Gradient Descent

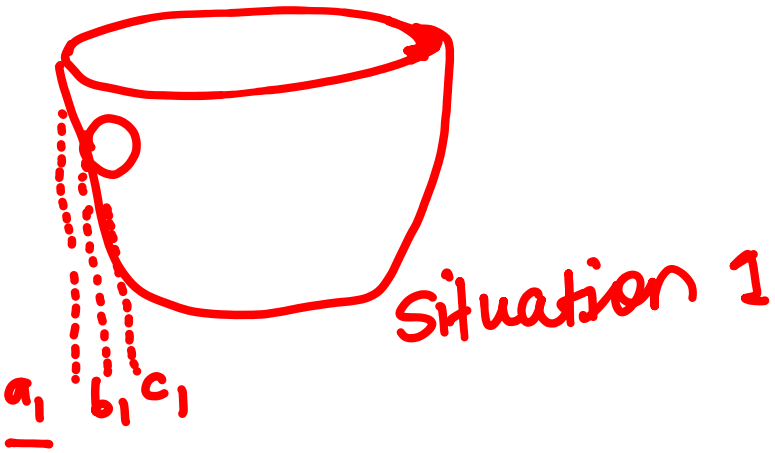
$y = \theta_0 + \theta_1 x$ ..... (θ_0, θ_1 = Regr^ param...)

↳ It is used to estimate Regr^ parameters.

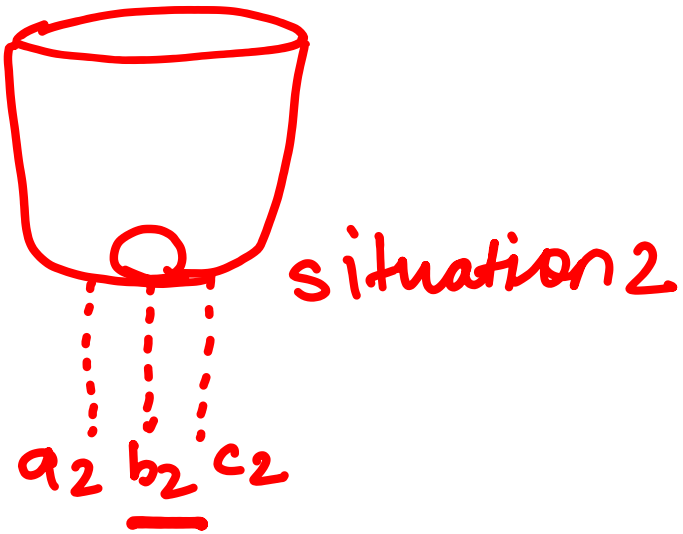
Why GD over OLS

→ OLS is expensive in terms of complexity (it involved matrix inversion)

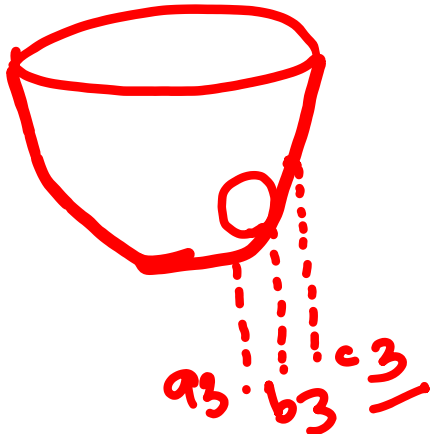
→ OLS fails on non linear trends.



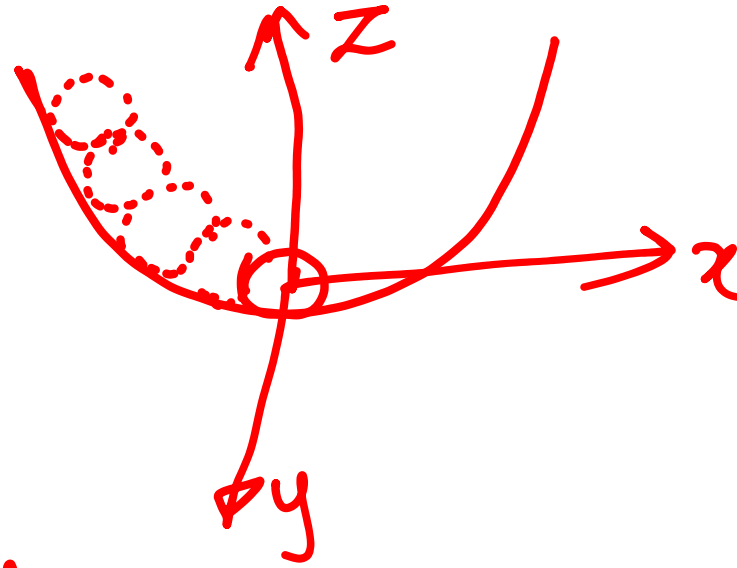
situation 1



situation 2



situation 3

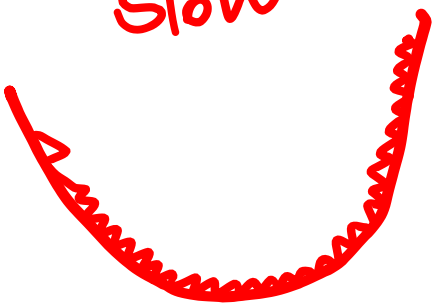


z = loss

x, y = co-ordinates

Learning Rate

slow



fast

Minimum of fun^n

$f(x) = x^4 - 10x^2 + 9$ ——— (1)

Step 1 find derivative of given fun^n

$$\frac{d}{dx}(f(x)) = \frac{d}{dx}(x^4 - 10x^2 + 9)$$

$$\frac{d}{dx}(f(x)) = 4x^3 - 20x$$

to find minima we need to equate it to 0

$$4x^3 - 20x = 0$$

$4x^3 - 20x = 0$

$4x^3 = 20x$

$x^2 = 5$

$x = \sqrt{5}$

$x = \pm 2.25$

Iterative method → G.D

$f(x) = x^4 - 10x^2 + 9$

Consider initial value of $x = 5$

Learning rate $\alpha = 0.001$

⟩ gradient of any fun is given by

$$\frac{d}{dx}(f(x)) = \frac{d}{dx}(x^4 - 10x^2 + 9)$$

$$= 4x^3 - 20x$$

1st iteration

$x = 5$ gradient $= 4(5)^3 - 20(5)$

$= 4(125) - 100$

$= 500 - 400$

We need to decrease x by α times gradient

New x = Prev(X) $- \alpha (4x^3 - 20x)$

$= 5 - 0.001\,(400)$

$= 5 - 0.4$

New $x = 4.6$

2nd iteration

New value of $x = prev(x) - \alpha(4x^3 - 20x)$

$= 4.6 - 0.001\left(4(4.6)^3 - 20(4.6)\right)$

$= 4.6 - 0.001\,(389.34 - 92)$

$= 4.6 - 0.001(297.34)$

$= 4.6 - 0.297$

$= 4.30$

New value of x =

$\sqrt{5} = 2.24$