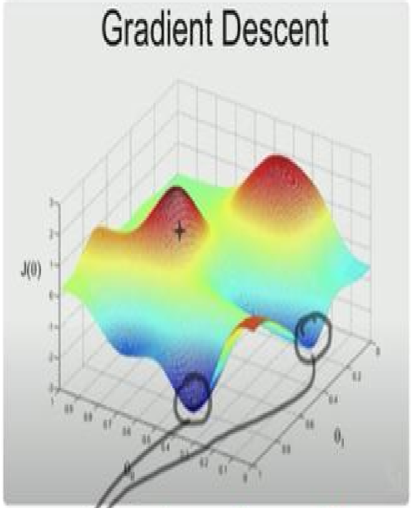
- Start at some point
- Take 360° look & figure out which direction baby step leads to lowest value
- Repeat



Lowest point

→ here θ_2 is not included to make a 3-D diagram

* Slope ⇒ Give rate of change
Derivative ⇒ Slope
-Derivative ⇒ -ve slope which means the rate of lowest descent.

'a := b'
↳ expression
$a := a + 1$
↳ assignment

Ex:- $f(x, y)$ → here unlike slope the function doesn't change in one direction like the above fig. We use gradient instead of slope (a vector that points in the direction of greatest rate of increase of the function

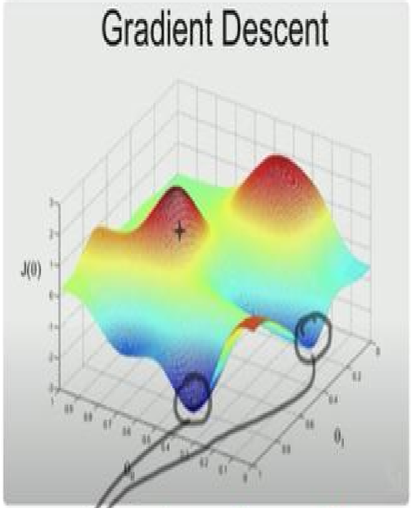
$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$$

If the gradient vector at $(2, 3)$ is $\nabla f = (4, 5)$, it means:

- Moving in the direction where x increases by 4 units and y increases by 5 units will make you climb the hill the fastest.

$x := x - \alpha\, \partial J(\theta)$] Gradient descent algo

- Start at some point
- Take 360° look & figure out which direction baby step leads to lowest value
- Repeat



Lowest point

→ here θ_2 is not included to make a 3-D diagram

* Slope ⇒ Give rate of change
Derivative ⇒ Slope
-Derivative ⇒ -ve slope which means the rate of lowest descent.

'a := b'
↳ expression
a := a+1
↳ assignment

Ex:- $f(x, y)$ → here unlike slope the function doesn't change in one direction like the above fig. We use gradient instead of slope (a vector that points in the direction of greatest rate of increase of the function

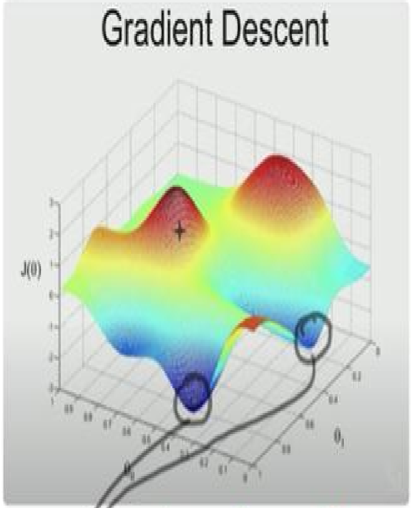
$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$$

If the gradient vector at $(2, 3)$ is $\nabla f = (4, 5)$, it means:

- Moving in the direction where x increases by 4 units and y increases by 5 units will make you climb the hill the fastest.

$x := x - \alpha \, \partial J(\theta)$] Gradient descent algo

- Start at some point
- Take 360° look & figure out which direction baby step leads to lowest value
- Repeat



Lowest point

→ here θ_2 is not included to make a 3-D diagram

* Slope ⇒ Give rate of change
Derivative ⇒ Slope
-Derivative ⇒ -ve slope which means the rate of lowest descent.

'a = b'
↳ expression
$a := a + 1$
↳ assignment

Ex:- $f(x, y)$ → here unlike slope the function doesn't change in one direction like the above fig. We use gradient instead of slope (a vector that points in the direction of greatest rate of increase of the function

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$$

If the gradient vector at $(2, 3)$ is $\nabla f = (4, 5)$, it means:

- Moving in the direction where x increases by 4 units and y increases by 5 units will make you climb the hill the fastest.

$x := x - \alpha\, \partial J(\theta)$] Gradient descent step