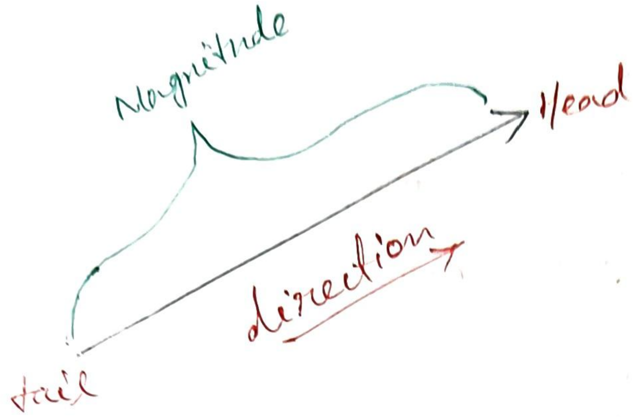
Linear Algebra



Scalar	Vector
> only magnitude eg: weight, height, speed 40kmph	> It has magnitude as well as direction. eg: velocity 40kmph south
> AB = 5	> $\overrightarrow{AB}$ (symbol)

Type of Vectors

① Zero Vector: Vector with zero magnitude.

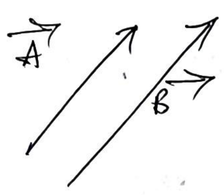
② Unit Vector: Vector having unit magnitude.

③ Coinitial Vector: Two or more vectors with same initial point

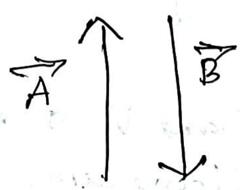
④ Collinear Vector: Two or more vectors lying on the same or parallel lines.

⑤ Equal Vectors: Two or more vectors with same magnitude and direction.

⑥ Negative Vectors: Vectors with same magnitude but opposite direction as that of the given vector.

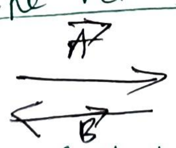


Like Vectors



Opposite Vectors

Unlike Vectors



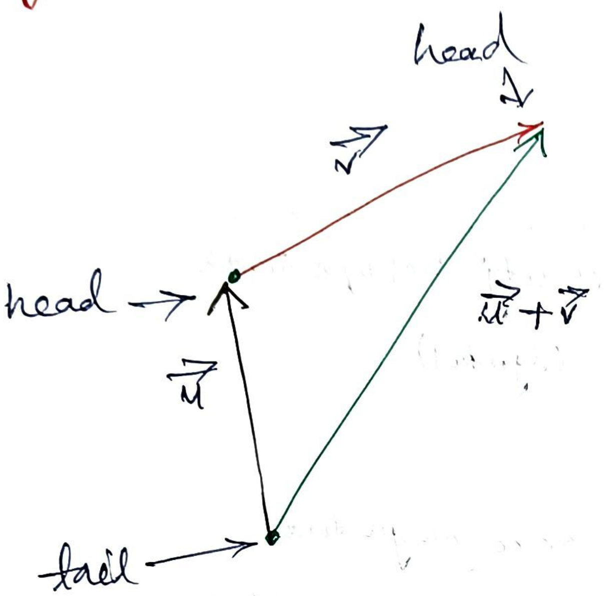
Equal Vectors.

Co-initial Vectors

> Vector moves in a space.
> Vector is always having magnitude and direction.

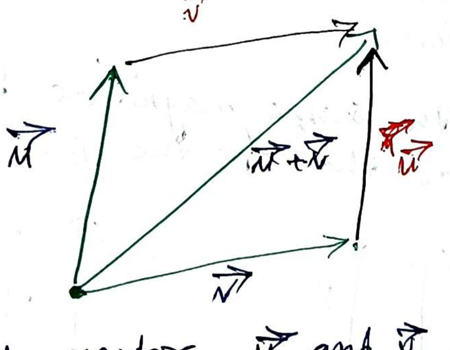
Graphical Methods for Vector Addition

Triangle Method or Head to Tail Method



① Place the vectors with the head of the previous vector $\vec{u}$ connected to the tail of the successive vector $\vec{v}$.
② The resultant vector $\vec{u}+\vec{v}$ is formed by connecting the tail of the first vector to the head of the last vector.

Parallelogram Method



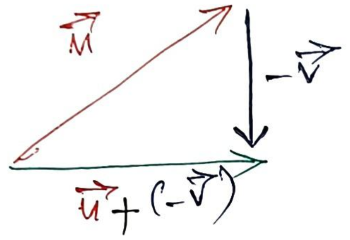
① Place both vectors, $\vec{u}$ and $\vec{v}$ at the same initial point
② Complete the parallelogram.
③ The diagonal of the parallelogram is the resultant vector $\vec{u}+\vec{v}$.

Subtraction of Vectors

> Subtracting a vector is the same as adding its negative:

$$\boxed{\vec{u} - \vec{v} = \vec{u} + (-\vec{v})}$$

(1) Switch the direction of the vector that is being subtracted.



(2) Arrange the two vectors from "tip to tail".

(3) Draw a resultant vector from the tail of the first vector to the tip of the second.

Properties of vector

(1) Cummulative, $\vec{u} + \vec{v} = \vec{v} + \vec{u}$

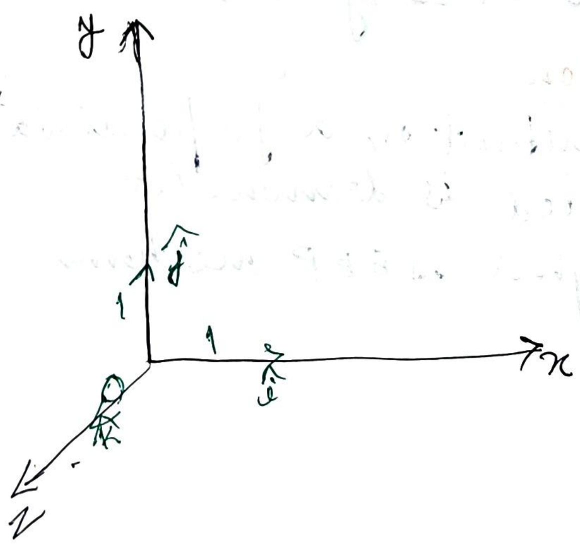
(2) Associative, $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$

(3) Existence of Identity,

$$\vec{v} + 0 = \vec{v}$$

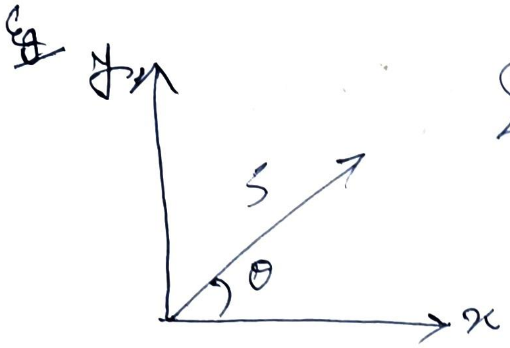
(4) Additive Inverse $= \vec{v} + (-\vec{v}) = \vec{0}$

Co-ordinate System



$\hat{i}$ = unit vector in x-directⁿ
$\hat{j}$ = unit vector in y directⁿ
$\hat{k}$ = unit vector in z directⁿ

Eg



$\left\{ \begin{array}{l} 3\hat{i} + 4\hat{j} \\ 4\hat{i} + 3\hat{j} \end{array} \right\}$ magnitude $= \sqrt{3^2+4^2} = 5$

> In data science, n-D space can be possible. We can visualise for 2D or 3D only.

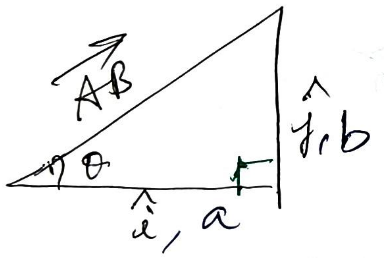
$$\boxed{\vec{AB} = a\hat{i} + b\hat{j} + c\hat{k}}$$

> Direction doesn't has magnitude.

Vector Sum

$\vec{a} = 2\hat{i} + 3\hat{j} + 4\hat{k}$

$\vec{a} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ Column Matrix



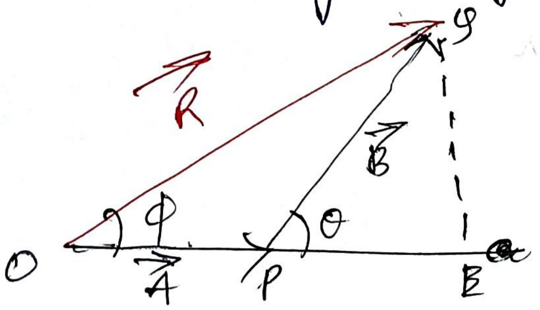
$$\boxed{|AB|^2 = |a|^2 + |b|^2}$$

a, b = magnitude.

> angle b/w $a\hat{i}$ & $b\hat{j}$ is $90°$.

Analytical Method of Vector Addition

To find the magnitude of resultant R, a perpendicular QE from Q on side OP produced is drawn. Let $\angle QPE = \theta$. Then, in right-angled $\triangle QEP$ we have



$$OQ^2 = OE^2 + QE^2$$

$$= \cancel{(OP+OR)^2 + QE^2}$$

$$= \cancel{OP^2 + P'}$$

$$= (OP+PE)^2 + QE^2$$

$$= OP^2 + PE^2 + 2OP.PE + QE^2$$

Here $PE^2 + QE^2 = QP^2$

$$\therefore OQ^2 = OP^2 + PQ^2 + 2OP.PE$$

In ΔPEQ,

$$\cos\theta = \frac{PE}{PQ}$$

$$PE = PQ\cos\theta$$

$$OQ^2 = OP^2 + PQ^2 + 2OP\,PQ\cos\theta$$

$$R^2 = A^2 + B^2 + 2AB\cos\theta$$

$$\boxed{R = A^2 + B^2 + 2AB\cos\theta} \Rightarrow \text{Magnitude}$$

Direction of the Resultant

$$\tan\phi = \frac{QE}{OE} = \frac{QE}{OP+PE}$$

$OP = A$ and $PE = B\cos\theta$

In ΔPEQ,

$$\sin\theta = \frac{QE}{PQ}$$

$$QE = PQ\sin\theta = B\sin\theta$$

$$\boxed{\tan\phi = \frac{B\sin\theta}{A + B\cos\theta}} \Rightarrow \text{Direction}$$

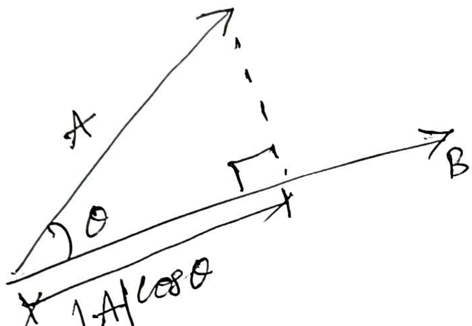
Dot Product of Vectors

Representation of vector :

① $\vec{a}$, $\vec{AB}$ (direction and magnitude)

② $\vec{a} = A\hat{i} + B\hat{j} + C\hat{k}$ (using own coordinate axes)

③ $\vec{AB} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ (column vector)



> Projection of vector A on the vector B. is called dot product.

$$\boxed{\vec{A} \cdot \vec{B} = |A|\,|B| \cos\theta}$$

> Dot product is a scalar.

So, we know

$$\vec{A} + \vec{B} = |A|^2 + |B|^2 + 2|A||B| \cos\theta$$

using, $|A||B| \cos\theta = \vec{A} \cdot \vec{B}$

$$\boxed{\vec{A} + \vec{B} = |A|^2 + |B|^2 + 2\vec{A} \cdot \vec{B}}$$ Transformed Formula.

Eg. $\vec{a} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ $\vec{b} = \begin{bmatrix} 1 \\ 7 \\ 8 \end{bmatrix}$

$$\vec{a} \cdot \vec{b} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 \\ 7 \\ 8 \end{bmatrix} =$$

$\vec{x_1} = a\hat{i} + b\hat{j} + c\hat{k}$ $\vec{x_2} = A\hat{i} + B\hat{j} + C\hat{k}$

Dot product $= \vec{x_1} \cdot \vec{x_2}$

$= aA + bB + cC$

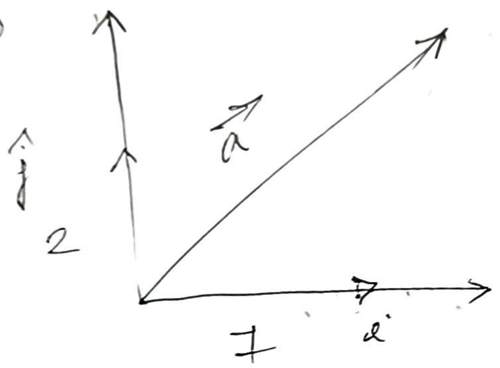
> If $\theta = 0$

$$|A|\,|B| \cos\theta = \vec{A}\cdot\vec{B} = |A|\,|B|$$

$$\cos 0^\circ = 1$$

Eg

①



$\vec{a} = 7\hat{i} + 2\hat{j}$

magnitude $= \sqrt{7^2+2^2} = \sqrt{53}$

② $\vec{A} = 3\hat{i} + 4\hat{j} + 7\hat{k}$

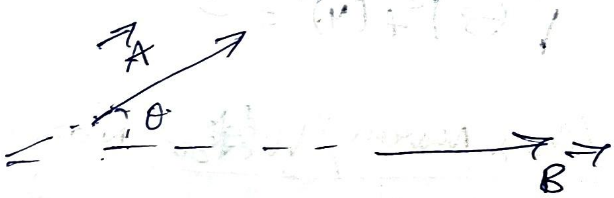
$\vec{B} = 7\hat{i} + 3\hat{k}$, $\cos\theta = 30^\circ$

$$\vec{A}+\vec{B} = \sqrt{|\vec{A}| + |\vec{B}|^2 + 2|\vec{A}|\,|\vec{B}|\cos\theta}$$

$$= \sqrt{(3^2+4^2+7^2) + (7^2+3^2) + 2\sqrt{(3^2+4^2+7^2)}\sqrt{(7^2+3^2)}\cdot\cos 30^\circ}$$

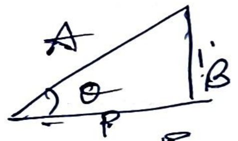
③ $\vec{A} = 2\hat{i} + 3\hat{j}$

$\vec{B} = 7\hat{i} + 6\hat{j}$



what is the projection of $\vec{A}$ on $\vec{B}$?

$\vec{A}\cdot\vec{B} = |\vec{A}|\,|\vec{B}|\cos\theta$



$\cos\theta = \frac{P}{|\vec{A}|}$

$\{|\vec{A}|\cos\theta = P\}$ (Projection)

$\vec{A}\cdot\vec{B} = |\vec{B}|\,P$

$= \sqrt{7^2+6^2}\,P$

$= \sqrt{85}\,P$

$\vec{A}\cdot\vec{B} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \begin{bmatrix} 7 \\ 6 \end{bmatrix} = 14+18 = 32$

$32 = \sqrt{85}\,P$

$\boxed{P = \frac{32}{\sqrt{85}}}$ ⇒ Projection.

Direction of projection,

$$P\left(\frac{\vec{B}}{|B|}\right) \quad \text{direction}$$

> cancel the magnitude from vector to get the direction.

Vector Norms

> Terminology to control the overfitting like L1 Norm, L2 Norm.

L2 Norm (Ridge Regression)

$3\hat{i} + 4\hat{j}$

Magnitude $= \sqrt{3^2 + 4^2} = 5 \quad \|x\|_2$ or $\|L\|_2$

L1 Norm (Lasso Regression)

~~$|3| + |4| = 7$~~

$-3\hat{i} + 4\hat{j}$

$\|L_1\| = |-3| + |4| = 7$

$\|L_2\| = \sqrt{(-3)^2 + (4)^2} = 5$

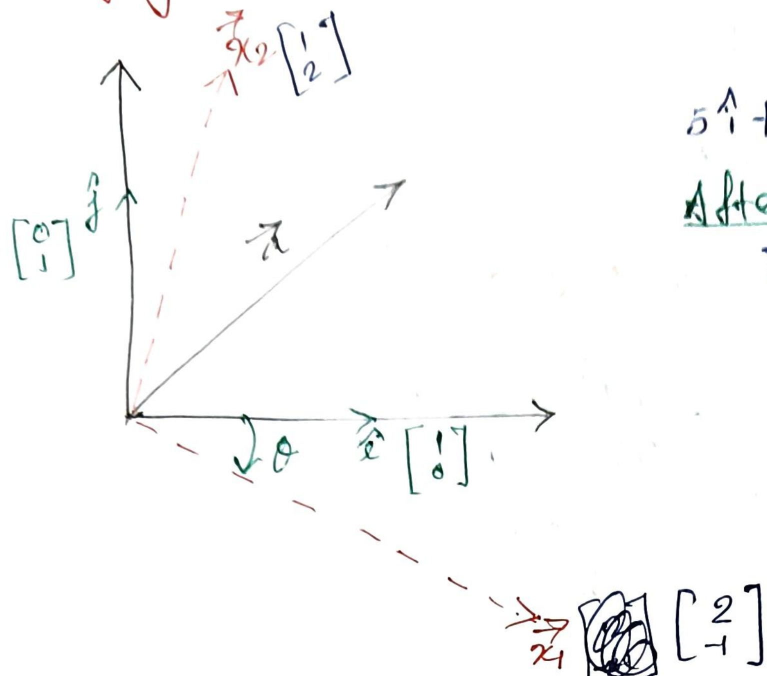
Infinite / Max Norm / Vector Norm

$\|x\|_\infty = 4 \qquad 3\hat{i} + 4\hat{j}$

maximum

$$\boxed{\text{P Norm} = \left(\sum_{i=1}^{N} |x_i|^P\right)^{1/P}}$$

changing coordinates



$5\hat{i} + 7\hat{j} = \begin{bmatrix}5\\7\end{bmatrix}$.

After rotation,

$\vec{x} = \lambda_1 \vec{x_1} + \lambda_2 \vec{x_2}$

We can find if the vectors are orthogonal or not by putting $\theta = 90°$.

$$x_1 x_2 = |\vec{x_1}|\,|\vec{x_2}| \cos\theta$$

$$\cos\theta = \frac{x_1 x_2}{|\vec{x_1}|\,|\vec{x_2}|}$$

$$= \frac{\begin{bmatrix}2\\-1\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix}}{\sqrt{5}\sqrt{5}} = 0$$

$\boxed{\theta = 90°}$

Projection

for x_1

$$\vec{a}\cdot\vec{x_1} = |\vec{a}|\,|\vec{x_1}|\cos\theta \quad \{|\vec{a}|\cos\theta = P\}$$

$$\vec{a}\cdot\vec{x_1} = P|\vec{x_1}|$$

$$\begin{bmatrix}5\\7\end{bmatrix}\cdot\begin{bmatrix}2\\1\end{bmatrix} = P\sqrt{5}$$

$$3 = P\sqrt{5}$$

Direction

$P = 3/\sqrt{5}$ (Projection)

$$\vec{P} = P\hat{x_1} = P\frac{\vec{x_1}}{|\vec{x_1}|} = \frac{3}{\sqrt{5}} \times \frac{1}{\sqrt{5}}\begin{bmatrix}2\\-1\end{bmatrix} = \frac{3}{5}\begin{bmatrix}2\\-1\end{bmatrix}$$

$\cancel{\vec{P} = 0.6}$

$\left[\lambda_1 = \frac{3}{5} = 0.6\right]$

Projection for x_2

$$\vec{a} \cdot \vec{x_2} = P|\vec{x_2}|$$

$$\begin{bmatrix} 5 \\ 7 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 19 = P\sqrt{5}$$

$P = 19/\sqrt{5}$ (magnitude)

Direction

$$P\hat{x_2} = \frac{19}{\sqrt{5}} \frac{\vec{x_2}}{|\vec{x_2}|} = \frac{19}{\sqrt{5}} \times \frac{1}{\sqrt{5}} \vec{x_2}$$

$$= \frac{19}{5} \vec{x_2}$$

$$\boxed{x_2 = \frac{19}{5}}$$

Transformed vector after changing coordinates,

$$\therefore \vec{a} = 0.6\,\vec{x_1} + \frac{19}{5}\vec{x_2}$$

Linearly Independent Vectors

> when no vector can be written as linear combinations of other ~~vs~~ vector.

Linear Combination

$$\vec{v_1} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \vec{v_2} = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} \quad \vec{v_3} = \begin{bmatrix} 0.5 \\ 1 \\ 1.5 \end{bmatrix}$$

$\vec{v_2} = 2\vec{v_1}$, $\vec{v_3} = \frac{1}{2}\vec{v_1}$ (Linearly Dependent)

→ one vector can be written in the form of other.

$$\boxed{k_1 \begin{bmatrix} a \\ b \\ c \end{bmatrix} + k_2 \begin{bmatrix} e \\ f \\ g \end{bmatrix} = 0}$$ {condition}

> In no cases, all the $k_j = 0$

Eg $\vec{v_1} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{v_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\vec{v_3} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

$k_1 = k_2 = k_3 = 0$ (Linearly Independent)

$k\vec{v_1} + k_2\vec{v_2} + k_3\vec{v_3} = 0$

Linearly Independent

$A_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $A_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, $A_3 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$

$2A_2 = A_3$

$k_1 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + k_2 \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$

$k_2 = 2k_1$

$k_1 A_1 + k_2 A_2 + k_3 A_3 = 0$

$k_2 = 2$, $k_3 = -1$

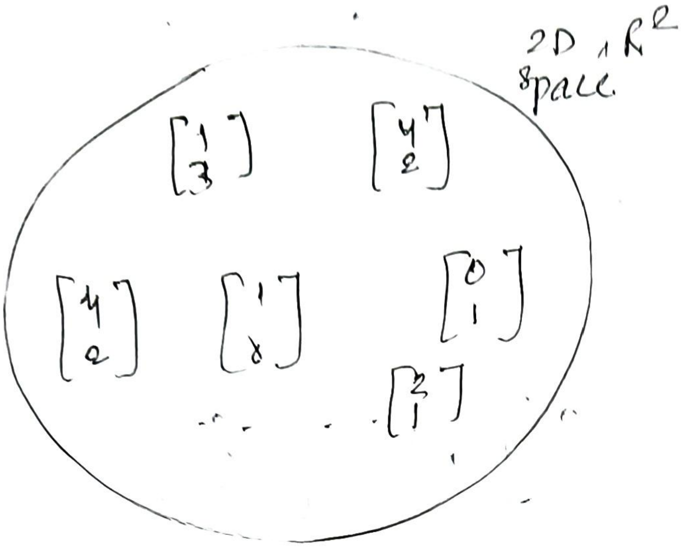
$k_1 A_1 + 2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = 0$

$k_1 A_1 = 0$

$k_1 = 0$, $k_2 = 2$, $k_3 = -1$

> A_1, A_2, A_3 are not linearly independent.

Basis Vectors



$\vec{v_1} = \begin{bmatrix}1\\0\end{bmatrix}, \quad v_2 = \begin{bmatrix}0\\1\end{bmatrix}$

$\begin{bmatrix}2\\1\end{bmatrix} = 2\begin{bmatrix}1\\0\end{bmatrix} + 1\begin{bmatrix}0\\1\end{bmatrix} = 2\vec{v_1} + \vec{v_2}$

$\begin{bmatrix}1\\3\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix} + 3\begin{bmatrix}0\\1\end{bmatrix} = \vec{v_1} + 3\vec{v_2}$

> If every vector ~~can~~ in a given space can be written as linear combination of some vectors and these vectors are independent to each other, then it is called as the Basis vector for the space.

> Space can be any space ($R, R^2, R^3 -- R^n$)

$\vec{v_1} = \begin{bmatrix}1\\0\end{bmatrix}, \quad \vec{v_2} \begin{bmatrix}0\\1\end{bmatrix}$

$K_1 \vec{v_1} + K_2 \vec{v_2} = 0$

if $K_1 = K_2 = 0$. ~~tif~~ K then only $\vec{v_1}$ & $\vec{v_2}$ are linearly dependent but this is not the case. So, $\vec{v_1}$ & $\vec{v_2}$ are linearly dependent.

> Basis vector are not ~~uni~~ unique.

Matrices

> It is collection of number and used as to represent various attributes.

Eg car = $\begin{bmatrix} 1000\ cc \\ 4\ doors \\ Red \end{bmatrix}$

> It is collection of array.

$1x - 1y = 1$
$2x + 1y = 5.$

$$\begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

Input vector   Output vector

After solving

$x = 2, y = 1$

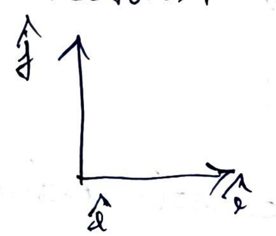
> Matrix is array of number which transform the vector to other values.

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} \xrightarrow{\text{Transformation}} \begin{bmatrix} 1 \\ 5 \end{bmatrix} \quad [\ \ ] \ [\ \ ]$$

$\vec{a} = 2\hat{i} + \hat{j}$, $\vec{u} = \hat{i} + 5\hat{j}$

Let us use Matrix A,

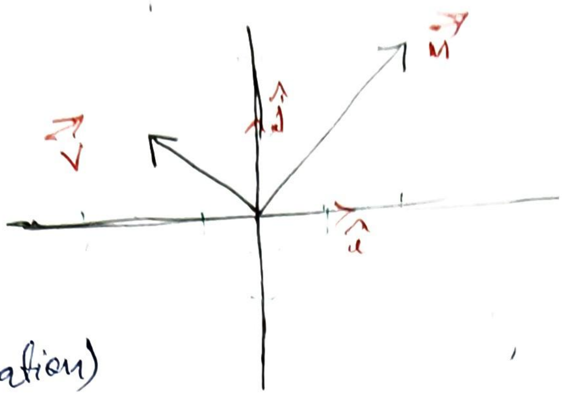
$A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}$ to transform the coordinate axes vectors.



$\hat{i} = 1(x)$
$\hat{j} = 1(y)$
$\hat{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} x$, $\hat{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} y$

$$\underset{A}{\begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}} \underset{\hat{i}}{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \rightarrow \vec{u}$$

$$\underset{A}{\begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}} \underset{\hat{j}}{\begin{bmatrix} 0 \\ 1 \end{bmatrix}} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \rightarrow \vec{v}$$



(change in rotation)

Transformation Properties

① Line will remain line
② Origin will be fixed

> The ~~colums~~ columns of the matrices are the transformed versions of Basis Vectors.

e.g
$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

$$\hat{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \hat{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{(Basis vector)}$$

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

same column

Types and ~~matrices~~ properties of matrices

> Usually, Matrices are used to store & represent the data.
> Matrix is a very natural approach for organizing data.
> Matrices is used to transform the vector.

In General Data is organized as below:
- Rows represent samples.
- Columns represent the values of the attributes/ features/ variables.
- This can be interchanged, however we will use Rows as samples and columns as features.

South Bombay

Eg Find whether the person is Male or Female?

	f_1: Height	f_2: Weight	f_3: Hair length	M/F
R_1	150 cm	45 kg	12"	F
R_2	155cm	60 kg	24"	F
R_3	170 cm	65 kg	6"	M
R_4	185cm	90 kg	8"	F
R_5	173 cm	75 kg	9"	F

Training Data

> Matrices are a set of array to organize the data.
> collection of vectors is also called a matrices.
> vectors are building blocks of matrices.

Eg $150\hat{i} + 45\hat{j} + 12\hat{k}$

> Matrices are made by vectors.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{bmatrix} 150 & 155 & 170 \\ 45 & 60 & 65 \\ 12 & 24 & 6 \end{bmatrix}$$ General way, In datascience Transpose it.

$$v_1 = \begin{bmatrix} 150 \\ 45 \\ 12 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 155 \\ 60 \\ 24 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 170 \\ 65 \\ 6 \end{bmatrix}$$

Addition of Matrices

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$$

> Size of matrices should be same for addition and subtraction.

Negative of matrix

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

Scalar Multiplication

$$2\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 4 & 8 \end{bmatrix}$$

Matrix Multiplication

Order of matrix = $m \times n$

m → no. of rows

n → no. of columns

$$\begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 6 & 7 & 8 \end{bmatrix}_{2\times 3}$$

$$= \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 6 & 7 & 8 \end{bmatrix}$$

$$= \begin{bmatrix} 1\times1+2\times6 & 1\times2+2\times7 & 1\times3+2\times8 \\ 4\times1+5\times6 & 4\times2+5\times7 & 1\times3+5\times8 \\ 1\times1+1\times6 & 9 & 4 \end{bmatrix}_{3\times3}$$

$$= \begin{bmatrix} 13 & 16 & 19 \\ 34 & 43 & 52 \\ 7 & 9 & 4 \end{bmatrix}$$

Types of Matrix

① Square Matrix

Rows = Columns

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 6 & 9 \end{bmatrix}$$

② Symmetric Matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 1 & 4 \\ 3 & 2 & 1 \end{bmatrix}$$

rows = column

Symmetrical.

③ Triangular Matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$$ rows = columns

upper triangular

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{bmatrix}$$ Lower triangular

④ Diagonal Matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$ rows = columns

> only diagonal elements will be present

⑤ Identity Matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$ rows = columns

> diagonal element will be 1

Properties of Addition of Matrix

① $A + B = B + A$ {Commutative}

② $A + (B + C) = (A + B) + C$ {Associative}

Properties of Matrix Multiplication

① $A \cdot B \neq B \cdot A$

② $A(BC) = (AB)C$

③ $(A+B)C = AC + BC$

④ $A(B+C) = AB + BC$

> There exists multiplicative identity.

$IA = AI = A$

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

$\cancel{A}B$

> $A_{3\times 2} \; B_{2\times 2} = (AB)_{3\times 2}$

This two should be equal ~~to~~ for multiplication.

Transpose of Matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

> Interchange rows and columns.

Properties

1. $(A^T)^T = A$
2. $(A+B)^T = A^T + B^T$
3. $(AB)^T = B^T A^T$
4. $(KA)^T = KA^T$

Data Representation

1. Storage

> the image is stored in the machine as a large matrix of pixel values across image.

2. Identification

> Several machine learning algorithm are deployed in order to "teach" the machine as to how to identify a particular image.

Data as MATRIX

Data matrix → Rows → Sample Points

Data matrix → columns → Features

We have understood now how a data can be represented as a matrix, now we have some questions:

① Are all features important?

② Can we reduce the size of data matrix? By removing unnecessary features.

③ How can we find correlation?

④ If yes, how do we find the linear relationship?

Identification of Independent features:

Car → Speed
→ Engine Power (hp)
→ Engine Capacity (cc)
→ Milage (KM/L)
→ Fuel type (P/D)
→ No. of Doors

> correlate using linear dependency and independency of the vectors.

{ Power = Torque × Angular Rotation }

Q How does one identify the no. of independent attributes?

→ Using Domain Knowledge

Mileage = f(Engine Power, Fuel Type, Engine Capacity)

> Now, we ask if the data itself will help us identify these relationships.

> If we don't have domain knowledge then with the help of rank of a matrix we can identify the relationship b/w the features.

Number of independent attributes:

Rank of a Matrix

> It refers to the number of LINEARLY INDEPENDENT rows or columns of matrix.

eg:

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 6 & 0 \\ 5 & 10 & 1 \end{bmatrix}$$

⇒ col2 = 2 × col1

⇒ col 3 is independent.

} Two independent columns
So, Rank = 2

∴ Rank = 2

> Rank work with reduced set of variables.

> We got number of independent features in matrix. Now, shall find out the number of linear dependence relationship?

* this can be found out by Null space & Nullity?

> Rank: No. of independent columns or rows.

$$\boxed{\text{Rank} = \text{no. of columns} - \text{no. of linear dependecies}}$$

> Use function in Python to calculate rank.

No. of Linear Dependecies

Null Space for Matrices

> the null space of a matrix A consists of all vectors B such that, $AB = 0$ & $B \neq 0$.

> No. of vectors in the null space = NULLITY.

> The SIZE $(m \times n)$ of the null space of a matrix provides us with no. of linear relations among features.

> $A = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \\ x_{m1} & \cdots & x_{mn} \end{bmatrix}$ is a data matrix, and there is one vector in the null space of A, i.e., $B = [\beta_1 \ldots \beta_m]^T$, then as per the definition, β_1 satisfies all the equation given below:

$$x_{11}\beta_1 + x_{12}\beta_2 + \ldots + x_{1n}\beta_n = 0$$

$$\vdots$$

$$x_{m1}\beta_1 + x_{m2}\beta_2 + \ldots + x_{mn}\beta_n = 0$$

Rank - Nullity Theorem :

Nullity of A + Rank of A = Total Features

no. of linear relationships — no. of independent features.

eg

① $A = \begin{bmatrix} 1 & 2 \\ 2 & 7 \\ 5 & 6 \end{bmatrix}$ → no. of columns = no. of features = 2

Rank of A = 2 (2 ~~rela~~ independent relationship b/w columns)

→ Here all features are linearly independent

Nullity + 2 = 2

$\boxed{\text{Nullity} = 0}$

> $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 5 & 6 & 7 \end{bmatrix} \begin{bmatrix} \; \\ \; \\ \; \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Number of vector is nullity.

② $A = \begin{array}{c} \begin{array}{ccc} f_1 & f_2 & f_3 \end{array} \\ \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 0 \\ 4 & 8 & 1 \end{bmatrix} \end{array}$

$f_2 = 2f_1$, f_3 → Two independent features

no. of features = 3

Rank = 2

Nullity = 1

Vector In Null Space

> We need to find the vectors in the null space of A which is, Non-zero.

$$A \cdot \beta = 0$$

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 0 \\ 4 & 8 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow b_1 + 2b_2 + b_3 = 0$$

$$2b_1 + 4b_2 = 0$$

$$\Rightarrow b_1 + 2b_2 = 0$$

$$\boxed{b_3 = 0}$$

> The null vector is

$$\beta = [b_1 \; b_2 \; b_3]^T$$

$$= [-2b_2 \; b_2 \; 0]^T$$

$$= b_2 [-2 \; 1 \; 0]^T$$

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} -2b_2 \\ b_2 \\ 0 \end{bmatrix} = b_2 \underbrace{\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}}_{\text{Null vector}}$$

Solving Linear Equations

Q.
$$3x+2y=10$$
$$2x+5y=12$$

$$\begin{bmatrix} 3 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 10 \\ 12 \end{bmatrix}$$

$$AX = B$$

$$\cancel{X = B^{-1}} \quad X = A^{-1}B$$

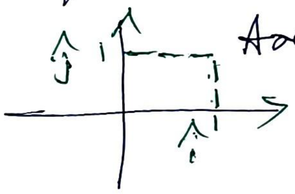
$$X = \begin{bmatrix} 3 & 2 \\ 2 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 10 \\ 12 \end{bmatrix}$$

$$\boxed{A^{-1} = \frac{1}{|A|} \text{Adj}' A}$$ Inverse of A

Determinant

> Area of a square matrix.

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



Area = 1×1 = 1

$$x = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \text{Det } x = ad-bc$$

Adjoint

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

$$A_{11} = (-1)^{(1+1)} \begin{vmatrix} e & f \\ h & i \end{vmatrix}$$

$$A_{12} = (-1)^{(1+2)} \begin{vmatrix} d & f \\ g & i \end{vmatrix}$$

$$A_{13} = (-1)^{(1+3)} \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

$$\text{Adj}' A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{21} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

> $AX = B$ $\quad A(m \times n),\ X(n \times 1),\ B = (m \times 1)$

categories = $m = n$ (Sample and feature are same)
$m > n$
$m < n$

Case 1:

$m = n$

A is full rank, $|A| \neq 0$
$\downarrow$
$x = A^{-1}b$
unique solution

A is not full rank

Consistent (∞ sol^n)

Inconsistent (No sol^n)

eg
$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 10 \end{bmatrix}$$

$|A| \neq 0$
rank$(A) = 2$ = no. of columns (2 columns are linearly independent)
$\therefore$ A ~~is~~ is full rank.

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 7 \\ 10 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 1.5 \\ 1 & -0.5 \end{bmatrix} \begin{bmatrix} 7 \\ 10 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$(x_1, x_2) = (1, 2)$

$A_{11} = (-1)^2 4$
$A_{12} = -2$
$A_{21} = (-1)^3 3$
$A_{22} = (-1)^2 2$

$$\text{Adj } A = \begin{bmatrix} 4 & -3 \\ -2 & 2 \end{bmatrix}^T = \begin{bmatrix} 4 & -2 \\ -3 & 2 \end{bmatrix}$$

eg. $\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 10 \\ 19 \end{bmatrix}$

$C_2 = 2C_1$

$|A| = 0$

rank$(A) = 1$, nullity $= 1$

∴ 1 linearly independent column.

$\begin{bmatrix} x_1 + 2x_2 \\ 3x_1 + 6x_2 \end{bmatrix} = \begin{bmatrix} 10 \\ 19 \end{bmatrix}$

one cannot find solⁿ (x_1, x_2)

$2R_1 = 3x_1 + 6x_2 = 19$

∴ Equations are inconsistent

eg $2x + 3y = 0$

$K_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + K_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0$

$K_1 = 0,\ K_2 = 0$

∴ Infinite solution.

~~x and~~ y

case 2: $m > n$

> Instead of identifying a solⁿ to $Ax - b = 0$, we identify on x such that $(Ax - b)$ is minimized.

> $(Ax - b) = e(m)$, there are m errors e_i, $i = 1:m$.

optimization

$$\text{minimize} \sum_{i=1}^{m} e_i^2$$

→ minimize $(Ax-b)^T (Ax-b)$

$\min[(Ax-b)^T (Ax-b)]$

$\min[(b^T - x^T A^T)(Ax-b)]$

$\min[(x^T A^T A x - 2b^T A x + b^T b) = f(x)$

Explanation

> $AX = B$

$X = A^{-1}B$

min $(AX - B) \approx 0$ (tending to 0)

$(AX-B)^2 = \Sigma e^2$

$A \cdot A^T = A^2 \ [\ \ ][\ \ ]$

> Solving the optimization problem as a function of x, will result in a solⁿ of x.

> the solⁿ to this optimization problem is obtained by differentiating f(x) w.r.t x & setting the ~~diff~~ differential to zero.

$\Delta f(x) = 0$

> Differentiating f(x) and setting the ~~diff~~ differential to zero results in

$2(A^TA)x - 2A^Tb = 0$

$\therefore \boxed{(A^TA)x = A^Tb}$

> Assuming that all the columns are linearly independent

$\boxed{x = (A^TA)^{-1} A^T b}$ → unconstrained optimization solⁿ.

eg: case 2

$$\begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.5 \\ 5 \end{bmatrix}$$

$m = 3, n = 2$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$$ after applying $x = (A^TA)^{-1}A^Tb$ formula.

Substituting:

$$\begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix} \neq \begin{bmatrix} 1 \\ -0.5 \\ 5 \end{bmatrix}$$ > with error.

e.g. $\begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$

$x = (A^T A)^{-1} A^T b$

$x_1 = 1, \; 2x_1 = 2, \; 3x_1 + x_2 = 5$

$3(1) + x_2 = 5$

$\Rightarrow \quad (x_1, x_2) = (1, 2)$

use $x = (A^T A^{-1}) A^T b$ and find the value of x.

Case 3: $m < n$

> Here the features are more than samples.
> Since the no. of features is greater then no. of equations, one can obtain/get multiple solutions. There is infinite solution case.
> How can we choose single solution from the set of infinite solutions?
> this is the example of constrained Optimization problems:

Optimize:

$$\min\left(\frac{1}{2} x^T x\right) \text{ such that } Ax = b$$

Let a Lagrangian function $f(x, \lambda)$

$$\min\left[f(x, \lambda) = \frac{1}{2} x^T x + \lambda^T (Ax - b)\right]$$

differentiate w.r.t x and set to zero, gives

$$x + A^T \lambda = 0$$

$$x = -A^T \lambda$$

multiplying by A

$$Ax = b = -AA^T \lambda$$

Hence, we get $\lambda = -(AA^T)^{-1} b$ assuming that all the rows are linearly independent.

and $\boxed{x = -A^T (AA^T)^{-1} b}$

e.g.:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$x = A^T (AA^T)^{-1} b$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -0.2 \\ -0.4 \\ 1 \end{bmatrix}$$ with some errors

Generalization of Concepts

- The cases cover all the scenarios one might encounter while solving linear equations.
- ~~Can~~ there be any form in which the results obtained for all the 3 cases can be generalized?
- The concept we use to generalize the solutions is called as ~~MOORE-PENSO~~ MOORE-PENROSE PSEUDO-INVERSE OF A MATRIX.

- The pseudo inverse is used as below:

$$AX = B \rightarrow \text{probs}$$

$$X = A^{+}B$$ {SINGULAR VALUE DECOMPOSITION} can be used to calculate the pseudo inverse (A^{+}).

Recap

$Ax = b$

- $m = n$
  - A is full rank, $|A| \neq 0$ → unique sol^n
  - A is not full rank
    - consistent solution (∞ sol^n)
    - Inconsistent solution (no solution)
- $m > n$ → probs? → $x = A^{+}B$
- $m < n$ → $x = A^{-1}b$

Vectors recap :

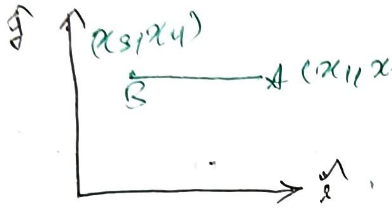
$\cancel{A + b\hat{j}}$ $\vec{A} = a\hat{i} + b\hat{j} + c\hat{k} \rightarrow 3D$

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \rightarrow 3D$$

Magnitude, $|A| = \sqrt{a^2 + b^2 + c^2}$

direction $= \frac{a}{|A|}\hat{i} + \frac{b}{|A|}\hat{j} + \frac{c}{|A|}\hat{k}$

Length of vectors



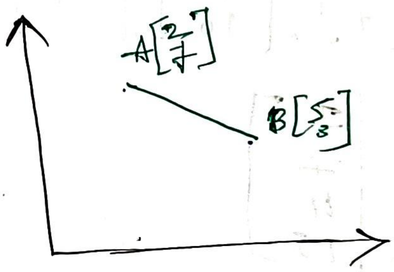
$A = x_1\hat{i} + x_2\hat{j}$

$B = x_3\hat{i} + x_4\hat{j}$

Distance b/w A & B

$= \sqrt{(x_3 - x_1)^2 + (x_4 - x_2)^2}$

eg $A = \begin{bmatrix} 2 \\ 7 \end{bmatrix}$, $B = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$



length $= \sqrt{(5-2)^2 + (3-7)^2}$

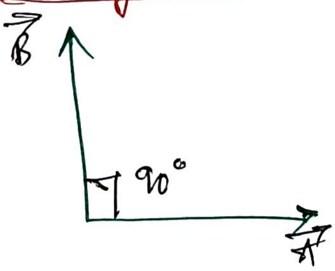
$= 5$ unit

direction =

$A, \quad \frac{2\hat{i}}{\sqrt{2^2 + 7^2}} + \frac{7\hat{j}}{\sqrt{2^2 + 7^2}}$

$B, \quad \frac{5\hat{i}}{\sqrt{5^2 + 3^2}} + \frac{3\hat{j}}{\sqrt{5^2 + 3^2}}$

Orthogonal vectors



$\vec{A} \cdot \vec{B} = 0$

$\cos\theta = 0$

$\theta = 90°$

$\boxed{A \cdot B = A^T B = 0}$

Eg.

$\vec{A} = a\hat{i} + b\hat{j} + c\hat{k}$

$\vec{B} = d\hat{i} + e\hat{j} + f\hat{k}$

$\vec{A} \cdot \vec{B} = ad + be + cf$

$= \cancel{\begin{bmatrix} a \\ b \\ c \end{bmatrix}} \begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} d \\ e \\ f \end{bmatrix} = ad + be + cf$

Basis vector

> It is a linear ~~dep.~~ independent vector

Que Find out the basis vector

$\begin{bmatrix} 6 \\ 3 \\ 8 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 9 \\ 4 \\ 7 \\ 10 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 8 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 14 \\ 7 \\ 12 \\ 17 \end{bmatrix}, \begin{bmatrix} 11 \\ 8 \\ 13 \\ 18 \end{bmatrix}, \begin{bmatrix} 7 \\ 0 \\ 1 \\ 2 \end{bmatrix},$

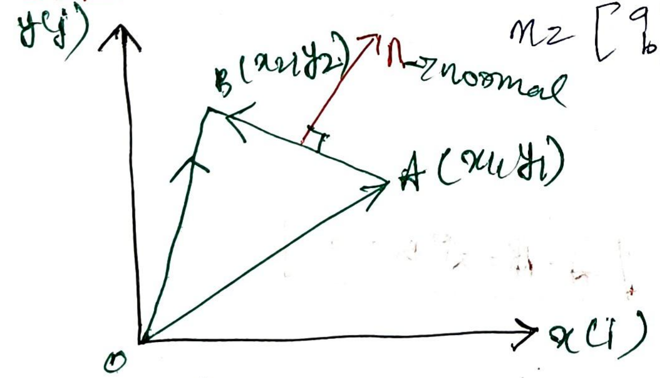
$\begin{bmatrix} 2 \\ -3 \\ -4 \\ -5 \end{bmatrix}, \begin{bmatrix} 7 \\ 7 \\ 4 \\ 15 \end{bmatrix},$

Ans wherever rank = 2, is basis vector.

→ Basis vector can ~~space~~ span the entire space.

Equation of line & plane

$n = \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = w.$



Eqn of line:

$ax + by + c = 0$

$w_1x + w_2y + c = 0$

In vector form:

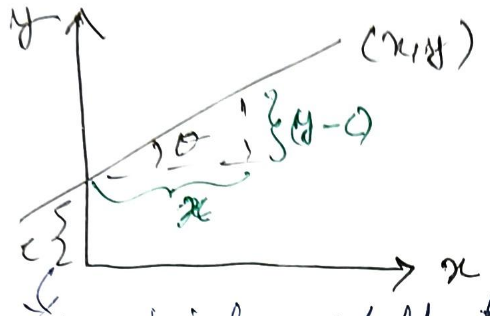
$[w_1, w_2] \begin{bmatrix} x \\ y \end{bmatrix} + c = 0$

$\text{Equ}^n = W^T X + c$

where $X = \begin{bmatrix} x \\ y \end{bmatrix}$

w = weights

Equ^n of line



y-axis intercept / offset

$\tan\theta = \frac{y}{x} = m$ (slope)

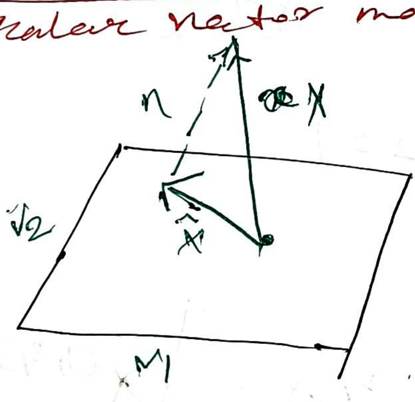
$m = \frac{y - c}{x}$

$y - c = mx$

$\boxed{y = mx + c}$

Projections

- PCA, Scalar vector machine.



> v_1, v_2 are basis vector

> We can define the projection $(\hat{X})$ of a vector (X) onto a lower dimension (2D).

$\hat{X} = c_1 v_1 + c_2 v_2$

using vector addition:

$X = c_1 v_1 + c_2 v_2 + n$

> Projections onto general directions (2D in this case)

$v_1^T n = 0$ {dot product will be zero as they are $\perp$ to each other}

$v_1^T(x - (c_1 v_1 + c_2 v_2)) = 0$

$v_1^T x - c_1 v_1^T v_1 = 0$

$$\hat{x} = \underbrace{\frac{v_1^T x}{v_1^T v_1}}_{c_1} v_1 + \underbrace{\frac{v_2^T x}{v_2^T v_2}}_{c_2} v_2$$

eg

$x = [1\ 2\ 3]^T$

Projecting this vector onto the space spanned by the vectors

$v_1 = [1\ -1\ -2]^T$

$v_2 = [2\ 0\ 1]^T$

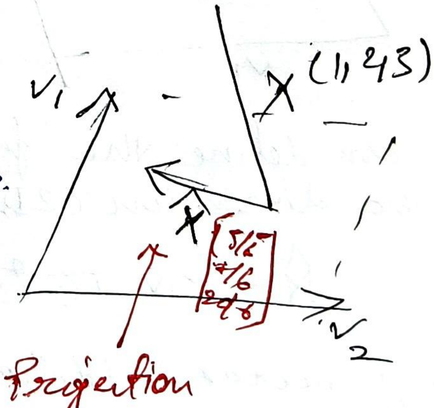
~~$c_1 = [1\ 2\ 3]$~~ $c_1 = \dfrac{v_1^T x}{v_1^T v_1}$, $c_2 = \dfrac{v_2^T x}{v_2^T v_2}$

$c_1 = -7/6$, $c_2 = 1$

$$\hat{x} = -\frac{7}{6}\begin{bmatrix}1\\2\\3\end{bmatrix} + 1\begin{bmatrix}1\\2\\3\end{bmatrix}$$

$$\hat{x} = \begin{bmatrix}5/6\\7/6\\20/6\end{bmatrix}$$ → component in 2D space.

Projected vector



> Reduced from 3D to 2D.

> Projection from 3D to 2D.

Hyperplanes & Halfspaces

> Geometrically, hyperplane is a geometric entity whose ~~direction~~ dimension is one less than that of its ambient space.

> 3D Space → 2D Space / (n-dimension to n-1 dimension).

> The hyperplane is usually described by an eqn as follows:

$$\boxed{x^T n + b = 0}$$

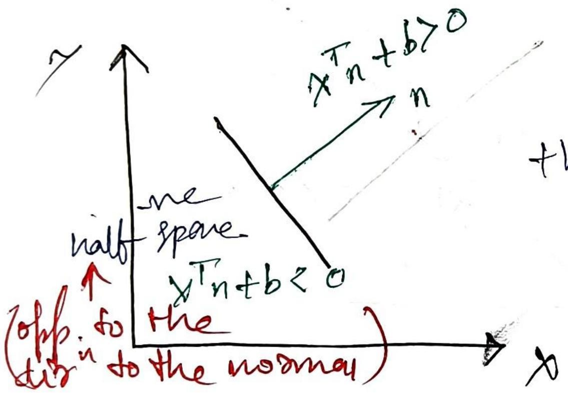
> Any plane above 2D is called hyperplane.

Halfspace

> We can analyse $x^T n + b = 0$ for 2 halfspace:

$x^T n + b = 0$ for all $x \in$ line

$x^T n + b > 0$ for all x is n direction



the half space ← (in the direction to the normal)

eg

$$n = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \; b = 4$$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$x^T n + b = 0$$

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 4 = 0$$

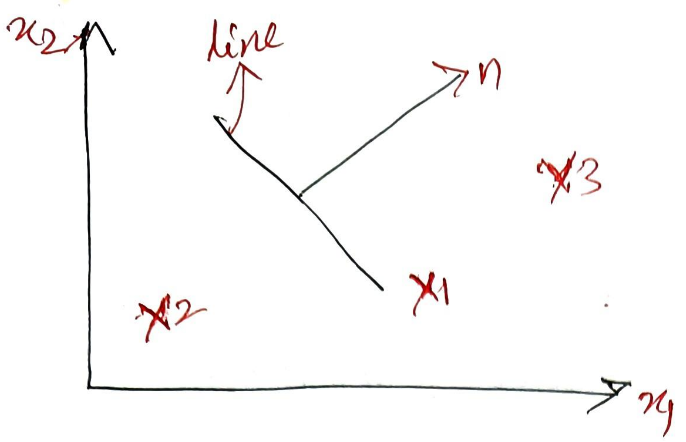
$$x_1 + 3x_2 + 4 = 0$$

$\therefore$ $x_1 + 3x_2 + 4 > 0$ ← Positive halfspace for all x values

$\therefore$ $x_1 + 3x_2 + 4 < 0$ ← Negative halfspace for all x values.

Q How will we determine the direction of normal?

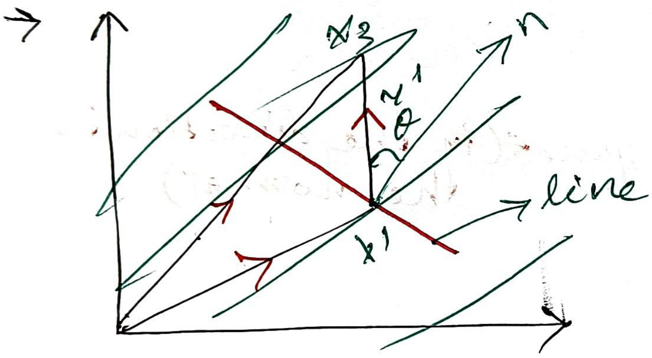
Ans Normal can be in direction, just the orientation will change.

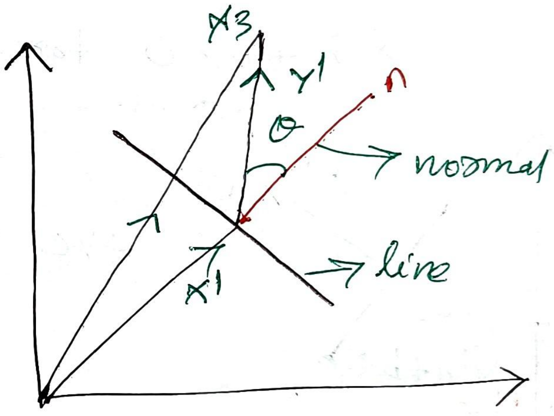


$X_1^T n + b = 0$ (X_1 lie on the line itself)

$X_2^T n + b =$ Below is the derivation

$X_3^T n + b =$ Below is the derivation.





$X_3 = X' + Y'$

$X_3^T n + b = (X' + Y')^T n + b$

$= X'^T n + Y'^T n + b$

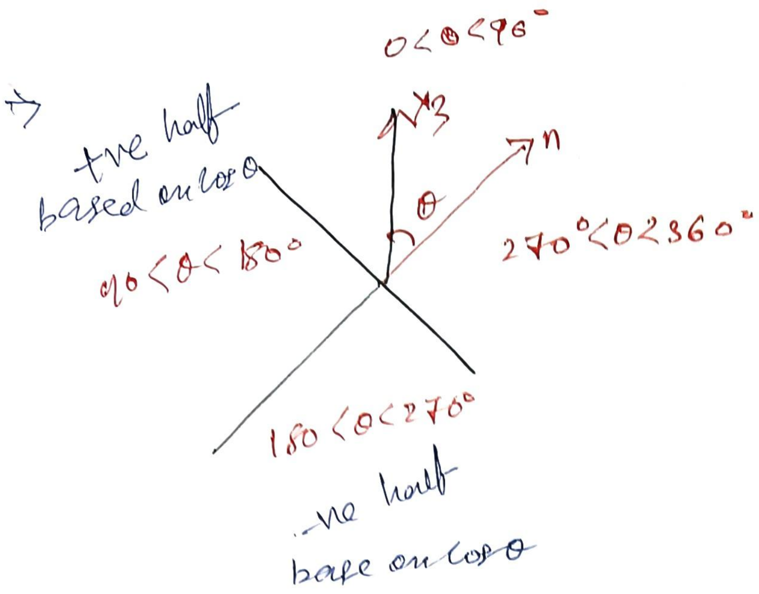
$= \underbrace{X'^T n + b}_{0} + Y'^T n$

(X' lie on the line)

$= Y'^T n$

$= Y_c \cos\theta$

→ θ will change depends on the vector direction.



Eigenvalues & Eigenvectors

> We say linear eqn of the form $AX = B$

Q. What is the geometrical interpretation of this eqn?

Ans. When vector X is operated on by A, we get vector B in a new direction.

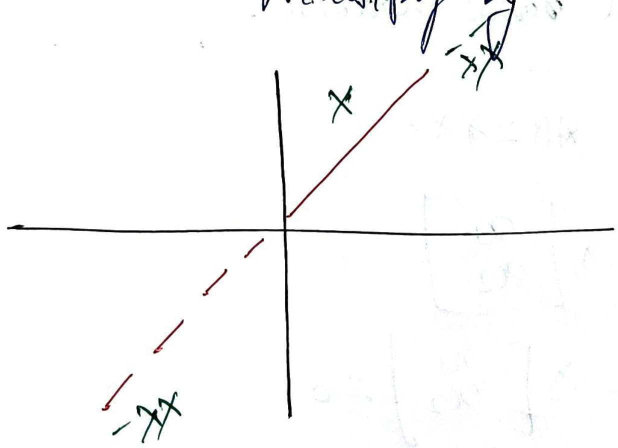
eg $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ = new vector
A 2×2, X 2×1, new vector 2×1 ↓ diff direction

$$X \longrightarrow \boxed{\quad A \quad} \longrightarrow B$$

Q. Are there directions for a matrix A such that, when the matrix operates on these directions they maintain their directions?

$$X \longrightarrow \boxed{\quad A \quad} \longrightarrow \lambda X$$

multiply by scalar



→ Transformed matrix, such that direction remain same.
→ Find a scalar such that magnitude changes but direction remain same.

> The mathematical answer to the previous question is,

$$\cancel{AX =}$$

$$\boxed{AX = \lambda X}$$

> if $0 < \lambda < 1$, then,
A shrinks

> if $\lambda > 1$, then
A elongates.

∴ The solutions (X) are known as eigenvectors (null space) and their corresponding λ are known as eigenvalues.

> Eigenvectors should have unit magnitude.

Find eigenvalues

$AX = \lambda X$, $A(n \times n)$, $X(n \times 1)$

$AX - \lambda X = 0$

$(A - \lambda I)X = 0$, X is in null space of $(A - \lambda I)$

Thus, the eigenvalues can be found by,

$|A - \lambda I| = 0$

> Rank of null space is atleast 1.
> Nullity is atleast 1.
> $A - \lambda I$ is not a full rank matrix.

eg $A = \begin{bmatrix} 8 & 7 \\ 2 & 3 \end{bmatrix}$, $AX = \lambda X$

$$\begin{bmatrix} 8 & 7 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 7 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\left\{ \begin{bmatrix} 8 & 7 \\ 2 & 3 \end{bmatrix} - \lambda \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{I = \text{Identity matrix}} \right\} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$|A - \lambda I| = 0$$

$$\begin{vmatrix} 8-\lambda & 7 \\ 2 & 3-\lambda \end{vmatrix} = 0$$

$$(8-\lambda)(3-\lambda) - 14 = 0$$

$\lambda = \lambda_1, \lambda_2$ (Eigen values)

$\lambda = (10, 1)$

Now, we will find eigen vectors,

> if $\lambda = 1$

$$\begin{bmatrix} 8 & 7 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} 8x_1 + 7x_2 \\ 2x_1 + 3x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$x_1 + x_2 = 0$$

$\therefore$ Thus the eigen vector (unit) corresponding to $\lambda = 1$ is

$$x = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$, unit magnitude.

> if $\lambda = 10$,

$$\begin{bmatrix} 8 & 7 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 10 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} 8x_1 + 7x_2 \\ 2x_1 + 3x_2 \end{bmatrix} = \begin{bmatrix} 10x_1 \\ 10x_2 \end{bmatrix}$$

$$7x_2 = 2x_1$$

$$7x_2 - 2x_1 = 0$$

$$x = \begin{bmatrix} 7/\sqrt{53} \\ 2/\sqrt{53} \end{bmatrix}$$, unit magnitude.

> Eigenvector should have unit magnitude.

Properties of Eigenvalues and Eigenvectors (they don't change direction)

> If the matrix is symmetrical (yield same matrix after transpose) then the eigenvalues are always real.
Then eigenvectors are also real.

> there will be always be 'n' linearly independent eigenvectors for symmetric matrices.

> Symmetric matrices have very important role in data science.

> Eigenvectors corresponding to zero eigenvalues span the null space.

> Eigenvectors corresponding to non-zero eigenvalues span the column space for symmetric matrices.

Eg

$$A = \begin{bmatrix} 0.36 & 0.48 & 0 \\ 0.48 & 0.64 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

eigenvalues:

$\lambda = (0, 1, 2)$

Nullity (No. of linear relationship)

Nullity = 1

eigenvectors:

$$u_1 = \begin{bmatrix} -0.8 \\ 0.6 \\ 0 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 0.6 \\ 0.8 \\ 0 \end{bmatrix}, \quad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$\therefore$ Rank = 2

check:

$$\begin{bmatrix} 0.36 & 0.48 & 0 \\ 0.48 & 0.64 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} -0.8 \\ 0.6 \\ 0 \end{bmatrix} = 0 \;,$$

> u_2 and u_3 can be used to represent all the matrices.