* Day-12 * 5th may, 2022

* Calculus *

1) Functions, Scalar Derivative, Definition, Intuition, Common Rules of Differentiation, chain Rule

2) Partial Derivatives, Gradient, Concept, Intuition, Properties, Directional Derivative.

3) Gradient of Vector Valued functions, Gradient of Matrices.

4) Linearization and Multivariate Taylor series

5) Optimization Using Gradient Descent, Constrained optimization and Lagrange multipliers.

6) Vector and Matrix Calculus, Jacobian.

7) Gradient Algorithm, Local/Global maxima of Minima, Saddle Point, Convex functions.

* Funⁿ :- It is generally return as F(x).

It means there is some mathematical formulation by which you are connecting two variables.

Generally, $y = f(x)$

Funⁿ can be any form. such linear or non-linear.

for. eg. $y = 2x+3$

y is in form of Function (2x+3), which is linear function.

* Constant Function :-

If $F(x) = k$, where k is constant is said to be constant Funⁿ.

* Power function :-

If $f(x) = x^n$ is said to be power funⁿ.

* Trignometric function :-

If $f(x) = \sin x, \cos x, \tan x, \cot x, \operatorname{cosec} x, \sec x$ are said to be Trignometric function.

* Exponential funⁿ

If $F(x) = e^x$ or $f(x) = a^x$ are said to be exponential function.

* Logarithmic function :-

If $F(x) = \log x$ is said to be logarithmic funⁿ.

* Derivatives :-

If $y = F(x)$ is 1st derivative can be given as

$y' = f'(x)$ or $dy = \frac{df(x)}{dx}$

* Derivative means Rate of change of Variable *

Ex.

$y = 2x + 3$

$\underline{dy} = \underline{2dx}$

* Derivatives of Some Standard Functions :-

A) Constant fun^n :-

IF $F(x) = K$ then $f'(x) = 0$.

B) Power fun^n :-

1) IF $F(x) = x^n$ then $f'(x) = nx^{n-1}$

2) IF $F(x) = \sqrt{x}$ then $F'(x) = \frac{1}{2\sqrt{x}}$

3) IF $F(x) = \frac{1}{x}$ then $F'(x) = \frac{-1}{x^2}$

4) IF $F(x) = x$ then $F'(x) = 1$.

C) Trignometric fun^n :-

1) $\frac{d}{dx} \sin x = \cos x$

2) $\frac{d}{dx} \cos x = -\sin x$

3) $\frac{d}{dx} \tan x = \sec^2 x$

4) $\frac{d}{dx} \cot x = -\text{cosec}^2 x$

5) $\frac{d}{dx} \sec x = \sec x \cdot \tan x$.

6) $\frac{d}{dx} \text{cosec} x = -\text{cosec} x \cdot \cot x$.

D) Exponential fun^n :-

1) $\frac{d}{dx} e^x = e^x$

2) $\frac{d}{dx} a^x = a^x \log a$

E) Logarithmic fun^n :-

1) $\frac{d}{dx} \log x = \frac{1}{x}$

2) $\frac{d}{dx} \log_a x = \frac{1}{x} \log_e a.$

F) Derivatives & Inverse Trignometric fun^{ns} :-

1) $\frac{d}{dx} \sin^{-1} x = \frac{1}{\sqrt{1-x^2}}$

2) $\frac{d}{dx} \cos^{-1} x = \frac{-1}{\sqrt{1-x^2}}$

3) $\frac{d}{dx} \tan^{-1} x = \frac{1}{1+x^2}$

4) $\frac{d}{dx} \cot^{-1} x = \frac{-1}{1+x^2}$

5) $\frac{d}{dx} \sec^{-1} x = \frac{1}{x\sqrt{x^2-1}}$

6) $\frac{d}{dx} \text{cosec}^{-1} x = \frac{-1}{x\sqrt{x^2-1}}$

Ex. $y = 2x+3$

If $x_1 = 1$ ~~they~~ then $y_1 = 5$

change x by small amount dx.

$\therefore x_2 = x + dx = 1 + 0.1 = 1.1$

then $y_2 = 2(1.1) + 3 = 5.2$

$\therefore \Delta y = y_2 - y_1 = 5.2 - 5 = 0.2$.

In limiting case Δy can be written as dy

dy – giving us the change in y. not rate.

To find rate of change of y, we to take derivative with respect to some variable. eq. n

$$\frac{dy}{dx} = \frac{d}{dx}(2x+3)$$

$$= \underline{\underline{2}}$$

Note :-

$$\boxed{\frac{dy}{dx} = f'(x)}$$

$$\boxed{dy = f'(x)dx}$$

2) $dy = ?$ $\quad y = x^2+5$

$$\underline{dy = 2x\,dx}$$

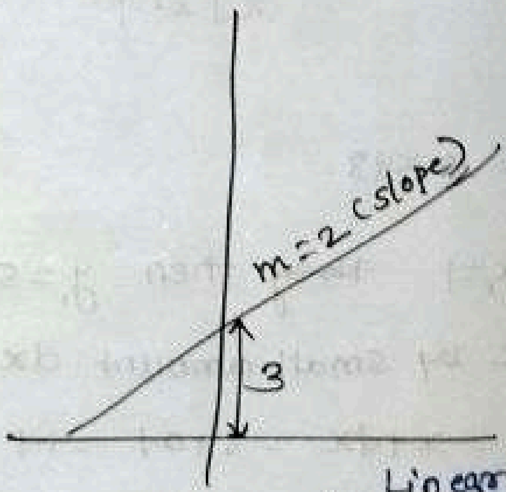
3) $y = \sin x + x^2$

$$= \cos x\,dx + 2x\,dx$$

* ~~Linear F~~ Linear fun^n:

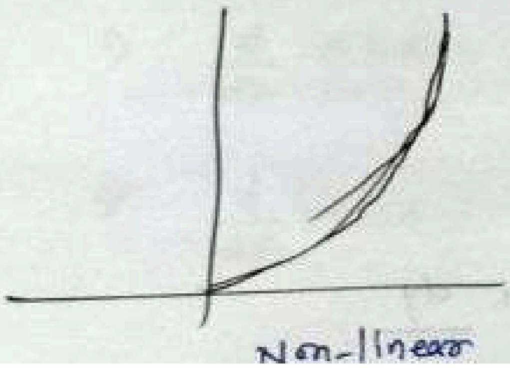
can be given as $\underline{y = mx + c}$

for eg. $\quad y = 2x+3.$



Linear.

* Non-linear fun^n.



Non-linear

e.g. $\underline{\underline{y = 2x^2 + 3}}$

y = 2x^2 + c

* The change in dy is proportional to change in dx.

i.e. $dy \propto dx$

$$\therefore \boxed{dy = F'(x)\,dx}$$

where $f'(x)$ = Differential. this gives idea, How the value of y is changing with respect to x.

* if $z = F(x, y)$.

$$\boxed{dz = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy}$$

eg. $z = 3x^2 y$

$\cancel{dz} = \cancel{6xy\,dx}$

$$dz = 6xy\,dx + 3x^2\,dy$$

$dz \rightarrow$ Total differential

$\left.\begin{matrix}\partial z\\ \partial x\\ \partial y\end{matrix}\right\} \rightarrow$ Partial differential.

* Some Common Rules :-

$$d(u.v) = u.dv + v.du$$
$$= u.v' + v.u'$$

eg. ① $y = x^2 \sin x$.

$$dy = 2x.\sin x + x^2 \cos x$$

② $y = x^2 \sin x \, e^x$

$dy = 2x \sin x \, e^x + x^2 \cos x \, e^x + x^2 \sin x \, e^x$

⑥ $d\left(\frac{u}{v}\right) = \frac{vu' - uv'}{v^2}$

Ex ① $y = \frac{x^2}{\sin x} = \frac{\sin x \, 2x - x^2 \cos x}{\sin^2 x}$

$= \frac{\sin x \, 2x}{\sin^2 x} - \frac{x^2 \cos x}{\sin^2 x}$

$\left(\frac{\cos}{\sin} = \cot\right)$ $= \frac{2x}{\sin x} - \frac{x^2 \cot x}{\sin x}$

$\left(\frac{1}{\sin} = \operatorname{cosec}\right)$ $= 2x \operatorname{cosec} x - x^2 \cdot \cot x \cdot \operatorname{cosec} x.$

* Function of function / chain rule

$\frac{dy}{dx} = \frac{dy}{dv} \times \frac{dv}{du} \times \frac{du}{dx}$ — chain Rule.

If $y = \sin^2(3x+5)$.

if $u = 3x+5$

$\therefore y = \sin^2 u$

if $v = \sin u$

$\therefore y = v^2$

By chain rule

$= \frac{dy}{dv} * \frac{dv}{du} * \frac{du}{dx}$

$= \frac{d}{dv}(y) * \frac{d}{du}(v) \times \frac{d}{dx}(u)$

Put values of y, v & u

$= \frac{d}{dv}(v^2) \times \frac{d}{du}(\sin u) \times \frac{d}{dx}(3x+5)$

$= 2v \times \cos u \times 3$

Put values of u & v

$= 2\sin u \times \cos(3x+5) \times 3$

$\cancel{dy = 6 \cdot \sin u \times \cos(3x+5)}$

$\boxed{= 3.2 \sin(3x+5) \cdot \cos(3x+5)}$

$\cancel{dy = 3 \sin(6x+10)}$ — ~~2 sin A cos~~

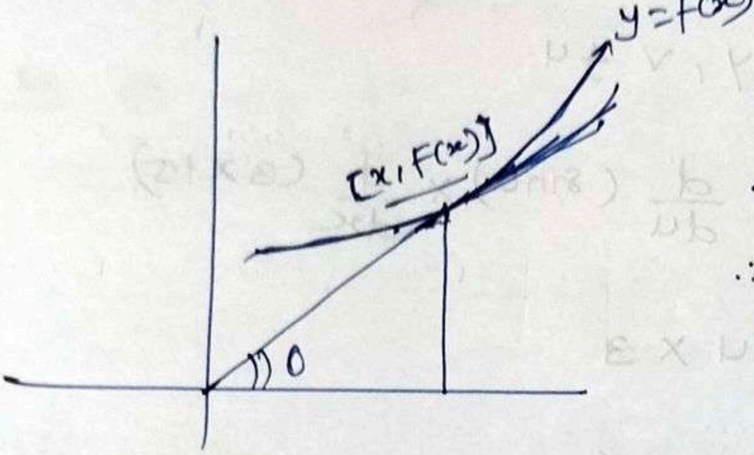
Explain.

$\cancel{2 \sin(A+B) \cos(A+B) = \sin(2A+2B)}$

$$y = f(x) \rightarrow dy = f'(x)\,dx$$

$f'(x)$ Also referred as a dy/dx

$$y' = \frac{df}{dx}$$

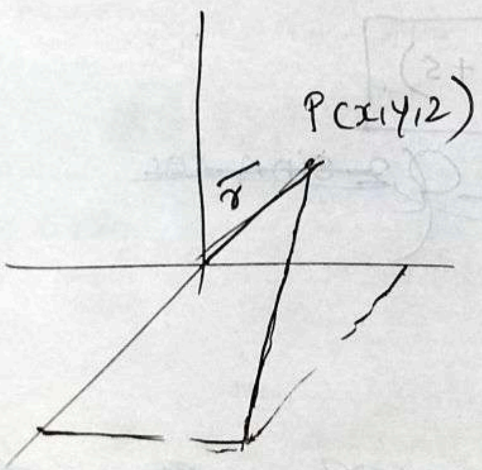


if $y = 3x^2$

Find slope at $x = 2$.

$$\therefore \left.\frac{dy}{dx}\right|_{x=2} = 6x = \underline{\underline{12}}$$

* If there is object in space whose posᵗⁿ is given as (x, y, z). how vector can return as.



Vector $\bar{r}$ can be given as

$$\bar{r} = xi + yj + zk$$

$$d\bar{r} = dx\,i + dy\,j + dz\,k \quad -(1)$$

If $U = f(x, y, z)$

$du = ?$

$$du = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz$$

$$= \left[\left(\frac{\partial f}{\partial x}\right)i + \left(\frac{\partial f}{\partial y}\right)j + \left(\frac{\partial f}{\partial z}\right)k\right] \cdot \left[dx\,i + dy\,j + dz\,k\right]$$

$$= \left[i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z} \right] \cdot f \cdot d\bar{r} \quad (d\bar{r} = \text{from} ①)$$

$$\boxed{du = \nabla f \cdot d\bar{r}}$$

where $\nabla = \left[i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z} \right]$

→ It is called Del operator.

* If you wand to find overall change of any function i.e du → there is other way which is taking the Gradient of that funⁿ and then taking the inner product with the differential element or position vector i.e $d\bar{r}$

i.e. $\boxed{du = \nabla F \cdot d\bar{r}}$

where du = Overall change in funⁿ

∇F = Gradient of function.

$d\bar{r}$ = position vector.

* When you multiplied del operator with any scalar function, then it become Gradient of that scalar funⁿ.

i.e ∇F. $\quad \nabla \rightarrow$ Del operator.

$F \rightarrow$ scalar funⁿ. for eg. $u = F(x, y, z)$

* Problem $U = 3x^2y + 3z^2$.

$\therefore \nabla u = (6xy)i + (3x^2)j + (6z)k$

$\nabla x = \frac{\partial}{\partial x}(3x^2 + y + z)$

Explaination

$$\nabla u = \left[\frac{\partial}{\partial x}(3x^2y + 3z^2)\, i\right] + \left[\frac{\partial}{\partial y}(3x^2y + 3z^2)\, j\right] + \left[\frac{\partial}{\partial z}(3x^2y + 3z^2)\right]$$

* Gradiant :- (∇u)

1) Gradiant tell that how the funⁿ is changing.
2) Gradiant tell the direction in which the rate of change of scalar funⁿ is maximum.

Ex $U = 3x^2y - 3zx^2 + 5xyz$

find the gradiant at (1,1,1)

$$\nabla u = [6xy - 6xz + 5yz]\,i + [3x^2 - 0 + 5xz]\,j + [0 - 3x^2 + 5xy]\,k$$

Put $x, y, z = 1$

$$= [6 - 6 + 5]\,i + [3 - 0 + 5]\,j + [-3 + 5]\,k = 5i + 8j + 2k$$

$du = \bar{\nabla} u \cdot d\bar{r}$

* If we looking for the maximum rate of change of u

then $\boxed{du = 0}$

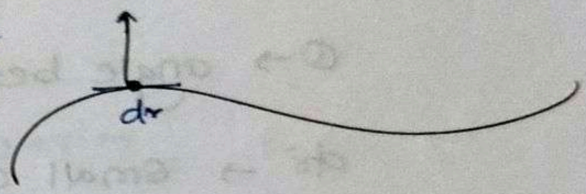
If the function is maximum or minimum then its differential will be $\boxed{\text{zero}}$

i.e. slope of funⁿ

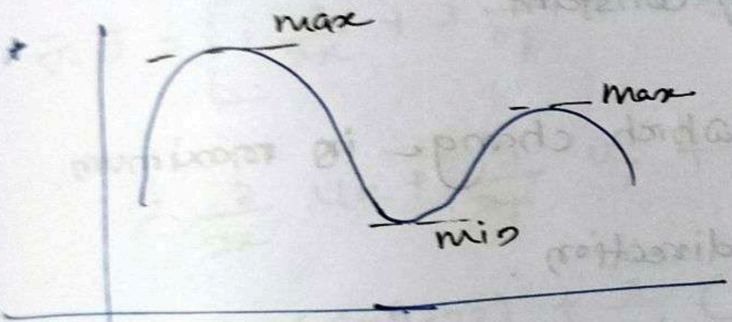
i.e. $\nabla u \, d\bar{r} = 0$

which mean. $\nabla u = 0$ or $d\bar{r} = 0$ or $\theta = 90^\circ$

$d\bar{r}$ will be on the surface



*



At the maximum point slope will be zero.

i.e $m = \frac{dy}{dx} = 0$.

if $y = 3x^2 - 2x + 5$

$\therefore \frac{dy}{dx} = 6x - 2 = 0$

$\therefore x = 2/6 = 1/3$

$\therefore \boxed{x = 1/3}$

$x = 1/3$ is the point at which. funⁿ have maximum value or minimum value

* Day - 13 *

9th may, 2022

Recalling :-

$du = \nabla u . d\bar{r}$

↳ Scalar

∇u → have both direction and magnitude.

In vector form - $|\nabla u| |d\bar{r}| . \cos\theta$

θ → angle between ∇u and $d\bar{r}$

$d\bar{r}$ → small change

$|d\bar{r}|$ = fixed / constant.

Finding the direction in which change is maximum.

∇u → is in normal direction.

When the angle betn Δu and dr is 0

i.e $\underline{\cos\theta = 1}$

Then there change in unit is maximum.

if $\theta = 90$ → change is zero.

if $\theta = 180$ → the magnitude will be the same but there is only change in sign.

* Gradient of vector valued function :-

It is generally called Divergence

Divergence of any vector :-

let $\bar{u} = (u_x)i + (u_y)j + (u_z)k$
↳ vector function.

$u_x = 3x^2 + 2z$

$u_y = -3z^3 y$

$u_z = 4xyz$

When we take ~~derivative~~ diversion

$$\bar{\nabla}.\bar{u} = \left[i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z}\right].[u_x i + u_y j + u_z k]$$

$$= \frac{\partial}{\partial x}u_x + \frac{\partial}{\partial y}u_y + \frac{\partial}{\partial z}u_z$$

$$= \frac{\partial}{\partial x}(3x^2+2z) + \frac{\partial}{\partial y}(-3z^3 y) + \frac{\partial}{\partial y}(4xyz)$$

$$= 6x + (-3z^3) + 4xy$$

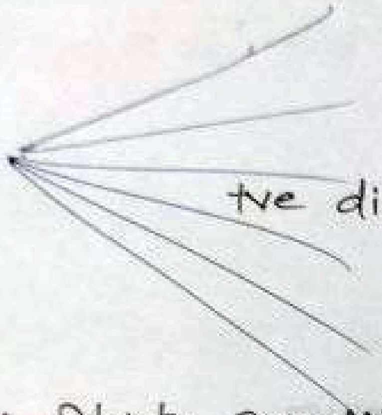
$$\boxed{\bar{\nabla}\bar{u} = 6x - 3z^3 + 4xy}$$ It is scalar function.

When we take diversion of any vector function, we will get an scalar.

* Physical meaning of diversion :-

Literal meaning of diversion → It is more like Spreading.

If it is moving away from Origin its called +ve Diversion.



+ve diversion.

Way by which something (for. e.g. traffic) "Spreading out" is called Diversion

* Reverse diversion is called Conversion or -ve diversion.

means. "Spreading in"

∇U = When ∇ operator operate with Scalar we get vector ⇒ which is called Gradient

$\nabla \vec{U}$ = When ∇ operator operate with vector we get Scalar ⇒ This is called Diversion

Ex:- $r = \sqrt{x^2+y^2+z^2}$ $\nabla r = ?$

$\rightarrow$

$$r = (x^2+y^2+z^2)^{1/2}$$

$$\nabla r = \frac{1}{2}(x^2+y^2+z^2)^{-1/2}[2xi+2yj+2zk]$$

$$= (x^2+y^2+z^2)^{-1/2}[xi+yj+zk]$$

$$= \frac{[xi+yj+zk]}{\sqrt{x^2+y^2+z^2}} = \frac{\bar{r}}{|\bar{r}|} = \hat{r}$$

* If you take gradient of scalar function we will get unit vector.

Ex:- $F(x,y,z) = e^x \cdot \sin y \cdot \ln(z)$

$\nabla f = ?$ $\nabla(\nabla f) = ?$

$$\nabla F = \left(i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z}\right) \cdot e^x \cdot \sin y \cdot \ln(z)$$

$$= \frac{\partial}{\partial x}(e^x \sin y \cdot \ln(z))i + \frac{\partial}{\partial y}(e^x \sin y \cdot \ln(z))j + \frac{\partial}{\partial z}(e^x \sin y \cdot \ln(z))k$$

$$= e^x \sin y \ln(z)\, i + e^x \cos y \cdot \ln(z)\, j + e^x \frac{\sin y}{z} k$$

$$= e^x \left[\sin y \ \ln(z)\, i + \cos y \ \ln(z)\, j + \frac{\sin y}{z} k \right]$$

$$\nabla.(\nabla.f) = \left(\frac{\partial}{\partial x} i + \frac{\partial}{\partial y} j + \frac{\partial}{\partial z} k \right) \left[e^x \left(\sin y \ \ln(z)\, i + \cos y \ \ln(z)\, j + \frac{\sin y}{z} k \right) \right]$$

$$= \frac{\partial}{\partial x} \left[e^x . \sin y \ \ln(z) \right] i + \frac{\partial}{\partial y} \left[e^x . \cos y \ \ln(z) \right] j + \frac{\partial}{\partial z} \left[e^x . \frac{\sin y}{z} k \right]$$

$$= \cancel{e^x . \sin y \ \ln(z)} - \cancel{e^x . \sin y \ \ln(z)} + e^x \sin y \left(\frac{-1}{z^2} \right)$$

$$\boxed{\nabla.(\nabla f) = - \frac{e^x . \sin y}{z^2}}$$

Answer is -ve which means the vector function is actually converging.

* Curl of Vector function :-

$\nabla . \bar{A} \rightarrow$ Divergence

$\nabla \times \bar{A} \rightarrow$ Curl

$\downarrow$

It will give you the assurance how the vector fun^n is turning around any point.

i.e there is change in direction.

Curl ie $\nabla \times \bar{A}$ - it will result in Vector function

$$\nabla \cdot \bar{A} = \begin{vmatrix} i & j & k \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ A_x & A_y & A_z \end{vmatrix}$$

$$= i\left[\frac{\partial}{\partial y}(A_z) - \frac{\partial}{\partial z}(A_y)\right] - j\left[\frac{\partial}{\partial x}(A_z) - \frac{\partial}{\partial z}(A_x)\right]$$

$$+ k\left[\frac{\partial}{\partial x}(A_y) - \frac{\partial}{\partial y}(A_x)\right]$$

It will give the rate at which the vector function is getting turn.

Ex

$A = 3x^2y\, i + 2yz\, k$

$$\nabla \times A = \begin{vmatrix} i & j & k \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 3x^2y & 0 & 2yz \end{vmatrix}$$

$$= i\left(\frac{\partial}{\partial y}(2yz)\right) - j\left[\frac{\partial}{\partial x}(2yz) - \frac{\partial}{\partial z}(3x^2y)\right]$$

$$+ k\left[0 - \frac{\partial}{\partial y}(3x^2y)\right]$$

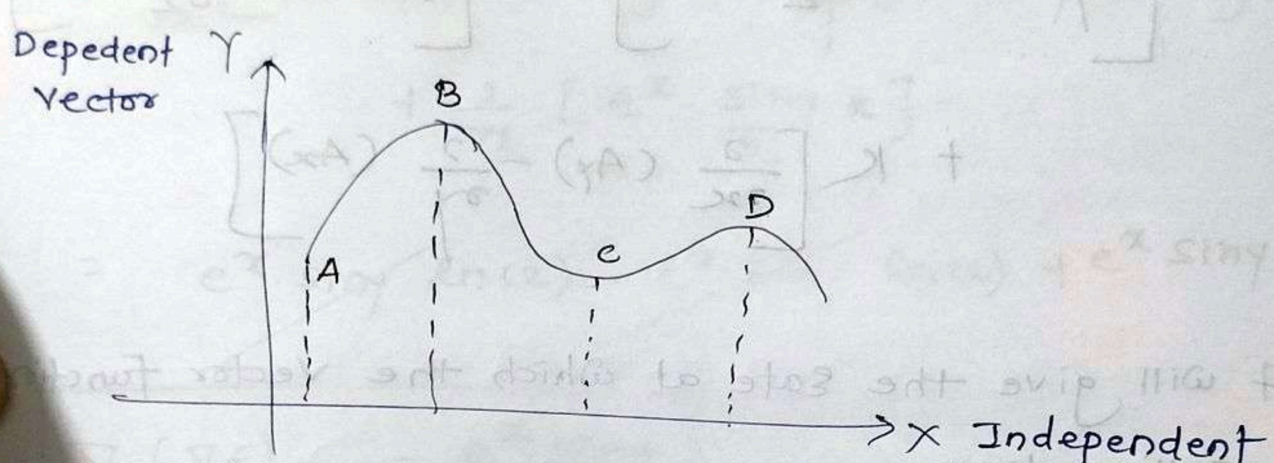
$$= 2z\, i - 0 + (-3x^2)k$$

$$= \underline{2z\, i - 3x^2 k}$$

$|\bar{\nabla}| = \sqrt{2^2+(-3)^2} = \sqrt{4+9} = \underline{\underline{\sqrt{13}}}$

This will tell us how the vector is changing in perticular direction.

* Increasing / Decreasing fun^n :-



A = Local minima

B = Local maxima

C = Local minima

D = Local maxima.

from point A & B function is increasing. i.e $\frac{dy}{dx} > 0$

If you increase x then function is increasing i.e the slope is $\underline{\underline{+ve}}$.

from point B & C function is decreasing i.e $\frac{dy}{dx} < 0$.

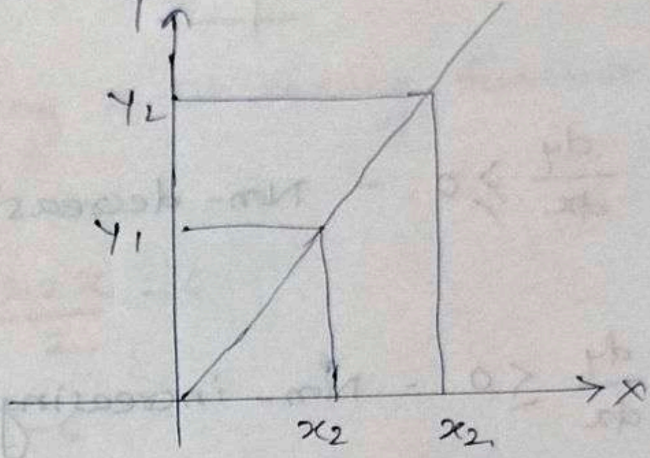
If you increase x then function is decreasing i.e the slope is $\underline{\underline{-ve}}$

Points A, B, C, D = are called point of Extremum.

i.e $\frac{dy}{dx} = 0$

Because tangent at all these points is parallel.
i.e $\theta = 0$, the slope is zero

* Show function is increasing :-



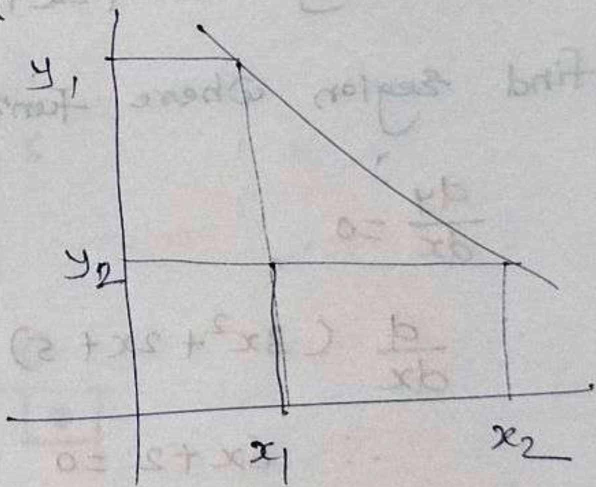
$y = f(x)$

If $x_1 < x_2$ & $f(x_1) < f(x_2)$.

Then function is called increasing Funⁿ.

If $x_1 < x_2$ But $f(x_1) > f(x_2)$

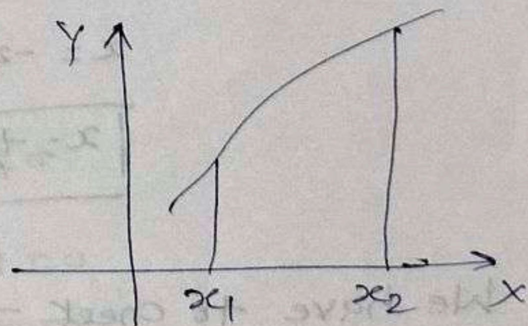
Then function is called decreasing Funⁿ



This is called zero derivative test.

Where we find function is increasing or decreasing without finding derivative.

* First Derivative Test :-

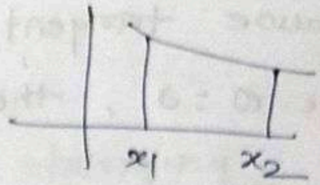


$y = f(x)$

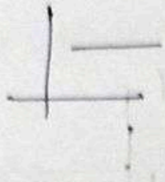
If $\frac{dy}{dx} > 0$

i.e The function is strictly increasing.

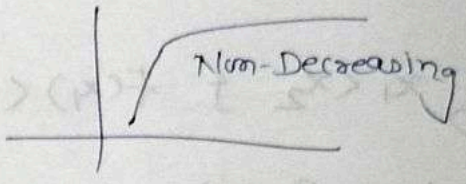
If $\frac{dy}{dx} < 0 \rightarrow$ function is strictly decreasing.



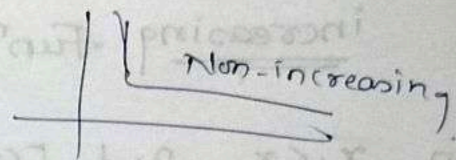
If $\frac{dy}{dx} = 0$



If $\frac{dy}{dx} \geq 0$ - Non-decreasing



If $\frac{dy}{dx} \leq 0$ - Non-increasing



Example :-

$$y = 3x^2 + 2x + 5$$

find region where funⁿ is increasing / Decreasing

$\rightarrow$

$$\frac{dy}{dx} = 0$$

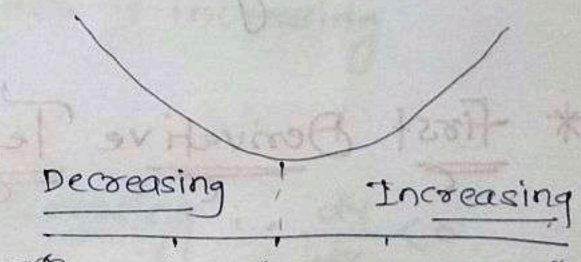
$$\therefore \frac{d}{dx}(3x^2 + 2x + 5) = 0$$

$$\therefore 6x + 2 = 0$$

$$\therefore 6x = -2$$

$$x = -2/6$$

$$\boxed{x = -1/3}$$



We have to check the point is maxima or minima.

$$\frac{dy}{dx} = 6x + 2$$

if $x = 0$ $\quad \frac{dy}{dx} = 2 > 0$

at $x=1$

$$\frac{dy}{dx} = 6(-1)+2 = -6+2 = -4<0$$

Example :- $y = \frac{x^4}{4} - 2x^3 + \frac{11x^2}{2} - 6x$

Find region for increasing & Decreasing function.

→ $$\frac{dy}{dx} = \frac{4x^3}{4} - 6x^2 + \frac{22x}{2} - 6$$

$$\frac{dy}{dx} = x^3 - 6x^2 + 11x - 6$$

for $x=1$ $\frac{dy}{dx} = 0$

$$\begin{array}{c|cccc} 1 & 1 & -6 & 11 & -6 \\ & & 1 & -5 & 6 \\ \hline & 1 & -5 & 6 & \boxed{0} \end{array}$$

$\therefore x^2 - 5x + 6 = 0$

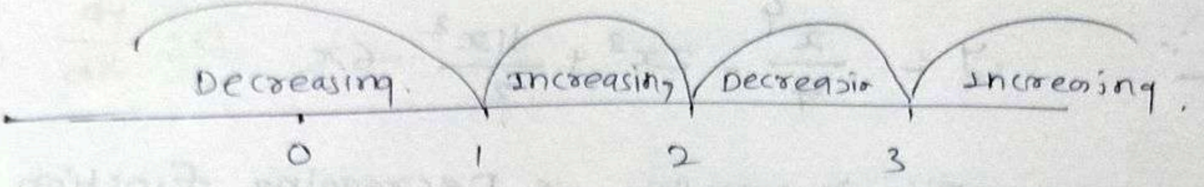
$\therefore x^2 - 2x - 3x + 6 = 0$

$\therefore x(x-2) - 3(x-2) = 0$

$\therefore x-2=0$ or $x-3=0$

$\boxed{x=2}$ or $\boxed{x=3}$

$\therefore \boxed{x = 1, 2, 3}$



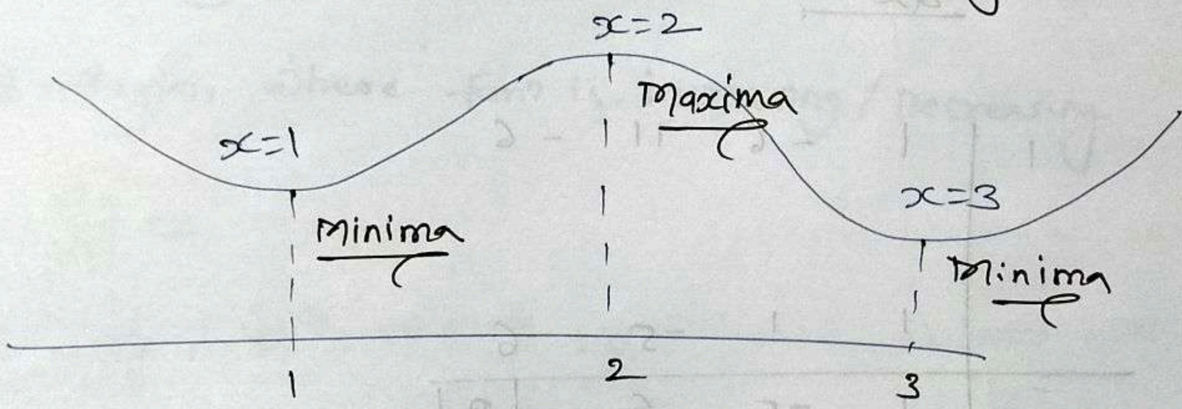
at $x=0$

$$\frac{dy}{dx} = x^3 - 6x^2 + 11x - 6 = 0 - 6 = \underline{\underline{-6}}$$

i.e fun^n is decreasing.

at $\boxed{x=1.5} \longrightarrow$ fun^n is increasing.

$\boxed{x=2.5} \longrightarrow$ fun^n is decreasing.



* Second Derivative Test :-

$$\frac{dy}{dx} = 0 \text{ for Both maxima \& minima}$$

To find the point is maximum or minimum we go for the second derivative test.

i.e,

$\frac{d^2y}{dx^2} > 0$ minima

$\frac{d^2y}{dx^2} < 0 \rightarrow$ maxima

$\frac{d^2y}{dx^2} = 0 \rightarrow$ Point of ~~Inflection~~. Inflation
i.e. Point where sign of slope changes

If the slope rate is +ve before inflation then slope rate becomes -ve after inflation.

$y = F(x)$

$\frac{dy}{dx} \Rightarrow$ shows the points of Extremum.

$\frac{d^2y}{dx^2} \Rightarrow$ shows the point is maxima / minima.

Suppose:

$y = F(x) \longrightarrow$ 1st Gen

$g = \frac{dy}{dx} \longrightarrow$ 2nd Gen (g is son y)

$h = \frac{dg}{dx} = \frac{d^2y}{dx^2} \longrightarrow$ 3rd Gen (h is son of g & Grandson of y

If $g = 0 \Rightarrow$ We get Extremum of y.

If $h = 0 \Rightarrow$ We get Extremum of g.
Also, get inflation of y.

* Day-12 *

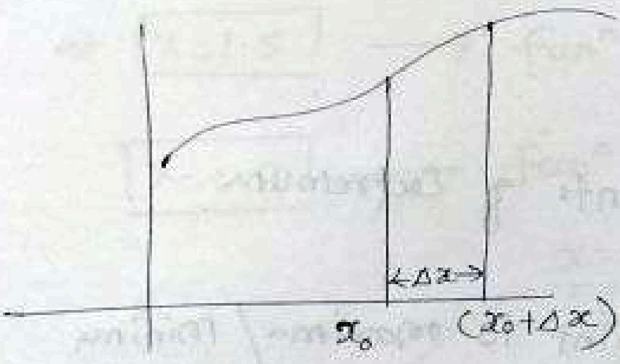
10th may, 2022
Tuesday

* Taylor Series :-

$y = f(x)$

$x = x_0 \longrightarrow$ called Intial value

$y = f(x_0) \longrightarrow$ called Intial function value



$y = f(x)$

Taylor Expansion :

$$y = f(x)\Big|_{x=x_0} + \frac{1}{1!}\frac{d}{dx} f(x)\Big|_{x=x_0} \Delta x + \frac{1}{2!}\frac{d^2}{dx^2} f^2(x)\Big|_{x=x_0} \Delta x^2$$

i.e. $$y = \sum_{i=0}^{\infty} \frac{1}{i!} f^{i}(x) (\Delta x)^{i}$$

Ex - $y = 3x^2+2$. Expand the y about $x=2$.

$$y = F(x)\Big|_{x=2} + \frac{1}{1!}\frac{d}{dx}F(x)\Big|_{x=2}\Delta x + \frac{1}{2!}\frac{d^2}{dx^2}F(x)\Big|_{x=2}\Delta x^2$$

$$\therefore F(x)_{x=2} = (3x^2+2)_{x=2} = 3(2^2)+2 = 14$$

$$\frac{1}{1!}\frac{d}{dx}(3x^2+2)\Delta x = 6x \text{ — at } x=2 = \underline{\underline{12\Delta x}}.$$

$$\frac{1}{2!}\frac{d^2}{dx}(3x^2+2) = \frac{1}{2}6 = 3\Delta x^2.$$

$$\therefore y = 14 + 12\Delta x + 3\Delta x^2$$

$$\therefore y = \underline{3\Delta x^2 + 12\Delta x + 14}$$

$$\Delta x = x - x_0$$

at $x_0 = 2$ $\quad \boxed{y = 14}$

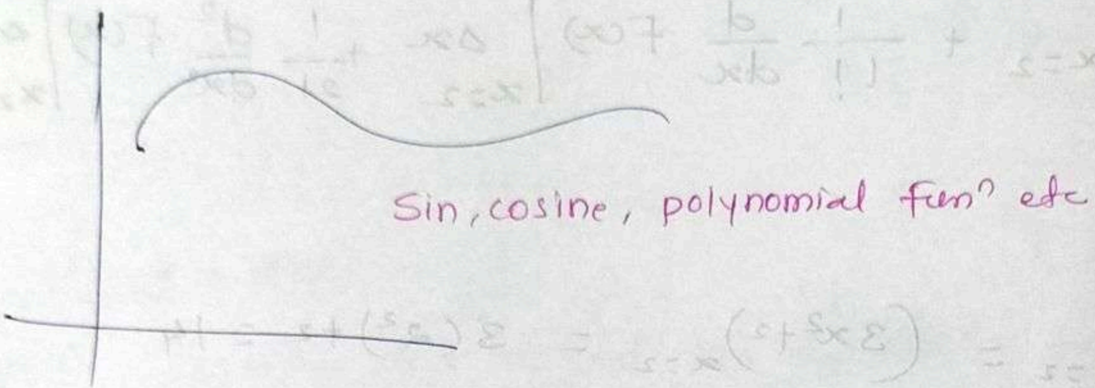
$\therefore \Delta x = 0$

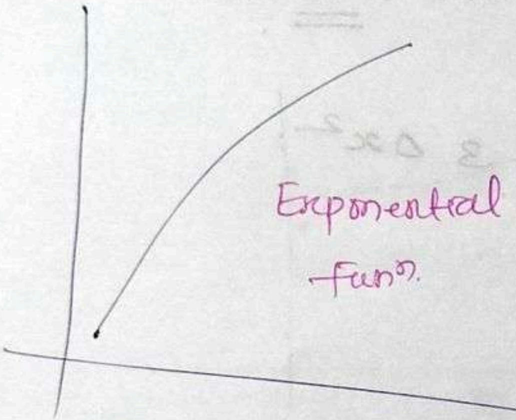
at $x_0 = 1$ $\quad y = 3 + 12 + 14 = 29$

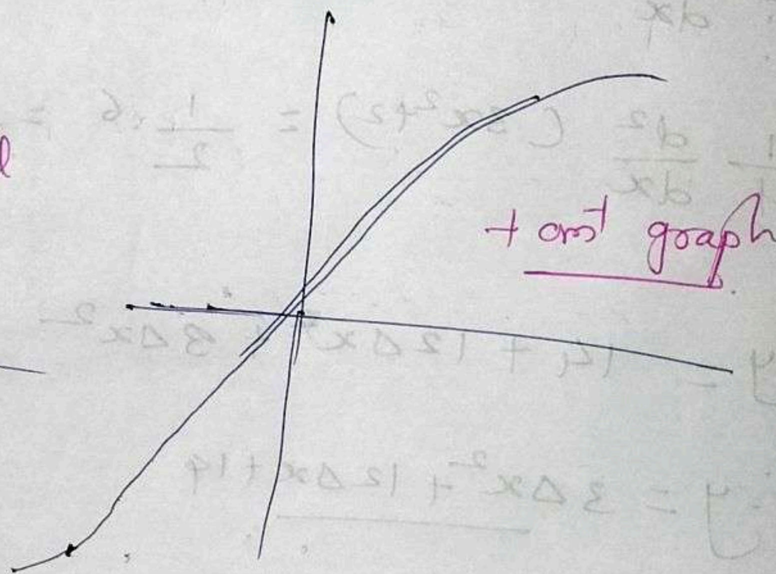
$\therefore \Delta x = 1$ $\quad \boxed{y = 29}$

* We can fit multiple function on same data
* Tayler series is used to approximate the function.

* Graph of some fun^n







* ~~Exponential~~ series - by Taylor series method

1) Exponential

$y = e^x$ at $x=0$

$$= \frac{d}{1!} f(x) + \frac{1}{1!}\frac{d}{dx} f(x)\Big|_{x=0} \Delta x + \frac{1}{2!}\frac{d^2}{dx^2} f(x)\Big|_{x=0} \Delta x^2 + \cdots$$

$$= e^x\Big|_{x=0} + 1\, e^x\Big|_{x=0} \Delta x + \frac{1}{2} e^x\Big|_{x=0} \Delta x^2 + \cdots$$

$$= 1 + \Delta x + \frac{\Delta x^2}{2} + \cdots$$

$\Delta x = x$

$$\boxed{\therefore\; y = 1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots}$$

2) Sin fun^n

$y = \sin x$ at $x = 0$.

$$y = F(x)\Big|_{x=0} + \frac{1}{1!}\frac{d}{dx}F(x)\Big|_{x=0}\Delta x + \frac{1}{2!}\frac{d^2}{dx}F(x)\Big|_{x=0}\Delta x^2 + \cdots$$

$$= \sin x\Big|_{x=0} + \frac{d}{dx}(\sin x)\Big|_{x=0}\Delta x + \frac{1}{2}\frac{d^2}{dx}(\sin x)\Big|_{x=0}\Delta x^2 + \cdots$$

$$= 0 + \cos x\Big|_{x=0}\Delta x - \frac{1}{2}\sin x\Big|_{x=0}\Delta x^2 + \cdots$$

$$= 0 + \Delta x - \frac{1}{2!}\times 0 - \frac{1}{3!}\Delta x^3 + \cdots$$

$\Delta x = x$

$$\boxed{\therefore\; y = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots}$$

* Taylor Series - For 2D

$$z = F(x, y)$$

$$z = F(x, y)\Big|_{\substack{x=x_0\\y=y_0}} + \frac{1}{1!}\frac{\partial F}{\partial x}\Big|_{\substack{x=x_0\\y=y_0}}\Delta x + \frac{1}{1!}\frac{\partial F}{\partial y}\Big|_{\substack{x=x_0\\y=y_0}}\Delta y + \frac{1}{2!}\frac{\partial^2 F}{\partial x^2}\Big|_{\substack{x=x_0\\y=y_0}}\Delta x^2$$

$$+ \frac{1}{2!}\frac{\partial^2 F}{\partial y^2}\Big|_{\substack{x=x_0\\y=y_0}}\Delta y^2 + \cdots$$

Ex - $Z = 3x^2 + y^2$

$$Z = f(x)\Big|_{\substack{x=0\\y=0}} + \frac{1}{1!}\frac{\partial}{\partial x}f\Big|_{\substack{x=0\\y=0}}\Delta x + \frac{1}{1!}\frac{\partial f}{\partial y}\Big|_{\substack{x=0\\y=0}}\Delta y$$

$$+ \frac{1}{2!}\frac{\partial^2 f}{\partial x^2}\Big|_{\substack{x=0\\y=0}}\Delta x^2 + \frac{1}{2!}\frac{\partial^2 f}{\partial y^2}\Big|_{\substack{x=0\\y=0}}\Delta y^2 + \dots$$

$$Z = (3x^2+y^2) + \frac{1}{1!}(6x+y^2)\Delta x + \frac{1}{1!}(3x^2+2y)\Delta y$$

$$+ \frac{1}{2!}(6+y^2)\Delta x^2 + \frac{1}{2!}(3x^2+2)\Delta y^2$$

at $x=0$ & $y=0$

$$Z = 0+0+0+\frac{1}{2}6\Delta x^2 + \frac{1}{2}2\Delta y^2$$

$$\therefore \boxed{Z = 3\Delta x^2 + \Delta y^2}$$

* Fourier Series is used when funⁿ is periodic in nature

* Maclauren Series is special case of Tayler Series. where intial point is origin.

* Linear function doesn't have 2nd derivative only have first derivative.

* Non-linear function have more than 1 derivatives.

* Day-13 * 24th may, 2022 Tuesday

* Jacobian matrix :-

It is the extension of $\frac{\partial f}{\partial x}$.

$F_1(x,y,z)$, $F_2(x,y,z)$, $F_3(x,y,z)$,

If then Jacobian matrix can be given as.

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} \end{bmatrix} \rightarrow \text{Jacobian matrix.}$$

Ex $\begin{bmatrix} F(x,y) \\ g(x,y) \end{bmatrix}$

$\rightarrow$ $x_0, y_0 \longrightarrow$ Intial Point

By Taylor series.

$$f(x,y)\Big|_{\substack{x_0\\y_0}} + \frac{1}{1!}\frac{\partial f(x,y)}{\partial x}\Big|_{\substack{x_0\\y_0}}\Delta x + \frac{1}{1!}\frac{\partial f(x,y)}{\partial y}\Big|_{\substack{x_0\\y=0}}\Delta y$$

$$g(x,y)\Big|_{\substack{x_0\\y_0}} + \frac{1}{1!}\frac{\partial g(x,y)}{\partial x}\Big|_{\substack{x_0\\y_0}}\Delta x + \frac{1}{1!}\frac{\partial g(x,y)}{\partial y}\Big|_{\substack{x_0\\y=0}}\Delta y$$

It can be written as

$$\begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} + \underbrace{\begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix}_{\substack{x_0=0 \\ y_0=0}}}_{\text{Jacobian Matrix}} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} + \dots$$

Ex.

$$f(x,y) = x^2+y^2+2xy$$
$$g(x,y) = x+y-2xy$$

find Linear expansion / Linearise fun^n at (x,y) $(1,1)$

$\rightarrow$ $\Delta x = 0.1$
$\Delta y = 0.2$.

$$f(x,y)_{1,1} = x^2+y^2+2xy$$
$$= 1+1+2 = 4$$

$$g(x,y)_{1,1} = x+y-2xy$$
$$= 1+1-2 = 0$$

$$J = \begin{bmatrix} \frac{df}{dx} & \frac{df}{dy} \\ \frac{dg}{dx} & \frac{dg}{dy} \end{bmatrix} = \begin{bmatrix} 2x+\cancel{y^2}+2y & 0+2y+2x \\ 1+0-2y & 0+1-2x \end{bmatrix}_{1,1}$$

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

$$= \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 4 & 4 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix}$$

$$= \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.4 + 0.8 \\ -0.1 - 0.2 \end{bmatrix}$$

$$= \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 1.2 \\ -0.3 \end{bmatrix} = \begin{bmatrix} 5.2 \\ -0.3 \end{bmatrix}$$

* Hessian Matrix :-

The Hessian matrix or Hessian is a square matrix of second-order partial derivatives of a scalar-valued fun?. or scalar-field.

It describes the local curvature of a function of many variables.

Hessian matrix is a square matrix.

It is represented a

$F(x,y)$

$$H = \begin{bmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} \\ \frac{\partial^2 F}{\partial y \partial x} & \frac{\partial^2 F}{\partial y^2} \end{bmatrix}$$

$F(x,y,z)$

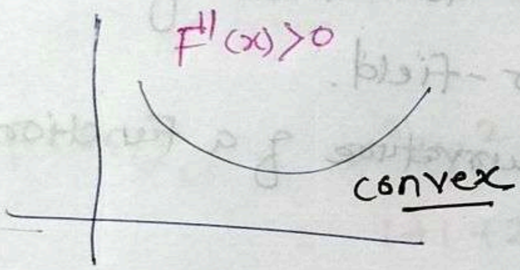
$$H = \begin{bmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} & \frac{\partial^2 F}{\partial x \partial z} \\ \frac{\partial^2 F}{\partial y \partial x} & \frac{\partial^2 F}{\partial y^2} & \frac{\partial^2 F}{\partial y \partial z} \\ \frac{\partial^2 F}{\partial z \partial x} & \frac{\partial^2 F}{\partial z \partial y} & \frac{\partial^2 F}{\partial z^2} \end{bmatrix}$$

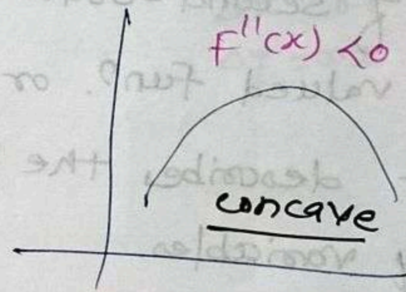
* Hessian matrix will give the curvature in multidimensional space.

* Hessian matrix are often used in machine learning and data science algorithms for optimizing a function of Interest.

* In convex optimization - Hessian matrix is positive, which is Semi-definite matrix.

* The hessian matrix g convex function is positive. Which means it can be zero or positive But can't be less than zero.





* If 2^{nd} derivative q any function is +ve then its a convex function.

* If 2^{nd} derivative q any function is -ve then its a convex function.

* Hessian matrix is ≥ 0 - for convex funⁿ.

* Hessian matrix is ≤ 0 for concave funⁿ.

* If $H > 0$ i.e convex fun. will give Local minima.

* If $H < 0$ i.e. concave fun^n. will give Local maxima

* Positive Definite matrix - will give Local minima

* Semi-positive Definite matrix will give Local maxima, Local minima, & Saddle point.

* 25th May, 2022 *

Wednesday

* Method for checking matrix positive definite or not.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{21} & a_{32} & a_{33} \end{bmatrix}$$

if $a_{11} > 0$

if $\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} > 0$

if $\det(A) > 0$

then the ~~posit~~ matrix is positive Definite matrix

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

$\rightarrow$

$a_{11} = 1$

$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ 2 & 1 \end{bmatrix} = 1+6 = 7$

$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & -3 & 2 \\ 2 & 1 & 4 \\ 1 & 2 & 4 \end{bmatrix} = 1(4-8) + 3(8-4) + 2(4-1)$

$= -4 + 12 + 6 = \underline{\underline{14}}$

Matrix A is positive definite matrix.

* If a_{11}, $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

One of these values is less than zero.

then the given matrix is Semi-positive definite matrix.

* All the eigen values of positive Definite matrix are positive

But all positive eigen values doesn't necessary that matrix is positive Definite.

* If the matrix is positive Definite then it will give local minima.

* cond^n for Negative definite matrix :

1) determint of odd order submatrix is <0

2) determinant of even order Submatrix is >0

e.g. $-ve \quad +ve \quad -ve$

$$\begin{bmatrix} \boxed{a_{11}} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

* Eigen-values of Negative definite matrix are negative.

For Local maxima -

$$\begin{cases} F_{xx} < 0 \leftarrow \boxed{1^{st} \text{ order}} \\ F_{xx} \cdot F_{yy} - F^2_{xy} > 0 \end{cases}$$

If the matrix is -ve Definite then it will give local maxima.

* If $F_{xx} \cdot F_{yy} - F^2_{xy} \leq 0$

Then this is the cond^n for saddle point.

* If all the eigen values are positive → test gives Local minima.

* If all the eigen values are negative → test gives Local maxima.

* If all the eigen values are mixed → test gives saddle point.

* If all the eigen values is zero → then test is inconclusive.

* Matrix Calculus :-

$$\bar{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \bar{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

X = set of variable. X is vector of variable

Y = set of function. Y is vector of function.

$\boxed{Y = Y(x)}$ → Y is a function of X.

Y can be written as.

$$Y = \begin{bmatrix} y_1(x_1, x_2, \cdots x_n) \\ y_2(x_1, x_2, \cdots x_n) \\ \vdots \\ y_m(x_1, x_2, \cdots x_n) \end{bmatrix}$$

X and Y are vector not <u>scalar</u>.

$$\frac{\partial Y}{\partial X} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \frac{\partial y_1}{\partial x_3} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_2}{\partial x_3} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots \\ \frac{\partial y_n}{\partial x_1} & \cdots \frac{\partial y_n}{\partial x_2} & \frac{\partial y_n}{\partial x_3} & \cdots & \frac{\partial y_n}{\partial x_n} \end{bmatrix}^T \rightarrow \text{It's a Jacobian Matrix.}$$

If y is scalar . $Y = [y]$

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial x_2} & \cdots & \frac{\partial y}{\partial x_n} \end{bmatrix}^T$$

If x is scalar.

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} \\ \frac{\partial y_2}{\partial x} \\ \vdots \\ \frac{\partial y_m}{\partial x} \end{bmatrix}^T.$$

Ex.

$$Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$y_1 = x_1^2 - x_2$$

$$y_2 = x_3^2 + 3x_2.$$

$$\rightarrow \frac{\partial y}{\partial x} = J^T = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \frac{\partial y_1}{\partial x_3} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_2}{\partial x_3} \end{bmatrix}$$

$$= \begin{bmatrix} 2x_1 & -1 & 0 \\ 0 & 3 & 2x_3 \end{bmatrix}^T = \begin{bmatrix} 2x_1 & 0 \\ -1 & 3 \\ 0 & 2x_3 \end{bmatrix}$$

* Some Properties *

1) $\frac{\partial}{\partial x}[CX] = C^T$

i.e. where $C = \begin{bmatrix} c_{11} & a_{12} & - & c_{1n} \\ c_{21} & a_{22} & - & c_{2n} \\ \vdots & & & \\ c_{n1} & - & - & c_{nn} \end{bmatrix}$ $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

$\downarrow$

It's matrix of constant.
i.e coeff of variable.

Ex

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$x_1 + 3x_2$

$2x_1 + 4x_2.$

$\frac{\partial y_1}{\partial x_1} = 1$ $\frac{\partial y_2}{\partial x_1} = 2$

$\frac{\partial y_1}{\partial x_2} = 3$ $\frac{\partial y_2}{\partial x_2} = 4$ $\Rightarrow \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

Hence proved

2) $\frac{\partial}{\partial X}[X^T C] = C$

3) $\frac{\partial}{\partial X}[X^T X] = 2X$

$\rightarrow \quad X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$X^T = [x_1 \ \ x_2]$

$X^T X = [x_1 \ x_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow [x_1^2 + x_2^2]$

$\frac{\partial}{\partial X}[X^T X] = \frac{\partial}{\partial X}[x_1^2 + x_2^2]$

$= \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} = 2\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$\boxed{\frac{\partial}{\partial X}[X^T X] = 2\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}$$

Hence Proved

4) $\frac{\partial}{\partial X}[X^T C X] = 2CX$

if C is symmetric

$\frac{\partial}{\partial X}[X^T C X] = CX + C^T X$ — if C isn't symmetric.

Ex - $g(x) = 2x_1 + 5x_2 + 12x_3$, $C = [2, 5 \ 12]$, $X = [x_1 \ x_2 \ x_3]$

$$g = \overset{C}{[2 \ \ 5 \ \ 12]} \overset{W = X^T}{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}$$

$g = C \cdot W$

$$\frac{\partial g}{\partial x} = \begin{bmatrix} \frac{\partial g}{\partial x_1} \\ \frac{\partial g}{\partial x_2} \\ \frac{\partial g}{\partial x_3} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_1}(2x_1 + 5x_2 + 12x_3) \\ \frac{\partial}{\partial x_2}(2x_1 + 5x_2 + 12x_3) \\ \frac{\partial}{\partial x_3}(2x_1 + 5x_2 + 12x_3) \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 12 \end{bmatrix} = C^T$$

26th May, 2022

Thursday

* Gradient Desecent Optimization :-

$F(x)$ — is objeetive function.

There are two possibilities, either you try to minimise it or maximise it

$$\begin{bmatrix} \text{min}: F(x) \rightarrow \text{Loss function} \\ \text{max}: F(x) \rightarrow \text{Gain function} \end{bmatrix}$$

We try to minimise the loss and maximise the Gain

$y = mx + c$ ——→ Eq^n of straight line

$y \Rightarrow$ Dependent Variable

$m \Rightarrow$ Slope

$x \Rightarrow$ Independent variable

$c \Rightarrow$ Intercept.

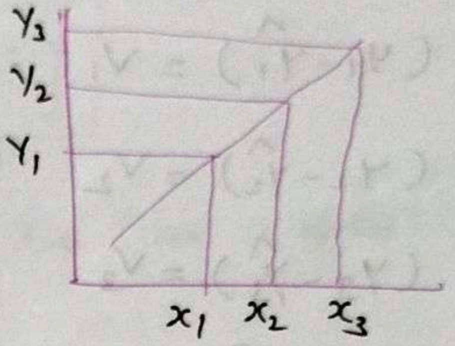
m & c are required for ~~finding~~ defining a model.

m & c are parameters.

values of x & y are changing but m & c are fixed for entire line.

Line is differentiate by

Parameter i.e. m & c



* Methods to find Parameter :-

* Least Squares : Regression (Linear)

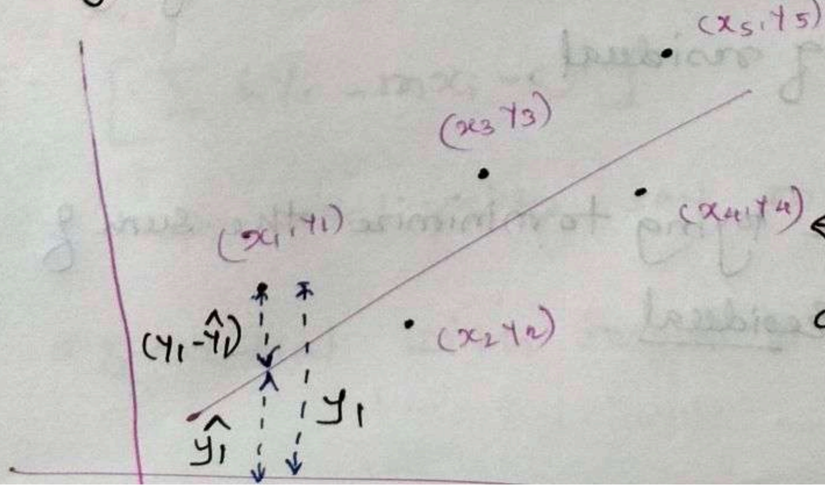
1> $y = mx + c$ – Linear model.

2> $y = ax_1 + bx_2 + c \rightarrow$ Linear model

3> $y = ax_1^2 + bx_2^2 + cx_1x_2 + d \rightarrow$ Linear model

4> $y = a^2x_1 + b^2 + x_2 + c^3 \Rightarrow$ Non-linear model

5> $y = e^{ax} \rightarrow$ Non-linear model



← These are observed values.

1: $y = mx + c$ — model.

$\hat{y}_1 = mx_1 + c$

$\hat{y} \rightarrow$ value of x_1 through model at x_1

$(y_1 - \hat{y}_1) = v_1$

$(y_2 - \hat{y}_2) = v_2$

$(y_3 - \hat{y}_3) = v_3$

$(y_4 - \hat{y}_4 = v_4$

— called Residual

Residual value → It is the difference between observed value of estimated value

→

$(\Sigma v_i)^2 \longrightarrow$ objective fun^n.

↓

This should be minimum.

$(\Sigma v_i)^2 = (v_1^2 + v_2^2 + v_3^2 + v_4^2 + v_5^2)$

We need to minimize the value of objective fun^n $\Sigma(v_i)^2$ of residual.

Here we are trying to minimize the sum of square of residual.

* objective funⁿ → min $[\sum v_i^2]$

$v_i = y_i - \hat{y}_i$

$$\boxed{v_i = y_i - (mx_i + c)}$$

$$\boxed{f = \min\left[\sum (y_i - mx_i - c)^2\right]}$$

* If we are changing the line means changing in m & c.

We have to optimise (minimum) the objective function by changing m & c.

We can take

$$\left.\begin{aligned} \frac{\partial F}{\partial m} &= 0 \\ \frac{\partial F}{\partial c} &= 0 \end{aligned}\right\}$$ This approach is called Linear Regression or Least Square method.

* Gradiant Descent method.

$$f = \left[\sum (y_i - mx_i - c)^2\right]$$

Let their intial values, say $m=1$ $c=1$

find differential with respect to m & c

$$\left.\frac{\partial f}{\partial m}\right|_{\substack{m=1 \\ c=1}} = \_\_\_\_\_\_\_\_\_\_$$ Here will gradiant with respect to m

$$\left.\frac{\partial f}{\partial c}\right|_{\substack{m=1 \\ c=1}} = \_\_\_\_\_\_\_\_\_\_$$ Here will gradiant with respect to c.

Gradiant is just like slope in 1D

We have to move in the direction where slope is minimise.

Ex - $y = mx + 3$

x	y
1	4.1
2	4.9
3	6.2

Let start, $m = 0.5$ - Intial value

When $m = 0.5$ $\hat{y}_1 = ?$

$$\hat{y}_1 = 0.5x + 3$$
$$= 0.5(1) + 3 \quad -- (x=1)$$
$$= 3.5$$

$$\hat{y}_2 = 0.5x + 3$$
$$= 0.5(2) + 3 \quad --. (x=2)$$
$$= 4$$

$$\hat{y}_3 = 0.5x + 3$$
$$= 0.5(3) + 3 \quad -- (x=3)$$
$$= 4.5$$

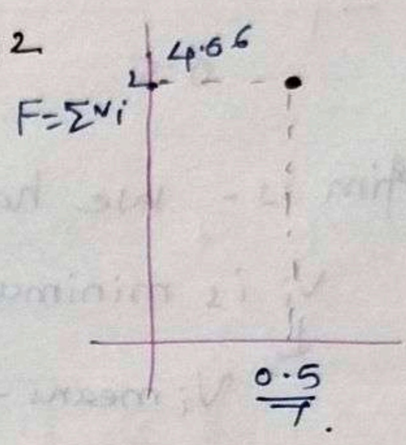
$\hat{y}_1, \hat{y}_2$ & $\hat{y}_3 \longrightarrow$ Predicted Value

$v_1 = y_1 - \hat{y}_1 = 4.1 - 3.5 = 0.6$

$v_2 = y_2 - \hat{y}_2 = 4.9 - 4 = 0.9$ } Residual value

$v_3 = y_3 - \hat{y}_3 = 6.2 - 4.5 = 1.7$

$$F = \sum_{i=1}^{3} v_i^2 = (0.6)^2 + (0.9)^2 + (1.7)^2$$
$$= 0.36 + 0.81 + 2.89$$
$$= 4.06$$



$* \; y = mx + 3$

$$\sum v_i^2 = \sum (y_i - mx_i - 3)^2$$

$$\frac{d}{dm} (y_i - mx_i - 3)^2$$
$$= [2(y_i - mx_i - 3)(-x_i)]$$

$$\text{Total Sum} = [2(4.1 - m(1) - 3)(-1)]$$
$$+ [2[4.9 - m(2) - 3](-2)]$$
$$+ [2[6.2 - m(3) - 3](-3)]$$

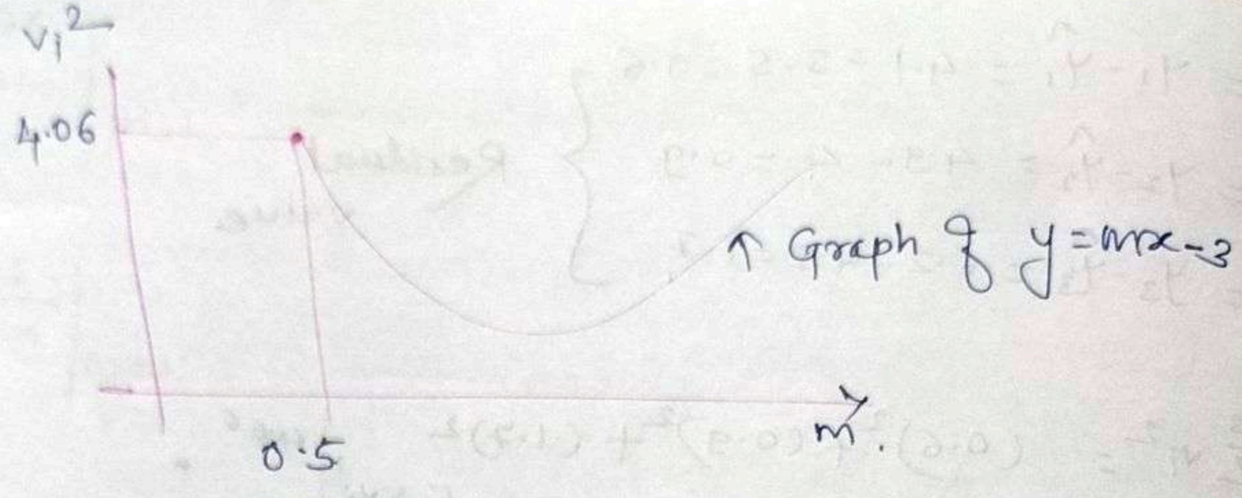
$$= 2(4.1 - m - 3)(-1) + 2(4.9 - 2m - 3)(-2)$$
$$+ 2(6.2 - 3m - 3)(-3)$$

$$= -2(1.1 - m) - 4(1.9 - 2m) - 6(3.2 - 3m)$$

$$= -2.2 + 2m - 7.6 + 8m - 19.2 + 18m$$

slope &

$$= -29 + 28m = \underline{-15} \; (\text{For } m = 0.5)$$ Residual



Aim is - We have to find the point at which

V_i is minimum.

↓

V_i means - Residual

$\boxed{* \text{ Slope, } m = \tan\theta}$

* <u>Learning Rate</u> :-

It is choosen by trial & <u>error</u>

α = Step size $= 0.01$ (Assumed).

$$\Delta m = \alpha \left(\frac{\partial}{\partial x} \sum v_i^2 \right)$$

$$= 0.01 \times (-15) = \underline{-0.15}$$

$$m_1 = m_0 - \Delta m = 0.5 - (-0.15)$$

$$= 0.65$$

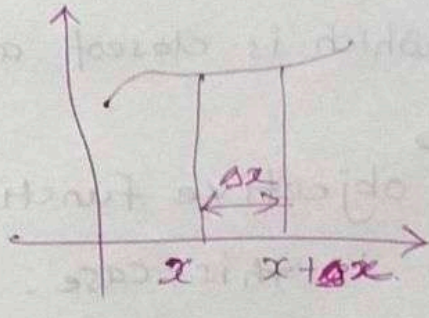
If $\alpha = 0.005$

$$m = m_0 - \alpha \nabla (v_i^2)$$

$$= 0.5 + 0.075 = 0.575$$

* Taylor Serie :-

$$F(x) = F(x)\Big|_{x=x_0} + F'(x)\Big|_{x=x_0} \Delta x$$



$$\cancel{F(x+\Delta x)} = \cancel{F(x+\Delta x)}$$

$$F(x+\Delta x) = F_x + F'(x)\,\Delta x$$

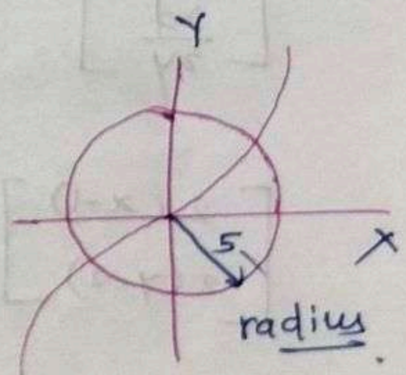
$$\therefore \boxed{F'(x) = \frac{F(x+\Delta x) - F(x)}{\Delta x}}$$

* 27th may, 2022 *

Friday

If $F(x) = x$. $[0 \le x \le 10]$

$F(x)\big|_{min} = 0$ $\qquad F(x)\big|_{max} = 10$



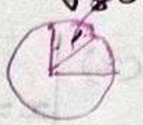
* Lagrange multiplier :-

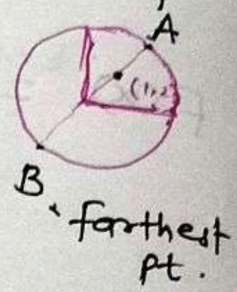
$y = F(x) = x^3$ — It is a objective function.

We want to maximize or minimise this function.

Constraint : $x^2 + y^2 = 25$

Ex - find the point on circle $x^2+y^2=80$ which is closest and farther from $(1,2)$.





$\Longrightarrow$ objective function is distance in this case.

Max or min $(d) = (x-1)^2 + (y-2)^2$

constraint - Point lies on the circle.

$$x^2+y^2=80$$

$$\nabla d = \cancel{K}\, \nabla c$$

where

$\nabla d \rightarrow$ Gradiant of objective function.

$\cancel{K} \rightarrow$ constant i.e. Lagrange multiplier.

$\nabla c \rightarrow$ Gradiant of constraint.

$$\therefore \begin{bmatrix} \frac{\partial d}{\partial x} \\ \frac{\partial d}{\partial y} \end{bmatrix} = K \begin{bmatrix} \frac{\partial}{\partial x} c \\ \frac{\partial}{\partial y} c \end{bmatrix}$$

$$\begin{bmatrix} 2(x-1) \\ 2(y-2) \end{bmatrix} = K \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

$2(x-1) = 2Kx \quad - ⓐ \rightarrow \quad x-1 = Kx \quad - ①$

$2(y-2) = 2Ky \quad - ⓑ \quad\quad y-2 = Ky. \quad - ②$

from ① $K = \dfrac{x-1}{x}$ from ② $= \dfrac{y-2}{y} = K$

$\therefore \frac{x-1}{x} = \frac{y-2}{y}$

$\therefore \cancel{xy} - y = \cancel{xy} - 2x$

$\therefore \boxed{y = 2x}$

$\therefore x^2 + y^2 = 80 \rightarrow x^2 + (2x)^2 = 80$

$x^2 + 4\,2x\,x^2 + 4x^2 = 80$

$x^2 + 8x\,x^2 + 4x^2 = 80$

$5x^2 + 8x^3 = 80$

$\therefore x^2 + (2x)^2 = 80$

$\therefore x^2 + 4x^2 = 80$

$\therefore 5x^2 = 80$

$x^2 = 80/5 = 16$

$\boxed{x = \pm 4}$

$\rightarrow x - 1 = kx$	$x - 1 = kx$
$4 - 1 = k(4)$	$-4 - 1 = k(-4)$
$\boxed{k = \frac{3}{4}}$	$-5 = -4k$ $\quad \cancel{k = 5/4}$
	$\boxed{k = \frac{5}{4}}$

$\rightarrow y - 2 = ky$

$y - 2 = \frac{3}{4}y - 4y - 8 = 3y$

$\boxed{y = 8}$

$\rightarrow y - 2 = ky$

$y - 2 = \frac{5}{4}y - 4y - 8 = 5y \quad \boxed{y = -8}$

I - find the dimension of box with maximum vol^m. If the total Surface area is 64 unit.

→

condition - Total Surface area = 64

$$2(lb + bh + lh) = 64$$

$$lb + bh + lh = 32$$

as $l = b = h$

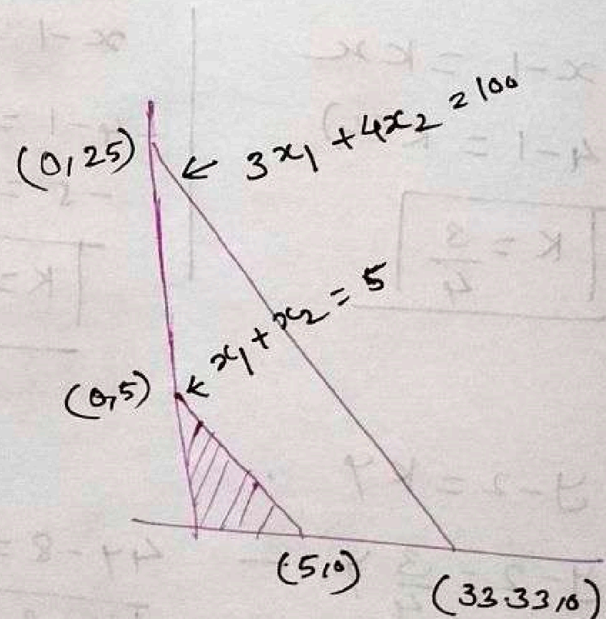
$$\therefore \quad 3l^2 = 32$$

$$l^2 = \frac{32}{3}$$

$$\boxed{l = \sqrt{\frac{32}{3}}}$$

Ex. $3x_1 + 4x_2 \leq 100$

$x_1 + x_2 \leq 5$



* If Objective fun^n is non-linear.
  - Use Lagrange's multipliers.
* If you have C_1 & C_2 selected → use PCA

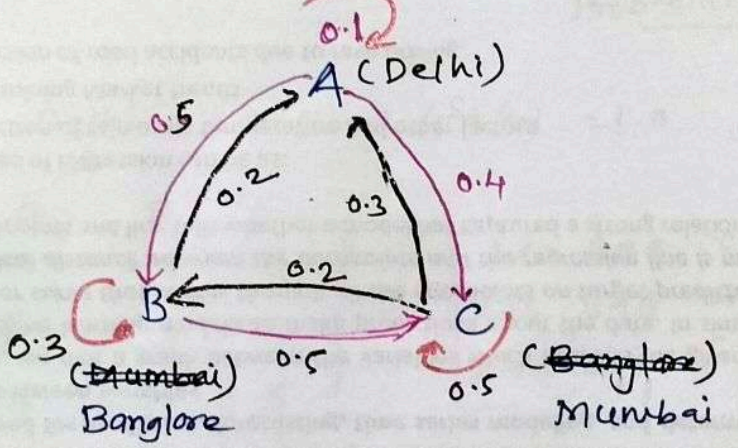
* 30^{th} may, 2022 *

Monday

* Markov chain / Process :-

Markov chain is Process or stochastic model in which there are sequence of events and each event is depends on previous event

For eg. - Suppose there are 3 cities A, B & C



Some one in city A - there is possibility that he go to city A or city B.

~~Person~~ Probability of person goes to mumbai from Delhi is 0.5. and for Delhi - Mumbai is 0.4

and there is also one possibility that he/she remains in the same city, don't go to any other city (Prob - 0.1)

IIIy, for Banglore

1) Banglore - Mumbai - 0.2

2) Banglore - Delhi - 0.5

3) Remain in Banglore - 0.3

IIIy, for Mumbai

1) Mumbai - Delhi - 0.3

2) Mumbai - Banglore - 0.2

3) Remain in Mumbai - 0.5

This process are called as "Markov chain"

Imp - The sum of total outward arrow (Probability) should be equal to 1.

i.e At Delhi → 0.5 + 0.4 + 0.1 = 1.0

$$T = \begin{array}{c|ccc} & A & B & C \\ \hline A & 0.1 & 0.2 & 0.3 \\ B & 0.5 & 0.3 & 0.2 \\ C & 0.4 & 0.5 & 0.5 \end{array} = \begin{bmatrix} AA & B\to A & C\to A \\ A\to B & B\to B & C\to B \\ A\to C & B\to C & C\to C \end{bmatrix}$$

A, B, C are states.

This matrix, T is also called as transition matrix.

and If take sum of each column, it is equal to 1.

Ex

$$T = \begin{array}{c|cc} & A & B \\ \hline A & 0.7 & 0.5 \\ B & 0.3 & 0.5 \end{array}_{2\times 2}$$ No. of states = 2.

if state $[0.4 \quad 0.6]$ → From this we can say that there is probability that particle has 40% chance hat it is state A + 60% chance that particle it is state B.

$\Rightarrow \begin{array}{cc} A & B \\ \end{array}$ $[1, 0]$ → Particle is in state A

$\Rightarrow$ $[0, 1]$ ⇒ Particle is in state B.

Ex ⇒ Today's Date 30^{th} may, 2022, Weather = Sunny.

Depending on the whether condⁿ, ~~we ca~~ what we exepect the temp of tomarroz

High probability is that tomarrow is sunny because today's whether is sunny.

$$T = \begin{array}{c|ccc} & S & C & R \\ \hline S & 0.5 & 0.4 & 0.3 \\ C & 0.3 & 0.3 & 0.5 \\ R & 0.2 & 0.3 & 0.2 \end{array}$$

State vector for 1^{st} June $= V \begin{bmatrix} 0.25 \\ 0.25 \\ 0.50 \end{bmatrix}$

What will be the whether condⁿ on 2^{nd} June is

T.V = Transition matrix x value error

$$= \begin{bmatrix} 0.5 & 0.4 & 0.3 \\ 0.3 & 0.3 & 0.5 \\ 0.2 & 0.3 & 0.2 \end{bmatrix} \begin{bmatrix} 0.25 \\ 0.25 \\ 0.140 \end{bmatrix}$$

$$= \begin{bmatrix} 0.375 \\ 0.400 \\ 0.225 \end{bmatrix}$$

markov matrix - you will get 1 as it's one of the eigen value. It's a Universal truth.

* Diagonalisation of Matrix :-

→ $\boxed{A = P D P^{-1}}$

Diagonalisation of matrix A.

P = matrix of eigen vector of A.

D = Diagonal matrix of eigen values of A

$$\therefore \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{array}{cc} v_1 & v_2 \\ \multicolumn{2}{c}{\begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix}} \end{array} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix}^{-1}$$

$v_1 = [v_{11} \; v_{12}]$ → Eigen vector for eigen value λ_1

$v_2 = [v_{21} \; v_{22}]$ → Eigen vector for eigen value λ_2.

* The

$$\therefore A^2 = [P D P^{-1}][P D P^{-1}]$$

$$\boxed{A^2 = P D^2 P^{-1}}$$

* IF any matrix (Transition matrix) is multiplied by its eigen vector then it gives $\underline{\underline{\lambda V}}$.

$\therefore TV = \lambda V$ is called Stationary State for ~~when~~ → $\lambda = 1$

$\boxed{\therefore TV = V}$ (As $\lambda = 1$ ~~For Stationary~~)

$$\begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.2 & 0.1 & 0.3 \\ 0.3 & 0.6 & 0.5 \end{bmatrix}$$

Eigen values for this matrix.

$$\lambda_1 = 1, \quad \lambda_2 = \frac{-1+\sqrt{21}}{20} \qquad \lambda_3 = \frac{1+\sqrt{21}}{20}$$

Eigen vector corresponding these eigen values.

$$V_1 = \begin{bmatrix} 9/13 \\ 19/39 \\ 1 \end{bmatrix} \quad V_2 = \quad V_3 \begin{bmatrix} -3+\sqrt{21}/6 \\ -3+\sqrt{21}/6 \\ 1 \end{bmatrix}$$

$\therefore \quad TV_1 = \lambda_1 V_1$

$\boxed{TV_1 = V_1}$ — for $\lambda_1 = 1$

$\therefore$ V_1 is called stationary state → sum of stationary state vector is $\frac{1}{T}$.

$$\begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.2 & 0.1 & 0.3 \\ 0.3 & 0.6 & 0.5 \end{bmatrix} \begin{bmatrix} 9/13 \\ 19/39 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.5\times\frac{9}{13} + 0.3\times\frac{19}{39} + 0.2\times 1 \\ 0.2\times\frac{9}{13} + 0.1\times\frac{19}{39} + 0.3\times 1 \\ 0.3\times\frac{9}{13} + 0.6\times\frac{19}{39} + 0.5\times 1 \end{bmatrix} = \begin{bmatrix} 9/13 \\ 19/39 \\ 1 \end{bmatrix}$$

Hence proved.

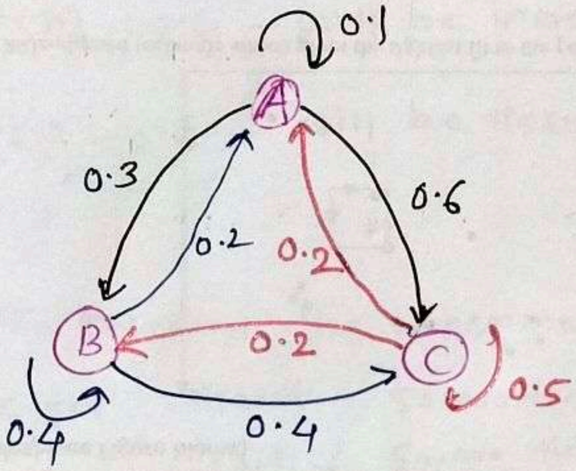
Markov chain is used to find

1) Stationary state
2) Last/Final state $\underline{\underline{T}}$

Ex. Transition matrix, T =

$$\begin{array}{c} \\ A \\ B \\ C \end{array}\begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.2 \\ 0.6 & 0.4 & 0.5 \end{bmatrix} \end{array}$$

A = Delhi   B = Mumbai   C = Banglore.



Digramatical Representation of Transition matrix

→ If you are sure that person is in Delhi, then the state vector is → $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

→ If you are sure that person is in Mumbai, then the state vector is → $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

→ If you are sure that person is in Banglore, then the state vector is → $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

If take coin and toss it 10 times.

Result	1	2	3	4	5	6	7	8	9	10
	T	H	H	T	T	H	T	H	T	T

Probability of getting head. $P(H) = \frac{4}{10} = 0.4$

Probability of getting Tail. $P(T) = \frac{6}{10} = 0.6$

$\therefore [0.4 \quad 0.6]$

If we increase the number of experiement.

$P(H) = 0.4$ - will be increase and close to 0.5

& $P(T) = 0.6$ - will be decrease and close to 0.5.

It is called limiting sense.

If we ~~took~~ perform the experiment 1,00,000 times

and we get Head - 50,000 times and

Tail - 50,000 times.

So the probability is $[1/2 \quad 1/2]$

We already know that the probability is $1/2$. So there is no need to perform Simulation.

- We can prove the same thing using Markov chain or Markov process.

n the case of No. of experiment is 10.

Relative frequency given by, $RF = \frac{n}{N}$.

$n \rightarrow$ No of outcomes

$N \rightarrow$ No of experiments.

$R.F(H) = 4/10 = 0.4 \qquad RF(T) = 6/10 = 0.6$

When N becomes ∞.

$\therefore \quad R.f = \frac{n}{\infty} \Rightarrow$ Probability otherwise its just relative frequency.

Ex. Transition matrix for coin problem

$$T = \begin{array}{c c} & \begin{array}{cc} H & T \end{array} \\ \begin{array}{c} H \\ T \end{array} & \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} \end{array}$$

Eigenvalues are $\lambda = 1 \quad \lambda_2 = 0$

Eigen vector are = $\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$\therefore$ If we consider $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ - But it's not stationary state

$\therefore$ So, stationary state for coin problem, is $\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}^T$

$$\boxed{\frac{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}{2} \Rightarrow \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}}$$

$\leftarrow$ Explanation How we get stationary state

. To clear confusion regarding previous problem for stationary stat.

consider 3x3 Transition matrix and it eigenvector

$$i_s = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

$\rightarrow$ $TV = \lambda V$

$$\begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}_{3\times 3} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

$\downarrow$

for Now its not stationary state

If ~~you~~ we want to convert this to stationary state.

$\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \rightarrow$ Sum of column is 4

then divide the vector by 4

$$\therefore \begin{bmatrix} 2/4 \\ 1/4 \\ 1/4 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.25 \\ 0.25 \end{bmatrix} \rightarrow \text{Now the Sum is } 1$$

$\downarrow$

Now, we call it as stationary state.