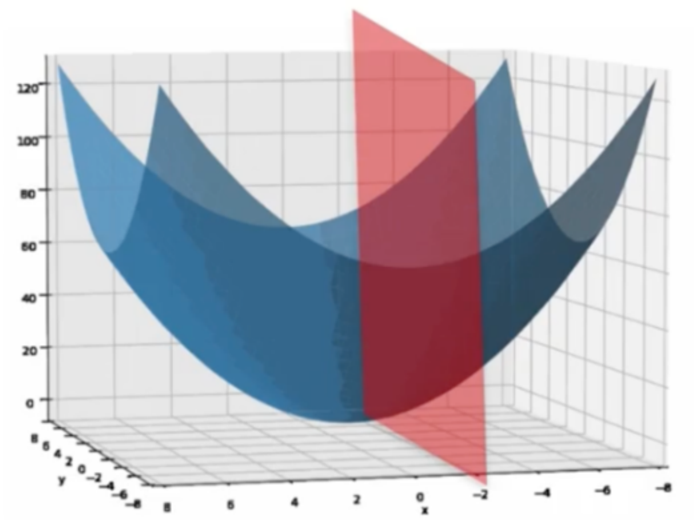
Day-83, Feb 21, 2025 (Falgun 9, 2081)

1. Partial Derivatives (function of one, two, three and n variables)
2. Gradients and Maxima/Minima
3. Optimization with Gradients (An Example)
4. Optimization with Gradients (Analyticals)
5. Optimization using Gradient Descent in One Variable

Slicing the Space



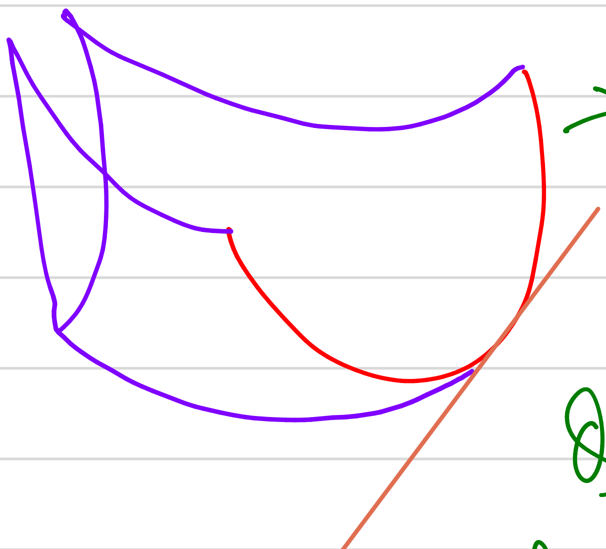
that you cut it with a plane like this, what do you get?

DeepLearning.AI

→ Cutting with a plane □ what happens?

Original Space

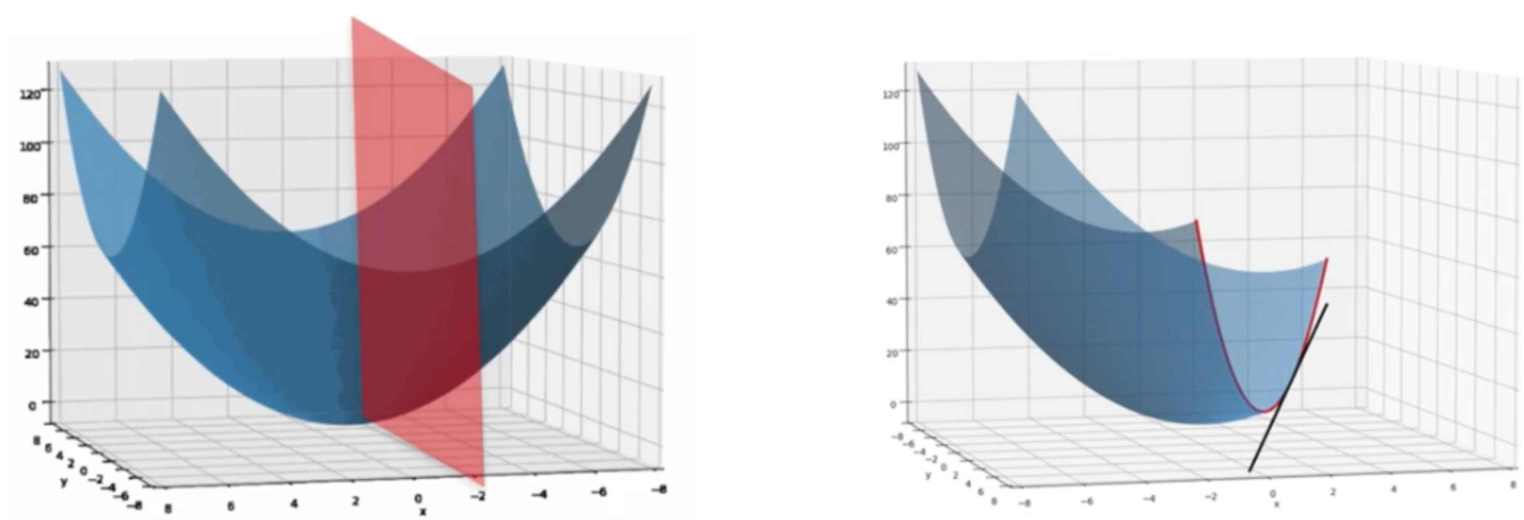
→ Bowl shape Space



→ partial Derivative

Q. Can we cut in another direction ?

Slicing the Space

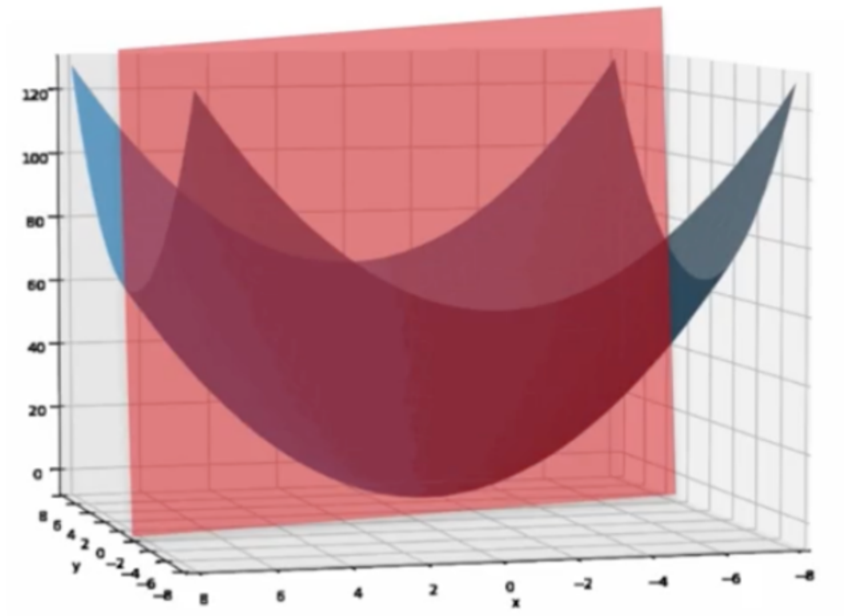


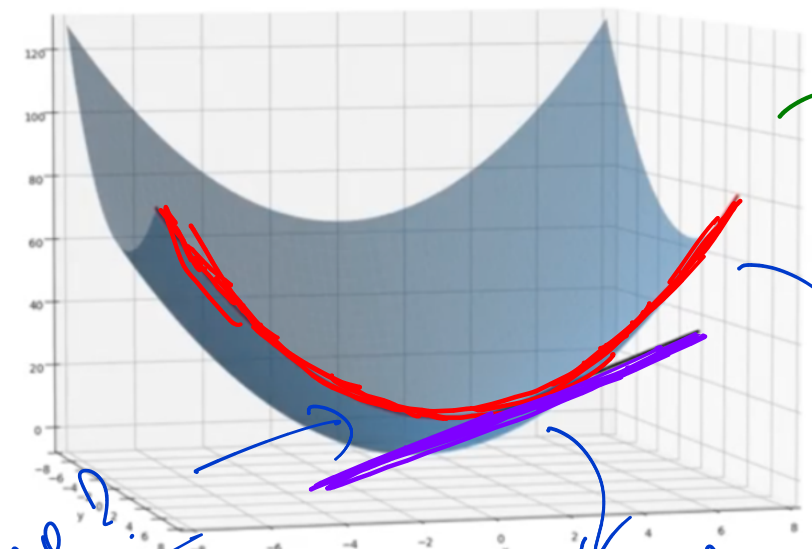
that's the partial derivative.

DeepLearning.AI

Slicing the Space

in another direction ? we get-





Slope?

tangent

$f(x,y) = x^2 + y^2$

Treat y as a constant

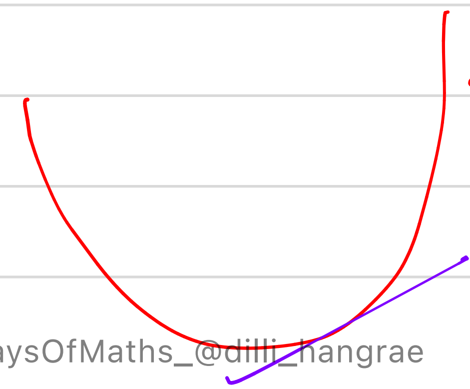
function of one variable 'x'?

$$\left[\frac{\partial f}{\partial x} = 0 + 2y. \right]$$

then you get also a red parabola and that red parabola is a function of one variable.

DeepLearning.AI

So, the slope of the tangent line? Answer is Partial Derivative (∂)



function of one variable x

slope? Partial Derivative

$f(x,y) = x^2 + y^2$ → constant

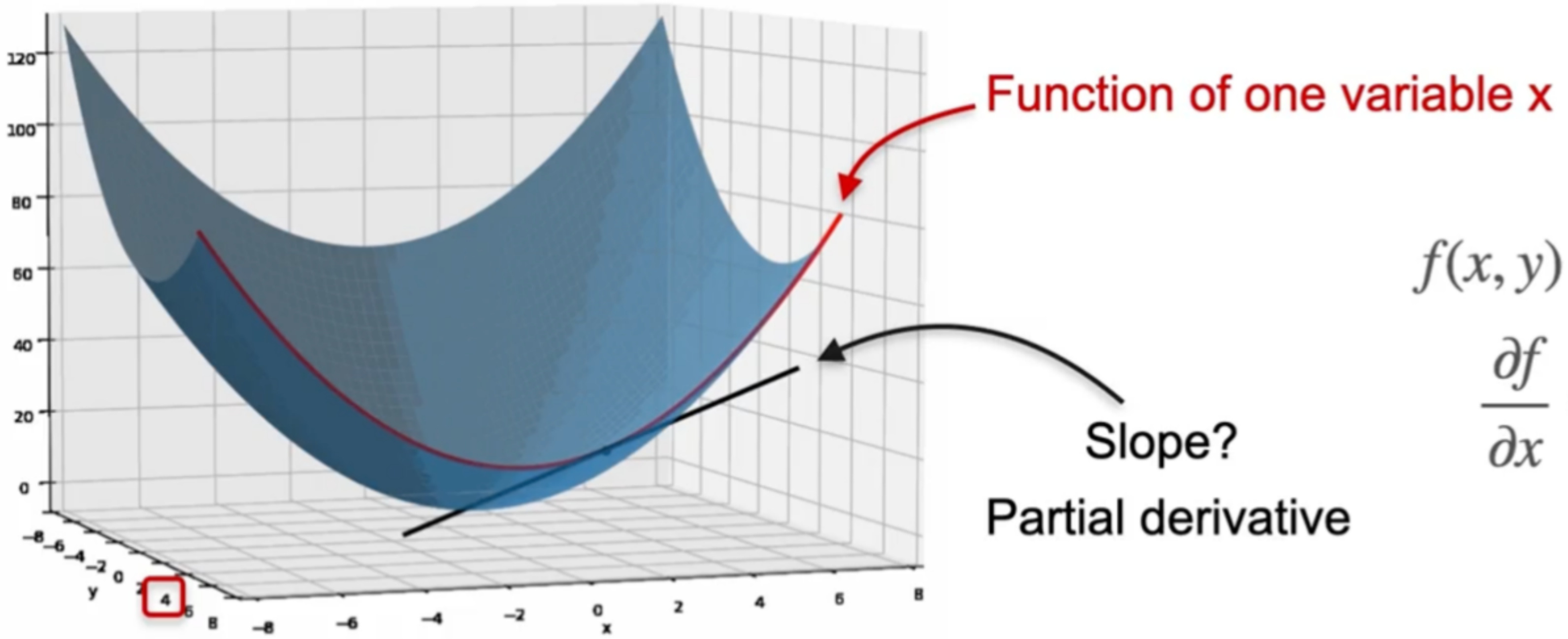
$$\frac{\partial f}{\partial y} = 2x + 0$$

Constant d is 0.

Partial Derivatives

$$f(x, y) = x^2 + y^2$$

Treat y as a constant



Constant

$$f(x, y) = x^2 + \boxed{y^2}$$

$$\frac{\partial f}{\partial x} = 2x + \boxed{0}$$

Derivative = 0

Treat x as Constant

$$\frac{\partial f}{\partial x} = x^2 + y^2$$

$$\Rightarrow 2y$$

Treat y as Constant.

$$\frac{\partial f}{\partial y} = x^2 + y^2$$

$$\Rightarrow 2x$$

$$f(x, y)$$

$$f_x = \frac{\partial f}{\partial x} \qquad f_y = \frac{\partial f}{\partial y}$$

Partial derivative of f with respect to x

$$f_x \Rightarrow \frac{\partial f}{\partial x}$$

Partial Derivative of y w.r.t to 'y'

$$f_y = \frac{\partial f}{\partial y}$$

Intro to Partial Derivatives

Partial Derivative notation $\quad f(x,y) = x^2 + y^2$

$\frac{\partial f}{\partial x} \Rightarrow \partial y$

$\frac{\partial f}{\partial y} \Rightarrow \partial x$

$f(x,y) = 3x^2 y^3$

$\frac{\partial f}{\partial x} = ?$

$\Rightarrow 3x^2 \times y^3$

Keeping y constant $\Rightarrow 6x y^3$

Intro To Partial Derivatives

$f(x,y) = 1 + y^2$

$f(x,y) = \boxed{x^2} + y^2$

$$\frac{\partial f}{\partial x} = 2x$$

$$\frac{\partial f}{\partial y} = 2y$$

TASK

Find partial derivative of f with respect to x

Step 1: Treat all other variables as a constant. In our case x.

Step 2: Differentiate the function using the normal rules of differentiation.

DeepLearning.AI

Partial Derivatives (More Examples)

$$f(x,y) = 3x^2$$

treat as constant coefficient

$$\frac{\partial f}{\partial x} = 3(2x)$$

$f(x,y) = 3x^2y^3$

keep y const.

TASK

Find partial derivate of f with respect to x

Step 1: Treat all other variables as a constant. In our case y.

Step 2: Differentiate the function using the normal rules of differentiation.

$\Rightarrow 3 \times 2x \times y^3$

$\Rightarrow 6xy^3$

$\frac{\partial f}{\partial y} = 3x^2y^3$

$\Rightarrow 3 \times 3y^2$

$\Rightarrow 9y^2x^2$

$\Rightarrow 9x^2y^2$

DeepLearning.AI

Example:

$f(x,y,z) \Rightarrow x^2y^3z^4 + x^2y^2 + x^3z^3 + y^4z^4$

$\frac{\partial f}{\partial x} \Rightarrow 2xy^3z^4 + 2xy^2 + 3x^2z^3 + y^4z^4$

Partial Derivatives (More Examples)

$$f(x,y) = 3x^2y^3$$

$$\frac{\partial f}{\partial y} = 3(x^2)(3y^2)$$

$$= 9x^2y^2$$

TASK

What is the partial derivate of f with respect to y?

Step 1: Treat all other variables as a constant. In our case x.

Step 2: Differentiate the function using the normal rules of differentiation.

DeepLearning.AI

Gradients:

$$f(x,y) = x^2 + y^2$$

Treat y as Constant

$$\frac{\partial f}{\partial x} = 2x$$

Treat x as Constant

$$\frac{\partial f}{\partial y} = 2y$$

Gradient

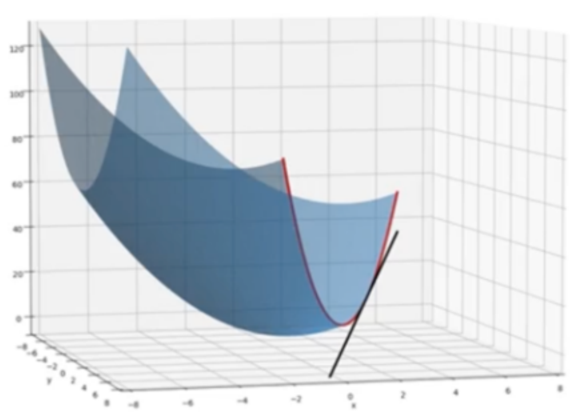
$$\begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

$$\nabla f \Rightarrow \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

Gradient

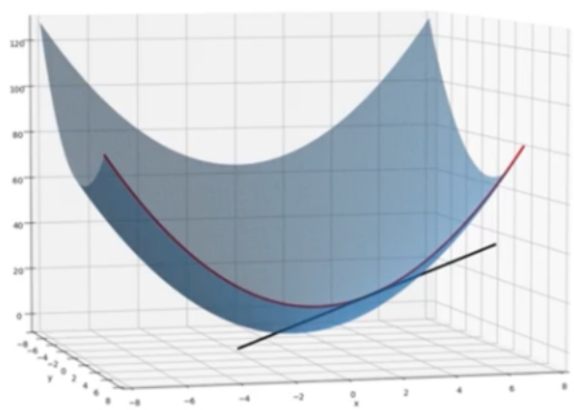
$$f(x,y) = x^2 + y^2$$

Treat y as a constant



$$\frac{\partial f}{\partial x} = 2x$$

Treat x as a constant



$$\frac{\partial f}{\partial y} = 2y$$

Gradient

$$\begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

$$f(x,y) = x^2 + y^2$$

The gradient of $f(x,y)$ is

$$\nabla f \Rightarrow \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

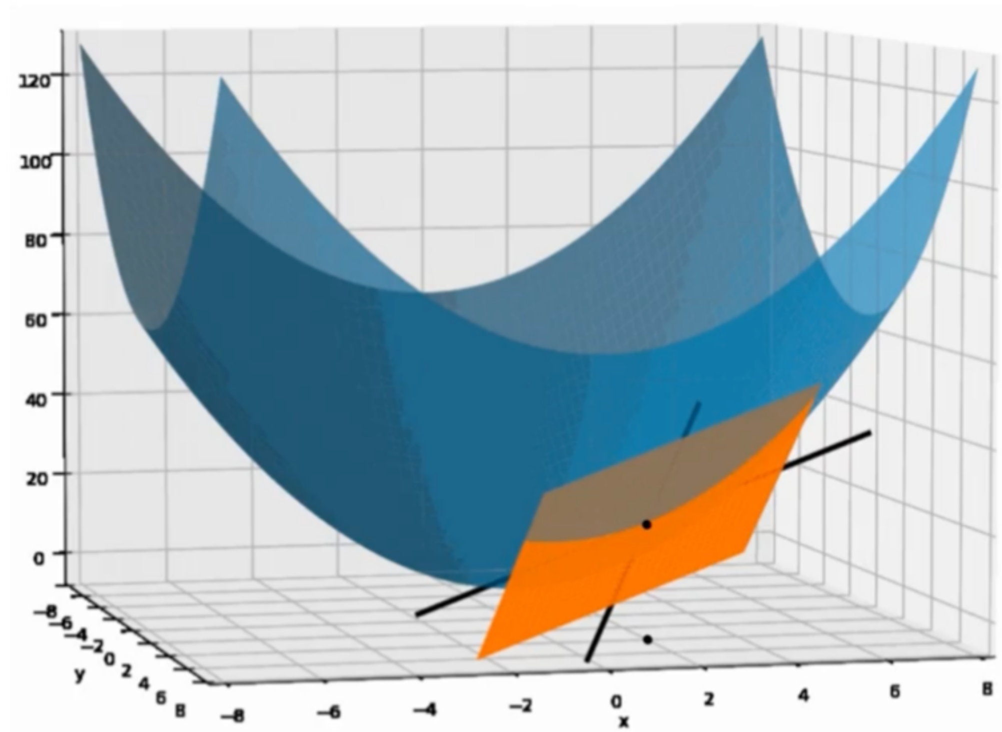
Task:

The gradient of $f(x,y)$ at $(2,3)$ is

$$\nabla f \Rightarrow \begin{bmatrix} 2\cdot 2 \\ 2\cdot 3 \end{bmatrix}$$

$$\boxed{\nabla f \Rightarrow \begin{bmatrix} 4 \\ 6 \end{bmatrix}}$$

Gradient



$$f(x,y) = x^2 + y^2$$

The gradient of $f(x,y)$ is: $\nabla f = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$

TASK

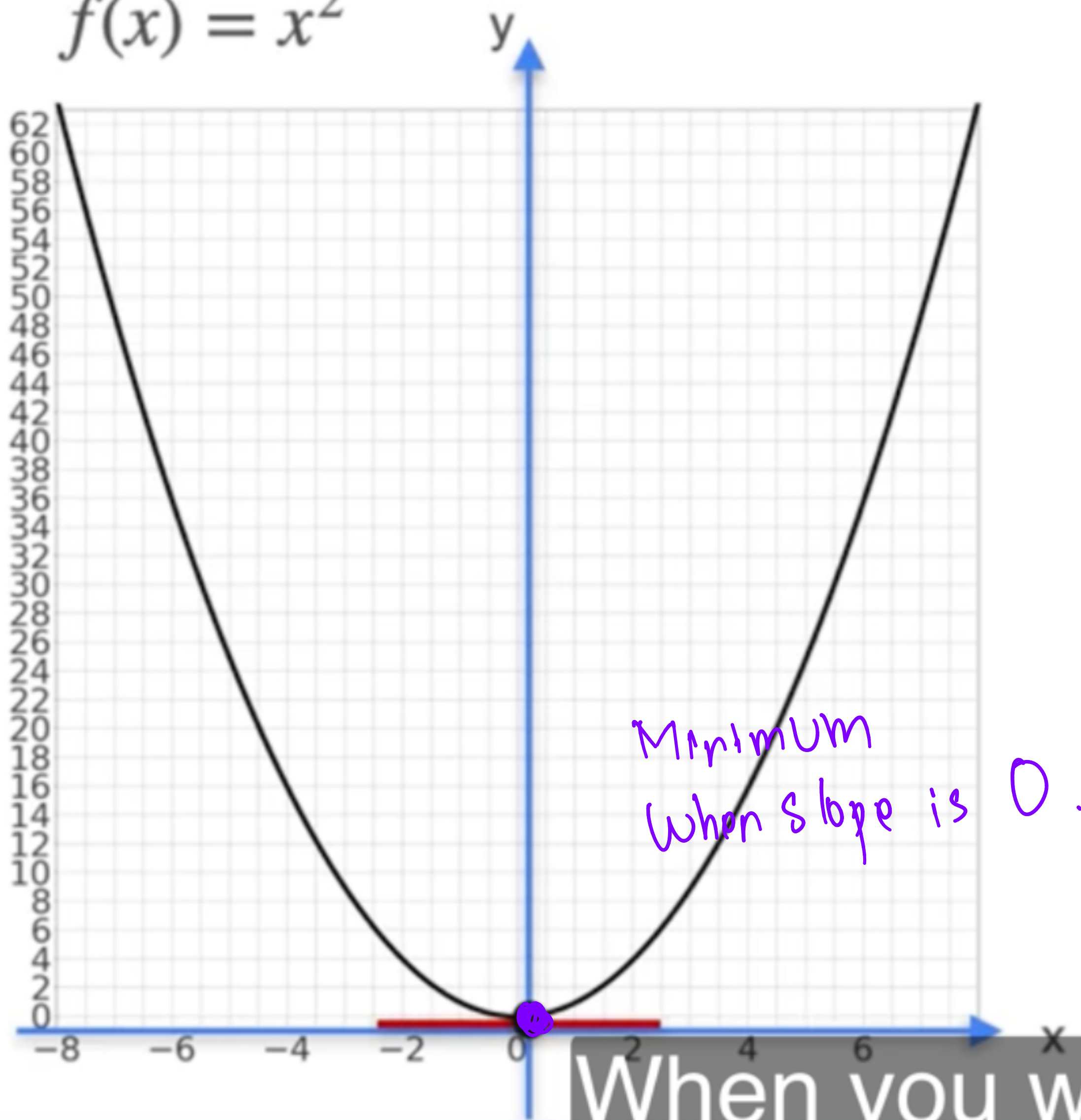
Find the gradient of $f(x,y)$ at $(2,3)$

DeepLearning.AI

So, if we have a function of many more variables, say a function of 12 variable, then we're just going to need for 12 slopes to be '0'. those corresponding to all partial derivatives.

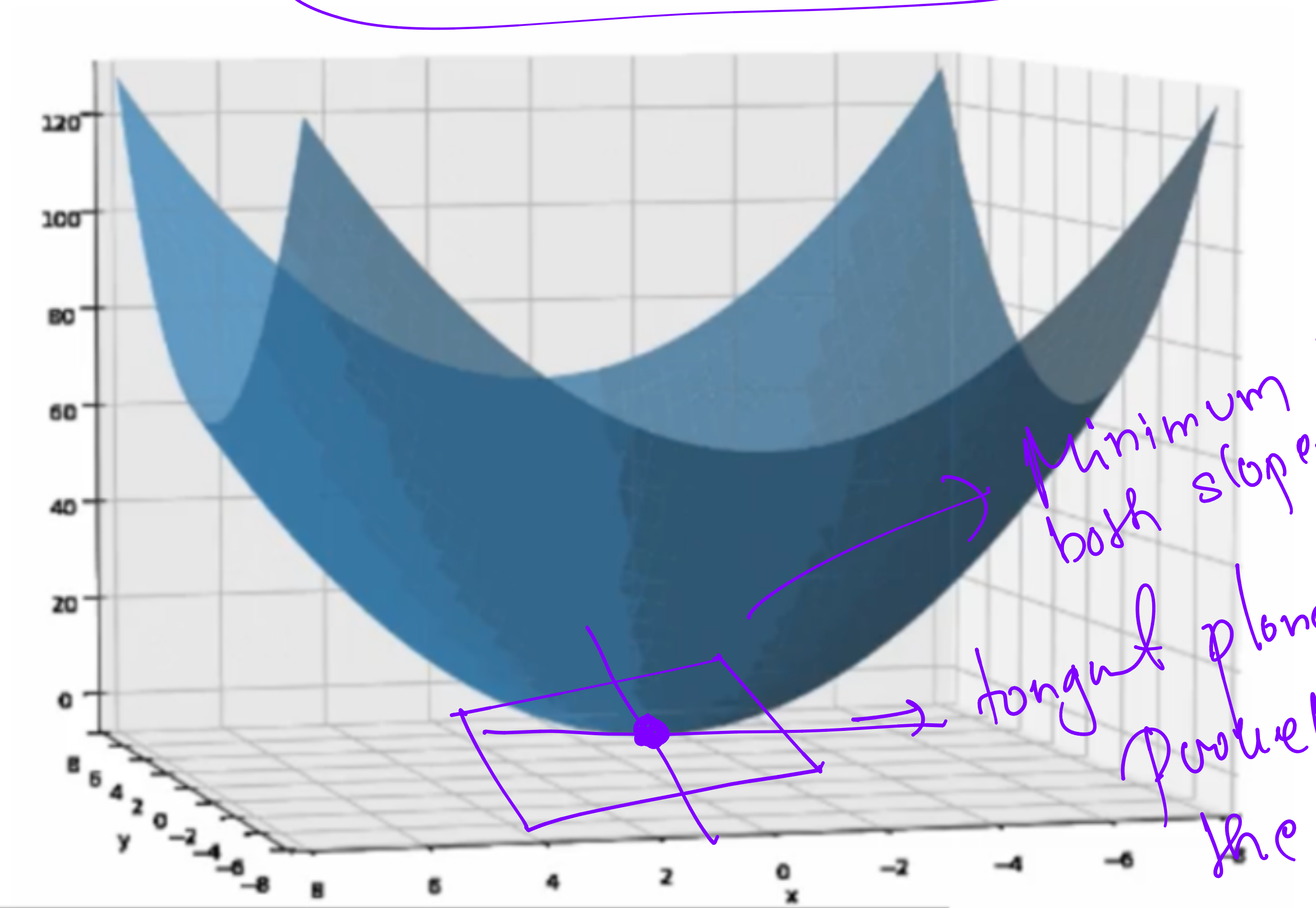
Functions of Two Variables

$f(x) = x^2$



Minimum when slope is 0.

$f(x, y) = x^2 + y^2$

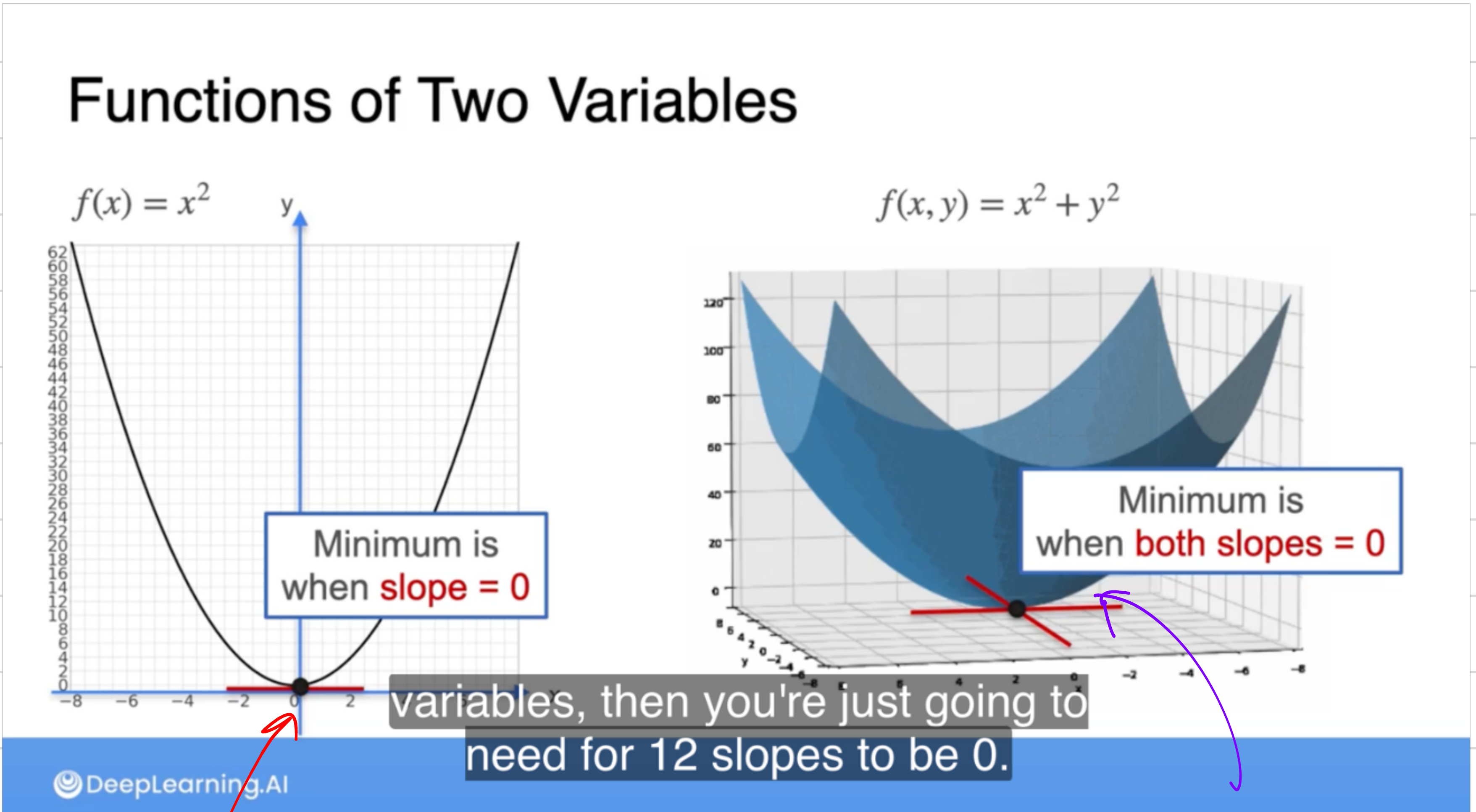


Minimum when both slopes is 0.

tangent plane is parallel to the floor

When you want to minimize this function, the minimum point is this point over here.

DeepLearning.AI



$f'(x) = 0$

$\partial x = 0$

$x = 0$.

$\frac{\partial f}{\partial x} = 0$ and $\frac{\partial f}{\partial y} = 0$

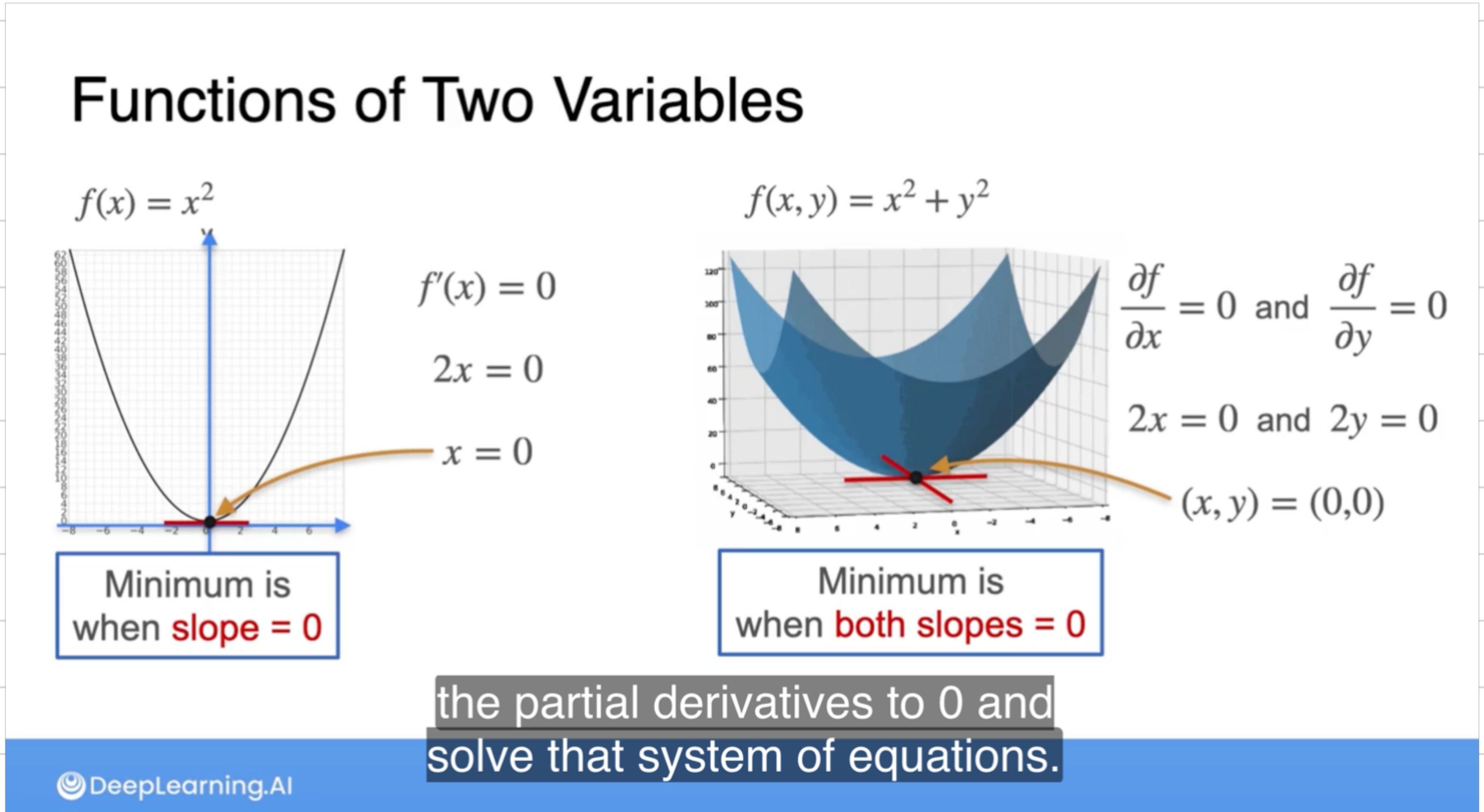
$\partial x = 0$ and $\partial y = 0$

$\boxed{(x, y) = (0, 0)}$

Now solving the system of equations, $\partial x = 0$ and $\partial y = 0$, was pretty.

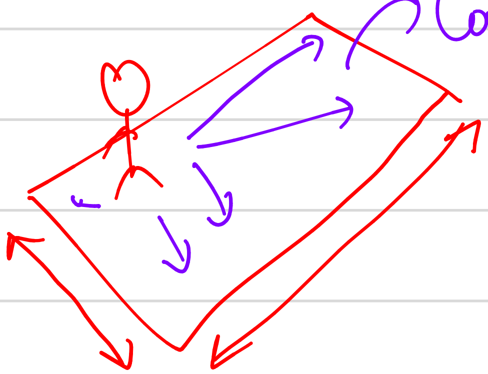
So, Basically to find the minimums and maximums all we have to do is set all the partial derivatives to 0 and solve the system of equations

So, for three variables it would be 3 slopes = 0.



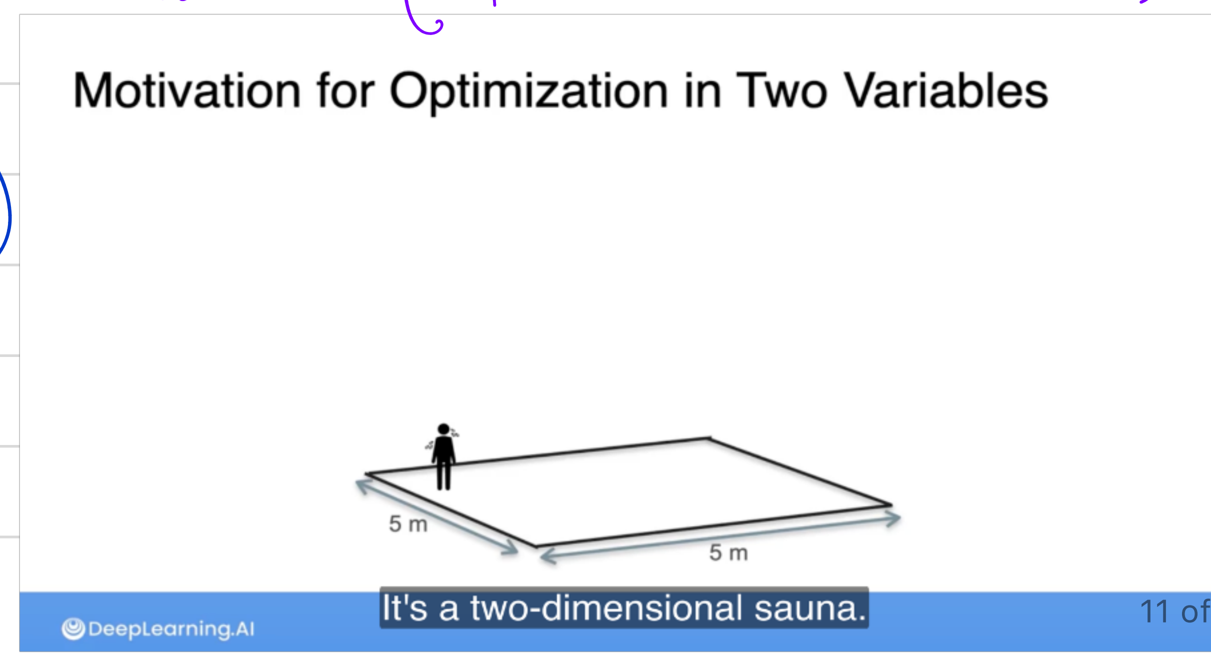
Motivation for Optimization in Two Variables

Can move in any direction inside a room (two-dimensional sauna).



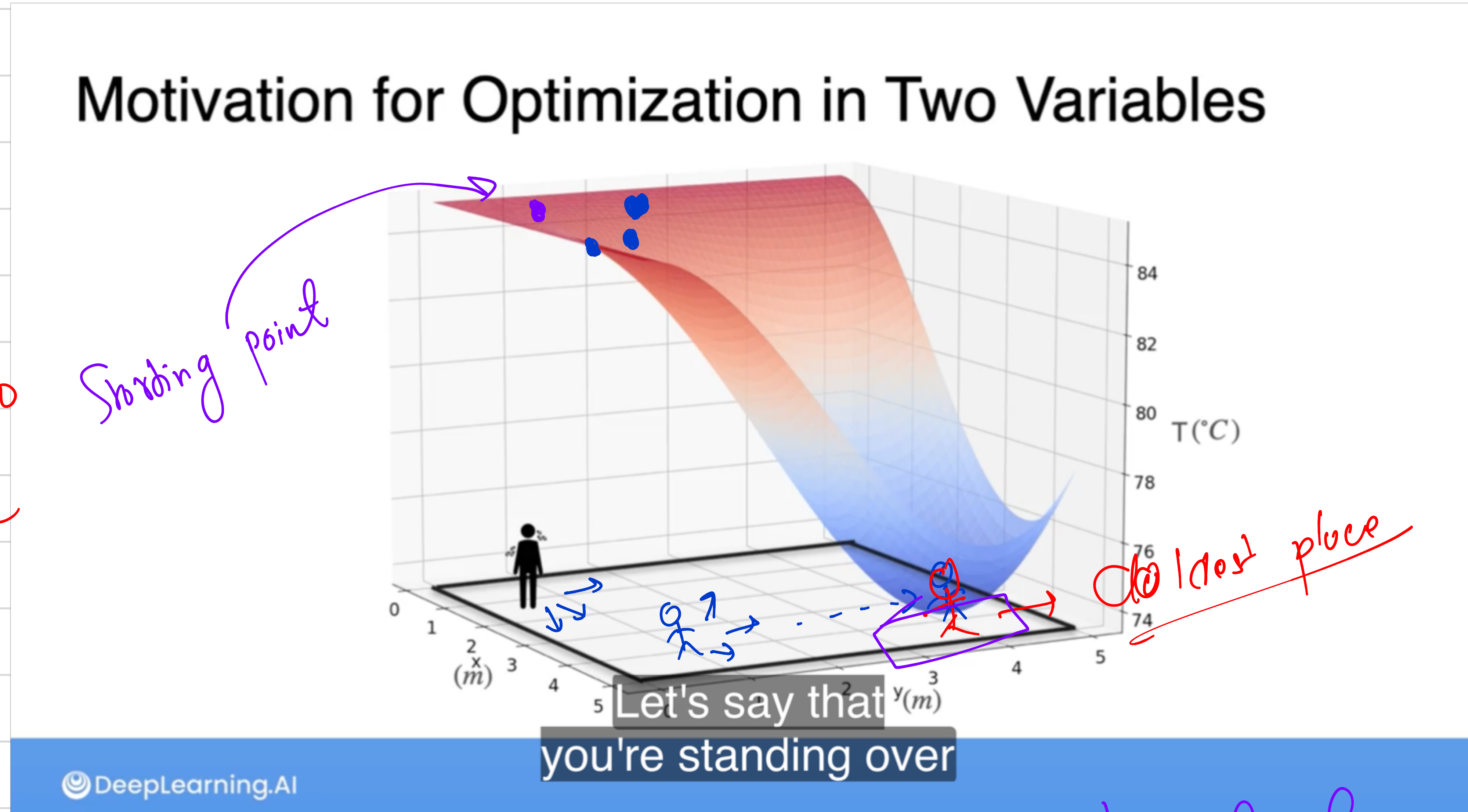
$$0.05 \leftrightarrow 0.05 + \alpha \cdot \frac{d}{dx} f(0.05)$$

ensure the 'steps' are small enough.



Red parts → Hotter
Cold parts → colder

So, if we move then how to know if we reached to the coldest place?

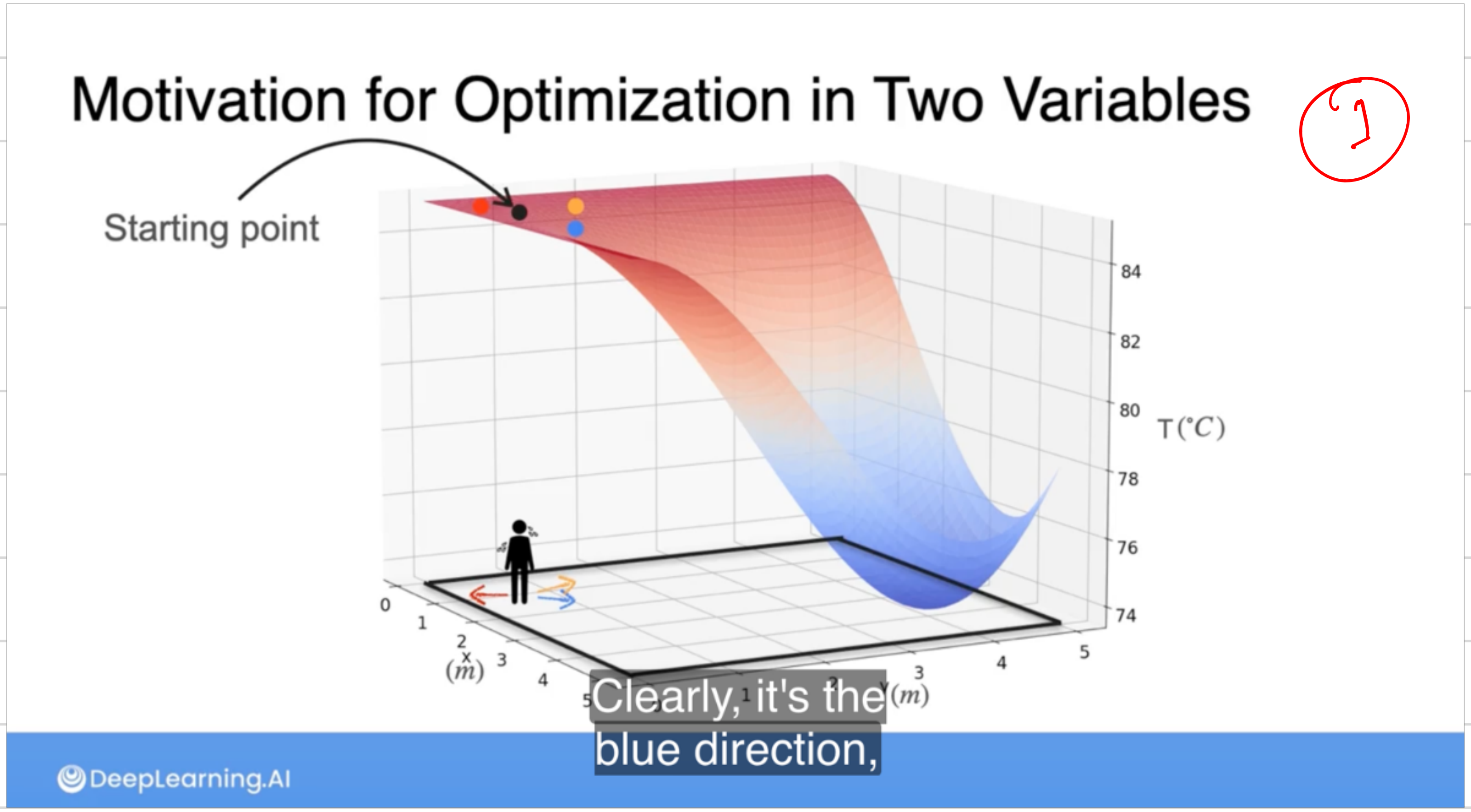


Answer: We have no more option to explore and feel hot nearby moving. OK if we take tangent the tangent plane is parallel to the ground.

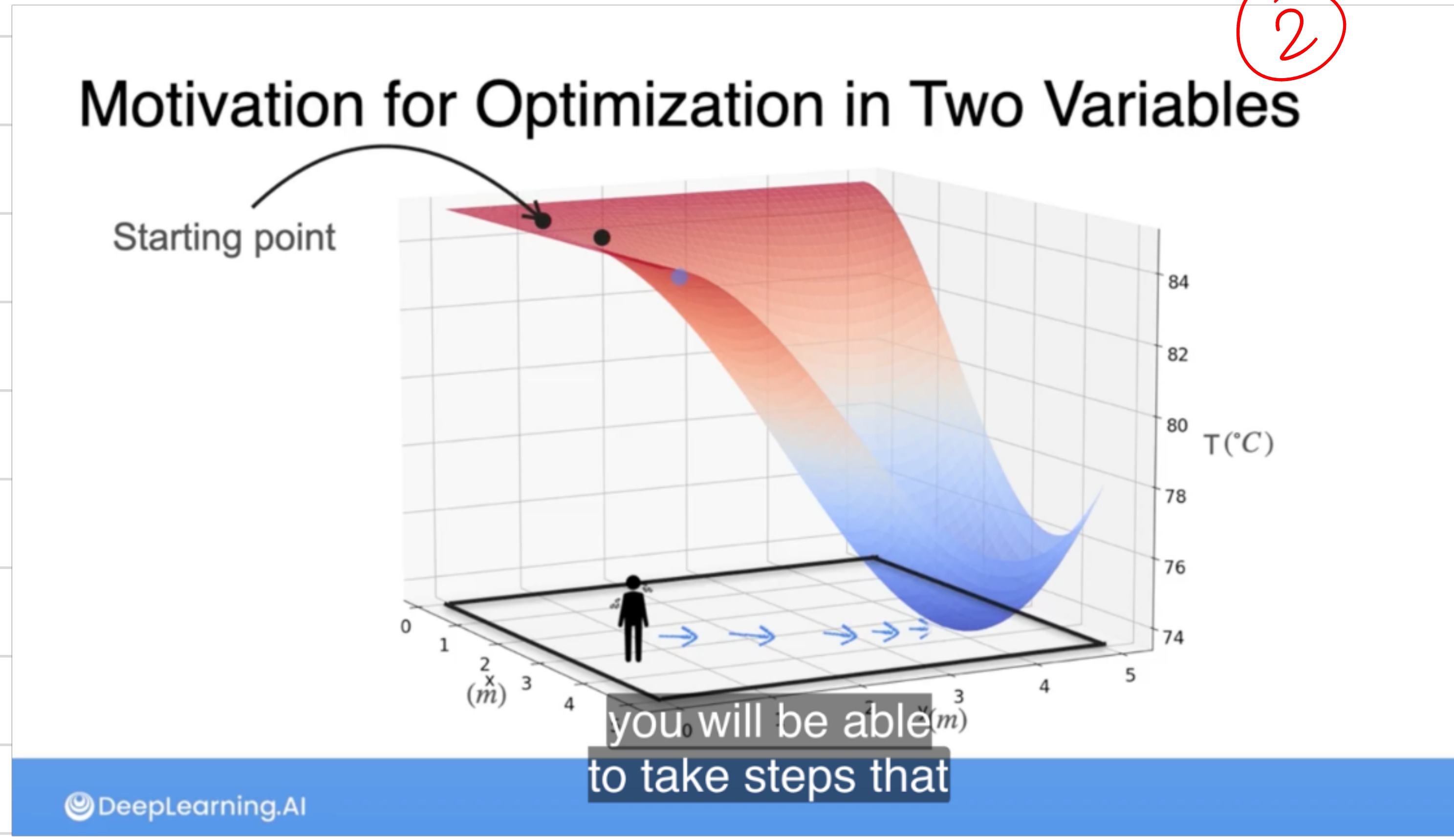
Mathematically, $T(^\circ C)$ $\frac{\partial T}{\partial x} = 0$

$$\frac{\partial T}{\partial y} = 0$$

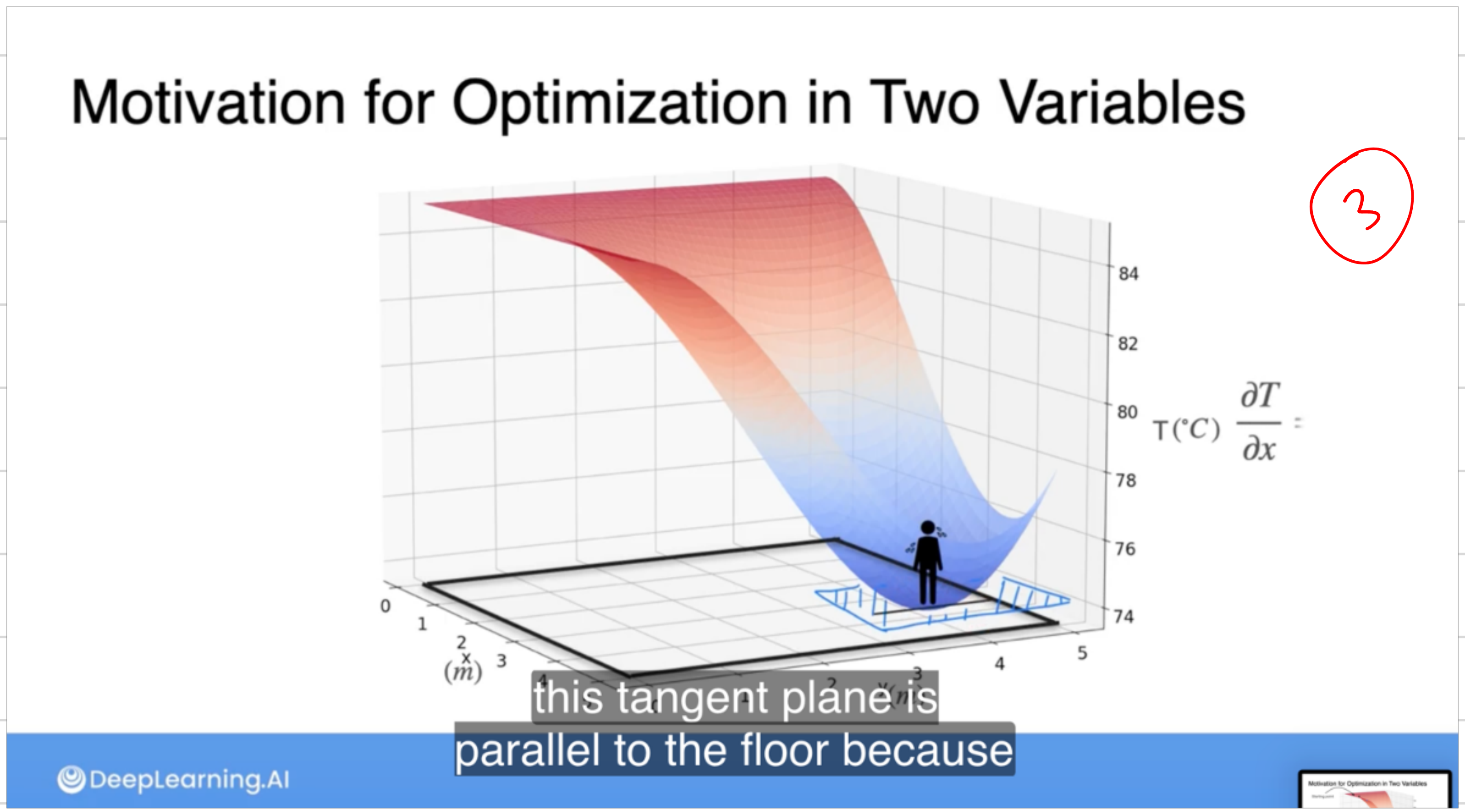
1



2



3



$\rightarrow \quad \dfrac{\partial T}{\partial x} = 0$

$\rightarrow \quad \dfrac{\partial T}{\partial y} = 0$

$$T = f(x,y) = 85 - \frac{1}{90} x^2(x-6)\, y^2(y-6)$$

$$\frac{\partial f}{\partial x} = 85 - \frac{1}{90} x^2(x-6)\,y^2(y-6)$$

y is constant –

$$\frac{\partial f}{\partial x} = -\frac{1}{90}\left[2x(x-6)\,y^2(y-6) + x^2y^2(y-6)\right]$$

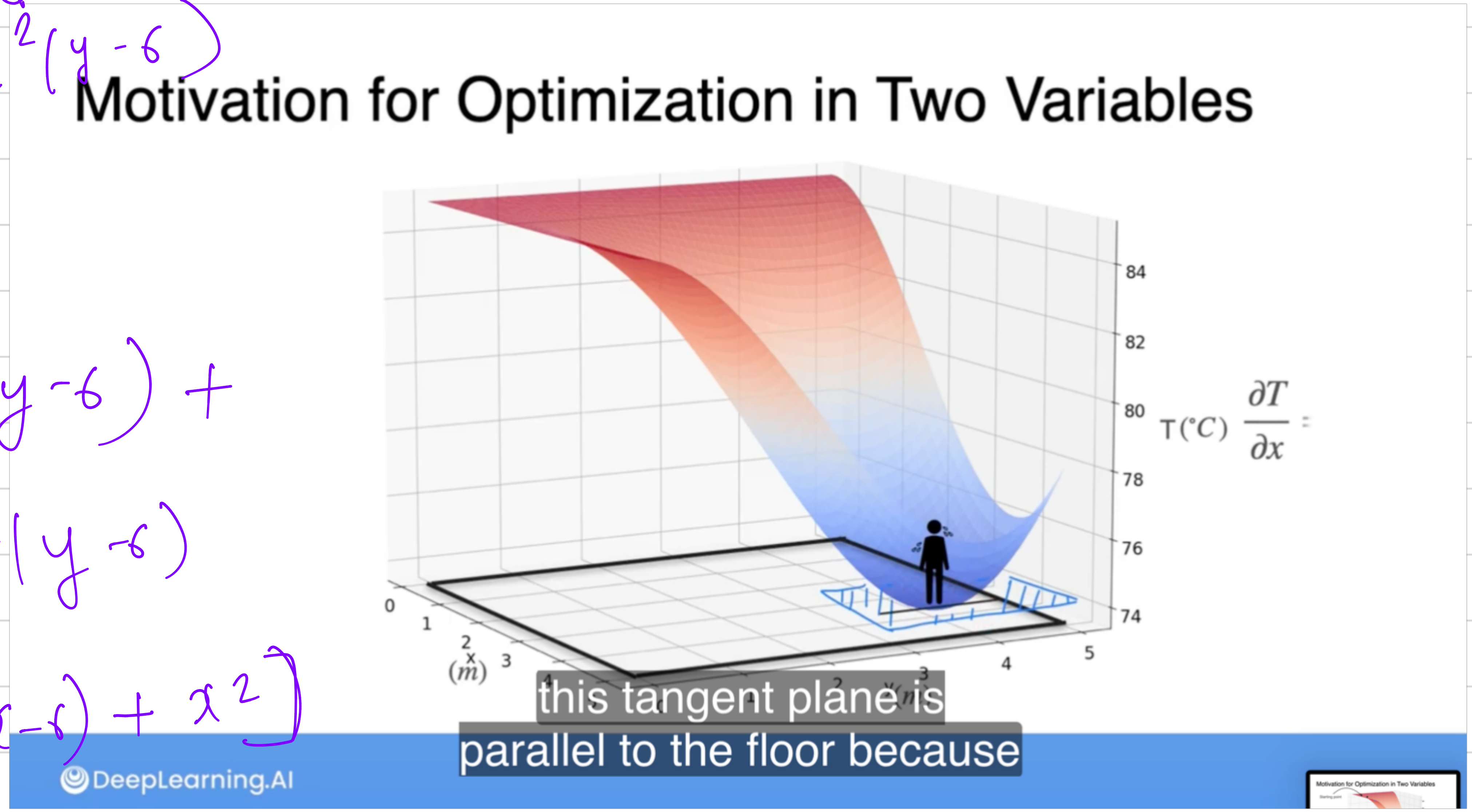
$$\Rightarrow -\frac{1}{90} y^2(y-6)\left[2x(x-6) + x^2\right]$$

$$\Rightarrow -\frac{1}{90} y^2(y-6)\left[2x^2 - 12x + x^2\right]$$

$$\Rightarrow -\frac{1}{90} y^2(y-6)\left[3x^2 - 12x\right]$$

$$\Rightarrow -\frac{1}{90} y^2(y-6)\,3x(x-4)$$

$$\boxed{\frac{\partial f}{\partial x} = -\frac{1}{30} y^2(y-6)\,x(x-4)}$$



$$\frac{\partial f}{\partial y} = 85 - \frac{1}{90} x^2 (x-6)(y-6)y^2$$

$$\Rightarrow -\frac{1}{90} x^2 (x-6) \left[2y(y-6) + y^2 \right]$$

$$\Rightarrow -\frac{1}{90} x^2 (x-6) \left[2y^2 - 16y + y^2 \right] \Rightarrow -\frac{1}{30} y (y-4) x^2 (x-6)$$

Because $g(y) = y^2 (y-6)$

$$\frac{d}{dy}\left[y^2 (y-6)\right] = y^2 \cdot \frac{d}{dy}(y-6) + (y-6) \cdot \frac{d}{dy}(y^2)$$

$$\Rightarrow y^2 (1) + (y-6)(2y)$$

$$\left[\frac{d}{dy}\left[y^2 (y-6)\right] \Rightarrow y^2 + 2y^2 - 12y \Rightarrow 3y^2 - 12y \right]$$

$\boxed{x = 0}$ → eqn (i)

$3x - 12 = 0$ — eqn (ii)

$\boxed{x = 4}$

$y - 6 = 0$

$\boxed{y = 6}$ → eqn (iii)

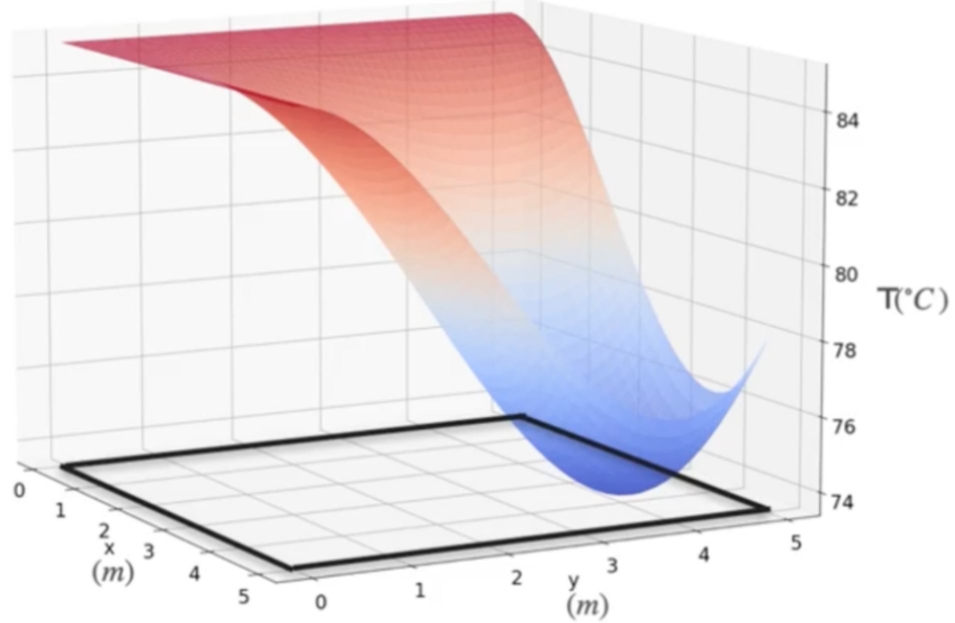
$x^2 = 0 \qquad \boxed{x = 0}$

$x - 6 \Rightarrow 0 \qquad \boxed{x = 6}$

$3y - 12 = 0 \qquad \boxed{y = 4}$

Motivation for Optimization in Two Variables

$$T = f(x, y) = 85 - \frac{1}{90}x^2(x-6)y^2(y-6)$$



$$\frac{\partial f}{\partial x} = -\frac{1}{90}\underbrace{x}_{x=0}\underbrace{(3x-12)}_{x=4}\underbrace{y^2}_{y=0}\underbrace{(y-6)}_{y=6} = 0$$

eqn (i), (ii), (iii)

$$\frac{\partial f}{\partial y} = -\frac{1}{90}\underbrace{x^2}_{x=0}\underbrace{(x-6)}_{x=6}\underbrace{y}_{y=0}\underbrace{(3y-12)}_{y=4} = 0$$

(i), (ii), (iii)

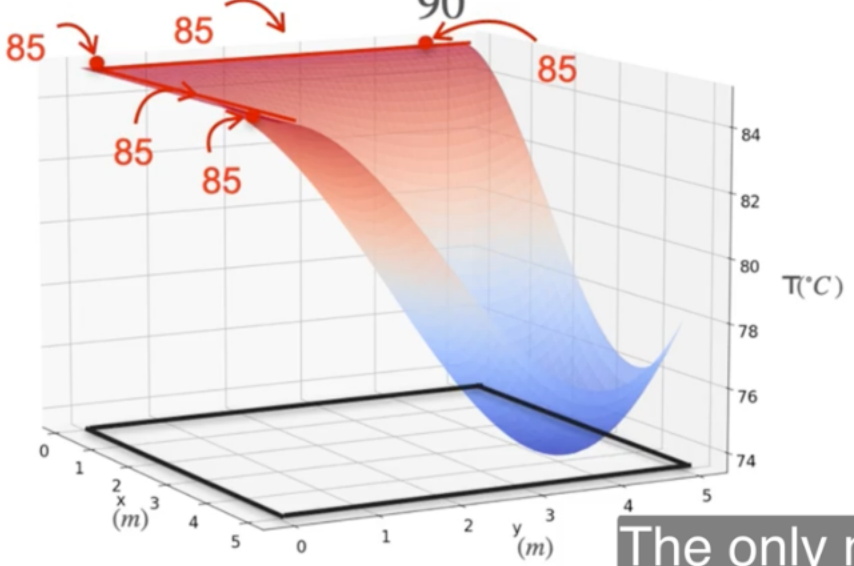
in which case y equals 4.

DeepLearning.AI

These are all the candidates points for the minima, then we have one point where the gradient is 0

Motivation for Optimization in Two Variables

$$T = f(x, y) = 85 - \frac{1}{90}x^2(x-6)y^2(y-6)$$



Candidate points for the minima

$x = 0$, $y = 0$	Maxima
$x = 0, y = 0$; $x = 0, y = 4$	Maxima
$x = 0, y = 6$	Outside
$x = 4, y = 0$	Maxima
$x = 4, y = 4$	
$x = 6, y = 0$; $= 6$	Outside

The only minimum is this one over here,

DeepLearning.AI

So we obtained

$$\frac{\partial f}{\partial x} = -\frac{1}{90} x(3x-12)\, y^2(y-6) = 0$$

put $x = 4$ and $y = 4$

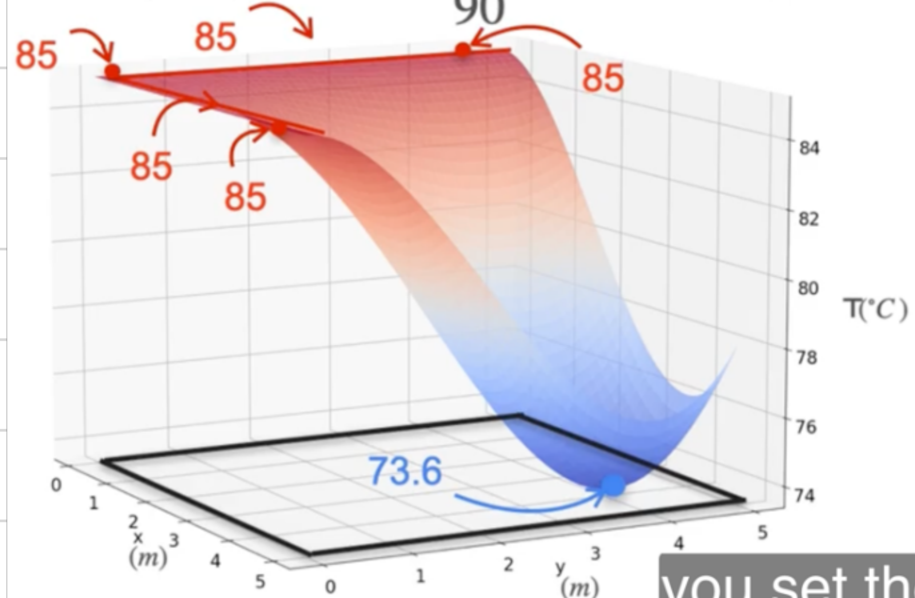
in $f(x,y)$ we get minimum —

At $x = 4$, $y = 4$ we have the minimum point which $73.6°C$.

$$\frac{\partial f}{\partial y} = -\frac{1}{90} x^2(x-6)\, y(3y-12)$$

Motivation for Optimization in Two Variables

$$T = f(x, y) = 85 - \frac{1}{90}x^2(x-6)y^2(y-6)$$



Candidate points for the minima

$x = 0$, $y = 0$	Maxima
$x = 0, y = 0$; $x = 0, y = 4$	Maxima
$x = 0, y = 6$	Outside
$x = 4, y = 0$	Maxima
$x = 4, y = 4$	Minimum
$x = 6, y = 0$; $= 6$	Outside

you set the derivative equal to zero,

DeepLearning.AI

Linear Regression: Analytical Approach:

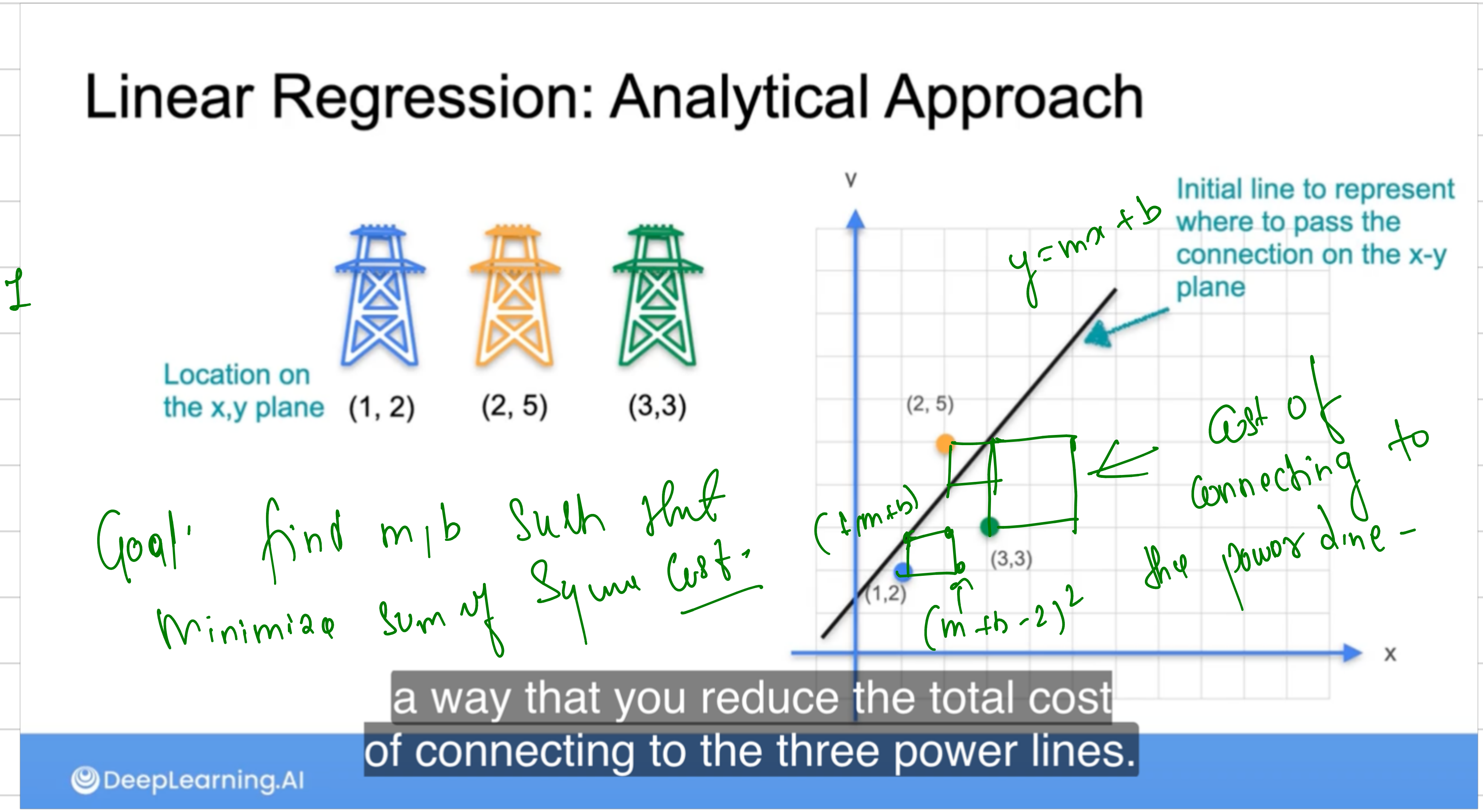
$$(x_1, y_1) = (1, 2)$$

$$y = mx + C \quad \because x = 1$$

$$y = m + C$$

$$2 = m + C$$

$$\text{or, } (m + C - 2)^2 = 0$$



$$(2, 5) = (x_2, y_2)$$

$$y = mx + C$$

$$\text{or, } 5 = 2x + C$$

$$\text{or, } (2x + C - 5)^2 = 0$$

$$(3, 3) = (x_3, y_3)$$

$$y = mx + C$$

$$\text{or, } 3 = 3m + C$$

$$\text{or, } (3m + C - 3)^2 = 0$$

Sum of square cost — minimize.

$(m+b-2)^2 \Rightarrow$

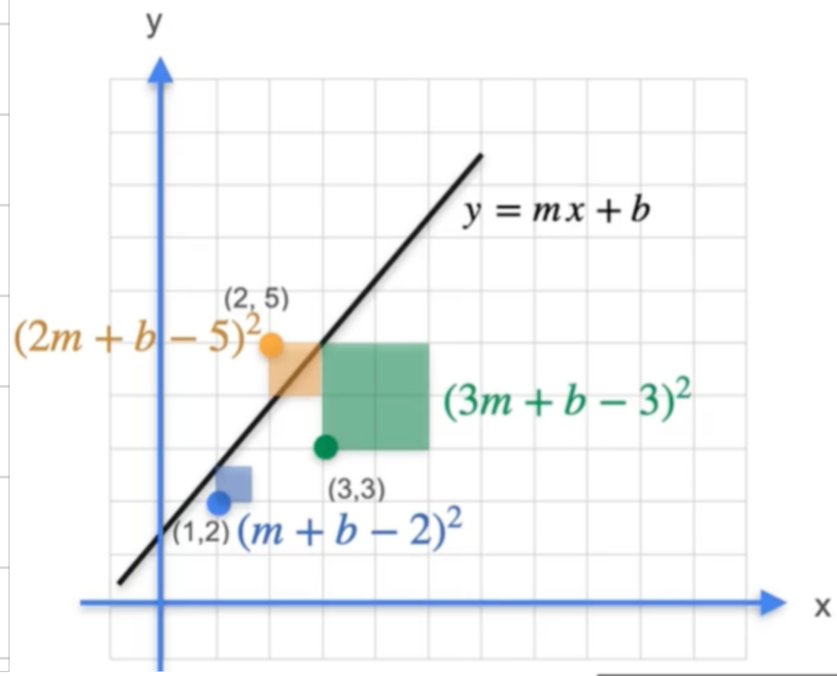
$(2m+b-5)^2$

$(3m+b-3)^2$

$(m+b+2)(m+b-2)$

$m(m+b-2)+b(m+b-2)+2$ $(m+b-$

Linear Regression: Analytical Approach

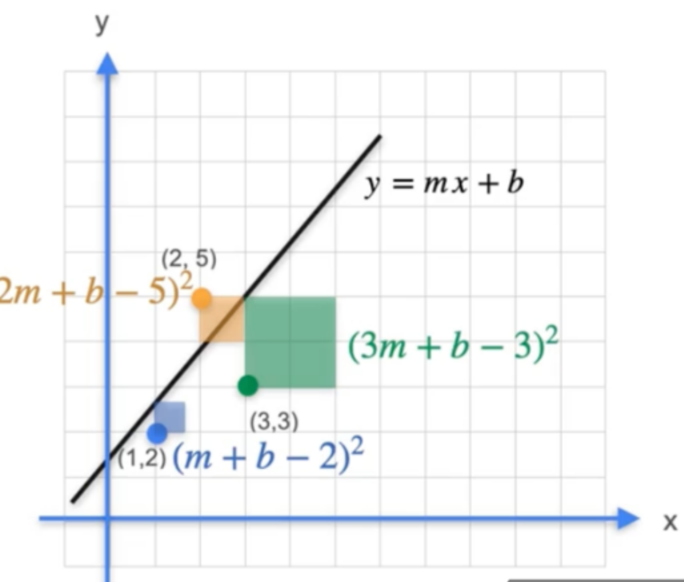


Goal: Minimize sum of squares cost

$$(m+b-2)^2 + (2m+b-5)^2 + (3m+b-3)^2$$

And therefore, the total cost is
he sum of these three numbers.

Linear Regression: Analytical Approach



Goal: Minimize sum of squares cost

$$(m+b-2)^2 + (2m+b-5)^2 + (3m+b-3)^2$$

$$m^2 + b^2 + 4 + 2mb - 4m - 4b$$

$$+4m^2 + b^2 + 25 + 4mb - 20m - 10b$$

$$+9m^2 + b^2 + 9 + 6mb - 18m - 6b$$

$$E(m,b) = 14m^2 + 3b^2 + 38 + 12mb - 42m - 20b$$

And if we join similar terms,
we get that E of m,b is 14m squared plus

DeepLearning.AI

$\frac{\partial E}{\partial m} = 0 \Rightarrow 28m + 12b - 42 = 0$

$\Rightarrow$

$\frac{\partial E}{\partial b} = 0 \Rightarrow 18b + 12m - 20 = 0$

$$\frac{\partial E}{\partial m} = 28m + 12b - 42 = 0$$

$$\frac{\partial E}{\partial b} = 6b + 12m - 20 = 0$$

$$(6b + 12m - 20) \times 2 = 0$$

$$12b + 24m - 40 = 0$$

$$-\ 12b + 28m - 42 = 0$$

$$4m - 2 = 0$$

$$\boxed{m = 0.5}$$

$$6b + 12(0.5) - 20 = 0$$

$$6b + 6 - 20 = 0$$

$$6b - 14 = 0$$

$$b = \frac{14}{6} = 7/3$$

$$\boxed{b = 7/3}$$

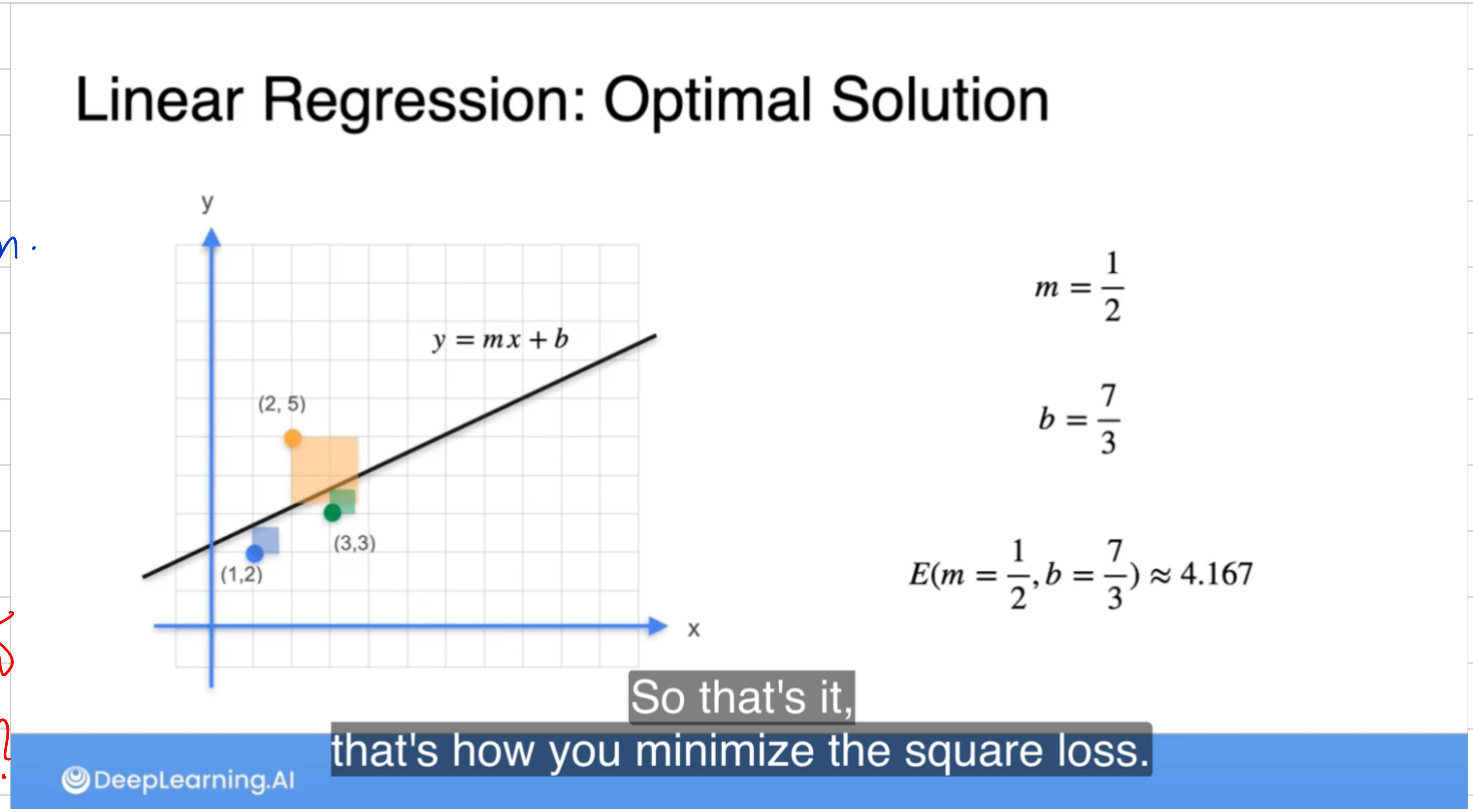
$$E\left(m = 1/2,\ b = 7/3\right) \approx 4.167$$

Example: $f(x, y, z) = x^2 + 2xyz + z^2$

$$\begin{bmatrix} 2x + 2yz \\ 2yz \\ 2xy + 2z \end{bmatrix} \begin{array}{l} \partial f/\partial x \\ \partial f/\partial y \\ \frac{\partial f}{\partial z} \end{array}$$

This problem is similar to Linear Regression.

Solving m's and b's for very large equations is expensive and slow. So, if there is faster way to do that?



Yes, Gradient Descent minimizes the cost functions effectively.

Examples ① $f(x,y) = x^2 y + 3x^2$

$$\frac{\partial f}{\partial x} \Rightarrow 2xy + 6x$$

$$\frac{\partial f}{\partial y} \Rightarrow y$$

② $f(x,y) = xy^2 + 2x + 3y$ $\therefore$ $\nabla f(x,y) = ?$

$$\nabla f(x,y) \Rightarrow \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \Rightarrow \begin{bmatrix} y^2 + 2 \\ 2xy + 3 \end{bmatrix}$$

③ Let $f(x,y) = x^2 + 2y^2 + 8y$. The min of f is:

$$\frac{\partial f}{\partial x} \Rightarrow 2x$$

$$\frac{\partial f}{\partial y} \Rightarrow 4y + 8$$

$\nabla f(x,y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$ Such that $\nabla f(x,y) = (0,0)$

Optimization using Gradient Descent in One Variable - Part 1.

$$f(x) = e^{x} - \log(x)$$

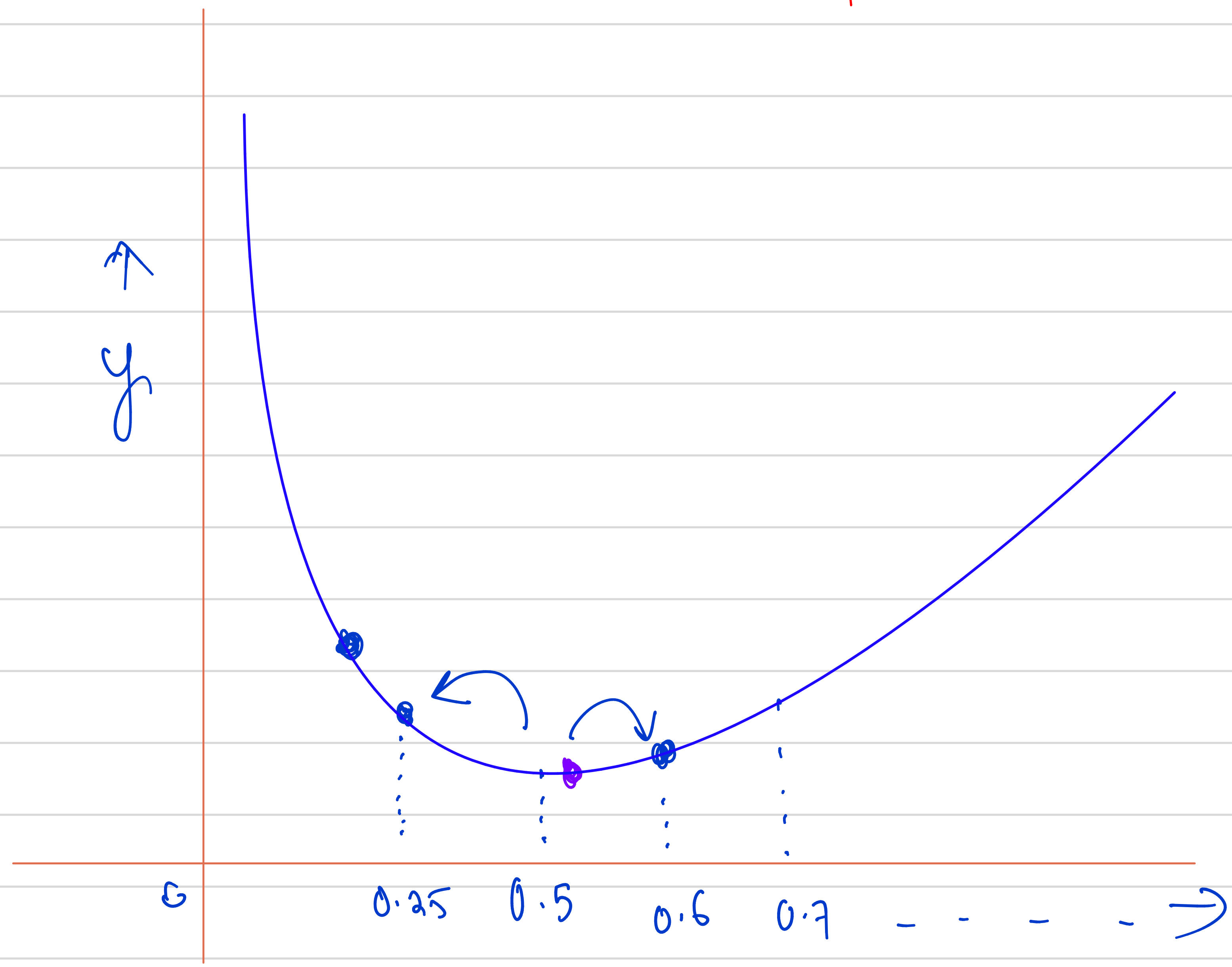
Minimum ?

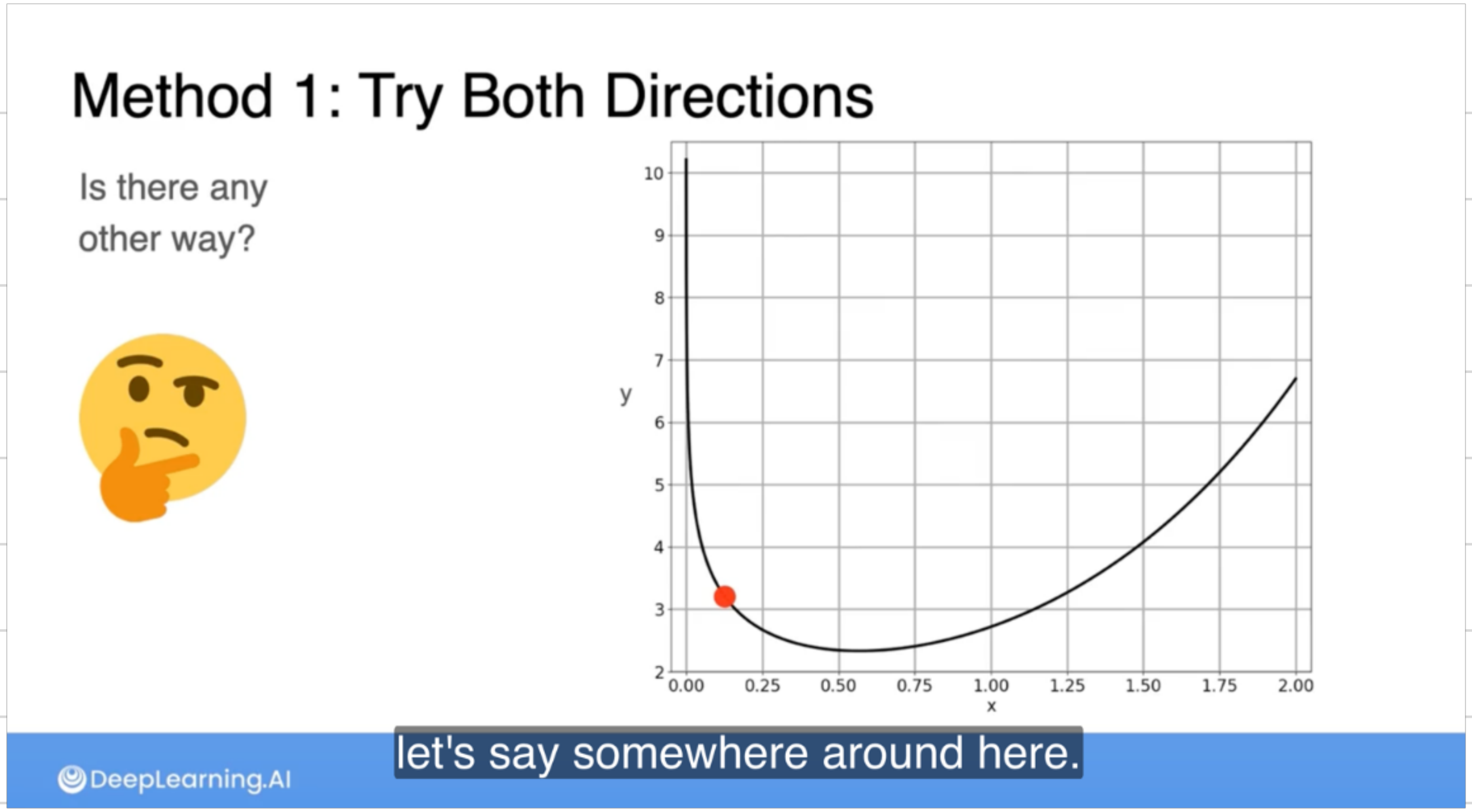
$$f'(x) = e^{x} - \frac{1}{x}$$

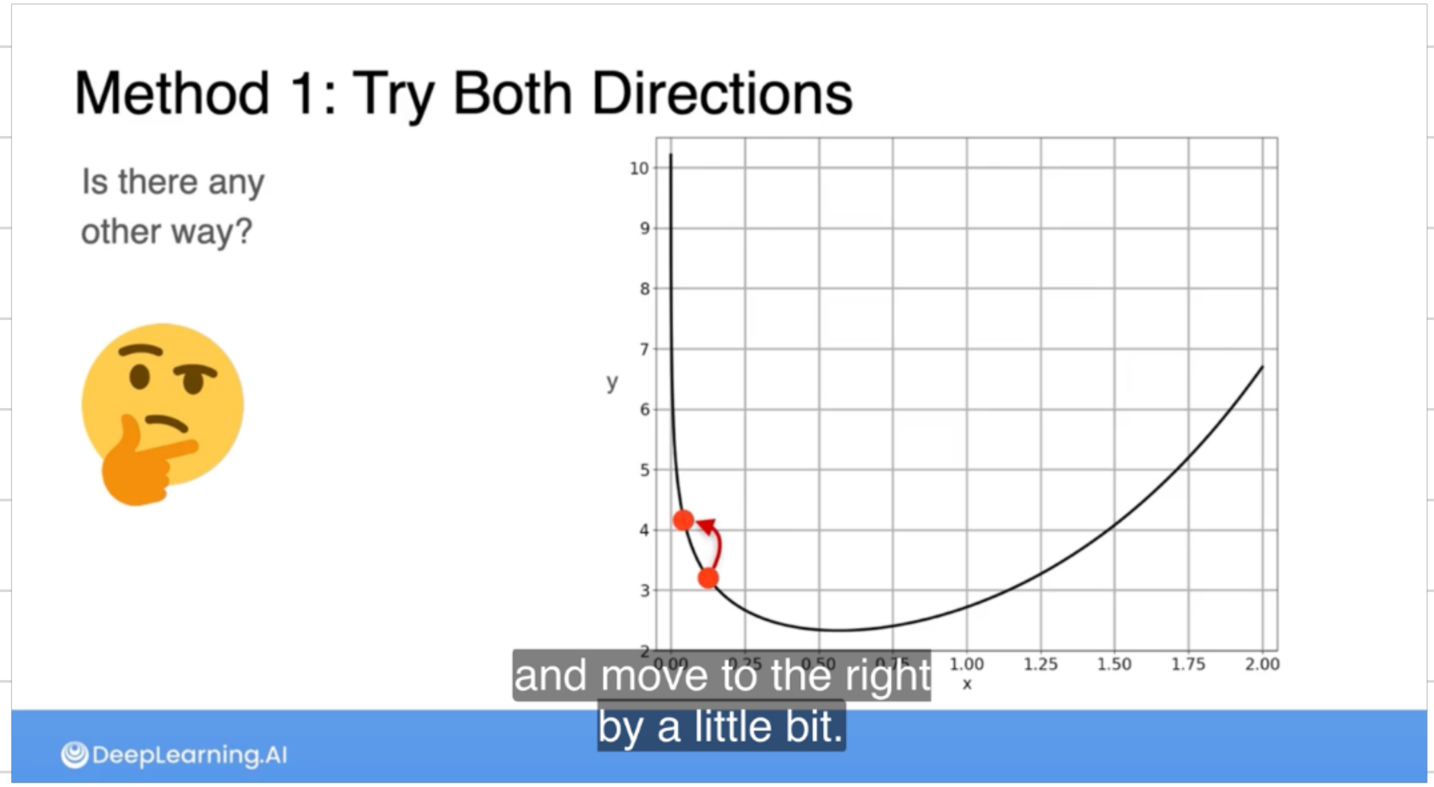
$$e^{x} - \frac{1}{x} = 0$$

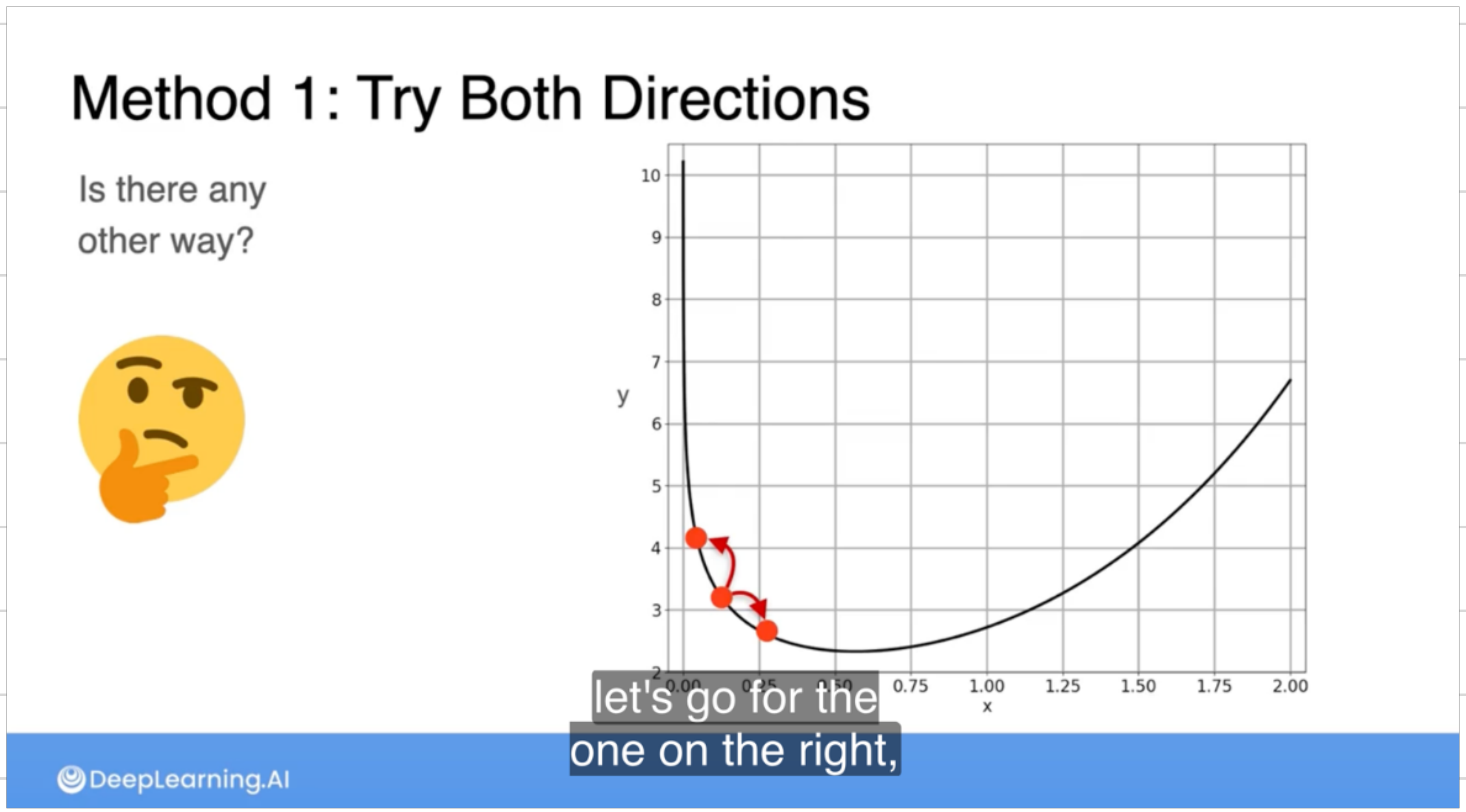
$$\Rightarrow \quad e^{x} = 1/x$$

$$x = 0.5671 \quad (\text{Omega Constant})$$

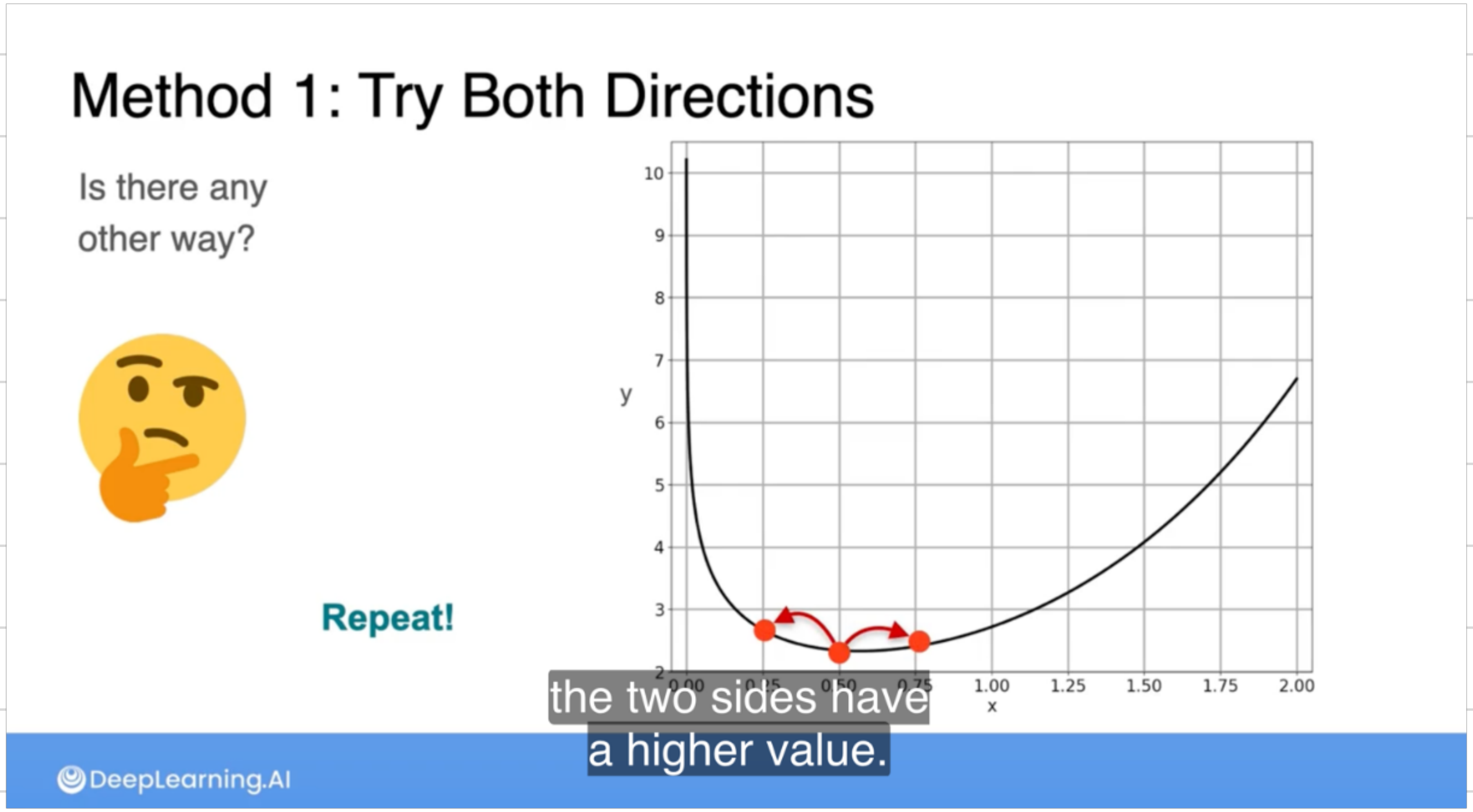


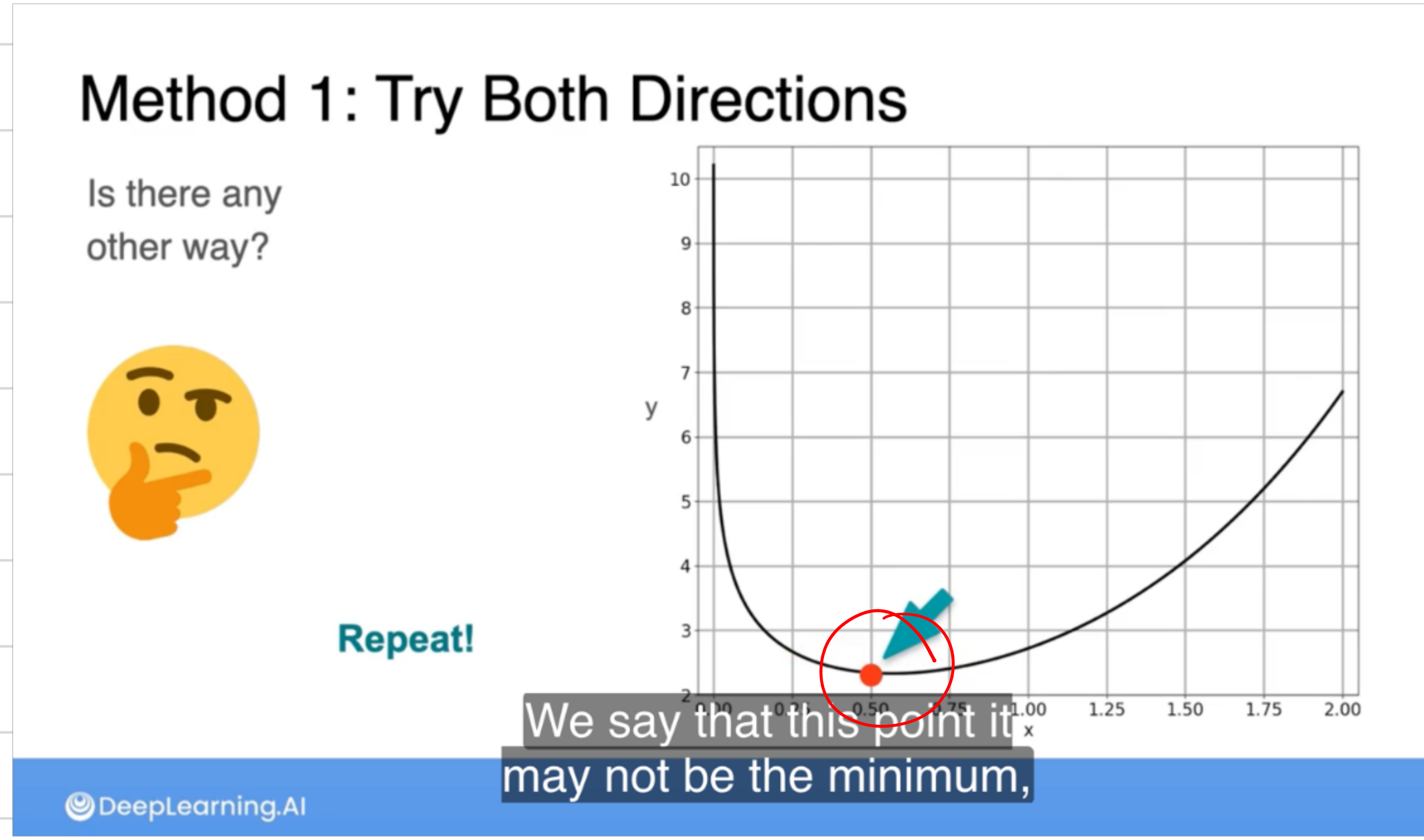

Method 1: Try Both Directions
Is there any other way?
DeepLearning.AI
let's say somewhere around here.


Method 1: Try Both Directions
Is there any other way?
DeepLearning.AI
and move to the right by a little bit.


Method 1: Try Both Directions
Is there any other way?
DeepLearning.AI
let's go for the one on the right,


Method 1: Try Both Directions
Is there any other way?
Repeat!
DeepLearning.AI
and this one wins.





Sources: Calculus for Machine Learning and Data Science Coursera Course.