector will be squashed into the line. The vector mapped into this line, cannot determine its location of origin i.e. we can't find A^{-1}

- **Rank, Column Space and Null Space**
  - Rank is the number of dimensions in the output after linear transformation.
  - If a linear transformation of 3D vectors lands on a line which is a 1-D we say that we have a rank-1 transformation. If we have a rank-2 transformation then 3-D vectors will land into a plane.



  - Column space is span of columns of matrix i.e. All possible solutions that one matrix A.
  - A null space is a set of all vectors that are mapped into $\vec{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ vector.

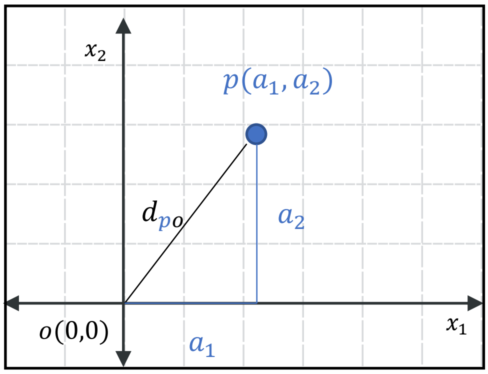
- **Distance between 2 points**



  - By Pythagoras theorem,
  - In 2D Space, $d_{pq} = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}$
  - In 3D space $p(a_1, a_2, a_3), q(b_1, b_2, b_3)$ $d_{pq} = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2}$
  - In ND Space $p(a_1, a_2, a_3, \cdots, a_n), q(b_1, b_2, b_3, \cdots, b_n)$

$$d_{pq} = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2}$$

- **Distance of a point from its origin**



  - By Pythagoras theorem,
  - In 2D Space, $d_{po} = \sqrt{{a_1}^2 + {a_2}^2}$
  - In 3D space $p(a_1, a_2, a_3)$ $d_{po} = \sqrt{{a_1}^2 + {a_2}^2 + {a_3}^2}$
  - In ND Space $p(a_1, a_2, a_3, \cdots, a_n)$

$$d_{po} = \sqrt{\sum_{i=1}^{n} {a_i}^2}$$

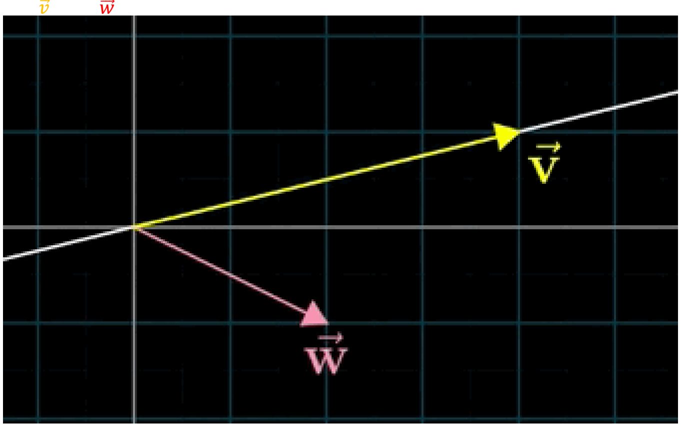
- **Dot Product or Scalar Product or Inner Product**
    - Two vectors can be multiplied using the "Dot Product". We mostly use dot product in ML
    - In matrix notation, the dot product of two vectors is the sum of their component wise products:

$$a = \begin{bmatrix}2\\3\\4\end{bmatrix} \quad b = \begin{bmatrix}4\\6\\7\end{bmatrix}$$

$a\,.\,b$ will be $a_x b_x + a_y b_y + a_z b_z = 2 * 5 + 3 * 6 + 4 * 7 = 56$

    - A dot product between two vectors is a projection of vector $\vec{w}$ on vector $\vec{v}$. Then, we multiply the length of projected vector $\vec{w}$ onto $\vec{v}$ and the length of $\vec{v}$.

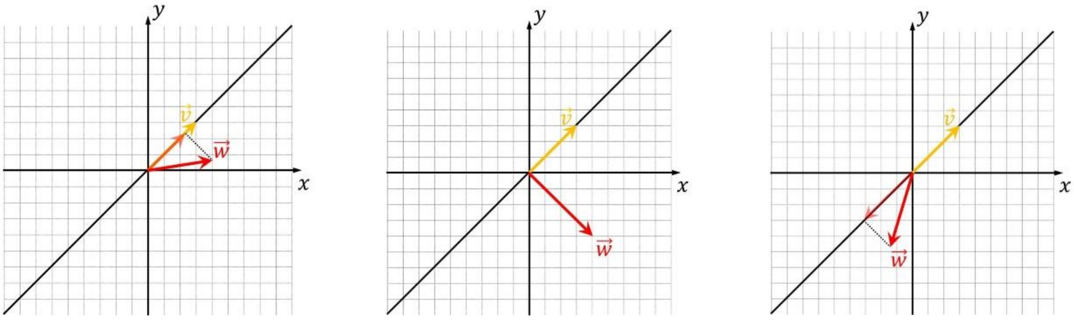
$$\underbrace{\begin{bmatrix}8\\2\end{bmatrix}}_{\vec{v}} \cdot \underbrace{\begin{bmatrix}-2\\-4\end{bmatrix}}_{\vec{w}} = (Length\ of\ projected\ \vec{w}) \cdot (Length\ of\ \vec{v}) \cdot$$



Gif Link

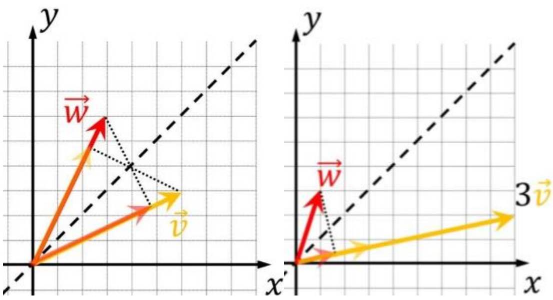
    - If projection is opposite side of $\vec{v}$ the dot product is negative and if they are perpendicular the dot product is zero

$\vec{v} \cdot \vec{w} > 0$ Similar directions - positive $\vec{v} \cdot \vec{w} = 0$ Perpendicular – equals to 0 $\vec{v} \cdot \vec{w} < 0$ Opposing directions - negative



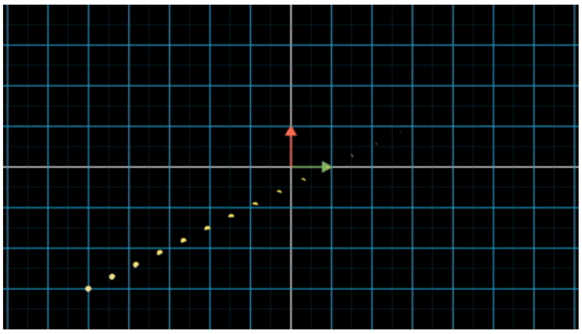
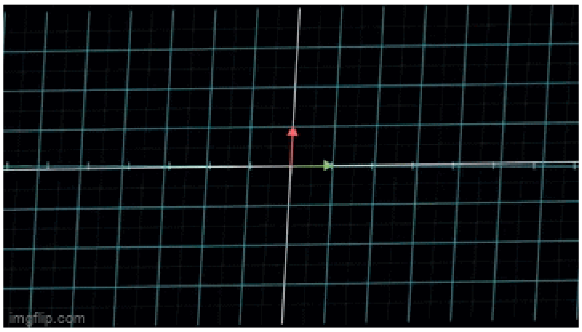
    - If $\vec{v}$ and $\vec{w}$ has same length, the projected length of both vectors on each other would be same. Thus, calculation sequence does not matter.

$$\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}$$



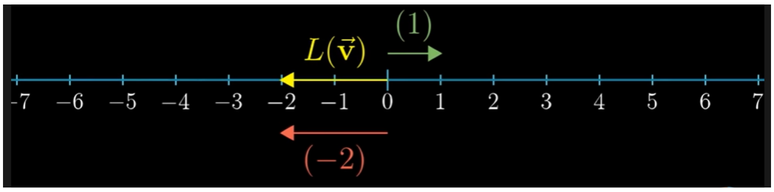
    - If $\vec{v}$ is 3 times longer than $\vec{w}$, we can interpret $3\vec{v}$ is scaling of $\vec{v}$. But scaling is actually scaling the length. Thus, $(3\vec{v}\ ) \cdot \vec{w} = 3(\vec{v} \cdot \vec{w})$

- The relation between 2 calculation methods comes with duality. For that let's see linear transformation from multiple dimension to single dimension i.e. the number line.
- Here A line with evenly spaced dots on it, will be mapped to a 1D line such that distance between dots on the mapped line will be equidistant. (property of the Linear Transformation). We can keep track of basis vectors that will be transformed

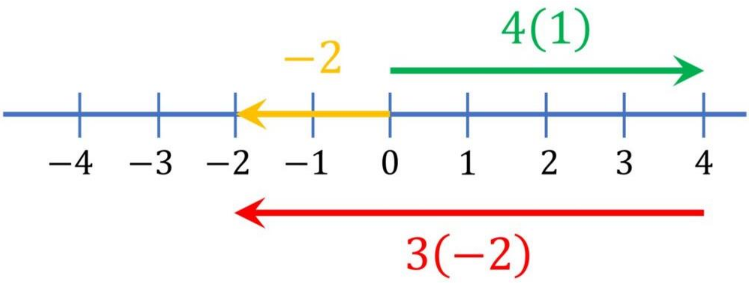
- By using basis vectors, we can calculate where any vector will land on transformed line.

  E.g. $\hat{\imath} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\hat{\jmath} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ change to $[1 \quad -2]$

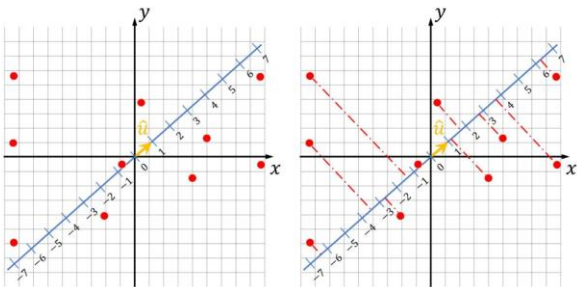


- A vector $\vec{v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ will be transformed to

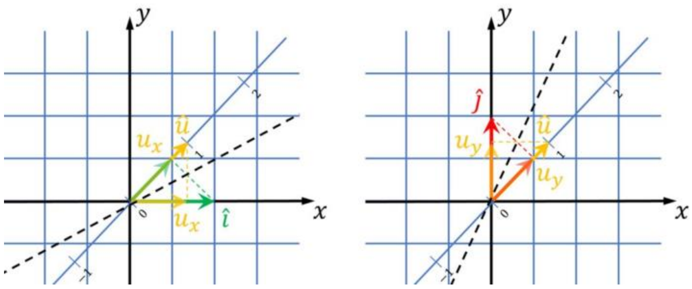
$$\overbrace{[1 \quad -2]}^{Transform} \underbrace{\begin{bmatrix} 4 \\ 3 \end{bmatrix}}_{Vector} = 4 \cdot 1 + 3 \cdot (-2) = 4 - 6 = -2$$



- For further clarification, it can be thought as how 2-D points will be projected onto a single line. The transformed single line has unit vector $\hat{u}$.



- The projection of Basis vector $\hat{\imath}$ and $\hat{\jmath}$ onto a unit vector $\hat{u}$ will be same as $\hat{u}$ will be a projected $\hat{\imath}$ and $\hat{\jmath}$ i.e. $x_{coordinate} = u_x$ and $y_{coordinate} = u_y$ of a vector $\hat{u}$.



- When we multiply vector (x, y) with u_x and u_y respectively, and sum the products, we get a position where (x, y) vector will land on a line defined with a vector $\hat{u}$ or the new co-ordinate on single line.
- The matrix vector products are dual with the dot product interpretation.

$$\begin{bmatrix} u_x & u_y \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = u_x \cdot x + u_y \cdot y$$

$$\begin{bmatrix} u_x \\ u_y \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = u_x \cdot x + u_y \cdot$$

- To calculate the length or norm of single vector using the dot product

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

  ⇨ e.g. $\vec{v} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$

  ⇨ $\vec{v} \cdot \vec{v} = 1 + 9 + 4 = 14$

  ⇨ $\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}} = \sqrt{14}$

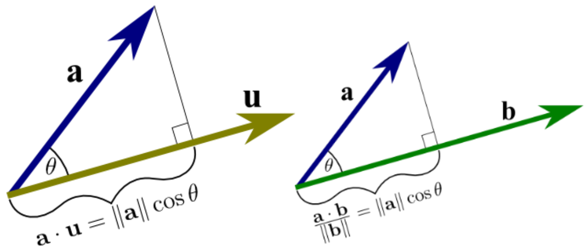
- To normalize vectors i.e. $\|\vec{v}\| = 1$ we need to divide the vector $\vec{v}$ by $\|\vec{v}\|$

  ⇨ $\hat{u} = \frac{\vec{v}}{\|\vec{v}\|} = \begin{bmatrix} \frac{1}{\sqrt{14}} \\ \frac{3}{\sqrt{14}} \\ \frac{2}{\sqrt{14}} \end{bmatrix}$

  ⇨ $\|\vec{u}\| = 1 = \frac{1}{14} + \frac{9}{14} + \frac{4}{14} = 1$

- We can also calculate angle between two vectors

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



- The projection of a on u is $\| a \| \, cos\theta$

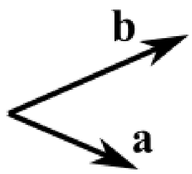
$$a \cdot u = \| a \| \, cos\theta$$

$$\frac{a.b}{\|b\|} = \| a \| \, cos\theta$$

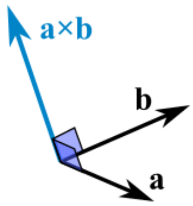
$$a \cdot b = \| a \| \| b \| \, cos\theta$$

$$cos\theta = \frac{a \cdot b}{\| a \| \| b \|}$$

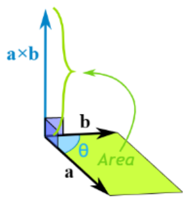
- **Cross Product or Vector Product**
  - Two vectors can be multiplied using the "Cross Product". Cross product for a vector is possible when row vector has dimension of 1x n and n x 1



  - The Cross Product a × b of two vectors is another vector that is at right angles to both



  - The magnitude (length) of the cross product equals the area of a parallelogram with vectors a and b for sides



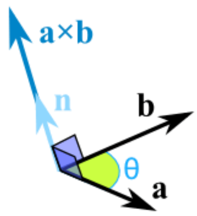
- It can be calculated as

$$\boxed{a\ X\ b = |a| * |b| * \sin\theta}$$

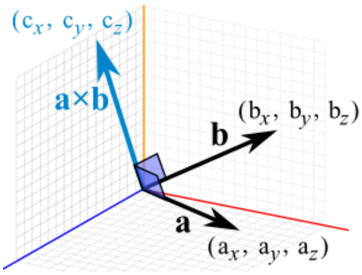
$$|a|_{origin} = \sqrt{{a_1}^2 + {a_2}^2 + \ldots, {a_n}^2}$$

$$|b|_{origin} = \sqrt{{b_1}^2 + {b_2}^2 + \ldots, {b_n}^2}$$

θ is the angle between a and b

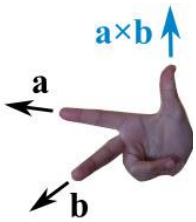


- It is to be noted that the cross product is a vector with a specified direction. The resultant is always perpendicular to both a and b.
- In case a and b are parallel vectors, the resultant shall be zero as $sin(0) = 0$
- When a and b start at the origin point (0,0,0), the Cross Product will end at:



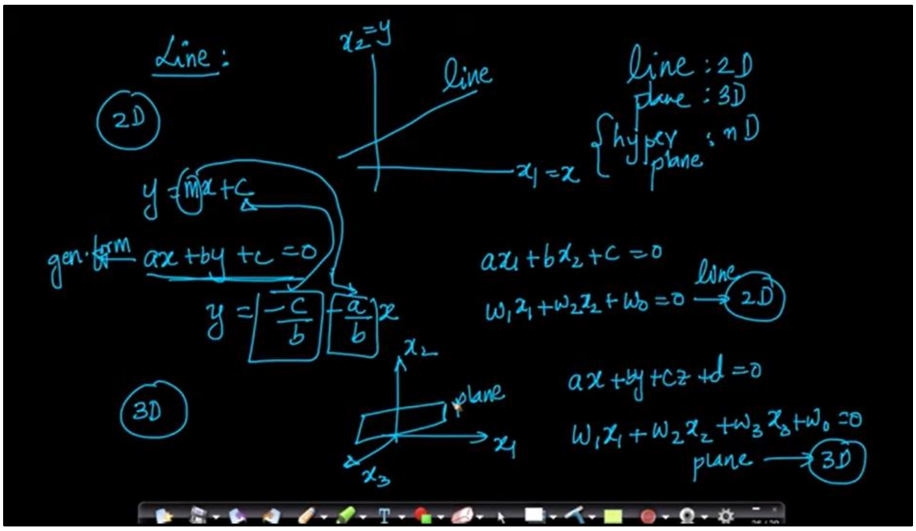
$$\begin{array}{c} \begin{matrix} [X & Y & Z] \end{matrix} \\ a = [2 \quad 3 \quad 4] \\ b = [5 \quad 6 \quad 7] \end{array}$$

$a\ X\ b$ will be
$$\begin{aligned} c_x &= a_y b_z - a_z b_y = 3 * 7 - 4 * 6 = -3 \\ c_y &= a_z b_x - a_x b_z = 4 * 5 - 2 * 7 = 6 \\ c_z &= a_x b_y - a_y b_x = 2 * 6 - 3 * 5 = -3 \end{aligned}$$

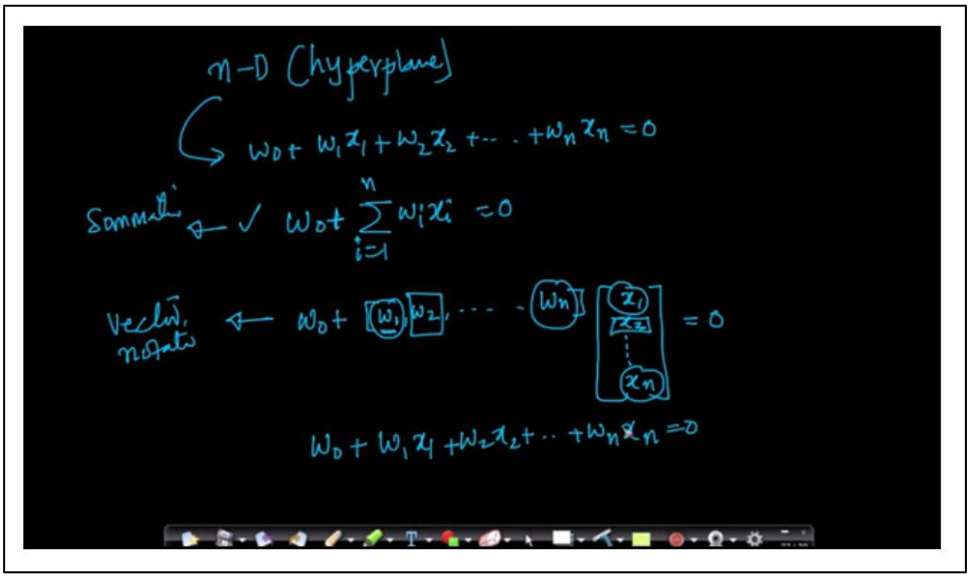


- The cross product could point in the completely opposite direction and still be perpendicular to the two other vectors, so we have the "Right Hand Rule"
- With your right-hand, point your index finger along vector a, and point your middle finger along vector b: the cross product goes in the direction of your thumb.

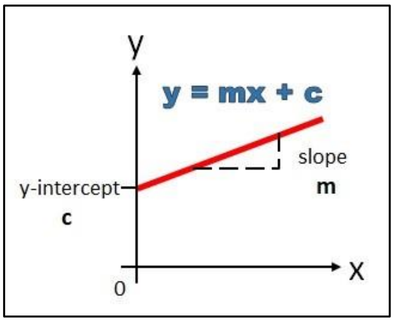
- **Equation of a line (2-D), Plane(3-D) and Hyperplane (n-D)**
  - Equation of a 2D line is $w_1 * x_1 + w_2 * x_2 + w0 = 0$ (w1, w2 are constants while as x1 x2 are dimensions)
  - Equation of a 3D plane is $w_1 * x_1 + w_2 * x_2 + w_3 * x_3 + w0 = 0$ (w1, w2, w3 are constants while as x1 x2 x3 are dimensions)
  - AWN, Line separates all 2D points in 2 regions, plane separates 3D Surface it into 2 Parts



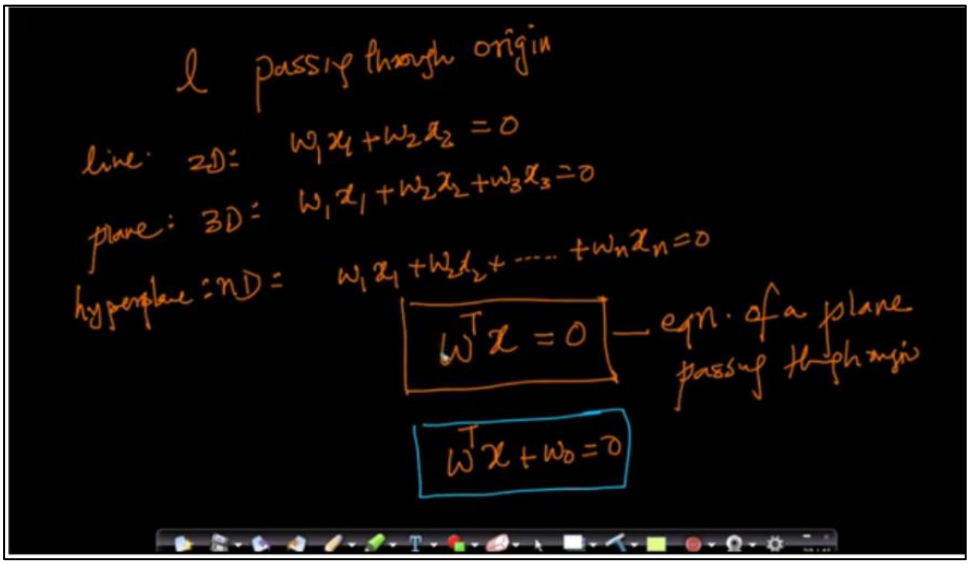
  - For n Dimensional plane, the equation would be, $W0 + \sum Wi.Xi = 0$
  - Consider we have two vectors; one is a vector of all constants as W. Another is column vector having x1, x2, x3 …. xn as X
  - The equation of hyperplane would be dot product of these two vectors $WT.X + w0$



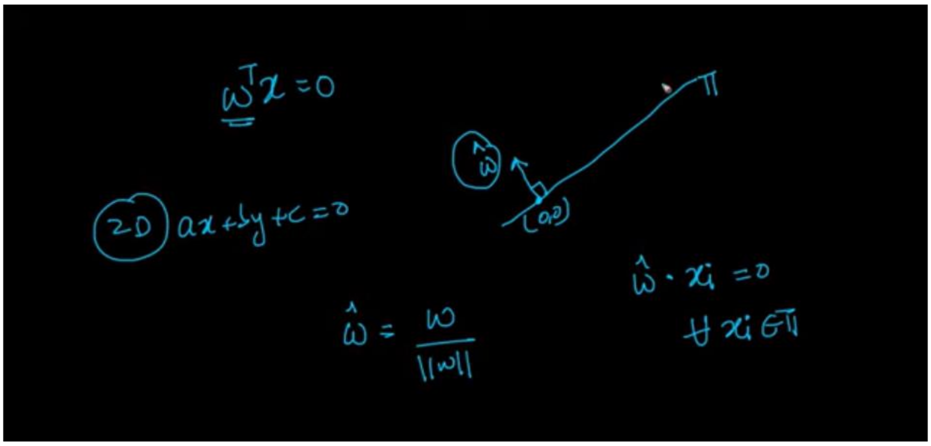
  - What is constant C or W0?
    - We will understand this by using 2D equation. General line equation is y = mx + c
    - m = slope of line & c is a y intercept
    - So, for a line passing through origin the c = 0



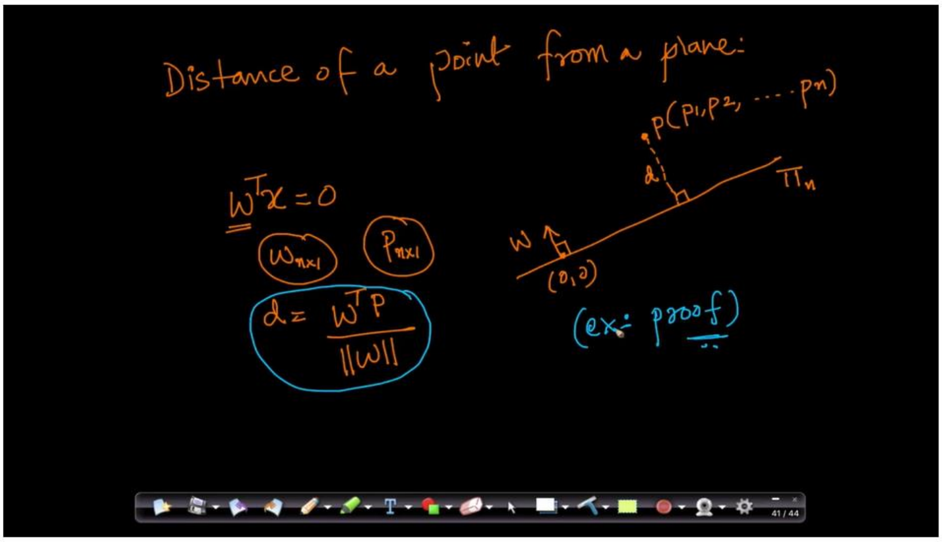
- We can summarise it as:



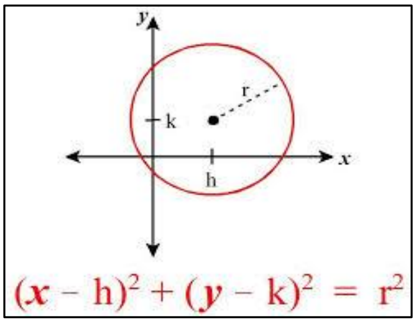
- Now, consider $W^TX = 0$ Geometrically, X represents a vector or a point of components x1, x2. W is a vector which is perpendicular to x
- i.e., $W^TX = 0$ signifies W and X are perpendicular to each other since there product is 0



- **Distance of a point from a Plane**
  - We know that $W^TX = 0$(W vector is perpendicular to X)
  - To measure a shortest distance of point X from plane the formula is $\frac{(W^T.X)}{\|W\|}$

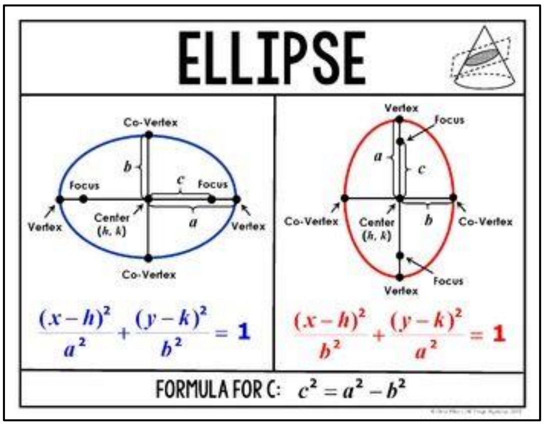


- **Circle**



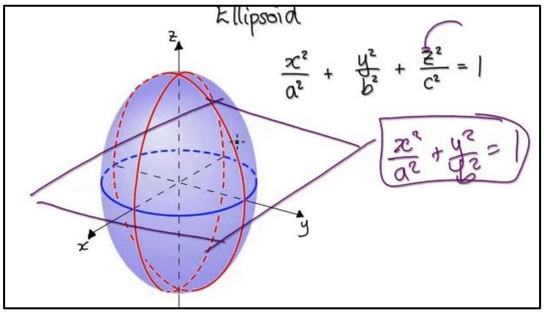
- In this figure we can see equation of circle if centre of circle is at 0,0. The equation becomes $r^2 = x^2 + y^2$
- Circle divides whole region into 2 parts. One inside & one outside
- Suppose we have a point p (d, c) and we want to check whether it is inside or outside the circle
  - If $d^2 + c^2 < r^2$: p lies inside the circle
  - If $d^2 + c^2 > r^2$: p lies outside the circle
  - If $d^2 + c^2 = r^2$: p lies on the circle
- For a sphere(3D) having x, y, z as its dimension's equation is $r^2 = x^2 + y^2 + z^2$
- For a Hypersphere (nD) having x_1, x_2, $x_3, \ldots x_n$ as its dimension equation is $r^2 = x_1^2 + x_2^2 + x_3^2 + \ldots + x_n^2$
- The point inside or outside concept remains same in for n dimensional space

- **Ellipse**



- Suppose we have a point p (d, c) and we want to check whether it is inside or outside the circle
  - If $(d - h)^2 / a^2 + (c - k)^2 / a^2 < 1$: p lies inside the ellipse
  - If $(d - h)^2 / a^2 + (c - k)^2 / a^2 > 1$: p lies outside the ellipse
  - If $(d - h)^2 / a^2 + (c - k)^2 / a^2 = 1$: p lies on the ellipse
- X and y in equation means any point with co-ordinate x, y

- **Ellipsoid(3D)**



- Ellipsoid is 3D variant of ellipse equation is same as ellipse extended for 3D. the point checking properties remain same for ellipsoid as that of ellipse
- Above concept can be extended to Hyper ellipsoid which is basically an n dimensional ellipse

- **Square & Rectangle**