



مجموعة جامعة
الجميع الذكية
للتعليم القابضة

MIDOCEAN
UNIVERSITY
BY EVERYONE'S SMART UNIVERSITY

Mathematical Foundations For Data Science and AI

Lecture 03

AI102

Calculus; Derivatives and Gradient Optimization

By
Dr. Khaled ELMenshawy

www.midocean.ae

Lecture

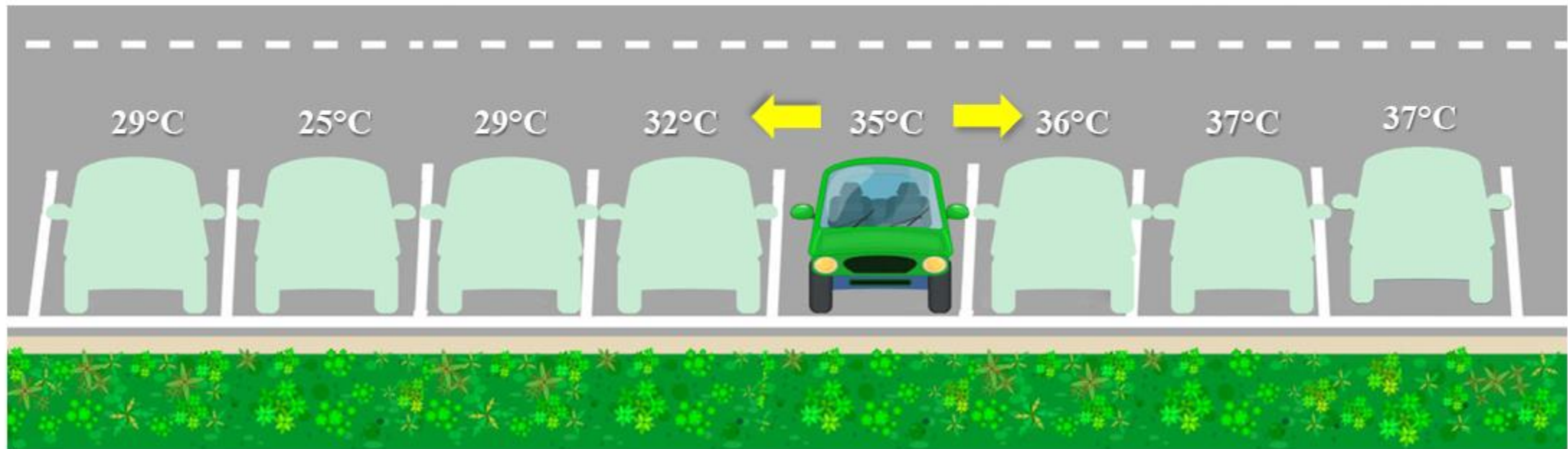
03

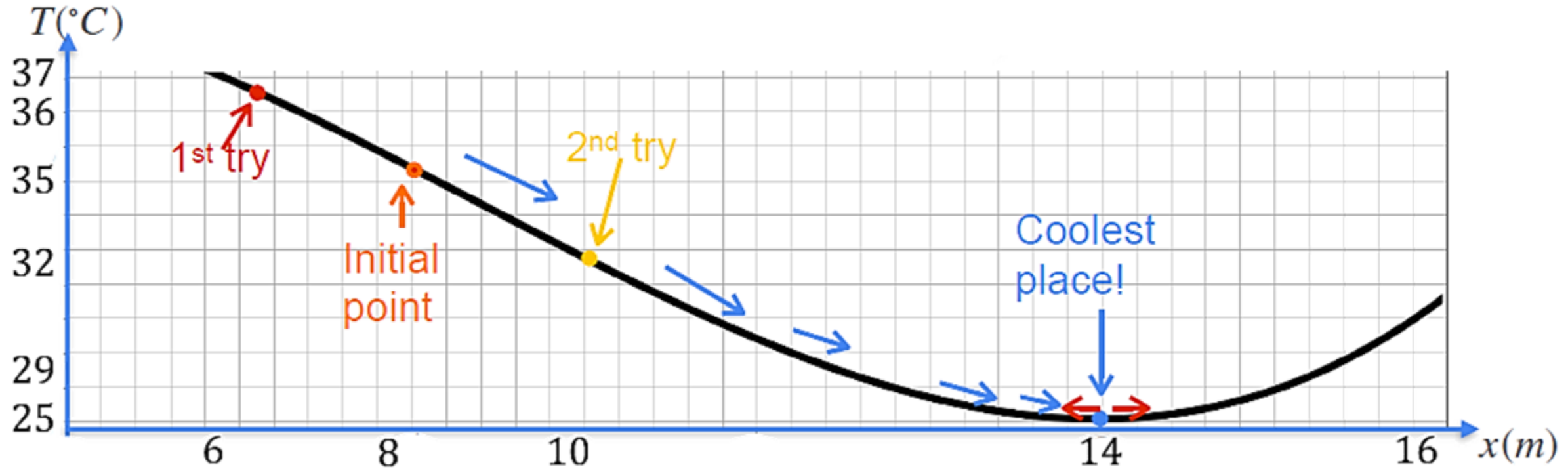
LEARNING OUTCOMES

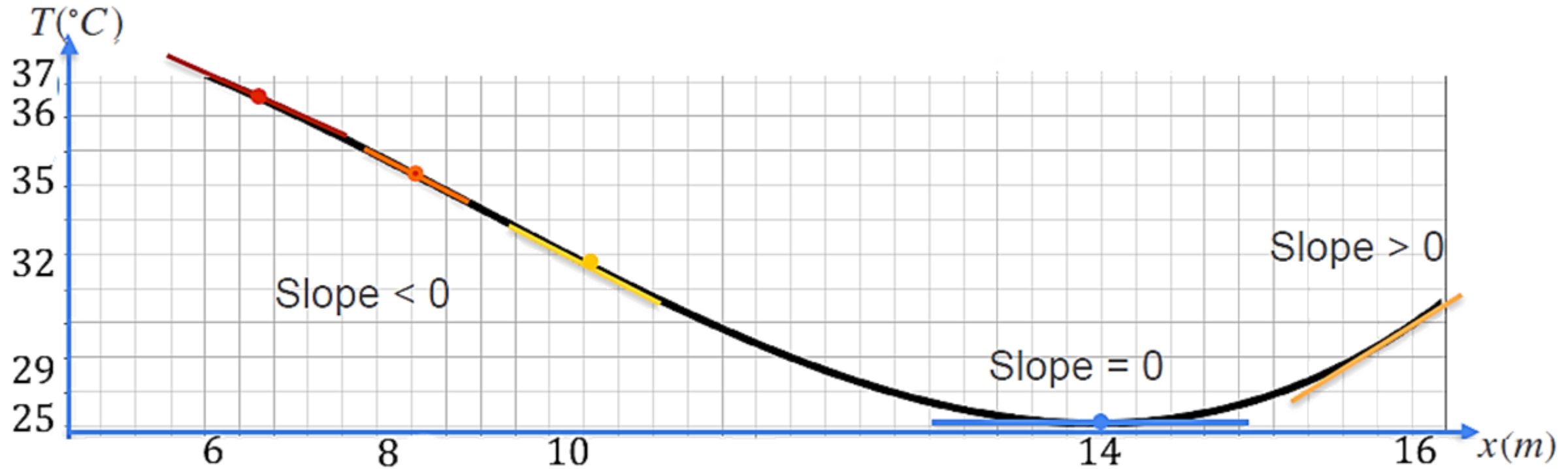
- Understanding Optimization.
- Understanding Square Loss function.
- Understanding Partial Derivatives.
- Understanding Function of two variables.
- Understanding Gradients.

What is Optimization?

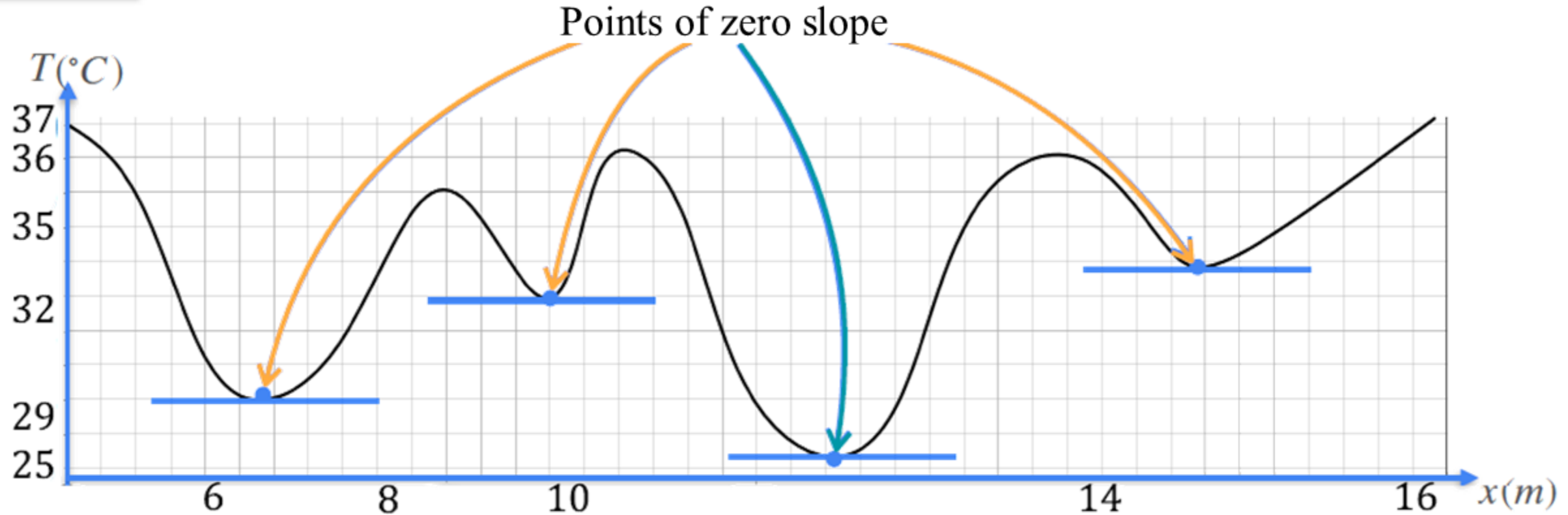
Optimization is finding the maximum or the minimum value of a function. This is very important in data science to find the model that best fits the used dataset, finding this model require calculating an error function that tells how far is the current model from an ideal one. When minimizing this error function then we have the best model.

Finding coldest spot for parking a car





Let's look at the slope of the tangent at all these points. Notice that, the colder spot is where the slope is zero.



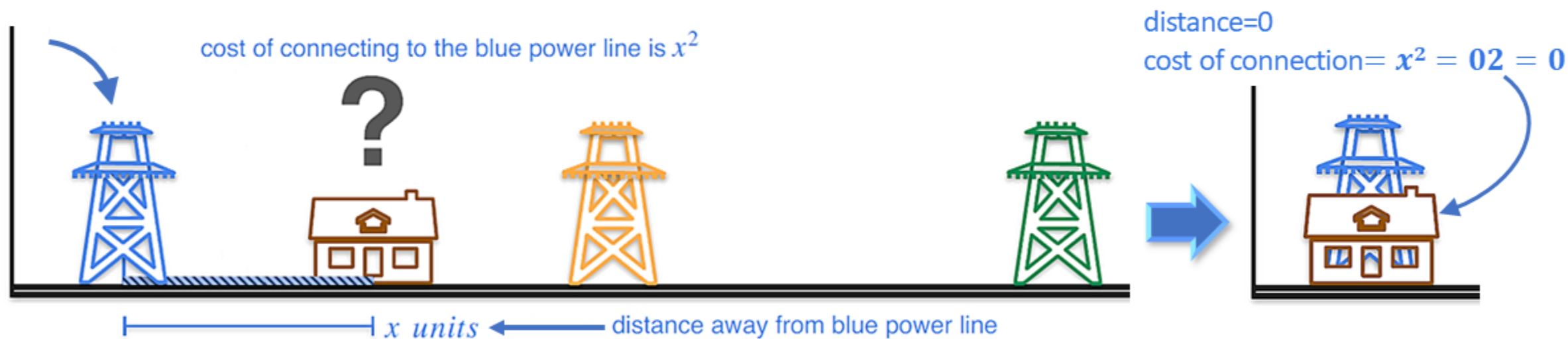
These points for the minimum are the points of zero slope and those are called **local minima**, and the absolute minimum is called the **global minimum**.

Locating a house in the perfect place, in order to minimize the cost of connecting this house to several power lines.

Why is this example important? Because the function you're minimizing is very similar to the squared error which is one of the most important functions used to train linear regression problems, and certain neural networks.

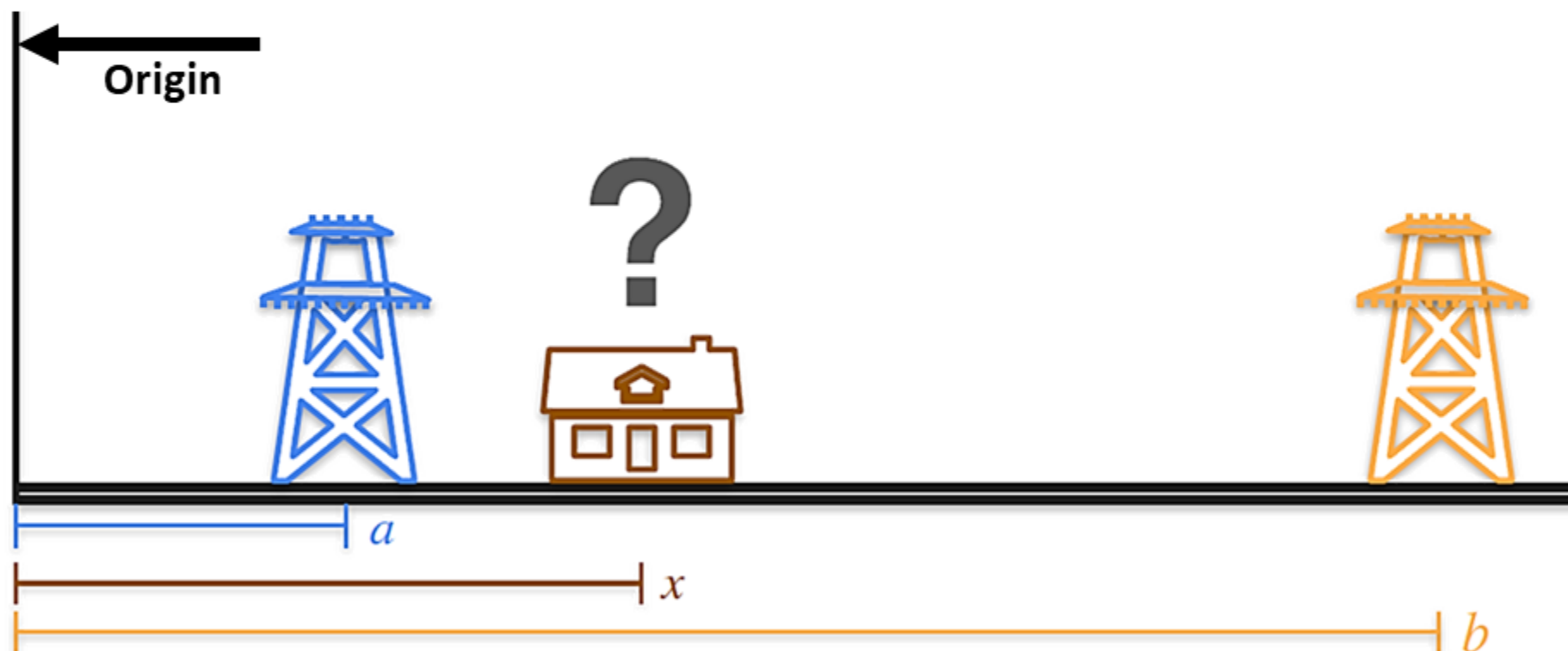


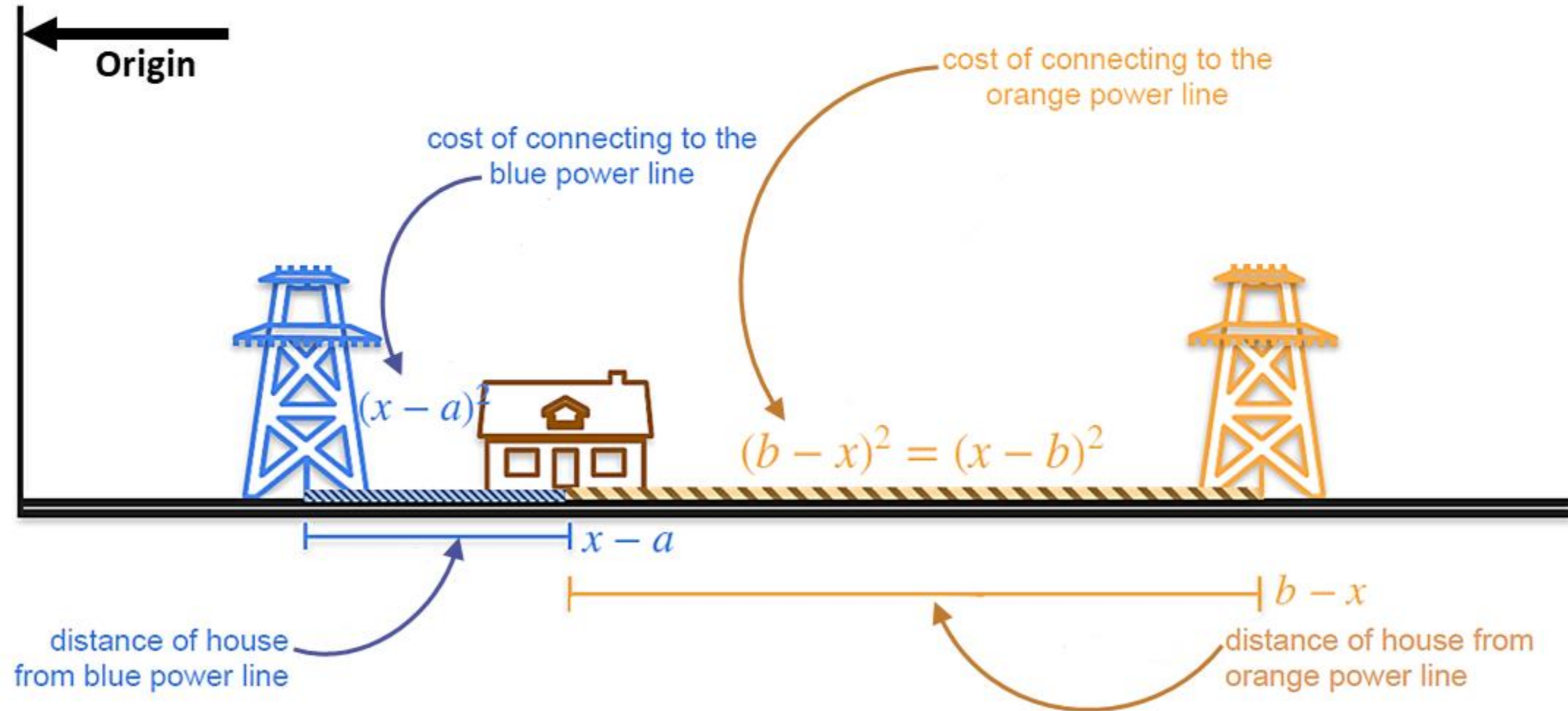
Starting with a very simple problem, let's say that there's only one power line and we want to get the power with the least amount of cost, where would you put the house?



The cost of connecting to this power line is going to be x^2 , that distance squared, So if the house is set right by the power line, then the distance from the power line is zero units, meaning that the cost of connecting to the powerline is going to be 0 squared which is 0.

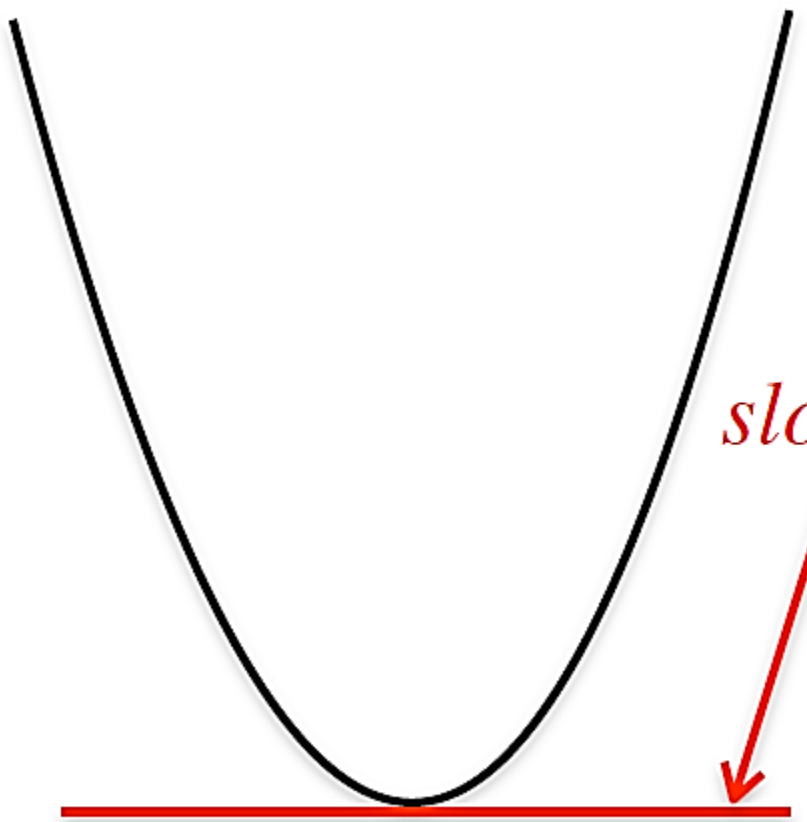
Minimize the cost when connecting the house to two power lines, a blue one and an orange one.





$$(x - a)^2 + (x - b)^2$$

slope = 0



$$\frac{d}{dx} [(x - a)^2 + (x - b)^2] = 0$$

$$2(x - a) + 2(x - b) = 0$$

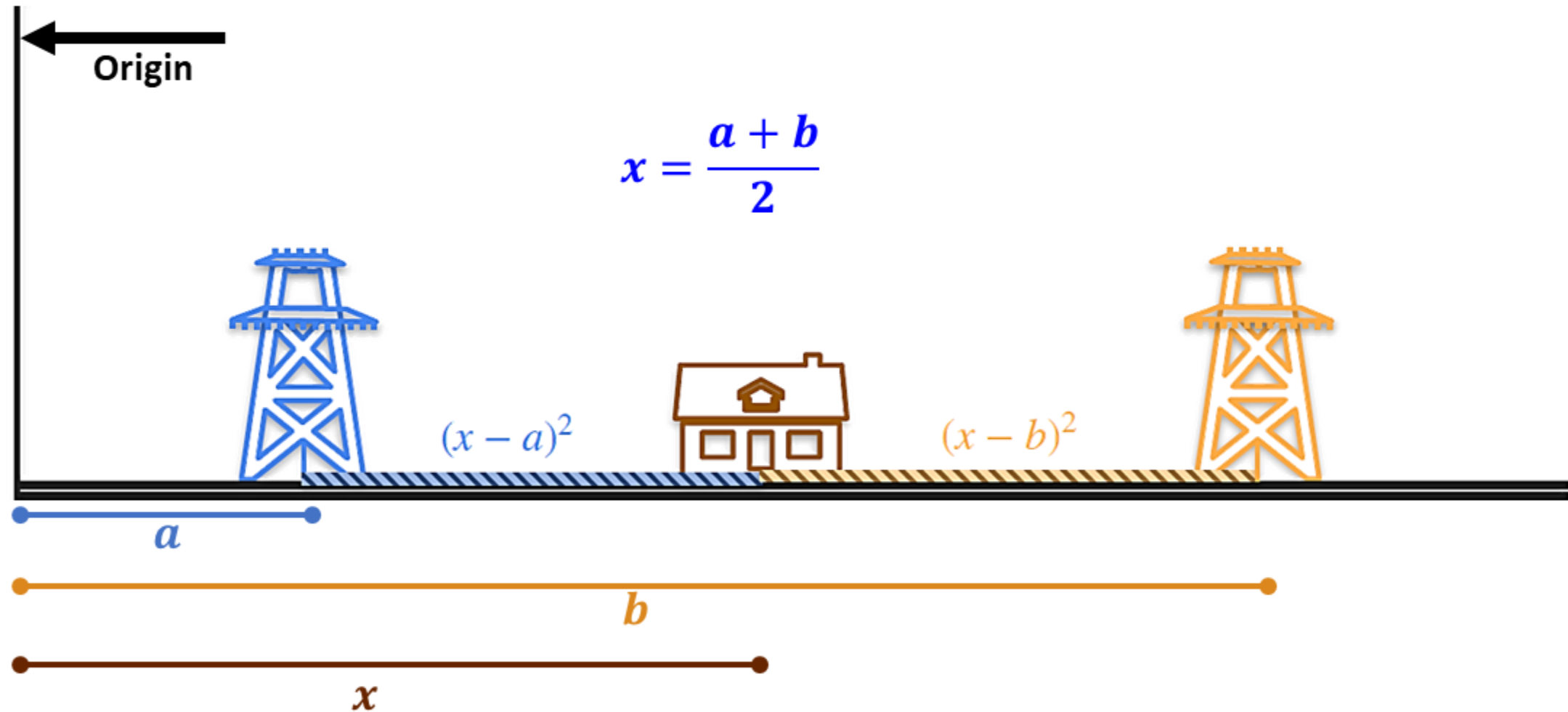
divide by 2

$$(x - a) + (x - b) = 0$$

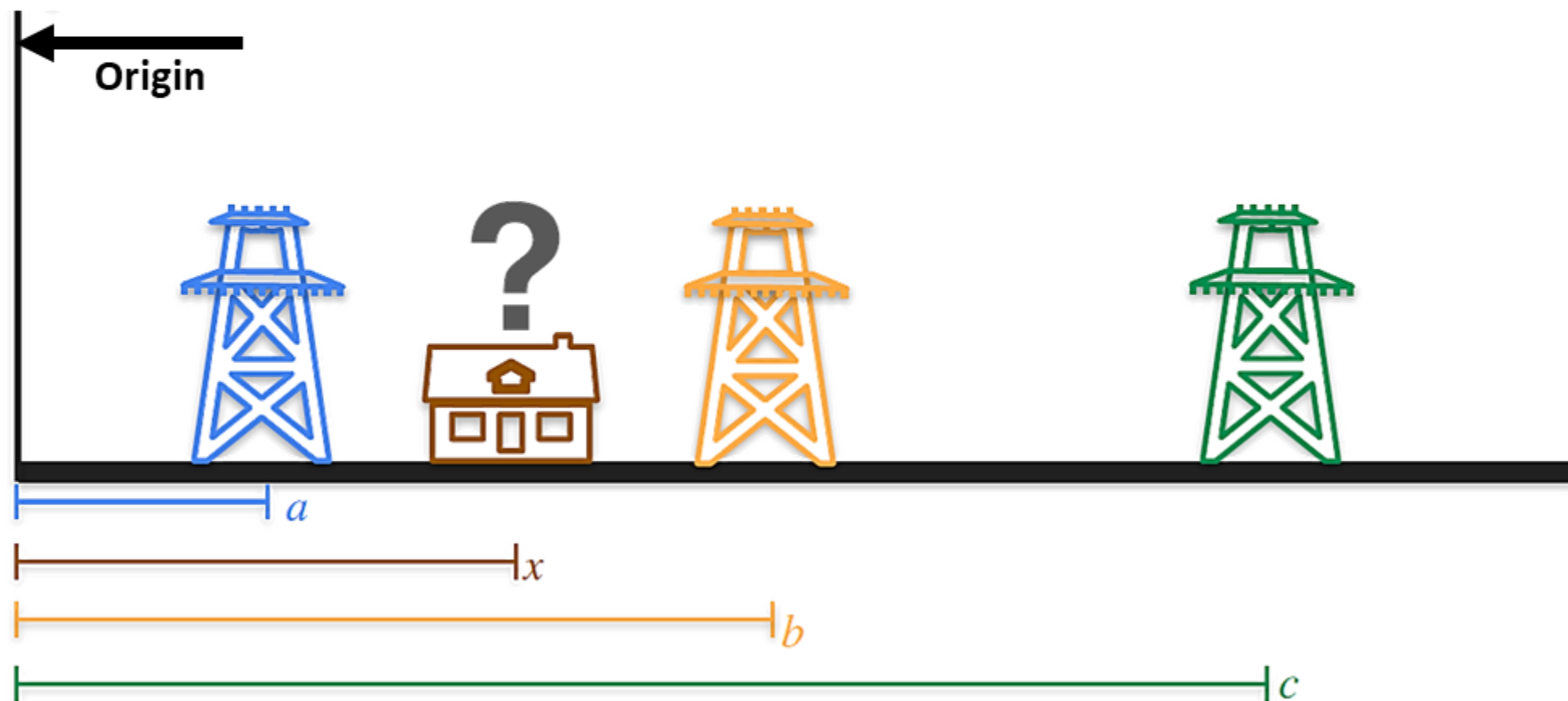
$$2x - a - b = 0$$

$$2x = a + b$$

$$x = \frac{a + b}{2}$$



Minimize the cost when connecting the house to three power lines, a blue one and an orange one.

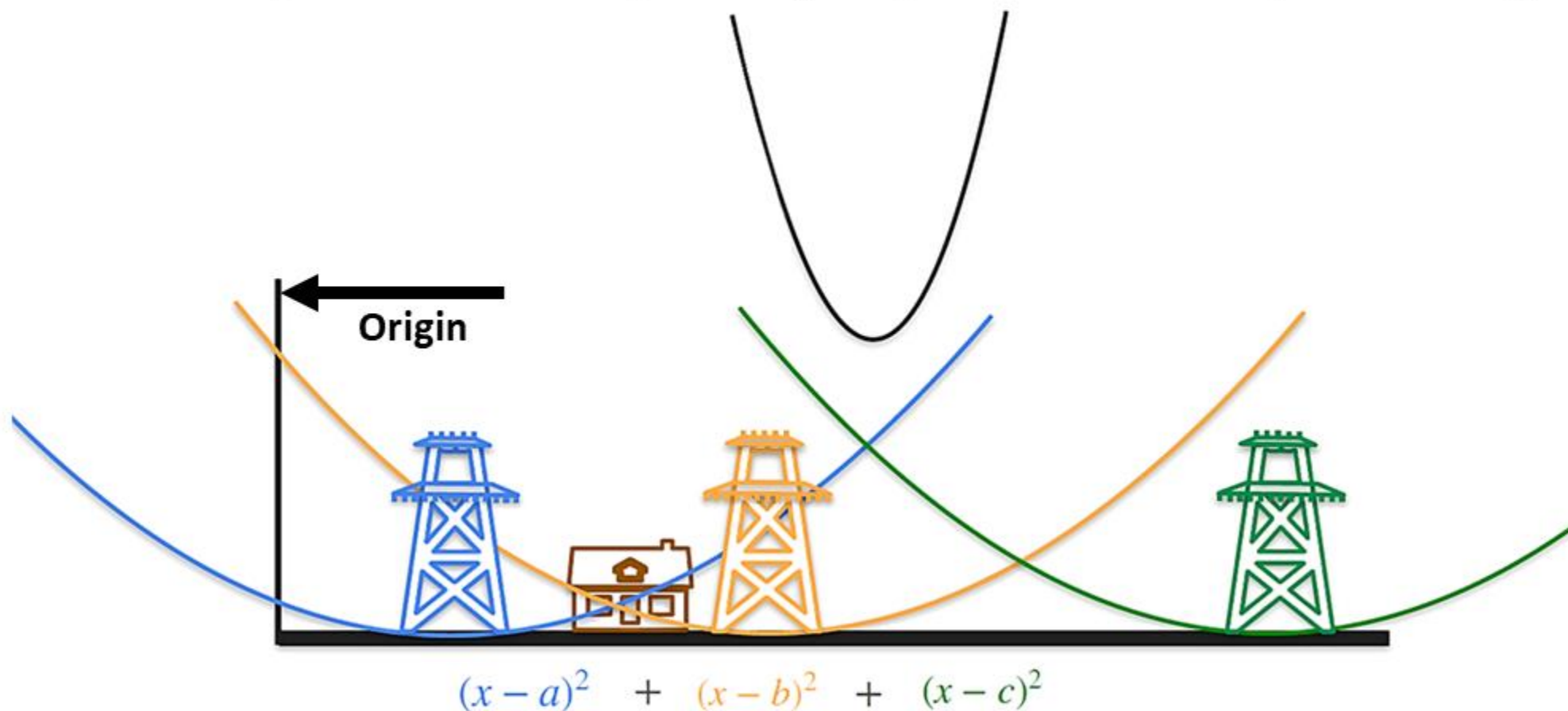


Origin

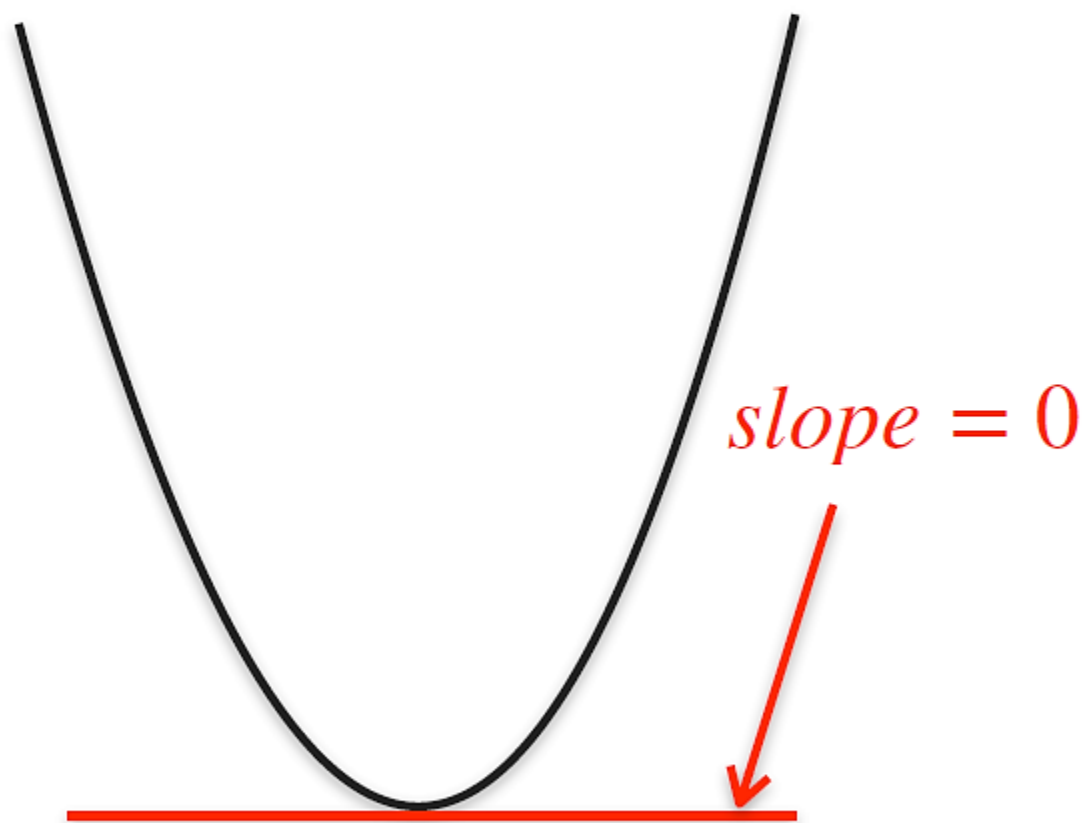
$$\text{Cost function} = (x - a)^2 + (x - b)^2 + (x - c)^2$$

distance of house
from blue power line $(x - a)$ $(b - x)$ distance of house from
orange power line $(c - x)$ distance of house
from green power
line

If you summed the heights of the three parabolas, you get the black parabola. And the point is to find that minimum point in the black parabola, the point where the slope of the tangent is zero.



$$(x - a)^2 + (x - b)^2 + (x - c)^2$$



$$\frac{d}{dx} [(x - a)^2 + (x - b)^2 + (x - c)^2] = 0$$

$$2(x - a) + 2(x - b) + 2(x - c) = 0$$

divide by 2

$$(x - a) + (x - b) + (x - c) = 0$$

$$3x - a - b - c = 0$$

$$3x = a + b + c$$

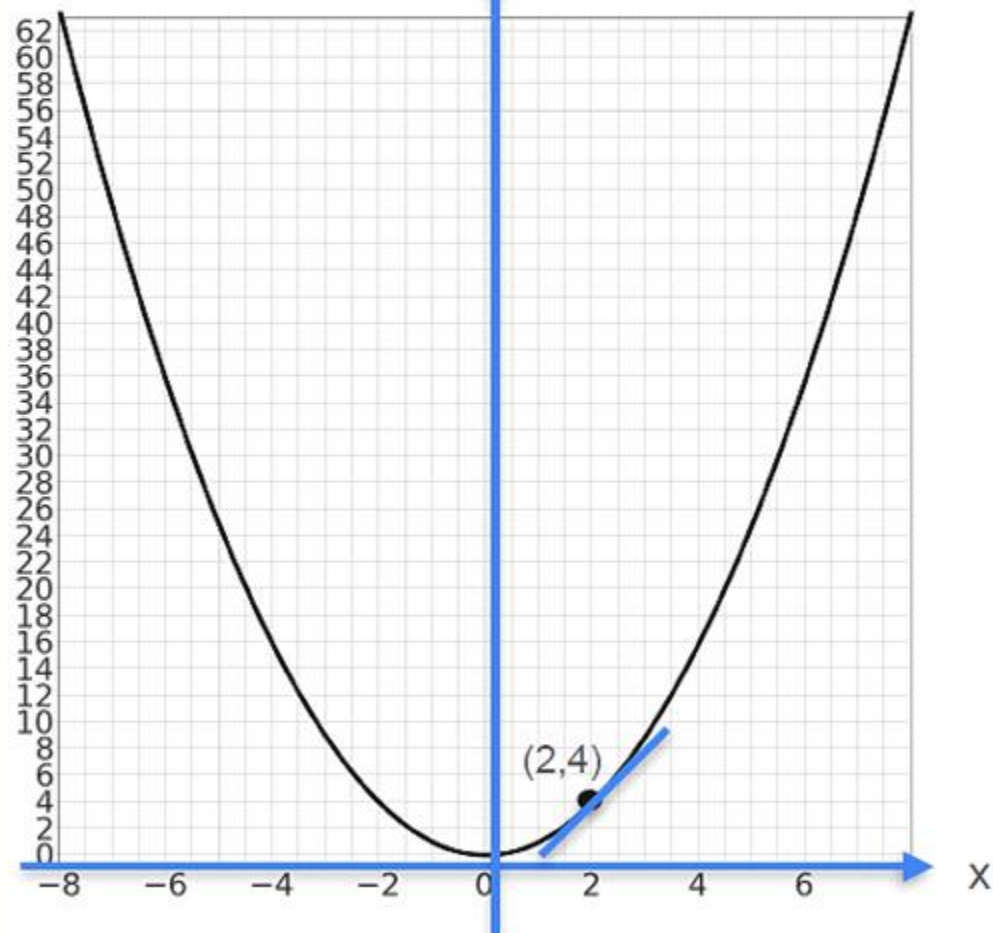
$$x = \frac{a + b + c}{3}$$

Square Loss function:

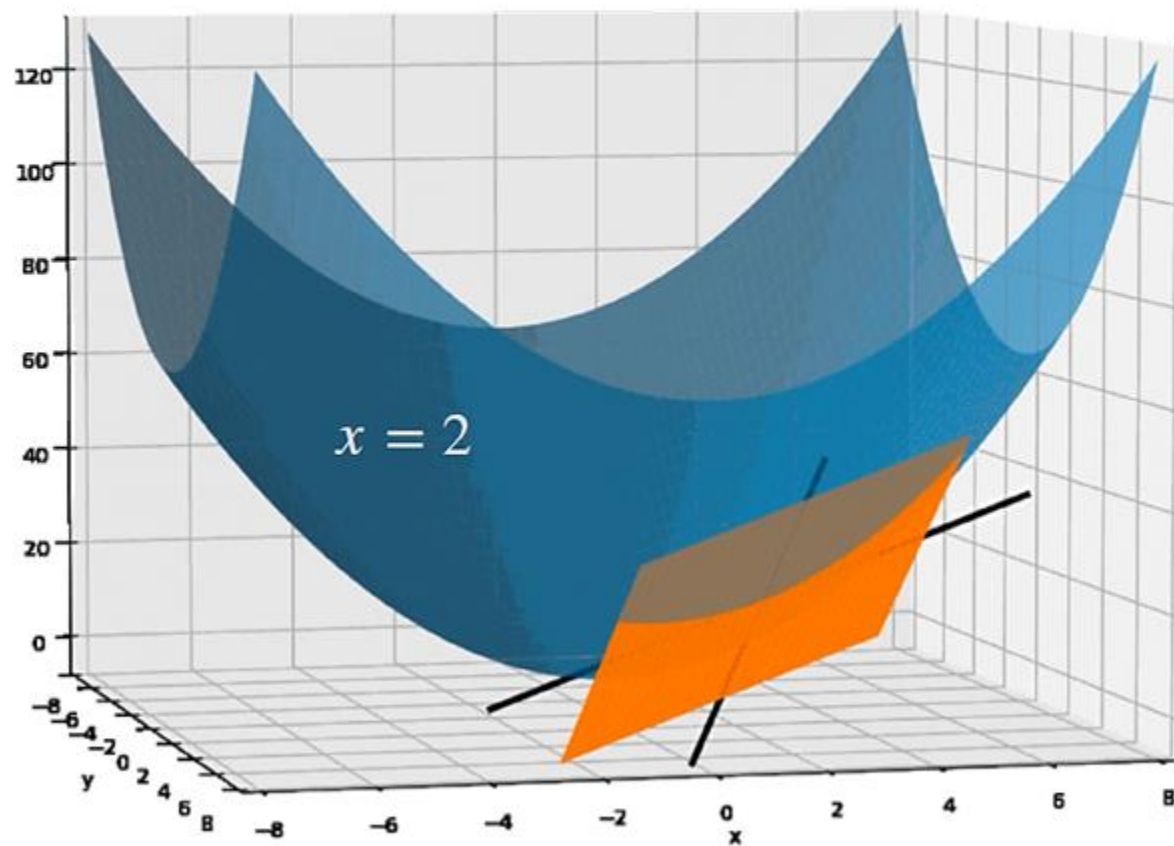
$$\textit{Minimize } (x - a_1)^2 + (x - a_2)^2 + (x - a_3)^2 + \cdots + (x - a_n)^2$$

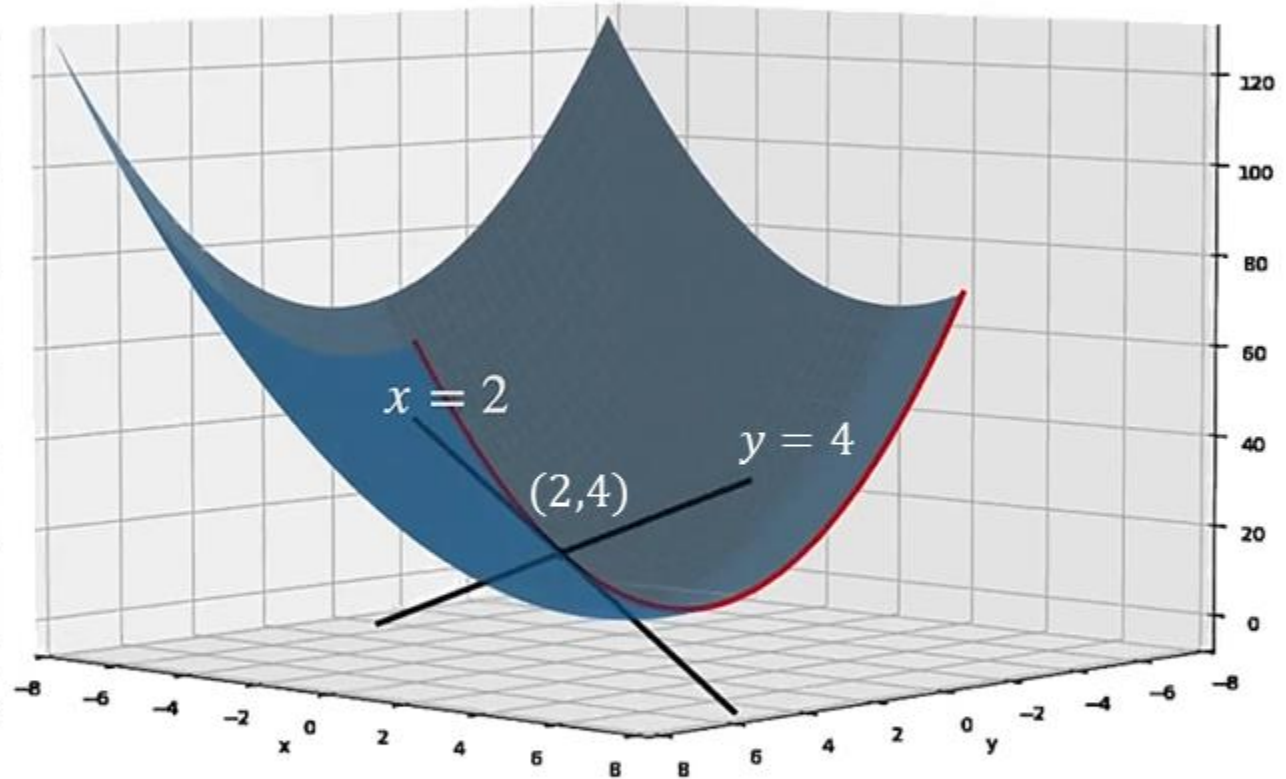
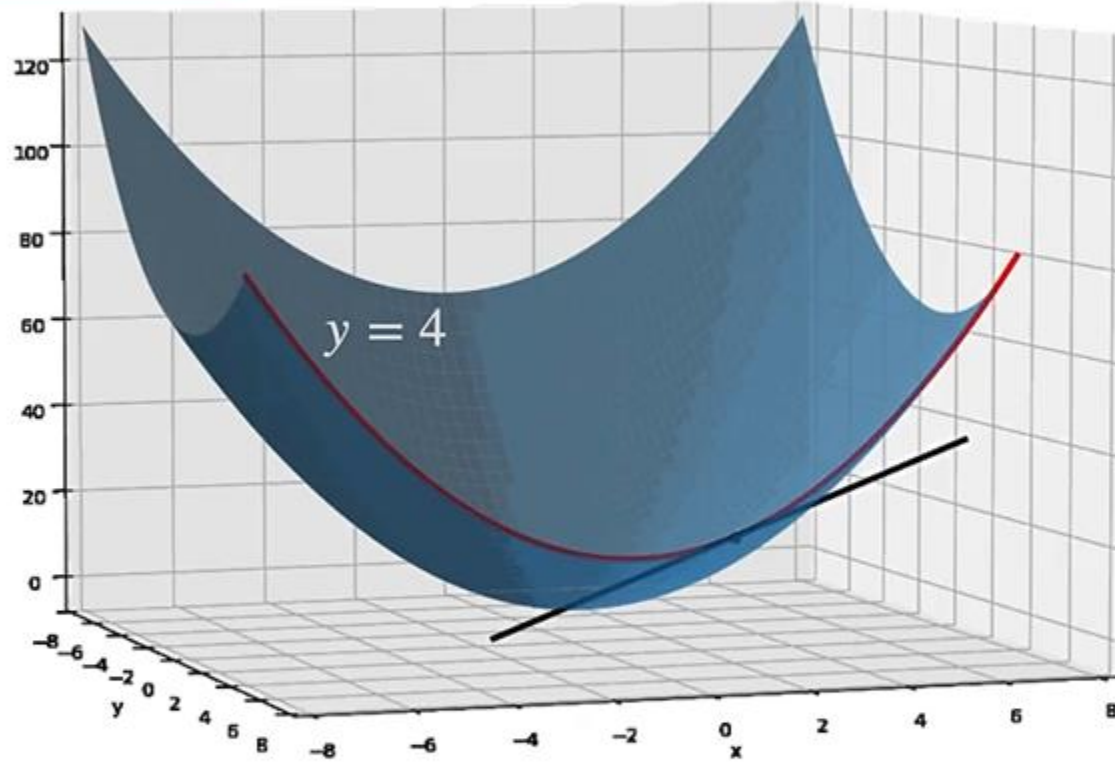
$$\textit{Solution: } x = \frac{a_1 + a_2 + a_3 + \cdots + a_n}{n}$$

$$f(x) = x^2$$



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

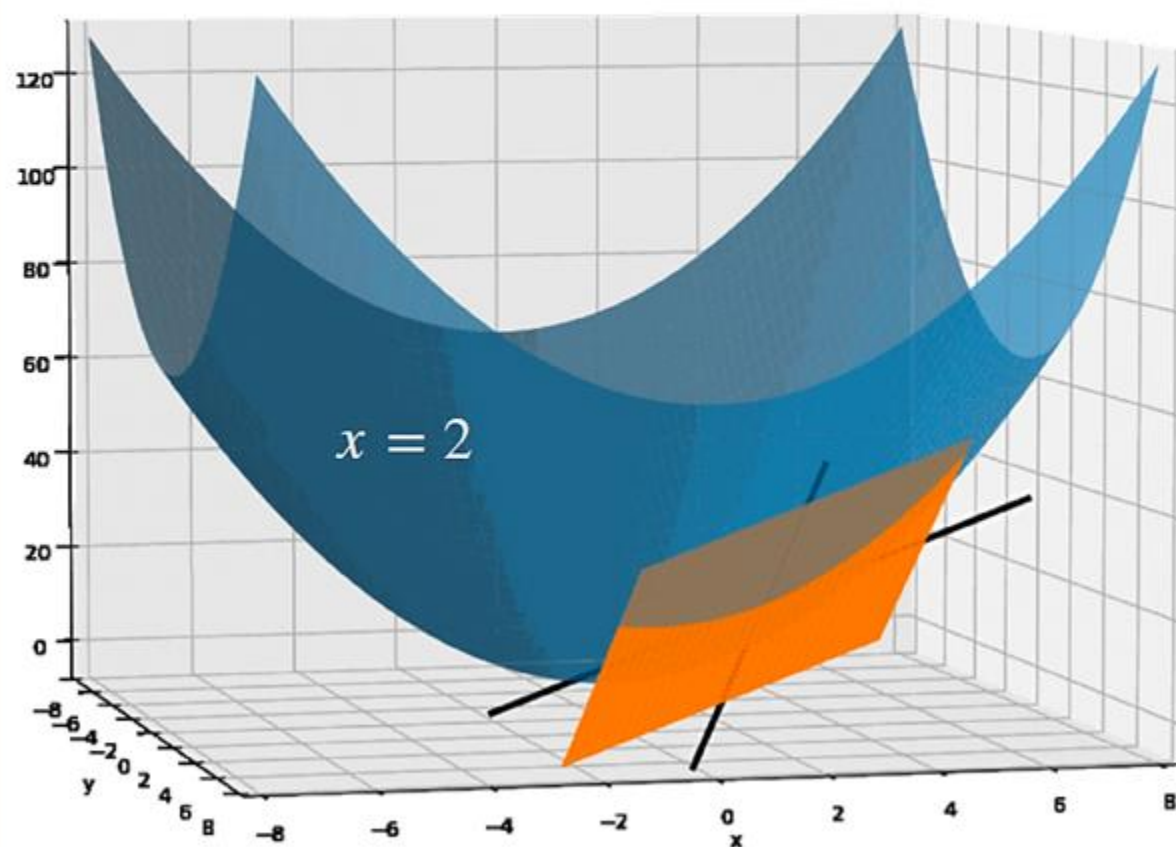




the function has two inputs and one output. Therefore, it needs to be plot in 3D. We're going to plot it over here. x and y are the two horizontal axis and the floor. z , or $f(x, y)$ is plot as the height.

Functions of Two Variables

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



$$y = 4, f(x, 4) = x^2 + (4)^2$$

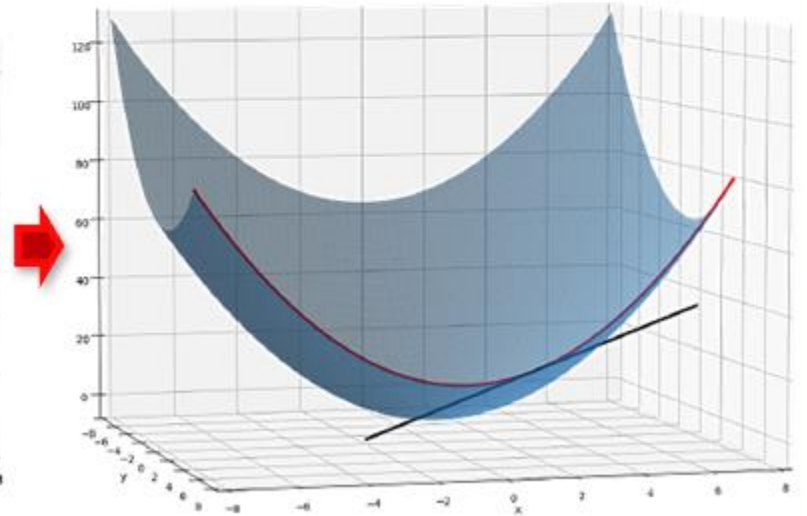
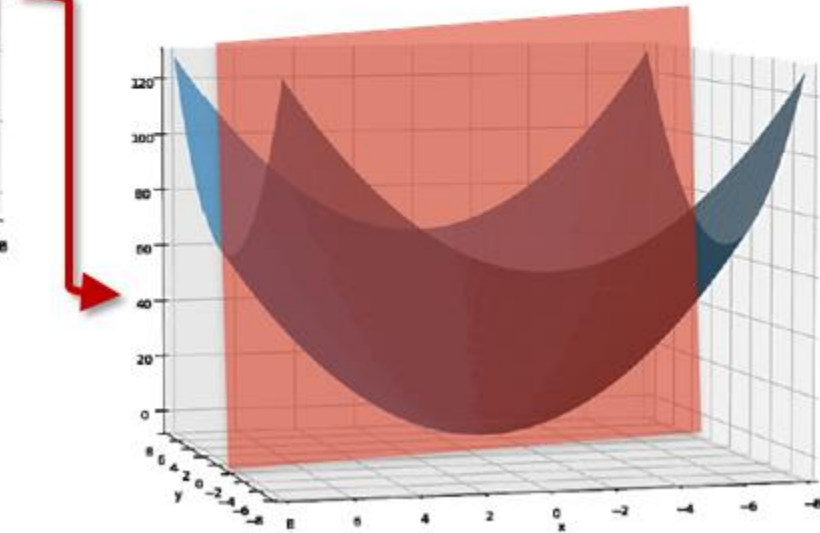
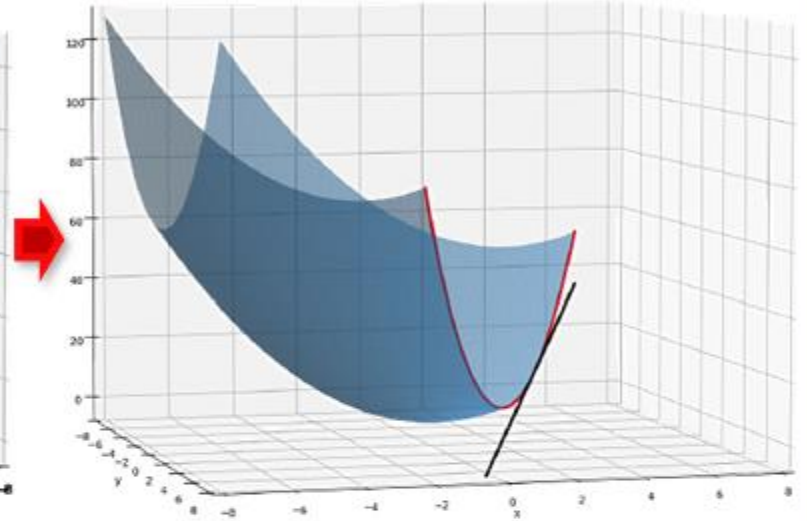
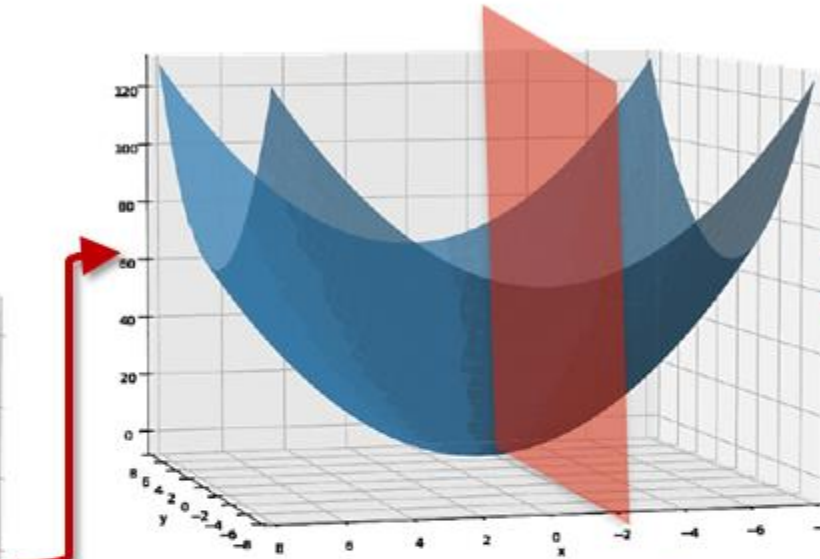
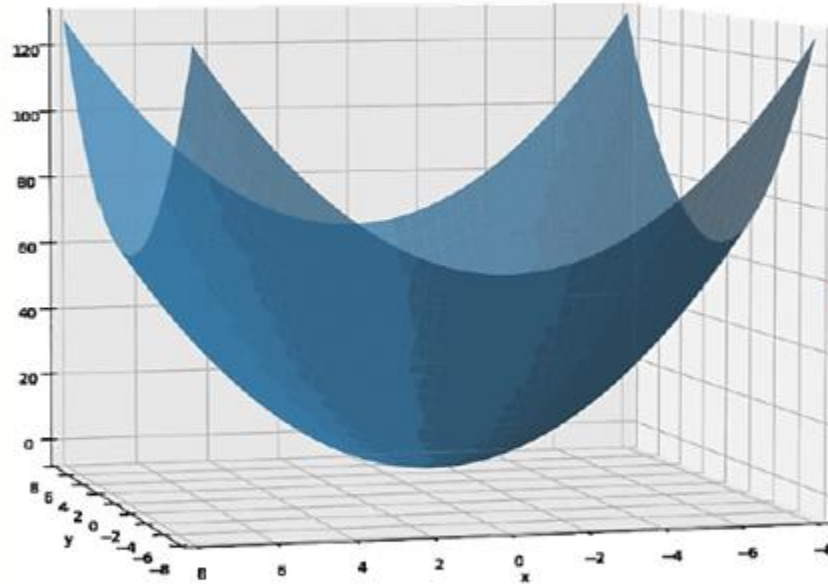
$$\frac{d}{dx}(f(x, 4)) = 2x$$

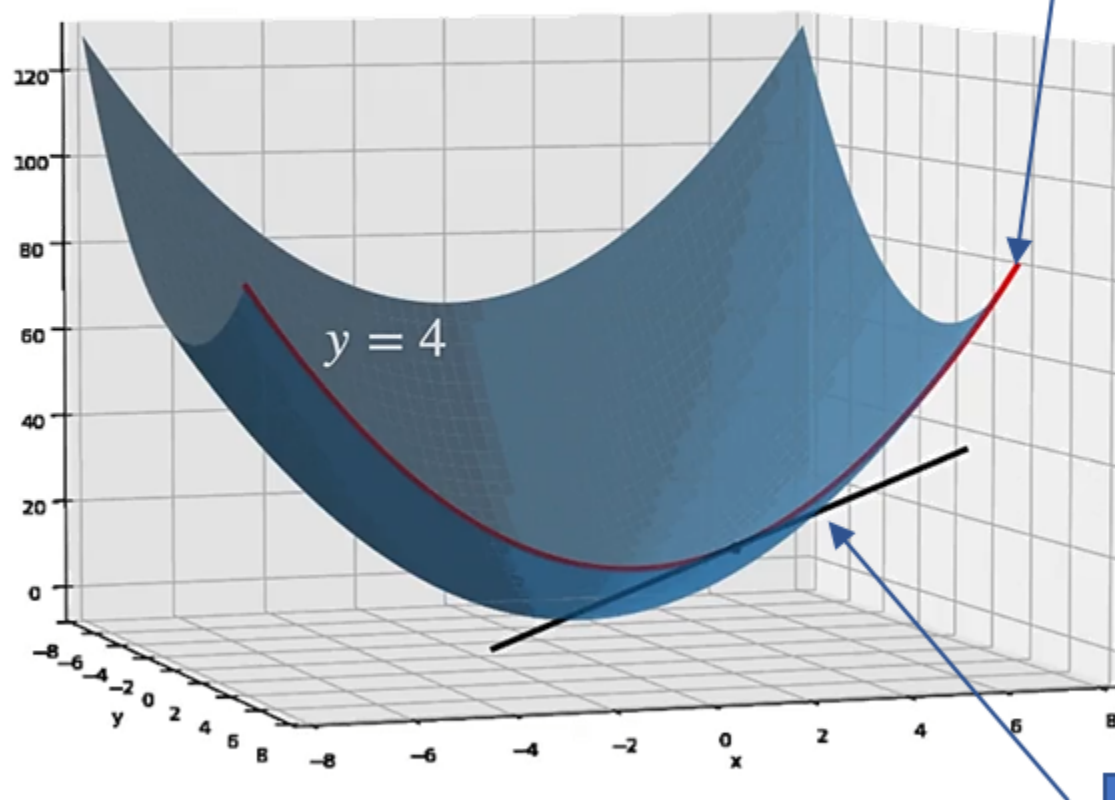
$$x = 2, f(2, y) = (2)^2 + y^2$$

$$\frac{d}{dy}(f(2, y)) = 2y$$

The tangent plane contains both tangent lines.

Partial Derivatives





Function of one variable

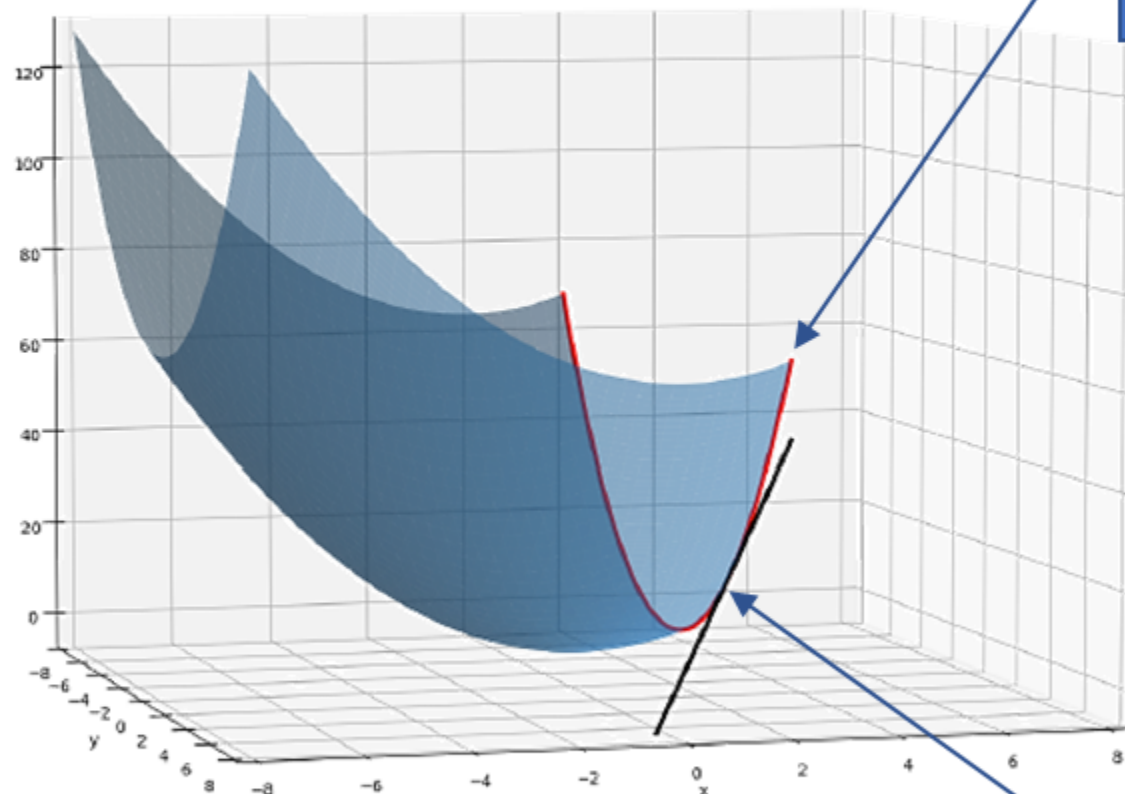
Treating y as a constant, now this function is only a function of x . So when differentiating this function with respect to x :

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

y is a constant (*derivative of a constant = 0*)

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

Slope of partial derivative



*Function of one
variable*

Treating x as a constant, now this function is only a function of y . So when differentiating this function with respect to y :

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

x is a constant (*derivative of a constant = 0*)

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

*Slope of partial
derivative*

$$f(x, y)$$



Partial derivative notation

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

*Partial derivative of f
with respect to x*

Partial derivative notation

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

*Partial derivative of f
with respect to y*

Summary**Solving function of two variables:**

Find partial derivative of f with respect to x

- Treat all other variables as constants. In this case y .
- Differentiate the function using the normal rules of differentiation.

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

Find partial derivative of f with respect to y

- Treat all other variables as constants. In this case x .
- Differentiate the function using the normal rules of differentiation.

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

Example

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

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

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

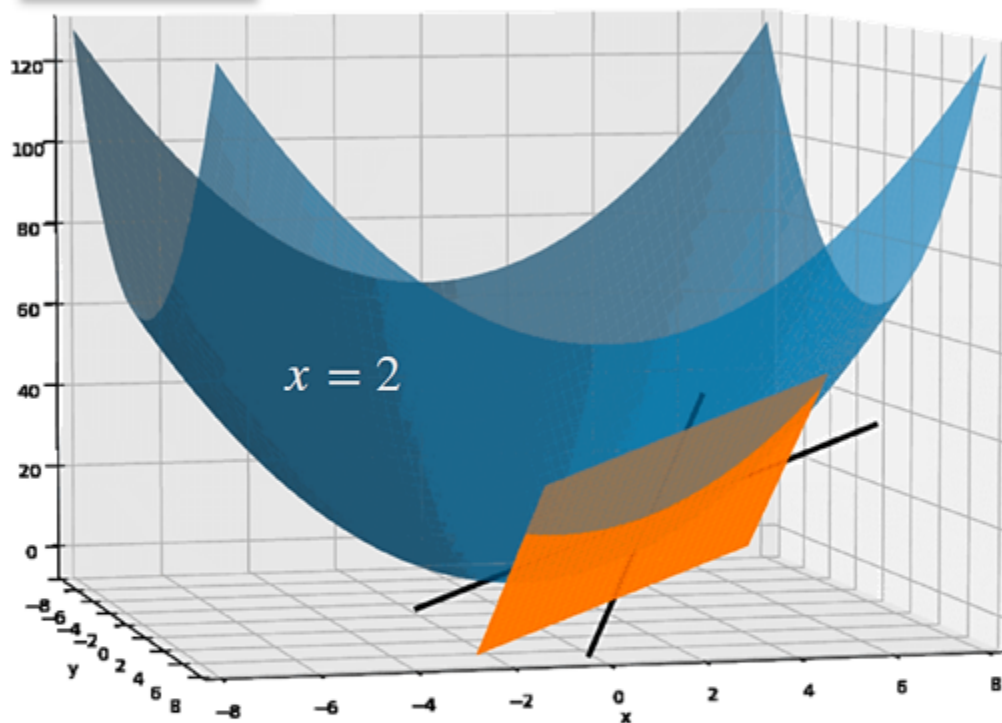
$$f_x = \frac{\partial f}{\partial x} = 6xy^3$$

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

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

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

$$f_x = \frac{\partial f}{\partial x} = 9x^2y^2$$



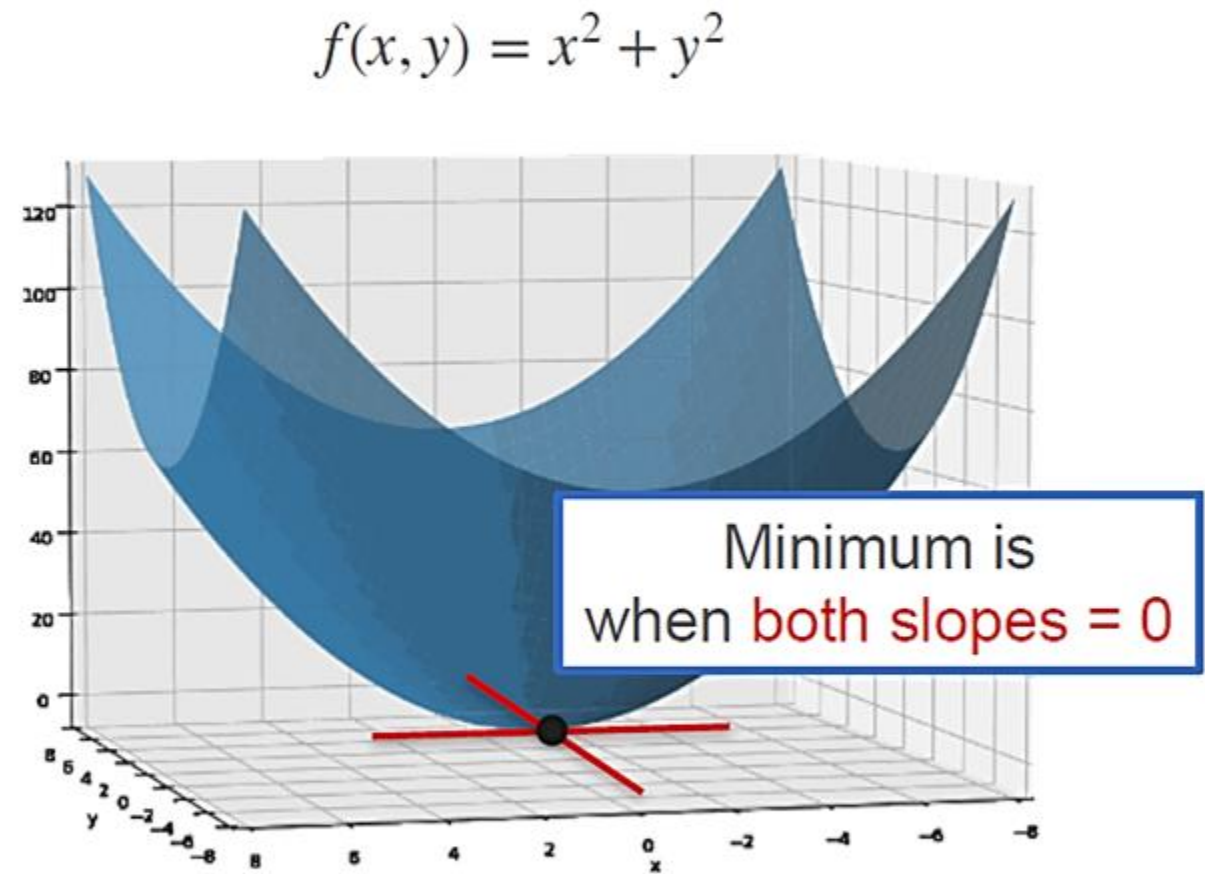
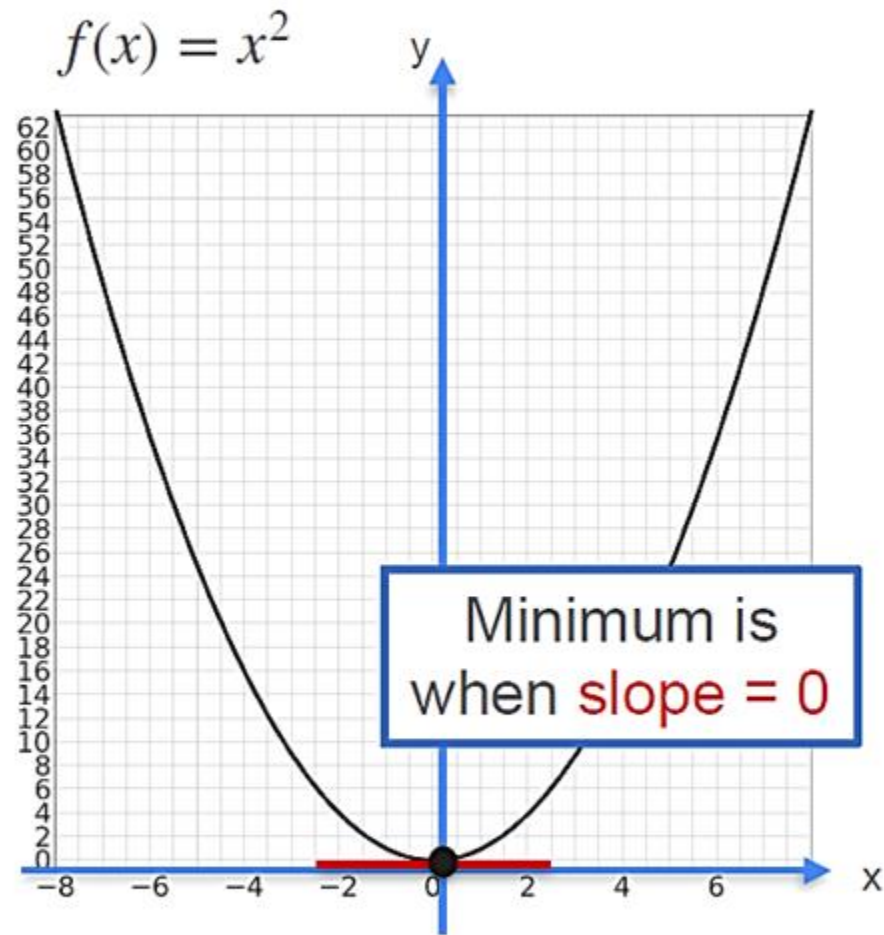
$$f(x, y) = x^2 + y^2$$

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

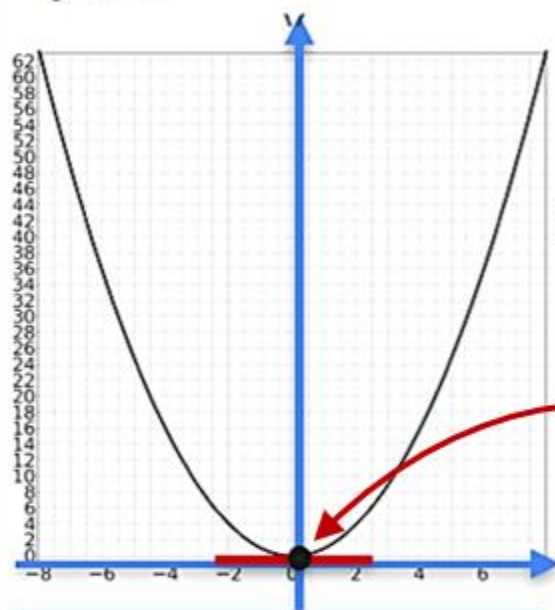
$$\frac{\partial f}{\partial y} = 2y$$

$$\text{The Gradient of } f(x, y) = \nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

The gradient is the vector containing these two partial derivatives, so basically the vector $2x, 2y$. In general, we call the vector using the nabla letter (∇) nabla (∇f) and that the collection of all partial derivatives with respect to all the variables in the function.



$$f(x) = x^2$$



$$f'(x) = 2x$$

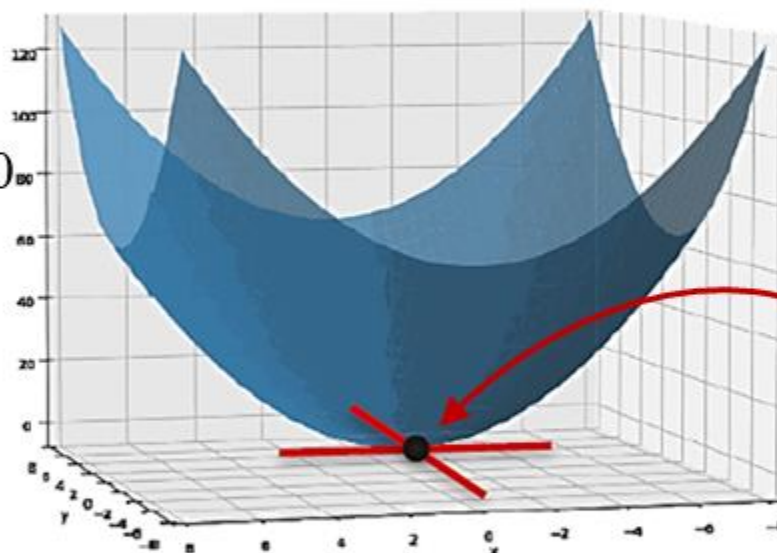
Minimum is when slope = 0

$$2x = 0$$

$$x = 0$$

Minimum is
when **slope = 0**

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



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

$$(x, y) = 0$$

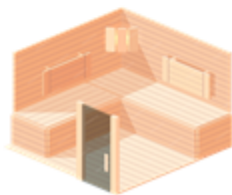
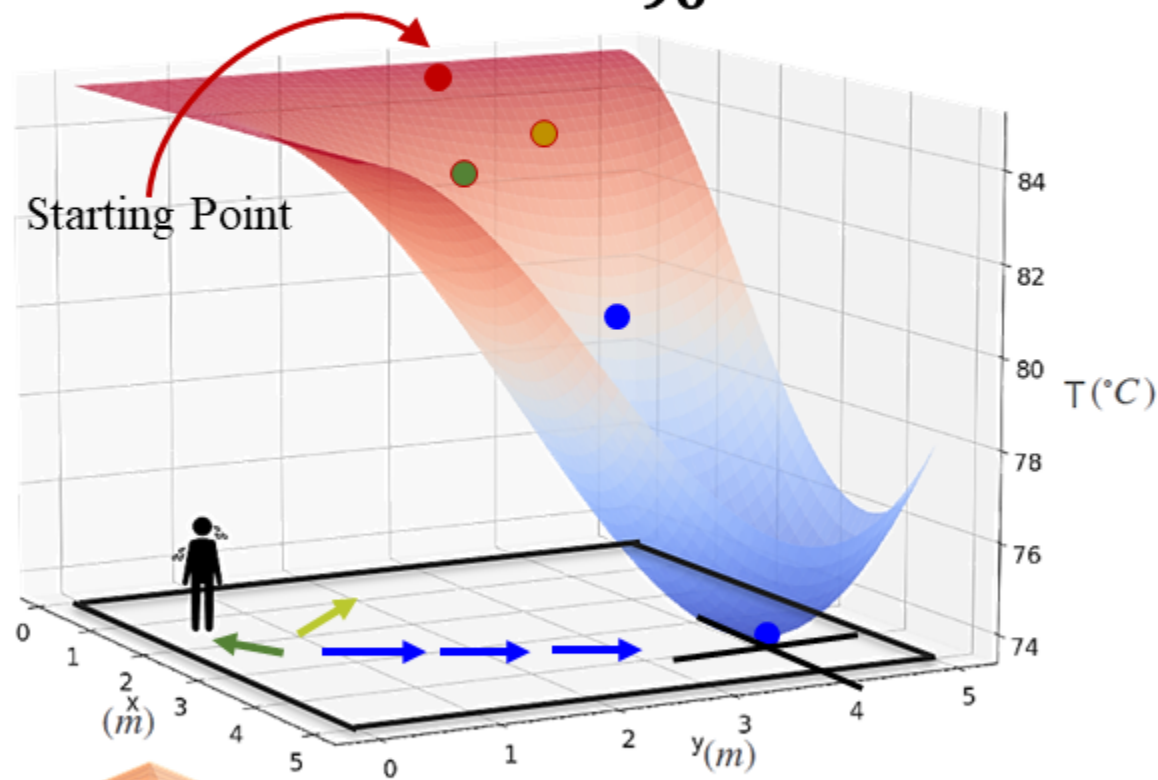
Minimum is
when **both slopes = 0**

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

Calculate $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$

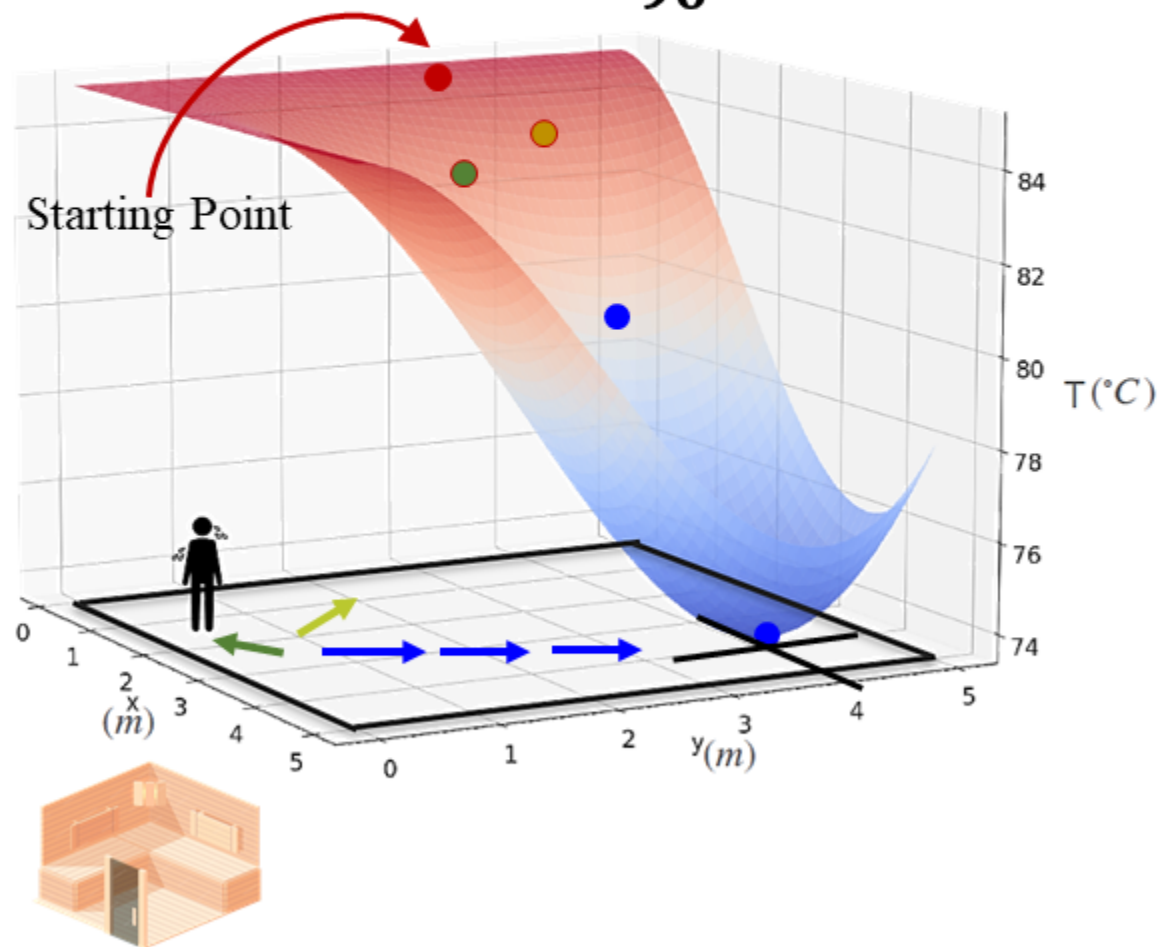
$$\frac{\partial f}{\partial x} = -\frac{1}{90} x (3x - 12) y^2 (y - 6)$$

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



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

Calculate $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$



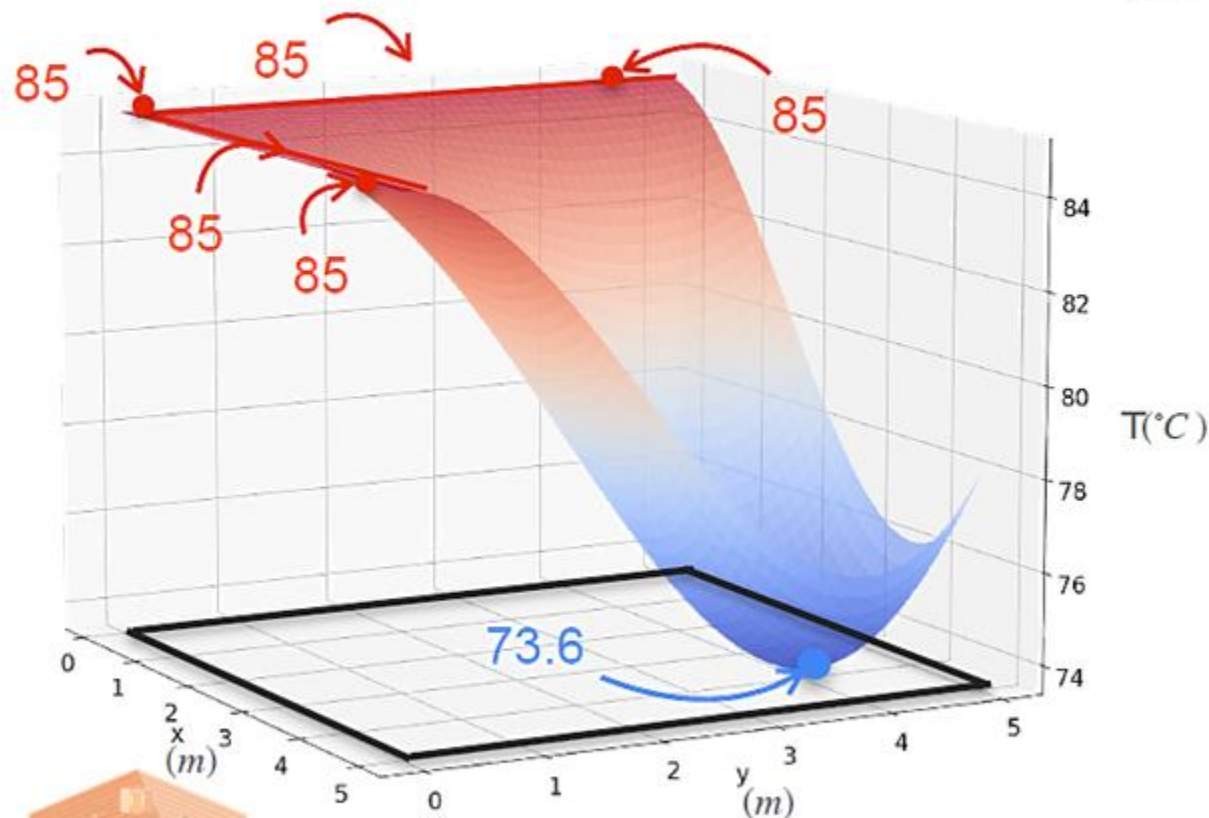
$$\frac{\partial f}{\partial x} = -\frac{1}{90} \boxed{x} \boxed{(3x - 12)} \boxed{y^2} \boxed{(y - 6)} = 0$$

$x = 0$ $x = 4$ $y = 0$ $y = 6$

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

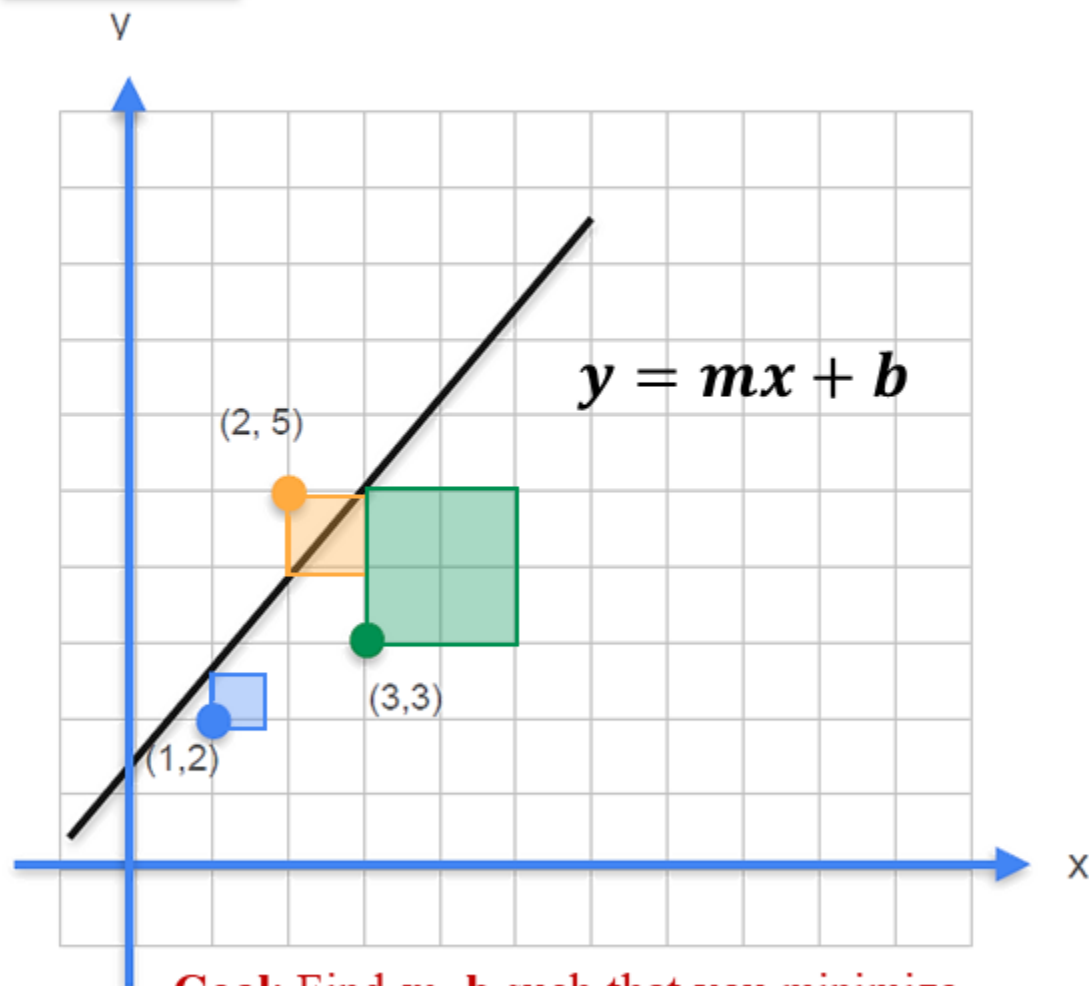
$x = 0$ $x = 6$ $y = 0$ $y = 4$

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



Points for Minima

x	y	
0	0	Maxima
0	4	Maxima
0	6	Outside
4	0	Maxima
4	4	Minima
6	0	Outside
6	6	Outside



Goal: Find m , b such that you minimize the sum of squares cost.



(1, 2)

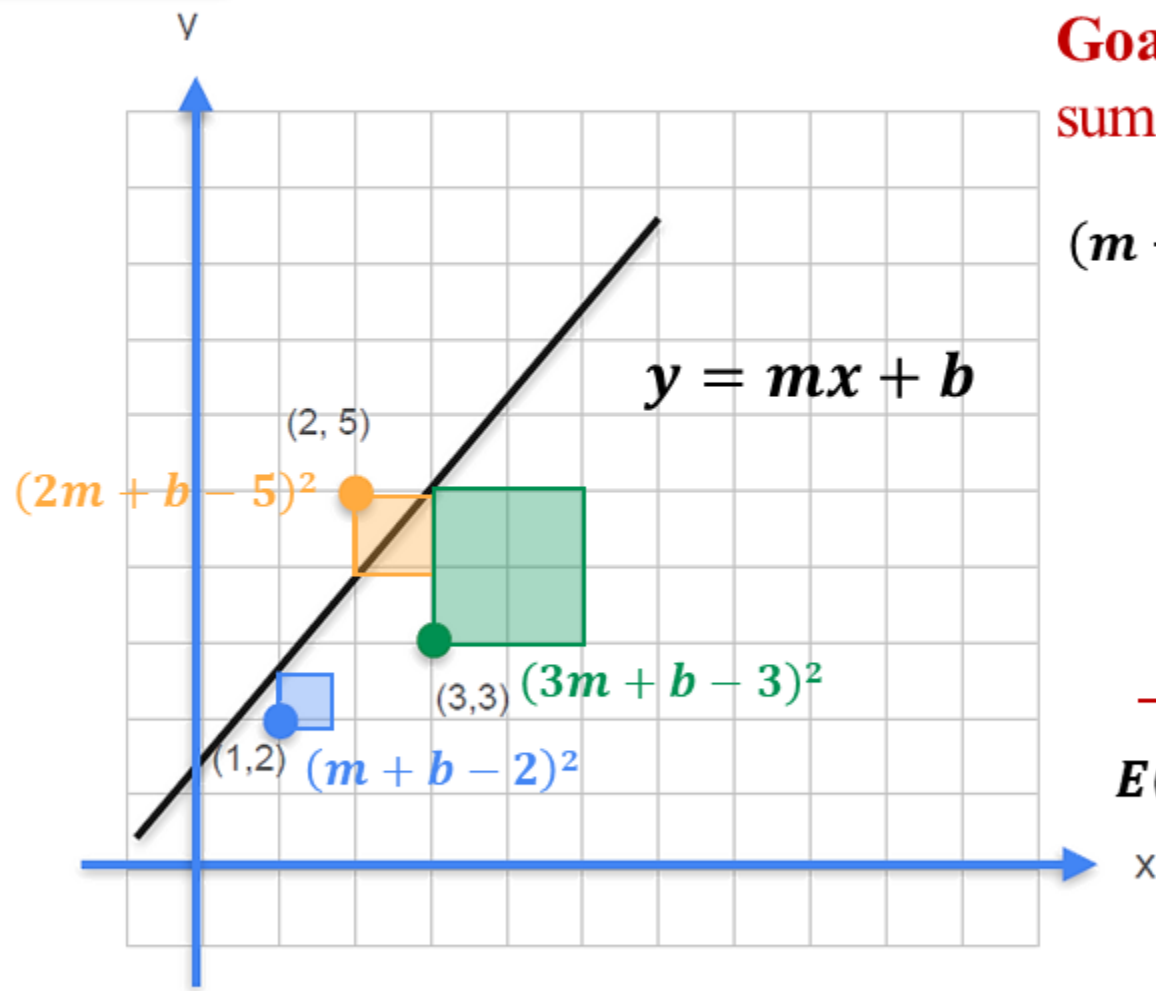


(2, 5)



(3, 3)

Finding the optimal place for a fiber line connection that goes in a straight line in such a way that you reduce the total cost of connecting to the three power lines by connecting a wire to touch the fiber line, but the wire needs to be parallel to the y-axis. And the cost for connection is the square of the length of the wires.



Goal: Find m , b such that you minimize the sum of squares cost.

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

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

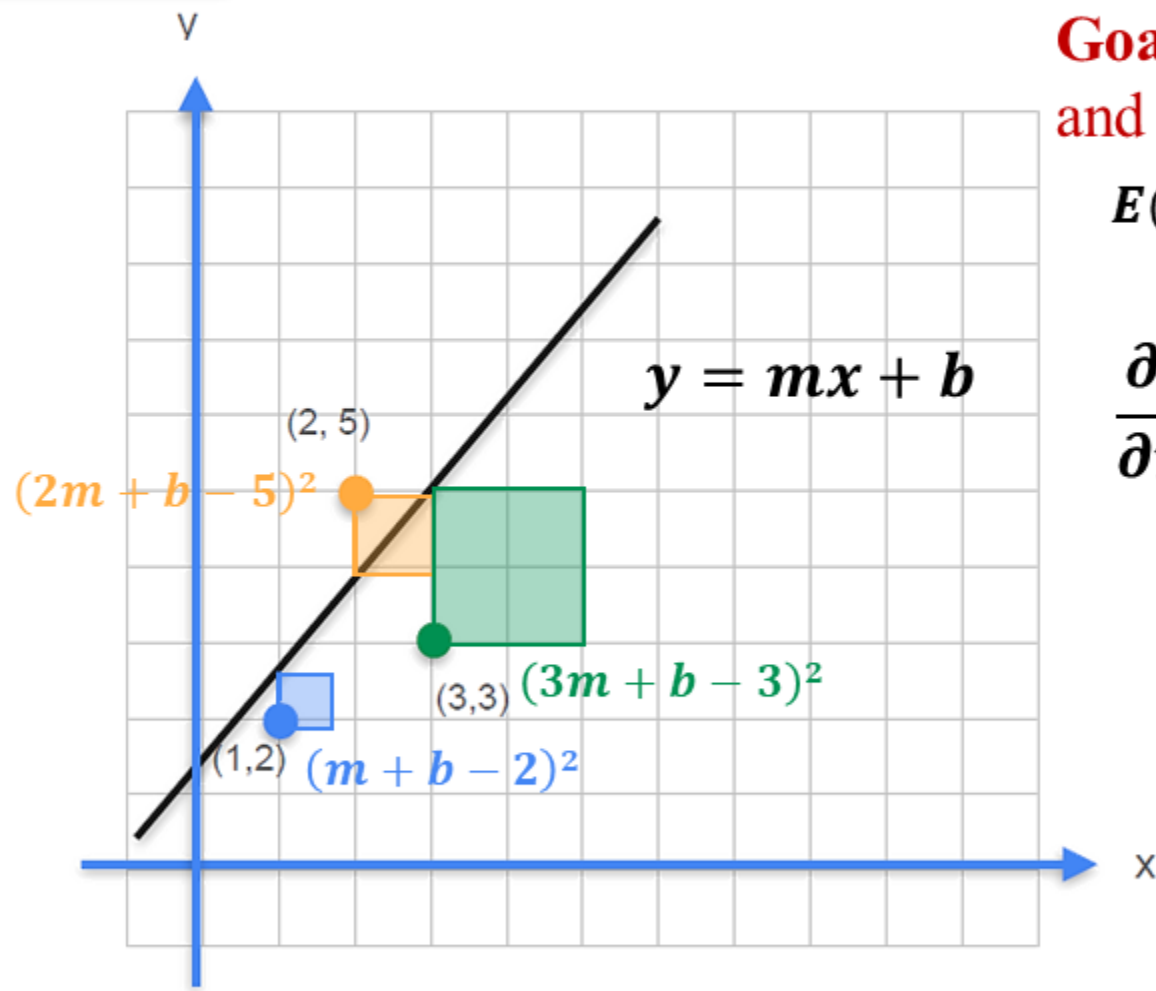
+

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

+

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

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



Goal: partial derivative of E with respect to m and b .

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

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

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

Calculate
 m and b

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

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

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

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

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

$$4m - 2 = 0$$

$$4m = 2$$

$$m = \frac{1}{2}$$

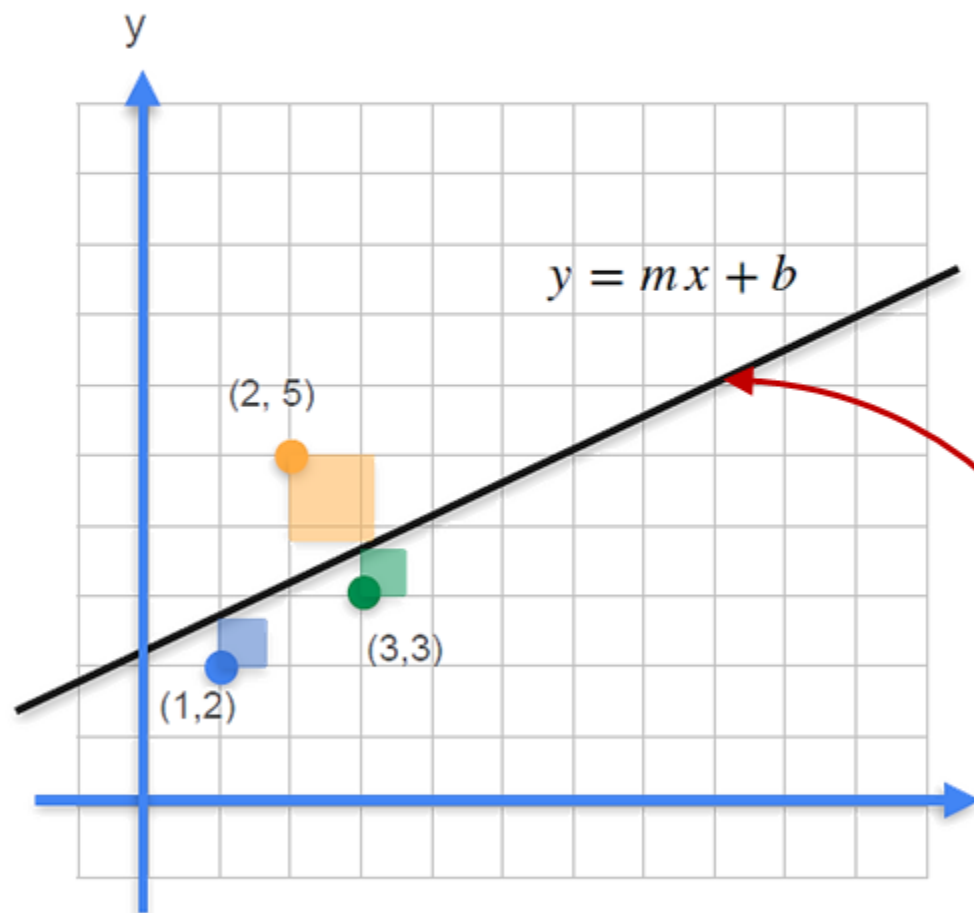
$$6b + 12\left(\frac{1}{2}\right) - 20 = 0$$

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

$$6b - 14 = 0$$

$$6b = 14$$

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



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

$$E(m, b) = \text{where } (m) = \frac{1}{2}, \text{ and } (b) = \frac{7}{3}$$

$$\therefore E(m, b) \cong 4.167$$

That is the minimum value of the cost.

Plotting the line of equation $\frac{1}{2}x + \frac{7}{3}$

