Linear Regression & Gradient Descend

- Supervised learning
  - simple regression
  - Multiple "

x	y
1	4
2	5
3	6
6	?

Linear Regression :

tries to fit a straight line into our data

- Hill Climbing → Steepest ascend

$$y = mx + c$$

$x_1 = 1$
$y_1' = 0.6(1) + 0.5$
$= 1.1$

$x_2 = 2$
$y_2' = 0.6(2) + 0.5$
$= 1.7$

$x_3 = 3$
$y_3' = 0.6(3) + 0.5$
$= 2.3$

We want to find the optimal values of m & c to get the best fitted straight line. Our predicted line needs to be as close as to the actual values as possible.

LOSS FUNCTION

$$\boxed{(y_1 - y_1')^2 + (y_2 - y_2')^2 + (y_3 - y_3')^2}$$

⇓

SSR
(Sum of Squared Residuals)
(Squared Error)

$$\text{avg. error} = \underbrace{\frac{(y_1 - y_1')^2 + (y_2 - y_2')^2 + (y_3 - y_3')^2}{n}}_{\Downarrow}$$

root mean square error

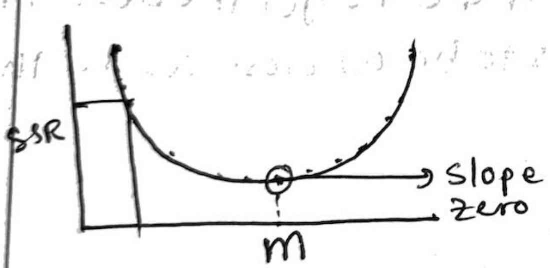
$$SSR = (y_1 - y_1')^2 + (y_2 - y_2')^2 + (y_3 - y_3')^2$$

$$= \{y_1 - (mx_1 + c)\}^2 + \{y_2 - (mx_2 + c)\}^2 + \{y_3 - (mx_3 + c)\}^2$$

$$+ \cdots + \{y_n - (mx_n + c)\}^2$$

$$m_{new} = m_{old} - \left\{ \frac{d}{dm_{old}}(SSR) \times \underbrace{\eta}_{\text{learning rate}} \right\}$$

Gradient Descent: tries to find slope where "m" is zero



Same হবে "c" এর জন্য

$$\frac{d(SSR)}{dm_1} = 2\{y_1 - (mx_1 + c)\}(-x_1) + 2\{y_2 - (m_1 x_2 + c)\}(-x_2)$$

$$+ 2\{y_3 - (m_1 x_3 + c)\}(-x_3)$$

$\Downarrow$ m_2

$= \cdots\cdots$ [here we put the values of m, x, y, c]

$= 0.4$

$m_{new} = m_{old}$ − step size

step size = slope × learning rate

{How do we know if we have found the right slope?}

$$\frac{d}{dx}(x^n) = nx^{n-1}$$

$$\frac{d}{dx}(fg) = fg' + gf'$$

$$\frac{d}{dx}(f/g) = \frac{gf' - fg'}{g^2}$$

$$m_{new} = m_{old} - \left\{ \frac{d(SSR)}{dm_{old}} \times \eta \right\}$$

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

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

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

$$\frac{d}{dx}(\cot x) = -\operatorname{cosec}^2 x$$

$$\frac{d}{dx}(\sec x) = \sec x \tan x$$

$$\frac{d}{dx}(\operatorname{cosec} x) = -\operatorname{cosec} x \cot x$$

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

$$\frac{d}{dx}(a^x) = a^x \ln a$$

$$\frac{d}{dx}(\ln x) = 1/x$$

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

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

Gradient Descent

⇒ two or more derivatives of the same function are called Gradients.

⇒ An algorithm which uses gradient to descent to the lowest point of a loss function.

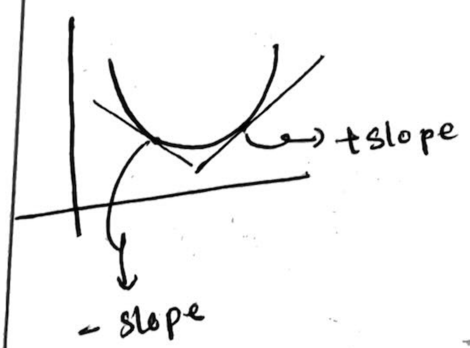
When we are only dealing with "m"; we have one unknown,
But if we have more complex function, for multiple features,
the function will be complicated; for example; polynomial.
These will increase possible directions.

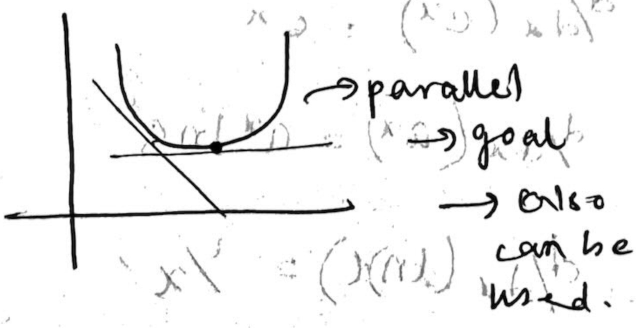
Example:

$y = mx + c$

$$\text{4 directions} \Rightarrow \begin{bmatrix} +m & +c \\ -m & -c \\ -m & +c \\ +m & -c \end{bmatrix} \rightarrow \text{no. of directions change exponentially.}$$

So trial & error will be computationally expensive.
So we solve the problem mathematically.





$m = m - \text{slope}$ will give a soln