S. No.	Date	Title	Page No.	Teacher's Sign

Applications :- ① Self driving Car.
② Google Home devices: Alexa and Siri.
③ Amazon & Apple's digital voice assistance
④ Text-to-Speech Synthesis
⑤ language identification.
⑥ English-to-French translation.
⑦ OCR
⑧ Arabic handwriting recognition.
⑨ Image Caption generation.
⑩ Video to textual description (Annotation to video)
⑪ Syntactic Parsing for NLP.

AI, ML, DL and Data Science

AI → m/c ability to think and learn on itself, It is capable of developing skills at its Practise & gain experience.
- It Purpose → support & enhance human activities.
- enable m/c to think without human intervention, & take decisions.

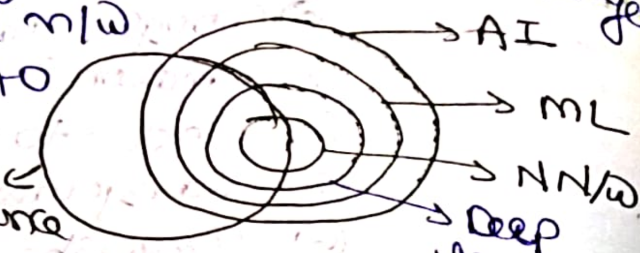
ML → m/c learning is a subset of AI.
→ Provide statistical tools to explore & analyze the data.
→ Method of data analysis which automates analytical model building.
→ branch of AI is based on idea that systems can learn from data, identify Patterns & make decisions with minimum human interventions.

Neural N/W → Make use of neurons that are used to transmit data in the form of I/P values & O/P values.
→ They are used to transfer data by using n/w of connections.

Deep learning → Composed of several hidden layer while neural n/w consists of upto 3 layers.

↓

Accuracy increases than layer is to be increased.



Computational Intelligence. (heart of AI Product)

→ DL is subset of ML.
→ Mimics the human brain.
→ In DL, Create architecture called multi-neural n/w architecture.
→ Idea of Deep neural n/w is to mimic human brain
→ eg:- Neural N/w & DLearning lies in heart of Product.

Data Science → work for all above AI, ML, DL using mathematical tools such as Statistics, Prob., linear algebra, Diff. Calculus etc.

References:-

① Charu C Aggrawal, "Neural N/w & Deep learning".
② Josh Patterson & Adam Gibson "Deep learning".

→ Automated image Sharpening
→ Automated image upscaling
→ WaveNet: generating human speech that can imitate anyone's voice.
→ Speech ~~recog~~ reconstruction from Silent video
→ Image autofill for missing regions.
→ Automated image Captioning.
→ Turning hand-drawn doodles into Stylized ~~art~~ artwork.

Scanned by CamScanner

3 layers of ANN :-

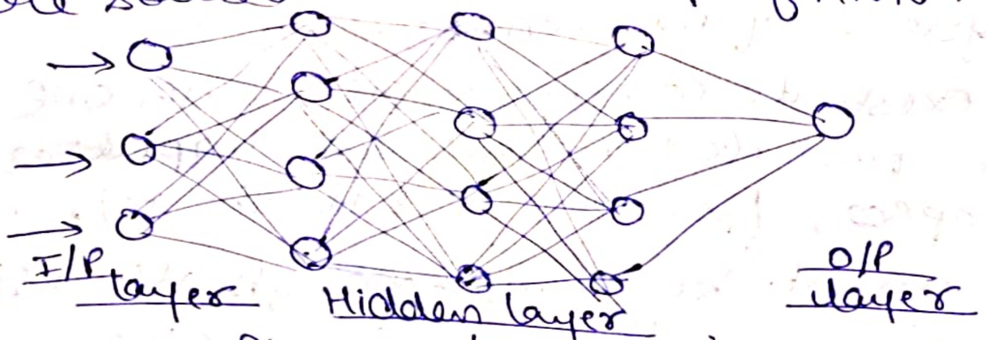
I/P layer
1 Hidden layer
O/P layer.

Deep learning
→ I/P layer
Where NN/w have multiple Hidden layer
→ O/P layer.

Deep learning :- It is a subset of ML. This got created for the m/c to learn to think how actually human brain learn to think.
→ we create architecture called multi-neural n/w architecture. Idea of deep neural

3 Technique used by Deep learning :-

① ANNs : Problems that involved with the no. are solved with the help of ANN.



② CNN : If the I/P is in images, then we use CNN. Advanced CNN known as Transfer learning.

③ Recurrent NN (RNN) : If data is time series kind of data, we use RNN.

Introduction To Neural Network

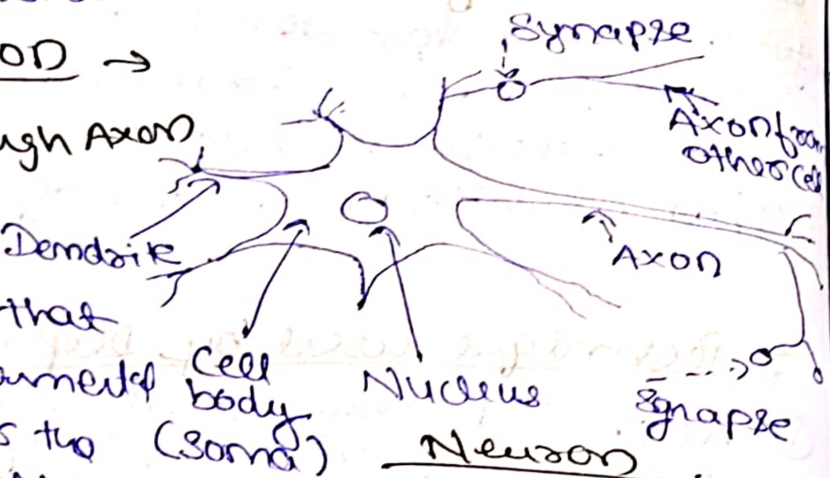
Introduction (Josh) :→ ANN are m/c learning techniques which simulates the mech of learning biological organisms.

→ The human nervous system contains cells, referred as neurons.

The Biological Neuron →

→ neuron send o/p through Axon

→ Dendrite's i/p to neuron



→ It is a nerve cell that provides the fundamental functional units for the nervous system of all animals.

→ Neurons exist to communicate with one another, & pass electro-chemical impulses across synapses, from one cell to the n/w.

→ impulse must be strong enough to activate the next neuron by release of chemicals across a synaptic bifurcation.

→ The strength of impulse must exceed a min threshold else chemicals will not be released.

→ The neuron made up of a nerve cell consisting of a soma (cell body) that has many dendrites but only one axon.

→ Axons are nerve fibers with a special cellular extension that comes from the cell body. the single axon can branch 1000 of times.

→ Dendrites are thin structures that arise from main cell body.

<u>Synapses</u> → Are connection Junction b/n axon & dendrites.

The synapses send signals from axon of a neuron to dendrite of another neuron.

<u>Dendrites.</u> → It have fibers branching out from the soma in a bushy n/w around the nerve cell.

→ It allow the cells to receive signals from connected neighbouring neurons.

→ each dendrite is able to perform multiplication by that dendrite's weight value.

→ Here, multiplication means an increase or decrease in ratio of synaptic <u>neurotransmitters</u> to signal chemicals introduced into the dendrites.

Chemical substance released out of neuro-fiber

<u>Axon</u> → It is single, long fiber extending from the main soma.

→ Its size is 1 cm in length & 100th times diameter of soma.

→ They ~~strength~~ stretch out longer distances than dendrites.

→ The axon will branch 100 of times & connect to other dendrites.

→ neurons are able to send electrochemical signal that travels along the cell's axon & activates synaptic connections with other neurons.

→ Info. flow across the biological neuron

→ Synapses that ↑ the potential & those that decrease the potential.

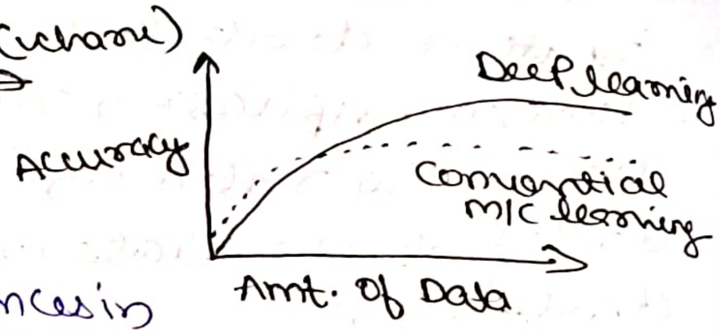
→ neurons also have been shown to form new connections over time & even migrate.
→ These combined mechanics of connection change drive the learning process in the biological brain.

Total Connections in human brain :-

→ Biological NN are roughly 86 billion neurons connected to many other neurons.
→ more than 500 trillion connections b/w neurons in human brain.

ANN → 80 billion neurons.

Human vs Computers → (chart)



→ Smaller data handled previously, but now "big data" era $\xrightarrow{\text{enabled}}$ advances in data collection technology.
→ GPU gave efficient processing on large data set.

Neural N/W (from Josh) → They are computational model that shares some properties with the animal brain in which many simple units are working in parallel with no centralized CU.
→ The behaviour of NN is shaped by its n/w architecture.
→ Network's architecture defined as :- No. of neurons.
No. of layers.
Types of connection b/w layers.

From biological to artificial :- The animal brain is fundamental components of the mind.

→ Study basic components of brain & understand them.

→ Research shown ways to map out functionality of the brain & track signals as they move through neurons.

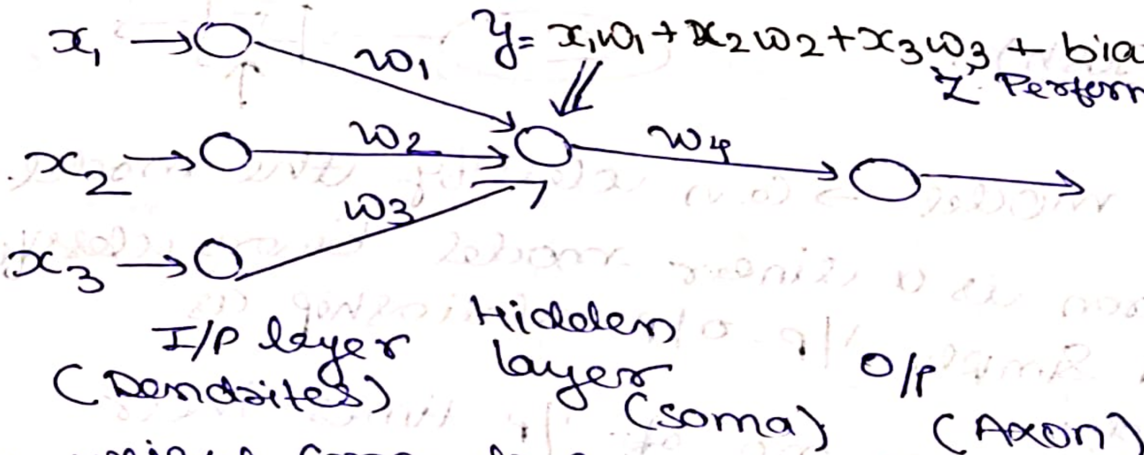
→ Having known how a biological neuron works, let's look at modeling the neuron with beginning of the Perceptron.

Perceptron → The Perceptron is a linear model used for binary classification. In the field of NNW the Perceptron is considered an artificial neuron.

$y = x_1w_1 + x_2w_2 + x_3w_3 + \text{bias}$

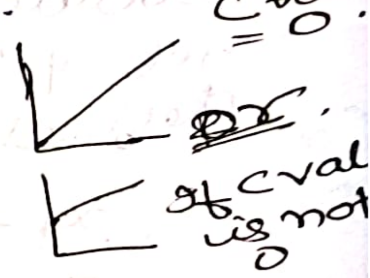
Z Perform = Activation function(y)

$= Z \cdot W_4$



→ weight corr. to every link depending on ↑ & ↓ of chemical components.

Bias components → $y = mx + c$



About to learn :-

① Forward Propagation
② Activation function
③ Backward Propagation
④ loss function
⑤ Cost function

① Summation } To Process is done in neuron.
② Activation

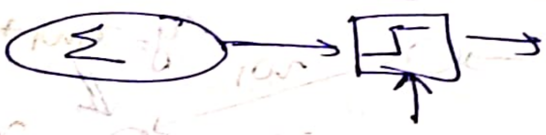
↓

Activation function :- ⓐ Sigmoid → $\frac{1}{1+e^{-y}}$
(for getting 'Z' value) ⓑ ReLU

increase in chemical substance (intensified)
↓
weight increase → y value increase
↓
For to do Action taken ← Value of Z = 1
changes are more.

<u>Binary classifier</u> → 0 or 1 is output

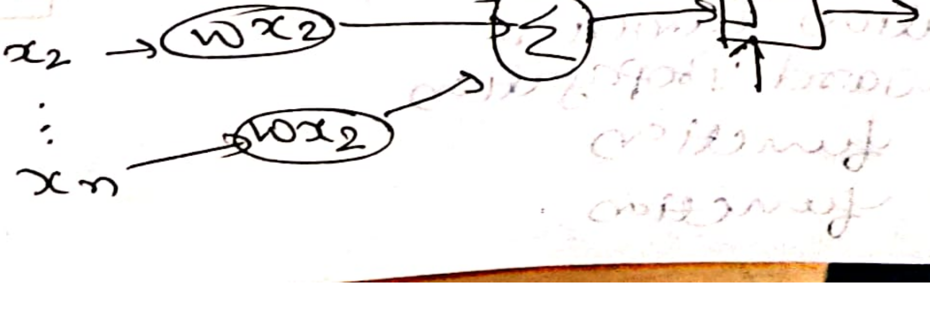
<u>Diagram (Josh)</u> :-



linear model → can classify the model.

→ Perceptron is a linear model binary classifier with a simple I/p-O/p relationship as depicted.

→ Summing n no. of I/p times their associated weights & then sending this "net I/p" to a step function with a defined threshold (0.5)



Neural N/W Phases → Training
→ learning.

→ With Perceptron, this is step function with a threshold value of 0.5. This function will O/P a single binary value (0 or 1) depending on I/P.

→ we can model the decision boundary & the classification O/P in the step function equation, as follows:- (By ReLU function).

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & x >= 0 \end{cases}$$

→ To Produce the net I/P to the activation function.

→ We take dot Product of the I/P & connection weights.

→ Summation in the left half

function Parameter	Description.
n	no. of I/P to Perceptor.
b	Bias termed & move away decision boundary from origin
w	

2 more to write.

→ The O/P of Step function (activation function) is the O/P for the Perceptron & gives us a classification of I/P value.

→ If bias value = -ve it forces the learned weights sum to be a much greater value to get a 1 classification O/P.

→ The bias term in this capacity moves the decision boundary around for the model.
→ I/P value do not affect

The Perceptron learning Algorithm :-

This algorithm will not terminate if the learning I/P is not linearly separable.
→ A linearly separable dataset is one for which we can find the values of a hyperplane that will clearly divide the 2 classes of Dataset.
→ The Perceptron learning algo initializes the weight vector with small random values or 0.08 at the beginning of training.
→ The Perceptron learning algo. takes each I/p record, & computes o/p classification to check against the actual classification label.
→ To Produce the classificatn, the columns (features) are matched up to weights where

n → no. of dimensions in both our I/P & weights.
→ The first I/P value is the bias I/P, which is always 1.0 bcoz we don't affect the bias I/P.
→ 1st weight is our bias term in diagram.
→ The dot Product of the I/p vector & the weight vector gives us the I/P

to our activation function.

→ If the classification is correct, no weight changes are made. If the classification is incorrect, the weight are adjusted accordingly.

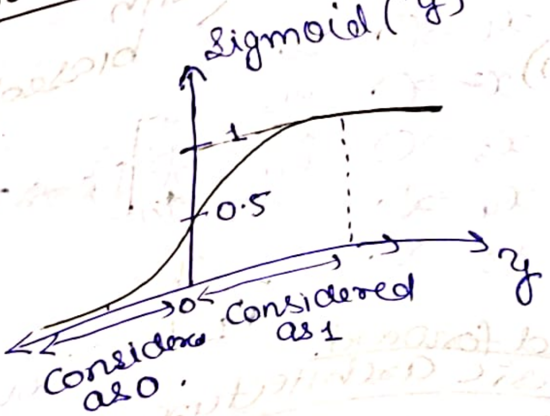
→ weights are updated b/w individual training examples in an 'online learning' fashion.

→ This loop continue until all the i/p eg. are correctly classified.

Types of Activation function :-

1. Linear
2. Non-linear
3. Sigmoid (logistic)
4. Tanh
5. ReLU (Rectified linear unit)

(I) Graph of Sigmoid :-



$f(x) = \frac{1}{1+e^{-y}}$ 0.5 threshold.

If $y = 10$ → $\times 10^{-5}$

$y = 100$ → $\times 10^{-44}$

If $y \uparrow$ we get ⇒ $f(x) = 1$ (Always Activate to take reaction.)

If $y \downarrow$ we get ⇒ $\frac{1}{\text{large value}} = 0$. (take NO action)

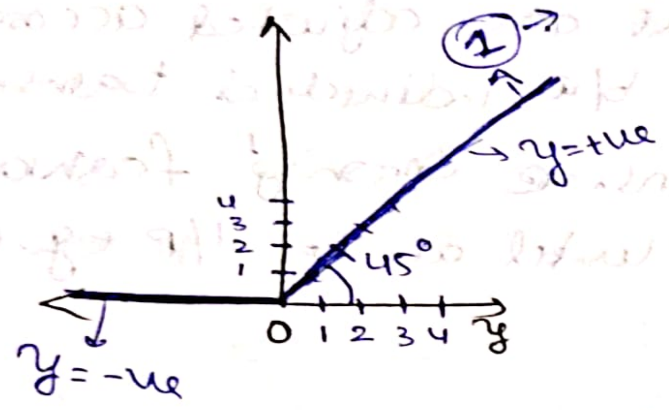
(II) ReLU. → $\max(y, 0)$

If $y = -ve$ $\max(-ve, 0)$ ⇒ 0.

$y = +ve$ $\max(+ve, 0)$ ⇒ +ve.

Scanned by CamScanner

ReLU :-



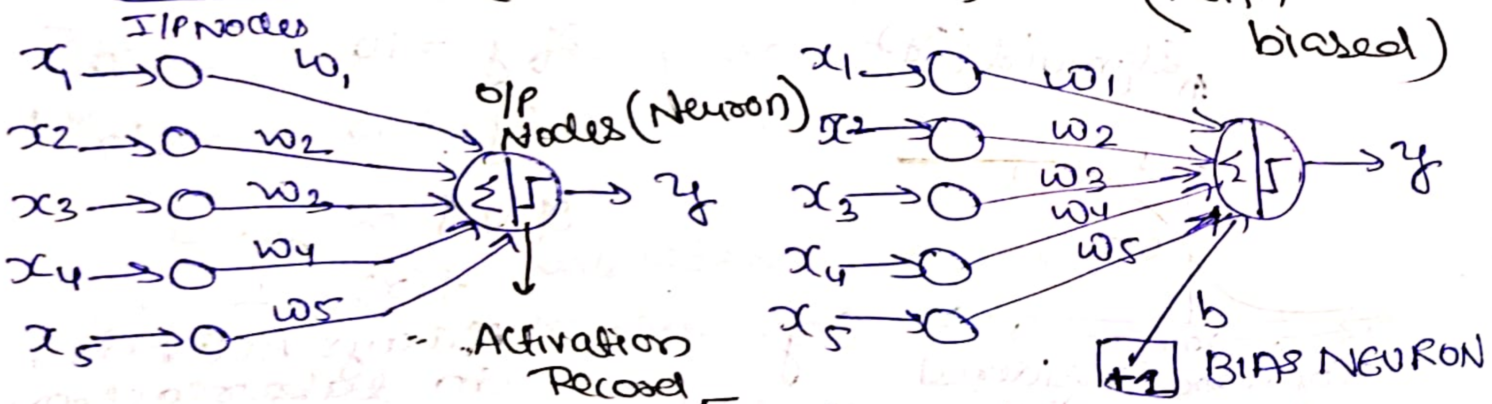
Charry

The Basic Architecture of N/w N/w :-

→ Single computational layer : The Perceptron :-

One would have historical cases in which the class variable is observed & other (current) cases in which the class variable has not yet been observed but needs to be Predicted.

(without biased) (with biased)



Feed forward
Single layer, Basic architecture of Perceptron

→ bias function / value allow value to shift either to left or right.

→ The I/P layer contains d nodes that transmit the d features $\bar{X} = [x_1 \ldots x_d]$ with edges of weight $\bar{W} = [w_1 \ldots w_d]$ to an O/P node.

- I/P layer does not Perform any Computation in its own.
- The linear function $\boxed{\overline{W} \cdot \overline{X} = \sum_{i=1}^{d} w_i \cdot x_i}$ is Computed at O/P node.
- Subsequently, the Sign of this real Value is used in order to Predict the dependent Variable of X.

∴, the Prediction $\hat{y}$ is computed as follows :

$\boxed{\hat{y} \rightarrow \text{Predicted value}}$ $\boxed{y \rightarrow \text{Actual value/observed value}}$

$$\hat{y} = \text{Sign}\{\overline{W} \cdot \overline{X}\} = \text{Sign}\{\sum_{j=1}^{d} w_j \cdot x_j\} \quad \text{Eq 1.1}$$

is always 1.

→ The Sign function maps a real value to either +1 or −1, which is appropriate for binary Classification.

training → Observed value.

learning →

→ Note → the Cap mark on top of the Variable y to indicate that it is Predicted value rather than an Observed Value.

→ The error of Prediction is ∴ $\boxed{E(\overline{X}) = (y - \hat{y})}$

→ In cases where the error value $E(\overline{X})$

when Predicted value = 0 ← $\boxed{\text{loss function} = (y-\hat{y})^2}$

→ In 'Perceptron' diagram, To reduce loss value we use optimizer ↳ more details at back

→ Sign function is applied to Convert aggregate value into a class label.

→ The Sign function serves the role of an activation function.

→ Basic m/c learning models can be represented easily as simple neural n/w architectures.

→ It is noteworthy that the Perceptron contains 2 layer

↓

I/p layer does not perform any computation & only transmits the feature value.

→ I/p layer is not included in count of the no.

⇒ Invariant part of Prediction → unaltered part of graph by adding **Bias components**.

For eg:- Setting in which feature variables are mean centred, but
mean of binary class Prediction from $\{-1, +1\}$ is not 0.

→ Additional bias variable b that capture this invariant part:-

$$\boxed{\hat{y} = \text{sign}\{\overline{W} \cdot \overline{X} + b\} = \text{sign}\left(\sum_{j=1}^{d} w_j x_j + b\right\} \quad \text{eq. 2}}$$

Biasing → weight of edge called as bias neuron.

→ Invariance definition: an entity, quantity, etc, i.e unaltered by a particular transformation.
use of additional i/p, which is called bias, of neuron improve properties of neuron. It allows moving the threshold of activation function.

→ bias value → activation function to be shifted to left or right, to better fit the data. It is always 1.

→ A weight represent the strength of connection b/w units. If the weight from node 1 to node 2 has greater magnitude, it means that neuron 1 > influence over neuron 2.

→ weights in NN can & will become +ve or -ve depending on training data.

→ training function is overall algorithm i.e used to train the N/W to recognize a certain I/P & map it to an O/P.

→ training process involves finding a set of weights in the n/w that proves to be goods or good enough at solving the specific problem.

→ learning function deals with individual weights & thresholds & decide how those would be manipulated.

→ Optimizer (using back Propagation)

① Gradient descent ② Stochastic Gradient descent

(change in weight value while doing Propagation).

using back Propagation :-

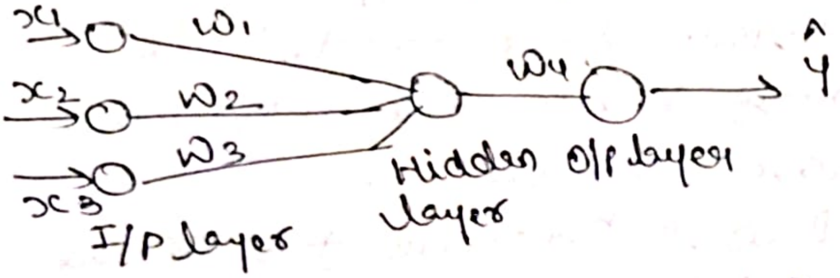
$$\boxed{W_{4new} = W_{4old} - \eta \cdot \frac{\partial L}{\partial W_{4old}}}$$

→ derivative of loss function

↓ learning rate (0.001).

| learning rate vary form (0.01 to 0.0001) |

In order to reach a point called as Global Minima therefore, ~~for~~ reduce the loss function.



$$w_{1new} = w_{1old} - \eta \cdot \frac{\partial L}{\partial w_{1old}}$$

$$w_{3new} = w_{3old} - \eta \frac{\partial L}{\partial w_{3old}}$$

w_2

w_1

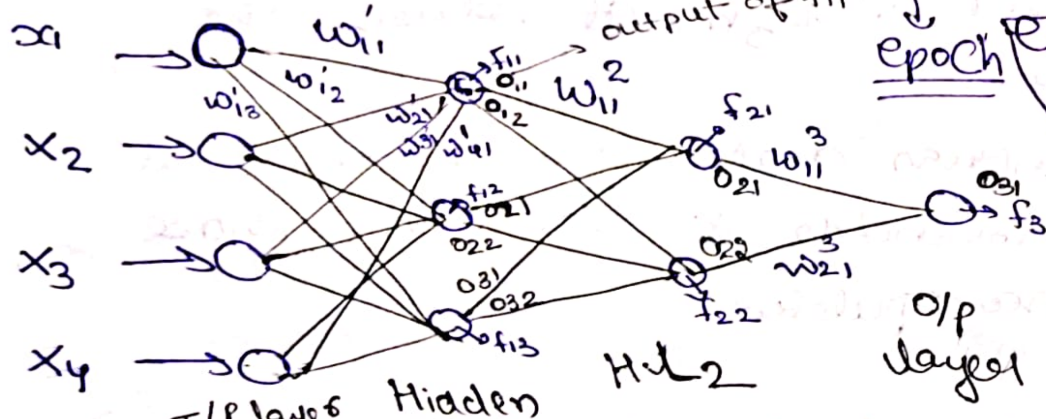
∴ Forward Propagation & Backward Propagation is used interchangebly to reduce the loss function to '0'. It is called as 'epoch'. (Forward + backward Propagation 1 iteration).

Multi layer Neural N/W :- In Deep learning we have M.L.Neural N/W.

iteration forward

epoch

Every diagram we have value of Bias = 1



Multilayer Neural N/W.

Backward iteration } epoch

w.r.t. weight & next layer

matrix formed	4×3	3×2	2×1

Loss function $= (y-\hat{y})^2$ reduces

Gradient Descent

Back

Single computational layer :-

→ we can still write the goal of the Perceptron algorithm in least - squares form with respect to all training instances in a data set D containing feature-label pair :

$$\boxed{\text{Minimize } wL = \sum_{(\bar{x},y)\in} (y-\hat{y})^2 = \sum_{(\bar{x},y)\in D} (y - \text{sign}(\bar{w}\cdot\bar{x}))^2}$$

cost function = $\sum_{i=1}^{n} (y-\hat{y})^2$, loss function
(w.r.t to many records) (w.r.t. to 1 record)

→ This type of Minimization objective function
↓
loss function.

→ Neural N/w learning algo. are formulated with the use of a loss function.

→ This loss function looks sim to least square regression

$\boxed{60{,}000 \text{ iterations}}$

→ Corr. loss is smooth & continous function of the Variable.

→ perceptron algorithm uses a smooth approximation of a gradient of this objective function w.r.t :-

$$\boxed{\nabla L_{smooth} = \sum_{(\bar{x},y)\in D} (y-\hat{y})\bar{x}}$$

→ The objective function is defined over the entire data, the training algo. of neural n/w works by feeding each i/p data instance x into the n/w one by one to create the prediction $\hat{y}$.

Weights are then updated, based on error value $E(x) = (y-\hat{y})$

Scanned by CamScanner

→ when data Point x is fed into the n/w, the weight vector w is updated as fn

$$\boxed{\bar{w} \leftarrow \bar{w} + \alpha(y-\hat{y})\bar{x}}$$

→ Parameter α regulates the learning rate of neural n/w.

→ Perceptron algo. repeatedly cycles through all the training eg in random order & iteratively adjusts the weights until Convergence is reached.

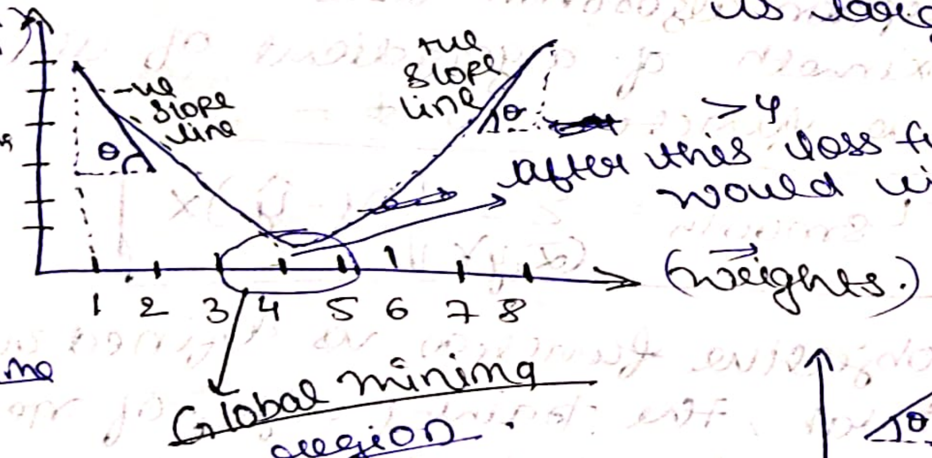
$$\boxed{\bar{w} \leftarrow \bar{w} + \alpha E(\bar{x})\bar{x}}$$

Gradient Decent : update the weights such a way that $\hat{y}$ to be ||r to y
find the max efficient weights to ~~get~~ be updated to best $\hat{y}$.

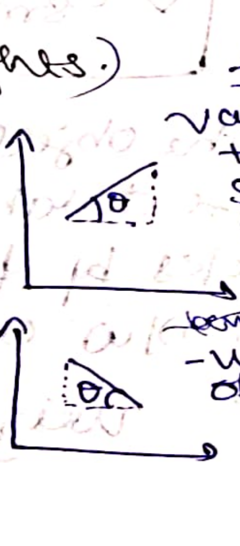
$$W_{new} = W_{old} - \eta \frac{\partial L}{\partial w} \qquad \eta \Rightarrow 0.0001$$

Loss function : $(y-\hat{y})^2$ ⟶ 1^{st} epoch diff is large.

Loss function approximate to '0' in global mining region except Neural N/W to work in Global minima Region



$\tan\theta$ = +ve values is +ve sloped line.



-ve slope of line.

→ when data Point x is fed into the n/w, the weight vector w is updated as fol

$$\boxed{\overline{w} \leftarrow \overline{w} + \alpha(y - \hat{y})\,\overline{x}}$$

→ Parameter α regulates the learning rate of neural n/w.

→ Perceptron algo. repeatedly cycles through all the training eg in random order & iteratively adjusts the weights until Convergence is reached.

$$\boxed{\overline{w} \leftarrow \overline{w} + \alpha E(\overline{x})\,\overline{x}}$$

Gradient Decent :- update the weights such a way that $\hat{y}$ to be ||r to y

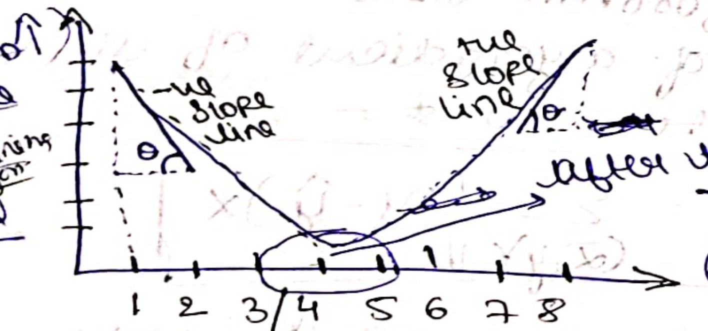
find the max efficient weights to ~~get~~ be updated to best $\hat{y}$.

$$\eta \Rightarrow 0.0001$$

$$W_{new} = W_{out} - \eta \frac{\partial L}{\partial w}$$

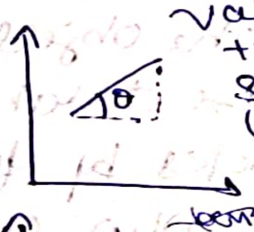
Loss function :- $(y-\hat{y})^2$ → 1^{st} epoch diff is large.

loss function approximate to '0' in global minima region except Neural N/W to work in global minima Region



>4 after this loss function would increase.

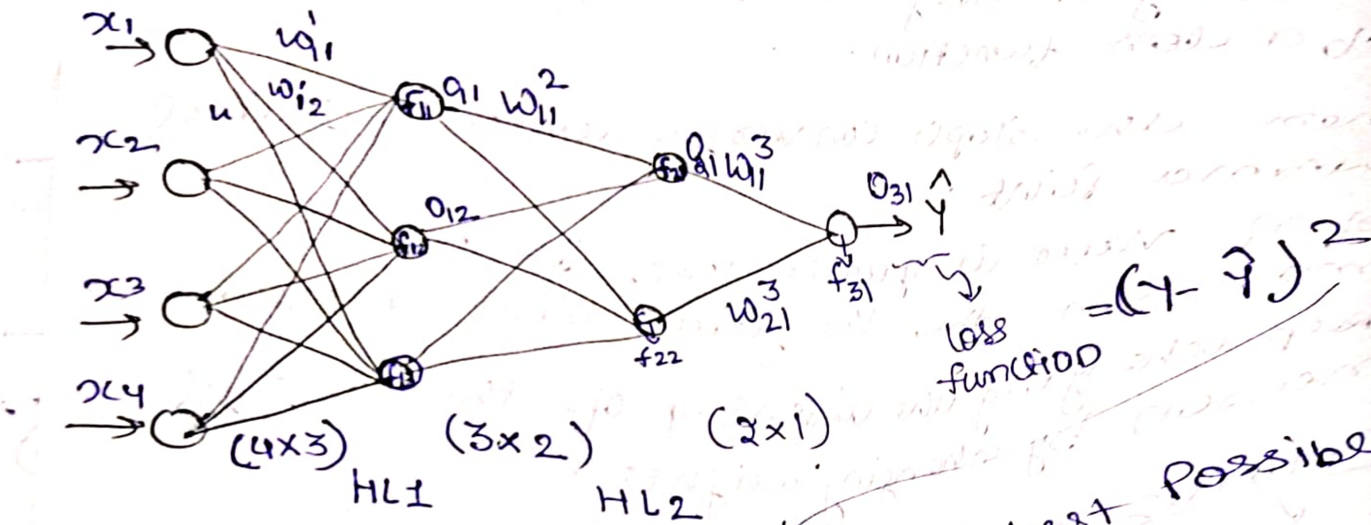
Global minima region.



$\tan\theta = +ve$ values is +ve slope of line.

-ve slope of line.

MNN : Multi-layer Neural N/W



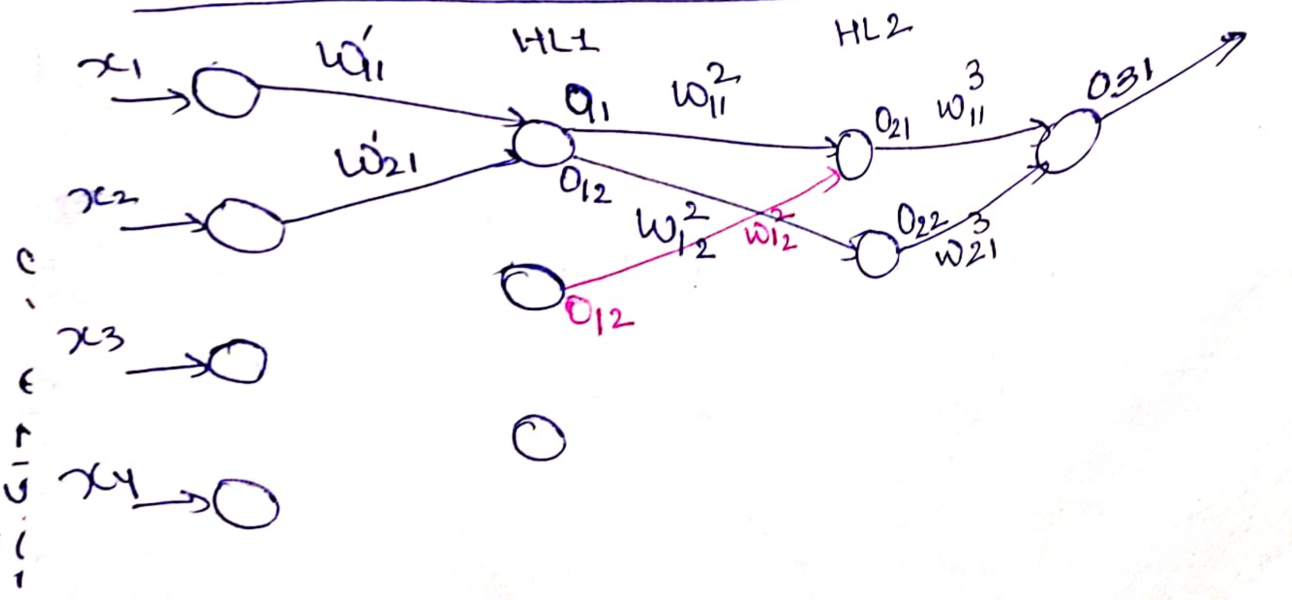
Ⓐ **Gradient Descent** : find best possible weights using backpropagation.

$$w_{new} = w_{old} - \eta \frac{\partial L}{\partial w}$$

$$\boxed{w^3_{11\,new} = w^3_{11\,old} - \eta \left|\frac{\partial L}{\partial w^3_{11}}\right|}$$

$O_{31} \rightarrow$ is impacted w.r.to weights.

Chain Rule on derivative :



$$\boxed{w^3_{11\,new} = w^3_{11\,old} - \eta \frac{\partial L}{\partial w^3_{11}}}$$

$$O_{31} = w^3_{11} \cdot O_{21} + w^3_{21} \cdot O_{22}$$

w^3_{11}.

$$\frac{\partial L}{\partial w^3_{11}} = \frac{\partial L}{\partial O_{31}} \times \frac{\partial O_{31}}{\partial w^3_{11}}$$ from chain rule of derivative.

$$w^3_{21\,new} = w_{21\,old} - \eta \frac{\partial L}{\partial w_{21}}$$

w^3_{21}

$$\frac{\partial L}{\partial w_{21}} = \frac{\partial L}{\partial O_{22}} \cdot \frac{\partial O_{22}}{\partial w_{21}}$$

w^2_{11}

$$w^2_{11\,new} = \longrightarrow \frac{\partial L}{\partial w^2_{11}}$$

$$\cancel{w^2_{11\,new} = \frac{\partial L}{} \cdot \frac{\partial L}{\partial w^2_{11}}}$$

layer no.

$$O_{21} = w^2_{11} * O_{11}$$

$$\frac{\partial L}{\partial w^2_{11}} = \frac{\partial L}{\partial O_{31}} * \frac{\partial O_{31}}{\partial O_{21}} * \frac{\partial O_{21}}{\partial w^2_{11}}$$

w^2_{12}

$$w^2_{12} = \longrightarrow \frac{\partial L}{\partial w^2_{12}}$$

$$O_{22} = w^2_{12}\, O_{12}$$

$$\frac{\partial L}{\partial w^2_{12}} = \frac{\partial L}{\partial O_{31}} * \frac{\partial O_{31}}{\partial O_{22}} * \frac{\partial O_{22}}{\partial w^2_{12}}$$

Vanishing gradient Problem

$$w^{1}_{11\,new} = w^{1}_{11\,old} - \eta \frac{\partial L}{\partial w^{1}_{11}}$$

Sigmoid → o/p of step 1.
(σ)

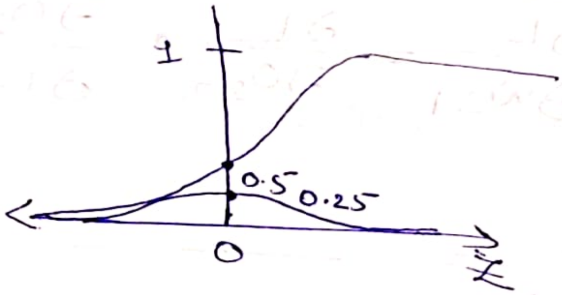
$$\boxed{\text{Sigmoid} = \frac{1}{1+e^{-z}}}$$

larger Value.

$$\boxed{e^{-5}}$$

$$z = \sum_{i=1} w_i x_i + b$$

$$\sigma = Act(z)$$



S. No.	Date	Title	Page No.	Teacher's Sign / Remarks

7 Jan

classmate

Date
Page

Quantum computing - ~~comp~~

- Fundamentals of quantum computing

→ what is quantum computing?

- when computation is done using the principles of quantum mechanics, it is called quantum computing.

→ Linear algebra : vector computation is done over complex no

- An Hilbert space H also called as Eucledian space is a normed vector space over the complex no. c.

- $|a\rangle$ → direct notation for vector space in quantum mechanics. It is also called as "ket" notation.

- Dirac's bracket notation : Z^* is used to denote complex conjugate. If $z = a+ib$ is a complex no then z^* or $\bar{z}$
$z^* = a - ib = \bar{z}$

- Ket vector is a column vector.

- The $|\psi\rangle$ corresponds to ~~vec~~ a column vector ψ.

- $|\psi\rangle$ → denotes "ket" while $\langle\psi|$ → denotes "bra".

- ψ^{+} is a conjugate transpose of a vector ψ.
dagger.

Eg :- $\psi = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ $\psi^{+} = (\bar{a}_1, \ldots \bar{a}_n)$

- In Direc's notation, the conjugate transpose of a "ket" $|\psi\rangle$ is called as a "bra" is written as $\langle\psi|$. Hence
$|\psi\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ and $\langle\psi| = (\bar{a}_1, \ldots, \bar{a}_n)$

- A bra $\langle\psi|$ corresponds to a row vector ψ^{+}.

• Inner product, tensor product, & various other computations can be done.

• <u>Inner Product b/w vectors $|\psi\rangle$ and $|\phi\rangle$</u>

$$|\psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix} \qquad |\phi\rangle = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_n \end{pmatrix}$$

then the inner product is $\langle\psi|\phi\rangle$ & defined to be the scalar obtained by multiplying the conjugate transpose. $\langle\psi| = (\bar{\psi}_1, \ldots \bar{\psi}_n)$ with $|\phi\rangle$.

$$\langle\psi|\phi\rangle = (\bar{\psi}_1 \cdots \bar{\psi}_n)\begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_n \end{pmatrix} = \sum_{i=1}^{n} \bar{\psi}_i \phi_i$$

$\langle x|y\rangle$ or $\langle x||y\rangle$ are same inner products of 2 vectors

Eg:- $|\psi\rangle = \begin{bmatrix} 2 \\ 6i \end{bmatrix}$, $|\phi\rangle = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

$$\langle\psi|\phi\rangle = [2 \quad -6i]\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \underline{\underline{6-24i}}$$

• <u>Tensor product b/w $|\psi\rangle$ and $|\phi\rangle$</u>

It is denoted by $|\psi\rangle \otimes |\phi\rangle = |\psi\rangle|\phi\rangle$

Eg:- $|\psi\rangle = \begin{bmatrix} 2 \\ 6i \end{bmatrix}$, $|\phi\rangle = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

$$\therefore |\psi\rangle \otimes |\phi\rangle = \begin{bmatrix} 2 \\ 6i \end{bmatrix} \otimes \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2\times 3 \\ 2\times 4 \\ 6i\times 3 \\ 6i\times 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \\ 18i \\ 24i \end{bmatrix}$$

$= \cancel{[illegible]}$

• <u>Norm (length) of the vector</u> :

• <u>$\psi^{+} \to$ Hermiltion Conjugate (adjoint) of a matrix ψ.</u>

if $\psi = \begin{bmatrix} 1 & 6i \\ 3i & 2+4i \end{bmatrix}$ then $\psi^{+} = \begin{bmatrix} 1 & -3i \\ -6i & 2-4i \end{bmatrix}$.

Note $\boxed{\psi^{+} = (\bar{\psi})^{T}}$

• Norm of the vector

$\| |\psi\rangle \| \rightarrow$ norm of vector $|\psi\rangle$

$$\boxed{\| |\psi\rangle \| = \sqrt{\langle\psi|\psi\rangle}}$$

• Normalization of vector $|\psi\rangle$ is given by $\dfrac{|\psi\rangle}{\| |\psi\rangle \|}$

Eg :- Find norm $|\psi\rangle = \begin{pmatrix} 1 \\ 2 \\ -i \end{pmatrix}$

$\| |\psi\rangle \| = \sqrt{\langle\psi|\psi\rangle}$

$\Rightarrow \langle\psi|\psi\rangle = (1 \ 2 \ i) \begin{pmatrix} 1 \\ 2 \\ -i \end{pmatrix} = 1+4+1 = 6$

$\therefore \| |\psi\rangle \| = \sqrt{6}.$

Normalization of $|\psi\rangle = \dfrac{|\psi\rangle}{\| |\psi\rangle \|} = \dfrac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ -i \end{pmatrix} = \begin{pmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ -i/\sqrt{6} \end{pmatrix} = |\phi\rangle$

$\| |\phi\rangle \| = \sqrt{\langle\phi|\phi\rangle} = \left(\frac{1}{\sqrt{6}} \ \frac{2}{\sqrt{6}} \ +\frac{i}{\sqrt{6}}\right) \begin{pmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ -i/\sqrt{6} \end{pmatrix} = \frac{1}{6} + \frac{4}{6} + \frac{1}{6} = \frac{6}{6} = 1$

Length = 1.

• Unit vector : A normalized vector or unit vector $|\psi\rangle$ a vector whose norm is 1.

• Orthogonal vector : 2 vectors $|\psi\rangle$ & $|\phi\rangle$ is said to be orthogonally if their inner product is 0.

$$\boxed{\langle\psi|\phi\rangle = 0}$$

Eg :- the vector $|\psi\rangle = \begin{pmatrix} i \\ i \end{pmatrix}$ and $|\phi\rangle = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$\Rightarrow \langle\psi|\phi\rangle = (-i \ -i) \begin{pmatrix} 1 \\ -1 \end{pmatrix} = -i + i = 0$

Problems :-

(1) If $|x\rangle = \begin{pmatrix} 1 \\ i \\ 2+i \end{pmatrix}$ and $|y\rangle = \begin{pmatrix} 2-i \\ 1 \\ 2+i \end{pmatrix}$. Find norm $\| |x\rangle \|$, $\langle x|y\rangle$, $\langle y|x\rangle$. Prove that $\langle x|y\rangle = \langle y|x\rangle^*$.

Ans

(a) $\| |x\rangle \| = \sqrt{\langle x|x\rangle}$

$$\langle x|x\rangle = \begin{pmatrix} 1 & -i & 2-i \end{pmatrix} \begin{pmatrix} 1 \\ i \\ 2+i \end{pmatrix}$$

$$= \cancel{(2+i) - i(1) + (2-i)^2} = 1 - i^2 + (2-i)(2+i)$$

$$= \cancel{2 + i - i + 4 - 4i + i^2} = 1 + 1 + 4 - i^2$$

$$= \cancel{6 - 4i - 1 = 5 - 4i} \quad = 1 + 1 + 4 + 1$$

$$= 7$$

$$\| |x\rangle \| = \sqrt{7}$$

(b) $$\langle x|y\rangle = \begin{pmatrix} 1 & -i & 2-i \end{pmatrix} \begin{pmatrix} 2-i \\ 1 \\ 2+i \end{pmatrix}$$

$$= 2 - i - i + 4 + 1 = 7 - 2i$$

(c) $$\langle y|x\rangle = \begin{pmatrix} 2+i & 1 & 2-i \end{pmatrix} \begin{pmatrix} 1 \\ i \\ 2+i \end{pmatrix} = 2 + i + i + 4 + 1$$

$$= 7 + 2i$$

(d) $\langle x|y\rangle = \langle y|x\rangle^{*}$ → congugate

$\langle x|y\rangle = 7 - 2i$ and $\langle y|x\rangle = 7 + 2i$

$\langle y|x\rangle^* = 7 - 2i$

$$\therefore \langle x|y\rangle = \langle y|x\rangle^*$$

Linearly Independent vectors :-

The set of vectors $|\psi_1\rangle, \ldots |\psi_n\rangle$ in $\mathbb{C}^n$ (complex nos., n → Dimension) are said to be linearly independent if there exists scalar complex nos. $a_1 \ldots a_n$ such that $a_1|\psi_1\rangle + a_2|\psi_2\rangle + \ldots + a_n|\psi_n\rangle = 0$

$$\Rightarrow a_1 = a_2 = a_3 \ldots = a_n = 0.$$

Eg:- The vector $|\psi_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ & $|\psi_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ linearly independent.

$a_1|\psi_1\rangle + a_2|\psi_2\rangle = 0$

$a_1\begin{pmatrix}1\\0\end{pmatrix} + a_2\begin{pmatrix}0\\1\end{pmatrix} = 0$

~~$a_1 + a_2 = 0$~~ $\begin{pmatrix}a_1\\a_2\end{pmatrix} = 0$

$\Rightarrow a_1 = 0, \; a_2 = 0.$

- When atleast one coefficient is not equal to zero, it becomes linearly dependent vectors.

ii) - Linearly dependent vectors : The set of vectors $|\psi_1\rangle, \ldots |\psi_n\rangle$ in $\mathbb{C}^n$ are said to be linearly dependent if there exists scalars $a_1 \ldots a_n$ not all of them are zero (0), such that $a_1|\psi_1\rangle + a_2|\psi_2\rangle + \ldots + a_n|\psi_n\rangle = 0$.

- Basis :- A vector obtained when one vector is linear combination of another vector.

$$|\psi\rangle = \begin{pmatrix}\psi_1\\\psi_2\end{pmatrix} = \psi_1\begin{pmatrix}1\\0\end{pmatrix} + \psi_2\begin{pmatrix}0\\1\end{pmatrix} \quad \text{in terms of } e^2.$$

in 2D complex space $\therefore \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}$ are called as std basis & is linearly independent.

Eg: The vectors $|\psi_1\rangle = \begin{pmatrix}1\\2\\-9\end{pmatrix}$ and $|\psi_2\rangle = \begin{pmatrix}-2\\-4\\18\end{pmatrix}$ are linearly dependent in $\mathbb{R}^3$.

$\Rightarrow a_1|\psi_1\rangle + a_2|\psi_2\rangle = \begin{pmatrix}0\\0\\0\end{pmatrix} \Rightarrow a_1\begin{pmatrix}1\\2\\-9\end{pmatrix} + a_2\begin{pmatrix}-2\\-4\\18\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$

Ex: The vectors $|\psi_1\rangle = \begin{pmatrix}1\\2\\3\end{pmatrix}, |\psi\;\rangle = \begin{pmatrix}2\\5\\7\end{pmatrix}$ and $|\psi_3\rangle = \begin{pmatrix}1\\3\\5\end{pmatrix}$ are linearly independent.

$$a_1\begin{pmatrix}1\\2\\3\end{pmatrix} + a_2\begin{pmatrix}2\\5\\7\end{pmatrix} + a_3\begin{pmatrix}1\\3\\5\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$$

i.e.
$$\begin{pmatrix} a_1+2a_2+a_3 \\ 2a_1+5a_2+3a_3 \\ 3a_1+7a_2+5a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \rightarrow \begin{matrix} a_1+2a_2+a_3=0 \\ 2a_1+5a_2+3a_3=0 \\ 3a_1+7a_2+5a_3=0. \end{matrix}$$

$$\Rightarrow a_1=0,\ a_2=0,\ a_3=0.$$

Basis : The set of vectors $|\psi_1\rangle, |\psi_2\rangle \dots |\psi_n\rangle$ is said to be basis ~~to be~~ for complex number, then every vector $|v\rangle$ can be uniquely expressed as a linear combination of $|\psi_1\rangle$, $|\psi_2\rangle \dots |\psi_n\rangle$ i.e.,

$$|v\rangle = a_1|\psi_1\rangle + a_2|\psi_2\rangle \dots + a_n|\psi_n\rangle.$$

where $a_1, a_2 \dots a_n$ are scalar.

Eg: The vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is the basis of $\mathbb{R}^2$

$$|v\rangle = \begin{pmatrix} a \\ b \end{pmatrix} \qquad \begin{pmatrix} a \\ b \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Note: A basis for complex no's consists of exactly n linearly independent vectors.

Orthonormal vectors

A basis is said to be orthonormal if each vector has norm 1 i.e., each pair of vectors are orthogonal.

Eg : The vectors $|\psi_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|\psi_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ basis in $\mathbb{R}^2$.

$$\| |\psi_1\rangle \| = \sqrt{\langle \psi_1 | \psi_1 \rangle} = \sqrt{1} = 1$$

$$\langle \psi_1 | \psi_1 \rangle = (1\ 0)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1$$

$$\langle \psi_2 | \psi_2 \rangle = (0\ 1)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1$$

$$\| |\psi_2\rangle \| = \sqrt{\langle \psi_2 | \psi_2 \rangle} = \sqrt{1} = 1$$

$$\langle \psi_2 | \psi_2 \rangle = (1\ 0)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0.$$

1) Eg: The vectors $|\psi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$ and $|\psi_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$ are orthonormal basis in $\mathbb{C}^2$

2) Eg: The vectors $|v_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$, $|v_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}$ are orthonormal basis in $\mathbb{C}^2$

$$\langle v_1|v_2\rangle = \begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}1/\sqrt{2}\\ -i/\sqrt{2}\end{pmatrix} = \tfrac{1}{2} - \tfrac{1}{2} = 0$$

$$\|\,|v_1\rangle\| = \sqrt{\langle v_1|v_1\rangle} \Rightarrow \langle v_1|v_1\rangle = \begin{pmatrix}1/\sqrt{2} & -i/\sqrt{2}\end{pmatrix}\begin{pmatrix}1/\sqrt{2}\\ i/\sqrt{2}\end{pmatrix} = \tfrac{1}{2}+\tfrac{1}{2} = 1$$

$$\langle v_2|v_2\rangle = \begin{pmatrix}1/\sqrt{2} & i/\sqrt{2}\end{pmatrix}\begin{pmatrix}1/\sqrt{2}\\ -i/\sqrt{2}\end{pmatrix} = \tfrac{1}{2} + \tfrac{1}{2} = 1.$$

Postulates of Quantum Mechanics.

Postulate 1 : Defination of quantum bits or qubits.

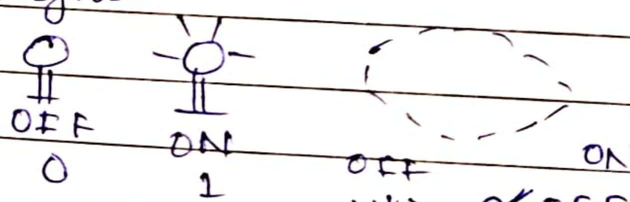
Postulate 2 : How qubit transform (evolve)

Postulate 3 : The effect of measurement.

Postulate : How qubits combine together into systems of qubits.

C bit	qubit
0 or 1	$|0\rangle$ or $|1\rangle$.

light



$$|\psi\rangle = \alpha\,\text{OFF} + \beta\,\text{ON}$$
$$= \alpha|0\rangle + \beta|1\rangle$$

1. Qubit : In quantum mechanics 2 possible states for qubit are $|0\rangle$ and $|1\rangle$.
These qubits can be expressed as orthonormal basis vectors in $\mathbb{C}^2$.

$$|0\rangle = \begin{pmatrix}1\\0\end{pmatrix} \text{ and } |1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$$

A generic qubit state $|\psi\rangle$ is represented by linear combination (or superposition) of $|0\rangle$ and $|1\rangle$

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

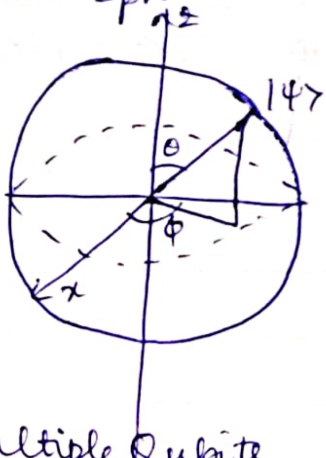
where α and β are complex nos. with $|\alpha^2| + |\beta^2| = 1$. When we measure qubit, we get 0 with prob $|\alpha^2|$ or we get 1 with prob $|\beta|^2$

Eg: $|\phi\rangle = \frac{1+i}{2}\,, |0\rangle + \frac{1}{\sqrt{2}}|1\rangle$

After measurement we get 0 with probability.

Block Sphere

The state $|\psi\rangle$ of a qubit is represented by a point by a point on the surface of a sphere unit radius called block Sphere



$$|\psi\rangle = \cos\theta/2 \quad |0\rangle + i^{i\phi}\sin\frac{\theta}{2}|1\rangle$$

$$\alpha = \cos\theta/2$$

$$\beta = i^{i\phi}\sin\frac{\theta}{2} \qquad |\alpha|^2 + |\beta|^2 = 1$$

Multiple Qubits

2 qubit system has 4 possible states,
$|00\rangle, |01\rangle, |10\rangle, |11\rangle$

A generic qubit state describing the 2-qubits is given classmate

Date
Page

tensor product

by, $|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$

$|00\rangle = |0\rangle \otimes |0\rangle = \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix} = (1\,0\,0\,0)^T$

$|01\rangle = |0\rangle \otimes |1\rangle = \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix} = (0\,1\,0\,0)^T$

$|10\rangle = |1\rangle \otimes |0\rangle = \begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} = (0\,0\,1\,0)^T$

$|11\rangle = |1\rangle \otimes |1\rangle = \begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix} = (0\,0\,0\,1)^T$

$$(\alpha_1|0\rangle + \beta_1|1\rangle) \otimes (\alpha_2|0\rangle + \beta_2|1\rangle)$$
$$= \begin{pmatrix}\alpha_1\\\beta_1\end{pmatrix} \otimes \begin{pmatrix}\alpha_2\\\beta_2\end{pmatrix}$$
$$= \begin{pmatrix}\alpha_1 & \alpha_2\\ \alpha_1 & \beta_2\\ \beta_1 & \alpha_2\\ \beta_1 & \beta_2\end{pmatrix}$$
$$= \alpha_1\alpha_2\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + \alpha_1\beta_2\begin{pmatrix}0\\1\\0\\0\end{pmatrix} + \beta_1\alpha_2\begin{pmatrix}0\\0\\1\\0\end{pmatrix} + \beta_1\beta_2\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$
$$= \alpha_1\alpha_2|00\rangle + \alpha_1\beta_2|01\rangle + \beta_1\alpha_2|10\rangle + \beta_1\beta_2|11\rangle.$$

Entangled states

States that cannot be written as the tensor product of n single qubits states are called entangled states or Bell state or EPR state (Einstein, Podolsky Rosen).

Eg: Consider the state
$r_1|00\rangle + r_2|11\rangle$ with $r_1 \neq 0$ and $r_2 \neq 0$.

Scanned by CamScanner

This cannot be written as

$(\alpha_1|0\rangle + \beta_1|1\rangle) \otimes (\alpha_2|0\rangle + \beta_2|1\rangle)$

The entangled states known as Bell states or EPR pairs.

$= (\alpha_1|0\rangle + \beta_1|1\rangle) \otimes (\alpha_2|0\rangle + \beta_2|1\rangle) = \alpha_1\alpha_2|00\rangle + \alpha_1\beta_2|01\rangle + \beta_1\alpha_2|10\rangle + \beta_1\beta_2|11\rangle.$

Eg: The following 2 qubit state.

$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$; $|\phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$

$|\psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$; $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ are entangled.

1) Prove that $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is in entangled state.

$\Rightarrow$ It is impossible to find $\alpha_1 \beta_1, \alpha_2 \beta_2$ such that

$(\alpha_1|0\rangle + \beta|1\rangle) \otimes (\alpha_2|0\rangle + \beta_2|1\rangle) = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$

Since $(\alpha_1|0\rangle + \beta_1|1\rangle) \otimes (\alpha_2|0\rangle + \beta_2|1\rangle)$

$= \alpha_1\alpha_2|00\rangle + \alpha_1\beta_2|01\rangle + \beta_1\alpha_2|10\rangle + \beta_1\beta_2|11\rangle$

If $\alpha_1\beta_2 = 0 \Rightarrow \alpha_1\alpha_2 = 0$ or $\beta_1\beta_2 = 0$ this is impossible

$|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}$, $|1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$ $|00\rangle = |0\rangle \otimes |0\rangle$

$= \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}$

2) Prove that $\frac{1}{\sqrt{2}}(|00\rangle + |01\rangle)$ is not in entangled state

$\Rightarrow |00\rangle = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}$ and $|01\rangle = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}$

$\Rightarrow \frac{1}{\sqrt{2}}(|00\rangle + |01\rangle) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\0\\0\end{pmatrix} = \begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\\0\end{pmatrix}$

$= \begin{pmatrix}1\\0\end{pmatrix} \otimes \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$

$= \begin{pmatrix}1\\0\end{pmatrix} \otimes \frac{1}{\sqrt{2}}\left(\begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\right) = |0\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$

3) prove that $\frac{1}{\sqrt{2}}(|00\rangle - |01\rangle)$ is not in entangled state.

4) Show that the following 2-qubit states are entangled

a) $1/\sqrt{2}(|00\rangle - i|11\rangle)$

b) $i/10(|00\rangle - \sqrt{99}/100|11\rangle)$

→ X gate

$|0\rangle \to |1\rangle$

$|1\rangle \to |0\rangle$

Z gate

$|0\rangle \to |0\rangle$

$|1\rangle \to -|1\rangle$

H-gate

$|0\rangle \to 1/\sqrt{2}(|0\rangle + |1\rangle)$

$|1\rangle \to 1/\sqrt{2}(|0\rangle - |1\rangle)$

$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ $H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

$X^T X = I$

$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $\Rightarrow Z^T Z = I$ and $H^T H = I$.

→ Y-gate

$|0\rangle \to i|1\rangle$

$|1\rangle \to -i|0\rangle$

$Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$

$Y|0\rangle = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ i \end{pmatrix} = i\begin{pmatrix} 0 \\ 1 \end{pmatrix} = i|1\rangle$

$Y|1\rangle = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -i \\ 0 \end{pmatrix} = -i\begin{pmatrix} 1 \\ 0 \end{pmatrix} = -i(|0\rangle)$

$e^{i\theta} = \cos\theta + i\sin\theta$.

$(|0\rangle + \beta|1\rangle) \to \boxed{Y} \to i(\alpha|1\rangle - \beta|0\rangle)$

$Y^T Y = I$

Phase - shift gate

$|0\rangle \to |0\rangle$

$|1\rangle \to e^{i\theta}|1\rangle$

$R_\theta = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}$, θ is any value.

$R_\theta|0\rangle = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle$

$R_\theta|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ e^{i\theta} \end{pmatrix} = e^{i\theta}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = e^{i\theta}|1\rangle$

$e^{i\theta} = \cos\theta + i\sin\theta$

$\theta = \pi$

$e^{i\pi} = -1$

When $\theta = \pi$ we get z-gate

Problem: 1) Write the following quantum gate in braket notation.

1. X 2. Z 3. H 4. Y 5. R_θ

① $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ = $|0\rangle$ $\langle 0|$

$\langle 1|$ $|1\rangle$

$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\langle 0| = (1 \ \ 0)$

$|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ $\langle 1| = (0 \ \ 1)$

$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}(1 \ \ 0) + \begin{pmatrix} 1 \\ 0 \end{pmatrix}(0 \ \ 1)$

$= |1\rangle\langle 0| + |0\rangle\langle 1|$

② $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

$= \begin{pmatrix} 1 \\ 0 \end{pmatrix}(1 \ \ 0) - \begin{pmatrix} 0 \\ 1 \end{pmatrix}(0 \ \ 1)$

$= |0\rangle\langle 0| - |1\rangle\langle 1|$

$|00\rangle$

$|00\rangle$

$\langle 01|$

$|01\rangle \quad |10\rangle \quad |11\rangle$

$= |0\rangle \otimes |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

$\rangle = |0\rangle \otimes |1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$

2-qubit controlled gates

Controlled-NOT gate:-

It flips the second bit if the first bit is 1, does nothing otherwise.

$$C_{NOT} = |00\rangle \rightarrow |00\rangle$$

control bit, target bit

$$|01\rangle \rightarrow |01\rangle$$
$$|10\rangle \rightarrow |11\rangle$$
$$|11\rangle \rightarrow |10\rangle$$

In matrix form, this is

$$C_{NOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

C_{NOT} is unitary or

$$C_{NOT}^{+} C_{NOT} = I$$

C_{NOT} is a generalization of the classical XOR gate:

$$|AB\rangle \rightarrow |A B \oplus A\rangle$$

The C_{NOT} gate cannot be decomposed into a .

Tensor Product of 2 single qubit states

C_{NOT} takes the unentangled ~~gates~~ qubit state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle$ to the entangled state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$

$$C_{NOT}\left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right) \otimes |0\rangle = C_{NOT}\left(\frac{1}{\sqrt{2}}(|00\rangle + |10\rangle)\right)$$

$$= C_{NOT}\left(\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\\0\end{pmatrix}\right)$$

$$= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix}1\\0\\1\\0\end{pmatrix} = \begin{pmatrix}1\\0\\0\\1\end{pmatrix}$$

$$= \frac{1}{\sqrt{2}}\left(\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\right) = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

The No-cloning principle :-

The no-cloning principle states we cannot copy or clone an unknown qubit

Let U is a unitary transformation that clones i.e. →

$U|a0\rangle) = |aa\rangle$ for all quantum states

Let $|a\rangle$ and $|b\rangle$ be 2 orthonormal quantum states,

consider, $|c\rangle = \frac{1}{\sqrt{2}}(|a\rangle + |b\rangle)$

$$U(|c0\rangle = U\left(\frac{1}{\sqrt{2}}|a0\rangle + |b0\rangle|\right)$$
$$= 1/\sqrt{2}(|aa\rangle + |bb\rangle)$$

But $U(|c0\rangle) = |cc\rangle$
$$= |c\rangle \otimes |c\rangle$$
$$= 1/2(|aa\rangle + |ab\rangle + |ba\rangle + |bb\rangle)$$
$$= \pm 1/2(|aa\rangle + |bb\rangle)$$

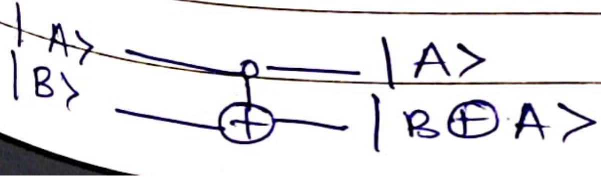
But $U(|c0\rangle) = |cc\rangle$

$$|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}$$

$$|1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$$

$$|+\rangle = 1/\sqrt{2}(|0\rangle + |1\rangle)$$
$$|-\rangle = 1/\sqrt{2}(|0\rangle - |1\rangle)$$
$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = |+\rangle$$
$$H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle = |-\rangle$$

CNOT



22/01/20

SWAP gate.

Simply exchange bit values.

$|00\rangle \rightarrow |00\rangle$

$|01\rangle \rightarrow |10\rangle$

$|10\rangle \rightarrow |01\rangle$

$|11\rangle \rightarrow |11\rangle$

In matrix form, $SWAP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

$|00\rangle = |0\rangle \otimes |0\rangle$ — tensor product

$= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$= \begin{bmatrix} 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 0\begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{bmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

w.k.t $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Tensor product for matrices

Consider 2 vectors, $|v\rangle = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ and $|w\rangle = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$

$$|v\rangle \otimes |w\rangle = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \otimes \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

$$= \begin{pmatrix} v_1 \cdot \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \\ v_2 \cdot \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} v_1 w_1 \\ v_1 w_2 \\ v_2 w_1 \\ v_2 w_2 \end{pmatrix}$$

Suppose we have 2 matrices.

$$M = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix}_{2\times 2} \qquad N = \begin{pmatrix} n_1 & n_2 \\ n_3 & n_4 \end{pmatrix}_{2\times 2}$$

The tensor product of M & N is defined as,

$$M \otimes N = \begin{pmatrix} m_1 \cdot \begin{pmatrix} n_1 & n_2 \\ n_3 & n_4 \end{pmatrix} & m_2 \cdot \begin{pmatrix} n_1 & n_2 \\ n_3 & n_4 \end{pmatrix} \\ m_3 \cdot \begin{pmatrix} n_1 & n_2 \\ n_3 & n_4 \end{pmatrix} & m_4 \cdot \begin{pmatrix} n_1 & n_2 \\ n_3 & n_4 \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} m_1 n_1 & m_1 n_2 & m_2 n_1 & m_2 n_2 \\ m_1 n_3 & m_1 n_4 & m_2 n_3 & m_2 n_4 \\ m_3 n_1 & m_3 n_2 & m_4 n_1 & m_4 n_2 \\ m_2 n_3 & m_3 n_4 & m_4 n_3 & m_4 n_4 \end{pmatrix}_{4\times 4}$$

1. Find X-gate & Z-gate

X-gate → not gate.

Z-gate → negate the amplitude of $|\psi_1\rangle$

$$X \otimes Z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

Quantum Measurement:

Quantum state: The collection of all ~~elements~~ relevant physical properties of quantum system. For eg: position, momentum, spin, polarization) is known as the state of the system.

Physical Support	Name	Information support	$\|0\rangle$	$\|1\rangle$
Photon	Polarization	Polarization	Vertical	Horizontal
Electron	Electronic spin	Spin	up ↑	down ↓

For eg; if we use the energy of an electron as our qubits $|0\rangle$ and $|1\rangle$, we could say that the ground state (lowest energy) is our qubit $|0\rangle$ and an excited state (highest energy) is our qubit $|1\rangle$.

ground state $= |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

excited state $= |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

We can ~~differ~~ define state $|+\rangle$ and $|-\rangle$ with the vectors

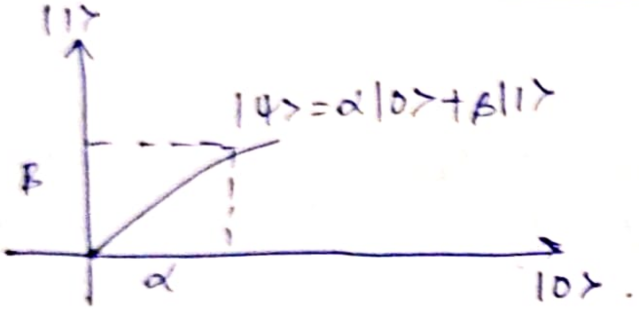
$|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\qquad$ $|-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$$|+\rangle = \frac{1}{\sqrt{2}}\left(\begin{pmatrix}1\\0\end{pmatrix}+\begin{pmatrix}0\\1\end{pmatrix}\right) = \frac{1}{\sqrt{2}}\left(|0\rangle+|1\rangle\right)$$

$$= \frac{1}{\sqrt{2}}\left(\begin{pmatrix}1\\0\end{pmatrix}-\begin{pmatrix}0\\1\end{pmatrix}\right) = \frac{1}{\sqrt{2}}\left(|0\rangle+|1\rangle\right)$$

05/01/20

Physical support	Information support	$\|0\rangle$	$\|1\rangle$
Photons	Polarization of photons	horizontal →	vertical ↑
Electrons	Spin of electron	up ↑	down ↓



$|0\rangle = OFF$

$|1\rangle = ON$

$|\psi\rangle = \alpha\, OFF + \beta\, ON.$

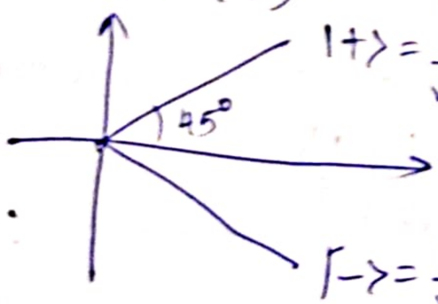
Quantum mechanics describes the behaviour of systems such as electrons, photons or molecules. We use mathematics to model this physical phenomenon.

Consider general quantum state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ where $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$. When we measure quantum state $|\psi\rangle$ we get either the result $|0\rangle$ with probability $|\alpha|^2$ or we get result $|1\rangle$ with probability $|\beta|^2$

Defn: Born's rule

$\{|0\rangle, |1\rangle\}$ $\quad \{|+\rangle, |-\rangle\}$

$\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$ $\quad \left\{\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}\right\}$



$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$

$|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$

$|\nearrow\rangle$ photons

$|\searrow\rangle$ electrons.

classmate

Date

Page

Born's rule:- suppose we have a quantum state $|\psi\rangle$ and orthonormal basis $\{|\phi_1\rangle, |\phi_2\rangle, \ldots |\phi_n\rangle\}$. Then we measure $|\psi\rangle$ w.r.t this orthonormal basis i.e, we ask the quantum system which one of these states it is in.

The probability of measuring state $|\phi_i\rangle$ is given by.

$$\boxed{P(\phi_i) = |\langle\phi_i|\psi\rangle|^2}$$

After the measurement the original state collapses in the measured state. i.e, we are left with one of the states $|\phi_1\rangle \ldots |\phi_n\rangle$

① Eg:- Given a quantum state

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$$

⇒ we measure $|\psi\rangle$ in standard computational basis $\{|0\rangle, |1\rangle\}$ then outcome state will be

$$\begin{cases} |0\rangle \text{ with probability } |\langle 0|\psi\rangle|^2 = |\alpha|^2 \\ |1\rangle \text{ —''— } |\langle 1|\psi\rangle|^2 = |\beta|^2 \end{cases}$$

② Eg:- What will be outcome if we measure $\{|+\rangle, |-\rangle\}$

$$\Rightarrow \begin{cases} |+\rangle, \text{ with probability } |\langle +|\psi\rangle|^2 = \frac{|\alpha+\beta|^2}{2} \\ |-\rangle, \text{ with probability } |\langle -|\psi\rangle|^2 = \frac{|\alpha-\beta|^2}{2} \end{cases}$$

③ Given a quantum state $|\psi\rangle = \frac{1+i}{2}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$

we measure $|\psi\rangle$ in std. computational basis $\{|0\rangle, |1\rangle\}$ then outcome state will be