# Gradients and Gradient Descent

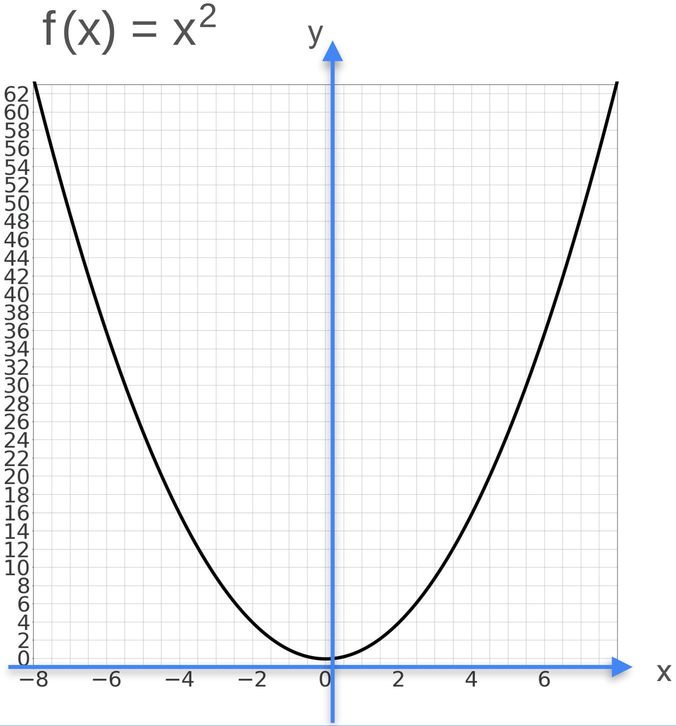
## Tangent planes

DeepLearning.AI

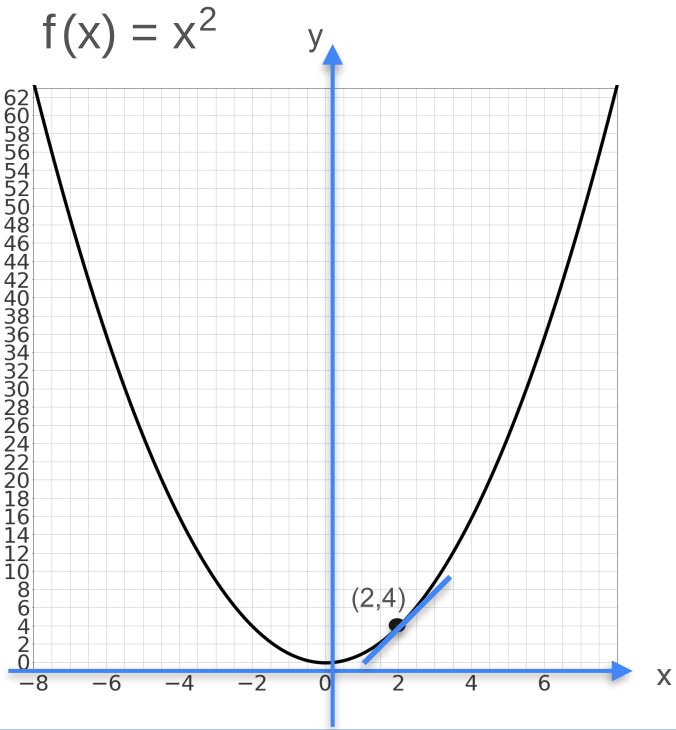
# Functions of Two Variables

# Functions of Two Variables

$f(x) = x^2$

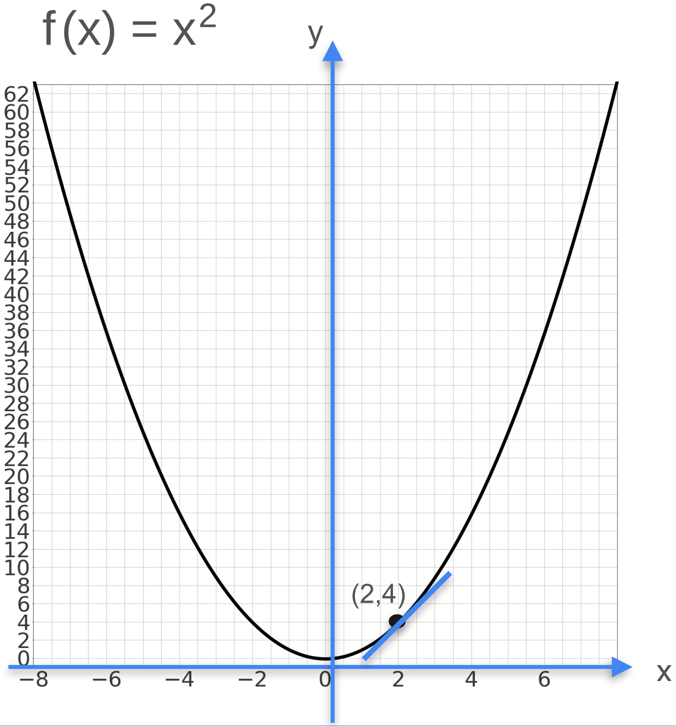


# Functions of Two Variables

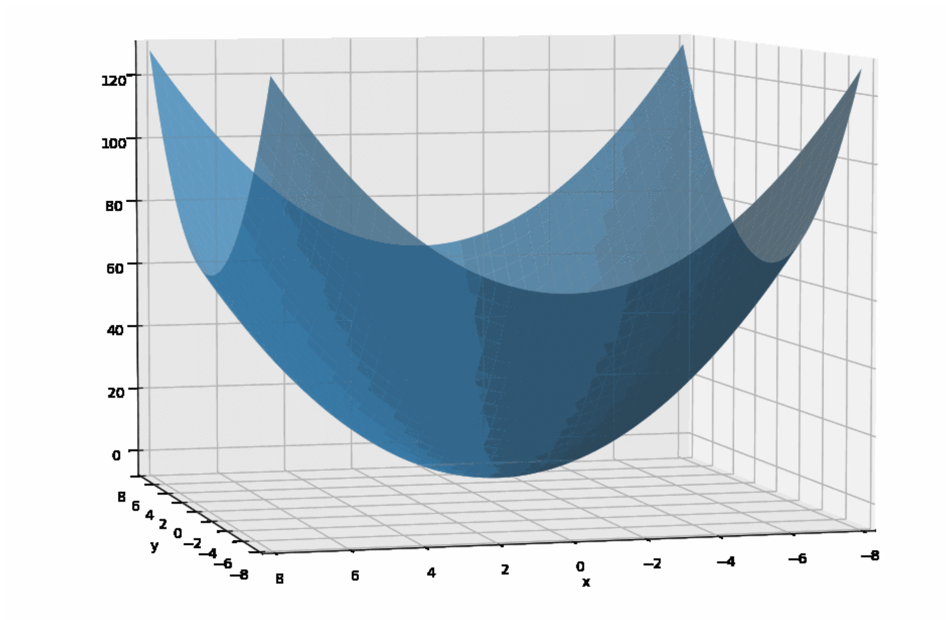


# Functions of Two Variables

$f(x) = x^2$

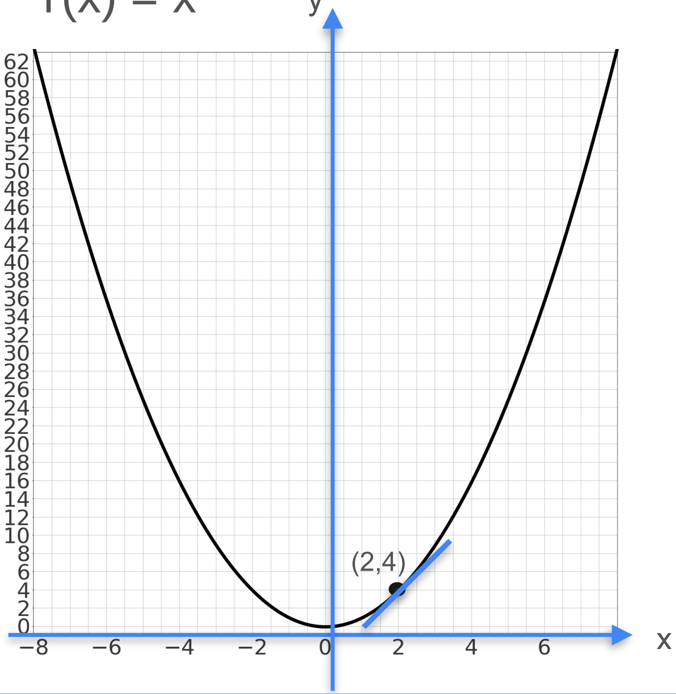



$f(x, y) = x^2 + y^2$

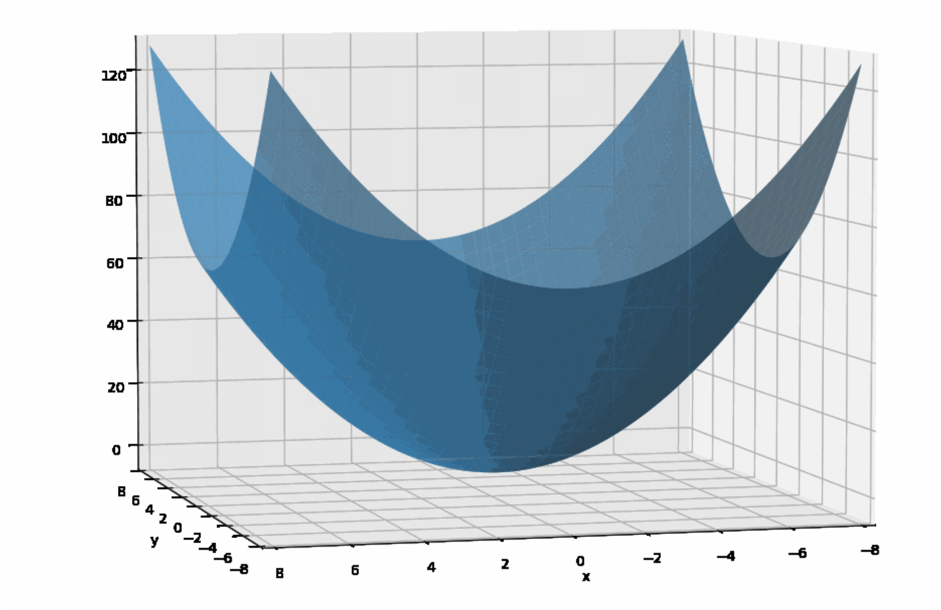


# Functions of Two Variables

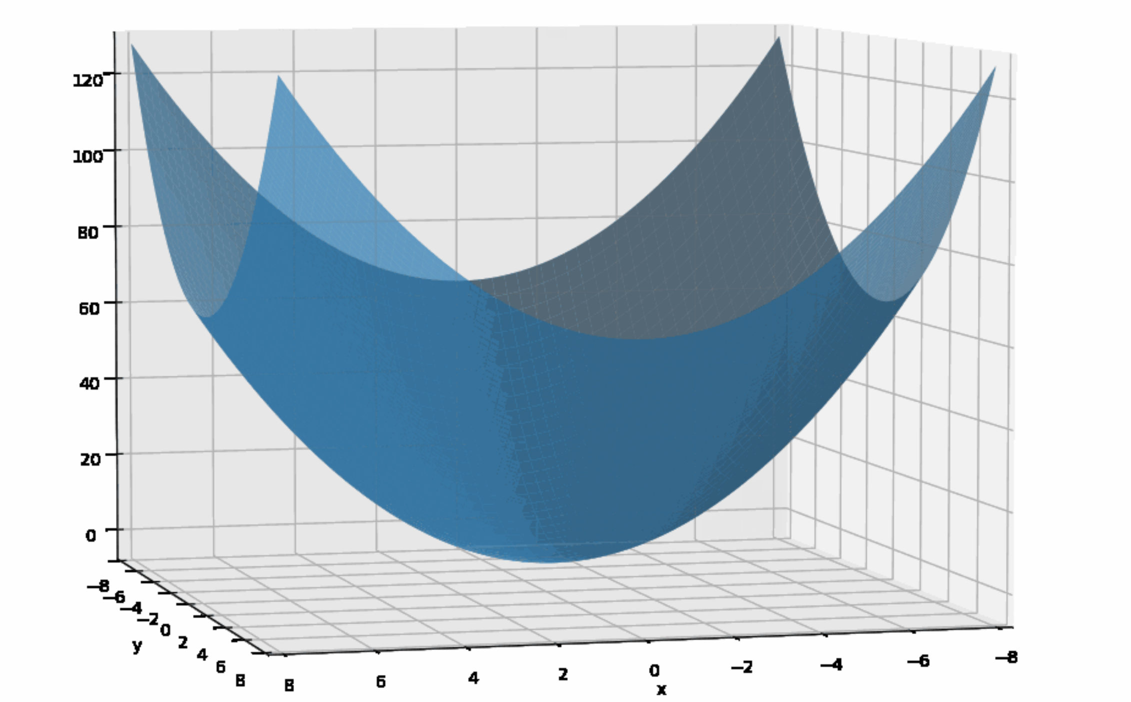
$f(x) = x^2$



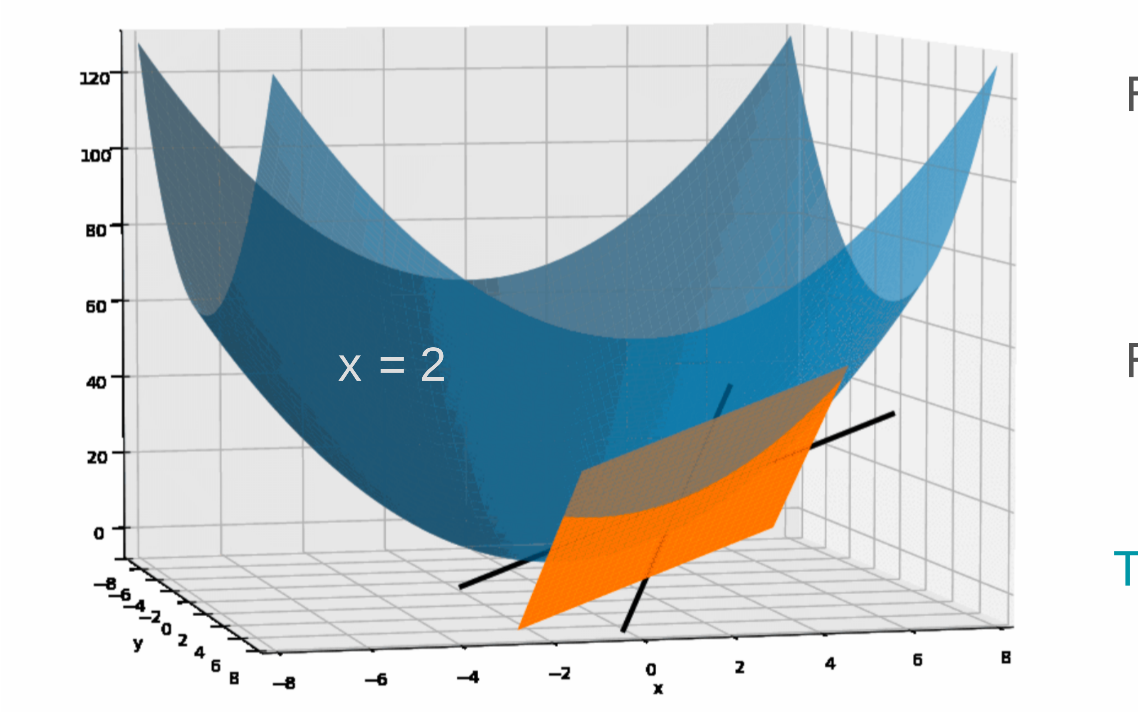

$f(x, y) = x^2 + y^2$



# Finding the Tangent Plane



# Finding the Tangent Plane




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

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

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

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

The tangent plane contains both tangent lines.

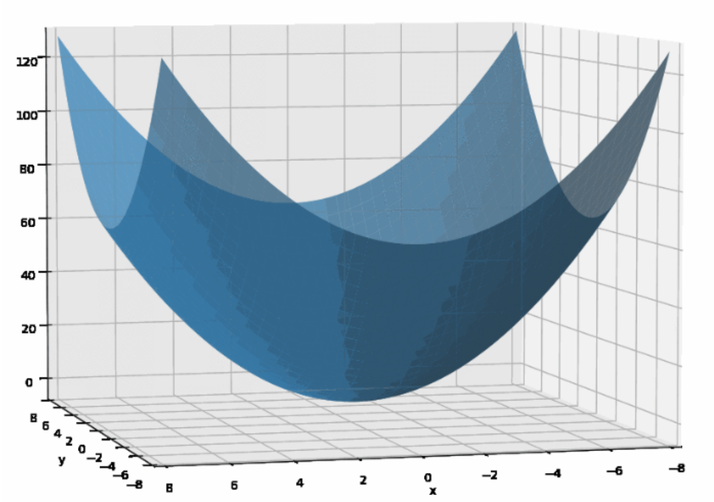
# Video 2: Introduction to Partial Derivatives

Example with the parabola, show tangent plane and slices

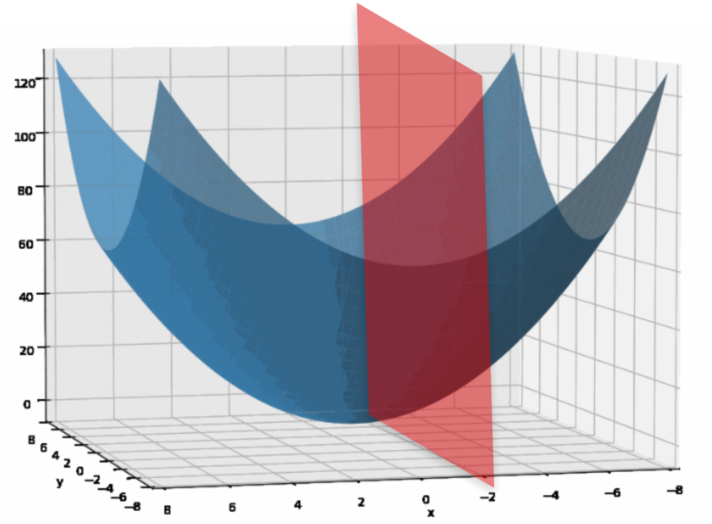
# Gradients and Gradient Descent

## Partial derivatives - Part 1

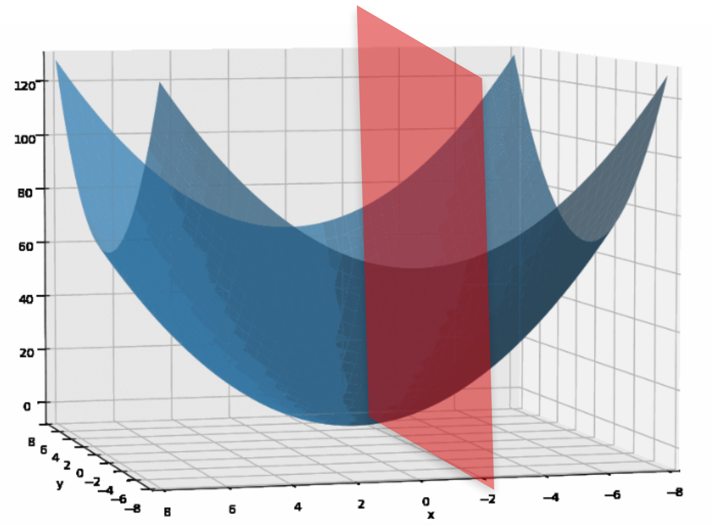
# Slicing the Space

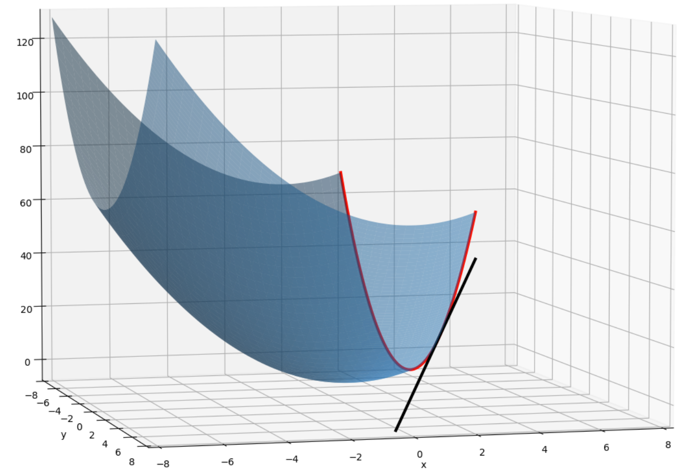


# Slicing the Space

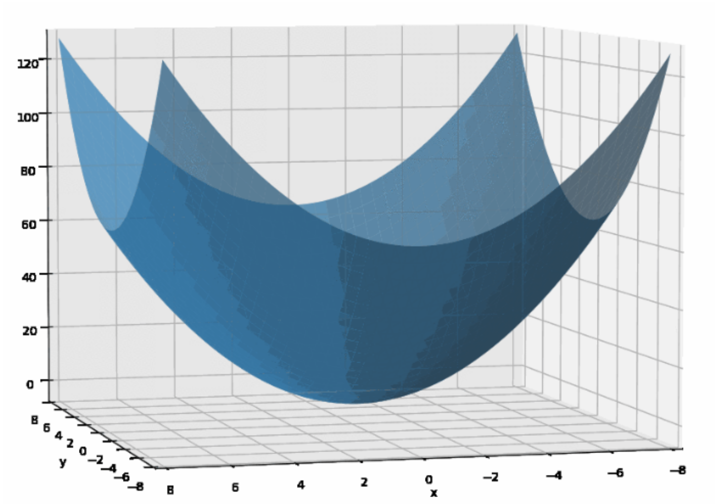


# Slicing the Space

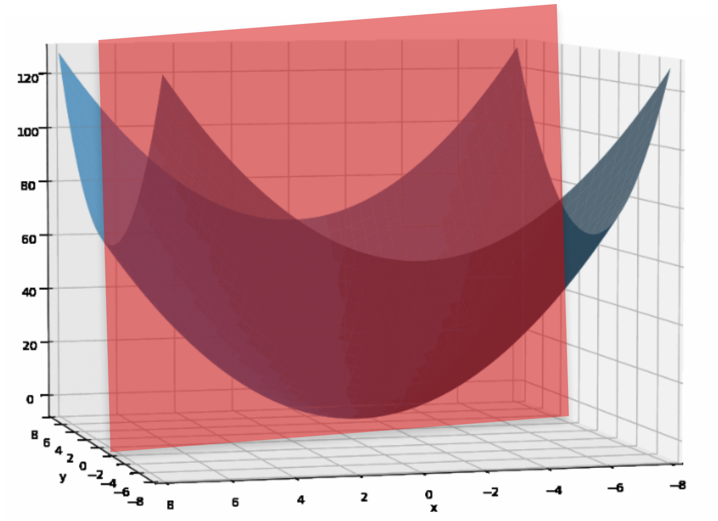


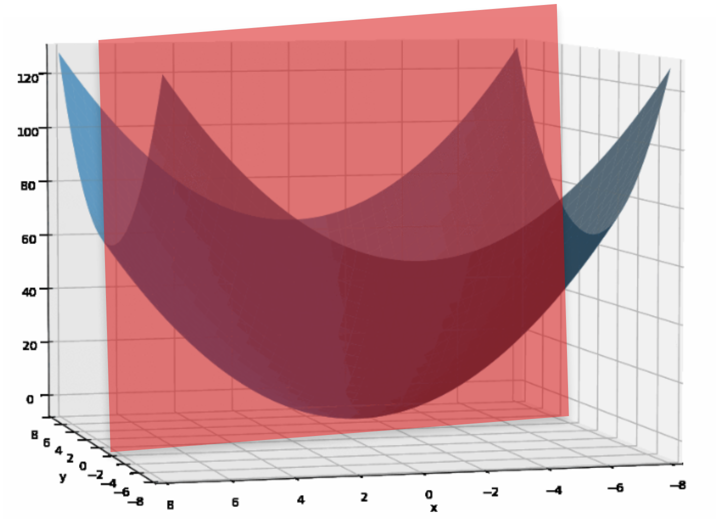
# Slicing the Space

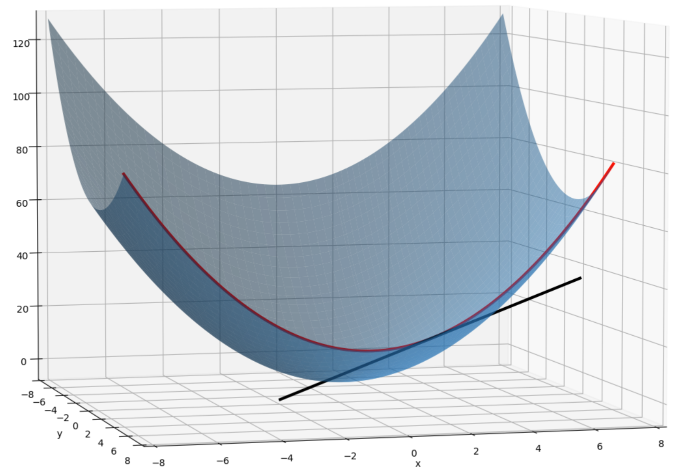


# Slicing the Space



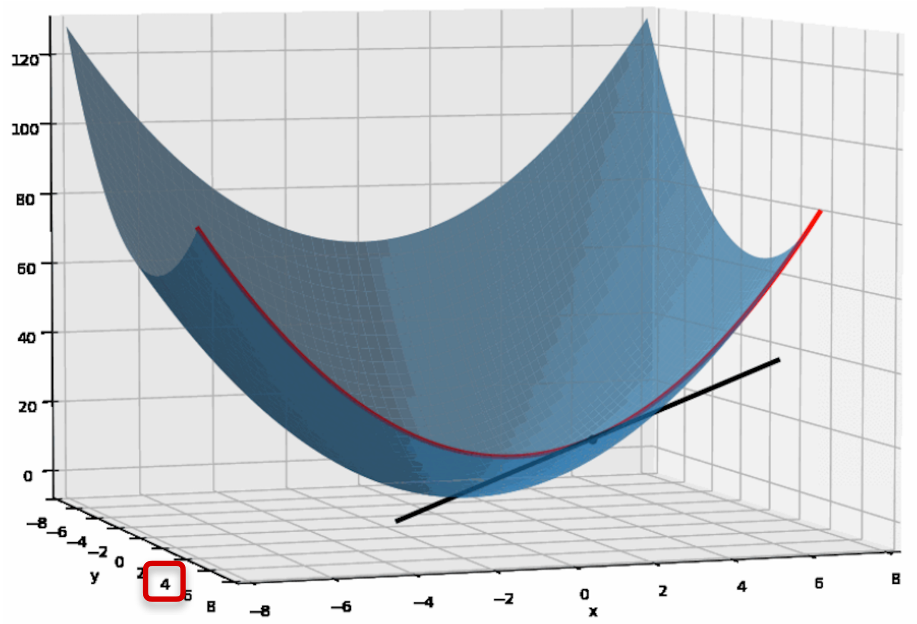
# Slicing the Space



# Partial Derivatives

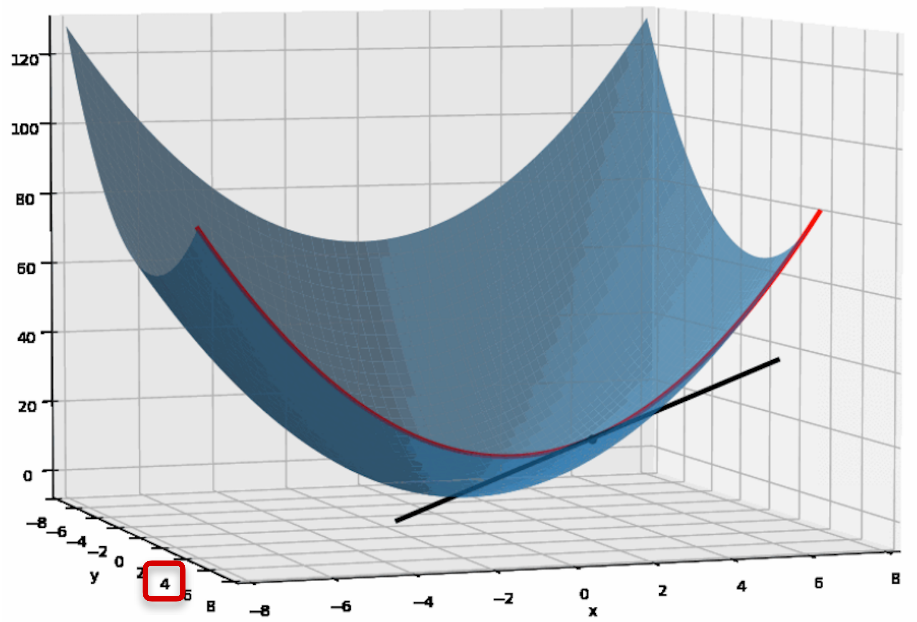
$$f(x, y) = x^2 + y^2$$



# Partial Derivatives

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

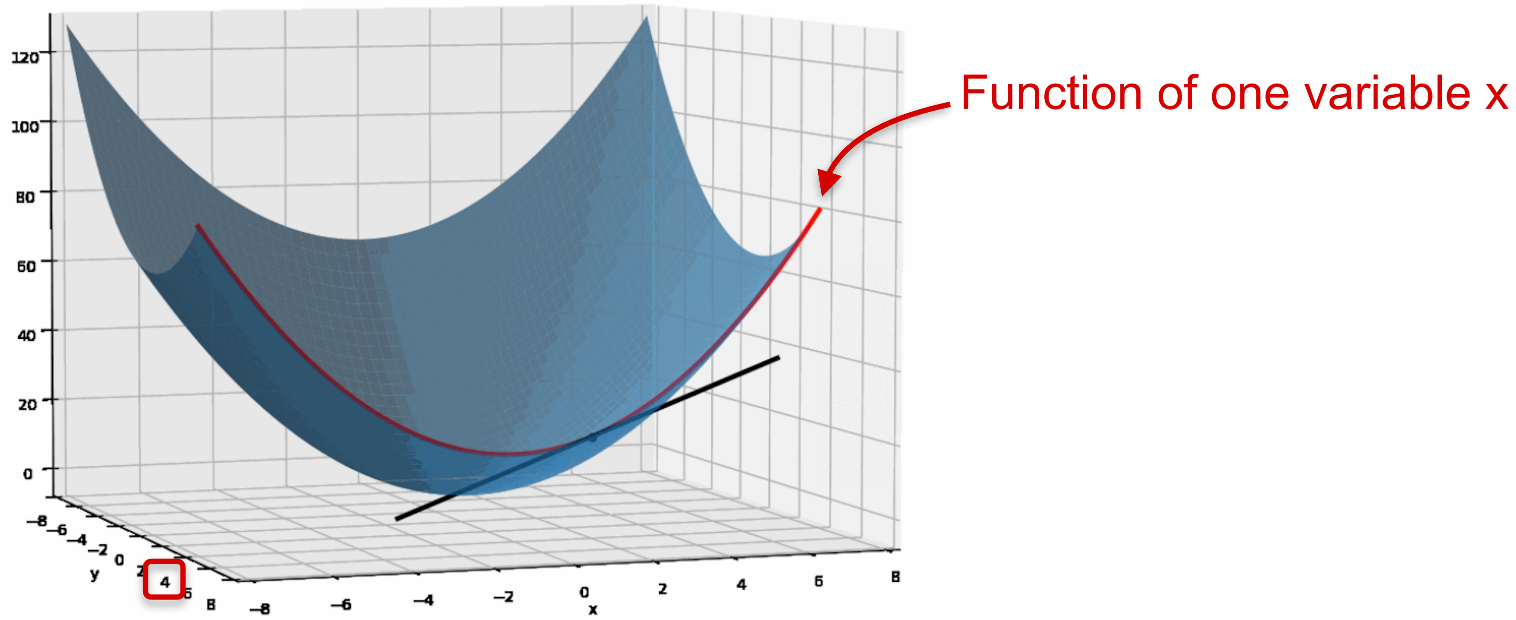
Treat y as a constant



# Partial Derivatives

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

Treat y as a constant

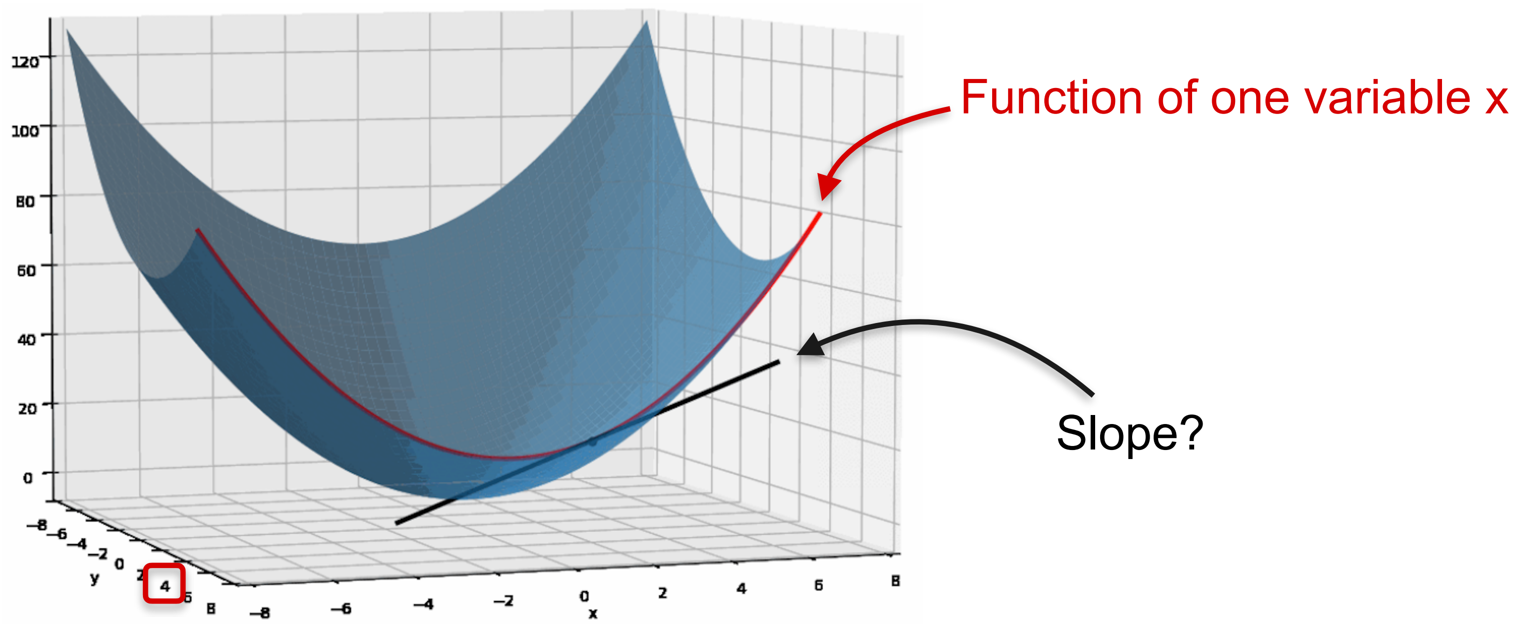



DeepLearning.AI

# Partial Derivatives

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

Treat y as a constant

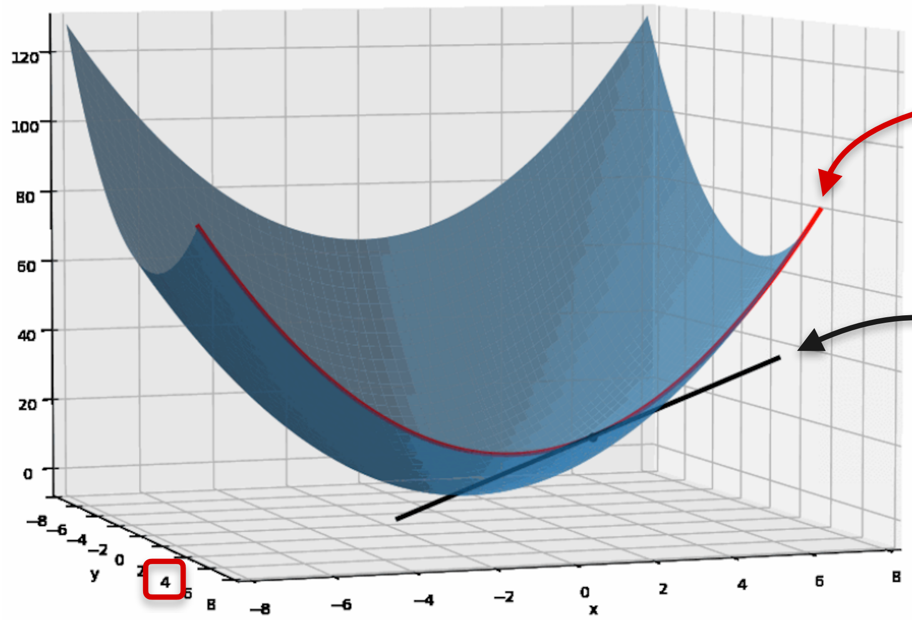



DeepLearning.AI

# Partial Derivatives

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

Treat y as a constant




Function of one variable x

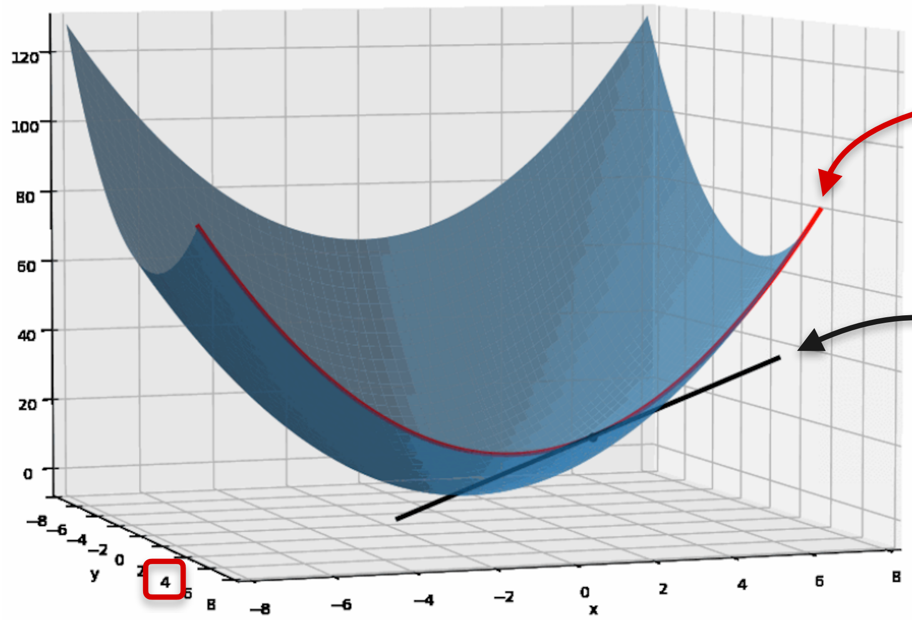
Slope?

Partial derivative

# Partial Derivatives

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

Treat y as a constant




Function of one variable x

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

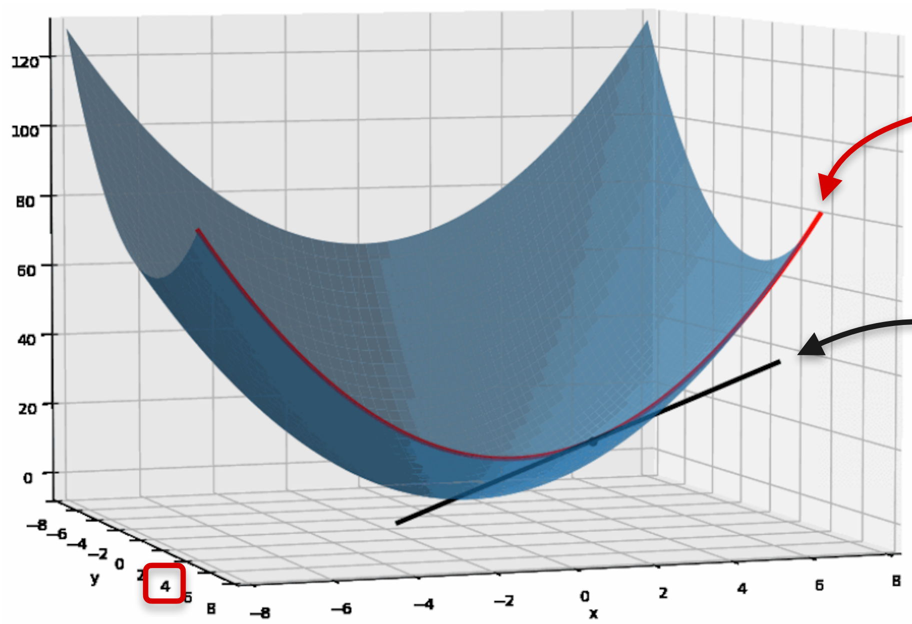
Slope?

Partial derivative

# Partial Derivatives

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

Treat y as a constant




Function of one variable x

Constant

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

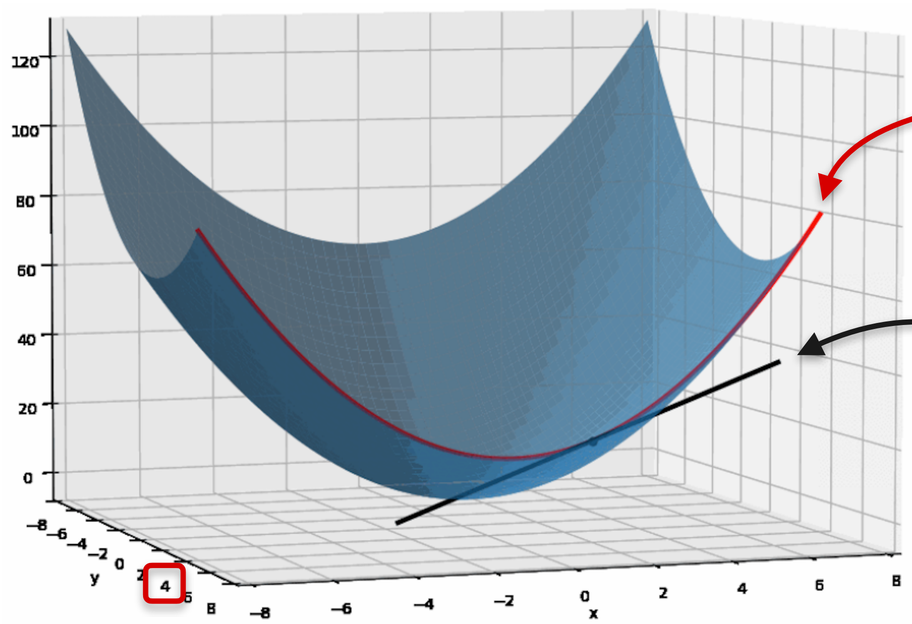
Slope?

Partial derivative

# Partial Derivatives

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

Treat y as a constant




Function of one variable x

Slope?

Partial derivative

Constant

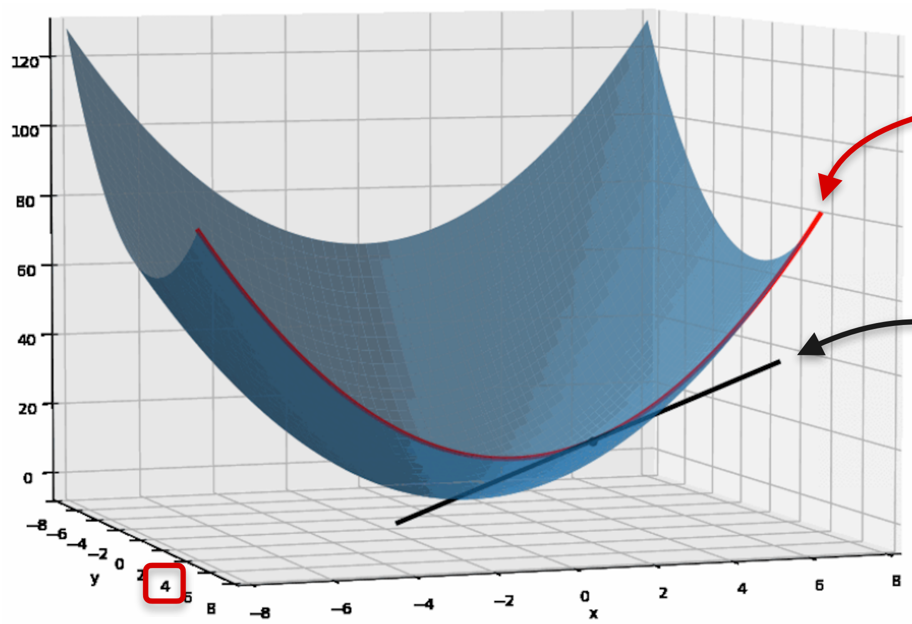
$f(x, y) = x^2 + \boxed{y^2}$

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

# Partial Derivatives

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

Treat y as a constant




Function of one variable x

Slope?

Partial derivative

Constant

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

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

Derivative = 0

# Partial Derivatives

# Partial Derivatives

$$x^2 + y^2$$

# Partial Derivatives

$$x^2 + y^2 \xrightarrow{\boxed{\text{Treat y as a constant}}} 2x$$

# Partial Derivatives

$$x^2 + y^2$$

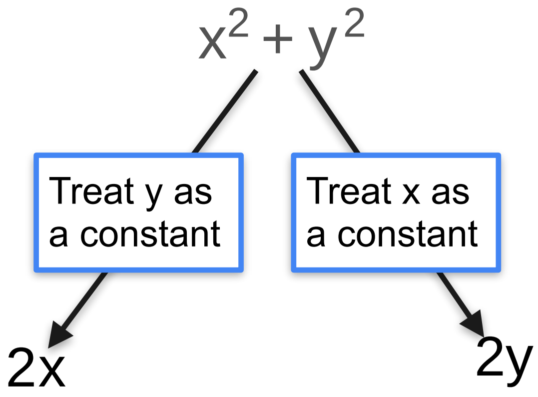
Treat y as a constant

Treat x as a constant

$2x$

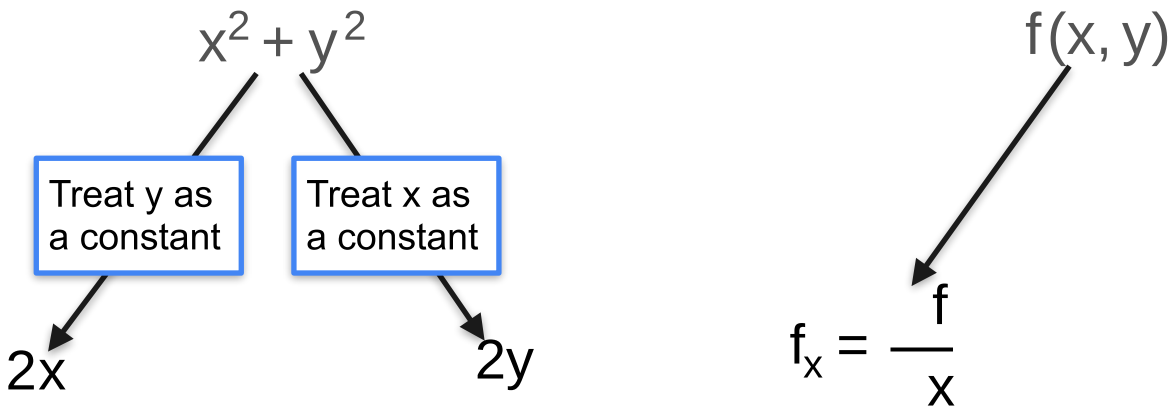
$2y$

# Partial Derivatives




$f(x, y)$

# Partial Derivatives



# Partial Derivatives

$$x^2 + y^2$$

Treat y as a constant → $2x$

Treat x as a constant → $2y$

$$f(x, y)$$

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

# Partial Derivatives

$x^2 + y^2$

Treat y as a constant → $2x$

Treat x as a constant → $2y$

$f(x, y)$

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

Partial derivative of f with respect to x

# Partial Derivatives

$x^2 + y^2$

Treat y as a constant

Treat x as a constant

$2x$

$2y$

$f(x, y)$

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

Partial derivative of f with respect to x

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

Partial derivative of f with respect to x

# Intro To Partial Derivatives

# Intro To Partial Derivatives

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

# Intro To Partial Derivatives

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

**TASK**

# Intro To Partial Derivatives

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

**TASK**

Find partial derivatives of $f$ with respect to $x$ and $y$

# Intro To Partial Derivatives

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

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

**TASK**

Find partial derivatives of $f$ with respect to $x$ and $y$

# Intro To Partial Derivatives

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

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

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

**TASK**

Find partial derivatives of $f$ with respect to $x$ and $y$

# Intro To Partial Derivatives

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

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

?

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

**TASK**

Find partial derivatives of $f$ with respect to $x$ and $y$

# Intro To Partial Derivatives

Partial derivative notation $f(x, y) = x^2 + y^2$

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

?

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

**TASK**

Find partial derivatives of f with respect to x and y

# Intro To Partial Derivatives

Partial derivative notation

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

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

?

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

**TASK**

Find partial derivatives of f with respect to x and y

# Intro To Partial Derivatives

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

**TASK**

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

?

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

# Intro To Partial Derivatives

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

**TASK**

Find partial derivative of f with respect to x

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

?

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

# Intro To Partial Derivatives

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

**TASK**

Find partial derivative of f with respect to x

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

?

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

**Step 1:**

# Intro To Partial Derivatives

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

**TASK**

Find partial derivative of f with respect to x

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

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

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

# Intro To Partial Derivatives

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

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

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

?

**TASK**

Find partial derivative of f with respect to x

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

**Step 2:**

# Intro To Partial Derivatives

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

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

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

**TASK**

Find partial derivative 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.

# Intro To Partial Derivatives

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

**TASK**

Find partial derivative 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.

# Intro To Partial Derivatives

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

**TASK**

Find partial derivative 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.

# Intro To Partial Derivatives

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

**TASK**

Find partial derivative 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.

# Intro To Partial Derivatives

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

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

**TASK**

Find partial derivative 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.

# Intro To Partial Derivatives

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

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

**TASK**

Find partial derivative of f with respect to x

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

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

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

# Intro To Partial Derivatives

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

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

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

**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.

# Intro To Partial Derivatives

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

**TASK**

Find partial derivative of f with respect to x

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

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

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

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

# Intro To Partial Derivatives

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

**TASK**

Find partial derivative of f with respect to x

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

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

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

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

# Intro To Partial Derivatives

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

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

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

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

**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.

# Intro To Partial Derivatives

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

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

## 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.

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

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

DeepLearning.AI

# Gradients and Gradient Descent

## Partial derivatives -Part 2

# Partial Derivatives (More Examples)

# Partial Derivatives (More Examples)

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

# Partial Derivatives (More Examples)

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

**TASK**

What is the partial derivate of $f$ with respect to $x$?

# Partial Derivatives (More Examples)

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

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

**TASK**

What is the partial derivate of $f$ with respect to $x$?

# Partial Derivatives (More Examples)

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

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

**TASK**

Find partial derivate of $f$ with respect to $x$

# Partial Derivatives (More Examples)

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

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

**TASK**

Find partial derivate of $f$ with respect to $x$

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

# Partial Derivatives (More Examples)

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

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

**TASK**

Find partial derivate of $f$ with respect to $x$

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

# Partial Derivatives (More Examples)

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

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

**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.

# Partial Derivatives (More Examples)

$$f(x, y) = 3x^2 \text{[illegible]}$$

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

**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.

# Partial Derivatives (More Examples)

$$f(x, y) = 3x^2 \text{[illegible]}$$

Constant coefficient

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

**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.

# Partial Derivatives (More Examples)

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

Constant coefficient

$$\frac{f}{x} = 3$$

**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.

# Partial Derivatives (More Examples)

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

$$\frac{f}{x} = 3$$

**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.

# Partial Derivatives (More Examples)

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

Differentiate with respect to x

$$\frac{f}{x} = 3$$

**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.

# Partial Derivatives (More Examples)

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

Differentiate with respect to x

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

**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.

# Partial Derivatives (More Examples)

$$f(x, y) = 3x^2 \text{ [illegible]}$$

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

**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.

# Partial Derivatives (More Examples)

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

treat as constant coefficient

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

**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.

# Partial Derivatives (More Examples)

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

treat as constant coefficient

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

**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.

# Partial Derivatives (More Examples)

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

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

**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.

# Partial Derivatives (More Examples)

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

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

$$= 6xy^3$$

**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.

# Partial Derivatives (More Examples)

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

**TASK**

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

# Partial Derivatives (More Examples)

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

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

**TASK**

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

# Partial Derivatives (More Examples)

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

$$\frac{f}{y} = \ ?$$

**TASK**

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

# Partial Derivatives (More Examples)

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

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

**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.

# Partial Derivatives (More Examples)

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

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

**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.

# Partial Derivatives (More Examples)

$$f(x, y) = 3 [illegible] y^3$$

$$\frac{f}{y} = 3$$

**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.

# Partial Derivatives (More Examples)

$$f(x, y) = 3 \quad y^3$$

$$\frac{f}{y} = 3$$

**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.

# Partial Derivatives (More Examples)

$$f(x, y) = 3\;\; y^3$$

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

**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.

# Partial Derivatives (More Examples)

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

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

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

**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.

# Partial Derivatives (More Examples)

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

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

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

**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.

# Gradients and Gradient Descent

## Gradients

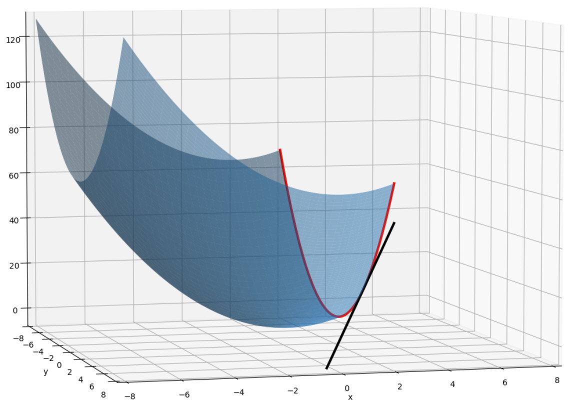
# Gradient

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

# Gradient

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

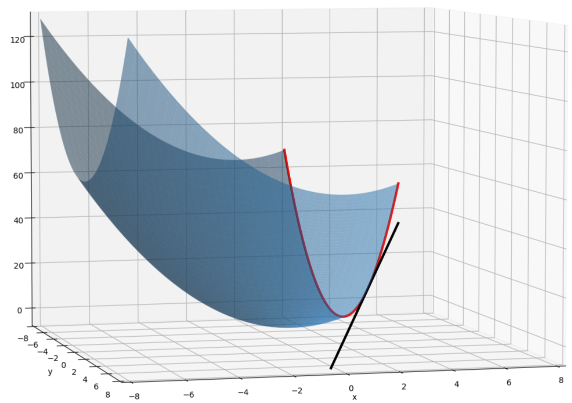
Treat y as a constant



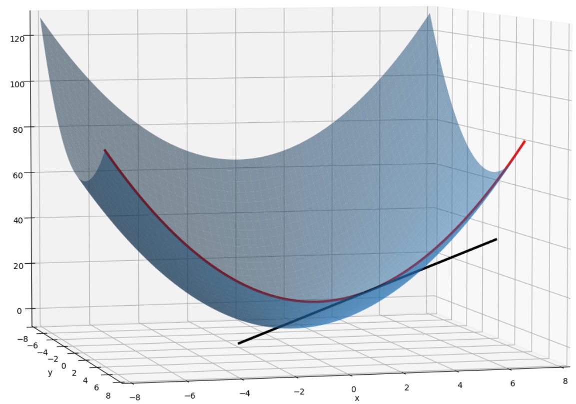
# Gradient

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

Treat y as a constant



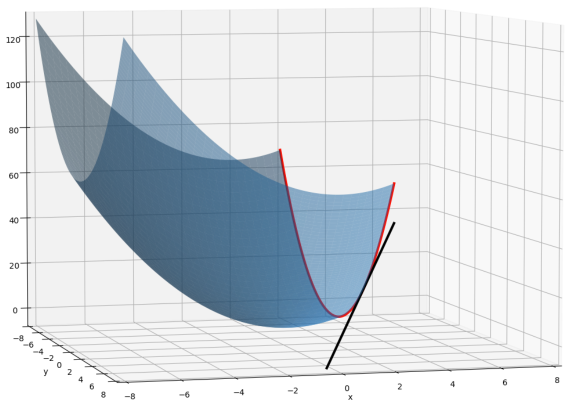
Treat x as a constant



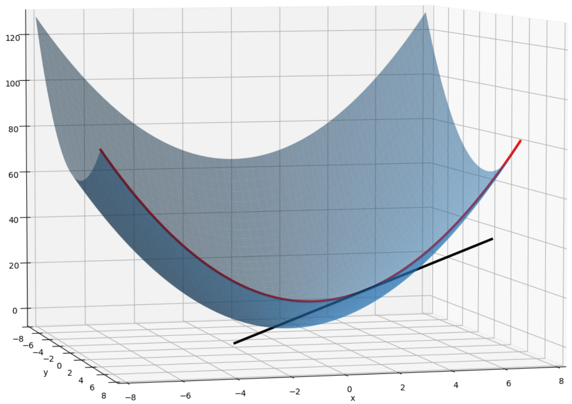
# Gradient

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

Treat y as a constant



Treat x as a constant

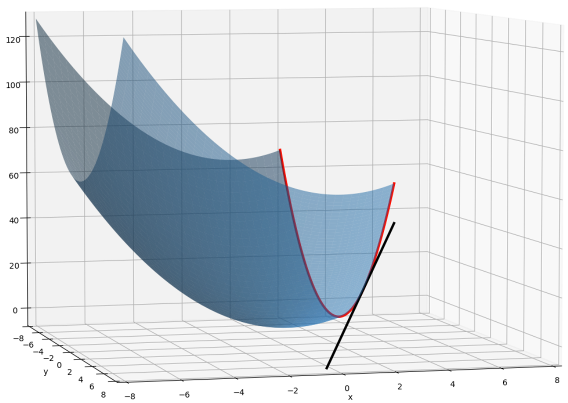


$$\frac{f}{x} = 2x$$

# Gradient

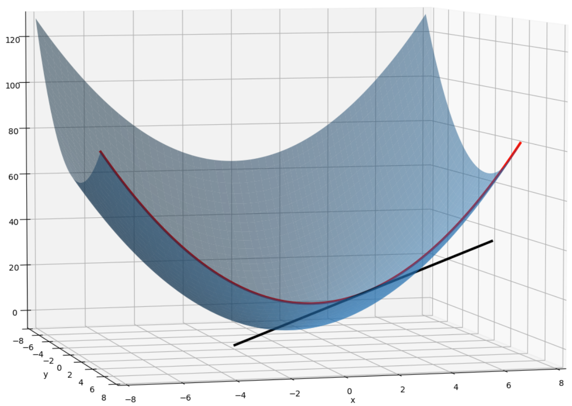
$$f(x, y) = x^2 + y^2$$

Treat y as a constant



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

Treat x as a constant



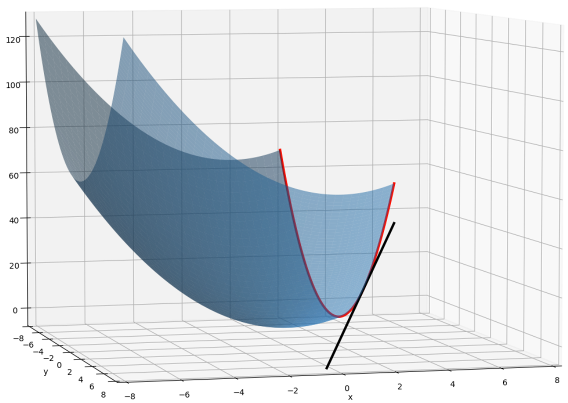
$$\frac{f}{y} = 2y$$

# Gradient

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

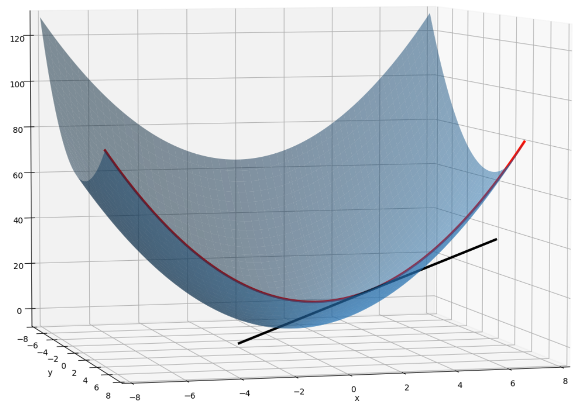
Gradient

Treat y as a constant



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

Treat x as a constant

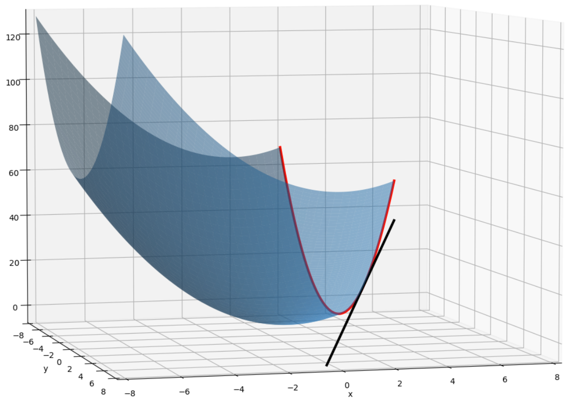


$$\frac{f}{y} = 2y$$

# Gradient

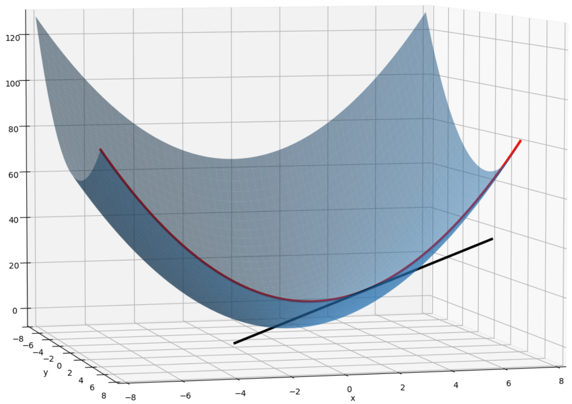
$$f(x, y) = x^2 + y^2$$

Treat y as a constant



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

Treat x as a constant



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

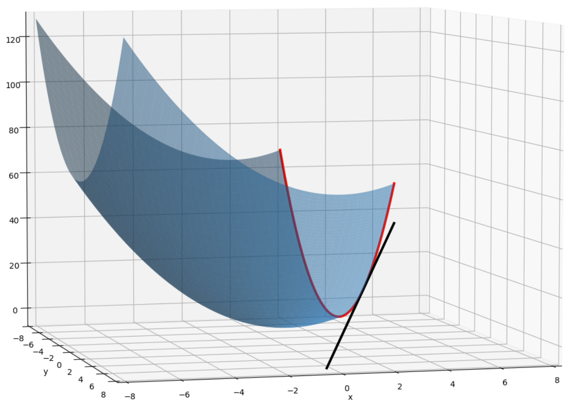
Gradient

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

# Gradient

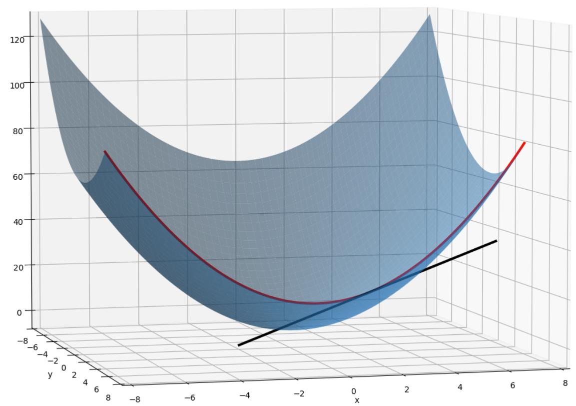
$$f(x, y) = x^2 + y^2$$

Treat y as a constant



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

Treat x as a constant



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

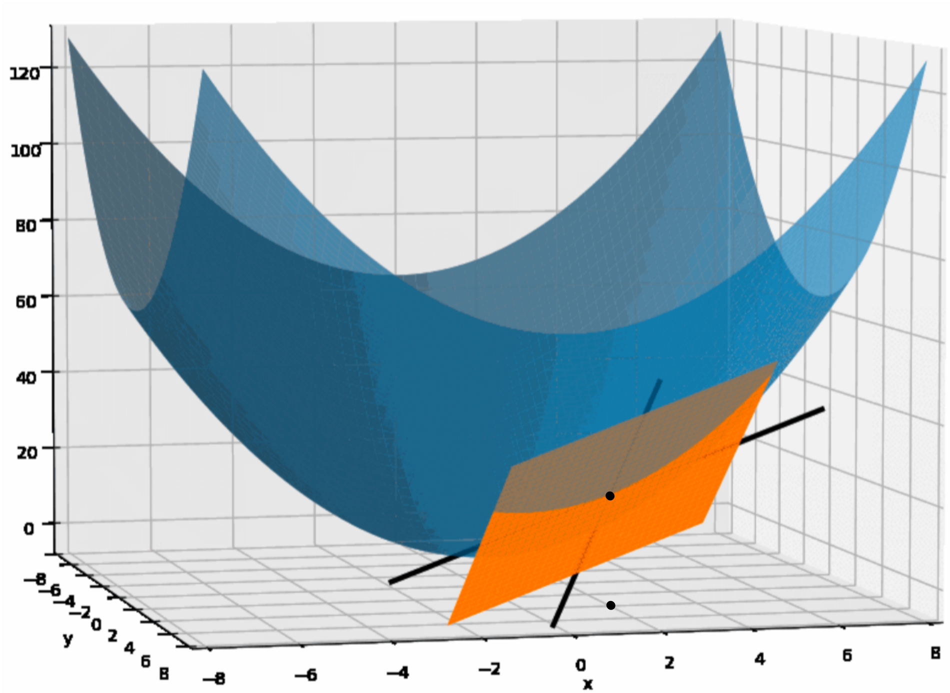
Gradient

2x

[ 2y]

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

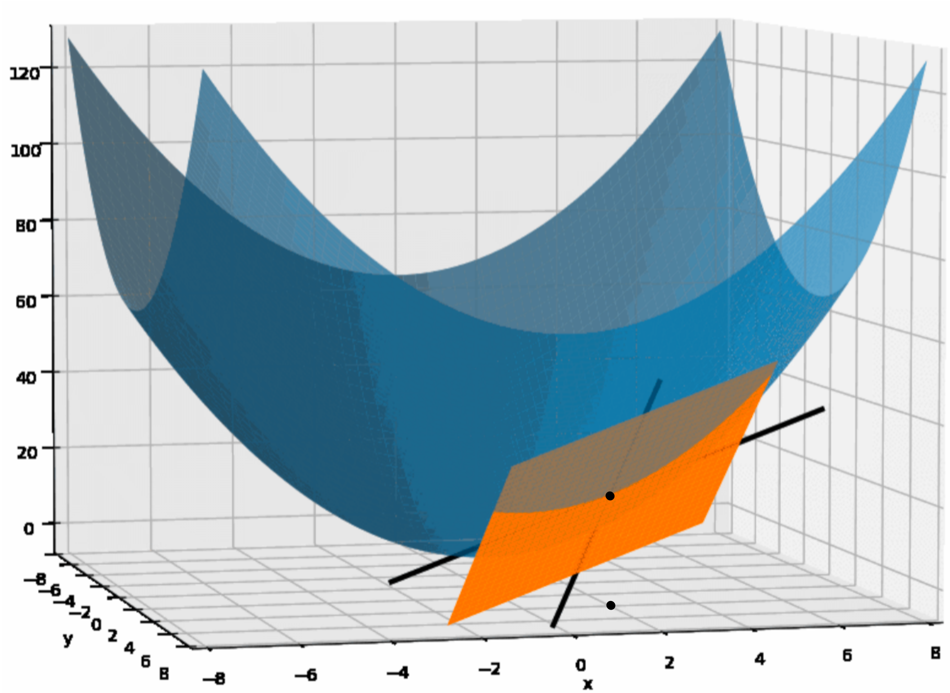
# Gradient




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

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

# Gradient



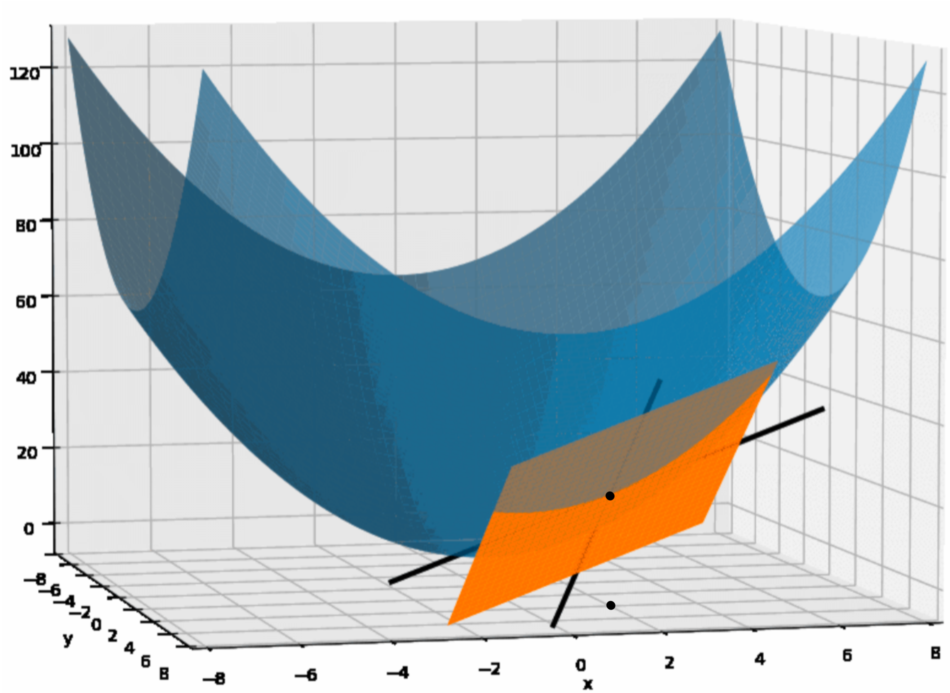

$f(x, y) = x^2 + y^2$

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

**TASK**

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

# Gradient




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

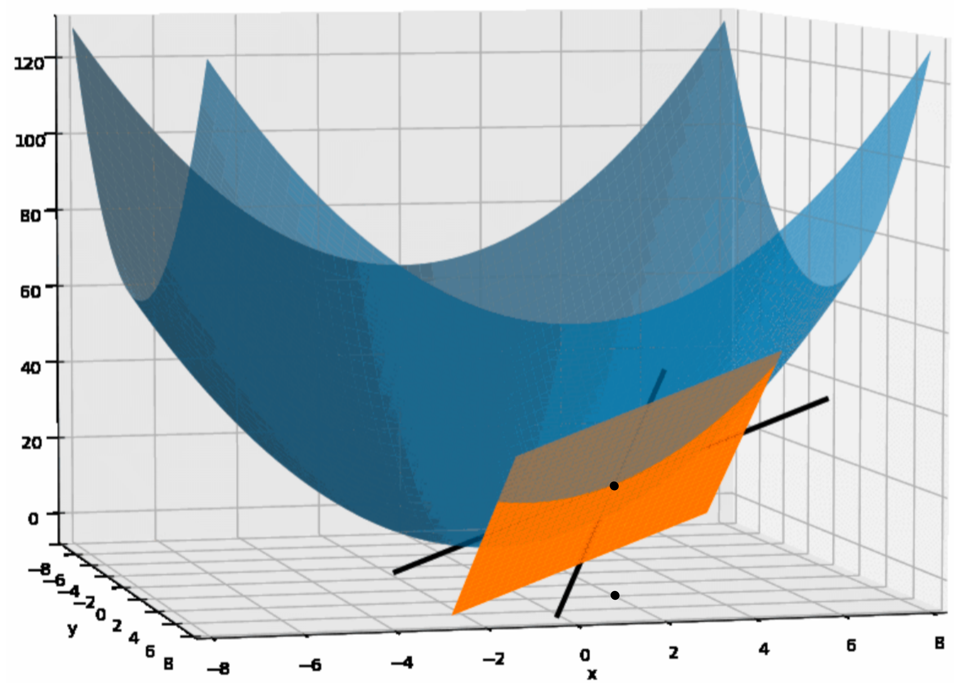
The gradient of $f(x,y)$ is: $f = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$

**TASK**

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

The gradient of $f(x,y)$ is given as:

# Gradient




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

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

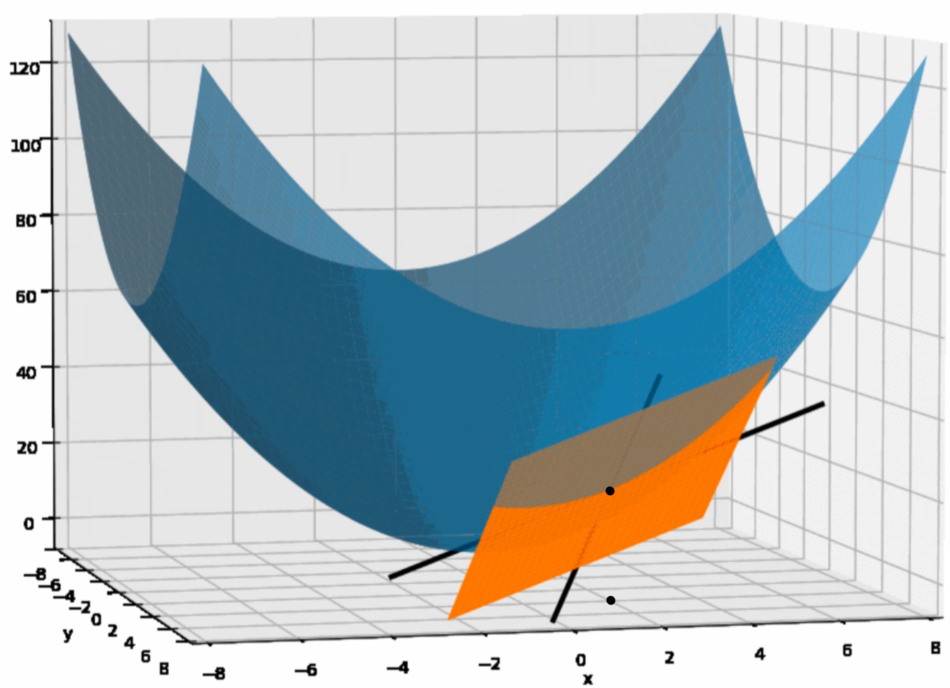
**TASK**

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

The gradient of $f(x,y)$ is given as:

$$f = \begin{bmatrix} 2 & 2 \\ 2 & 3 \end{bmatrix}$$

# Gradient




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

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

**TASK**

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

The gradient of $f(x, y)$ is given as:

$$f = \begin{bmatrix} 2 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

DeepLearning.AI

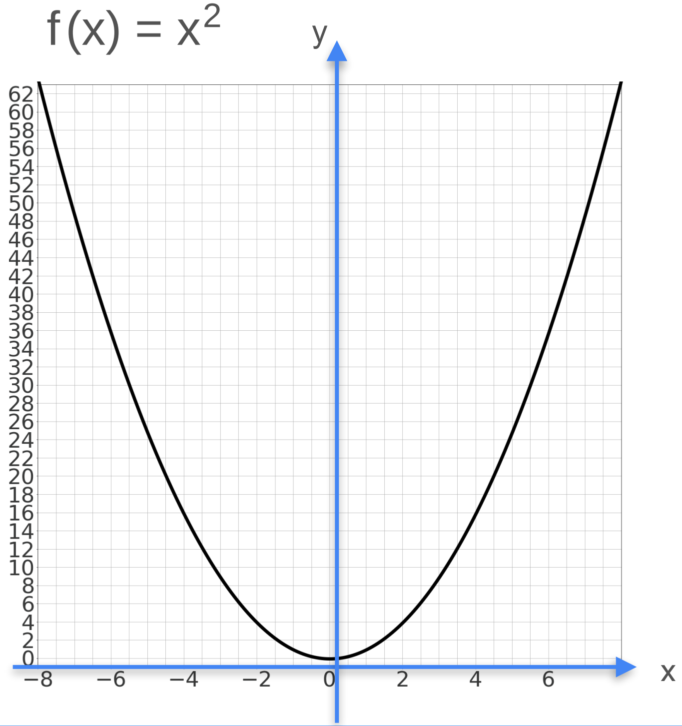
# Gradients and Gradient Descent

## Gradients and maxima/ minima

# Functions of Two Variables

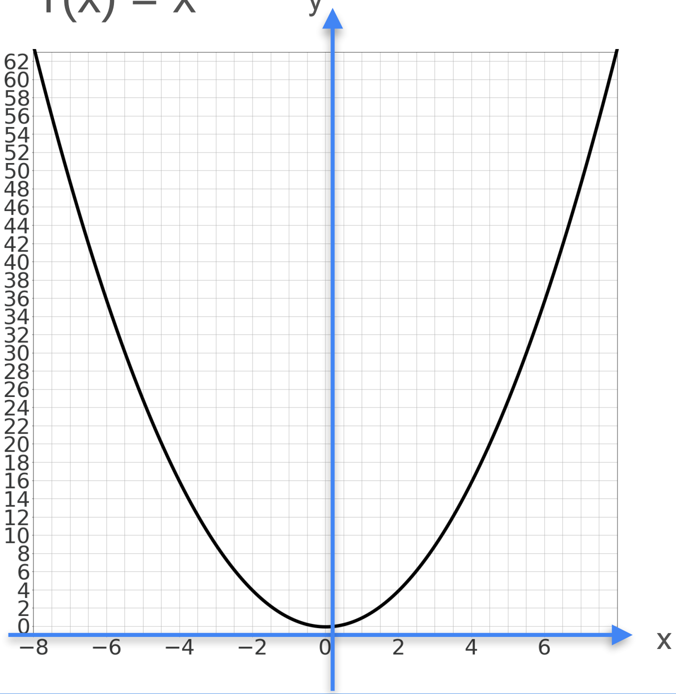
# Functions of Two Variables

$f(x) = x^2$

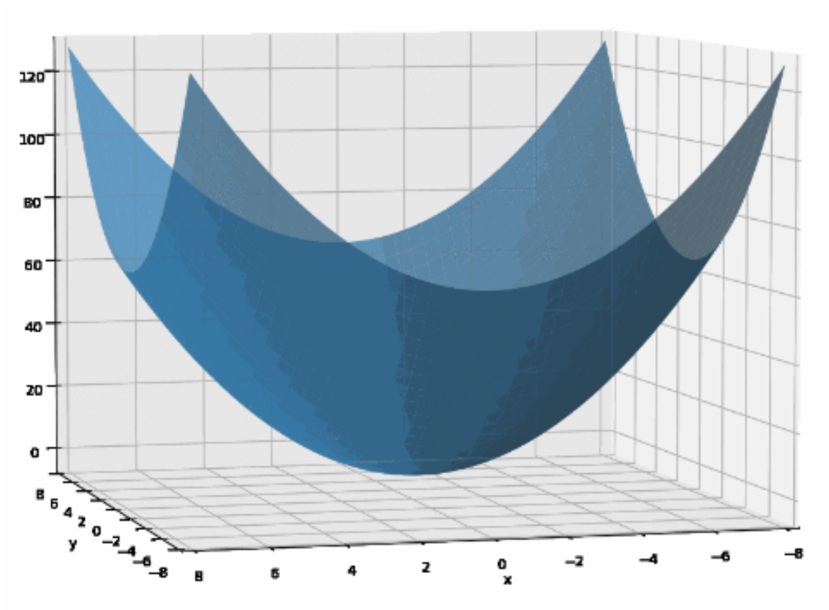


# Functions of Two Variables

$f(x) = x^2$

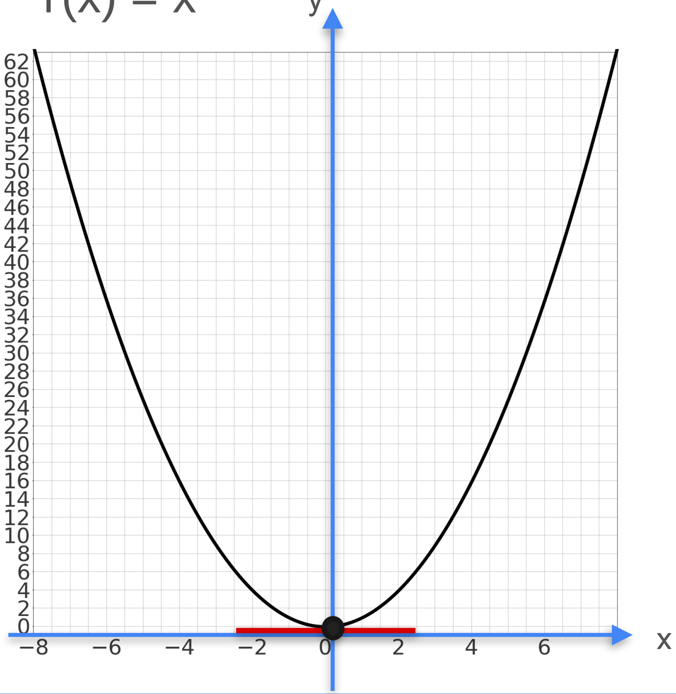



$f(x, y) = x^2 + y^2$

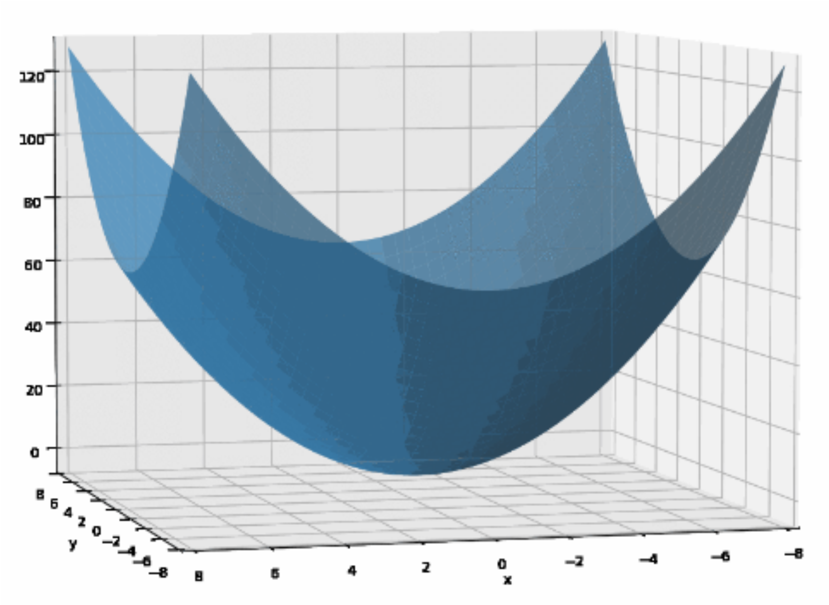


# Functions of Two Variables

$f(x) = x^2$

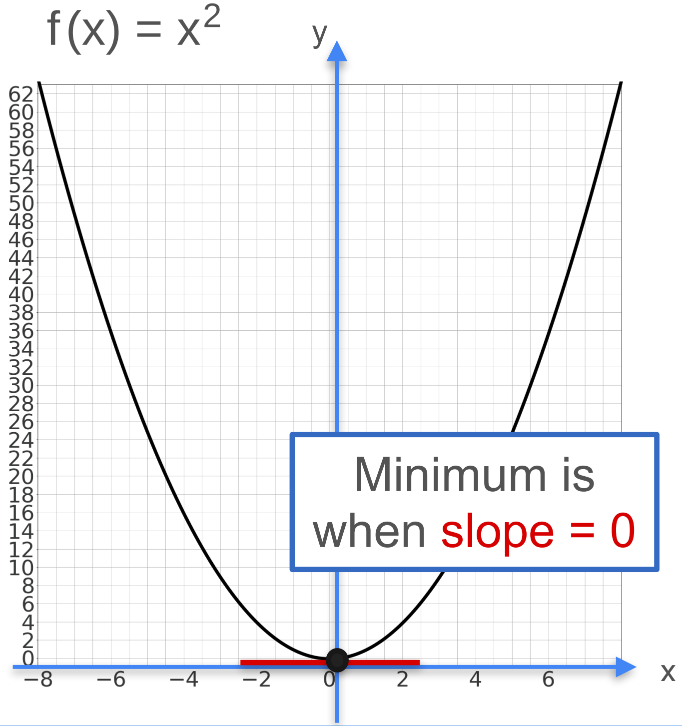



$f(x, y) = x^2 + y^2$

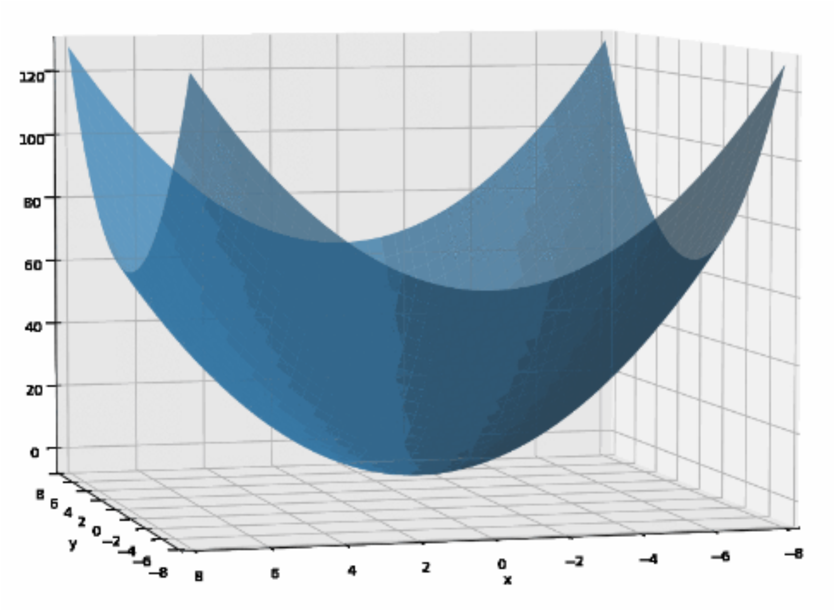


# Functions of Two Variables

$f(x) = x^2$

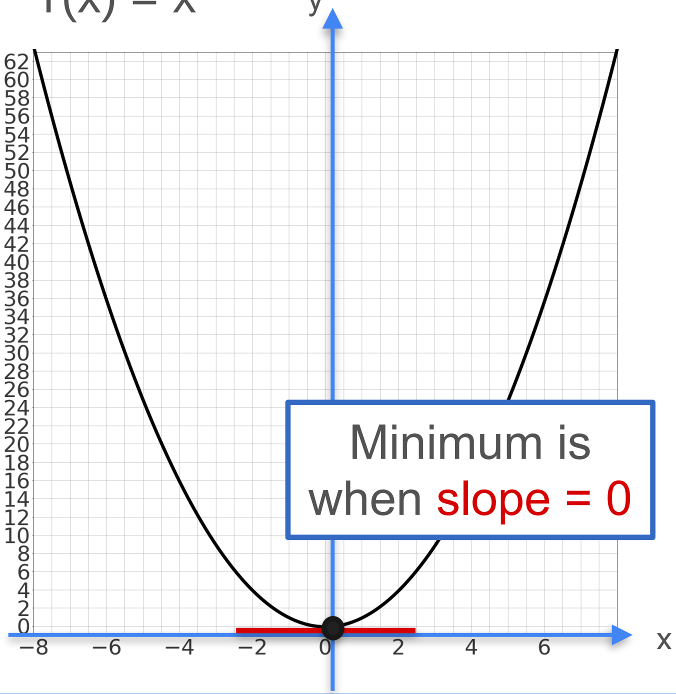



$f(x, y) = x^2 + y^2$

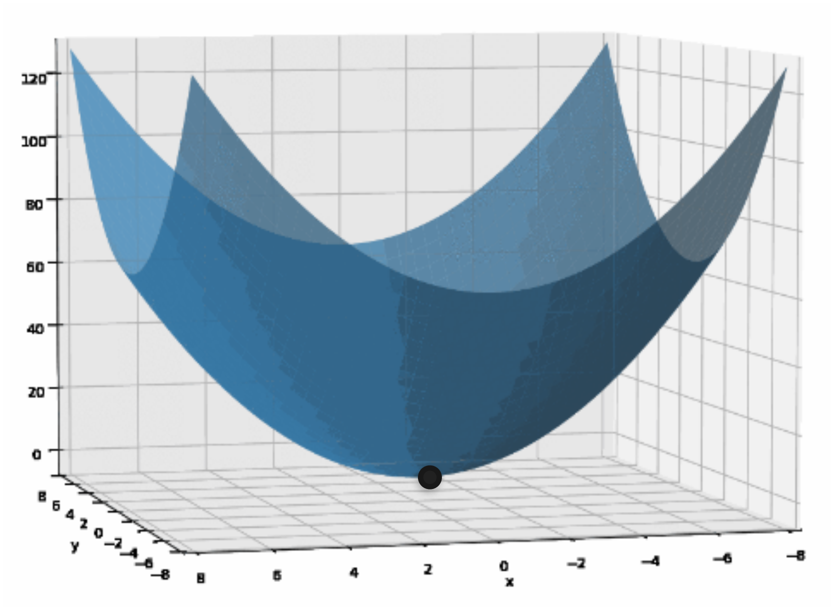


# Functions of Two Variables

$f(x) = x^2$

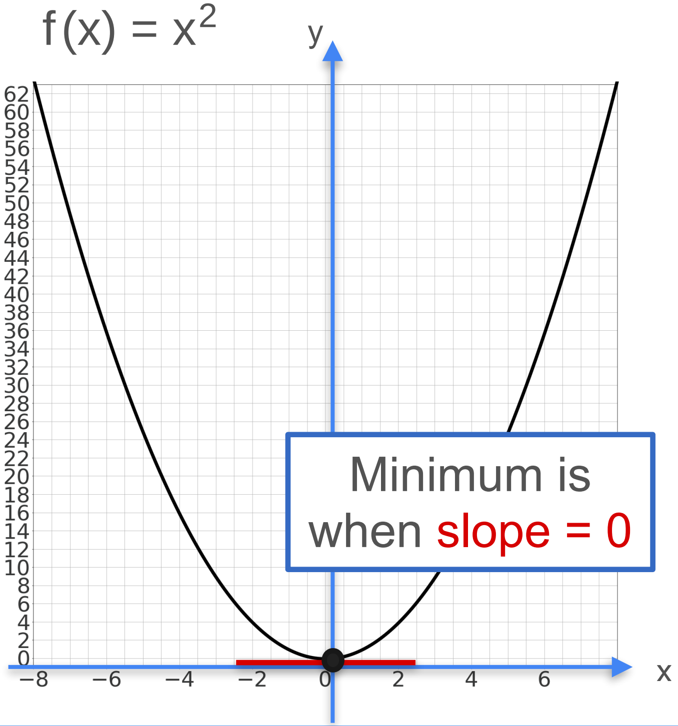


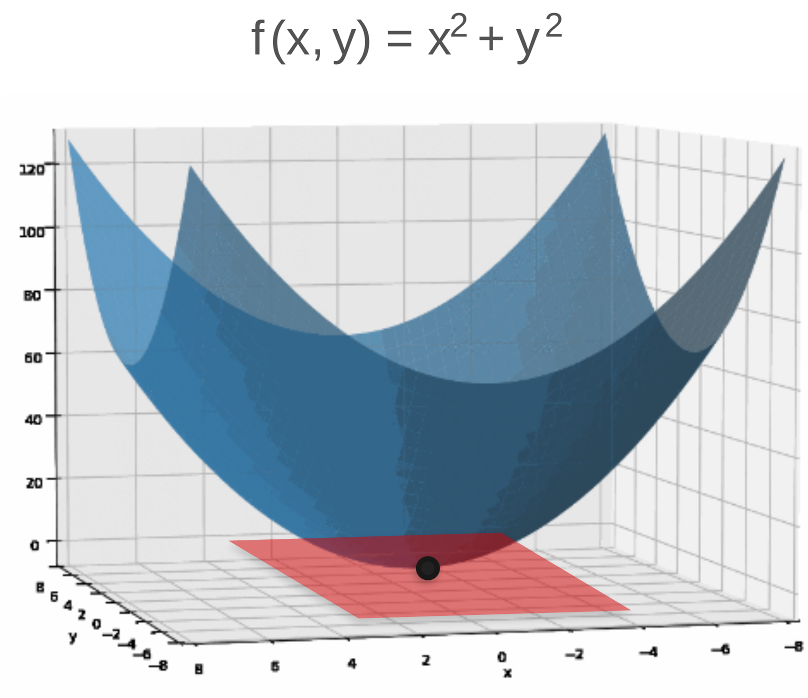

$f(x, y) = x^2 + y^2$




DeepLearning.AI

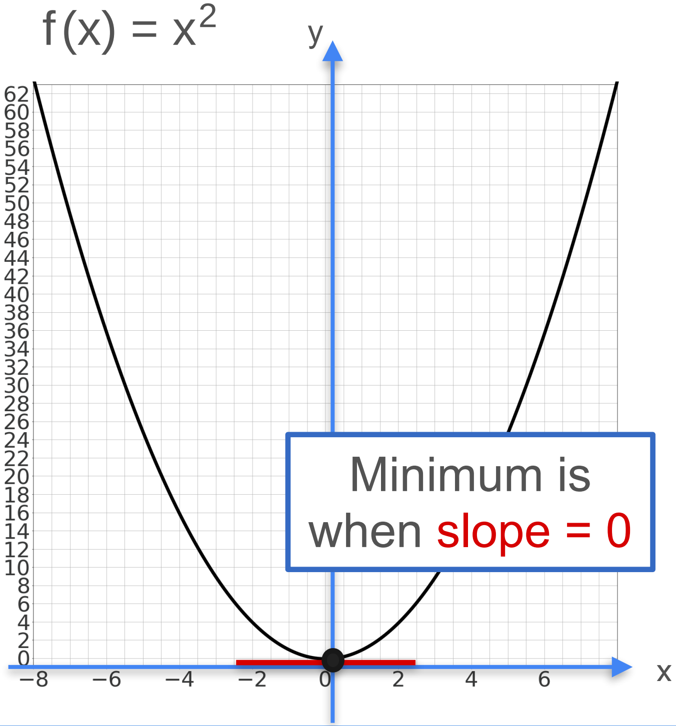
# Functions of Two Variables



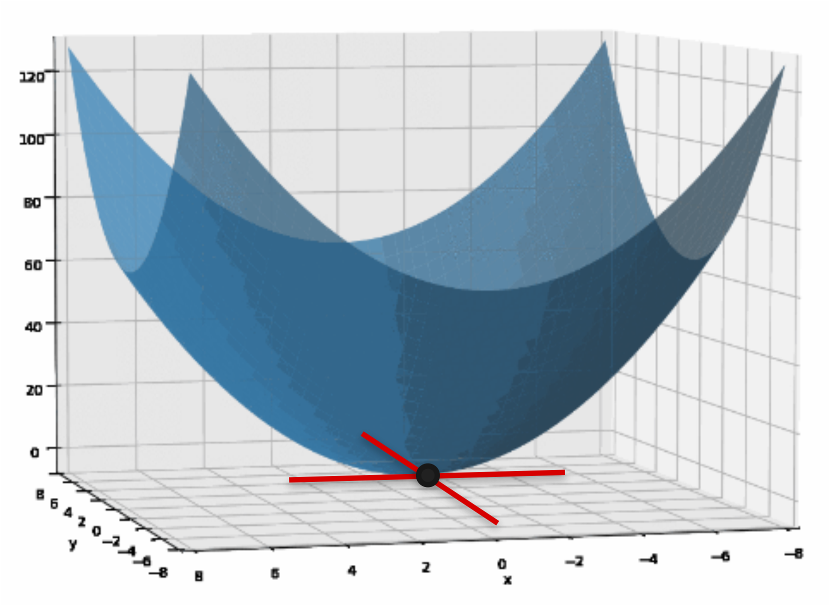



# Functions of Two Variables

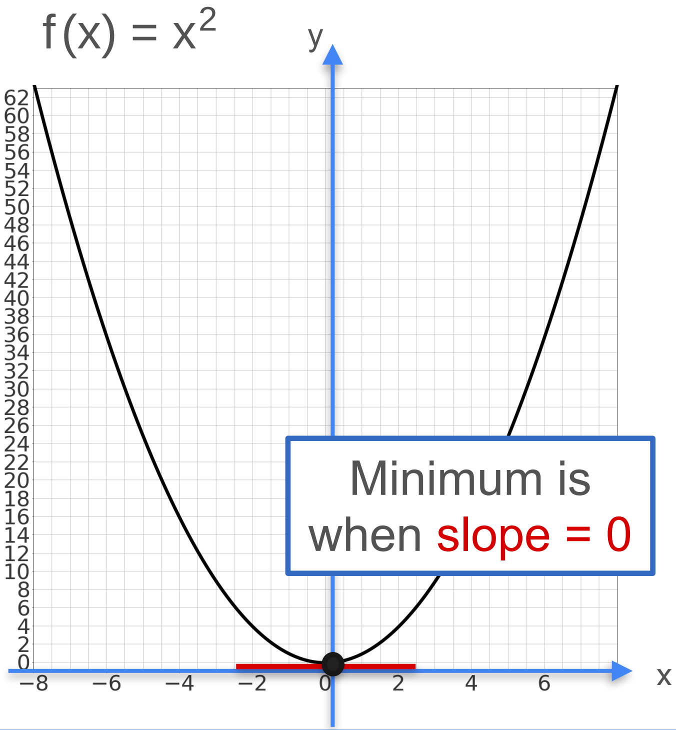
$f(x) = x^2$

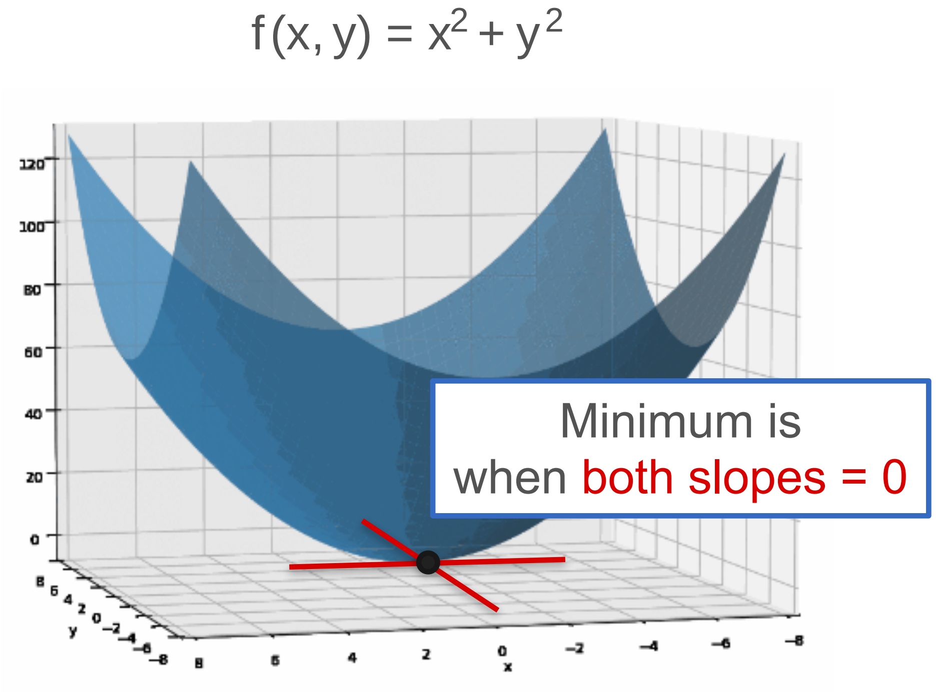



$f(x, y) = x^2 + y^2$



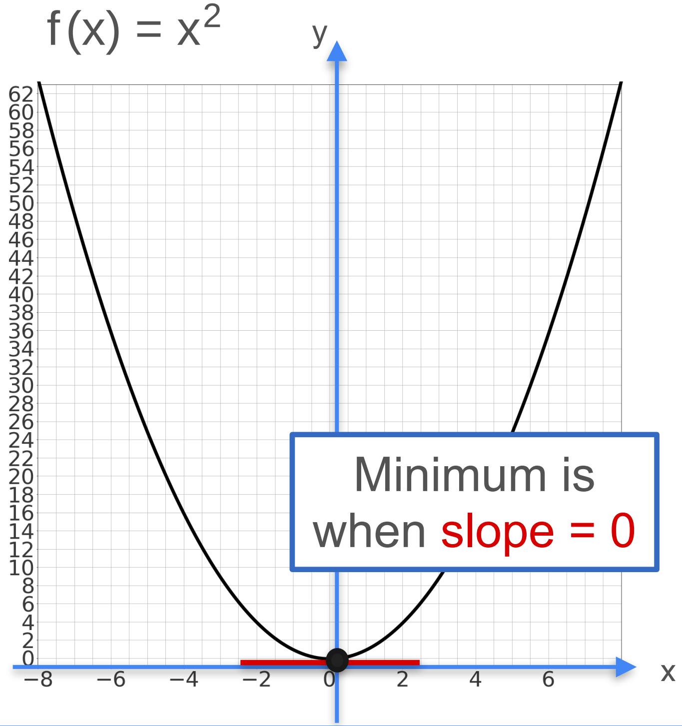
# Functions of Two Variables



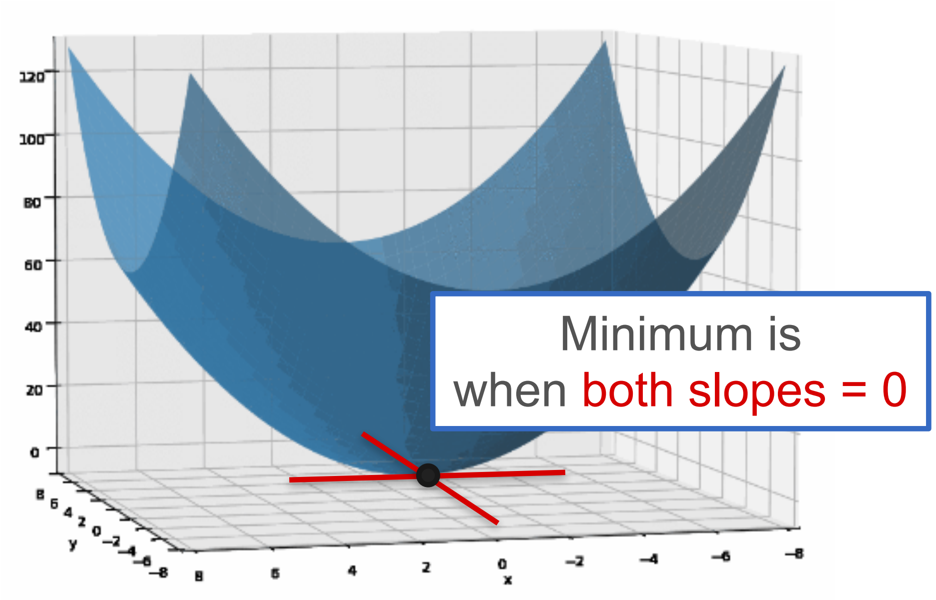



# Functions of Two Variables

$f(x) = x^2$

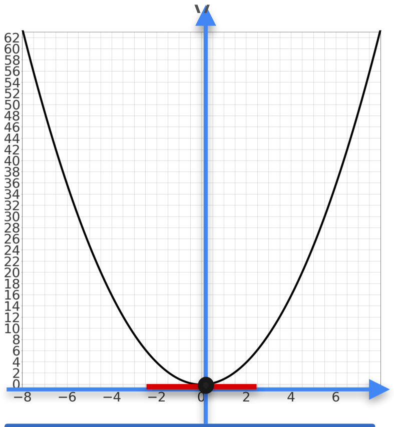



$f(x, y) = x^2 + y^2$



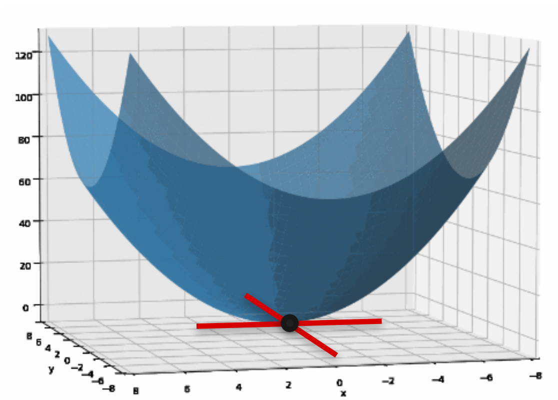
# Functions of Two Variables

$f(x) = x^2$




Minimum is when slope = 0

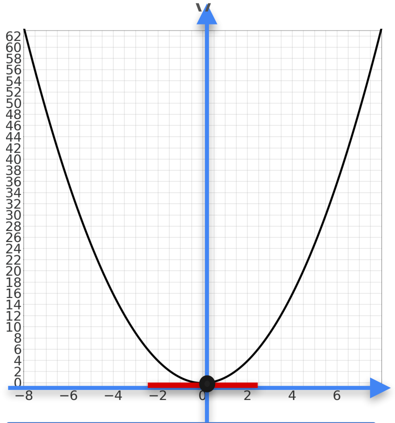
$f(x, y) = x^2 + y^2$




Minimum is when both slopes = 0

# Functions of Two Variables

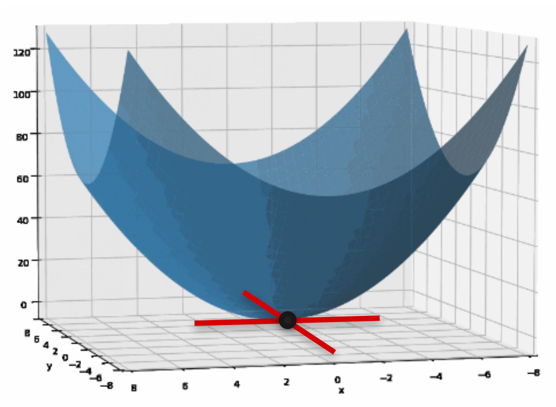
$f(x) = x^2$



$f'(x) = 0$

Minimum is when slope = 0

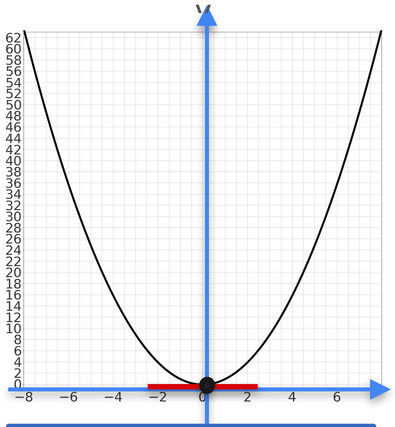
$f(x, y) = x^2 + y^2$



Minimum is when both slopes = 0

# Functions of Two Variables

$f(x) = x^2$

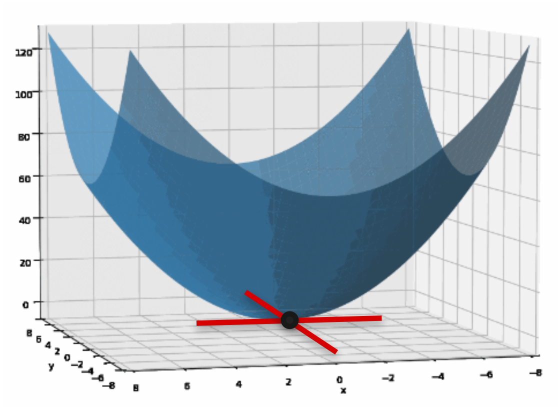



Minimum is when slope = 0

$f\ (x) = 0$

$2x = 0$

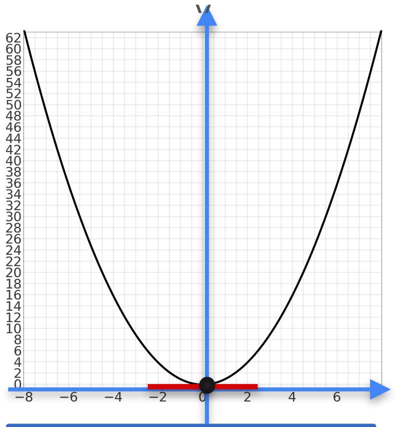
$f(x, y) = x^2 + y^2$



Minimum is when both slopes = 0

# Functions of Two Variables

$f(x) = x^2$



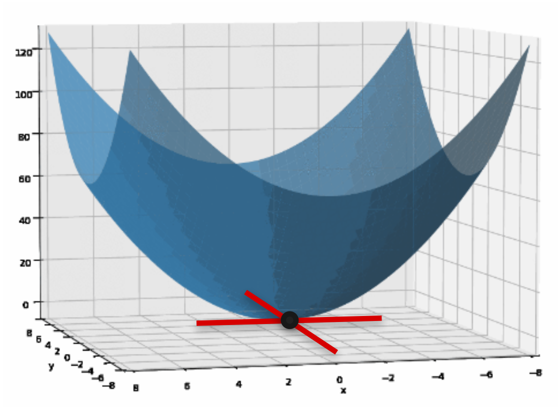

Minimum is when slope = 0

$f\ (x) = 0$

$2x = 0$

$x = 0$

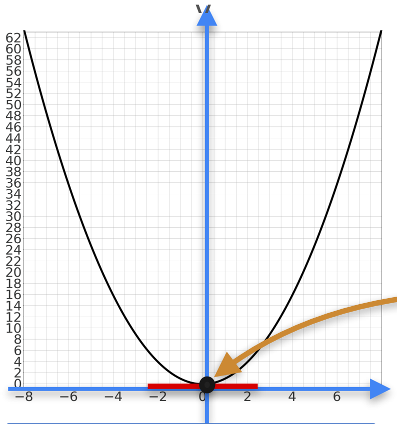
$f(x, y) = x^2 + y^2$



Minimum is when both slopes = 0

# Functions of Two Variables

$f(x) = x^2$



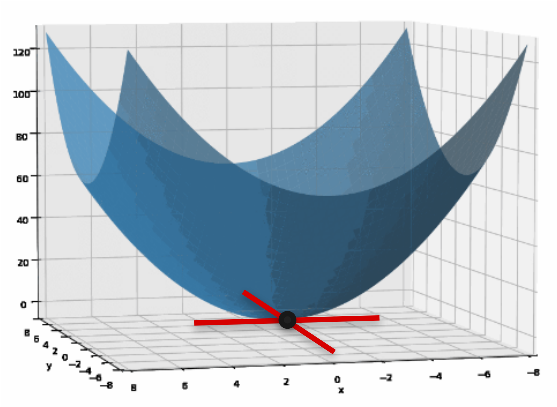

$f(x) = 0$

$2x = 0$

$x = 0$

Minimum is
when slope = 0

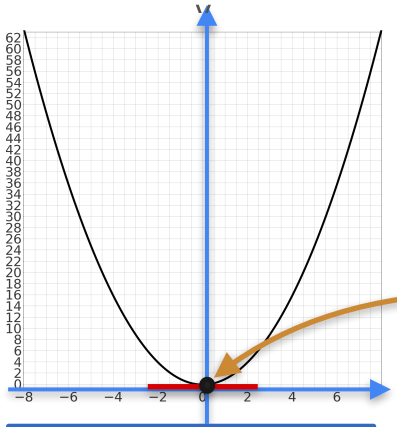
$f(x, y) = x^2 + y^2$




Minimum is
when both slopes = 0

# Functions of Two Variables

$f(x) = x^2$



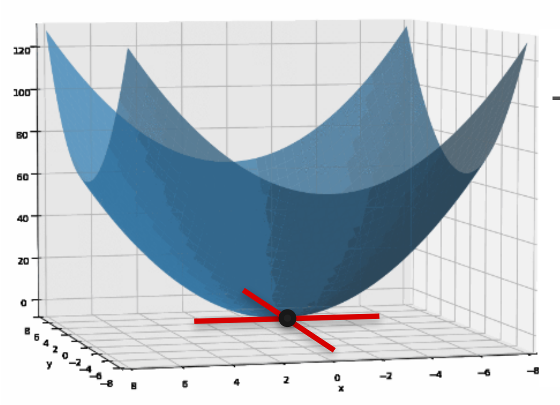
$f'(x) = 0$

$2x = 0$

$x = 0$

Minimum is when slope = 0

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

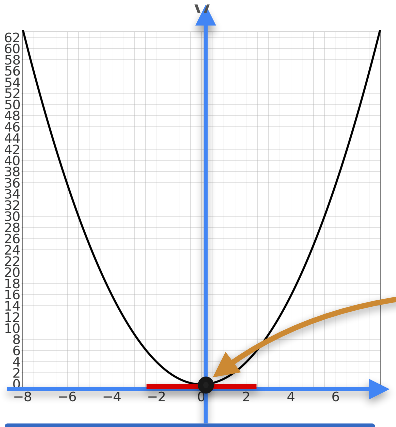


$\frac{f}{x} = 0$ and $\frac{f}{y} = 0$

Minimum is when both slopes = 0

# Functions of Two Variables

$f(x) = x^2$



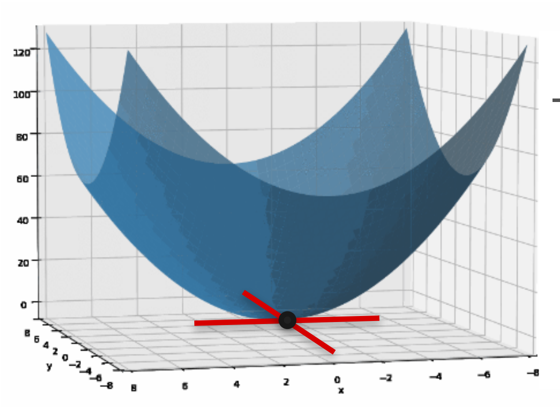
$f\ (x) = 0$

$2x = 0$

$x = 0$

Minimum is when slope = 0

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



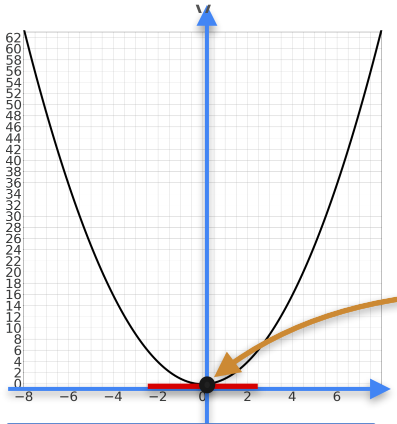
$\frac{f}{x} = 0$ and $\frac{f}{y} = 0$

$2x = 0$ and $2y = 0$

Minimum is when both slopes = 0

# Functions of Two Variables

$f(x) = x^2$



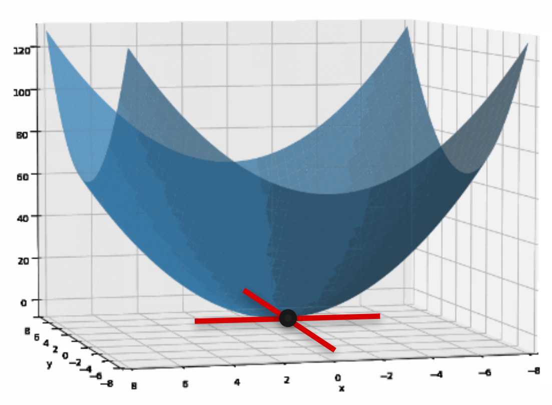
$f(x) = 0$

$2x = 0$

$x = 0$

Minimum is when slope = 0

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



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

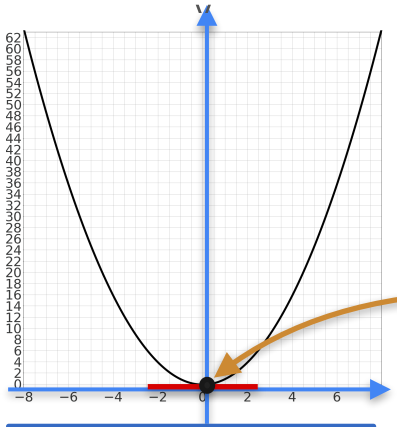
$2x = 0$ and $2y = 0$

$(x, y) = (0, 0)$

Minimum is when both slopes = 0

# Functions of Two Variables

$f(x) = x^2$



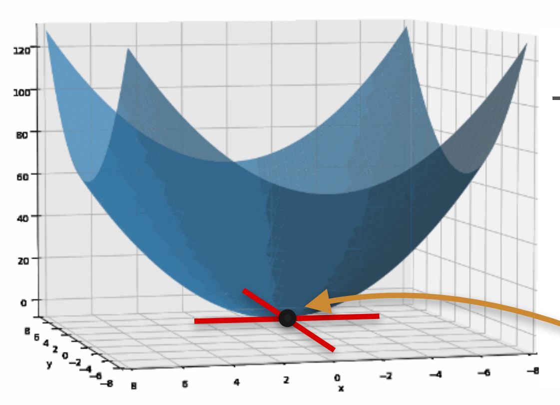

$f(x) = 0$

$2x = 0$

$x = 0$

Minimum is
when slope = 0

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



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

$2x = 0$ and $2y = 0$

$(x, y) = (0,0)$

Minimum is
when both slopes = 0

DeepLearning.AI

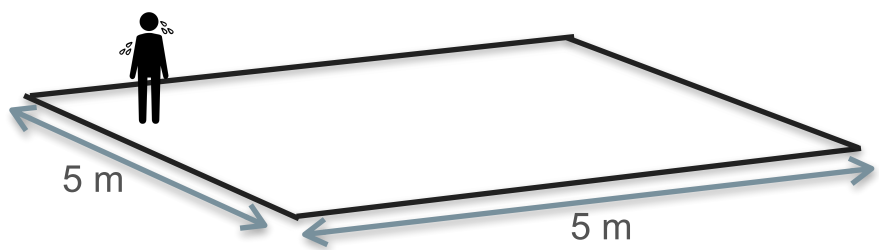
# Gradients and Gradient Descent

## Optimization with gradients: An example

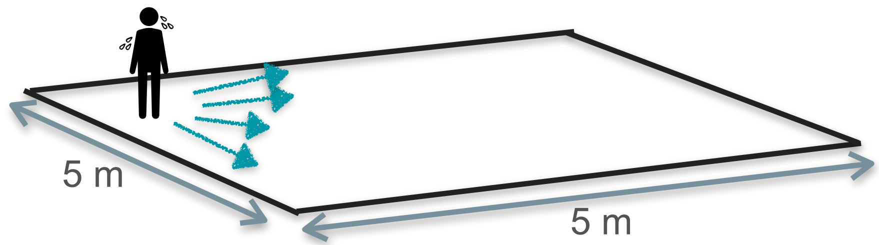
# Motivation for Optimization in Two Variables

# Motivation for Optimization in Two Variables

# Motivation for Optimization in Two Variables

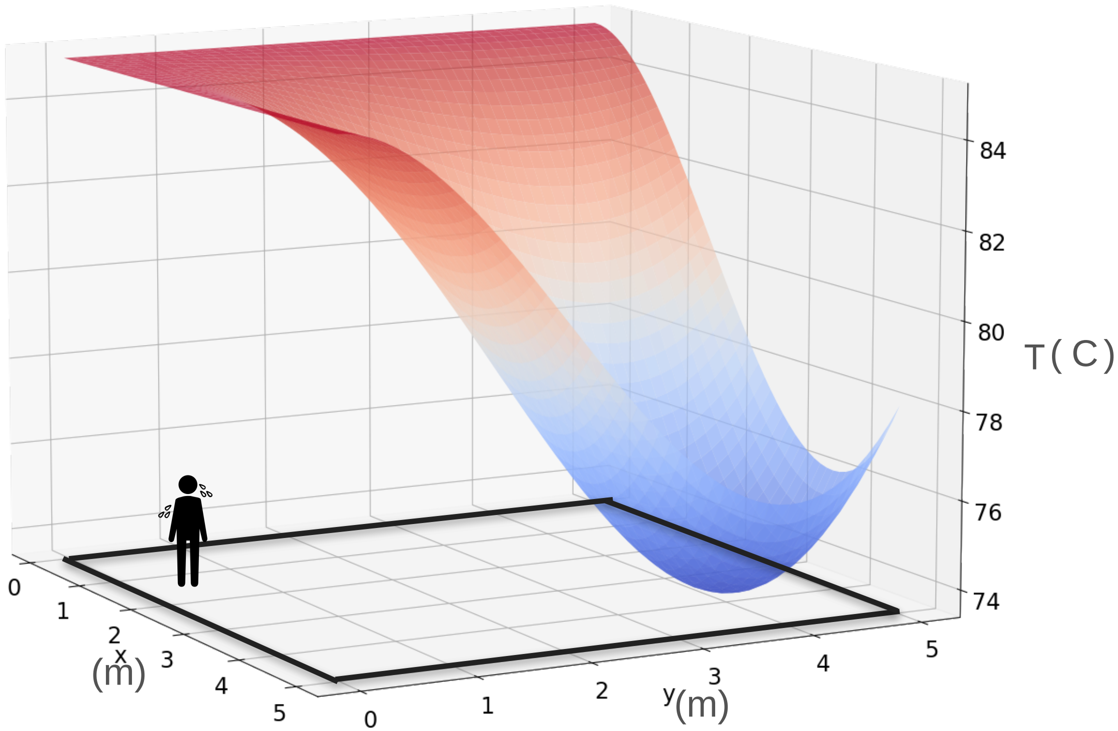


# Motivation for Optimization in Two Variables

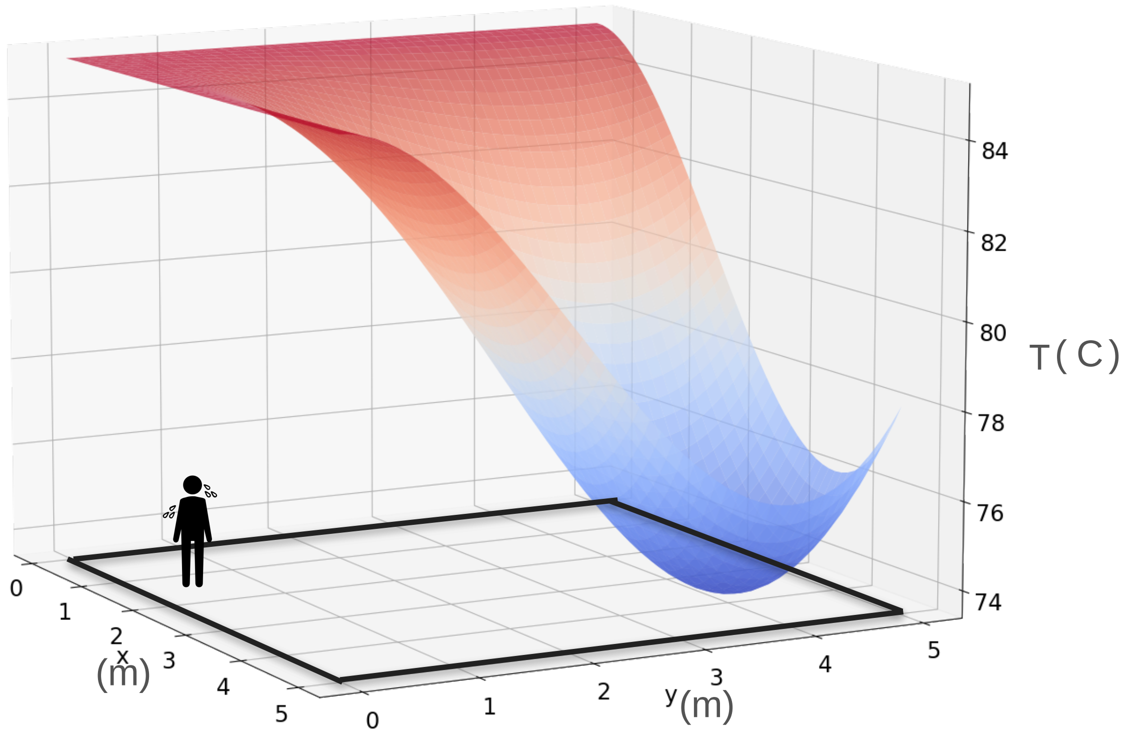



DeepLearning.AI

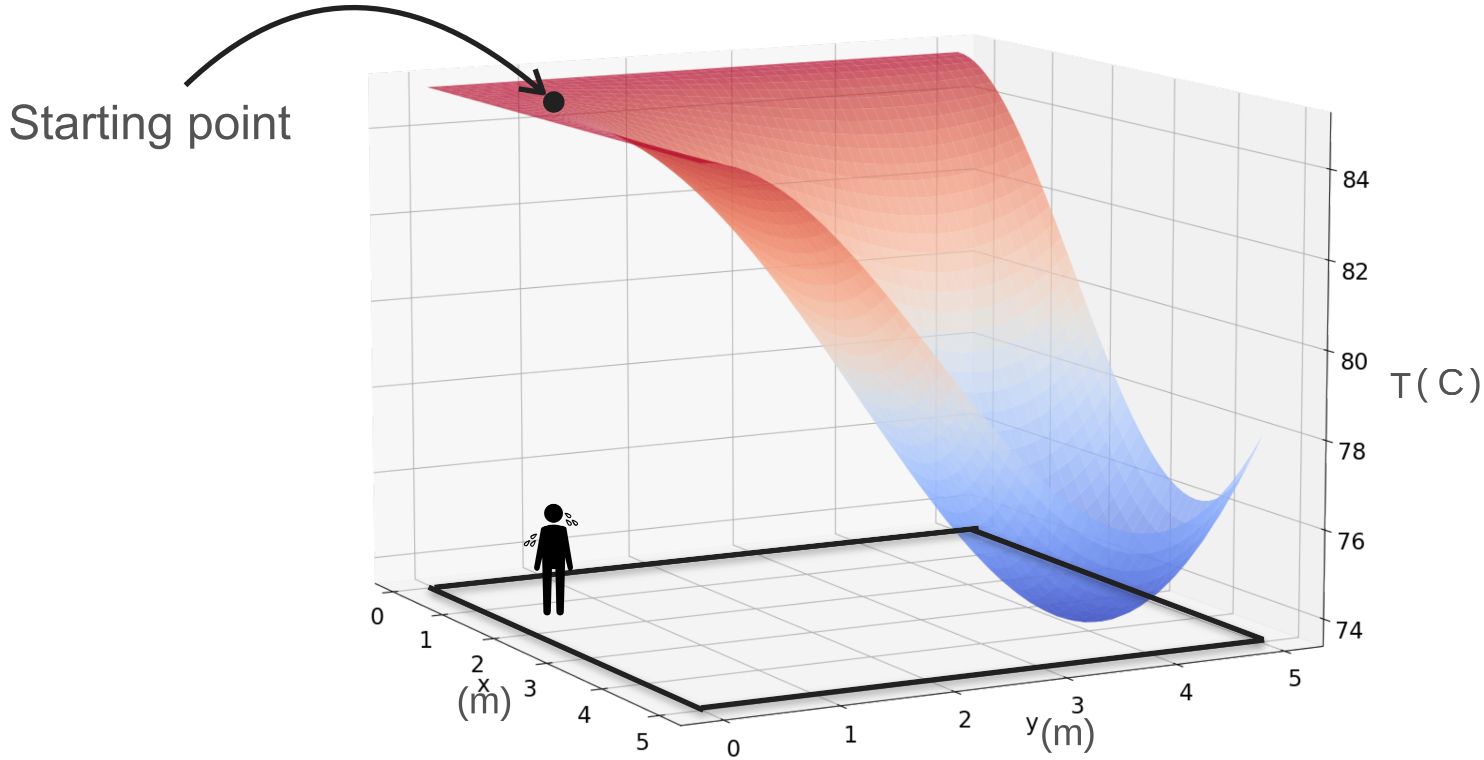
# Motivation for Optimization in Two Variables



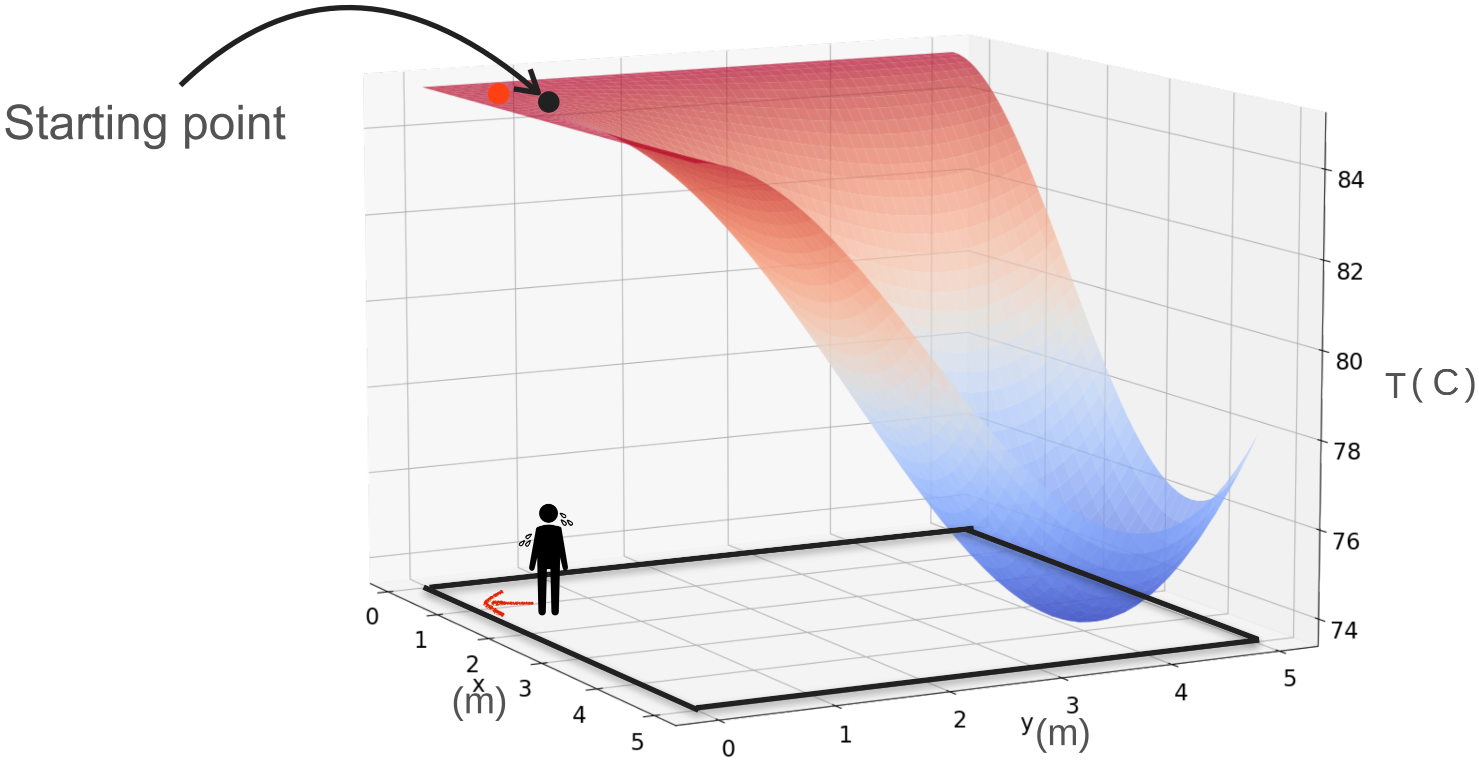
# Motivation for Optimization in Two Variables



# Motivation for Optimization in Two Variables

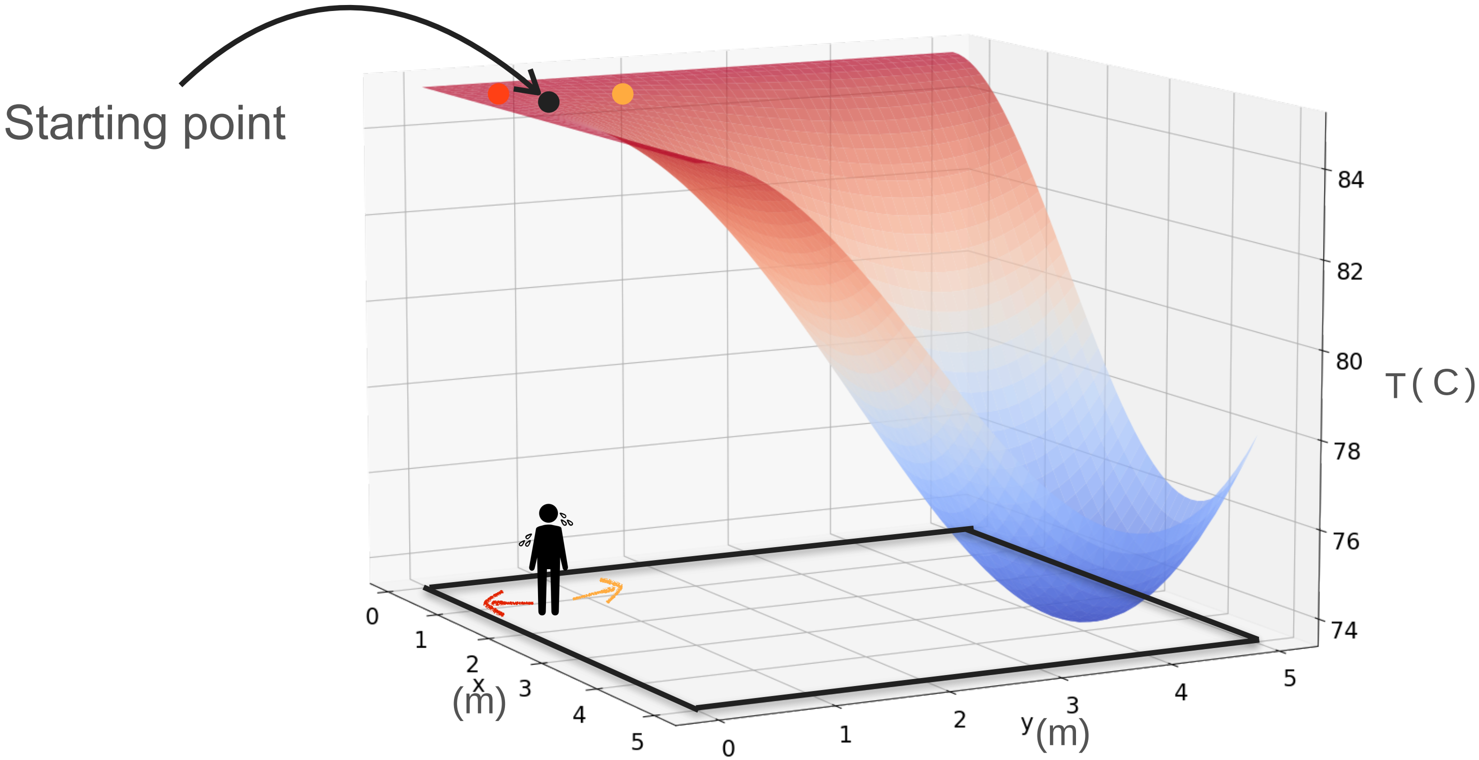


# Motivation for Optimization in Two Variables

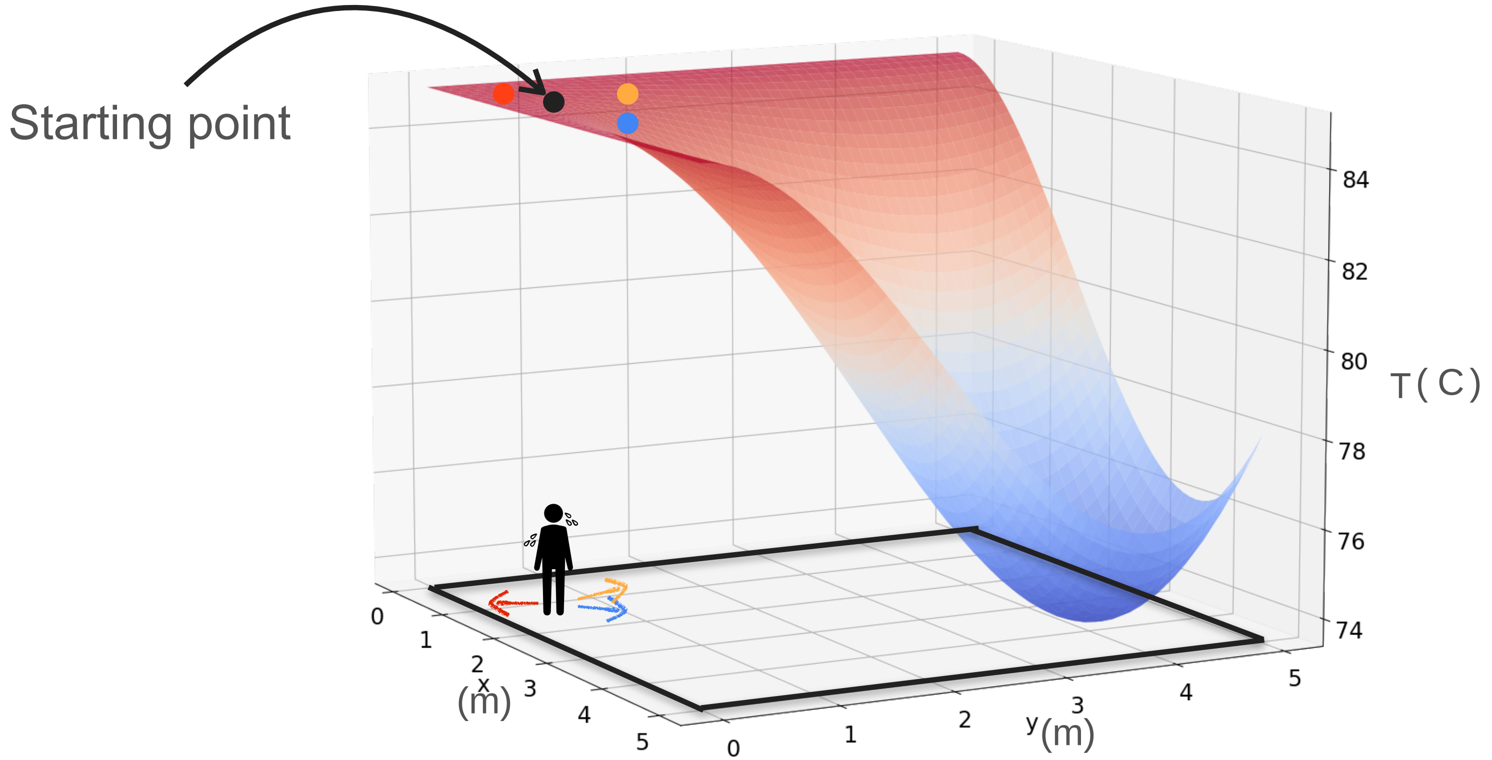



DeepLearning.AI

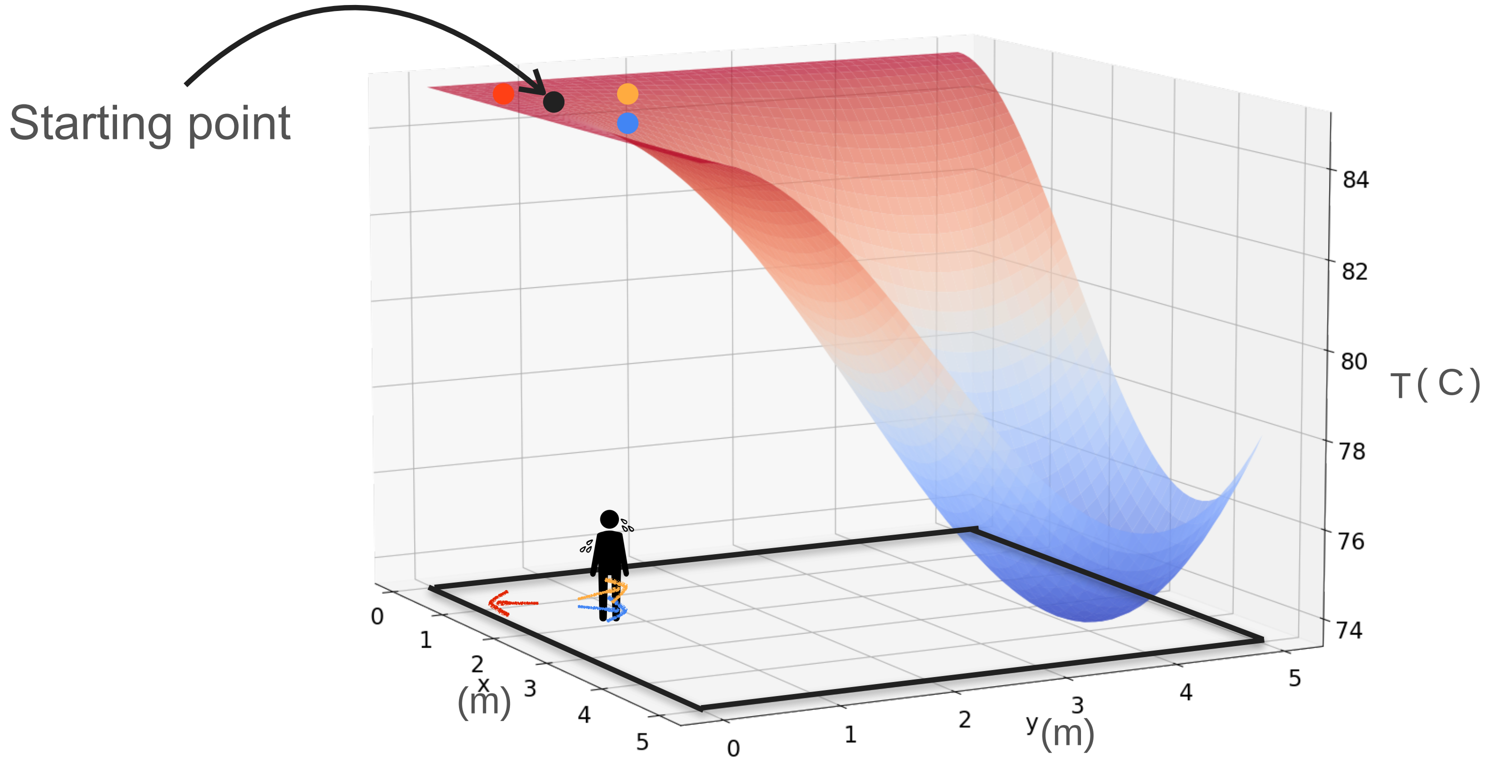
# Motivation for Optimization in Two Variables



# Motivation for Optimization in Two Variables

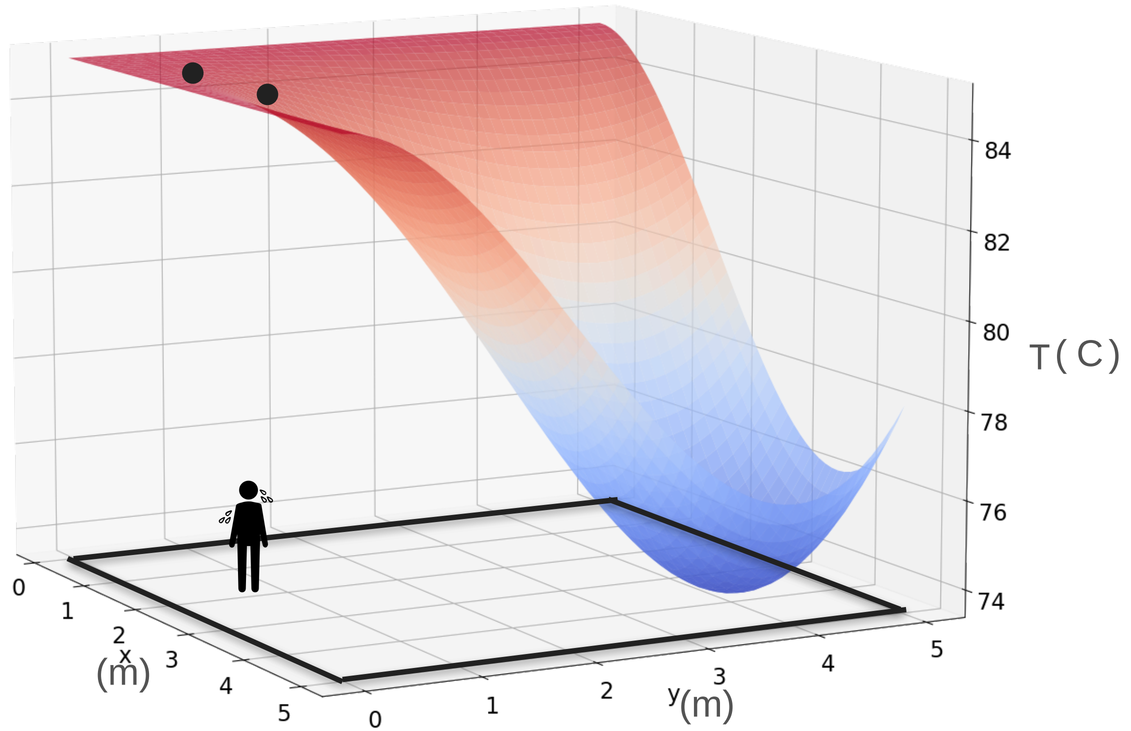


# Motivation for Optimization in Two Variables

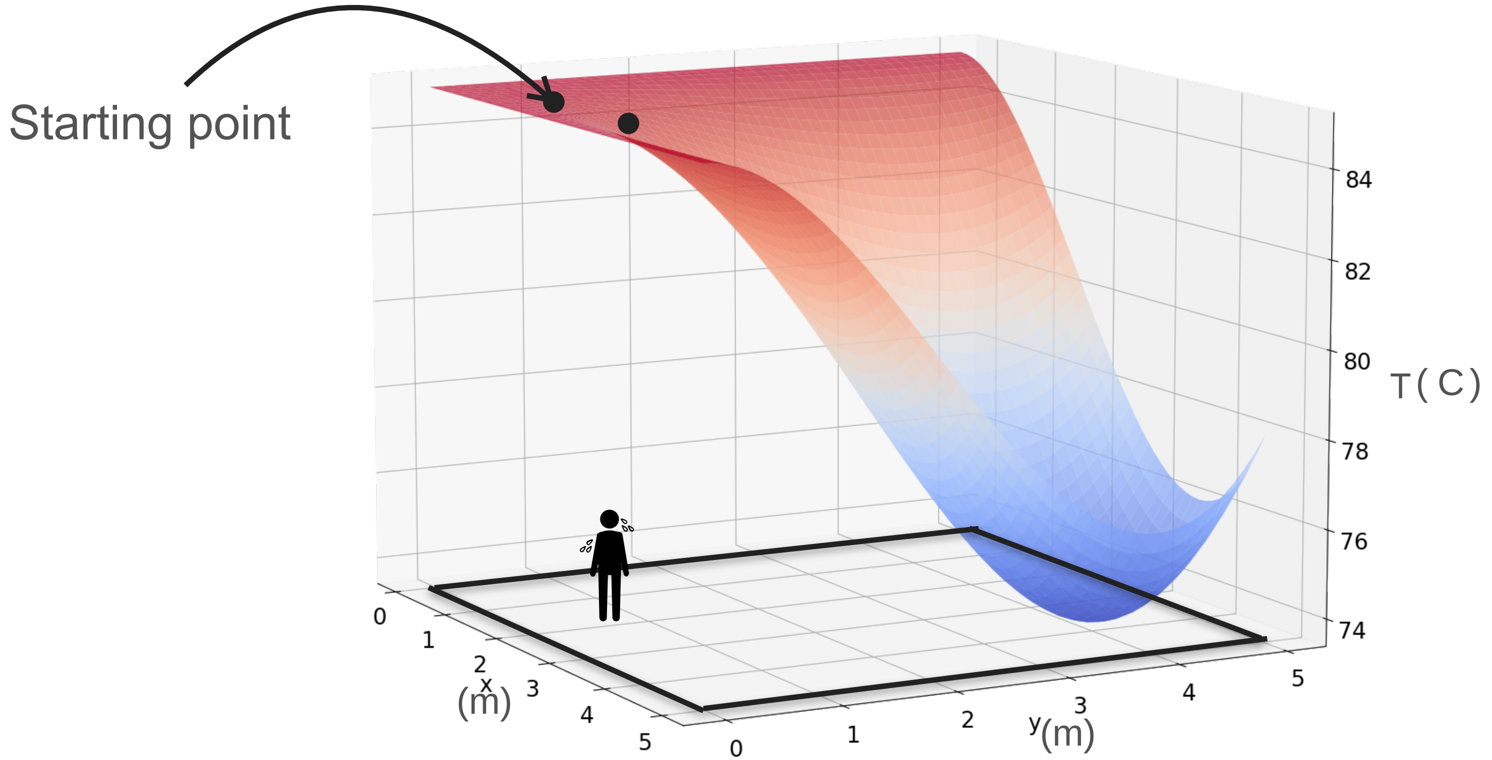


# Motivation for Optimization in Two Variables

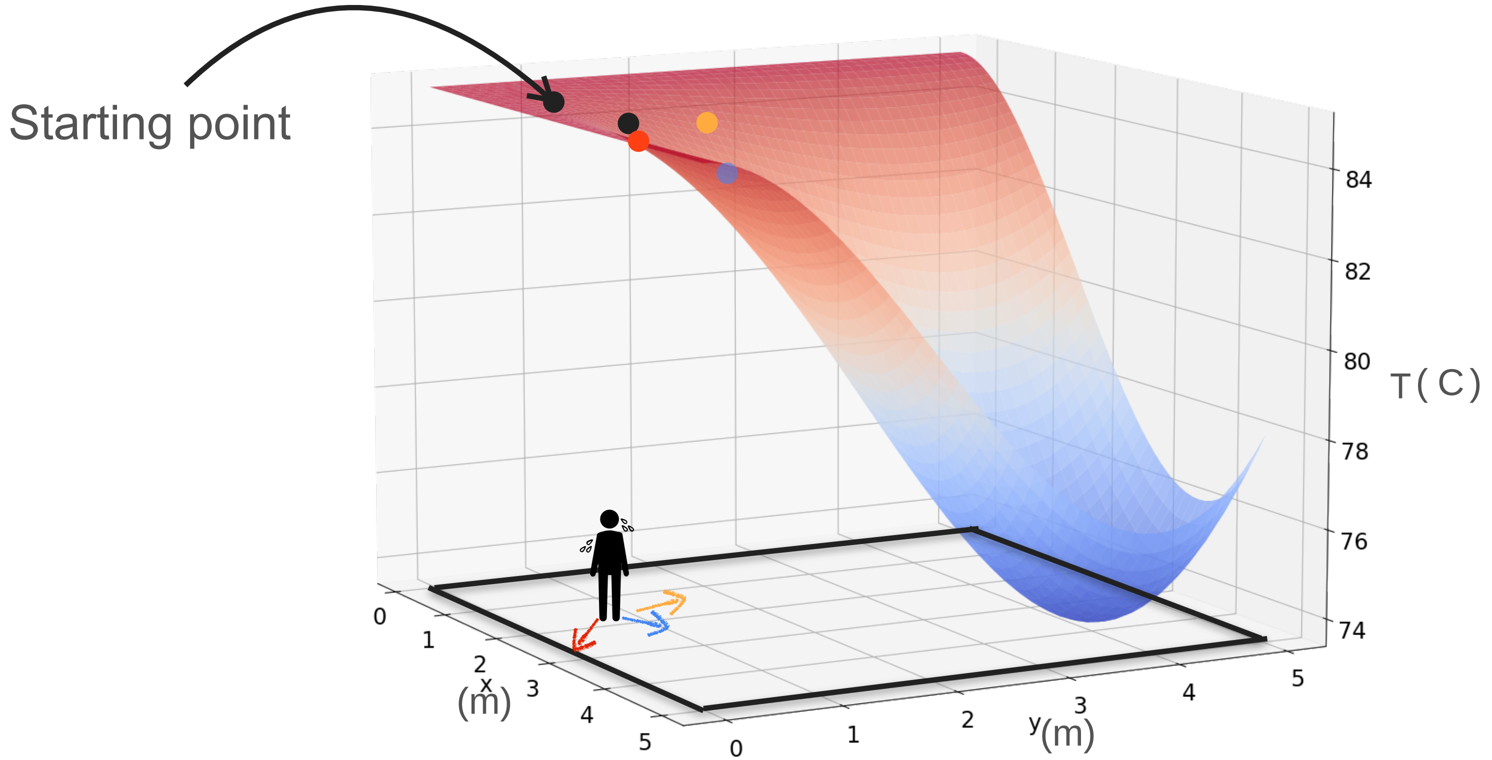
Starting point



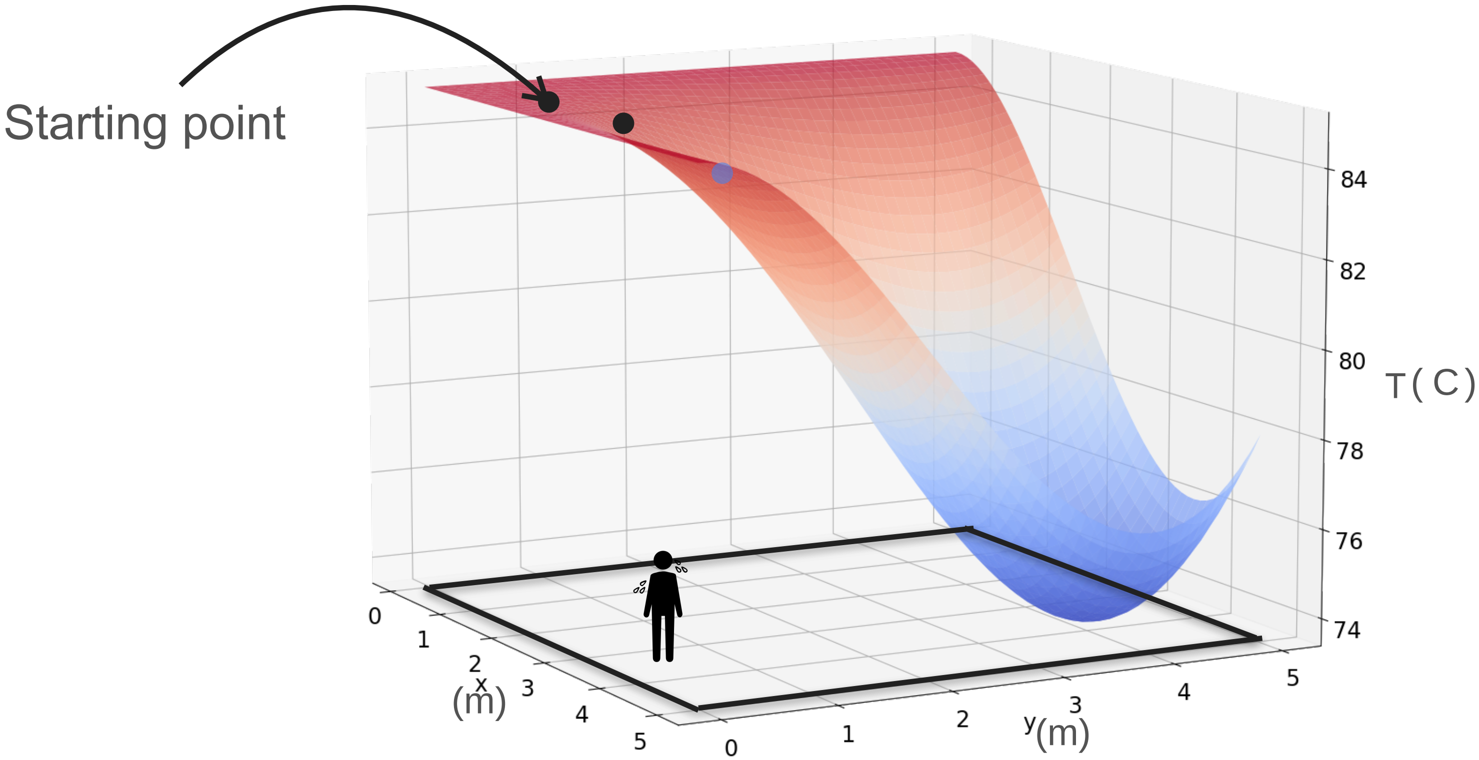
# Motivation for Optimization in Two Variables



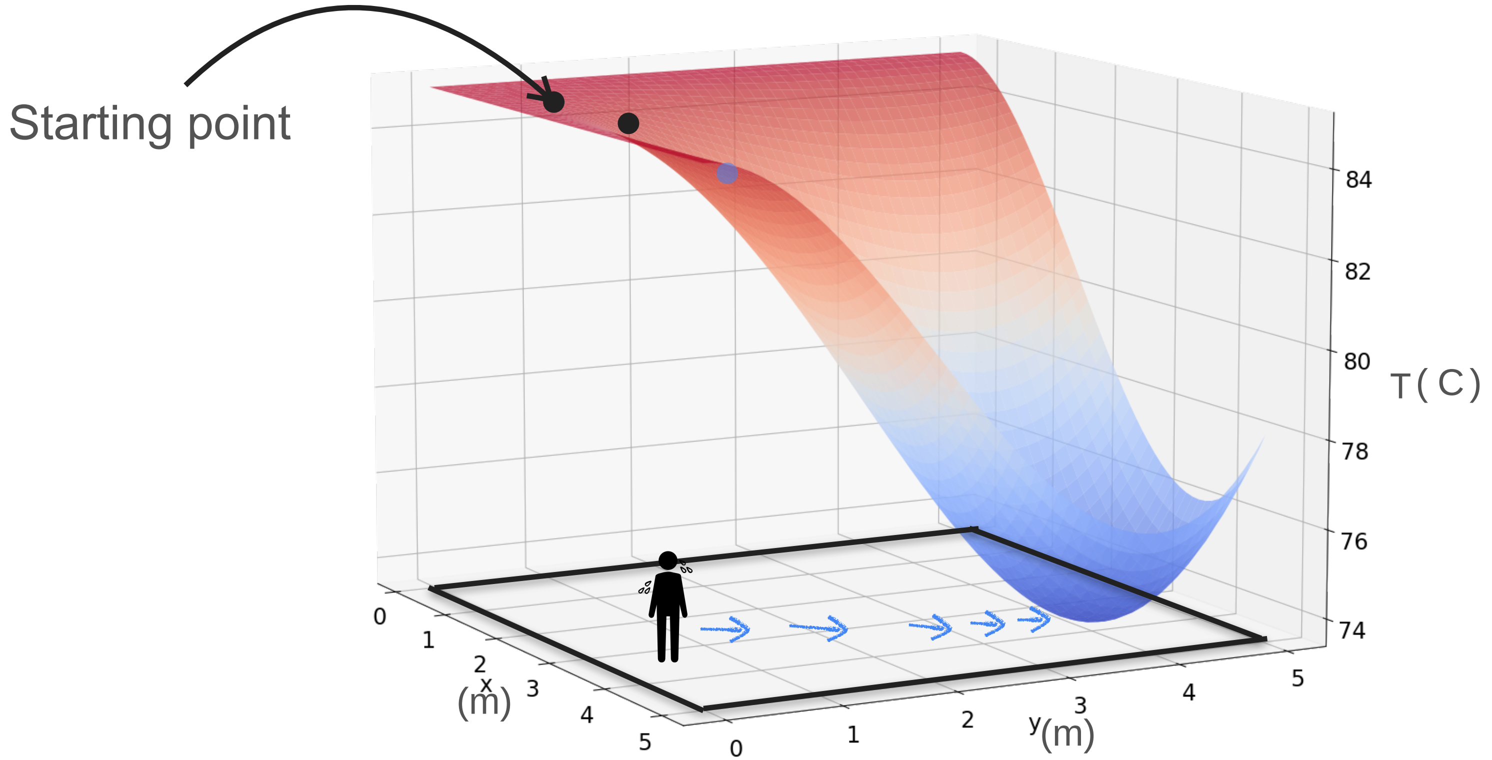
# Motivation for Optimization in Two Variables



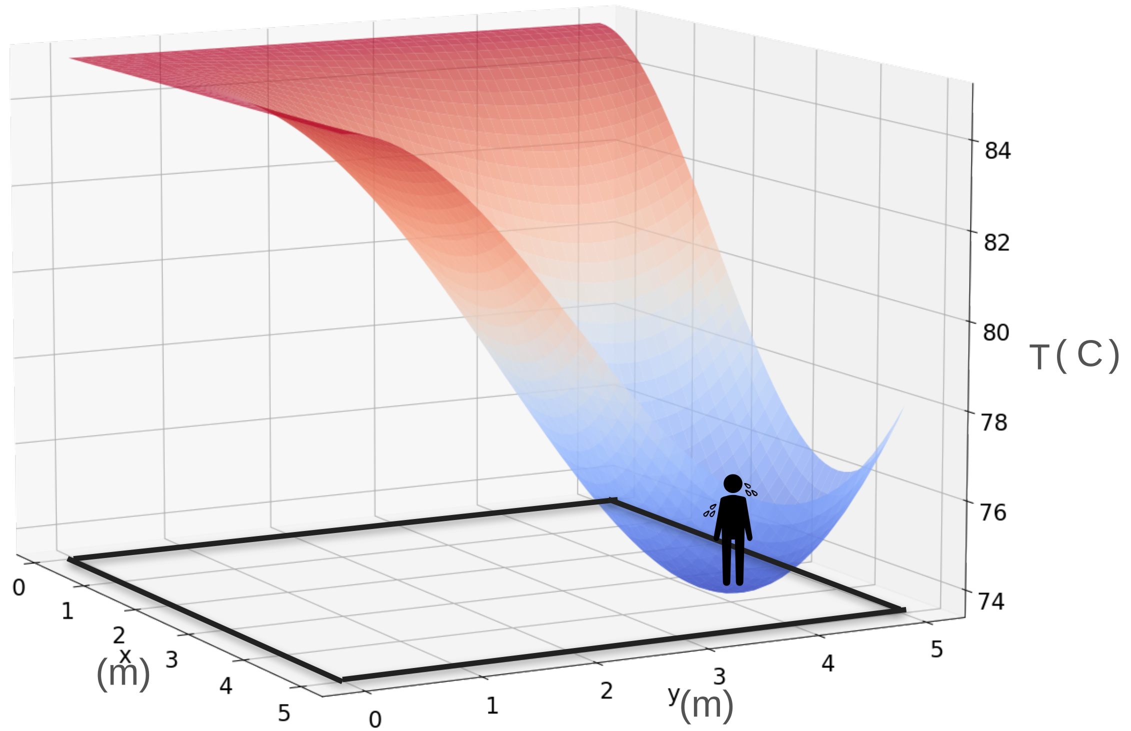
# Motivation for Optimization in Two Variables



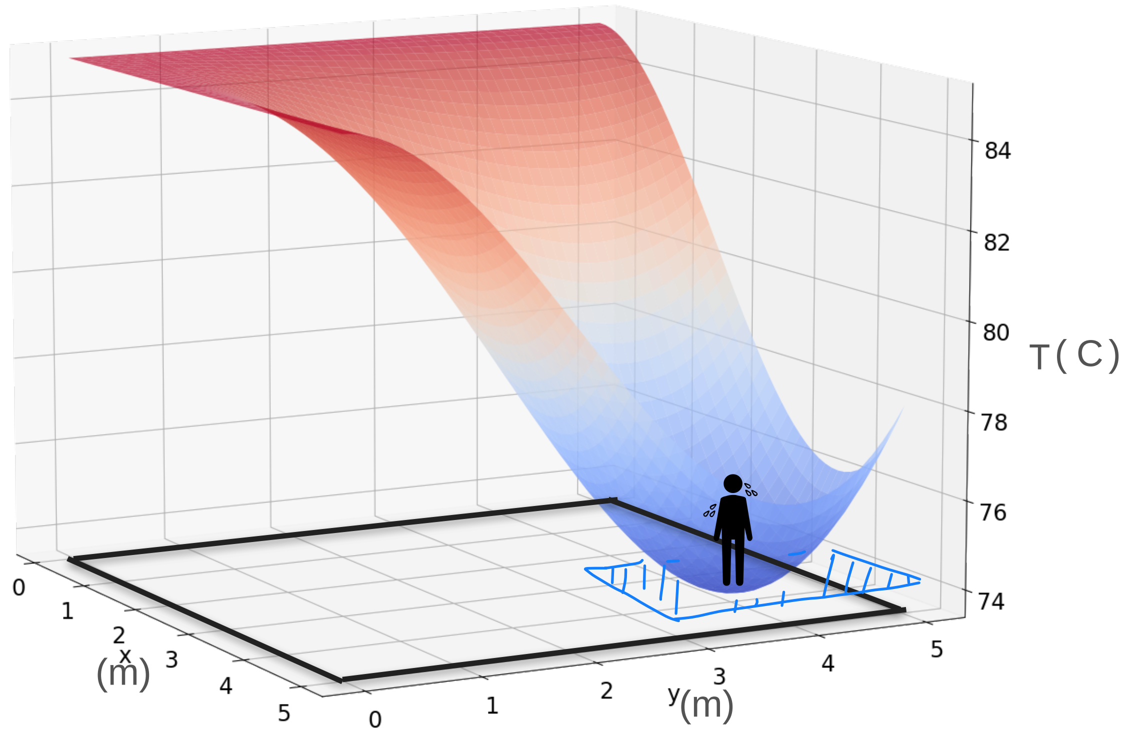
# Motivation for Optimization in Two Variables



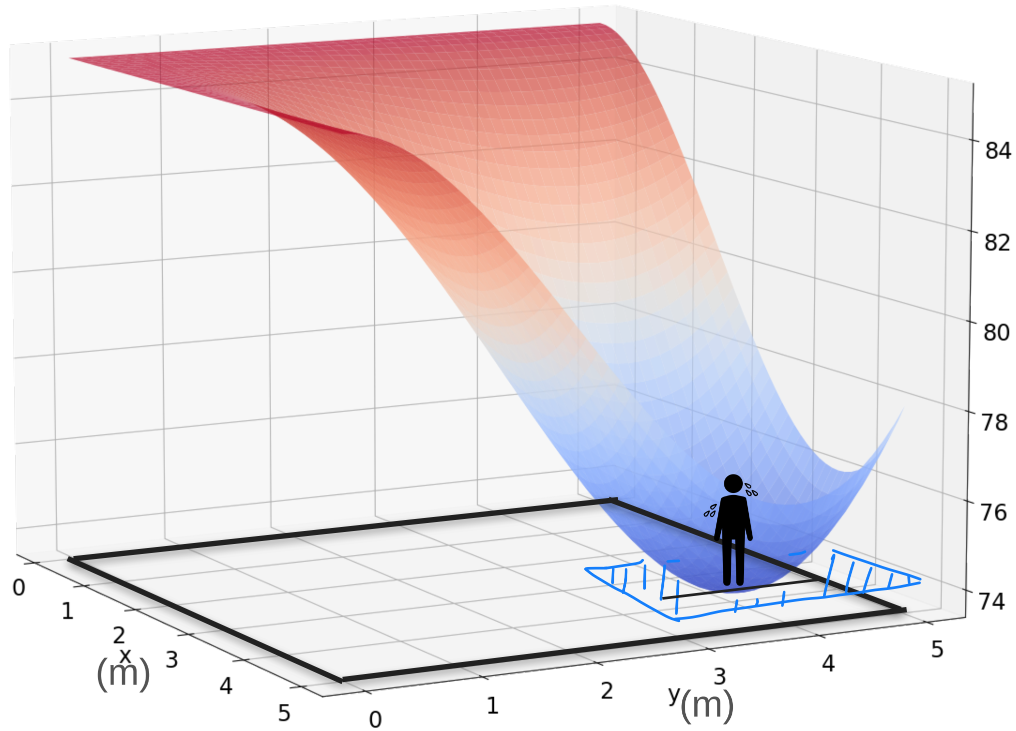
# Motivation for Optimization in Two Variables



# Motivation for Optimization in Two Variables

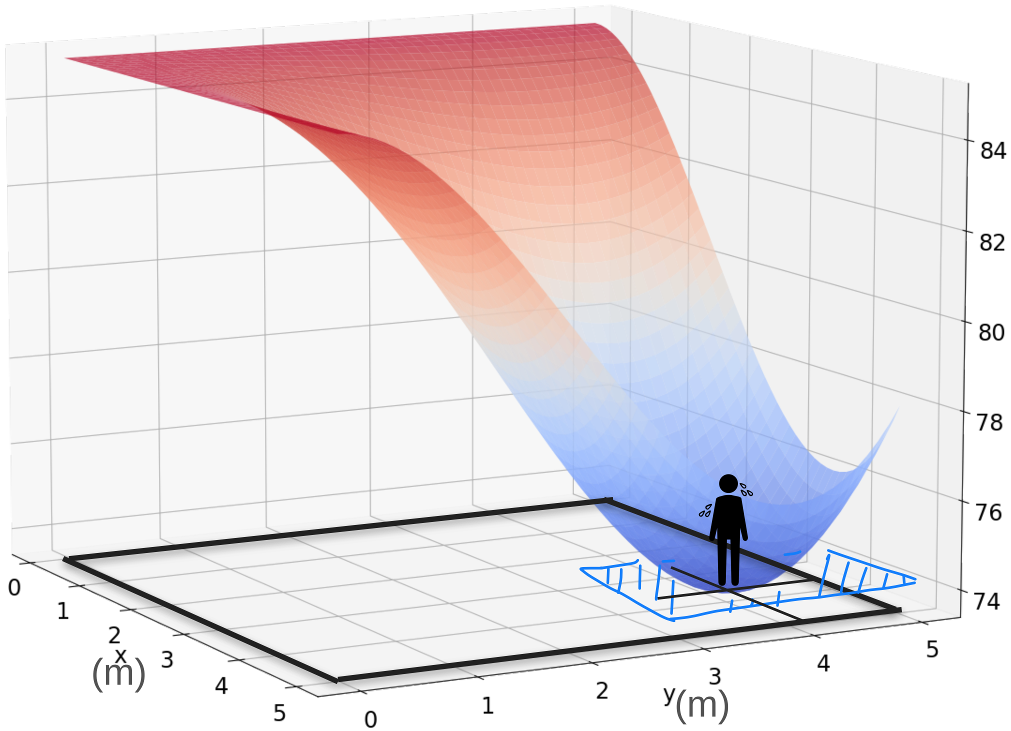


# Motivation for Optimization in Two Variables




$$T(C)\ \frac{T}{x} = 0$$

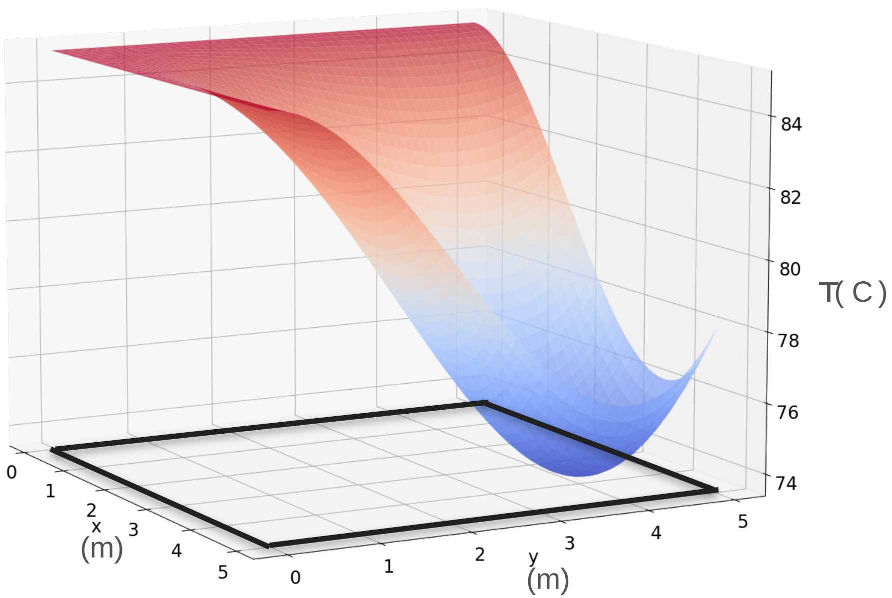
# Motivation for Optimization in Two Variables




$$T(C) \quad \frac{T}{x} = 0$$

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

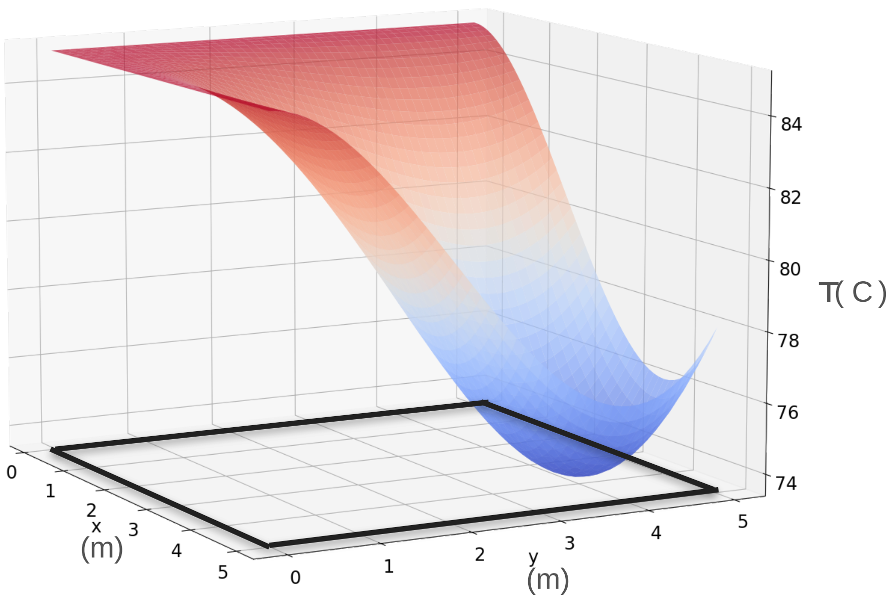
# Exercise




DeepLearning.AI

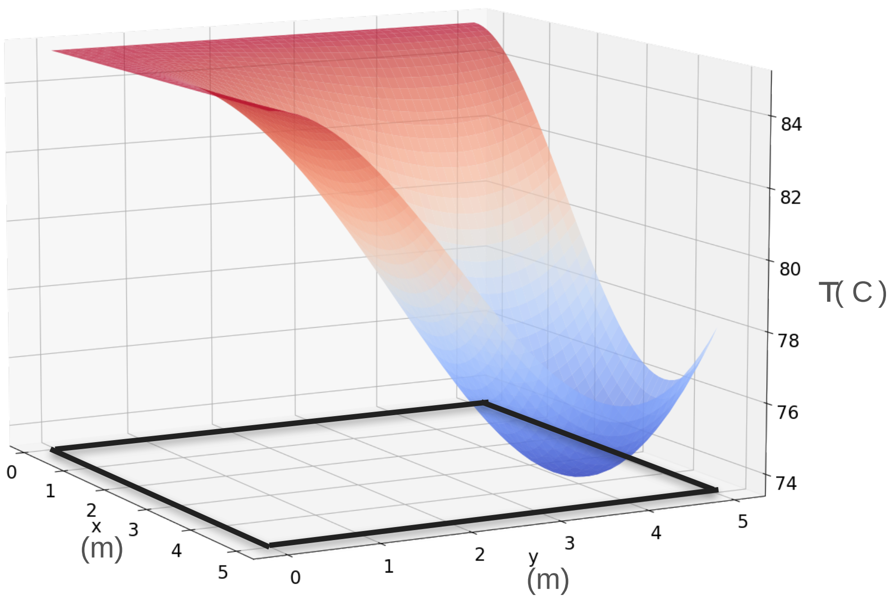
# Exercise

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



# Exercise

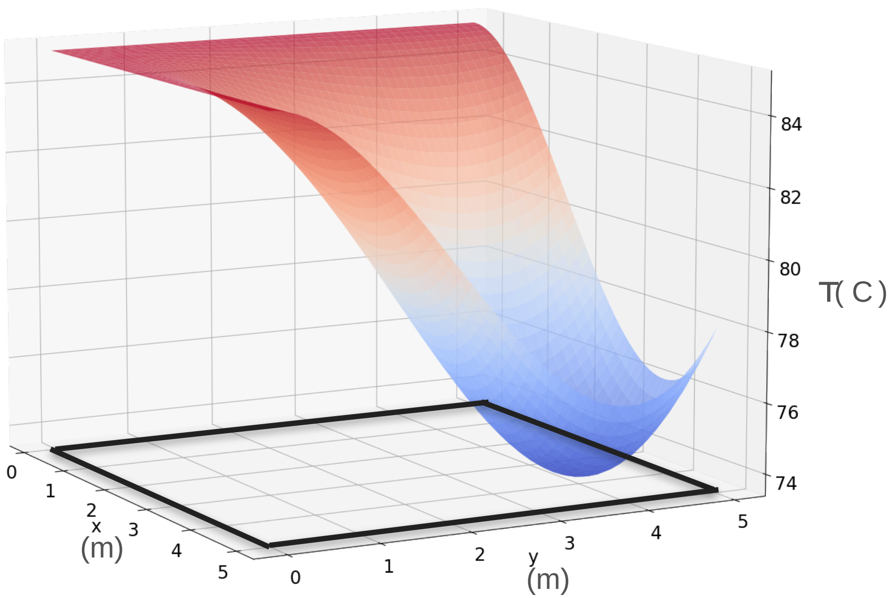
$$T = f(x, y) = 85 \quad \frac{1}{90}x^2(x \quad 6)y^2(y \quad 6)$$




Try and calculate $\frac{f}{x}$

# Exercise

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

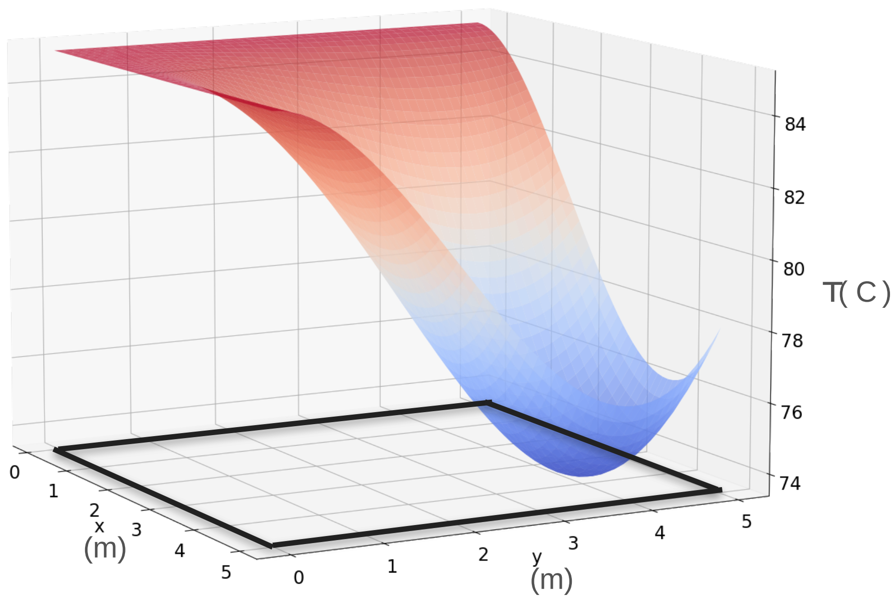



Try and calculate $\frac{f}{x}$

and $\frac{f}{y}$

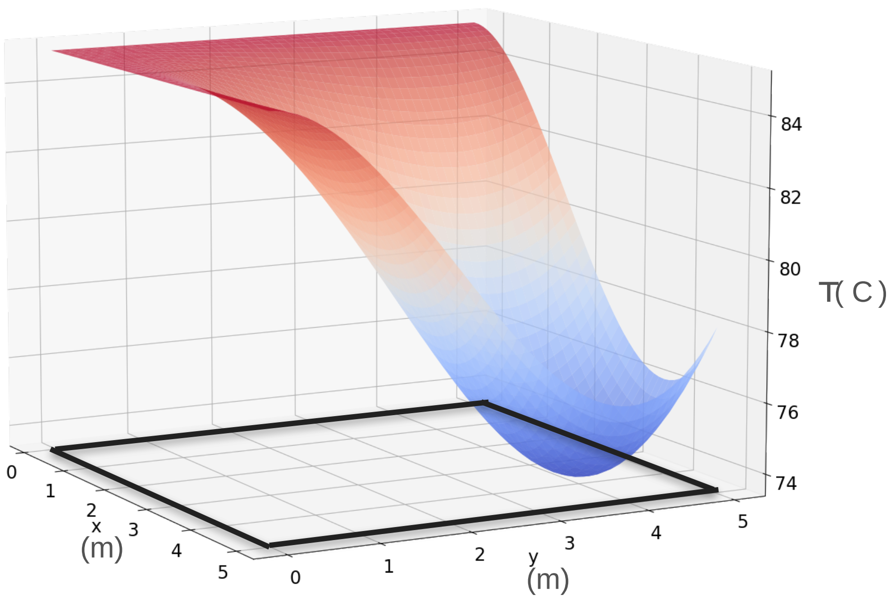
# Motivation for Optimization in Two Variables

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



# Motivation for Optimization in Two Variables

$$T = f(x, y) = 85 \quad \frac{1}{90}x^2(x \quad 6)y^2(y \quad 6) \qquad \frac{f}{x} = \quad \frac{1}{90}x(3x \quad 12)y^2(y \quad 6)$$

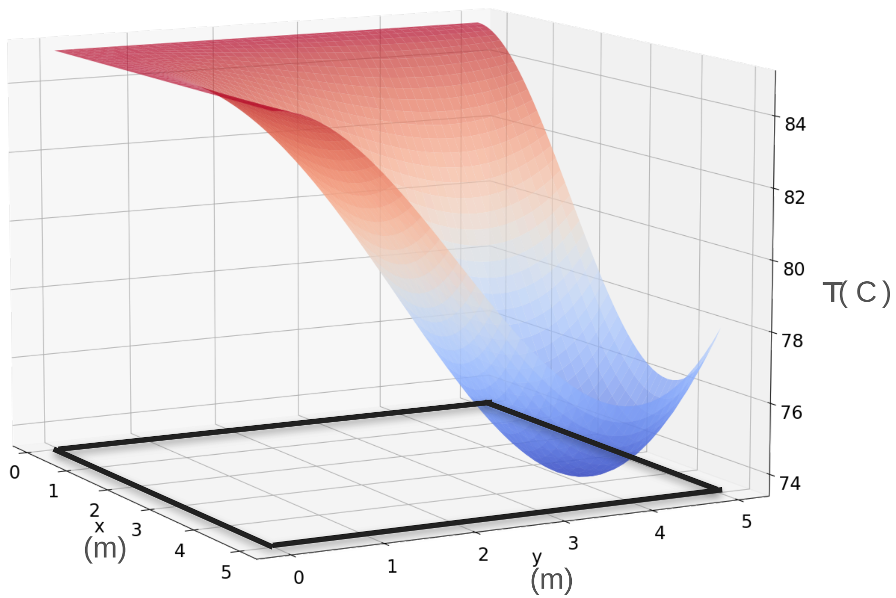


# Motivation for Optimization in Two Variables

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

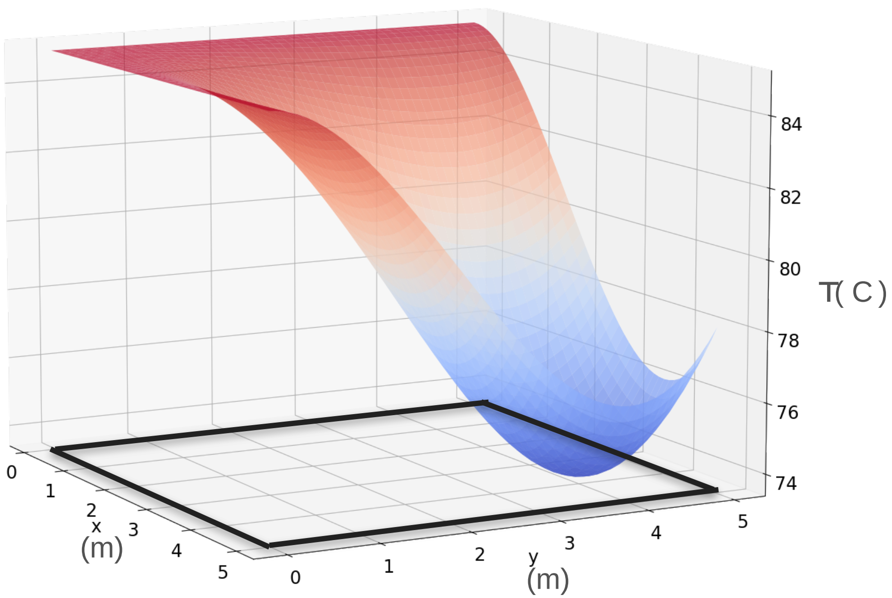
$$\frac{f}{x} = \quad \frac{1}{90} x(3x \quad 12) y^2 (y \quad 6)$$

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



# Motivation for Optimization in Two Variables

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

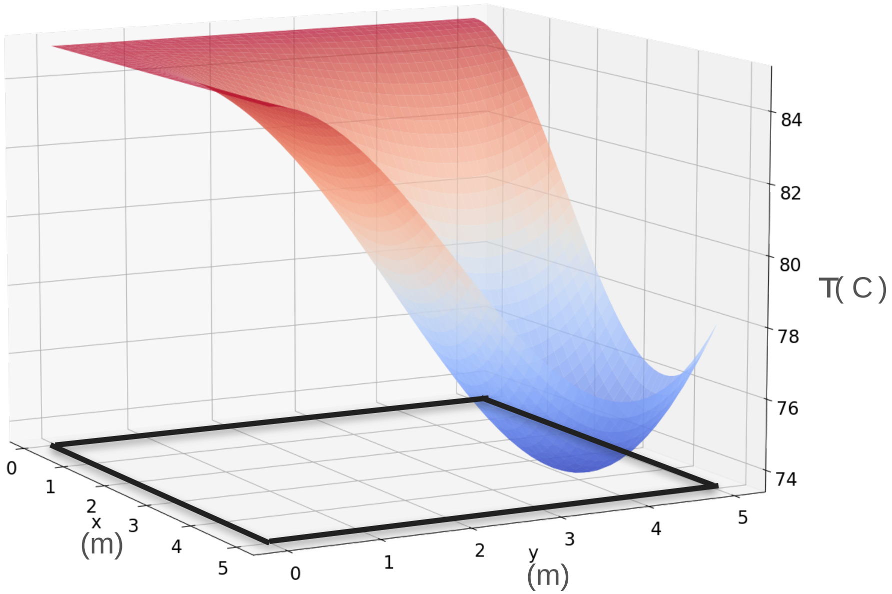



$$\frac{f}{x} = \quad \frac{1}{90}x(3x \quad 12)y^2(y \quad 6) = 0$$

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

# Motivation for Optimization in Two Variables

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



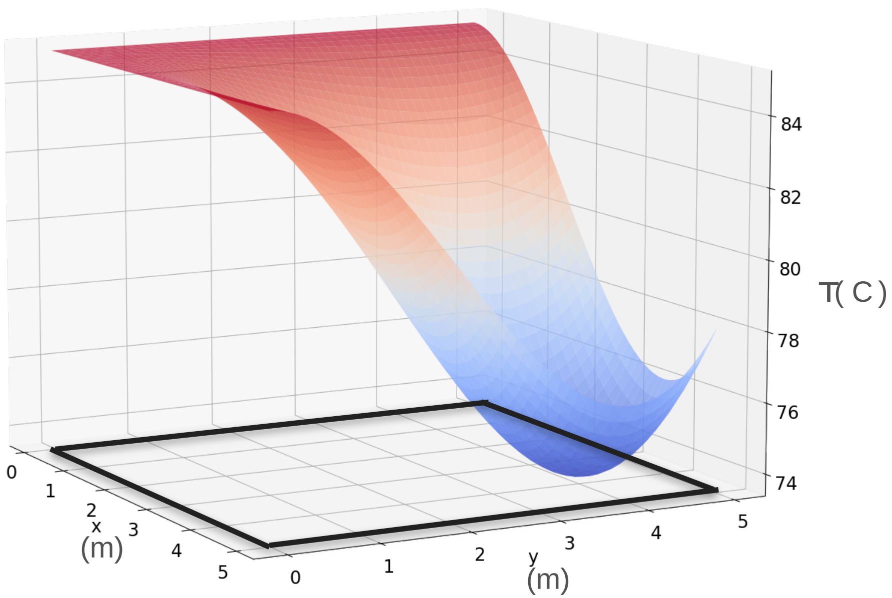

$$\frac{f}{x} = \quad \frac{1}{90}\boxed{x}(3x \quad 12)y^2(y \quad 6) = 0$$

$x = 0$

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

# Motivation for Optimization in Two Variables

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



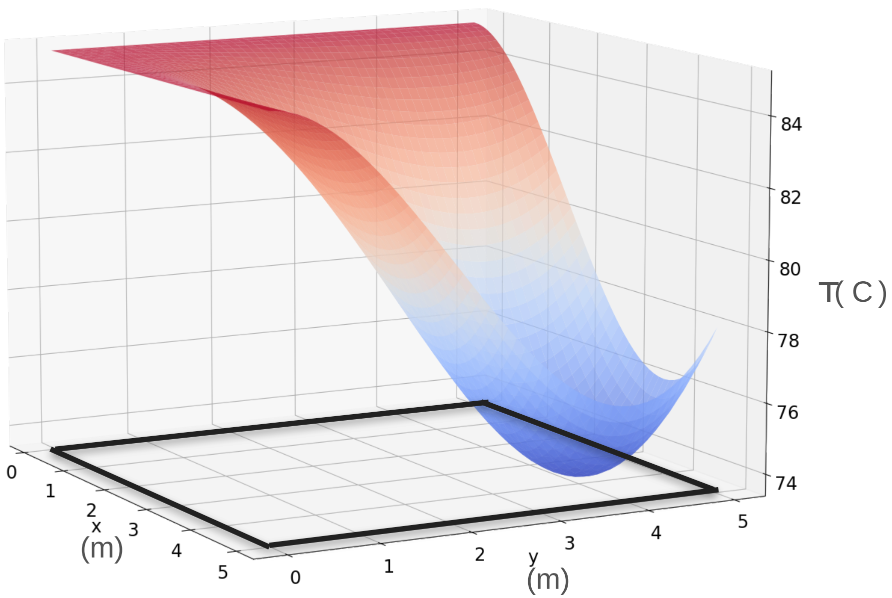

$$\frac{f}{x} = \quad \frac{1}{90} \boxed{x}\boxed{(3x \quad 12)}y^2(y \quad 6) = 0$$

$x = 0$ $\quad$ $x = 4$

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

# Motivation for Optimization in Two Variables

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



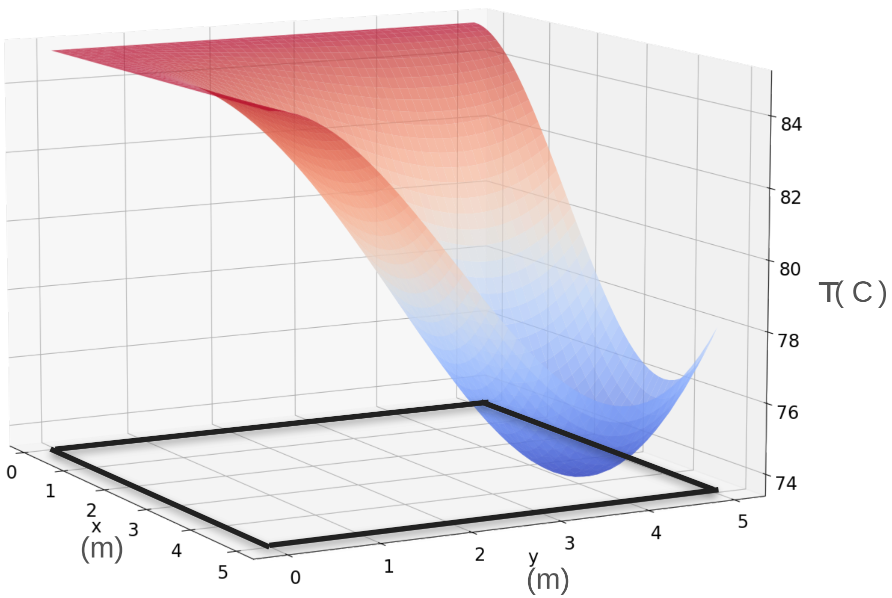

$$\frac{f}{x} = \quad \frac{1}{90}\boxed{x}\boxed{(3x \quad 12)}\boxed{y^2}(y \quad 6) = 0$$

$x = 0$ $\quad x = 4$ $\quad y = 0$

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

# Motivation for Optimization in Two Variables

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

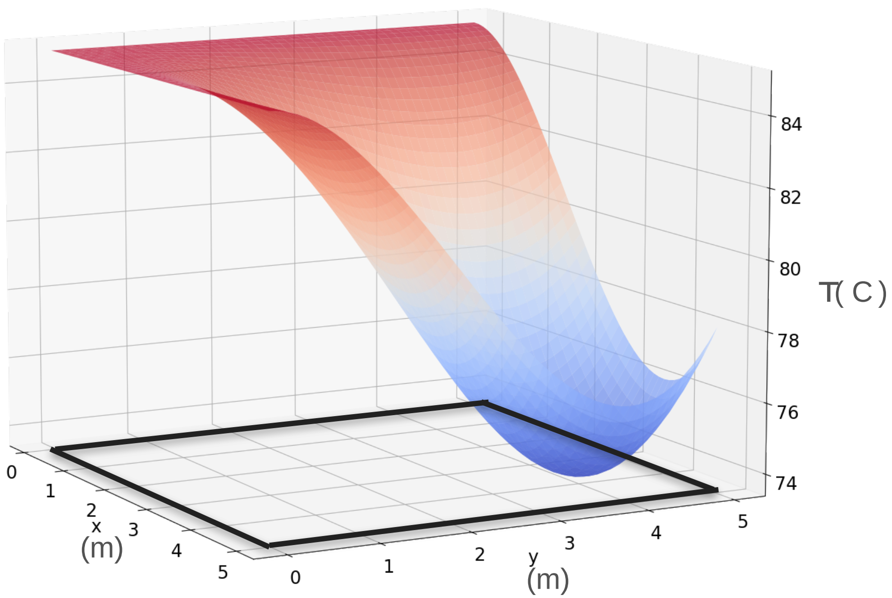



$$\frac{f}{x} = \quad \frac{1}{90}\underbrace{\boxed{x}}_{x\,=\,0}\underbrace{\boxed{(3x \quad 12)}}_{x\,=\,4}\underbrace{\boxed{y^2}}_{y\,=\,0}\underbrace{\boxed{(y \quad 6)}}_{y\,=\,6} = 0$$

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

# Motivation for Optimization in Two Variables

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

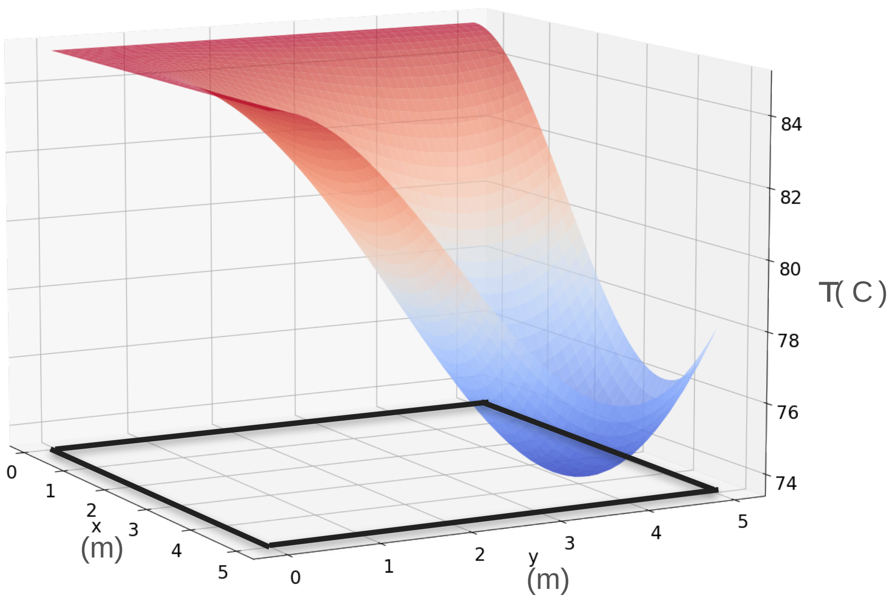



$$\frac{f}{x} = \frac{1}{90}\underbrace{\boxed{x}}_{x=0}\underbrace{\boxed{(3x \quad 12)}}_{x=4}\underbrace{\boxed{y^2}}_{y=0}\underbrace{\boxed{(y \quad 6)}}_{y=6} = 0$$

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

# Motivation for Optimization in Two Variables

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



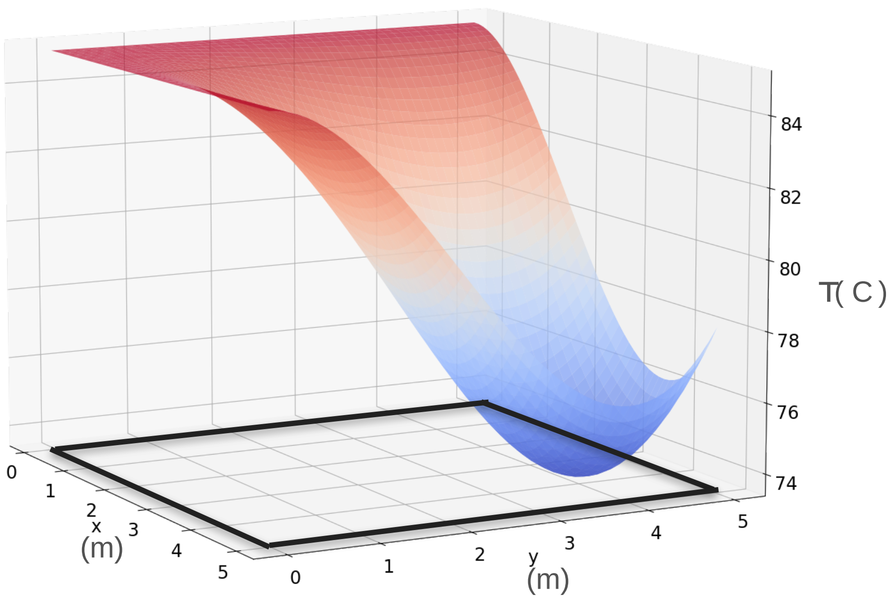

$$\frac{f}{x} = \frac{1}{90}\underbrace{\boxed{x}}_{x = 0}\underbrace{\boxed{(3x \quad 12)}}_{x = 4}\underbrace{\boxed{y^2}}_{y = 0}\underbrace{\boxed{(y \quad 6)}}_{y = 6} = 0$$

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

DeepLearning.AI

# Motivation for Optimization in Two Variables

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

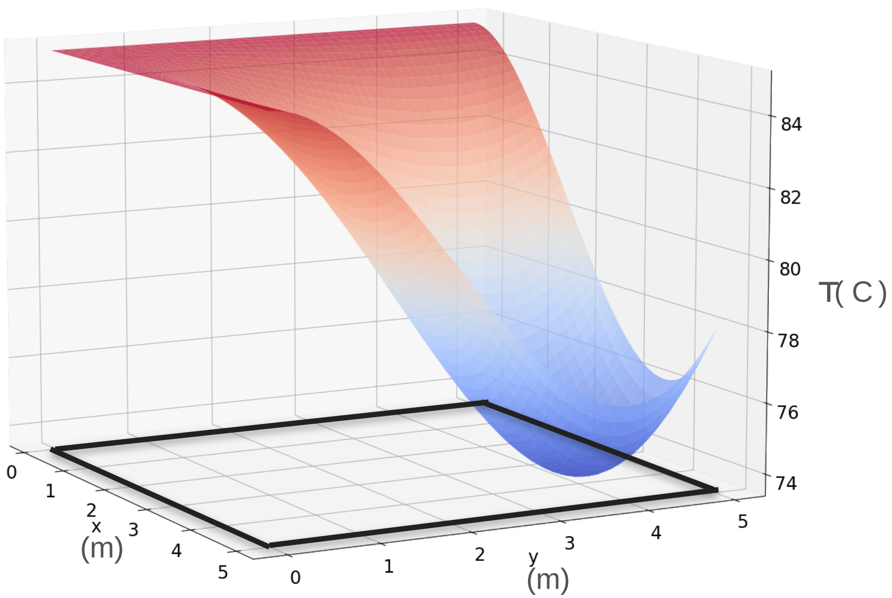



$$\frac{f}{x} = \frac{1}{90}\underbrace{\boxed{x}}_{x=0}\underbrace{\boxed{(3x \quad 12)}}_{x=4}\underbrace{\boxed{y^2}}_{y=0}\underbrace{\boxed{(y \quad 6)}}_{y=6} = 0$$

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

# Motivation for Optimization in Two Variables

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




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

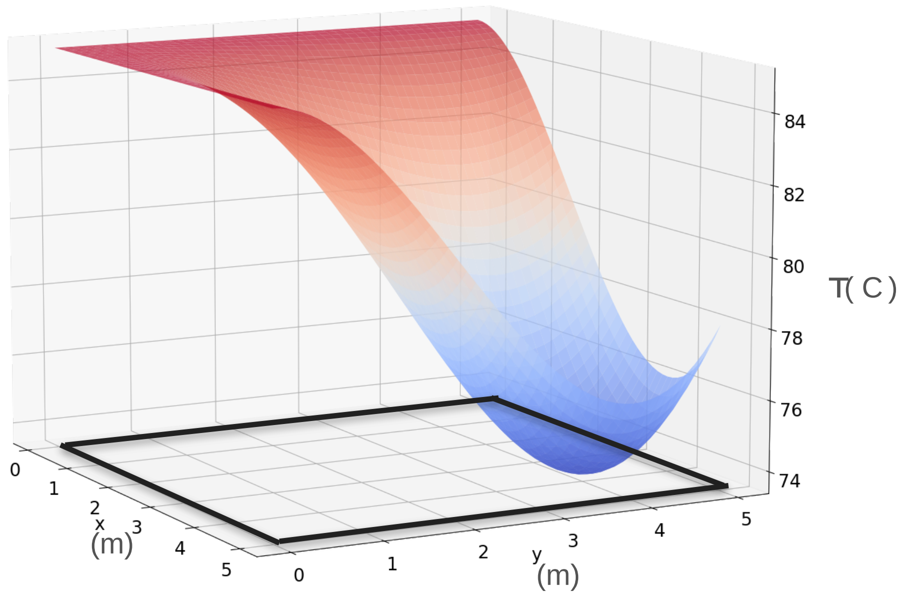
$x = 0$ $x = 4$ $y = 0$ $y = 6$

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

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

# Motivation for Optimization in Two Variables

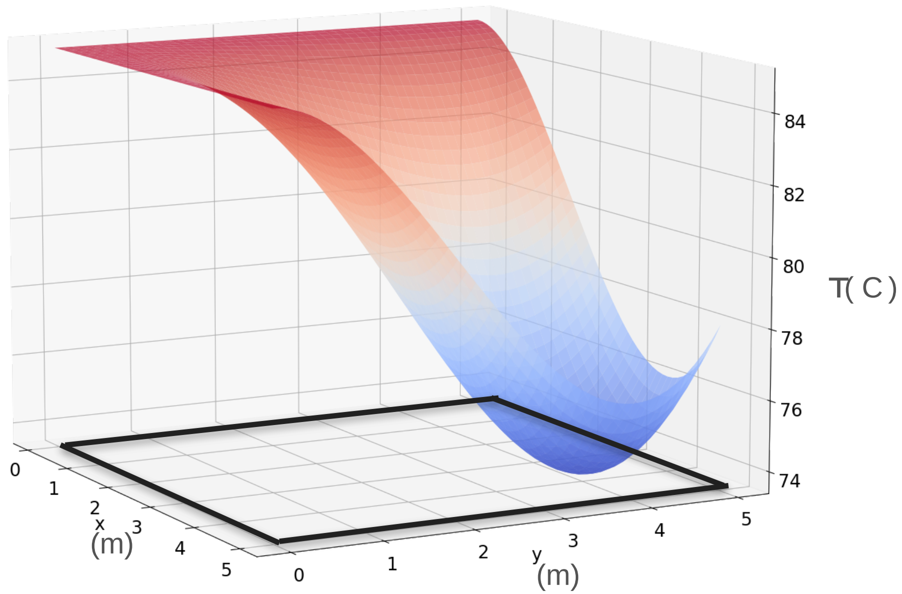
$$T = f(x, y) = 85 \quad \frac{1}{90}x^2(x \quad 6)y^2(y \quad 6)$$



# Motivation for Optimization in Two Variables

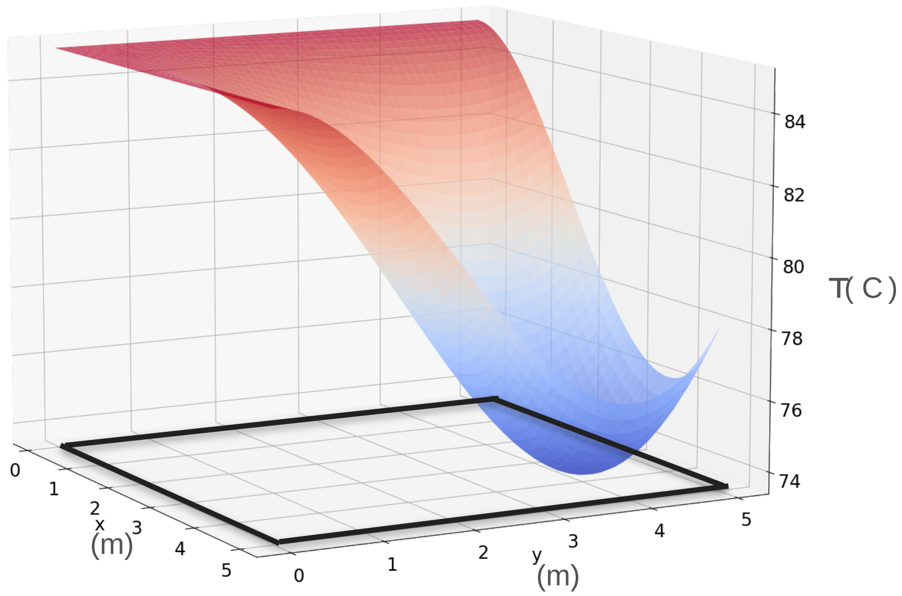
$$T = f(x, y) = 85 - \frac{1}{90}x^2(x - 6)y^2(y - 6)$$

Candidate points for the minima



# Motivation for Optimization in Two Variables

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




Candidate points for the minima

$x = 0$

$y = 0$

$x = 0, y = 0$

$x = 0, y = 4$

$x = 0, y = 6$

$x = 4, y = 0$

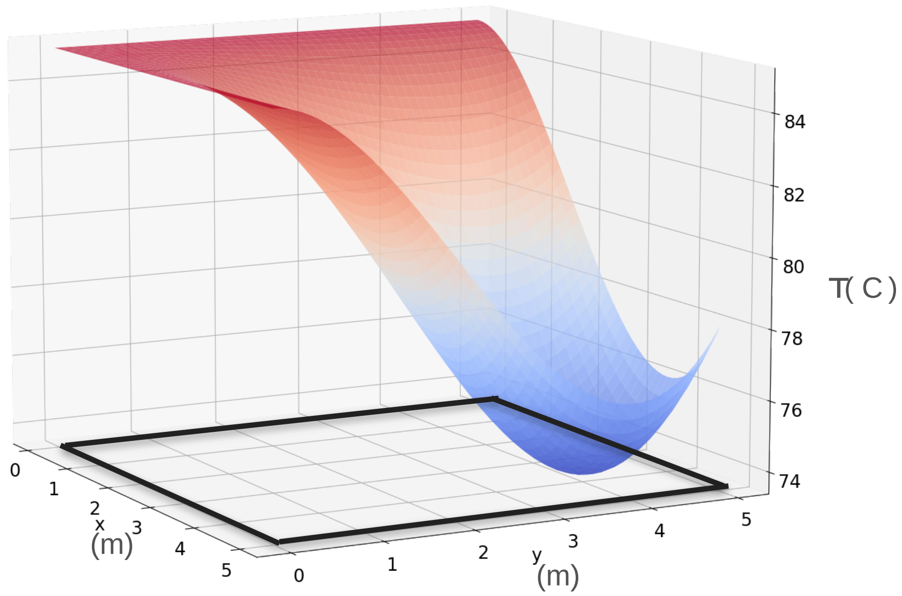
$x = 4, y = 4$

$x = 6, y = 0$

$x = 6, y = 6$

# Motivation for Optimization in Two Variables

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




Candidate points for the minima

$x = 0$

$y = 0$

$x = 0, y = 0$

$x = 0, y = 4$

$x = 0, y = 6$ Outside

$x = 4, y = 0$

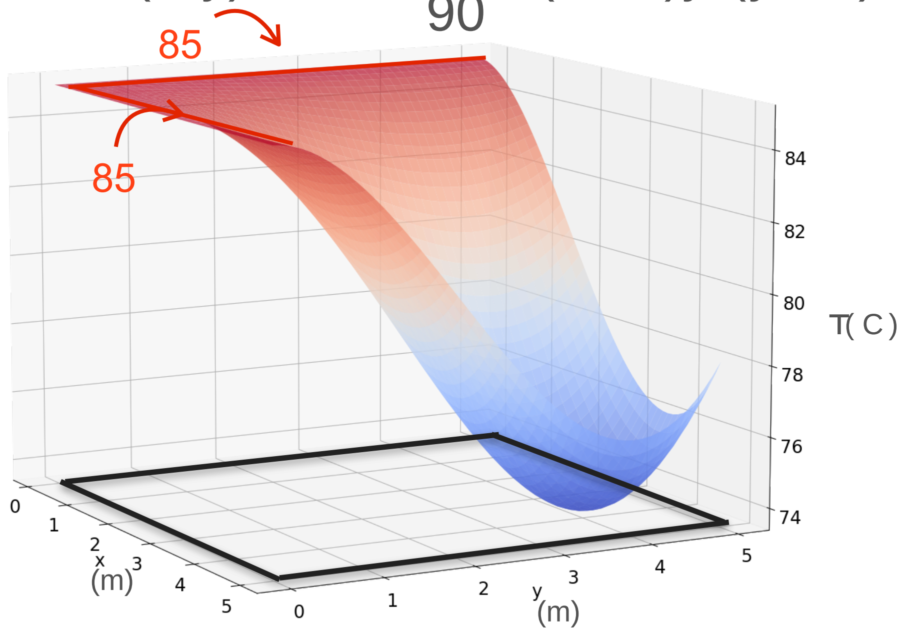
$x = 4, y = 4$

$x = 6, y = 0$

$x = 6, y = 6$ Outside

# Motivation for Optimization in Two Variables

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




Candidate points for the minima

$x = 0$
$y = 0$

Maxima

$x = 0, y = 0$

$x = 0, y = 4$

$x = 0, y = 6$ Outside

$x = 4, y = 0$

$x = 4, y = 4$

$x = 6, y = 0$

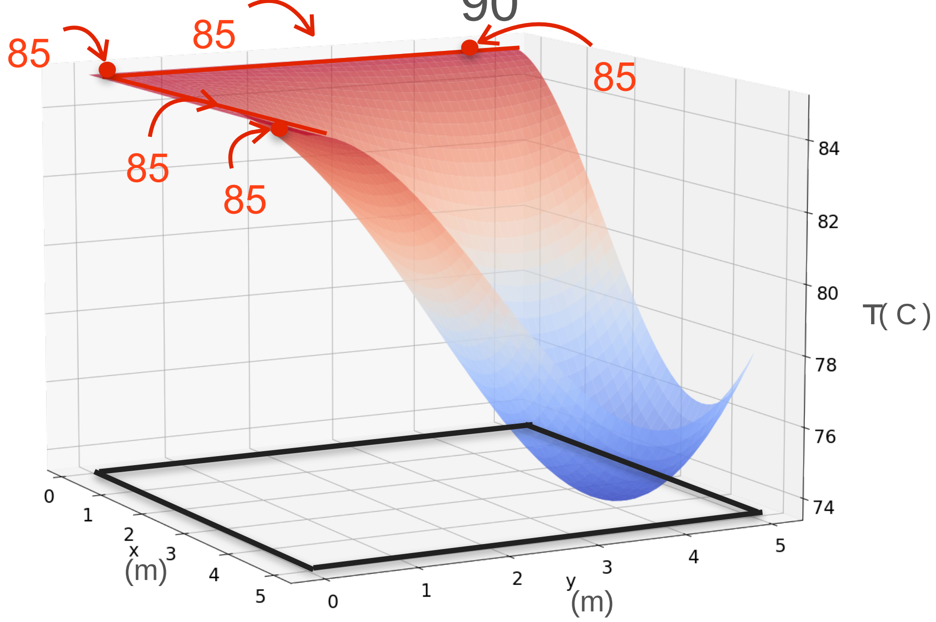
$x = 6, y = 6$

Outside

DeepLearning.AI

# Motivation for Optimization in Two Variables

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

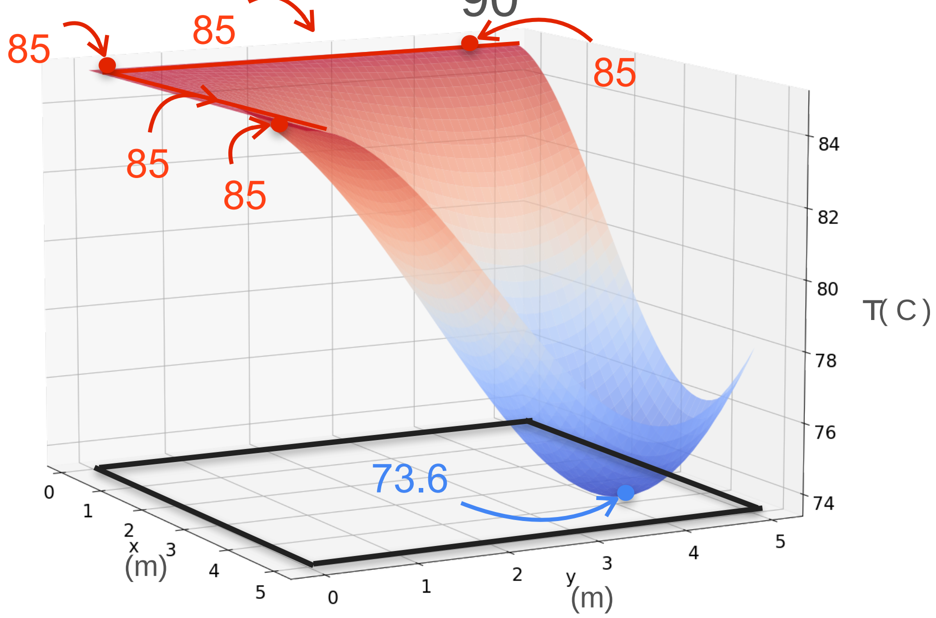



Candidate points for the minima

| | |
|---|---|
| $x = 0$<br>$y = 0$ | Maxima |
| $x = 0, y = 0$<br>$x = 0, y = 4$ | Maxima |
| $x = 0, y = 6$ | Outside |
| $x = 4, y = 0$ | Maxima |
| $x = 4, y = 4$ | |
| $x = 6, y = 0$<br>$x = 6, y = 6$ | Outside |

# Motivation for Optimization in Two Variables

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



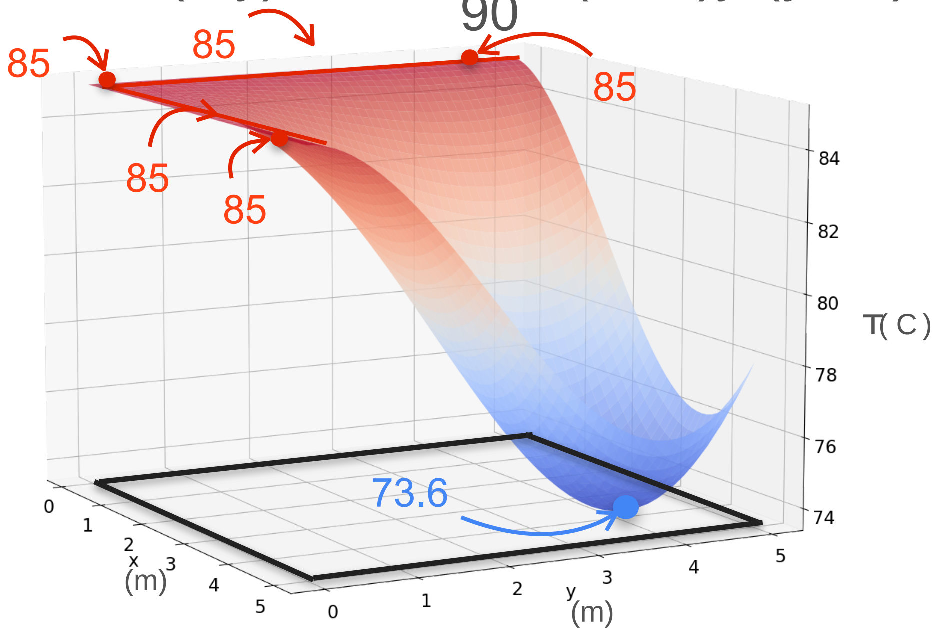

Candidate points for the minima

| | |
|---|---|
| $x = 0$, $y = 0$ | Maxima |
| $x = 0, y = 0$<br>$x = 0, y = 4$ | Maxima |
| $x = 0, y = 6$ | Outside |
| $x = 4, y = 0$ | Maxima |
| $x = 4, y = 4$ | |
| $x = 6, y = 0$<br>$x = 6, y = 6$ | Outside |

DeepLearning.AI

# 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<br>y = 0 | Maxima |
| x = 0, y = 0<br>x = 0, y = 4 | Maxima |
| x = 0, y = 6 | Outside |
| x = 4, y = 0 | Maxima |
| x = 4, y = 4 | Minimum |
| x = 6, y = 0<br>x = 6, y = 6 | Outside |

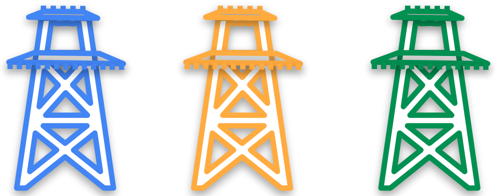
DeepLearning.AI

# Gradients and Gradient Descent

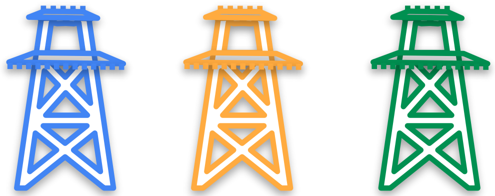
## Optimization using gradients - Analytical method

# Linear Regression: Analytical Approach

# Linear Regression: Analytical Approach

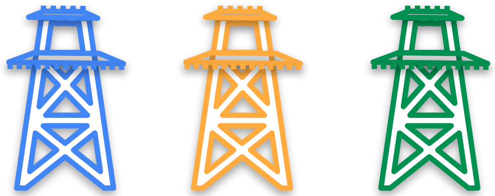


# Linear Regression: Analytical Approach



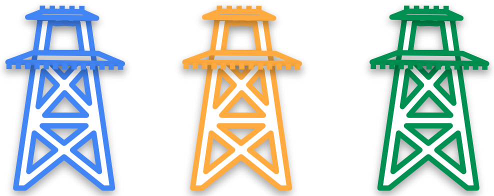
(1, 2) (2, 5) (3,3)

# Linear Regression: Analytical Approach

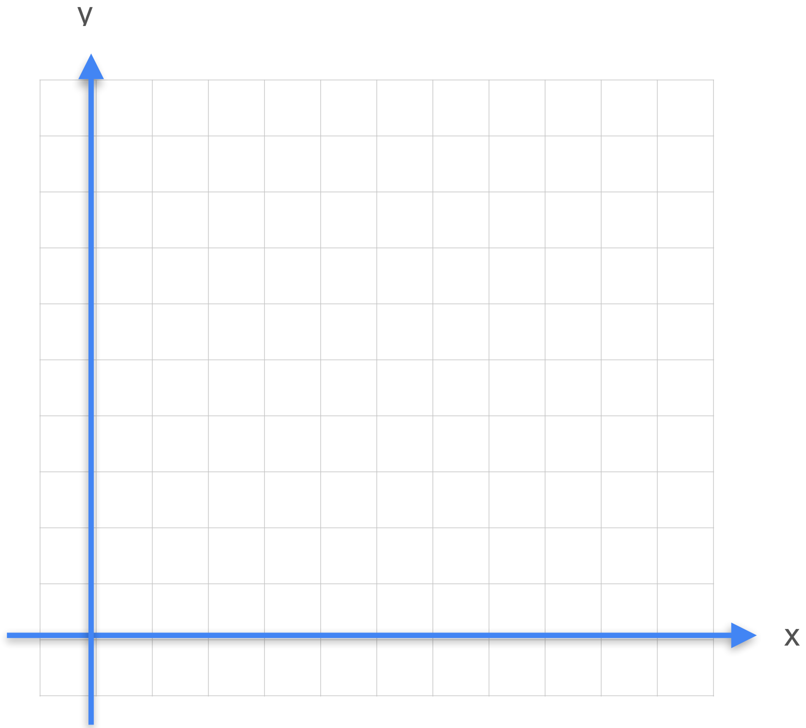


Location on the x,y plane (1, 2) (2, 5) (3,3)

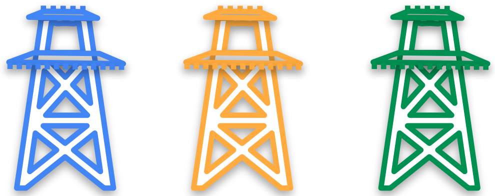
# Linear Regression: Analytical Approach



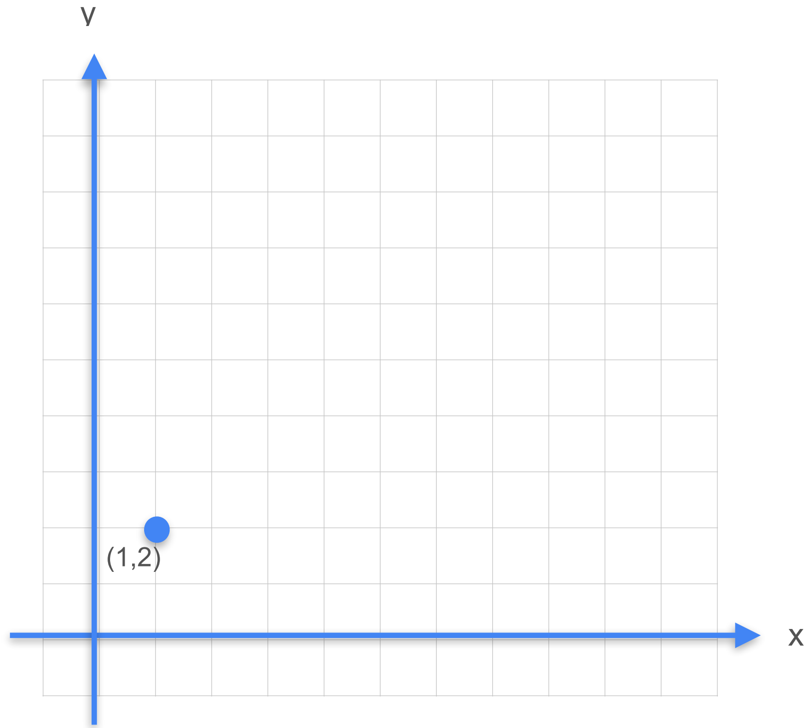
Location on the x,y plane (1, 2) (2, 5) (3,3)



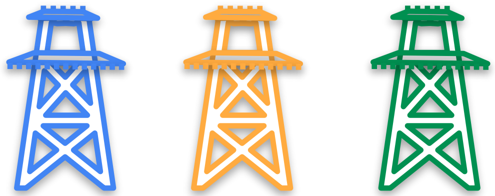
# Linear Regression: Analytical Approach



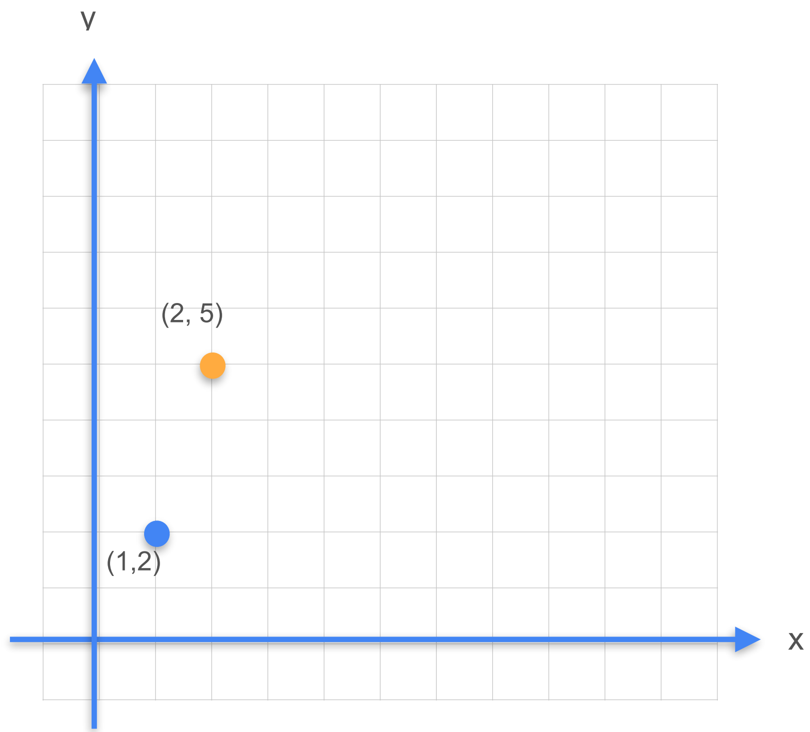
Location on the x,y plane (1, 2) (2, 5) (3,3)



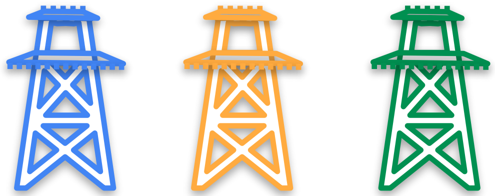
# Linear Regression: Analytical Approach



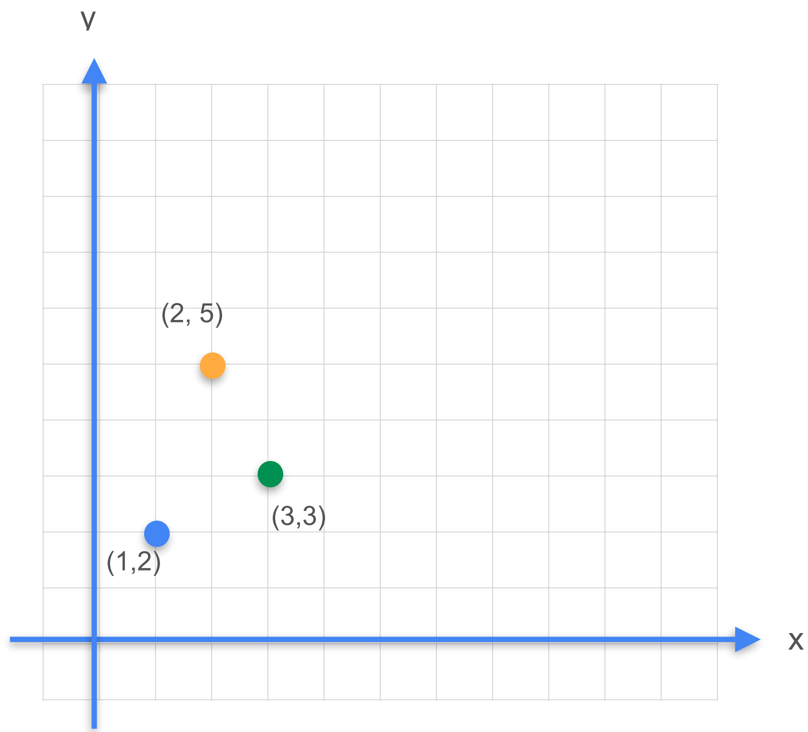
Location on the x,y plane (1, 2) (2, 5) (3,3)



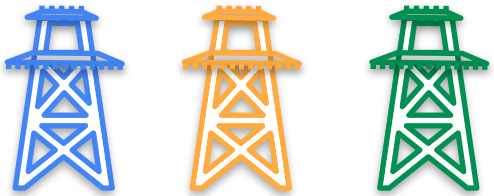
# Linear Regression: Analytical Approach



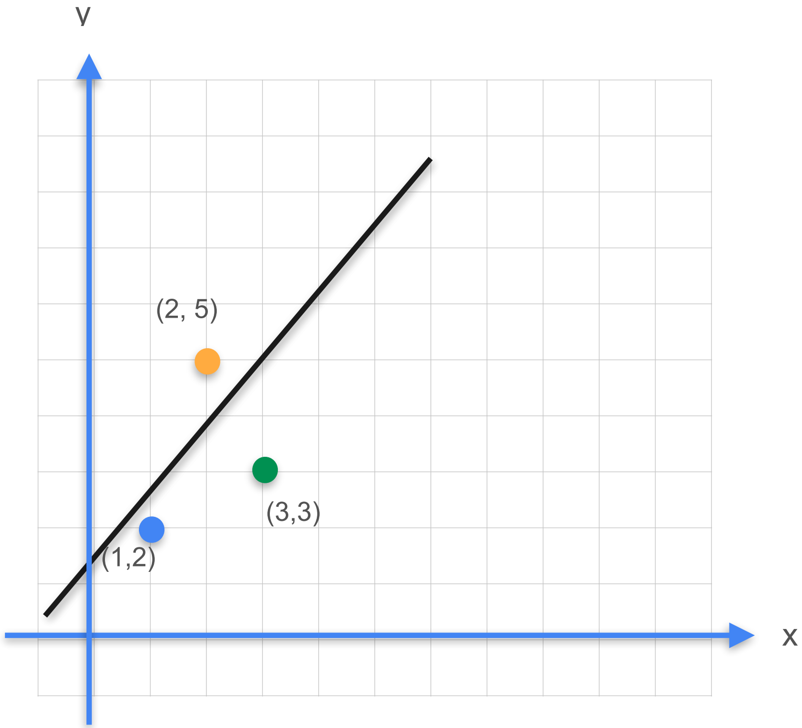
Location on the x,y plane (1, 2) (2, 5) (3,3)



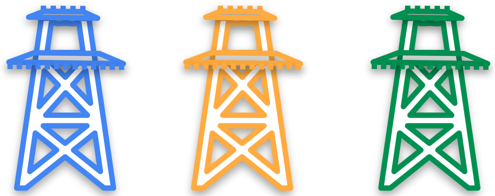
# Linear Regression: Analytical Approach



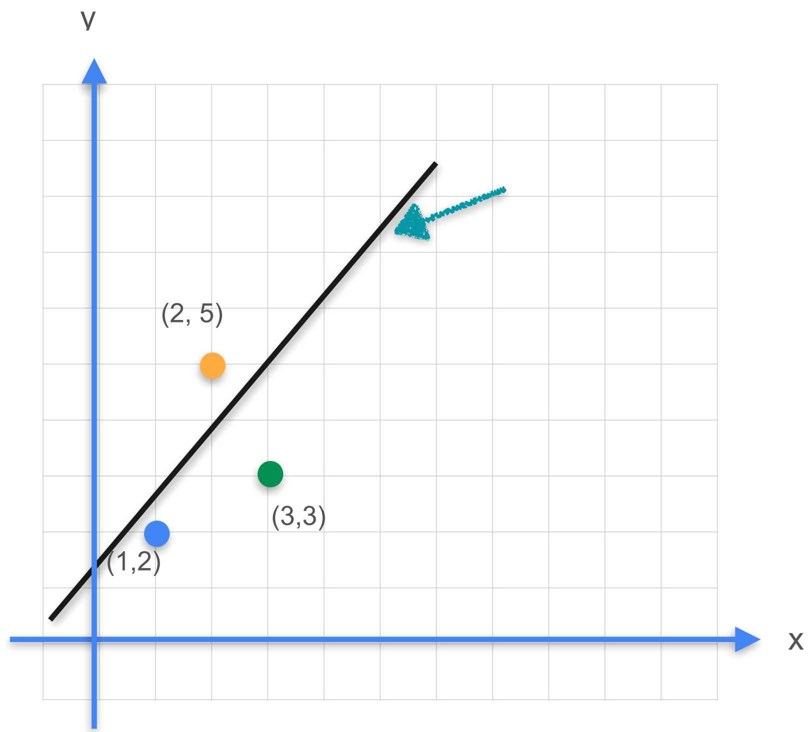
Location on the x,y plane (1, 2) (2, 5) (3,3)



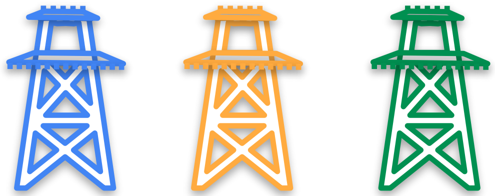
# Linear Regression: Analytical Approach



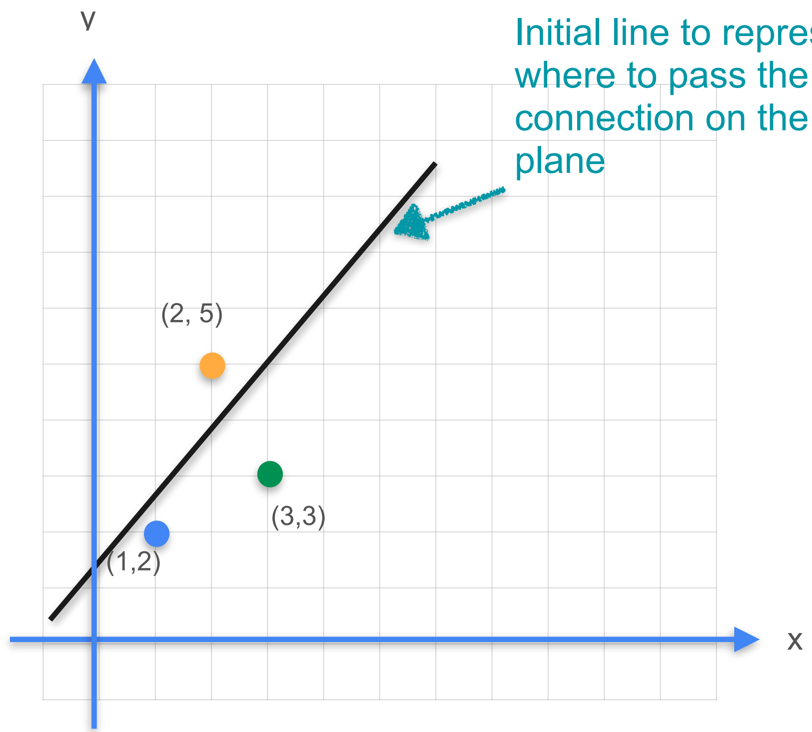
Location on the x,y plane (1, 2) (2, 5) (3,3)



# Linear Regression: Analytical Approach

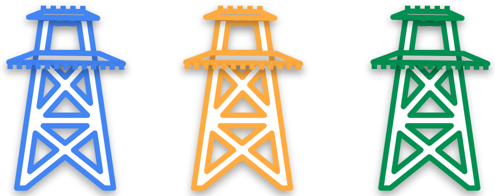


Location on the x,y plane (1, 2) (2, 5) (3,3)

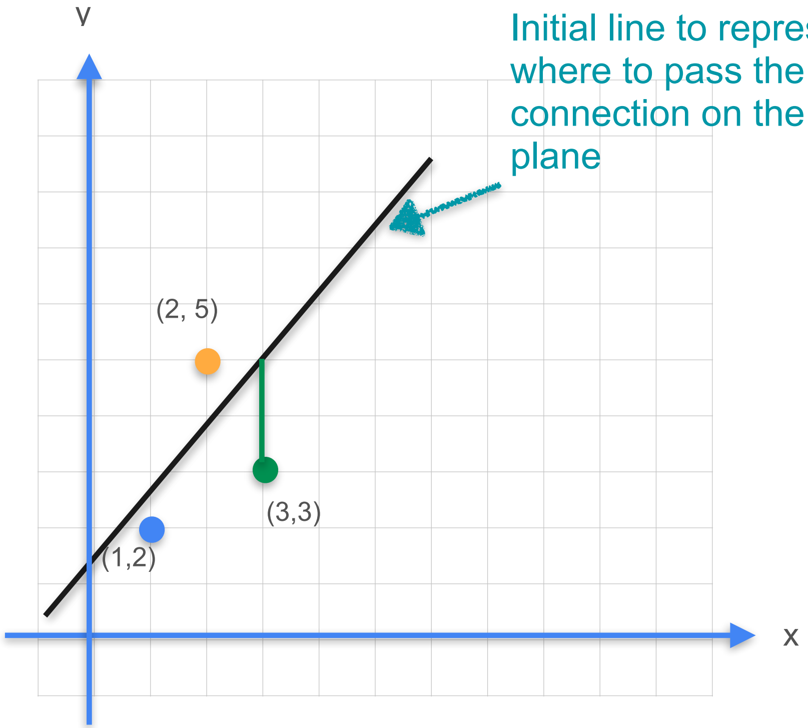



Initial line to represent where to pass the connection on the x-y plane

# Linear Regression: Analytical Approach

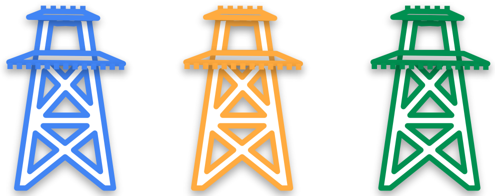


Location on
the x,y plane (1, 2) (2, 5) (3,3)

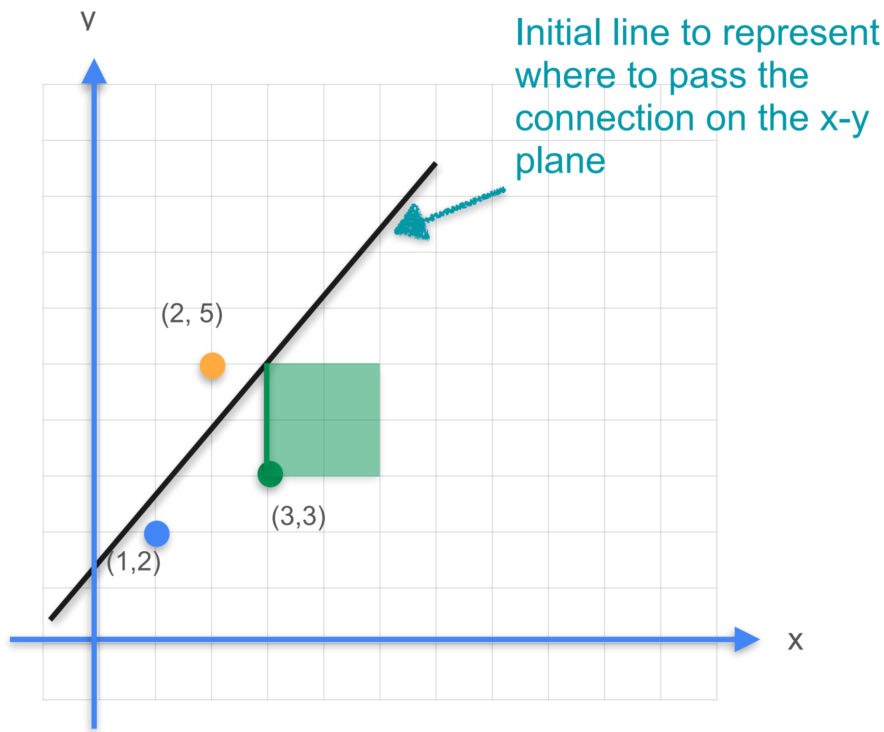



Initial line to represent where to pass the connection on the x-y plane

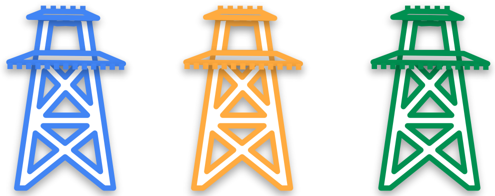
# Linear Regression: Analytical Approach



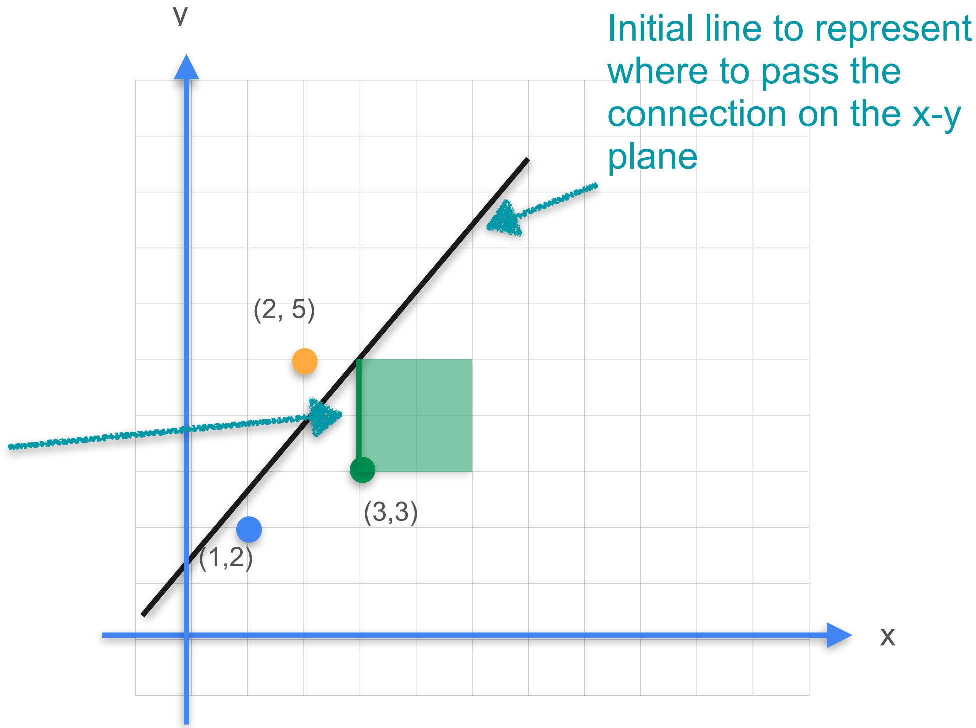
Location on the x,y plane (1, 2) (2, 5) (3,3)



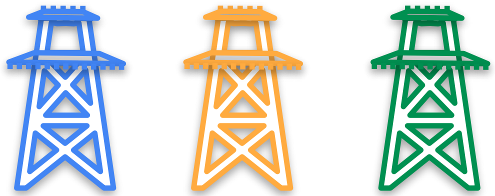
# Linear Regression: Analytical Approach

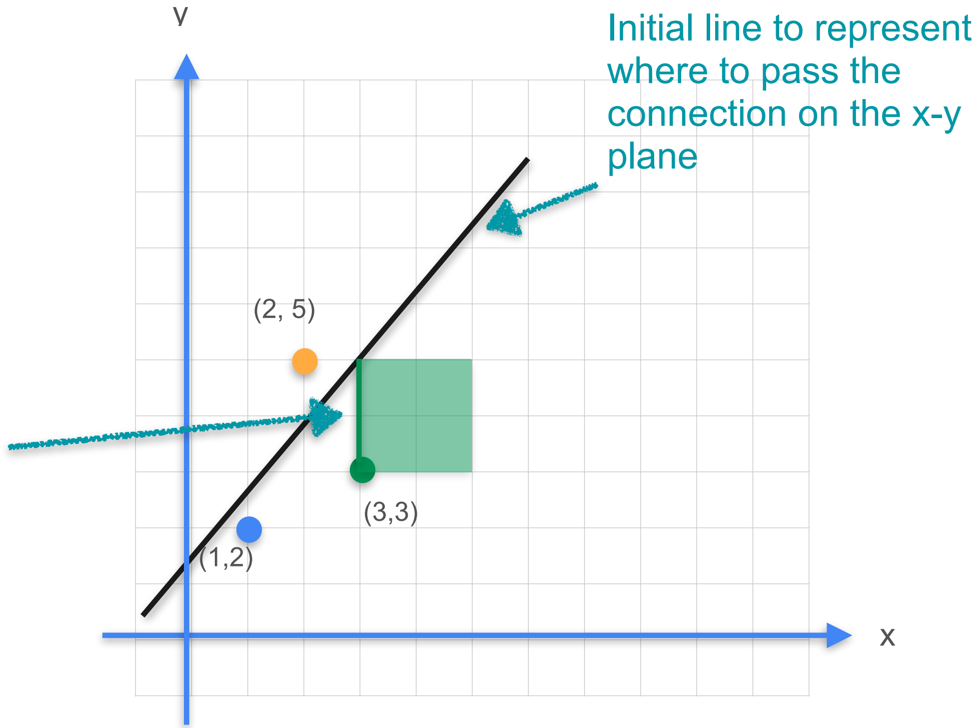


Location on the x,y plane (1, 2) (2, 5) (3,3)

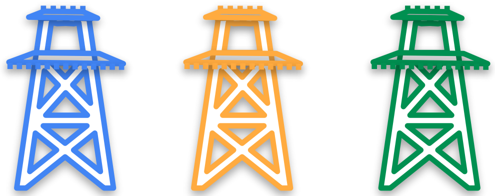


# Linear Regression: Analytical Approach



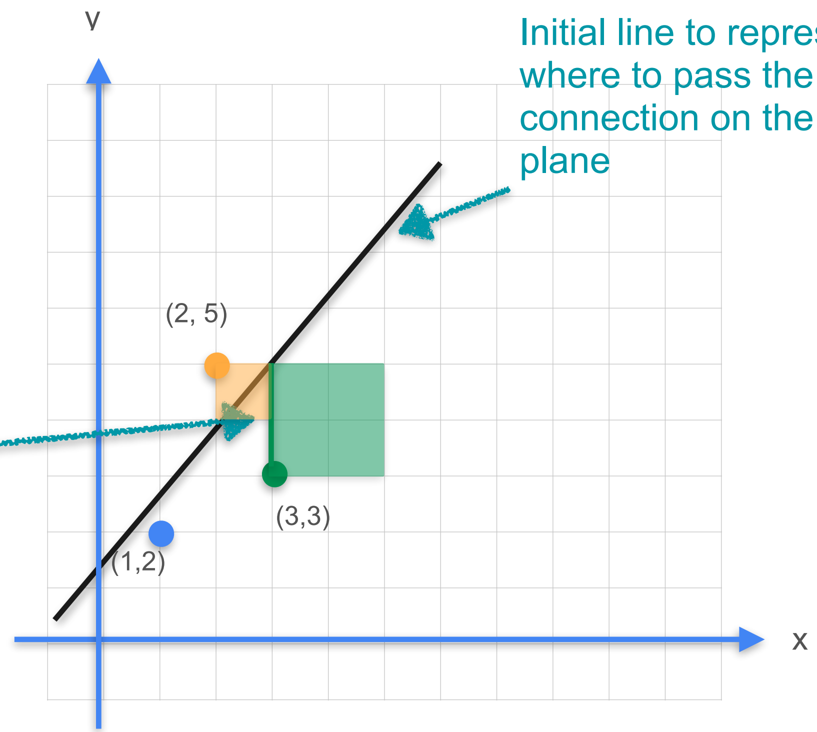



# Linear Regression: Analytical Approach



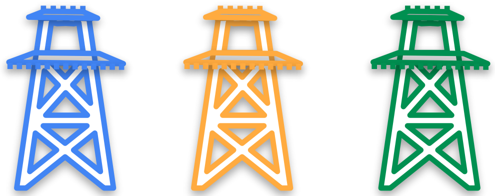
Location on the x,y plane (1, 2) (2, 5) (3,3)

The cost of connecting connection to the powerline



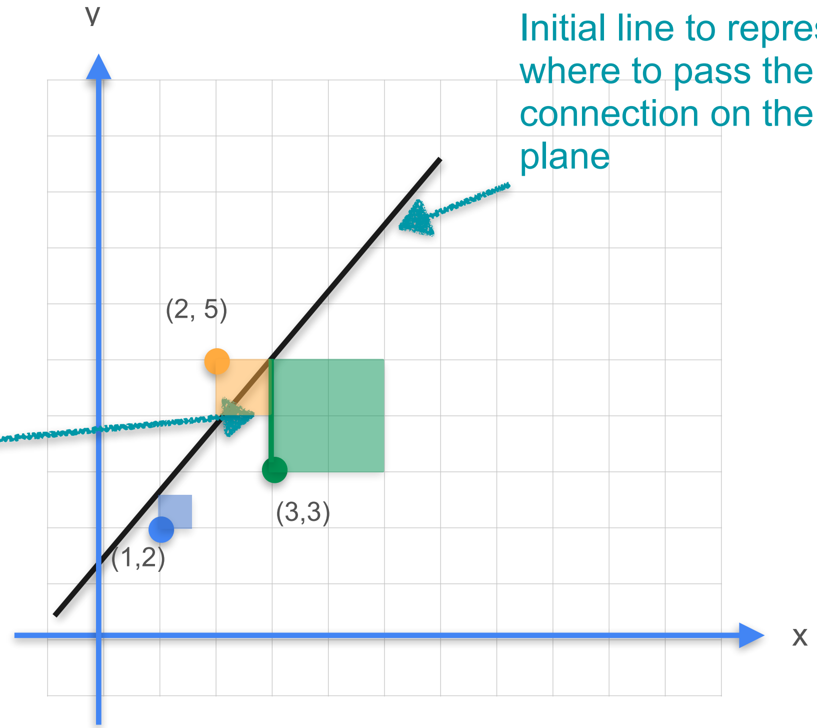

Initial line to represent where to pass the connection on the x-y plane

# Linear Regression: Analytical Approach



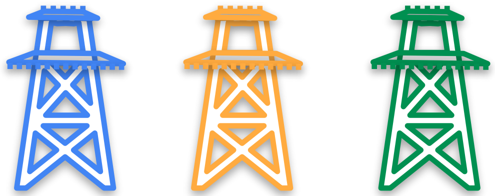
Location on the x,y plane (1, 2) (2, 5) (3,3)

The cost of connecting connection to the powerline



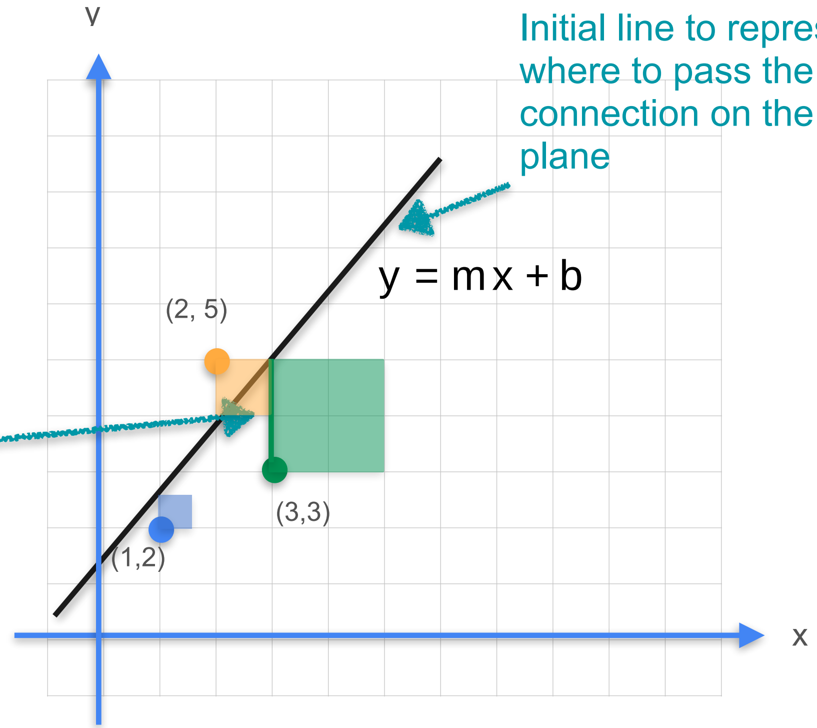

Initial line to represent where to pass the connection on the x-y plane

# Linear Regression: Analytical Approach



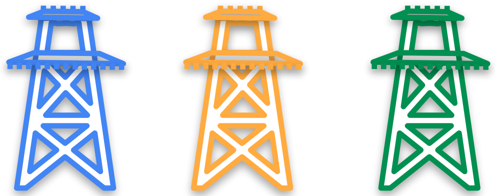
Location on the x,y plane (1, 2) (2, 5) (3,3)

The cost of connecting connection to the powerline

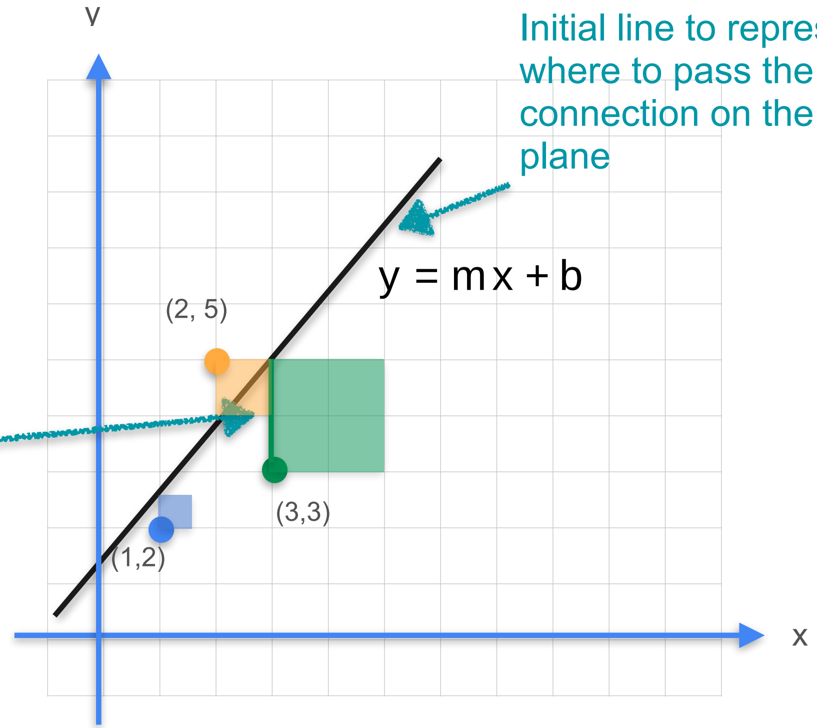



Initial line to represent where to pass the connection on the x-y plane

# Linear Regression: Analytical Approach



Location on the x,y plane (1, 2) (2, 5) (3,3)

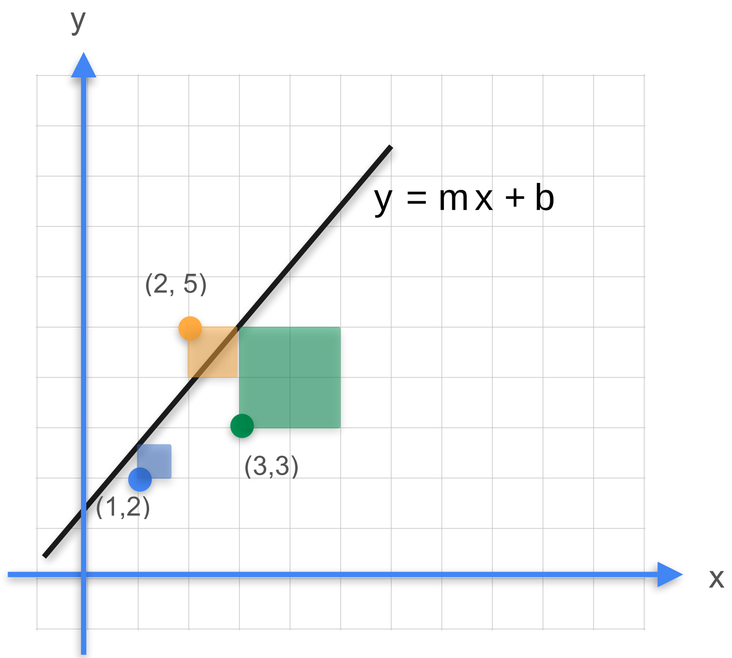



Initial line to represent where to pass the connection on the x-y plane

The cost of connecting connection to the powerline

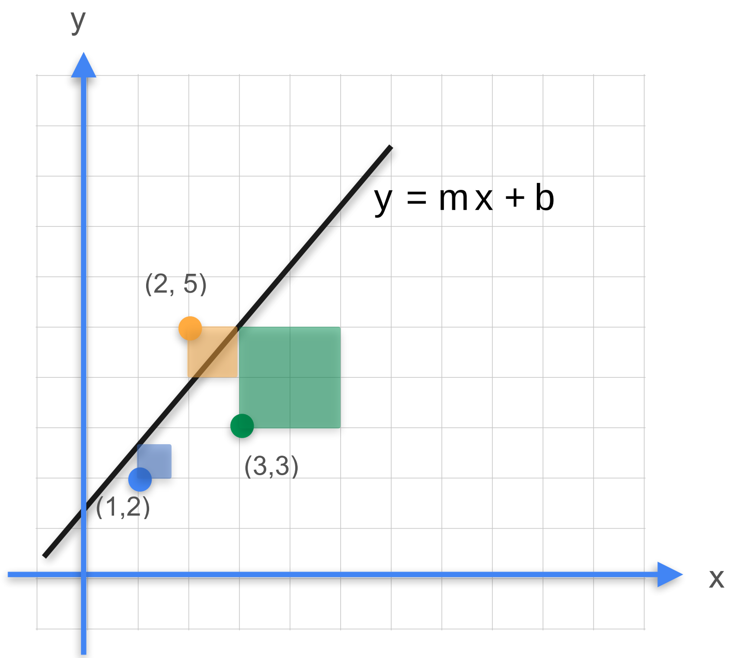
**Goal: Find m, b such that you minimize sum of squares cost**

# Linear Regression: Analytical Approach



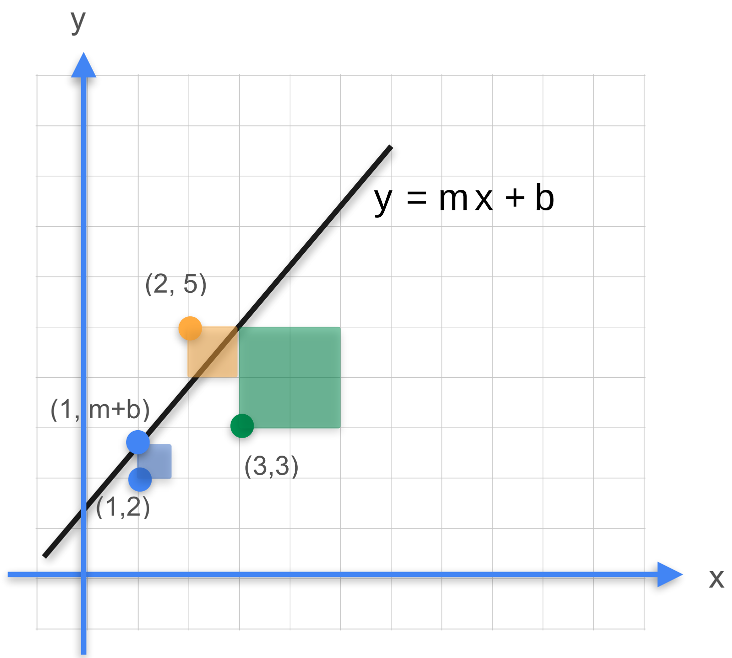

DeepLearning.AI

# Linear Regression: Analytical Approach



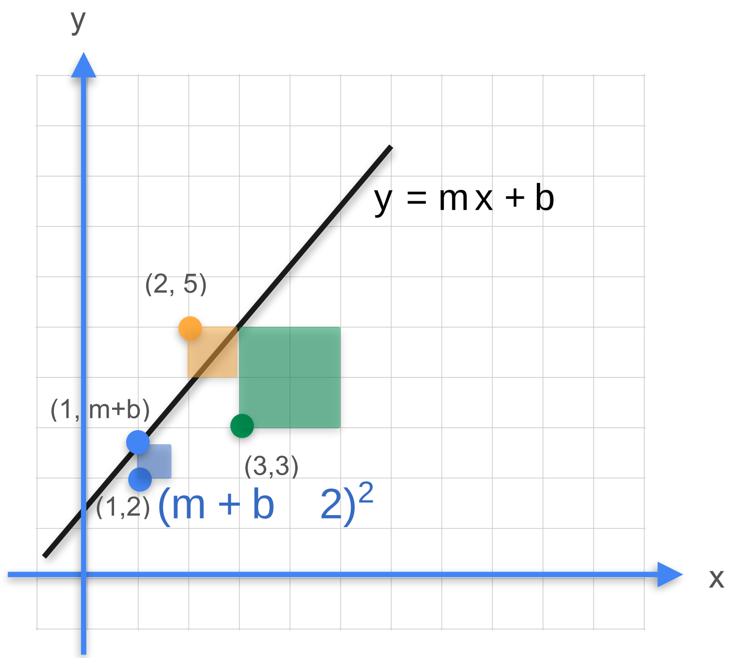

**Goal: Minimize sum of squares cost**

# Linear Regression: Analytical Approach



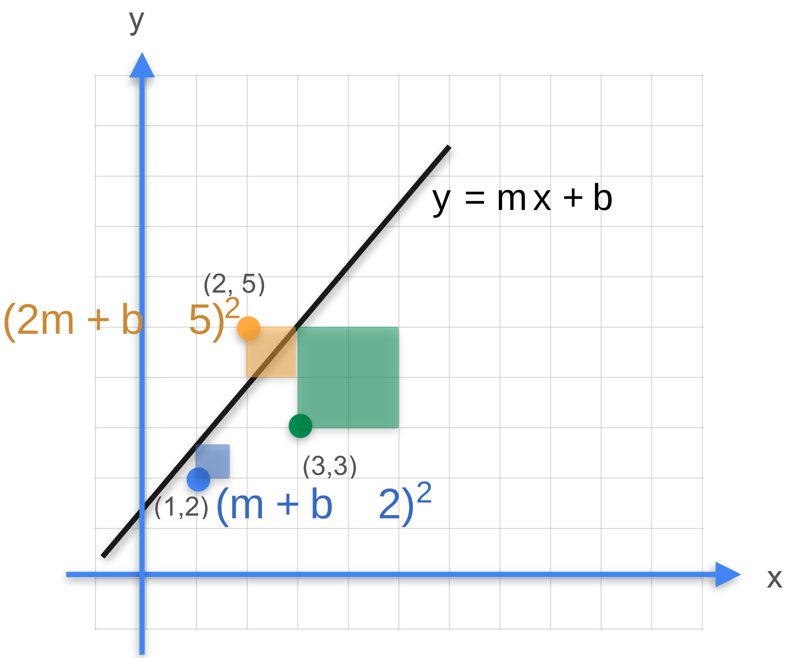

**Goal: Minimize sum of squares cost**

# Linear Regression: Analytical Approach



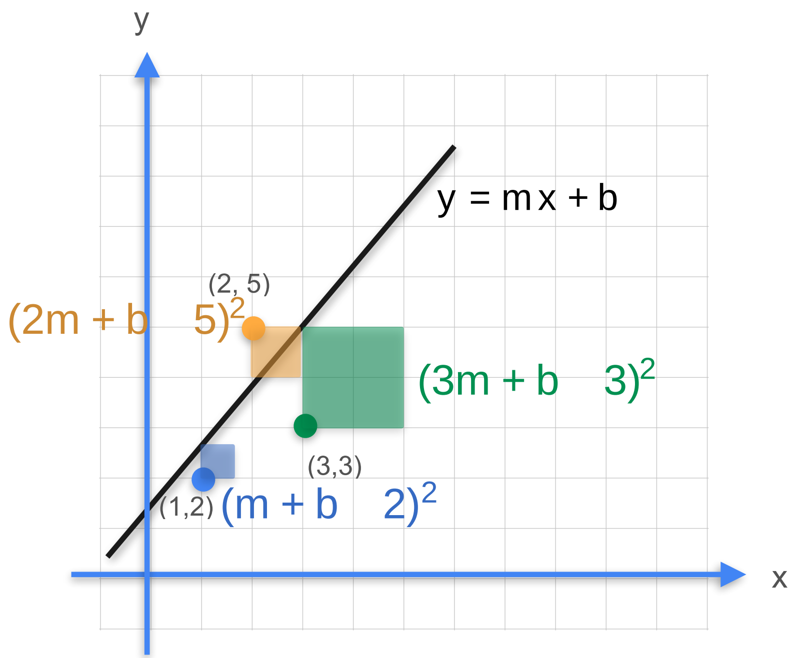

**Goal: Minimize sum of squares cost**

# Linear Regression: Analytical Approach



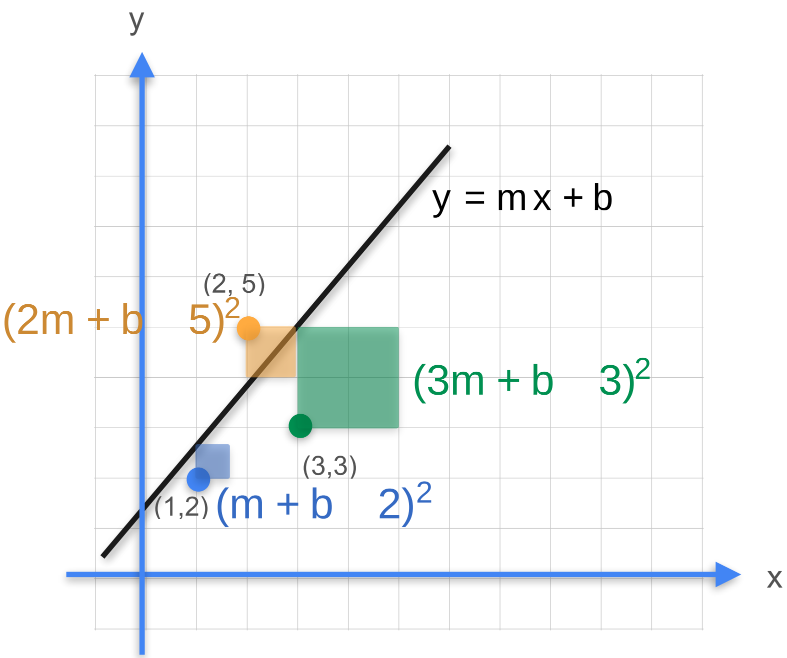

**Goal: Minimize sum of squares cost**

# Linear Regression: Analytical Approach




**Goal: Minimize sum of squares cost**

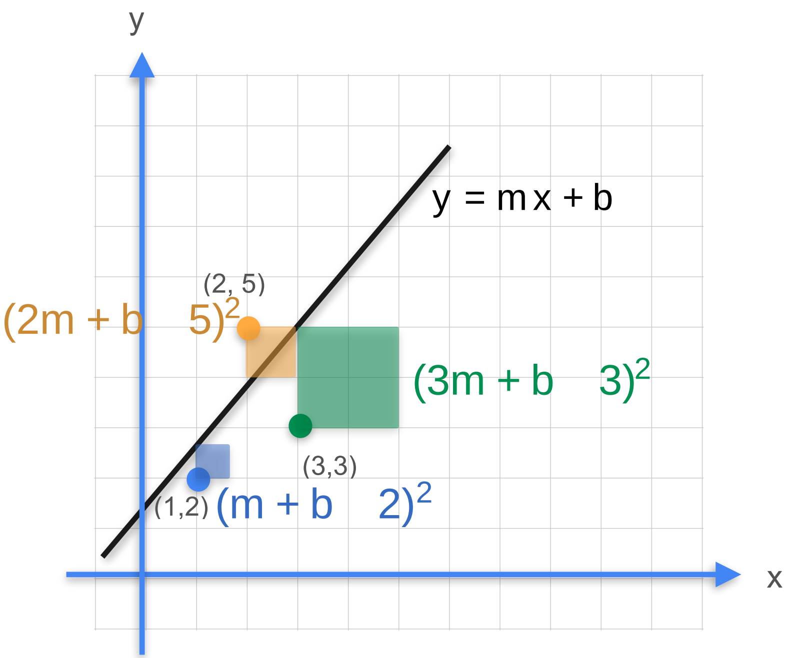
# Linear Regression: Analytical Approach




**Goal: Minimize sum of squares cost**

$$(m + b \quad 2)^2 + (2m + b \quad 5)^2 + (3m + b \quad 3)^2$$

# Linear Regression: Analytical Approach

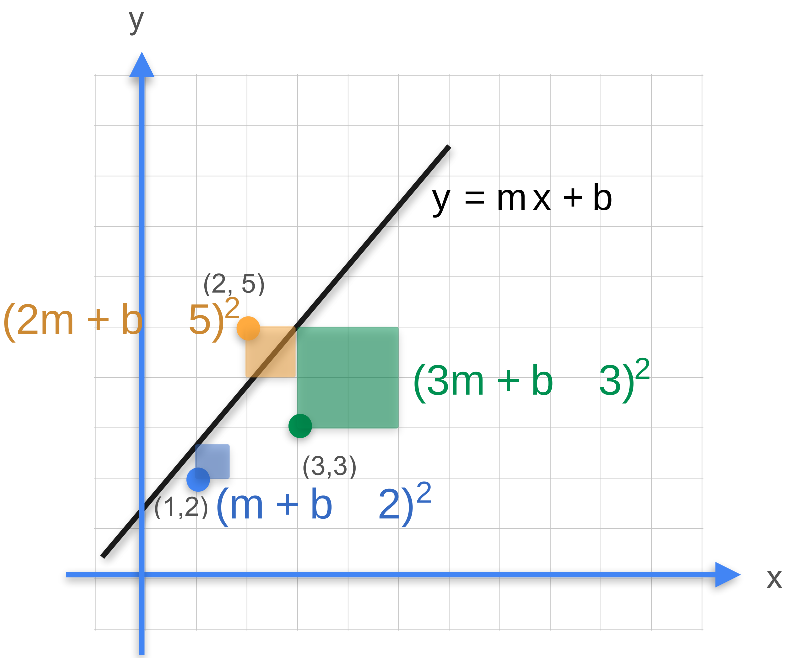



**Goal: Minimize sum of squares cost**

$$(m + b \quad 2)^2 + (2m + b \quad 5)^2 + (3m + b \quad 3)^2$$

$$m^2 + b^2 + 4 + 2mb \quad 4m \quad 4b$$

# Linear Regression: Analytical Approach



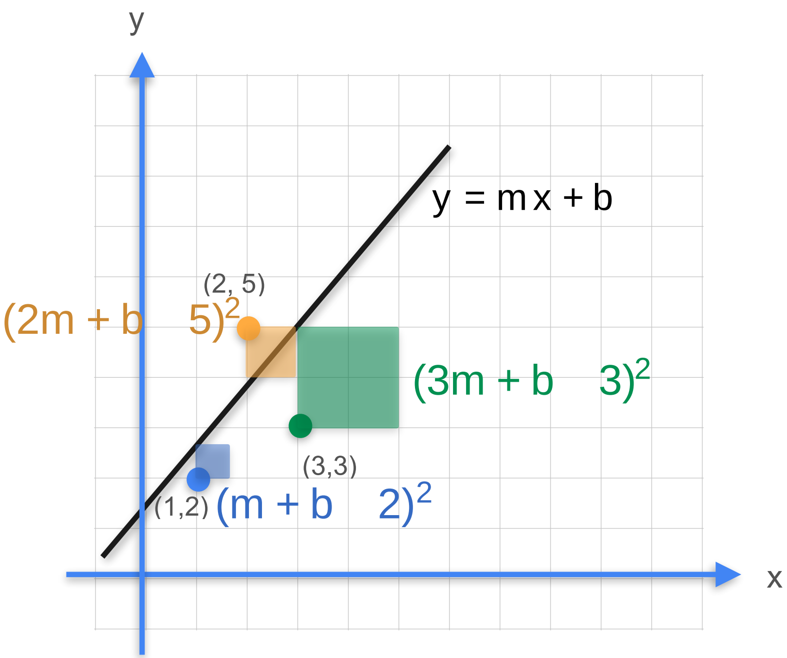

**Goal: Minimize sum of squares cost**

$$(m + b \quad 2)^2 + (2m + b \quad 5)^2 + (3m + b \quad 3)^2$$

$$m^2 + b^2 + 4 + 2mb \quad 4m \quad 4b$$

$$+4m^2 + b^2 + 25 + 4mb \quad 20m \quad 10b$$

# Linear Regression: Analytical Approach




**Goal: Minimize sum of squares cost**

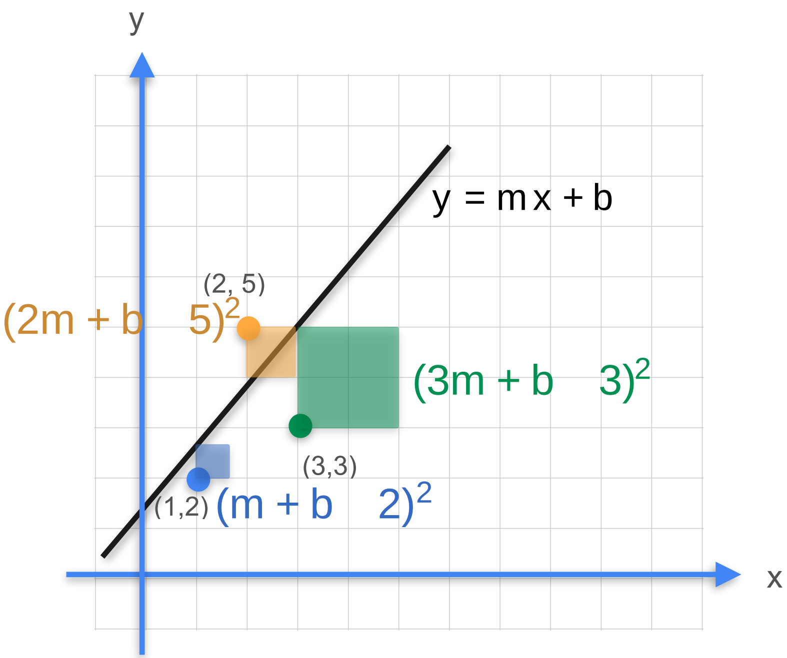
$$(m + b \quad 2)^2 \;+\; (2m + b \quad 5)^2 \;+\; (3m + b \quad 3)^2$$

$$m^2 + b^2 + 4 + 2mb \quad 4m \quad 4b$$

$$+4m^2 + b^2 + 25 + 4mb \quad 20m \quad 10b$$

$$+9m^2 + b^2 + 9 + 6mb \quad 18m \quad 6b$$

# Linear Regression: Analytical Approach




**Goal: Minimize sum of squares cost**

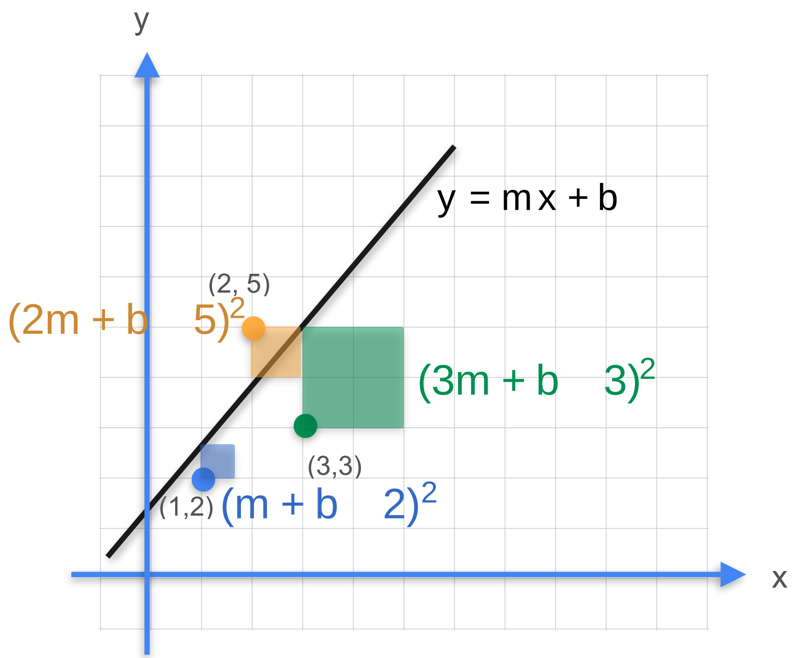
$$(m + b \quad 2)^2 + (2m + b \quad 5)^2 + (3m + b \quad 3)^2$$

$$m^2 + b^2 + 4 + 2mb \quad 4m \quad 4b$$

$$+4m^2 + b^2 + 25 + 4mb \quad 20m \quad 10b$$

$$+9m^2 + b^2 + 9 + 6mb \quad 18m \quad 6b$$

# Linear Regression: Analytical Approach




**Goal: Minimize sum of squares cost**

$$(m + b \quad 2)^2 + (2m + b \quad 5)^2 + (3m + b \quad 3)^2$$

$$m^2 + b^2 + 4 + 2mb \quad 4m \quad 4b$$

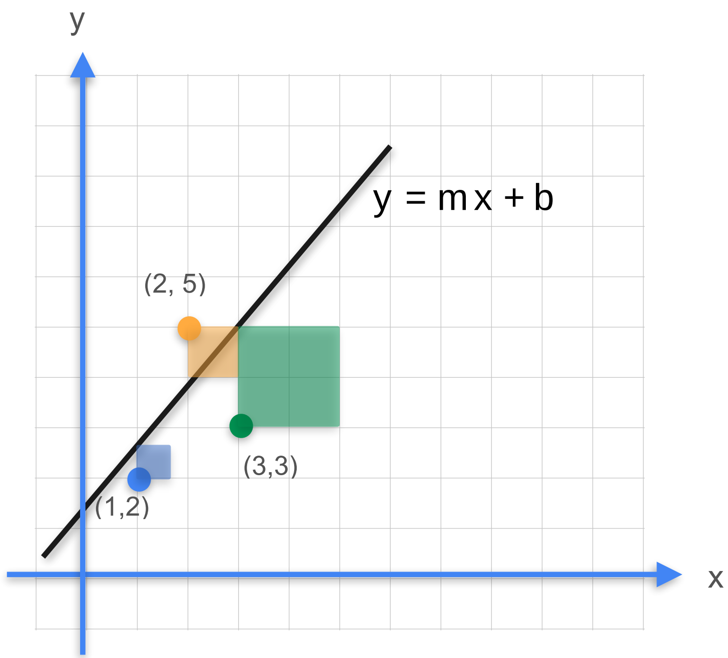
$$+4m^2 + b^2 + 25 + 4mb \quad 20m \quad 10b$$

$$+9m^2 + b^2 + 9 + 6mb \quad 18m \quad 6b$$

---

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

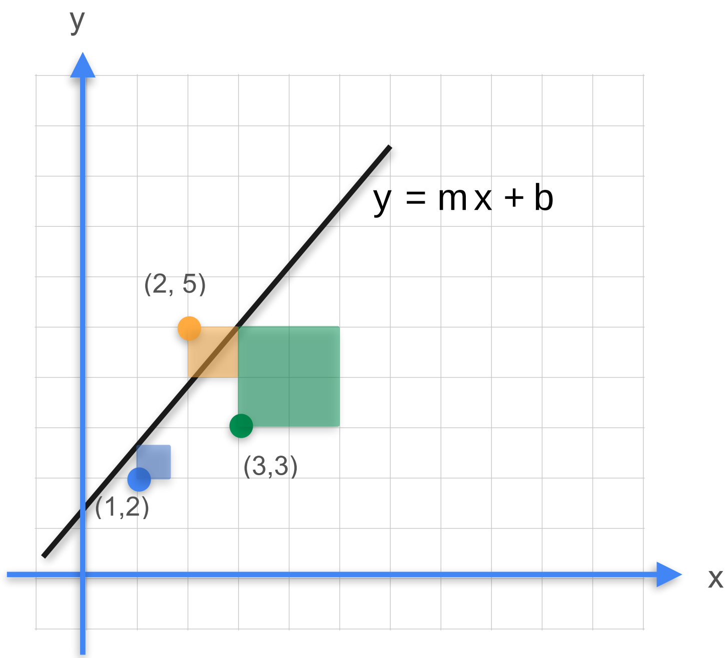
# Linear Regression: Analytical Approach




**Goal: Minimize sum of squares cost**

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

# Linear Regression: Analytical Approach

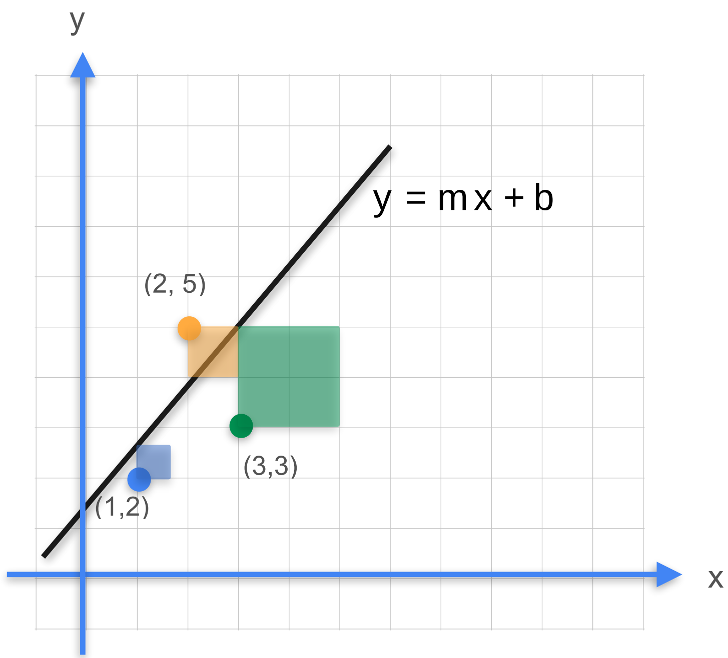



**Goal: Minimize sum of squares cost**

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

$$\frac{E}{m} = 0$$

# Linear Regression: Analytical Approach



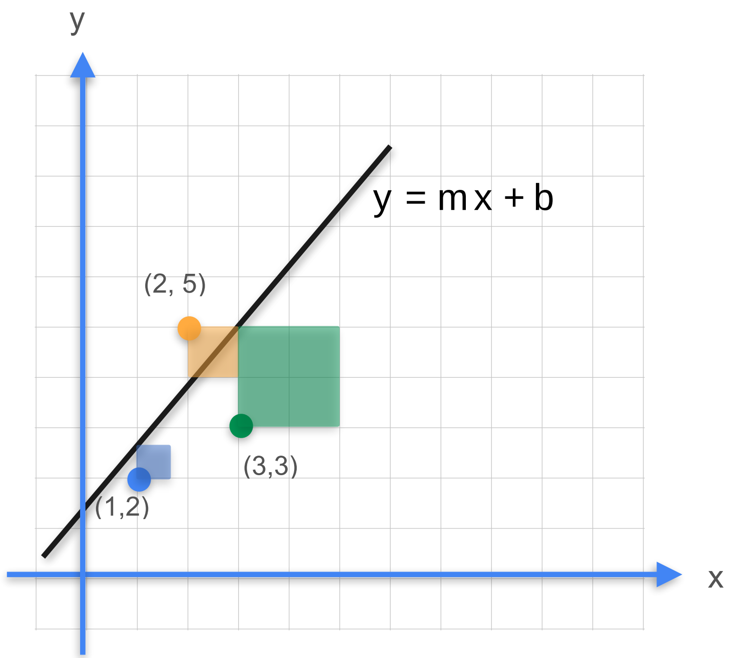

**Goal: Minimize sum of squares cost**

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

$$\frac{E}{m} = 0$$

$$\frac{E}{b} = 0$$

# Linear Regression: Analytical Approach




**Goal: Minimize sum of squares cost**

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

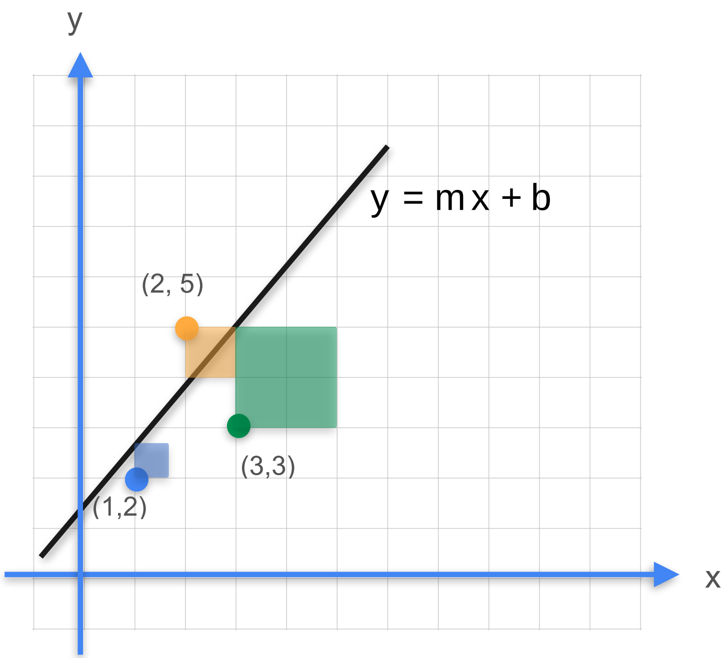
$$\frac{E}{m} = 0$$

$$\frac{E}{b} = 0$$

**Quiz:**

**Find the partial derivative of $E$ with respect to $m$**

# Linear Regression: Analytical Approach



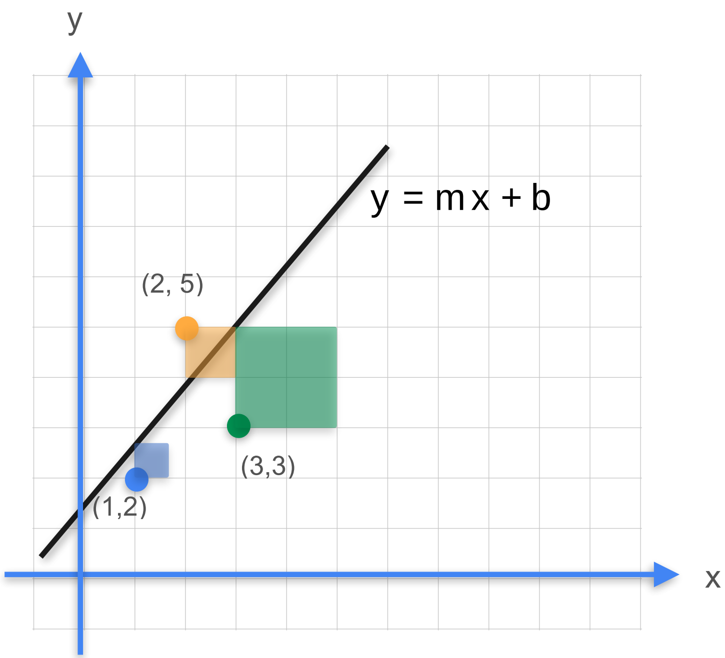

**Goal: Minimize sum of squares cost**

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

$$\frac{E}{m} = 28m + 12b \quad 42$$

$$\frac{E}{b} =$$

# Linear Regression: Analytical Approach




**Goal: Minimize sum of squares cost**

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

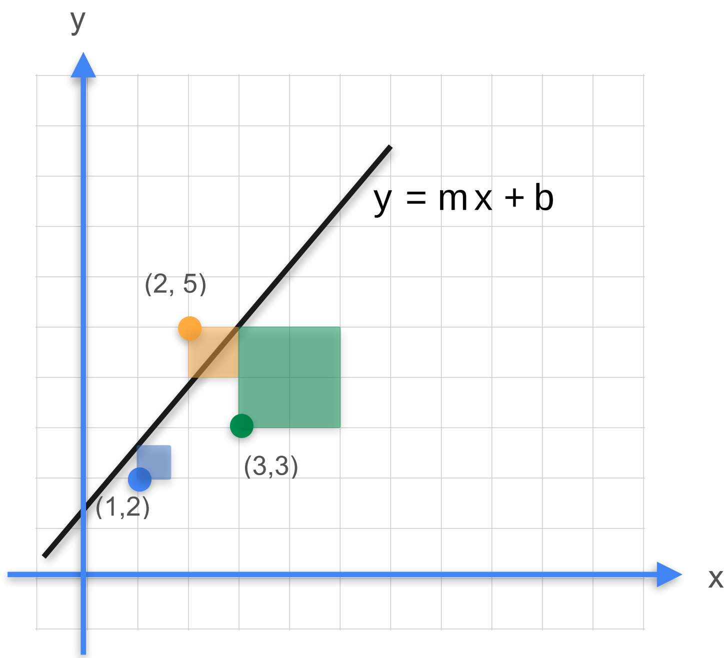
$$\frac{E}{m} = 28m + 12b \quad 42$$

$$\frac{E}{b} =$$

**Quiz:**

**Find the partial derivative of** E **with respect to** b

# Linear Regression: Analytical Approach



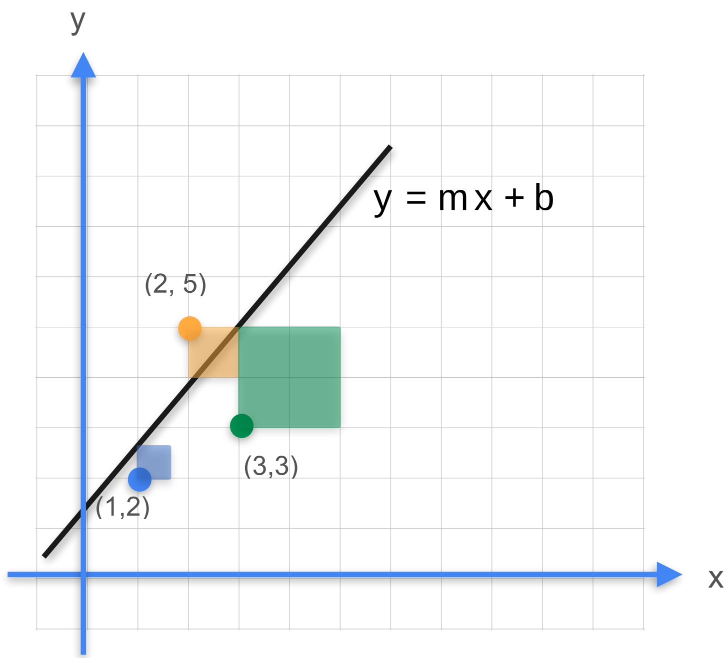

**Goal: Minimize sum of squares cost**

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

$$\frac{E}{m} = 28m + 12b \quad 42$$

$$\frac{E}{b} = 6b + 12m \quad 20$$

# Linear Regression: Analytical Approach

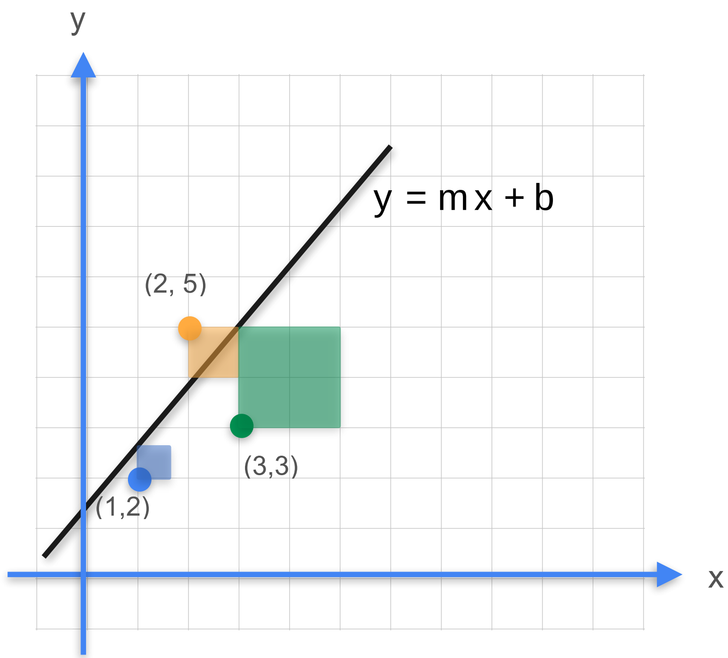



**Goal: Minimize sum of squares cost**

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

$$\boxed{\begin{aligned} \frac{E}{m} &= 28m + 12b \quad 42 \\ \frac{E}{b} &= 6b + 12m \quad 20 \end{aligned}}$$

# Linear Regression: Analytical Approach



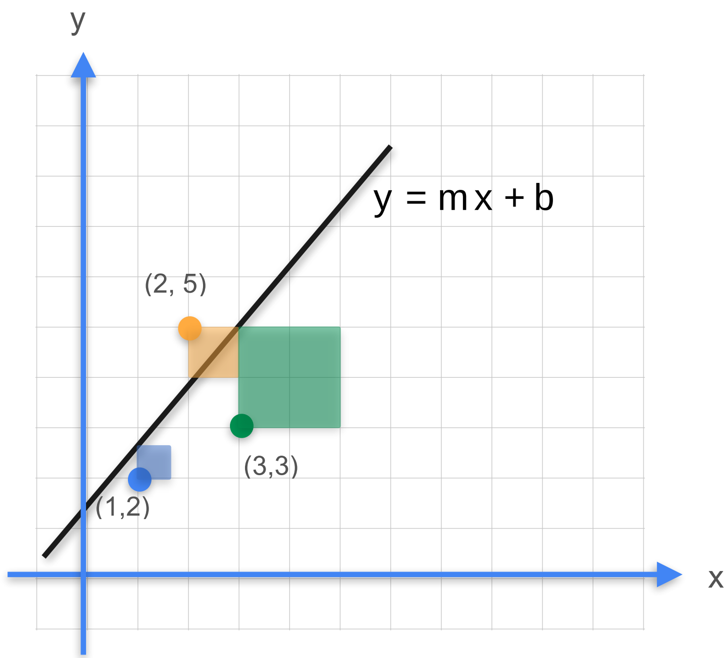

**Goal: Minimize sum of squares cost**

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

$$\frac{E}{m} = 28m + 12b \quad 42 \qquad = 0$$

$$\frac{E}{b} = 6b + 12m \quad 20$$

# Linear Regression: Analytical Approach



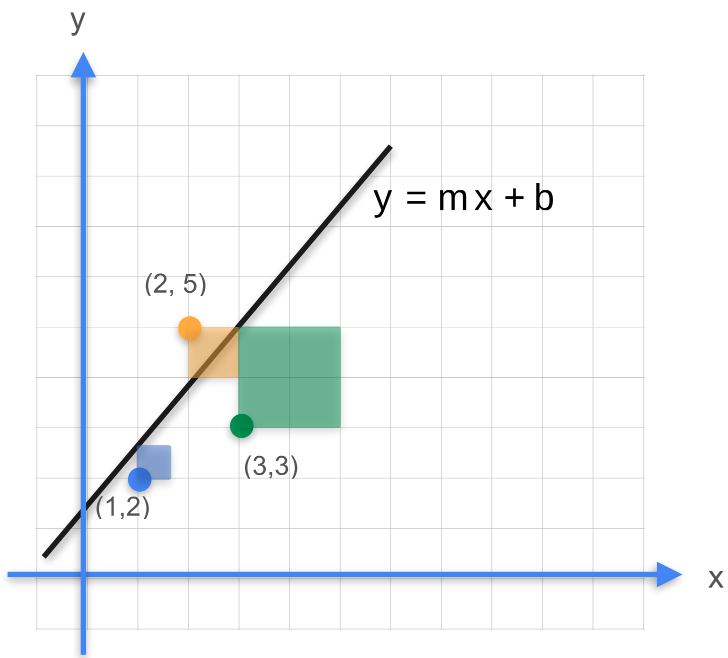

**Goal: Minimize sum of squares cost**

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

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

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

# Linear Regression: Analytical Approach




**Goal: Minimize sum of squares cost**

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

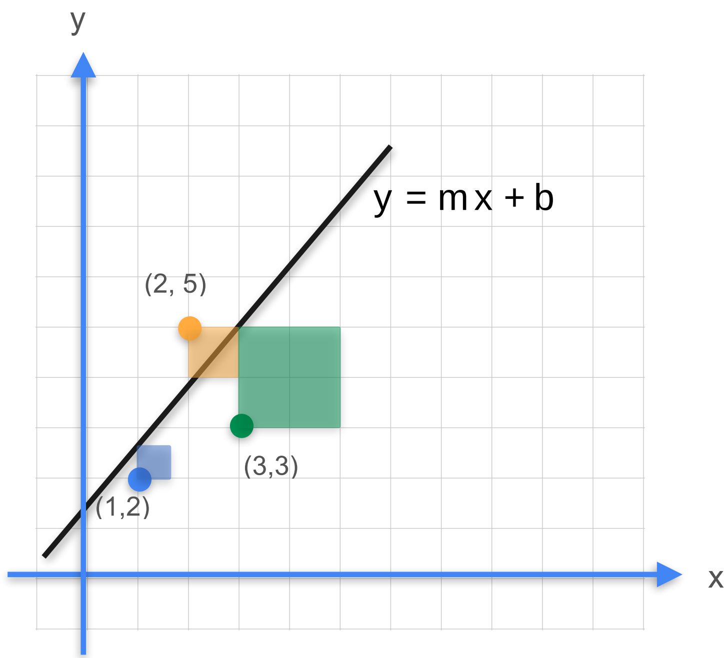
$$\frac{E}{m} = \boxed{28m + 12b \quad 42} = 0$$

$$\frac{E}{b} = \boxed{6b + 12m \quad 20} = 0$$

$m =$

DeepLearning.AI

# Linear Regression: Analytical Approach




**Goal: Minimize sum of squares cost**

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

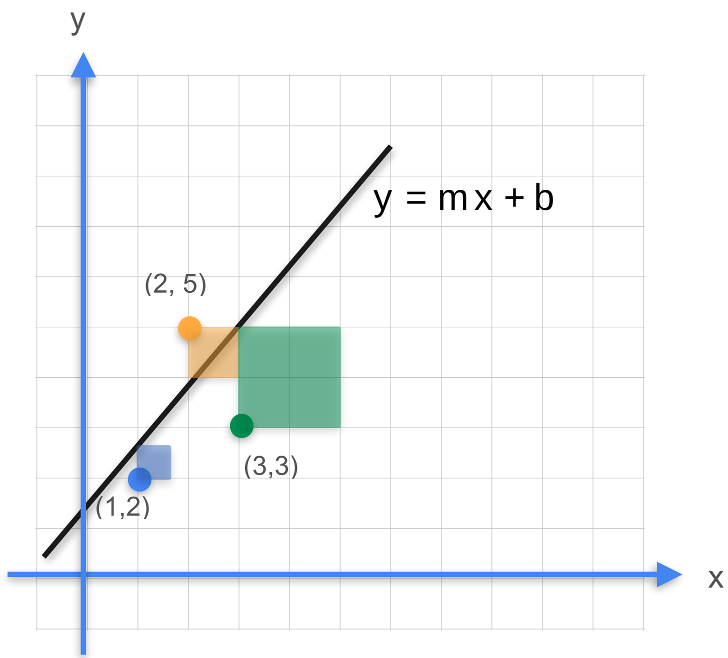
$$\frac{E}{m} = 28m + 12b \quad 42 = 0$$

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

$m =$

$b =$

# Linear Regression: Analytical Approach




**Goal: Minimize sum of squares cost**

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

$$\frac{E}{m} = \boxed{28m + 12b \quad 42} = 0$$

$$\frac{E}{b} = \boxed{6b + 12m \quad 20} = 0$$

m =
b =
?

# Linear Regression: Analytical Approach

**Goal: Minimize sum of squares cost**

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

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

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

$m =$

$b =$

?

# Linear Regression: Analytical Approach

**Goal: Minimize sum of squares cost**

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

$$\frac{E}{m} = \boxed{28m + 12b \quad 42} = 0$$

$$\frac{E}{b} = \boxed{6b + 12m \quad 20} = 0$$

$m =$
$b =$ ?

# Linear Regression: Analytical Approach

**Goal: Minimize sum of squares cost**

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

$$\frac{E}{m} = \boxed{28m + 12b \quad 42} = 0$$

$$\frac{E}{b} = \boxed{6b + 12m \quad 20} = 0$$

$m =$

$b =$ **?**

$$12b + 24m \quad 40 = 0$$

# Linear Regression: Analytical Approach

**Goal: Minimize sum of squares cost**

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

$$\frac{E}{m} = \boxed{28m + 12b \quad 42} = 0$$

$$\frac{E}{b} = \boxed{6b + 12m \quad 20} = 0$$

$m =$

$b =$ ?

$$12b + 24m \quad 40 = 0$$

$$4m \quad 2 = 0$$

# Linear Regression: Analytical Approach

**Goal: Minimize sum of squares cost**

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

$$\frac{E}{m} = \boxed{28m + 12b \quad 42} = 0$$

$$\frac{E}{b} = \boxed{6b + 12m \quad 20} = 0$$

$m =$

$b =$ **?**

$$12b + 24m \quad 40 = 0$$

$$4m \quad 2 = 0$$

$$m = \frac{2}{4} = 0.5$$

# Linear Regression: Analytical Approach

**Goal: Minimize sum of squares cost**

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

$$\frac{E}{m} = \boxed{28m + 12b \quad 42} = 0$$

$$\frac{E}{b} = \boxed{6b + 12m \quad 20} = 0$$

$m =$

$b =$ **?**

$$12b + 24m \quad 40 = 0$$

$$4m \quad 2 = 0$$

$$m = \frac{2}{4} = 0.5$$

$$6b + 12(0.5) \quad 20 = 0$$

# Linear Regression: Analytical Approach

**Goal: Minimize sum of squares cost**

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

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

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

$m =$

$b =$ ?

$$12b + 24m \quad 40 = 0$$

$$4m \quad 2 = 0$$

$$m = \frac{2}{4} = 0.5$$

$$6b + 12(0.5) \quad 20 = 0$$

$$6b + 6 \quad 20 = 0$$

# Linear Regression: Analytical Approach

**Goal: Minimize sum of squares cost**

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

$$\frac{E}{m} = \boxed{28m + 12b \quad 42} = 0$$

$$\frac{E}{b} = \boxed{6b + 12m \quad 20} = 0$$

$m =$

$b =$ ?

$$12b + 24m \quad 40 = 0$$

$$4m \quad 2 = 0$$

$$m = \frac{2}{4} = 0.5$$

$$6b + 12(0.5) \quad 20 = 0$$

$$6b + 6 \quad 20 = 0$$

$$6b \quad 14 = 0$$

# Linear Regression: Analytical Approach

**Goal: Minimize sum of squares cost**

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

$$\frac{E}{m} = \boxed{28m + 12b \quad 42} = 0$$

$$\frac{E}{b} = \boxed{6b + 12m \quad 20} = 0$$

$m =$

$b =$ **?**

$$12b + 24m \quad 40 = 0$$

$$4m \quad 2 = 0$$

$$m = \frac{2}{4} = 0.5$$

$$6b + 12(0.5) \quad 20 = 0$$

$$6b + 6 \quad 20 = 0$$

$$6b \quad 14 = 0$$

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

# Linear Regression: Analytical Approach

**Goal: Minimize sum of squares cost**

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

$$\frac{E}{m} = \boxed{28m + 12b \quad 42} = 0$$

$$\frac{E}{b} = \boxed{6b + 12m \quad 20} = 0$$

$m =$

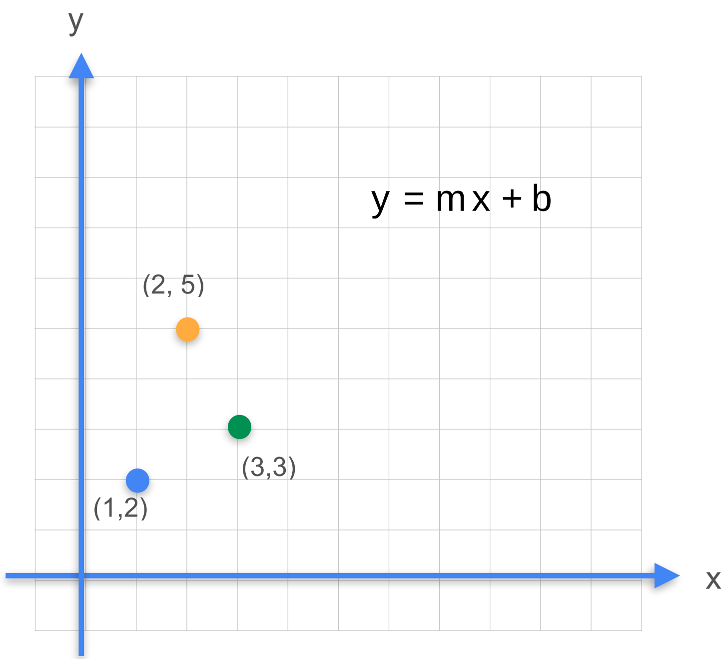
$b =$ ?

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

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

$$E(m = \frac{1}{2}, b = \frac{7}{3}) \quad 4.167$$

# Linear Regression: Optimal Solution

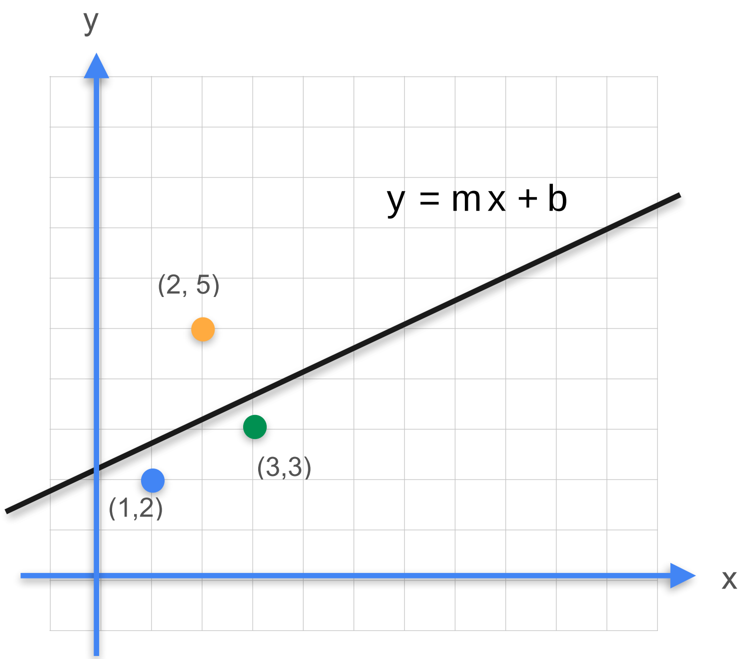



$$m = \frac{1}{2}$$

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

$$E(m = \frac{1}{2}, b = \frac{7}{3}) \quad 4.167$$

# Linear Regression: Optimal Solution

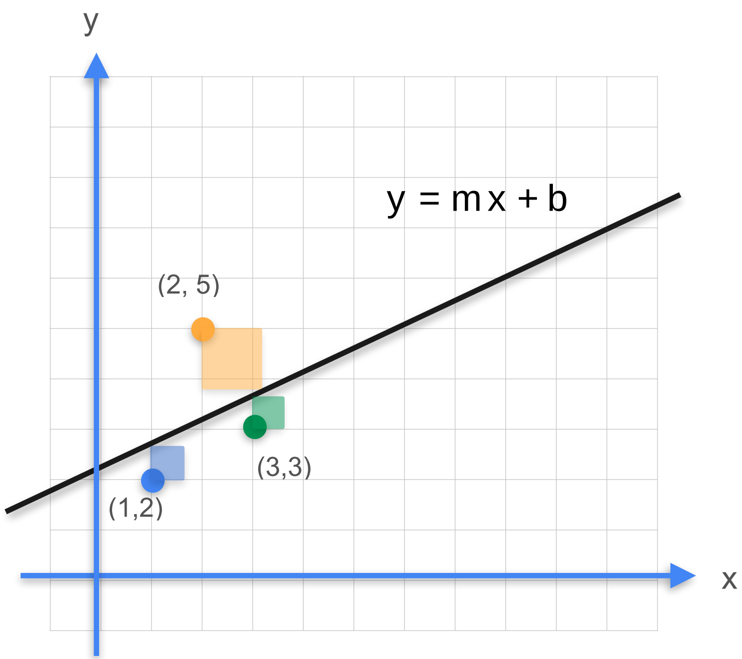



$$m = \frac{1}{2}$$

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

$$E(m = \frac{1}{2}, b = \frac{7}{3}) \quad 4.167$$

# Linear Regression: Optimal Solution




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

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

$$E(m = \frac{1}{2}, b = \frac{7}{3}) \quad 4.167$$

# Linear Regression: Gradient Descent

**Goal: Minimize sum of squares cost**

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

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

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

$m =$

$b =$ ?

**Gradient Descent to the rescue**

DeepLearning.AI

# Gradients and Gradient Descent

## Optimization using Gradient Descent in one variable - Part 1

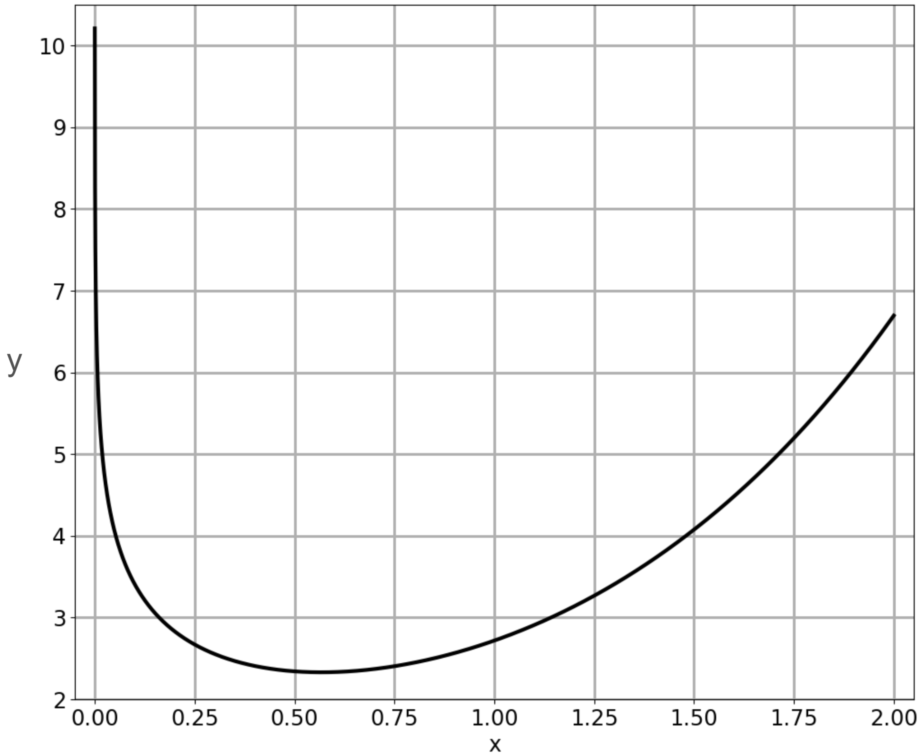
# Hard To Optimize Functions

# Hard To Optimize Functions

$f(x) = e^x \quad \log(x)$

# Hard To Optimize Functions

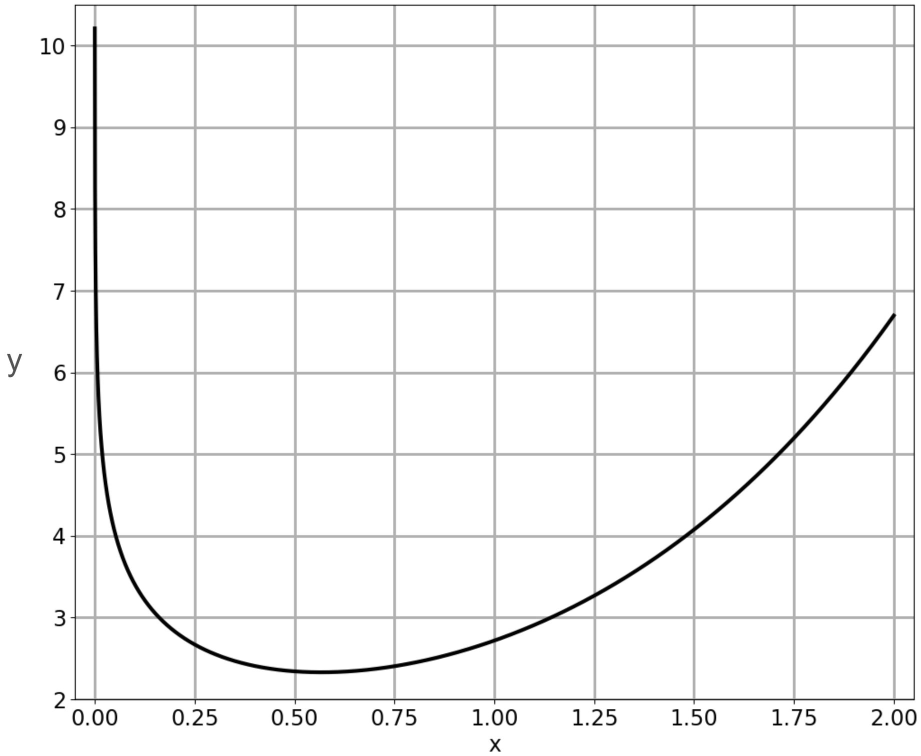
$$f(x) = e^{x} \quad \log(x)$$



# Hard To Optimize Functions

$f(x) = e^x \quad \log(x)$

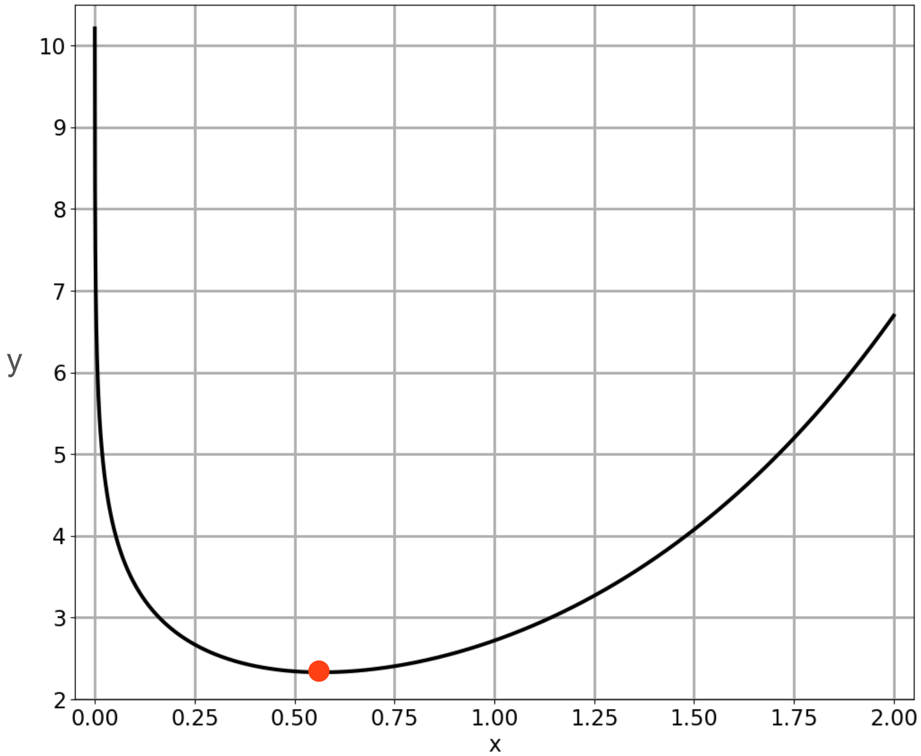
Minimum?



# Hard To Optimize Functions

$f(x) = e^x \quad \log(x)$

Minimum?

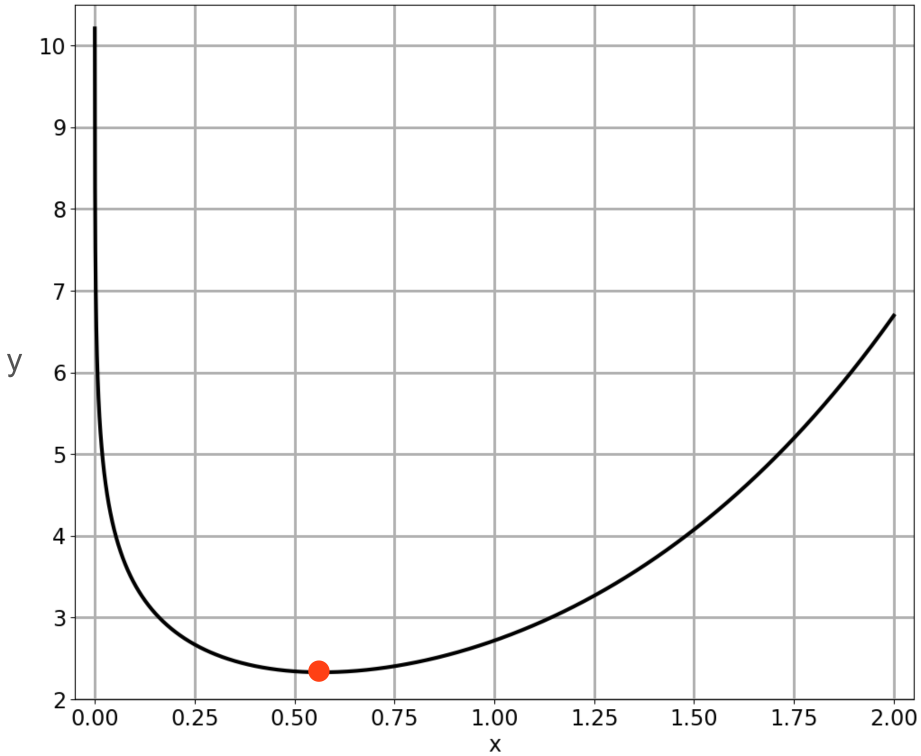


# Hard To Optimize Functions

$f(x) = e^x \quad \log(x)$

Minimum?

$f(x) = 0$

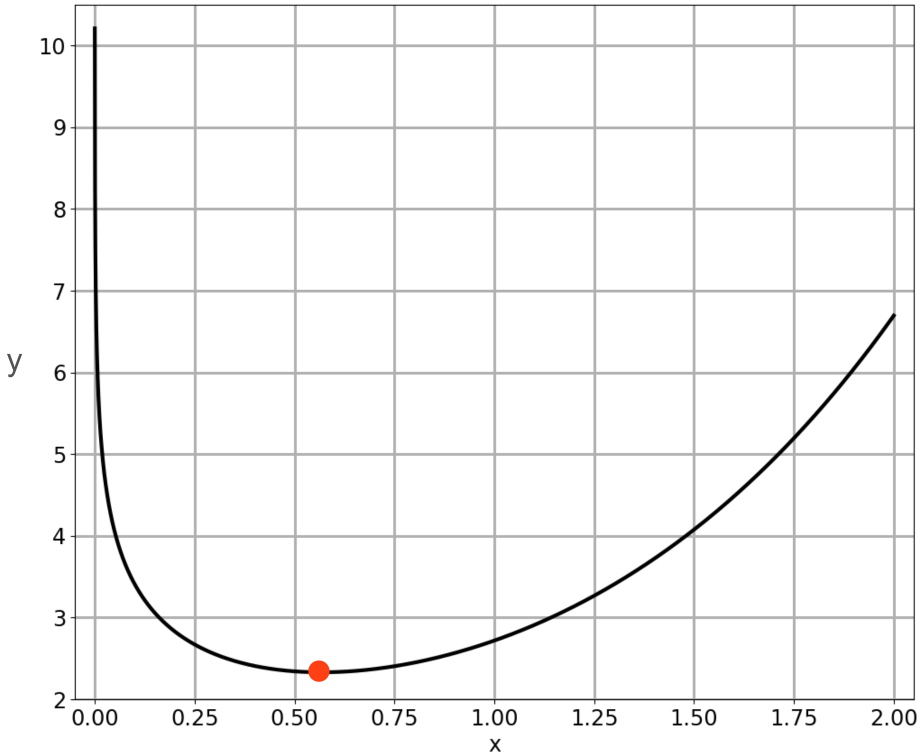


# Hard To Optimize Functions

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

Minimum?

$$f\,(x) = e^{x} \quad \frac{1}{x} = 0$$

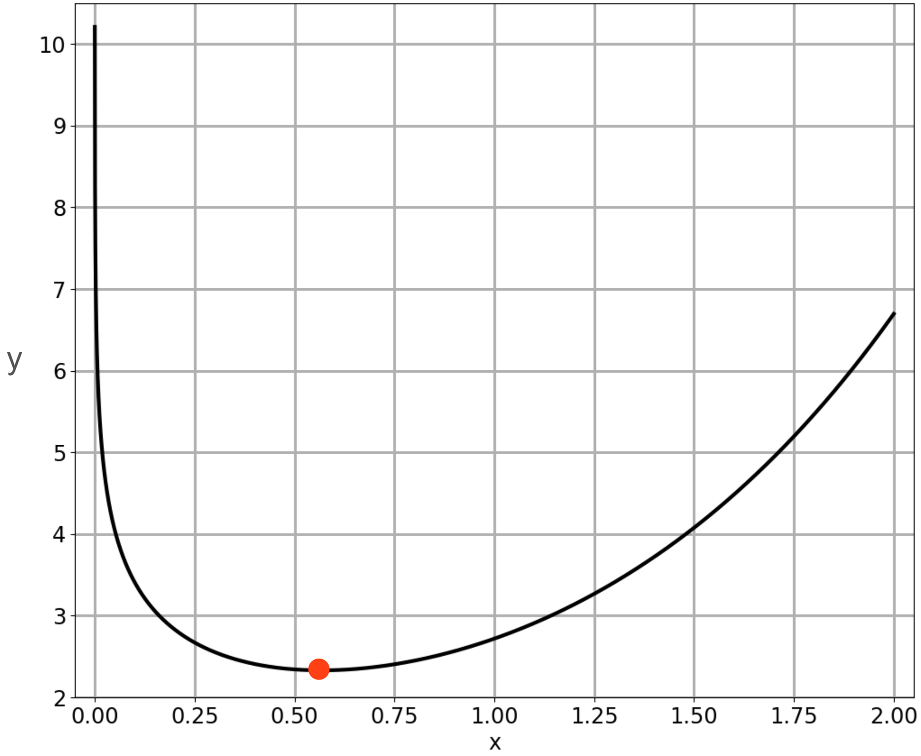


# Hard To Optimize Functions

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

Minimum?

$$f\ (x) = \boxed{e^x} \quad \frac{1}{x} = 0$$

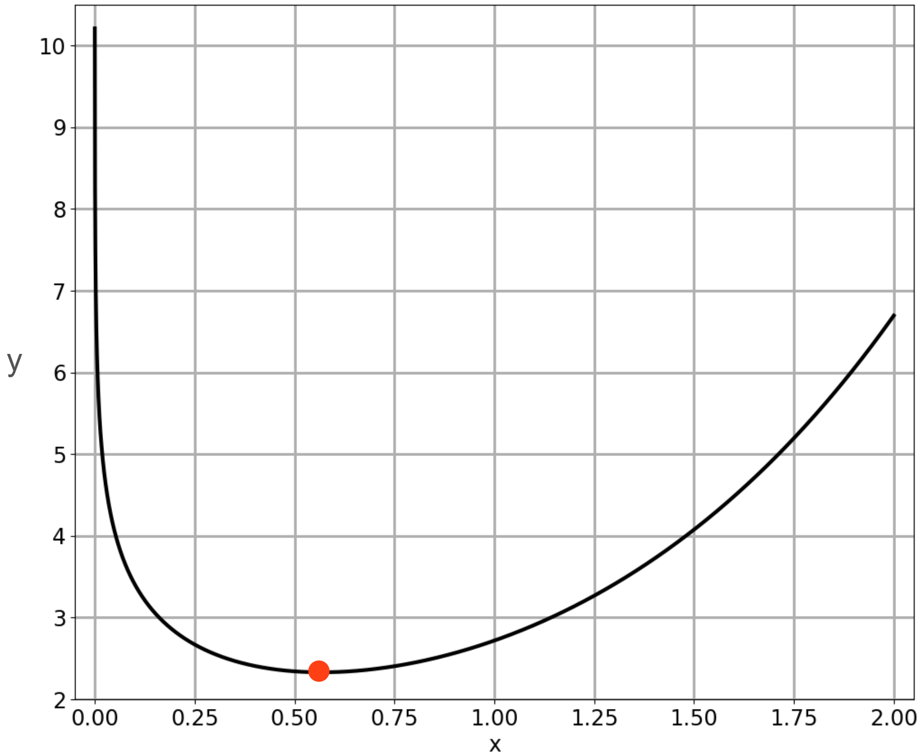


# Hard To Optimize Functions

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

Minimum?

$$f(x) = \boxed{e^x} \quad \boxed{\frac{1}{x}} = 0$$

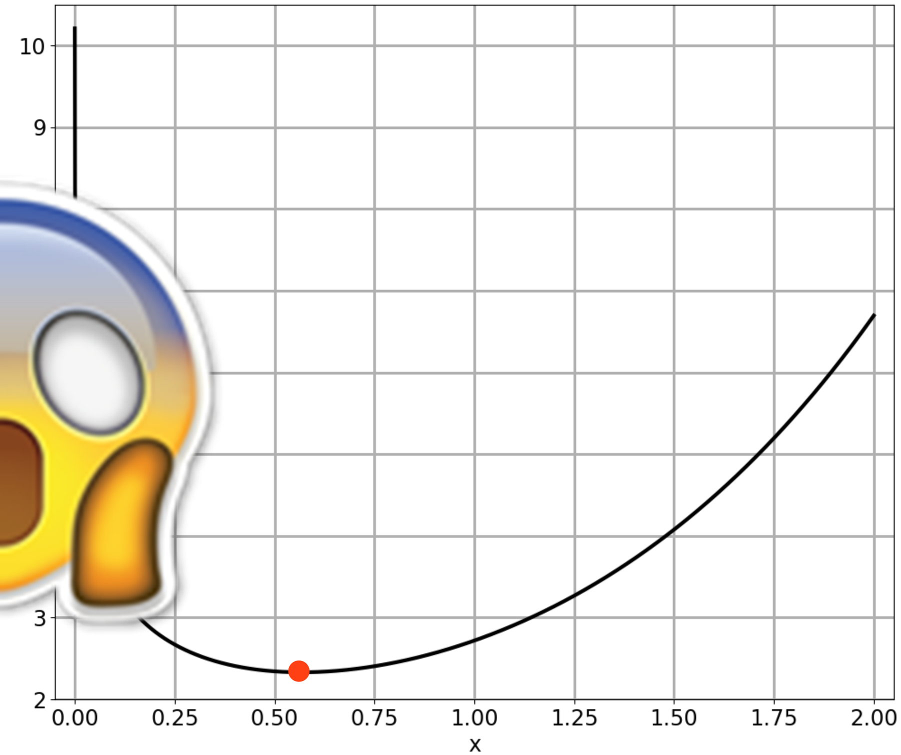


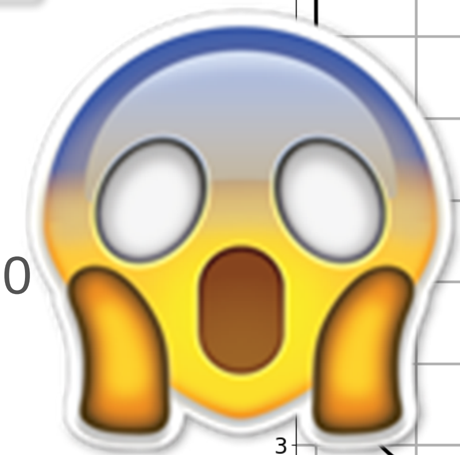
# Hard To Optimize Functions

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

Minimum?

$$f(x) = \boxed{e^x}\ \boxed{\frac{1}{x}} = 0$$

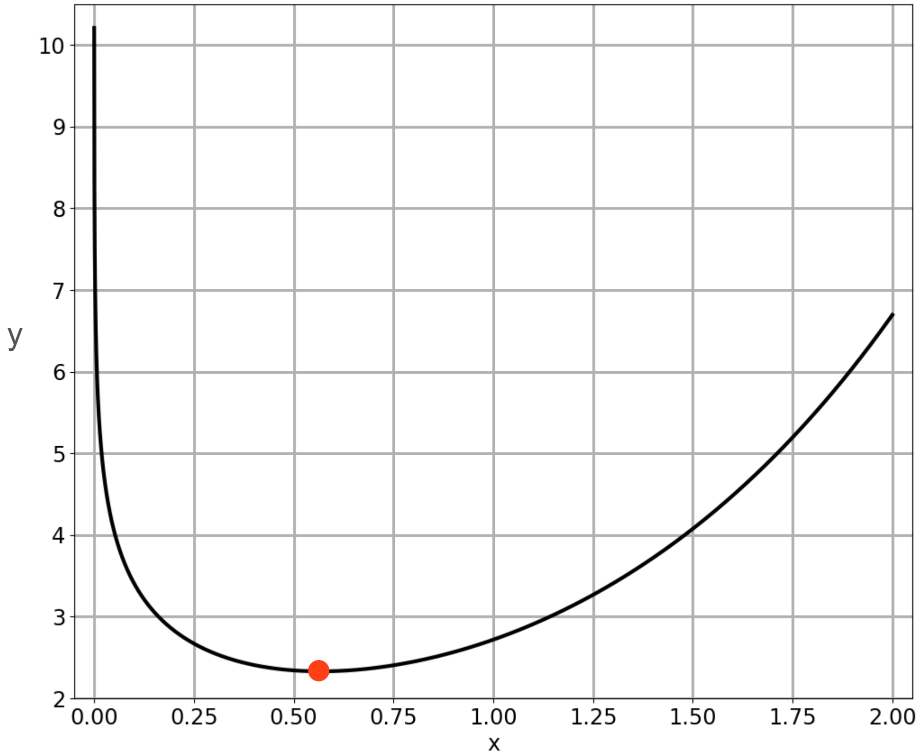


# Hard To Optimize Functions

$$f(x) = e^x \quad \log(x)$$

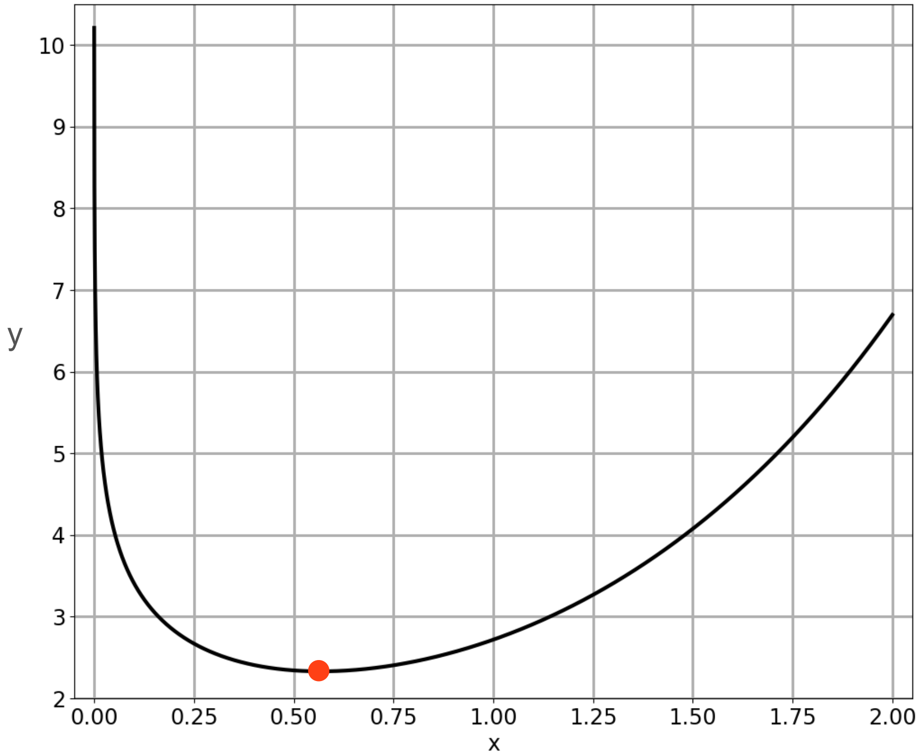
$$f\ (x) = e^x \quad \frac{1}{x}$$



# Hard To Optimize Functions

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

$$f\,(x) = e^{x} \quad \frac{1}{x} = 0$$

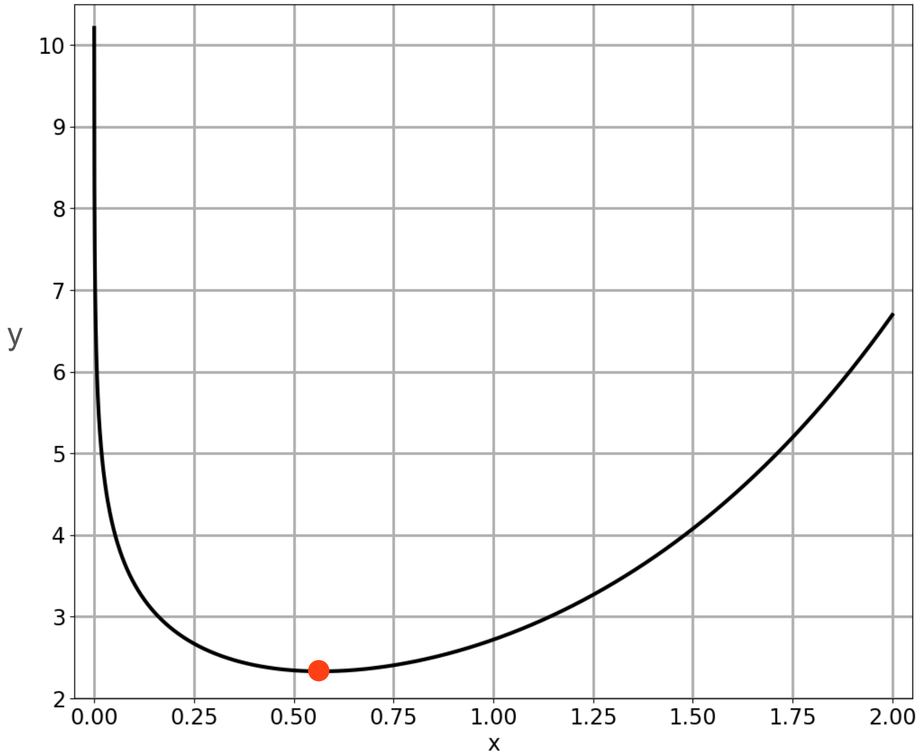


# Hard To Optimize Functions

$$f(x) = e^x \quad \log(x)$$

$$f(x) = e^x \quad \frac{1}{x} = 0$$

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



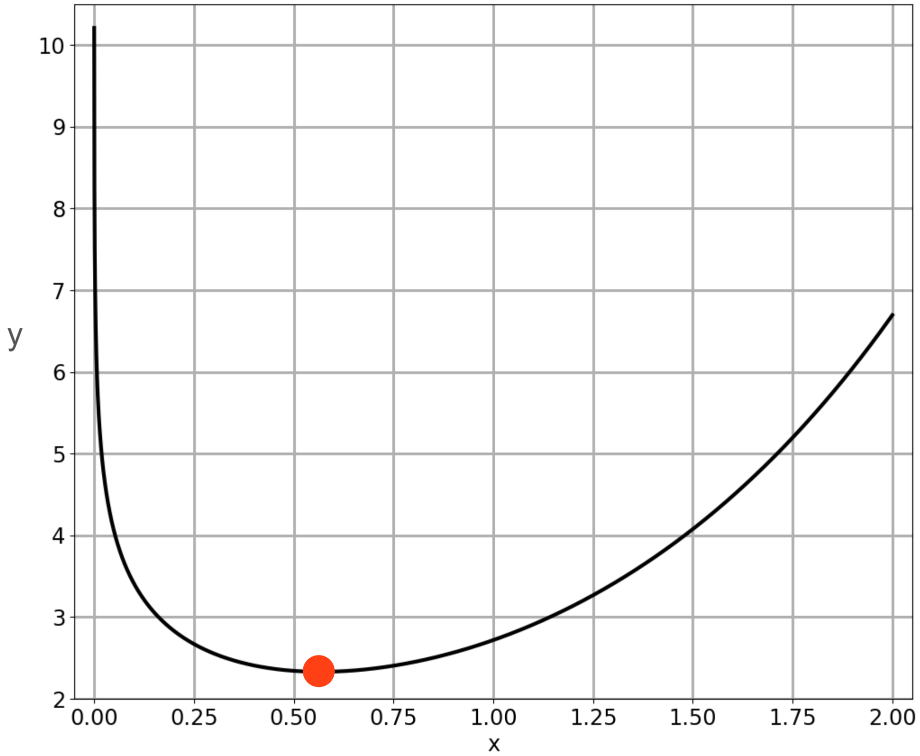
# Hard To Optimize Functions

$$f(x) = e^x \quad \log(x)$$

$$f(x) = e^x \quad \frac{1}{x} = 0$$

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

Solution: $x = 0.5671..$



# Hard To Optimize Functions

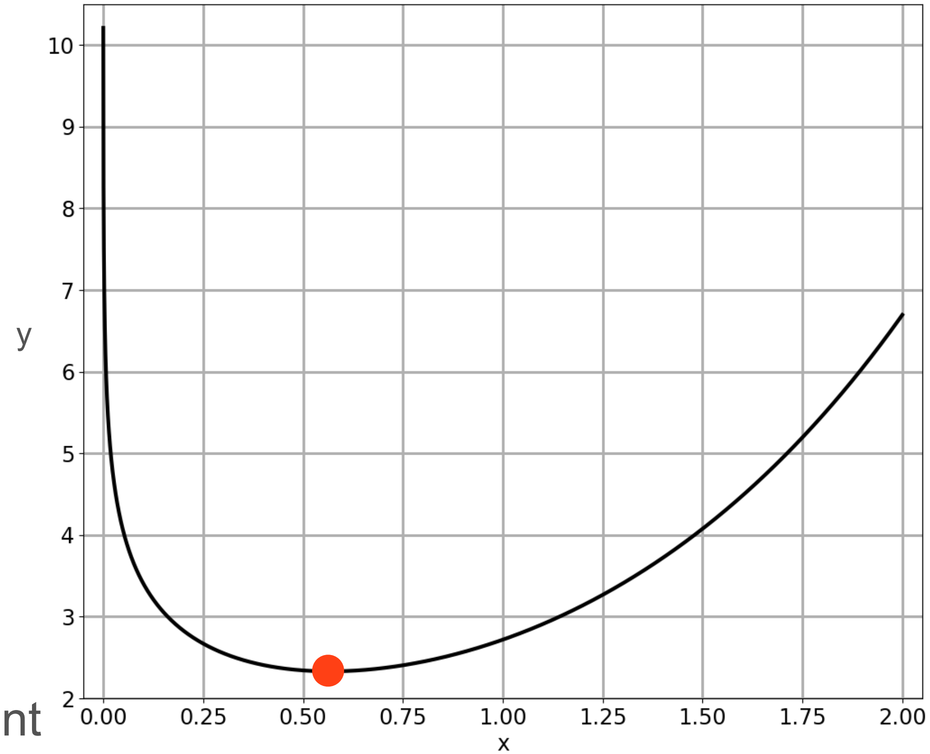
$$f(x) = e^x \quad \log(x)$$

$$f\ (x) = e^x \quad \frac{1}{x} = 0$$

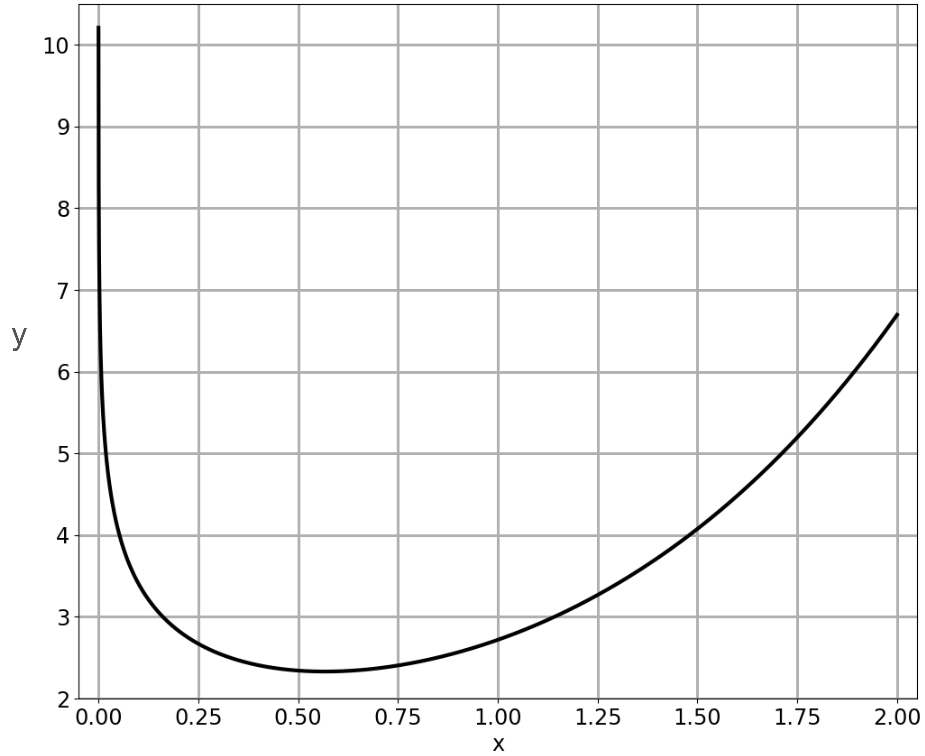
$$\Rightarrow \quad e^x = \frac{1}{x}$$

Solution: $x = 0.5671..$

Also known as the Omega constant

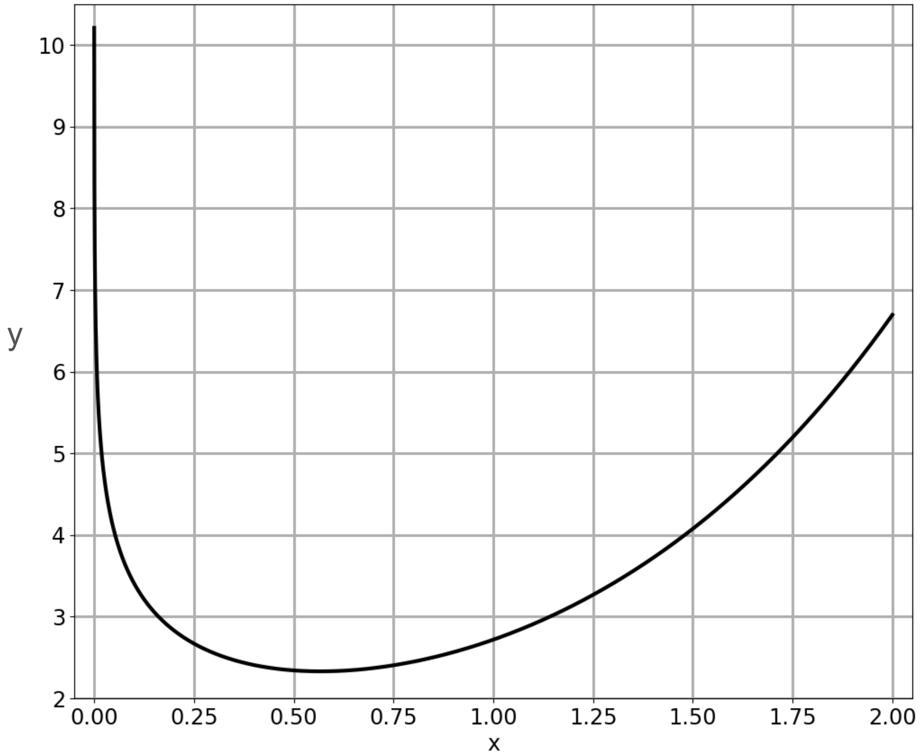


# Method 1: Try Both Directions



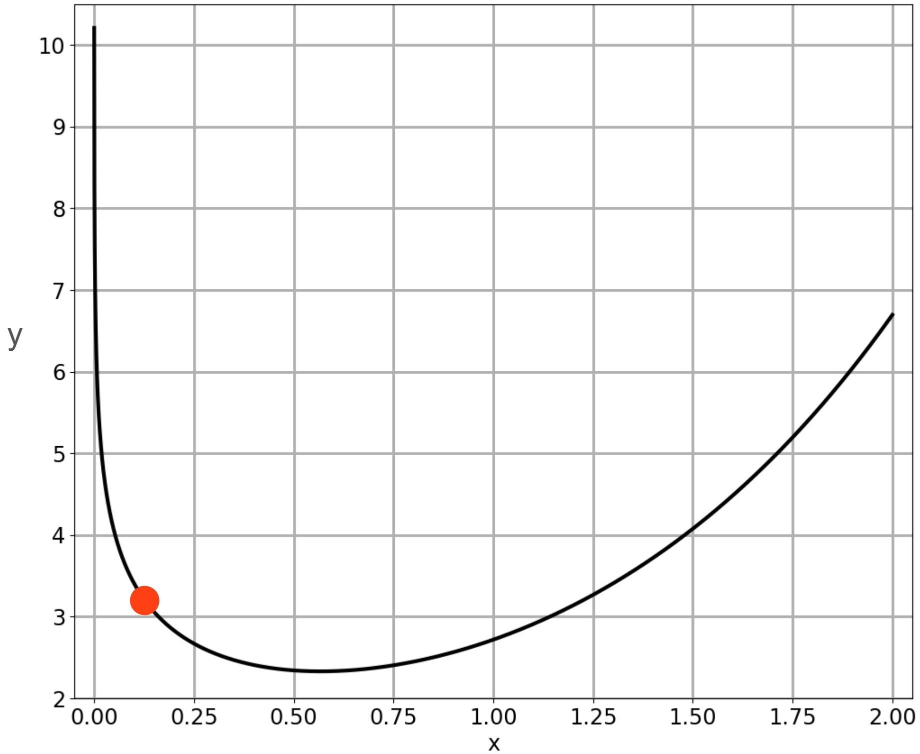
# Method 1: Try Both Directions

Is there any other way?



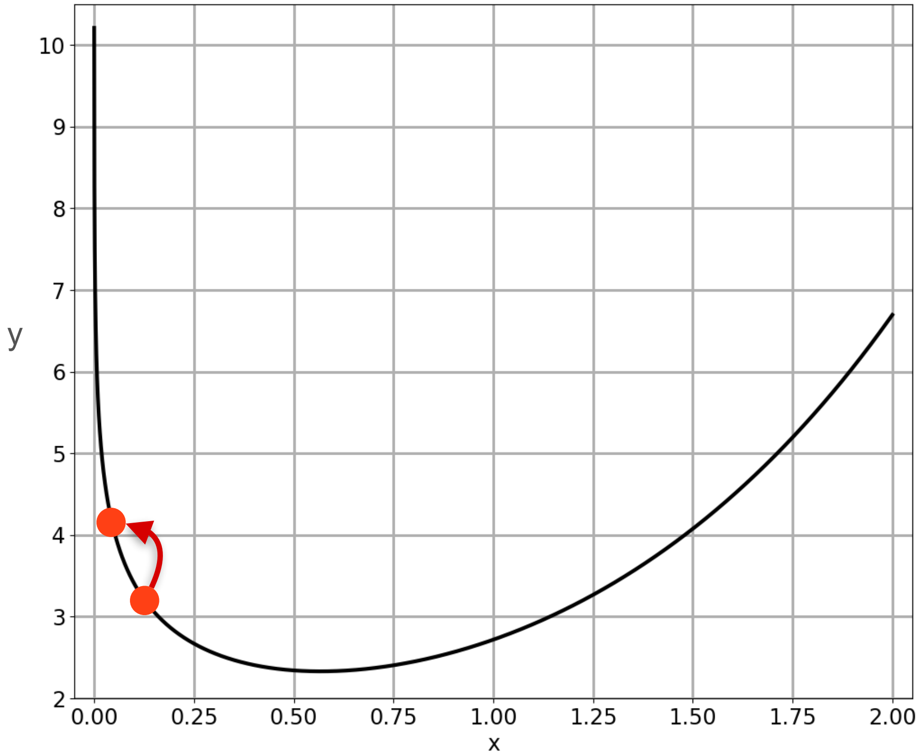
# Method 1: Try Both Directions

Is there any other way?



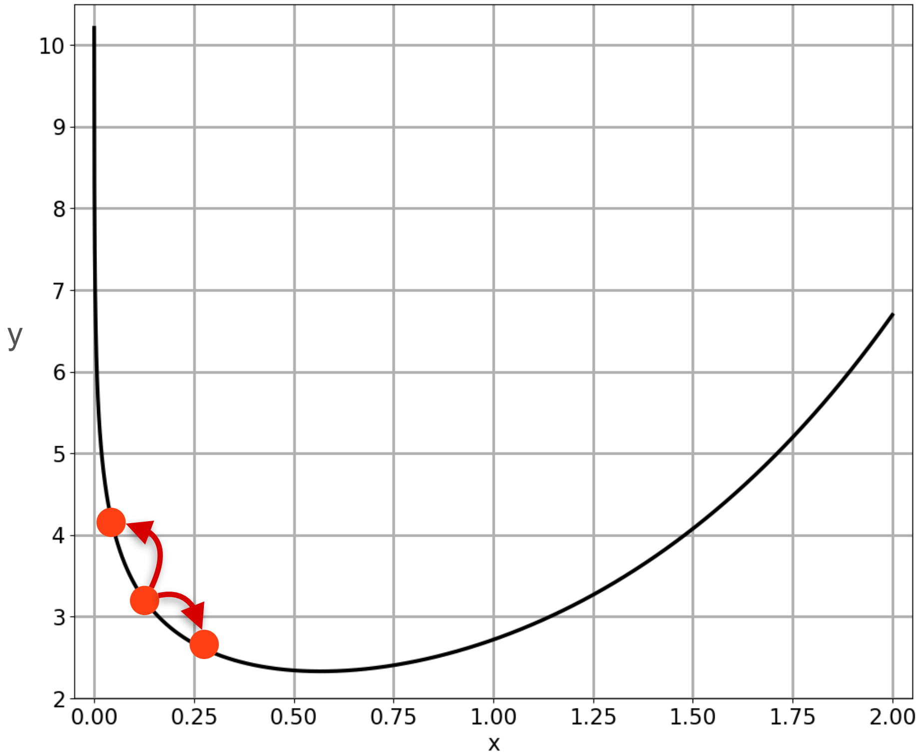
# Method 1: Try Both Directions

Is there any
other way?



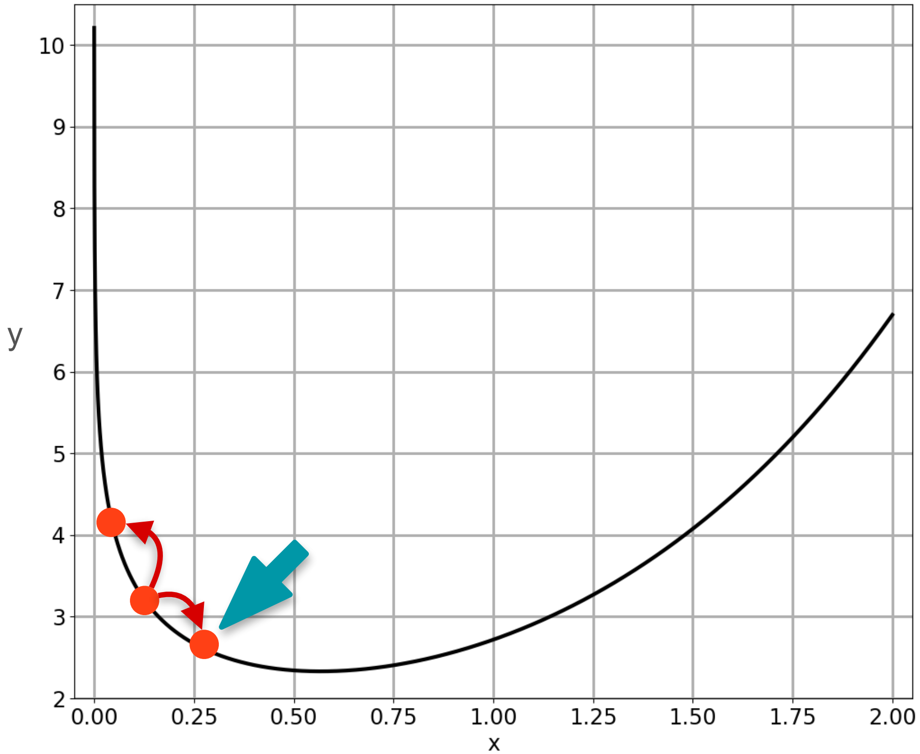
# Method 1: Try Both Directions

Is there any
other way?



# Method 1: Try Both Directions

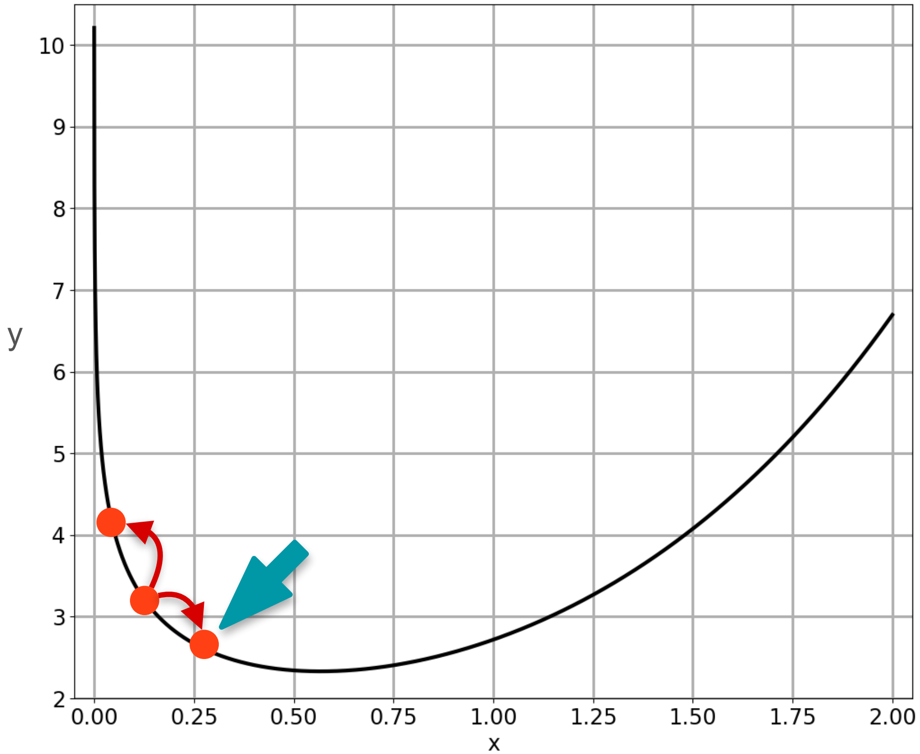
Is there any
other way?



# Method 1: Try Both Directions

Is there any other way?

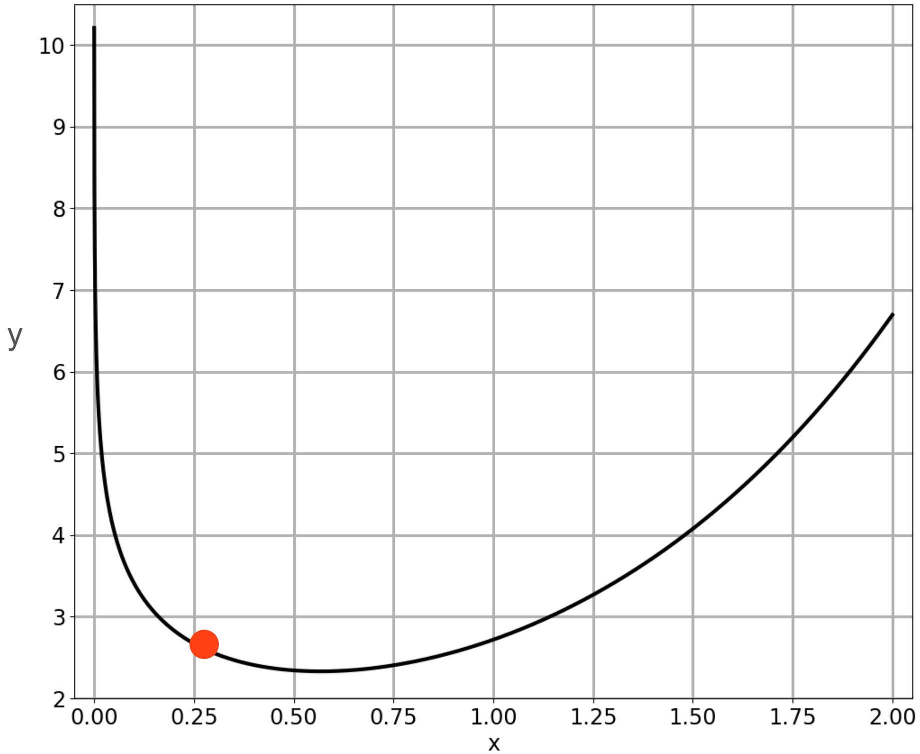
Repeat!



# Method 1: Try Both Directions

Is there any
other way?

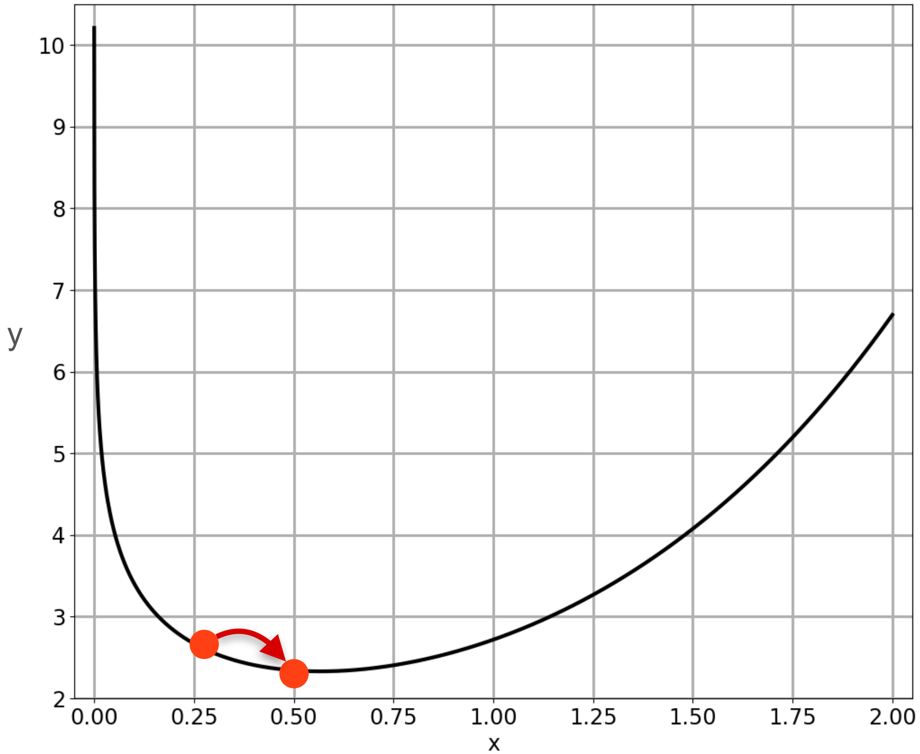
Repeat!



# Method 1: Try Both Directions

Is there any
other way?

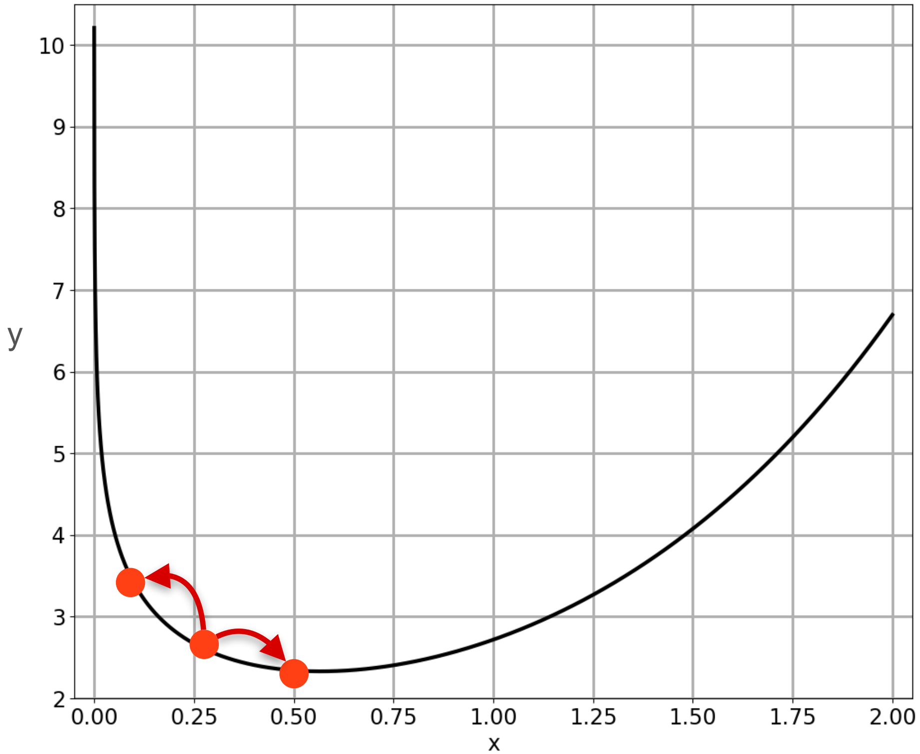
Repeat!



# Method 1: Try Both Directions

Is there any other way?

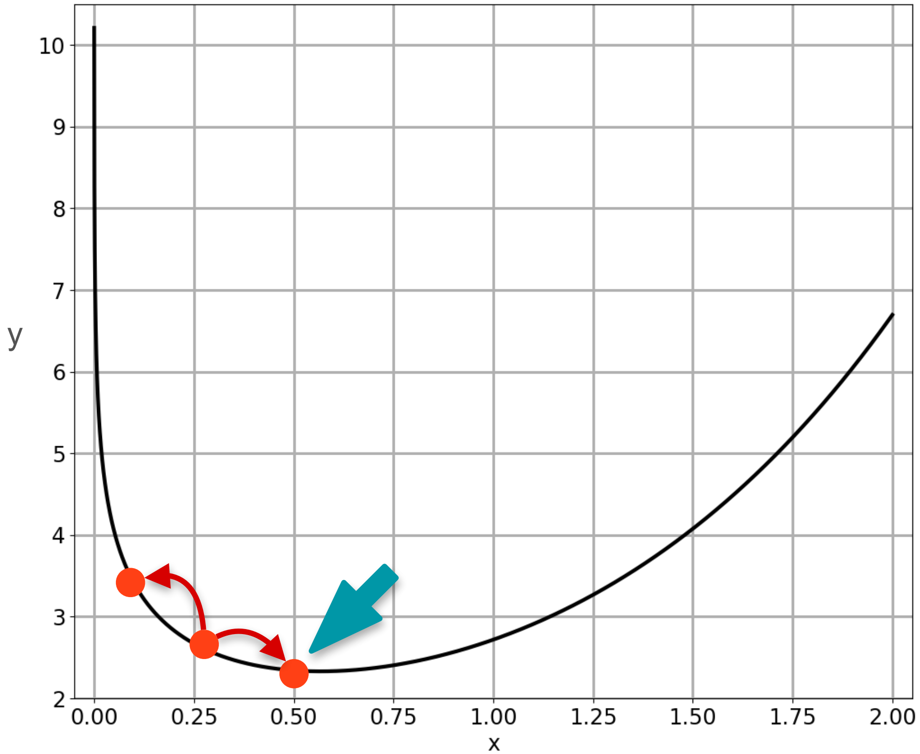
Repeat!



# Method 1: Try Both Directions

Is there any other way?

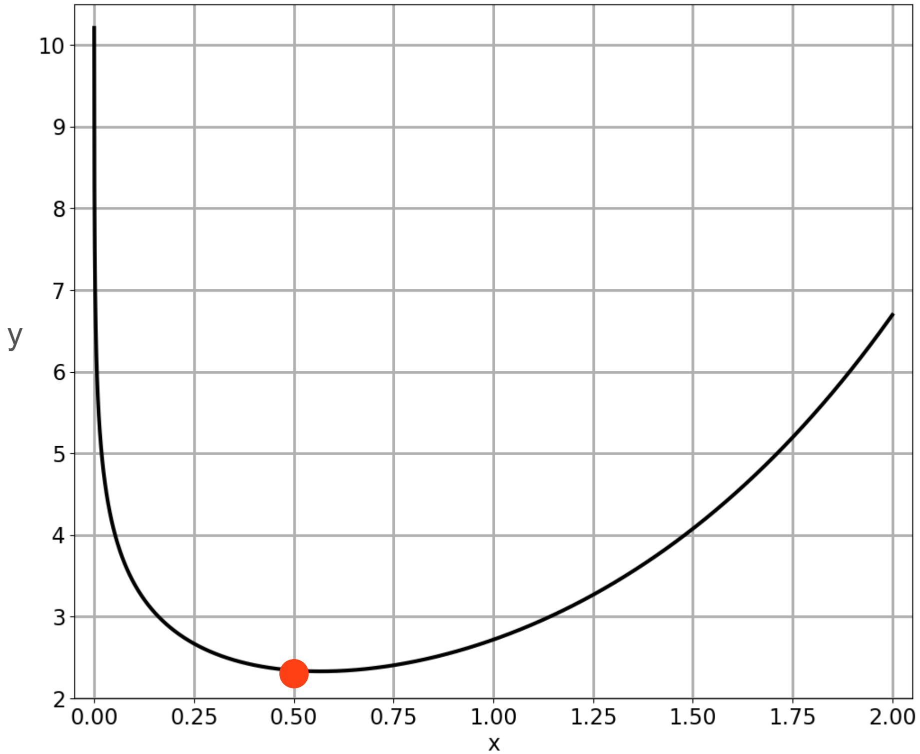
Repeat!



# Method 1: Try Both Directions

Is there any other way?

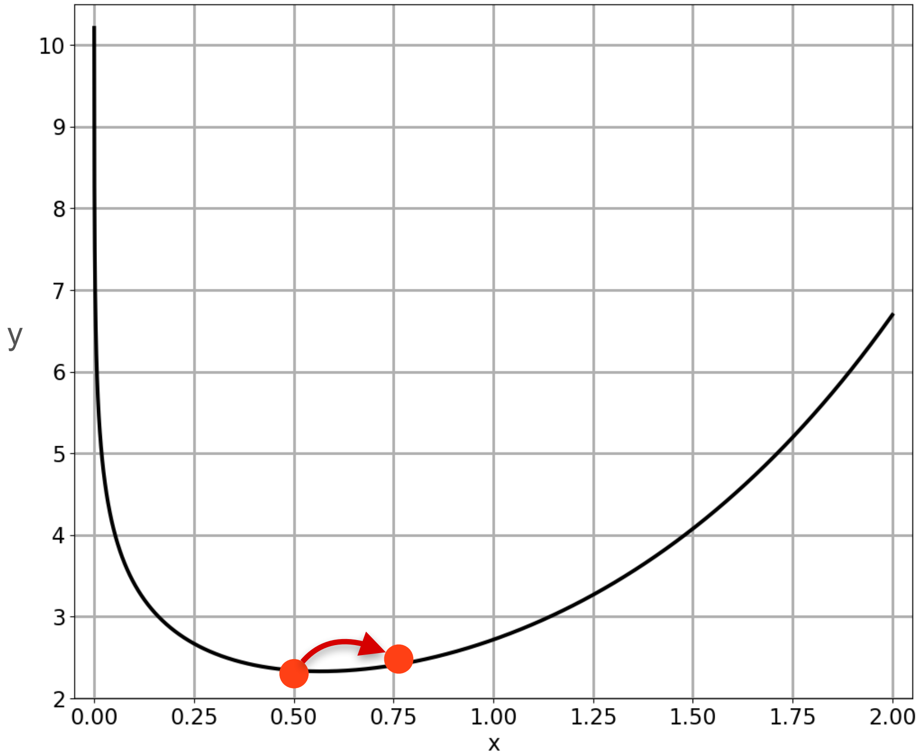
Repeat!



# Method 1: Try Both Directions

Is there any
other way?

Repeat!

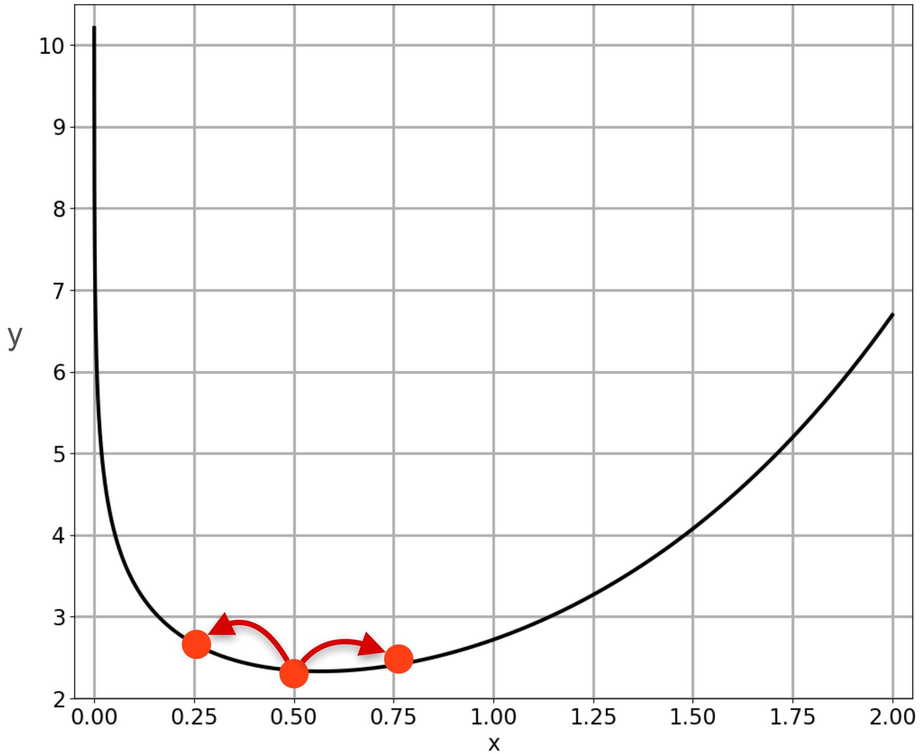



DeepLearning.AI

# Method 1: Try Both Directions

Is there any other way?

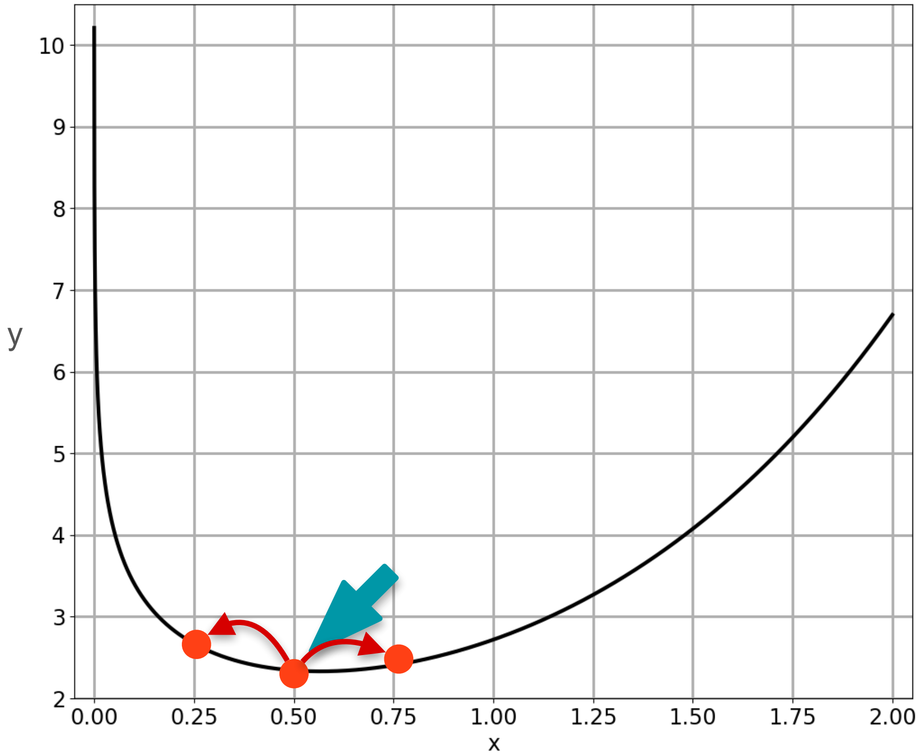
Repeat!



# Method 1: Try Both Directions

Is there any other way?

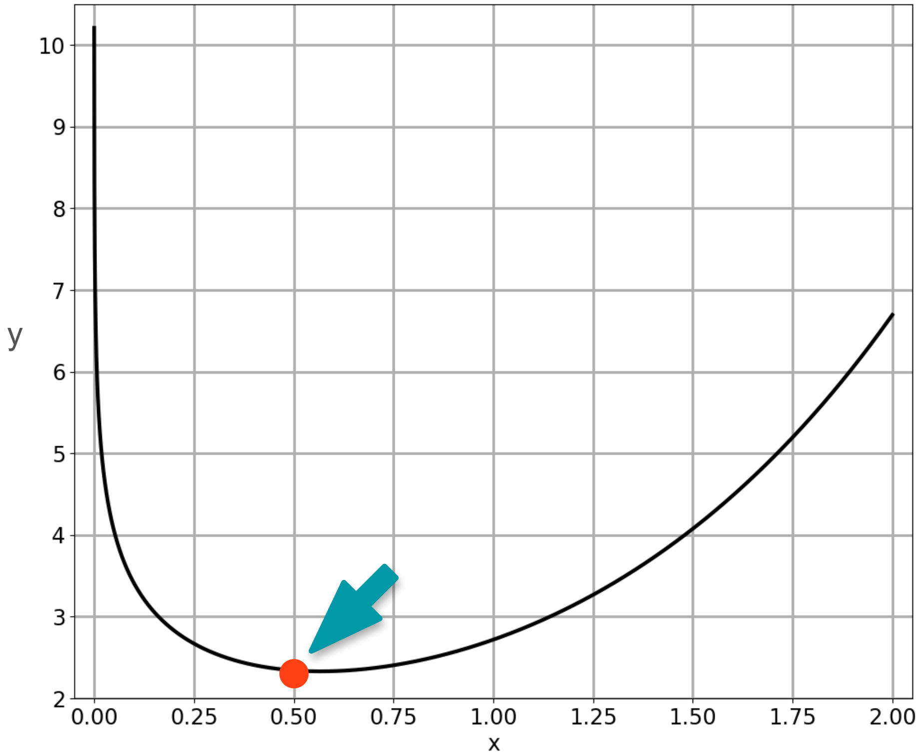
Repeat!



# Method 1: Try Both Directions

Is there any other way?

Repeat!

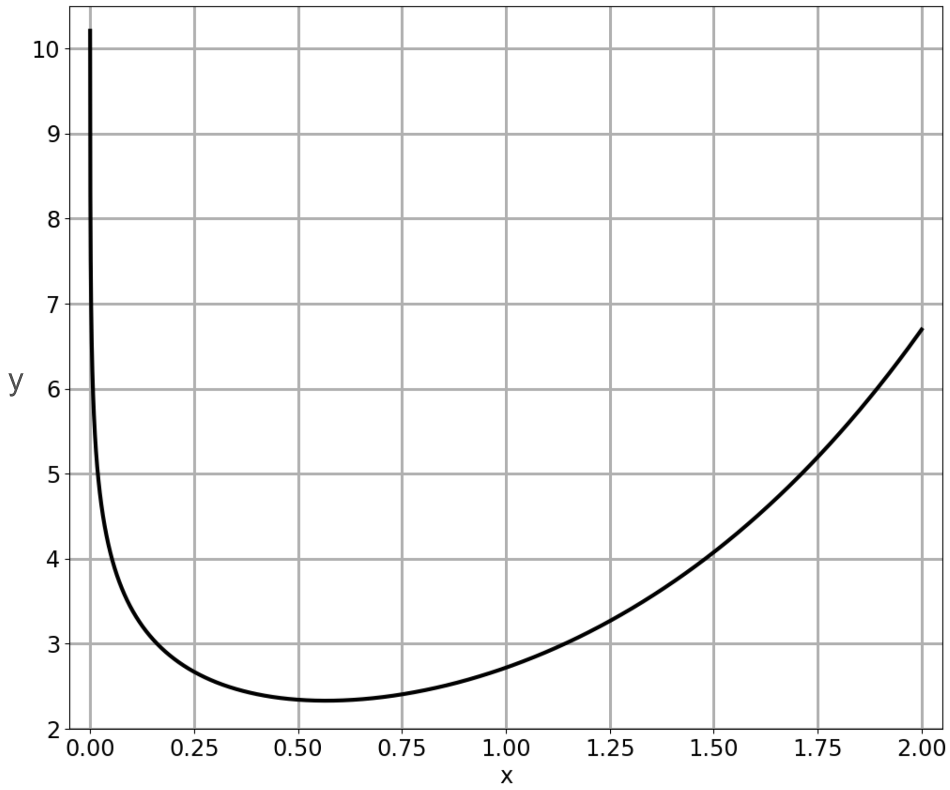


DeepLearning.AI

Gradients and Gradient Descent

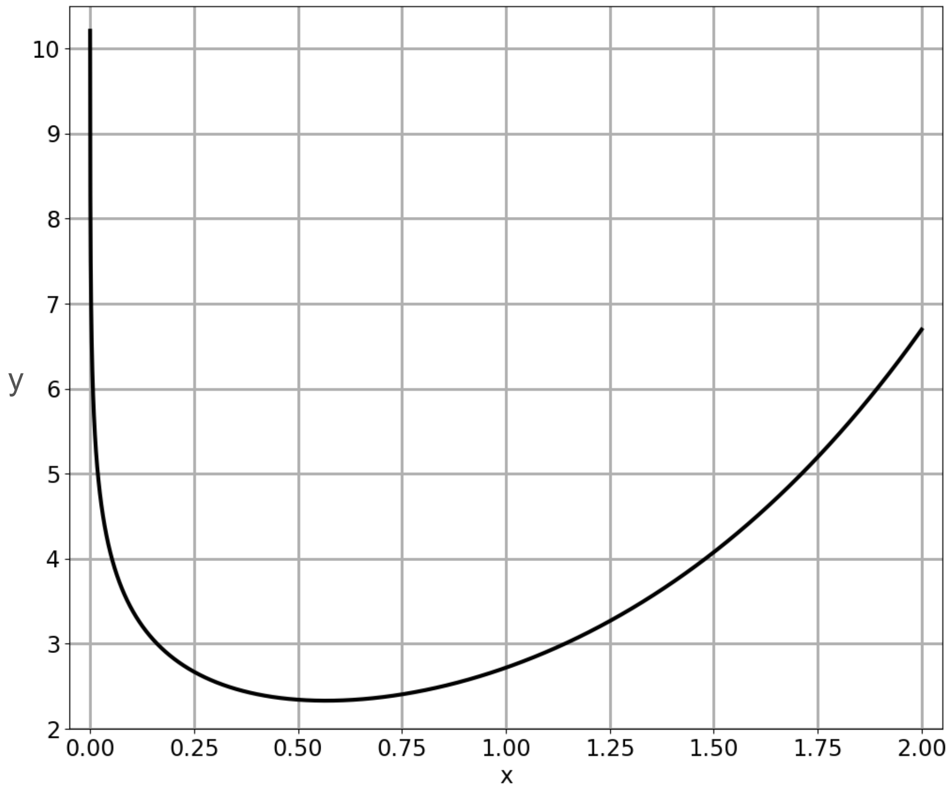
# Optimization using Gradient Descent in one variable - Part 2

# Method 2: Be Clever



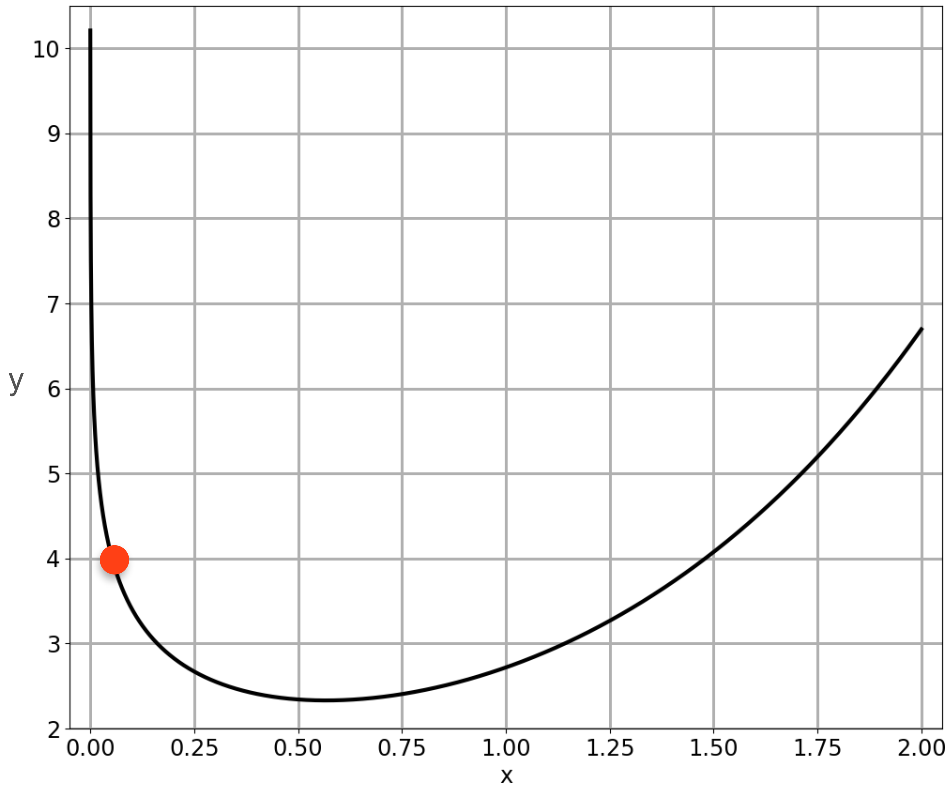
# Method 2: Be Clever

Try something smarter…



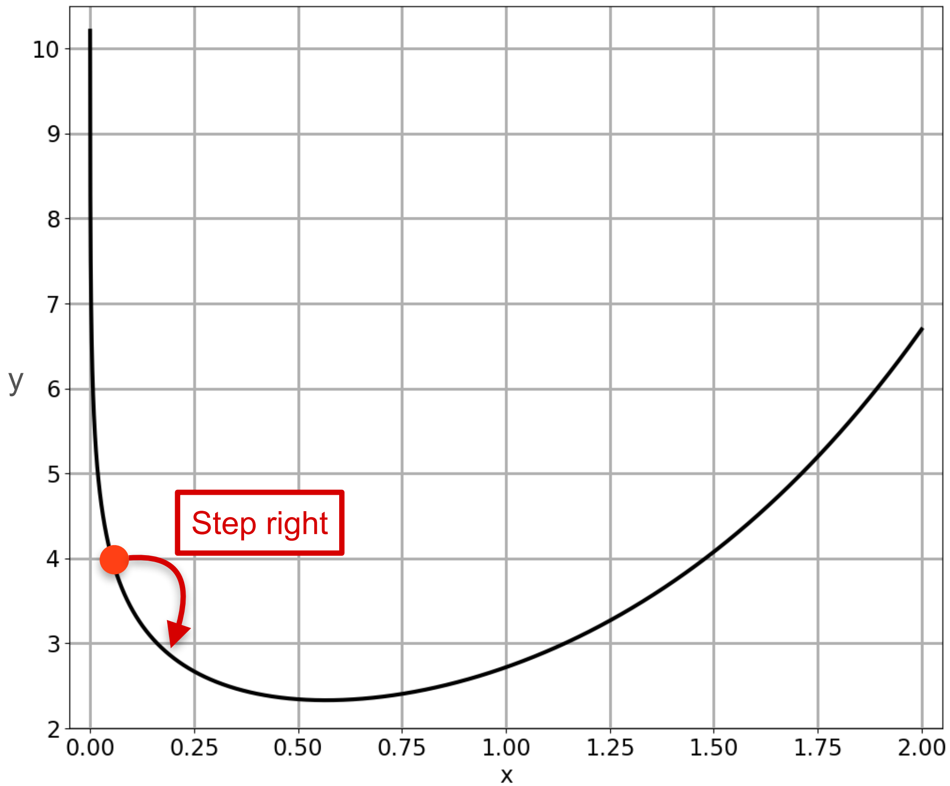
# Method 2: Be Clever

Try something smarter…



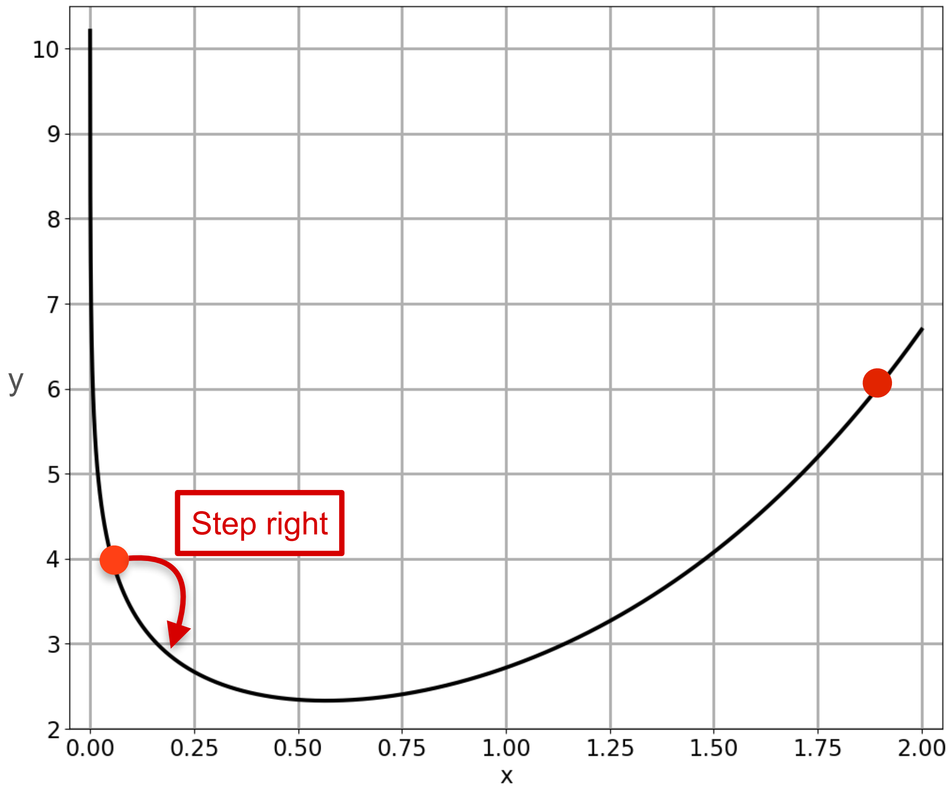
# Method 2: Be Clever

Try something smarter…



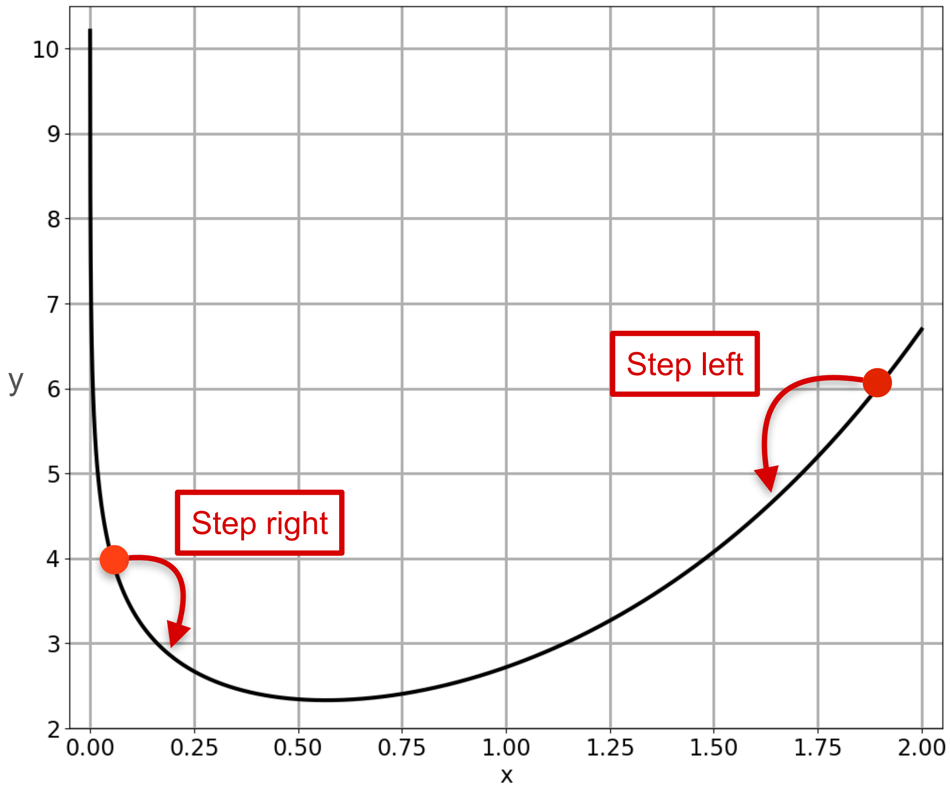
# Method 2: Be Clever

Try something smarter…



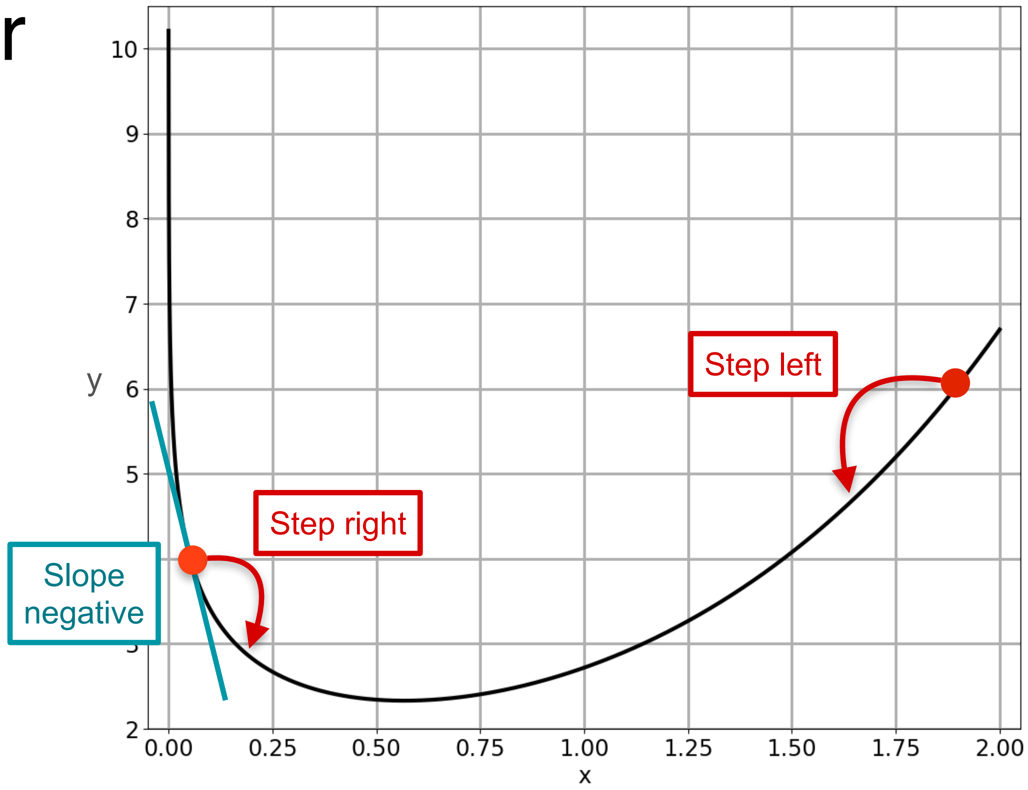
# Method 2: Be Clever

Try something smarter…



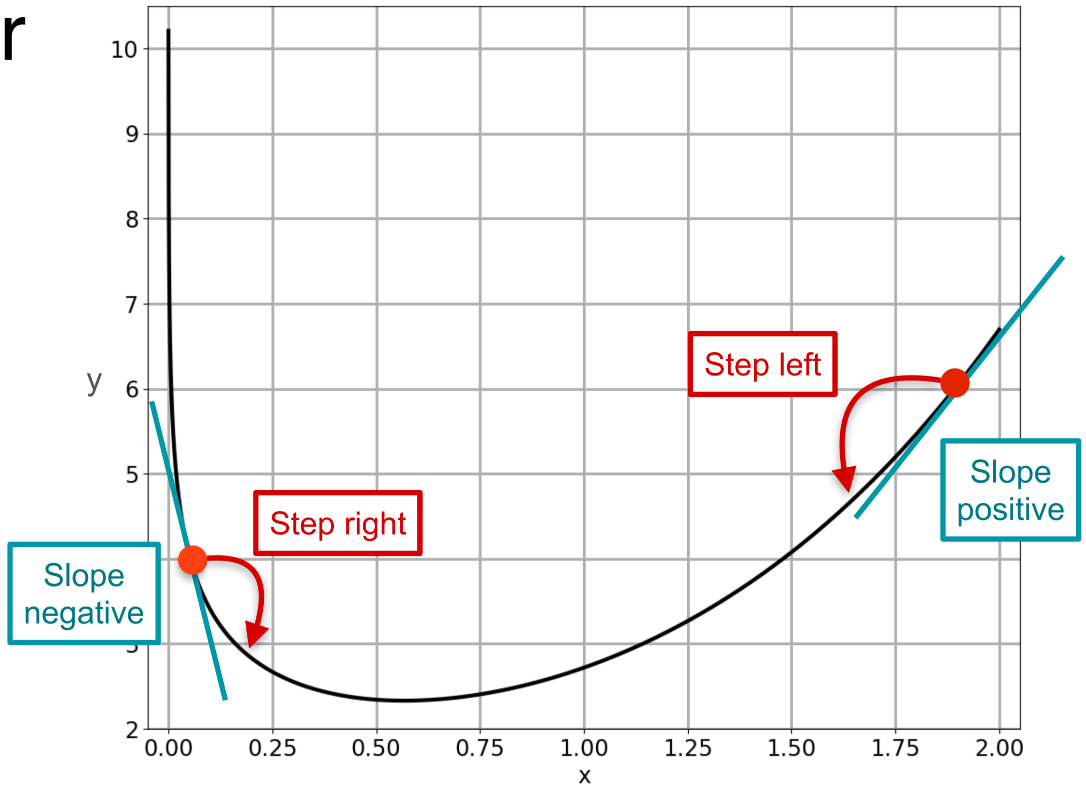
# Method 2: Be Clever

Try something smarter…



# Method 2: Be Clever

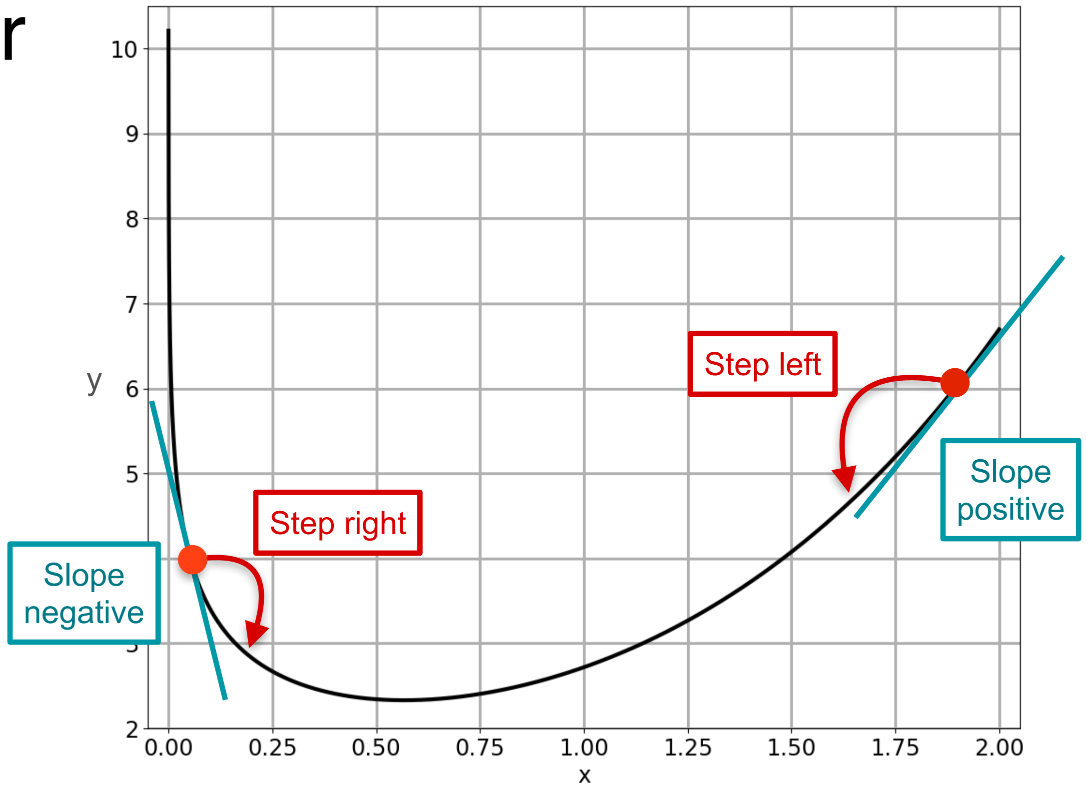
Try something smarter…



# Method 2: Be Clever

Try something smarter…

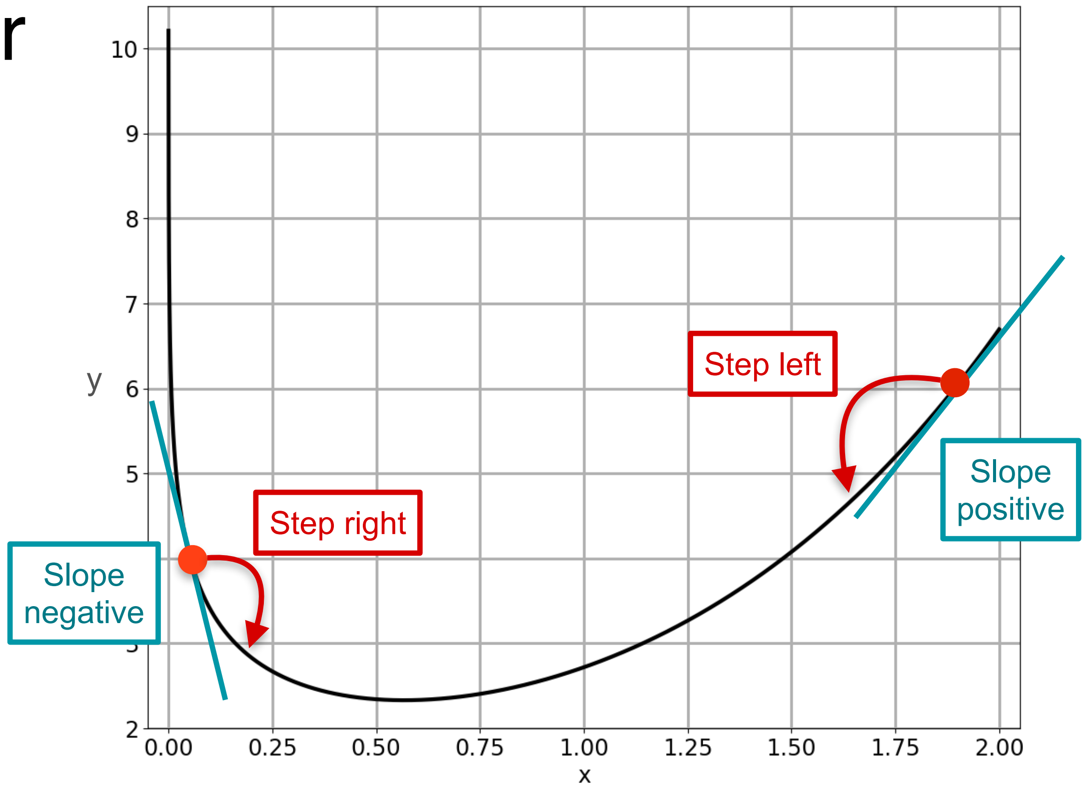
new point



# Method 2: Be Clever

Try something smarter…

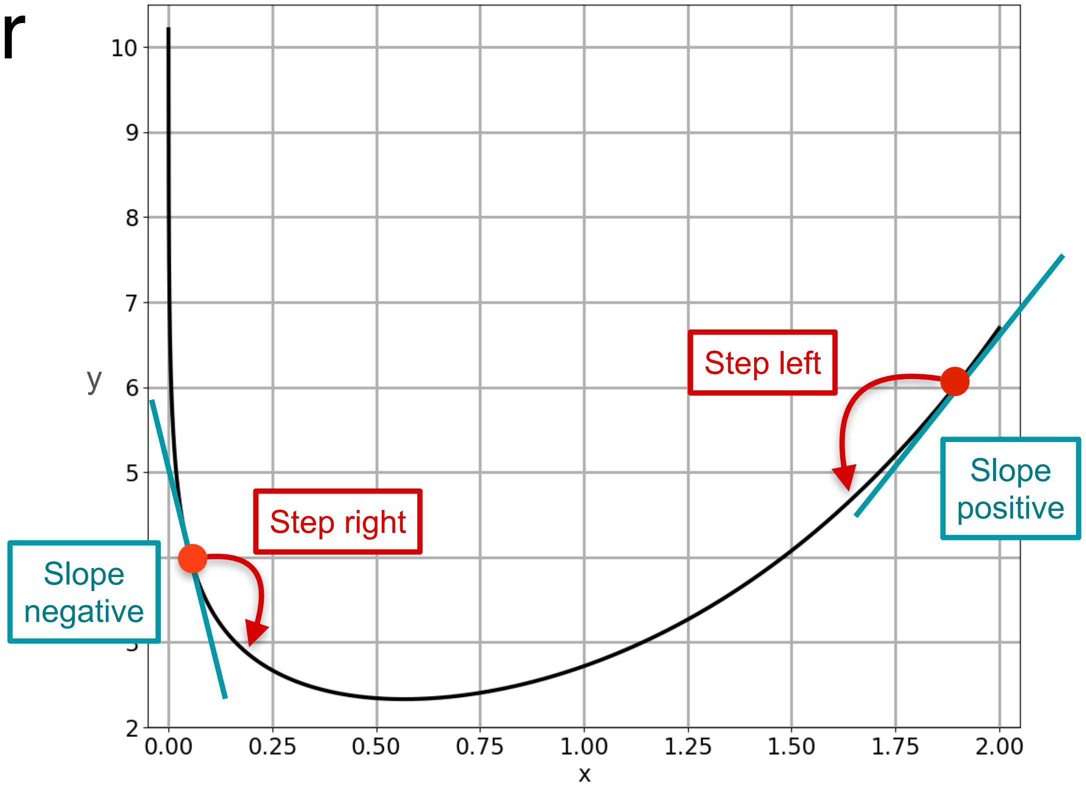
new point = old point



# Method 2: Be Clever

Try something smarter…

new point = old point - slope

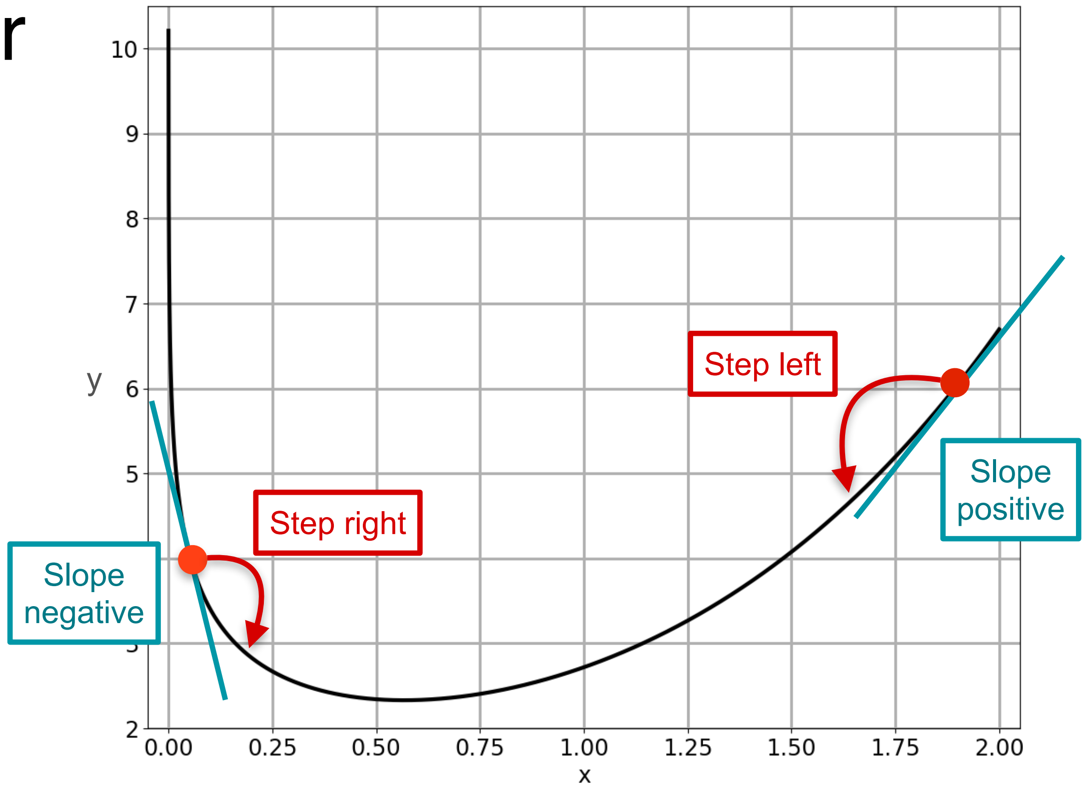


# Method 2: Be Clever

Try something smarter…

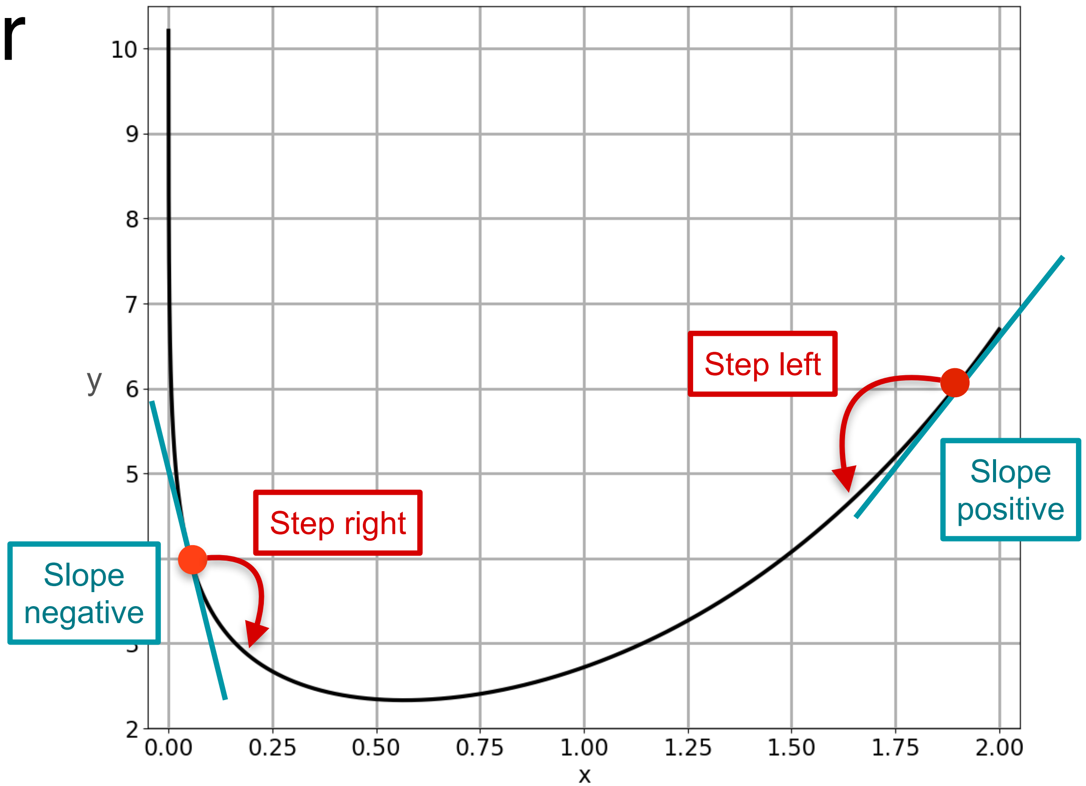
new point = old point - slope

$x_1$



# Method 2: Be Clever

Try something smarter…




new point = old point - slope

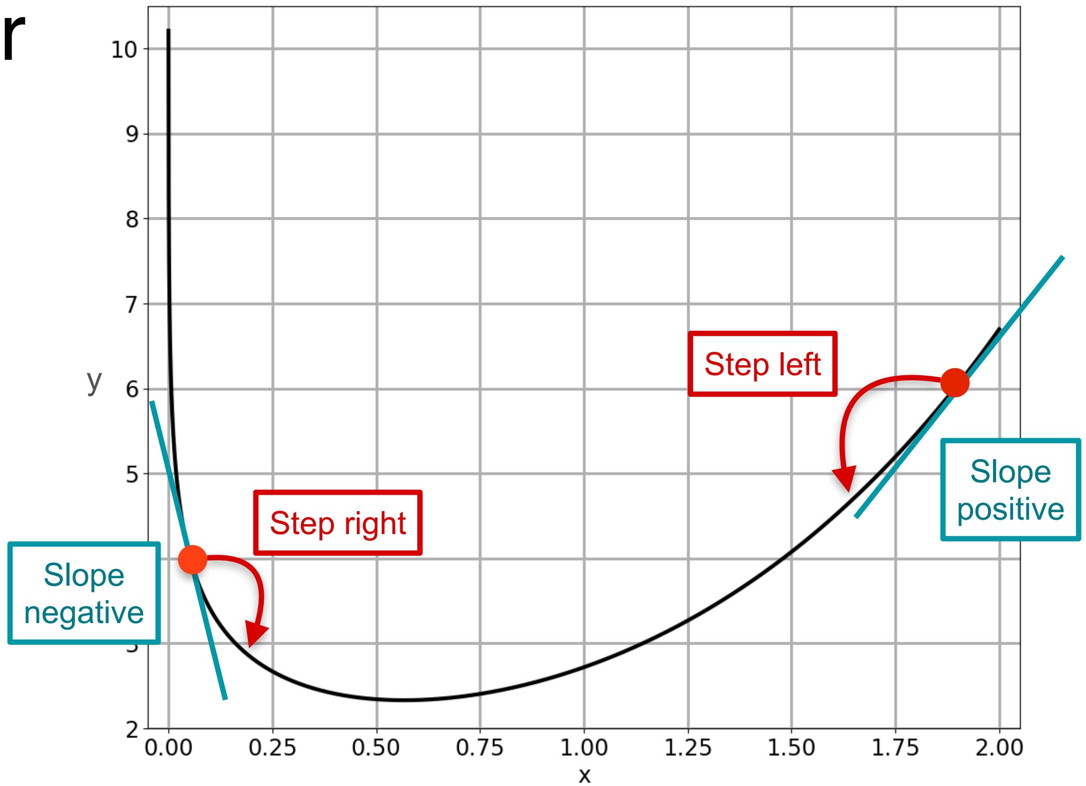
$$x_1 = x_0$$

# Method 2: Be Clever

Try something smarter…

new point = old point - slope

$$x_1 = x_0 \quad f\,(x_0)$$

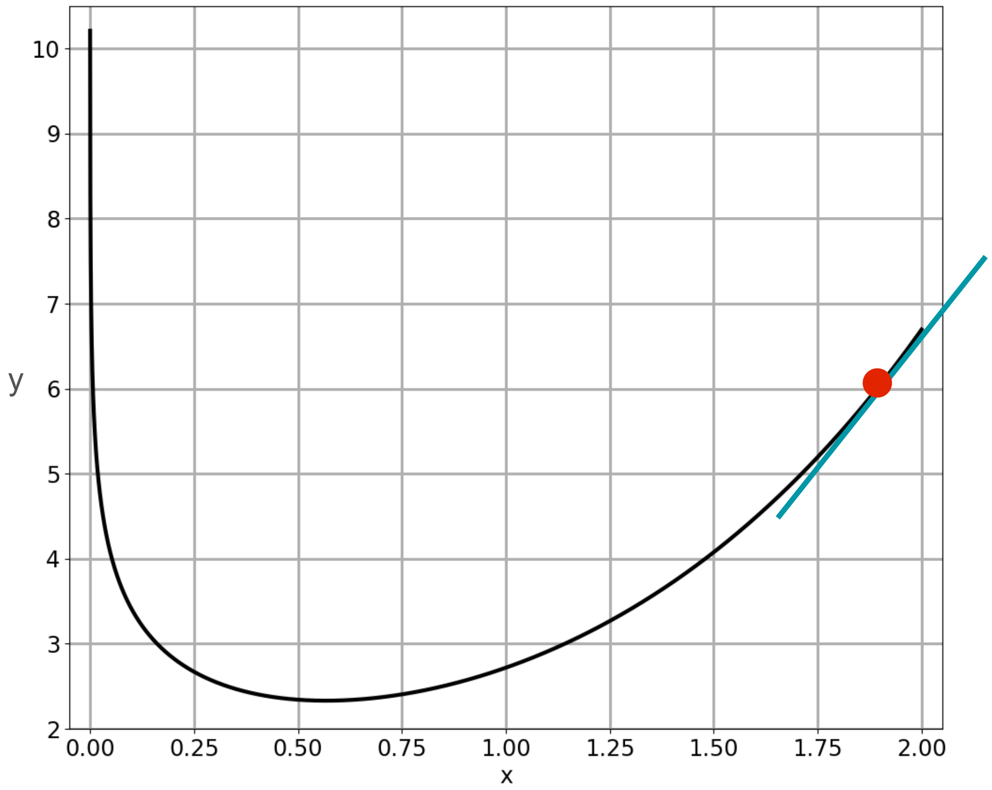


# Method 2: Be Clever

Try something smarter…

new point = old point - slope

$$x_1 = x_0 \quad f\,(x_0)$$

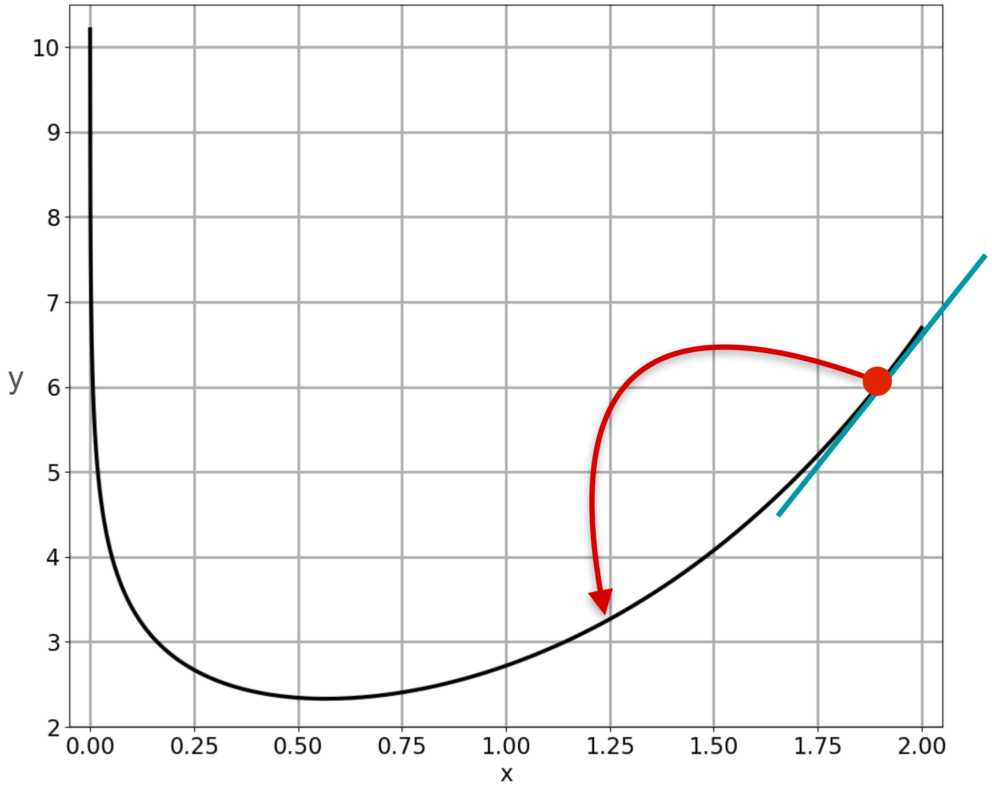


# Method 2: Be Clever

Try something smarter…

new point = old point - slope

$$x_1 = x_0 \quad f\,(x_0)$$

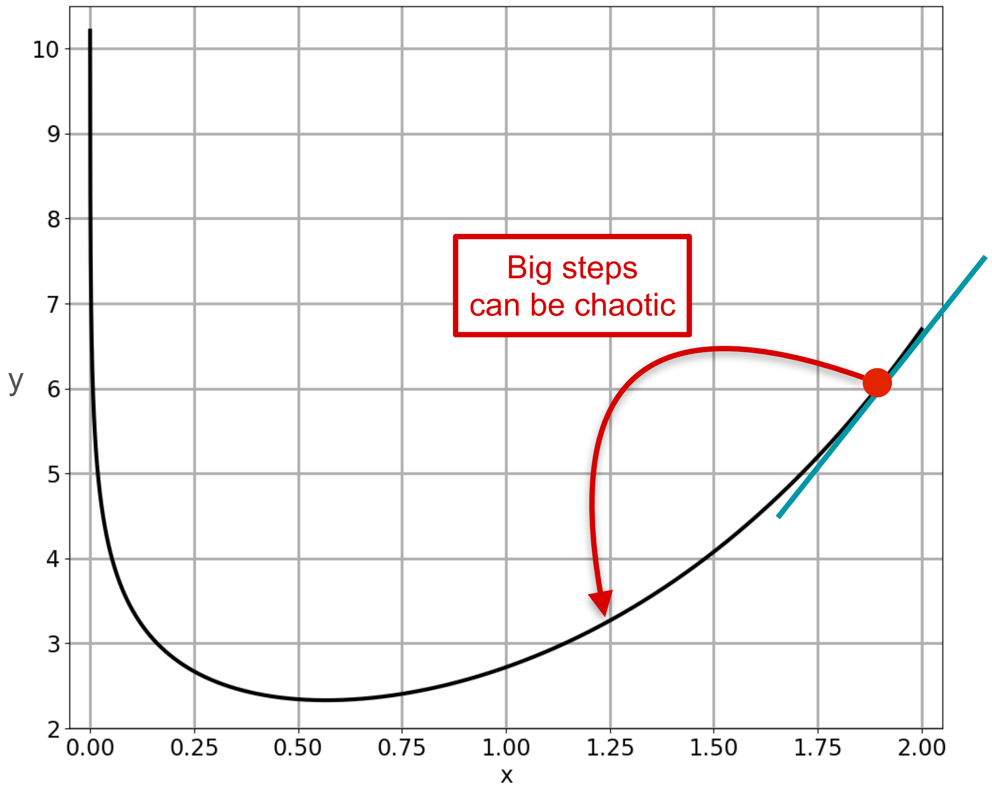


# Method 2: Be Clever

Try something smarter…

new point = old point - slope

$$x_1 \;=\; x_0 \qquad f\,(x_0)$$

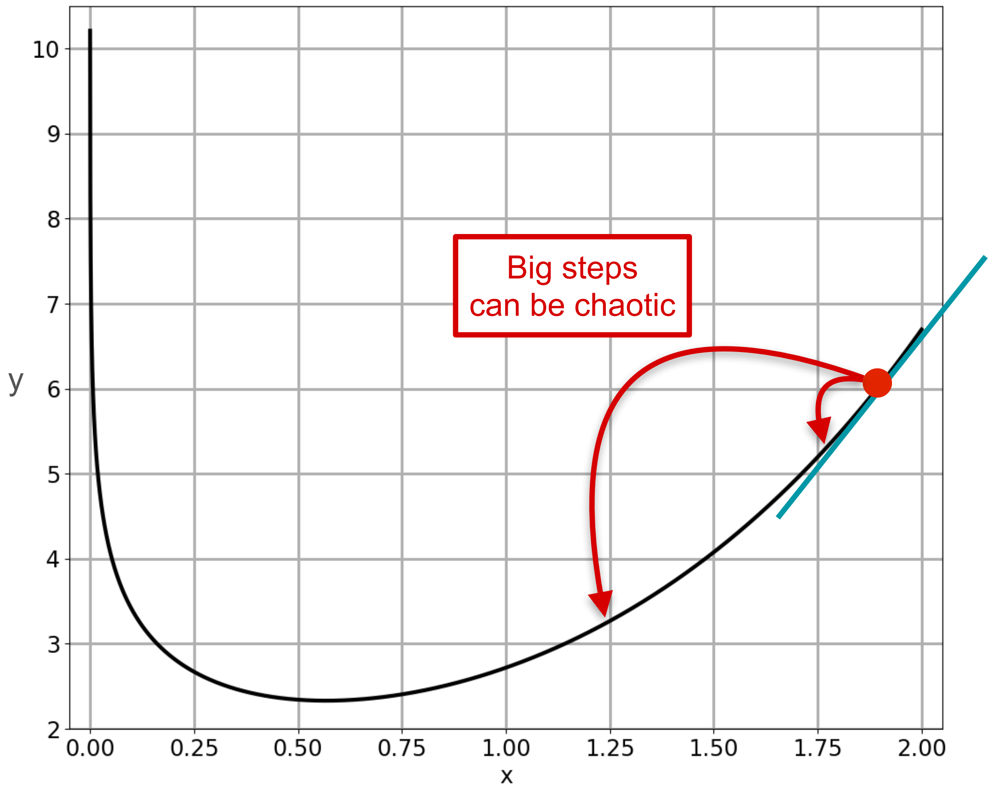


# Method 2: Be Clever

Try something smarter…

new point = old point - slope

$$x_1 = x_0 \quad f'(x_0)$$



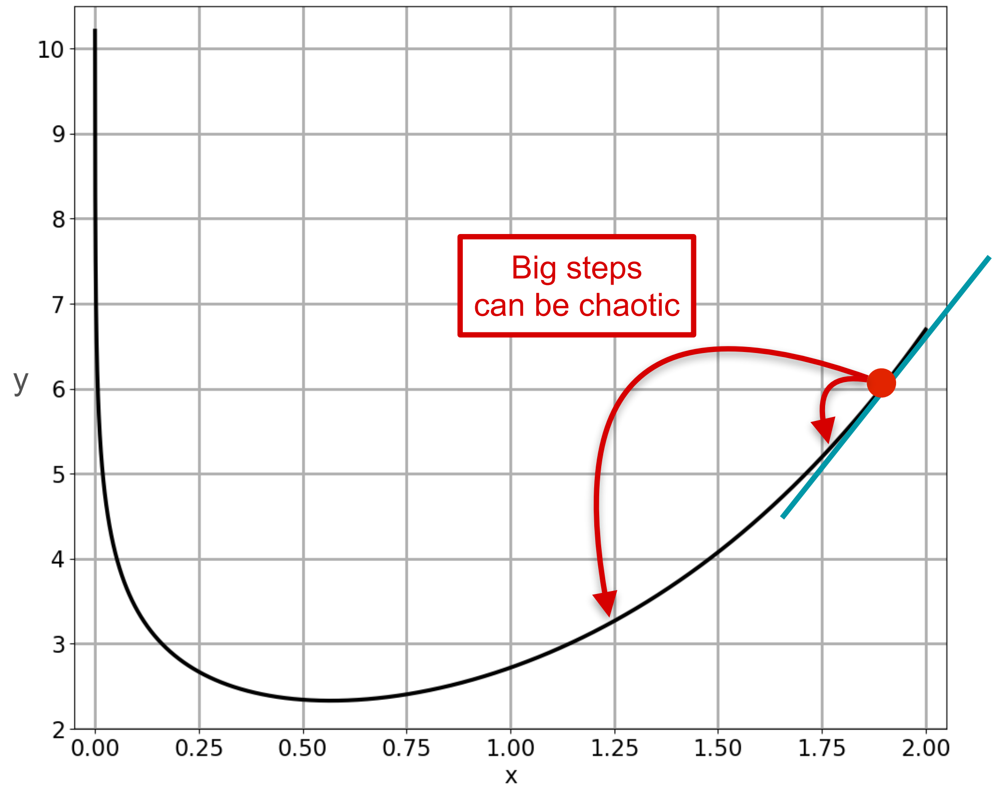
# Method 2: Be Clever

Try something smarter…

new point = old point - slope

$$x_1 = x_0 \quad f\,(x_0)$$

$$x_1 = x_0 \quad 0.01\,f\,(x_0)$$



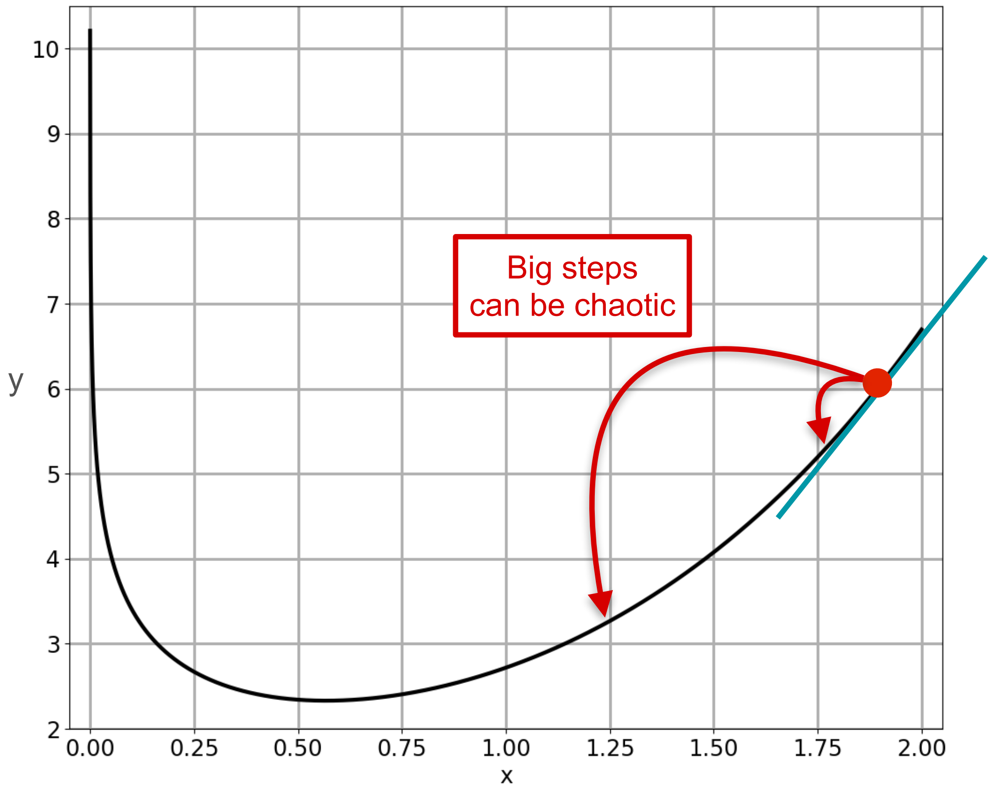
# Method 2: Be Clever

Try something smarter…

new point = old point - slope

$$x_1 = x_0 \quad f(x_0)$$

$$x_1 = x_0 \qquad f(x_0)$$



# Method 2: Be Clever

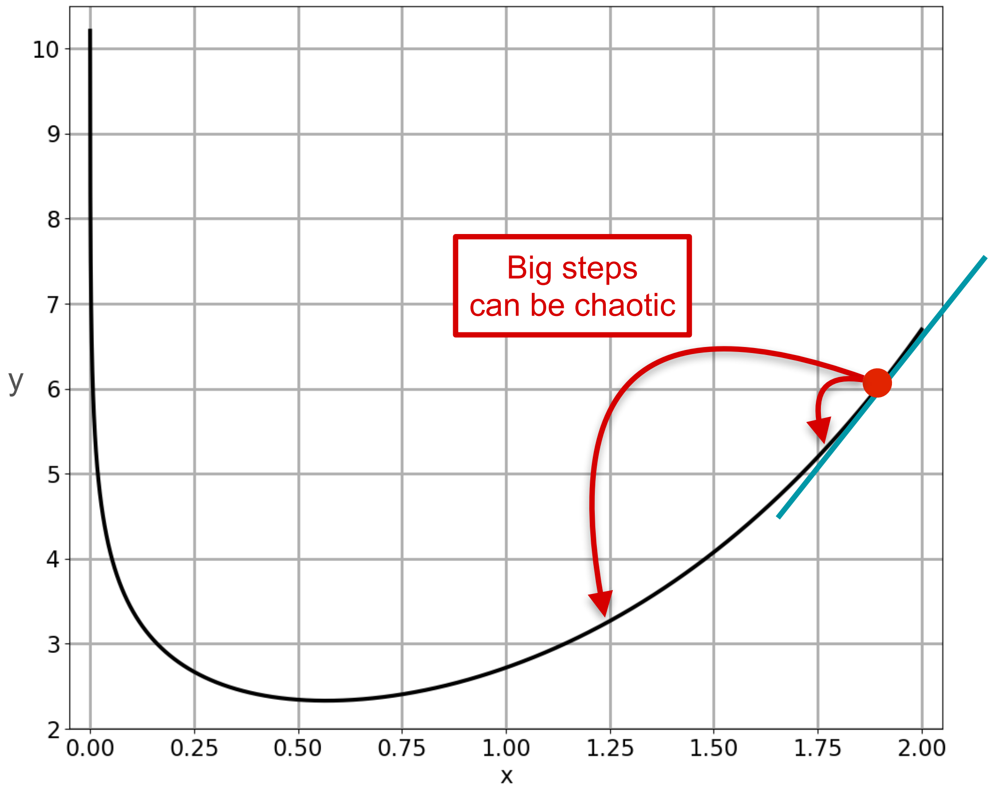
Try something smarter…

new point = old point - slope

$$x_1 = x_0 \quad f(x_0)$$

$$x_1 = x_0 \qquad \quad f(x_0)$$

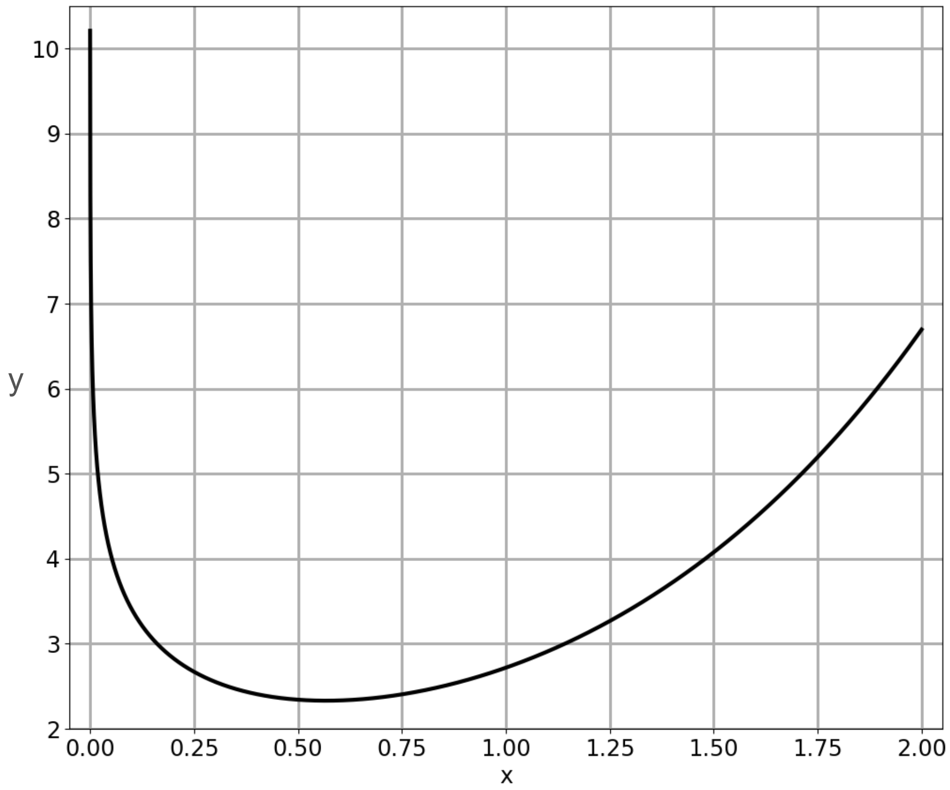
Learning rate



# Method 2: Be Clever

Try something smarter…

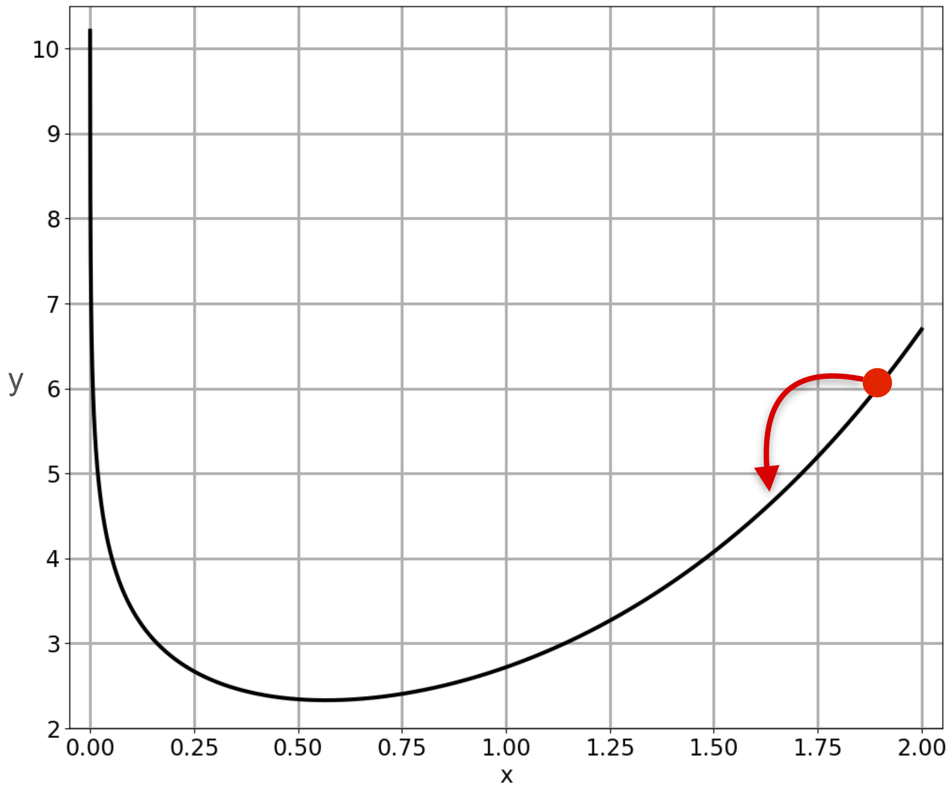
$$x_1 = x_0 \qquad f(x_0)$$



# Method 2: Be Clever

Try something smarter…

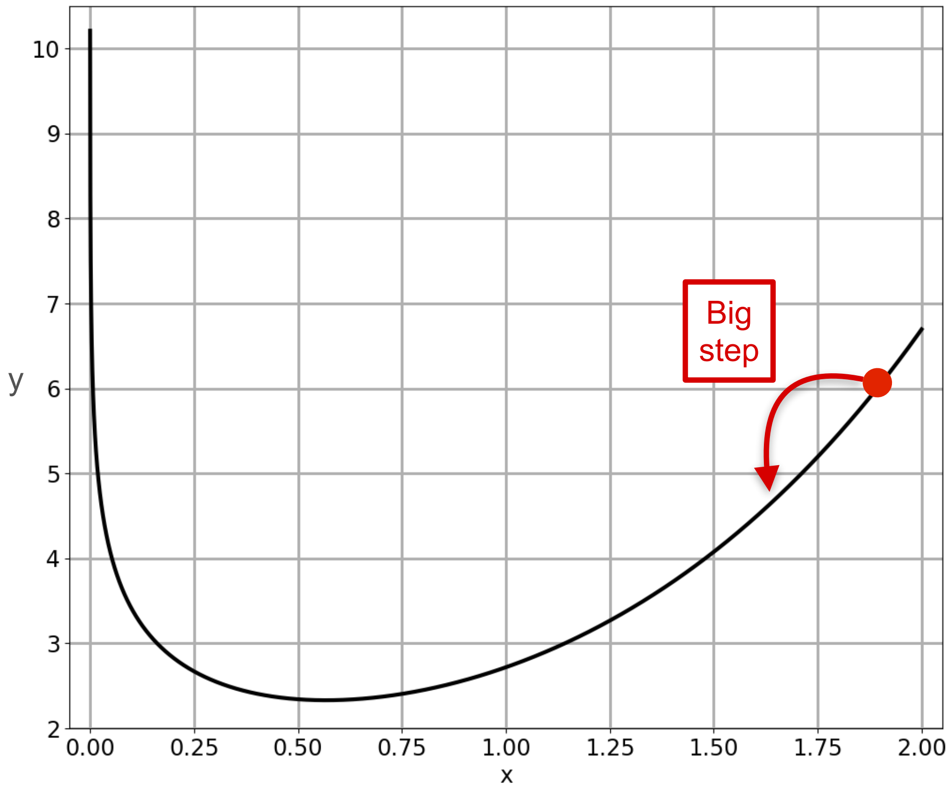
$$x_1 \quad = \quad x_0 \qquad\qquad f\,(x_0)$$



# Method 2: Be Clever

Try something smarter…

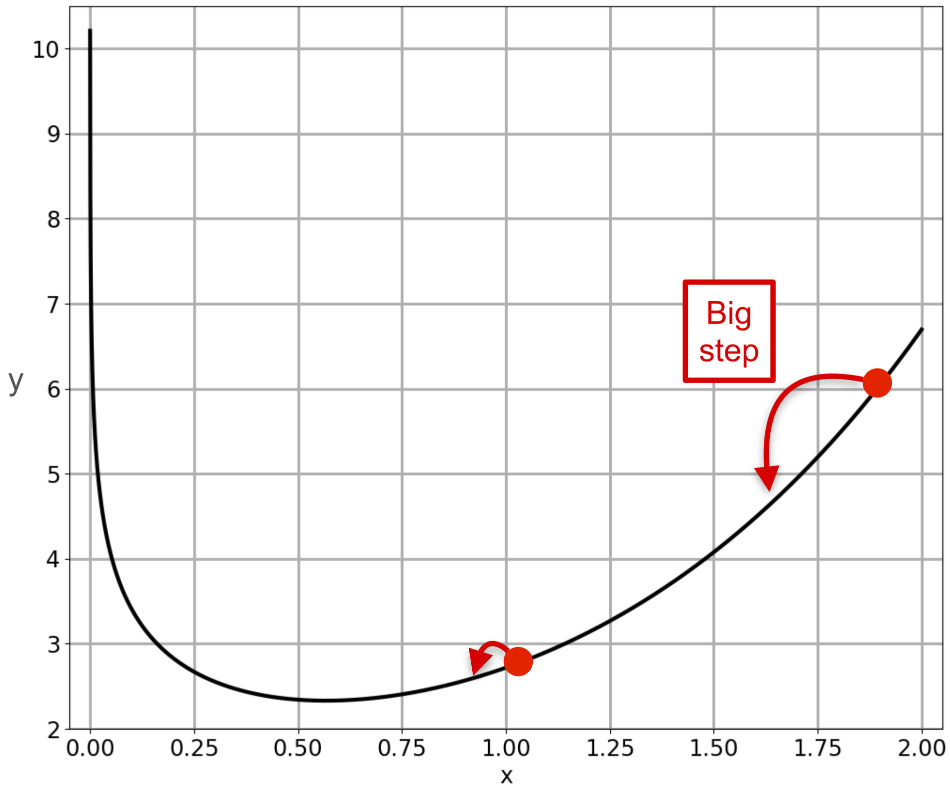
$$x_1 = x_0 \qquad f(x_0)$$



# Method 2: Be Clever

Try something smarter…

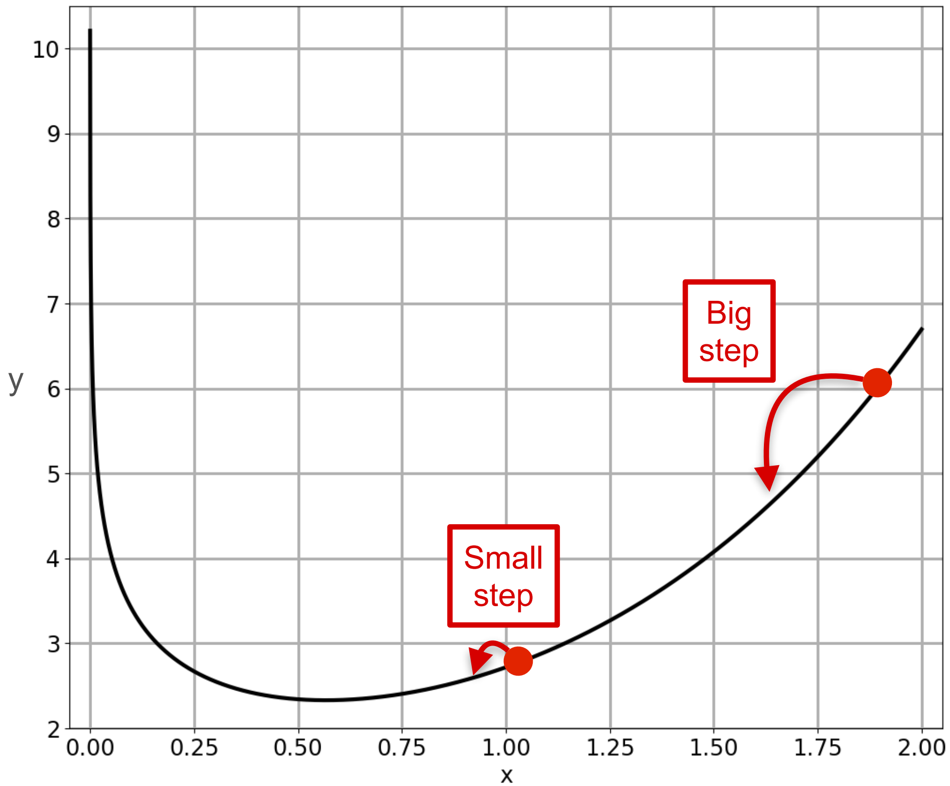
$x_1 = x_0 \qquad f(x_0)$



# Method 2: Be Clever

Try something smarter…

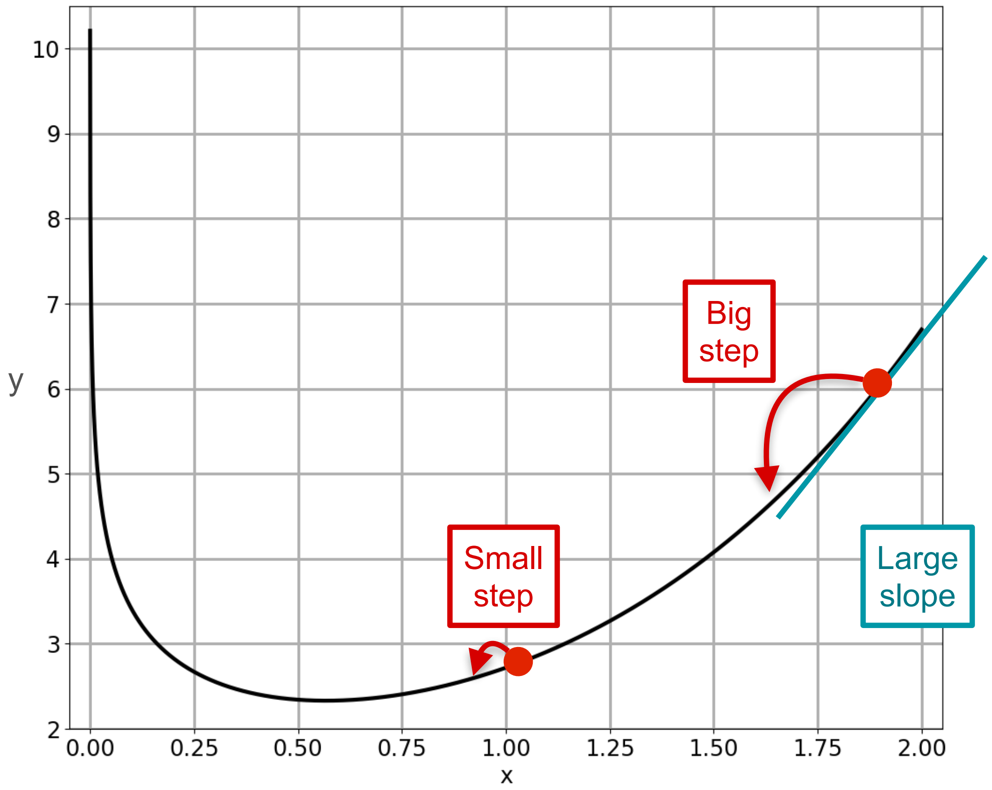
$$x_1 = x_0 \qquad f(x_0)$$



# Method 2: Be Clever

Try something smarter…

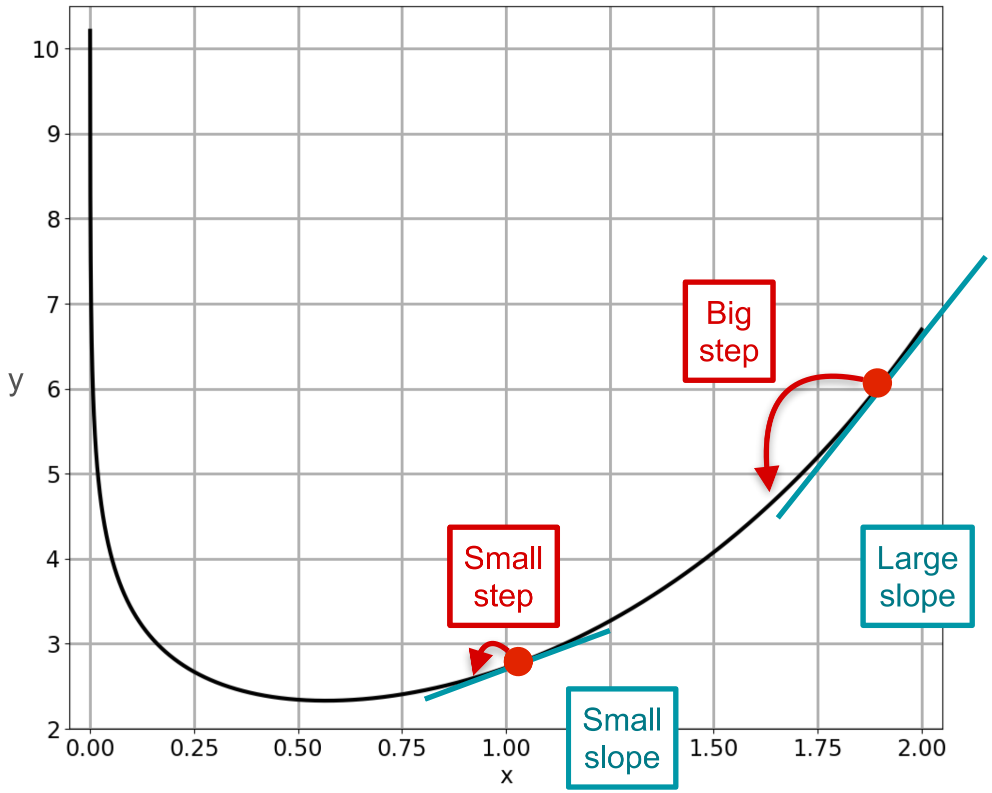
$$x_1 = x_0 \qquad\qquad f(x_0)$$



# Method 2: Be Clever

Try something smarter…

$x_1 = x_0 \qquad f(x_0)$

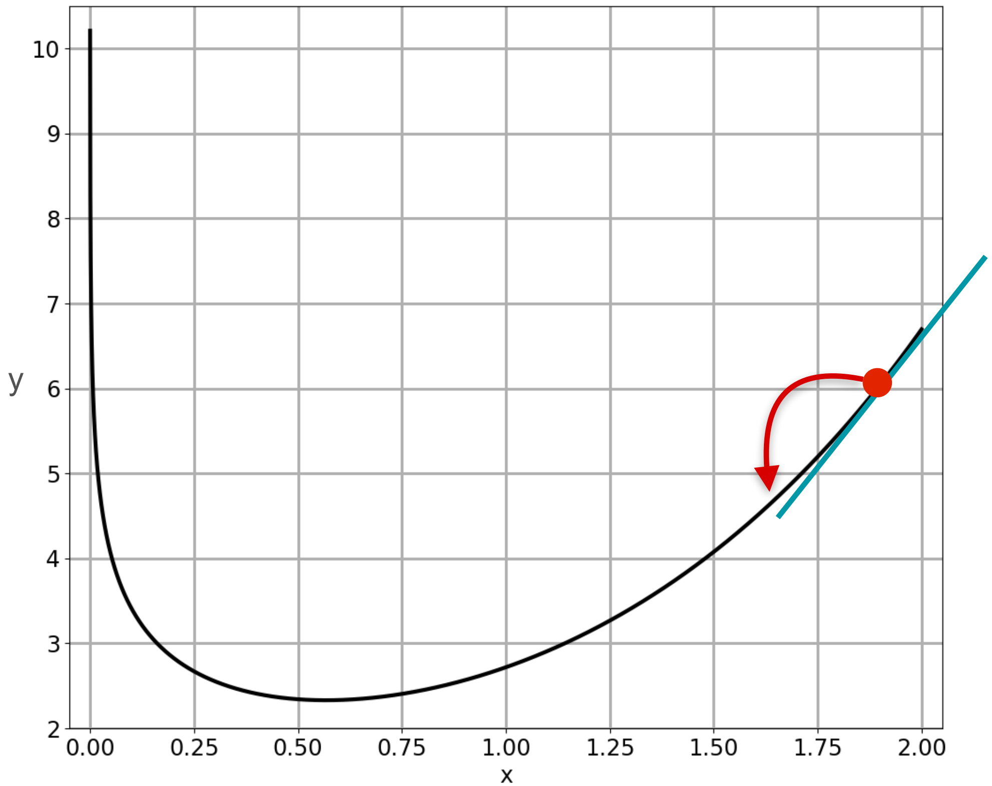


# Method 2: Be Clever

Try something smarter…

$$x_1 = x_0 \qquad\qquad f'(x_0)$$

Gradient descent

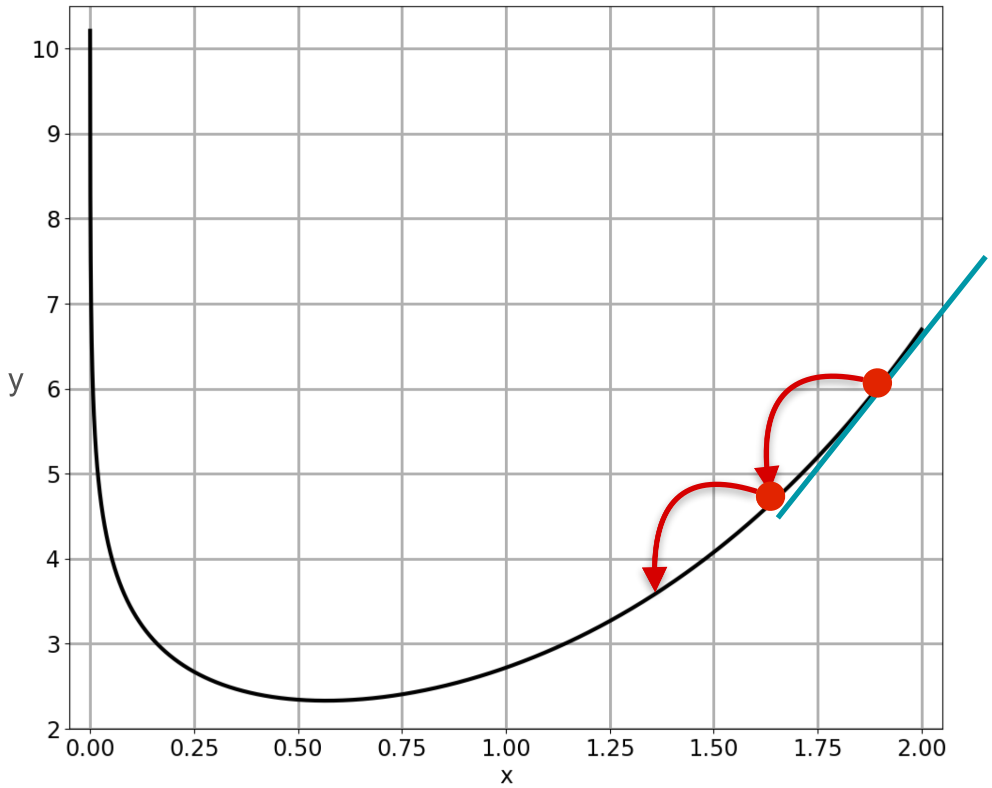


# Method 2: Be Clever

Try something smarter…

$$x_2 = x_1 \qquad f(x_1)$$

## Gradient descent

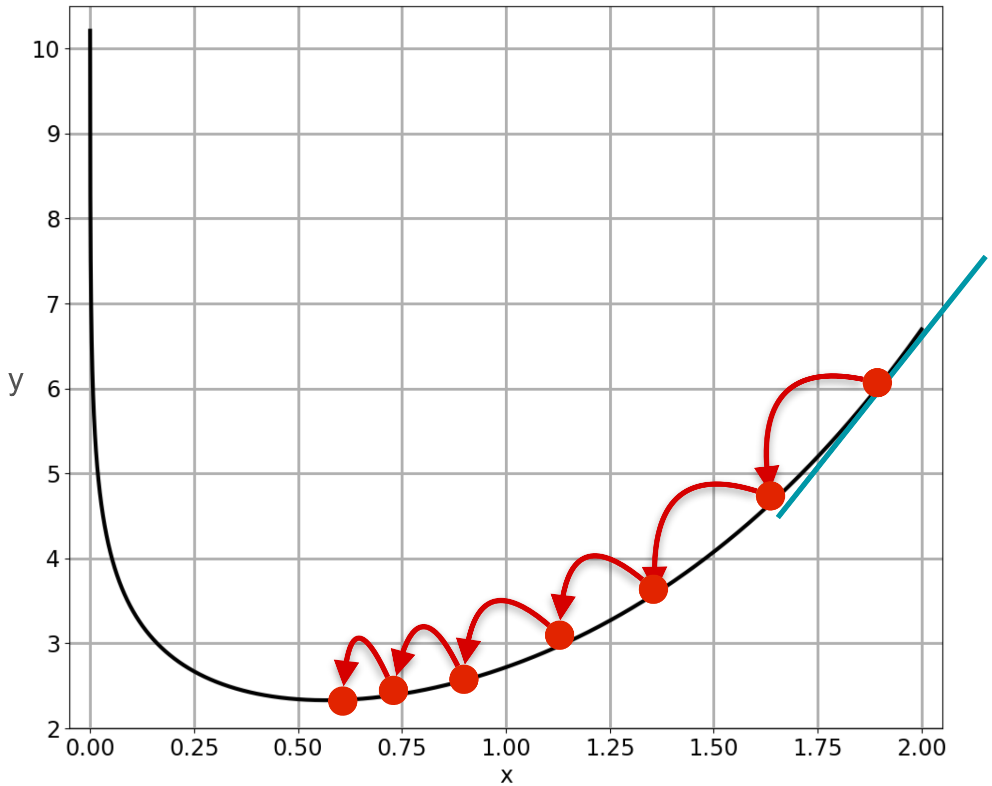


# Method 2: Be Clever

Try something smarter…

$x_{20} = \ x_{19} \qquad \qquad f\,(x_{19})$

Gradient descent



# Gradient Descent

Function: $f(x)$ Goal: find minimum of $f(x)$

Step 1:

Define a learning rate

Choose a starting point $x_0$

Step 2:

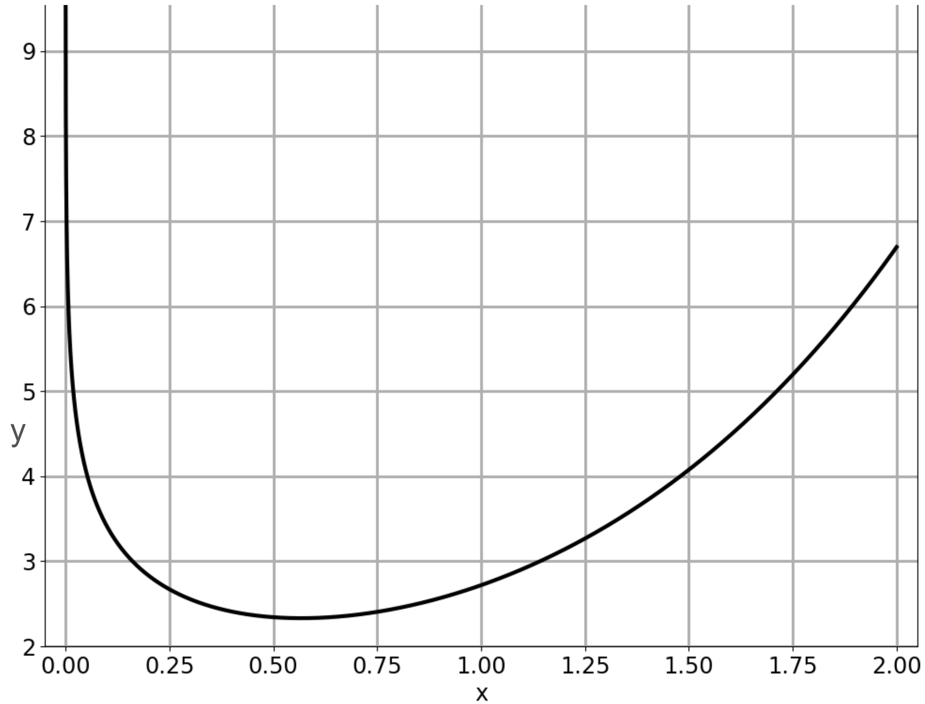
Update: $x_k = x_{k\ 1} \quad a f\ (x_{k\ 1})$

Step 3:

Repeat Step 2 until you are close enough to the true minimum $x^*$

# Gradient Descent

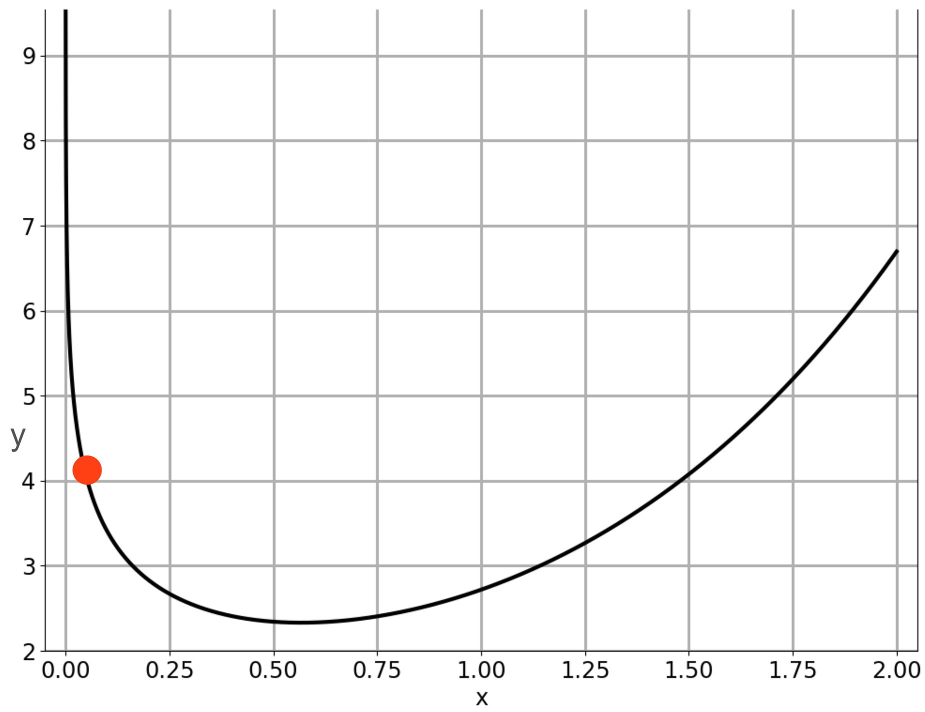
$f(x) = e^x \quad \log(x)$ $\qquad$ $f(x) = e^x \quad \frac{1}{x}$



# Gradient Descent

$f(x) = e^x \quad \log(x)$ $\qquad$ $f\,(x) = e^x \quad \frac{1}{x}$

Start: $x = 0.05$ $\qquad$ Rate: $\quad = 0.005$



# Gradient Descent

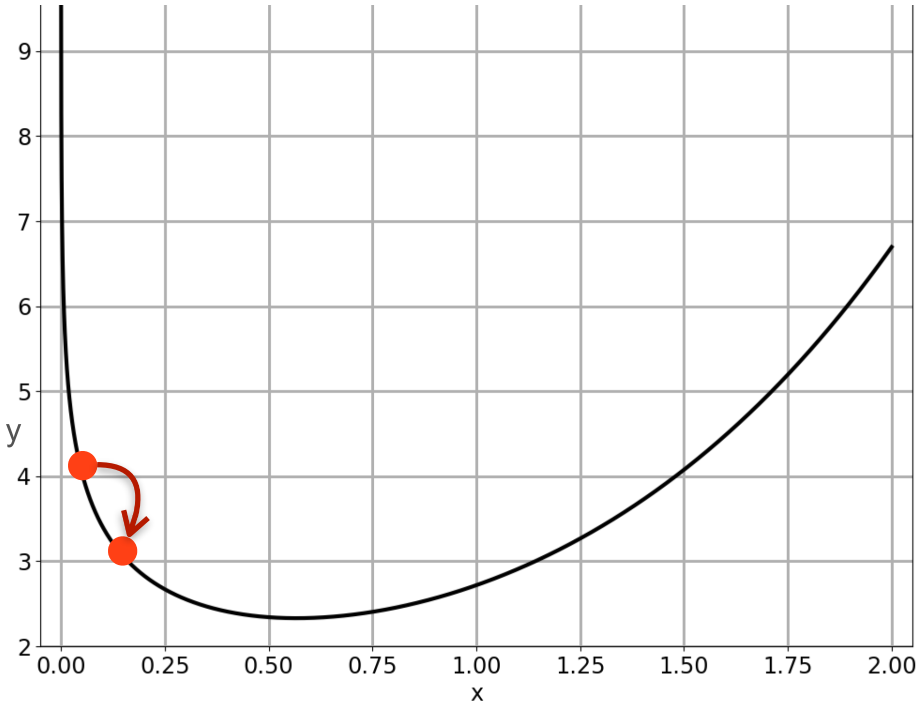
$f(x) = e^x \quad \log(x)$ $\qquad$ $f(x) = e^x \quad \frac{1}{x}$

Start: $x = 0.05$ $\qquad$ Rate: $= 0.005$

Find:

$f(0.05) = \quad 18.9$

Move by $0.005 f(0.05) \qquad x \quad 0.1447$



# Gradient Descent

$f(x) = e^x \quad \log(x)$ $\qquad$ $f(x) = e^x \quad \frac{1}{x}$

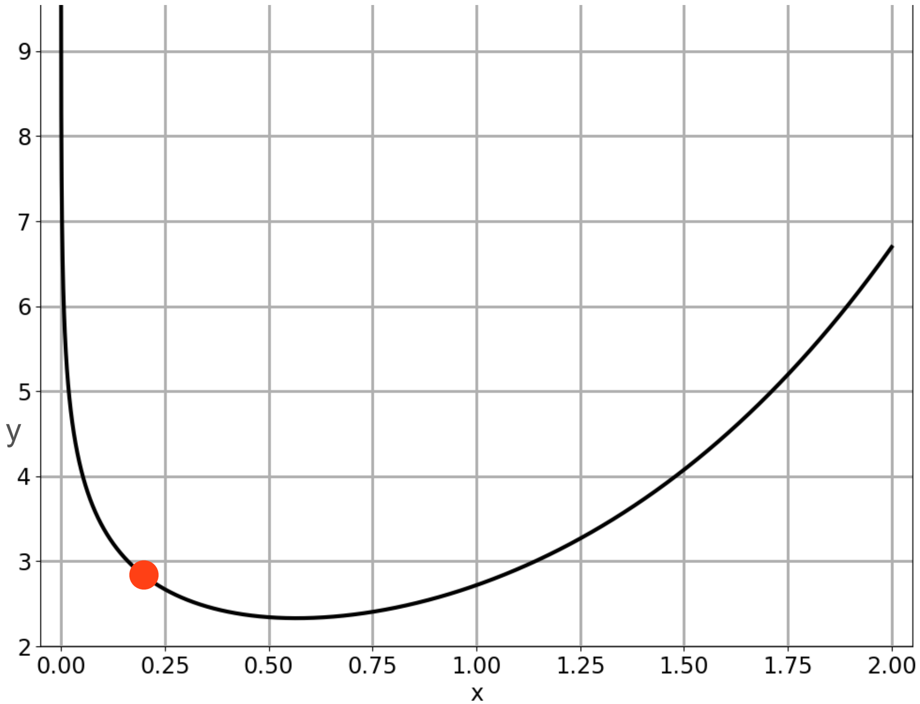
Start: $x = 0.05$ $\qquad$ Rate: $= 0.005$

Find:
$f(0.05) = \quad 18.9$

Move by $\quad 0.005\, f(0.05) \qquad x \quad 0.1447$

Find:
$f(0.1447) = \quad 5.755$

Move by $\quad 0.005\, f(0.05) \qquad x \quad 0.1735$



# Gradient Descent

$f(x) = e^x \quad \log(x)$ $\qquad$ $f(x) = e^x \quad \frac{1}{x}$

Start: $x = 0.05$ $\qquad$ Rate: $= 0.005$

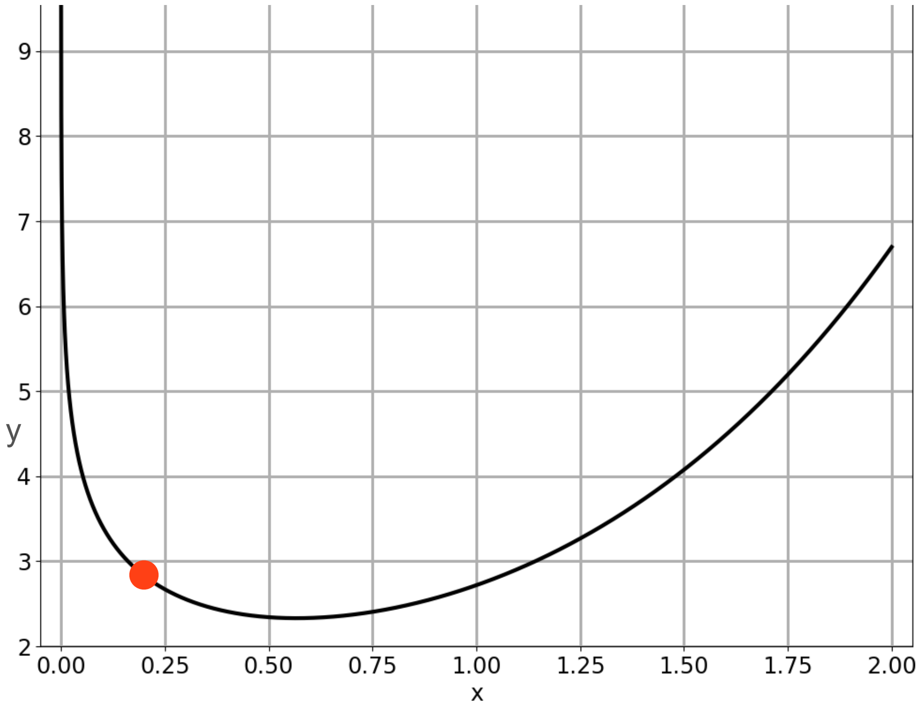
Find:
$f(0.05) = \quad 18.9$

Move by $\quad 0.005 \, f(0.05) \qquad x \quad 0.1447$

Find:
$f(0.1447) = \quad 5.755$

Move by $\quad 0.005 \, f(0.05) \qquad x \quad 0.1735$

Repeat!



# Gradient Descent

$f(x) = e^x \quad \log(x)$ $\qquad$ $f(x) = e^x \quad \frac{1}{x}$

Start: $x = 0.05$ $\qquad$ Rate: $= 0.005$

Find:

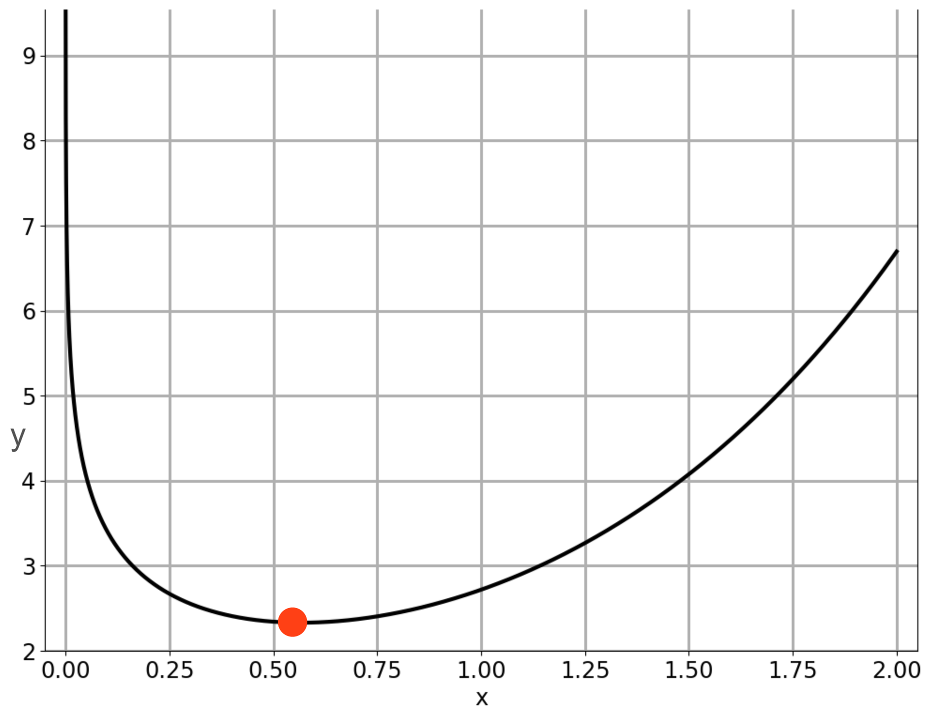
$f(0.05) = \quad 18.9$

Move by $\quad 0.005\, f(0.05) \qquad x \quad 0.1447$

Find:

$f(0.1447) = \quad 5.755$

Move by $\quad 0.005\, f(0.05) \qquad x \quad 0.1735$




Repeat!

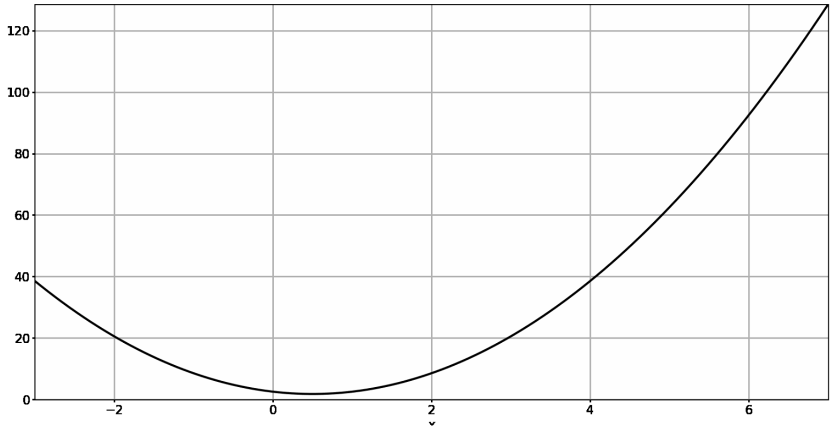
DeepLearning.AI

# Gradients and Gradient Descent

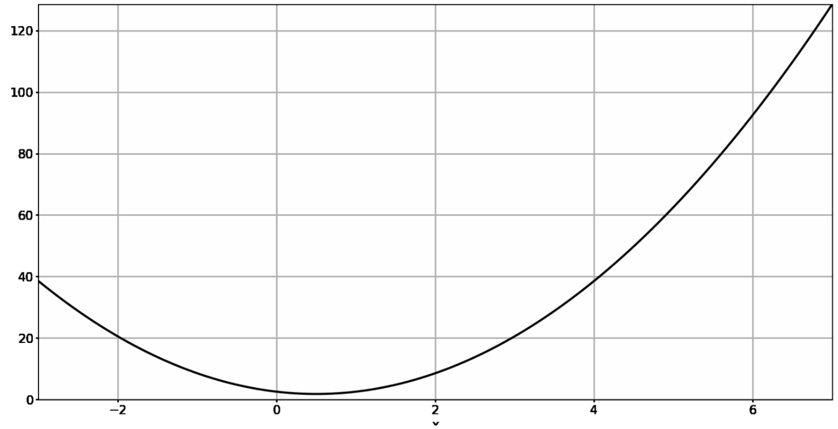
## Optimization using Gradient Descent in one variable - Part 3

# What Is a Good Learning Rate?

Too large



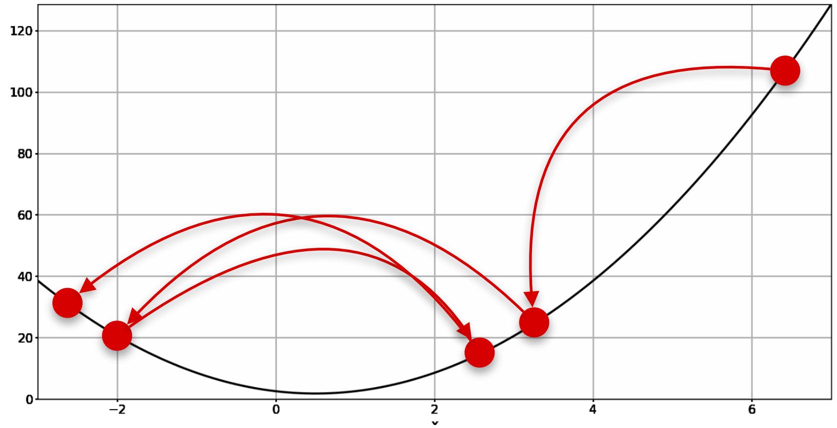

Too small



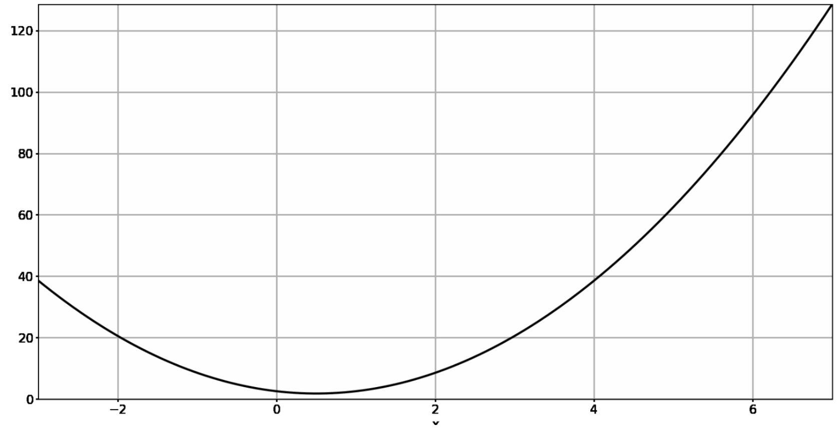

DeepLearning.AI

# What Is a Good Learning Rate?

Too large

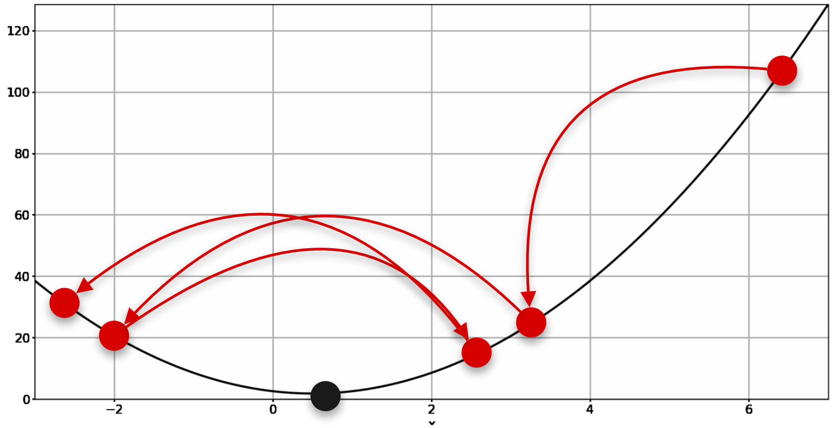



Too small

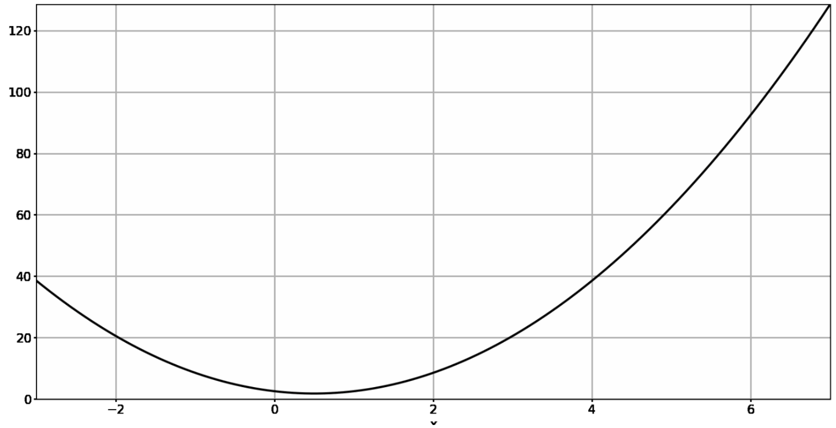


# What Is a Good Learning Rate?

Too large

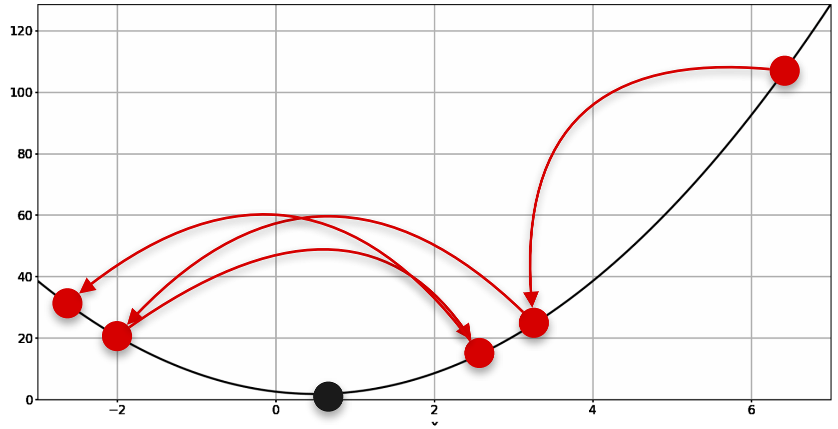



Too small

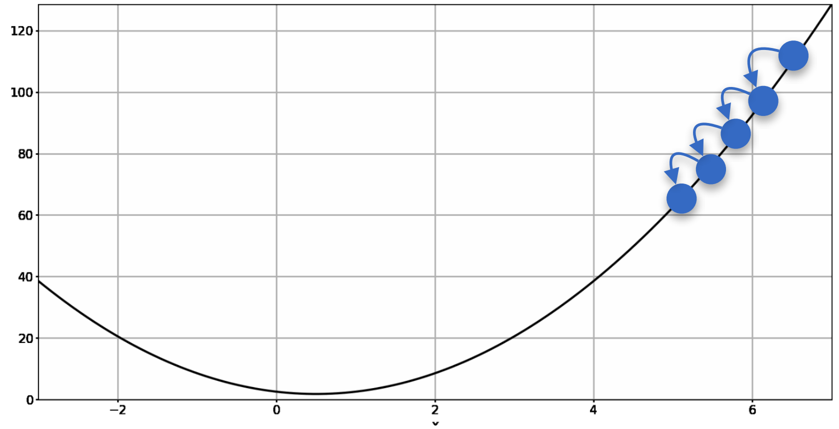


# What Is a Good Learning Rate?

Too large

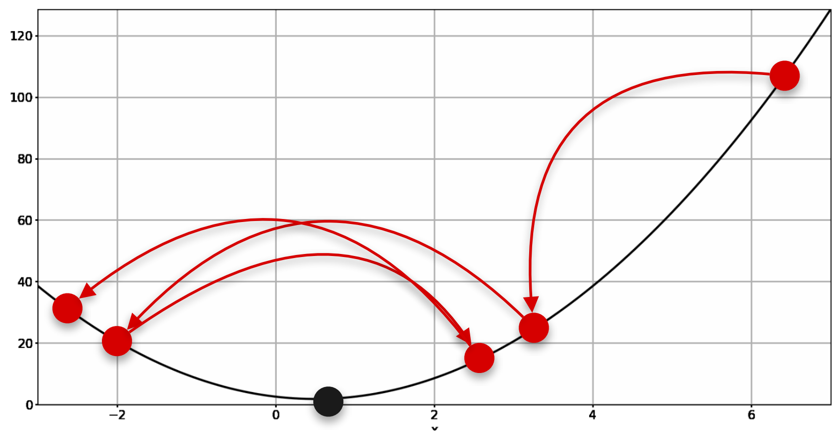



Too small

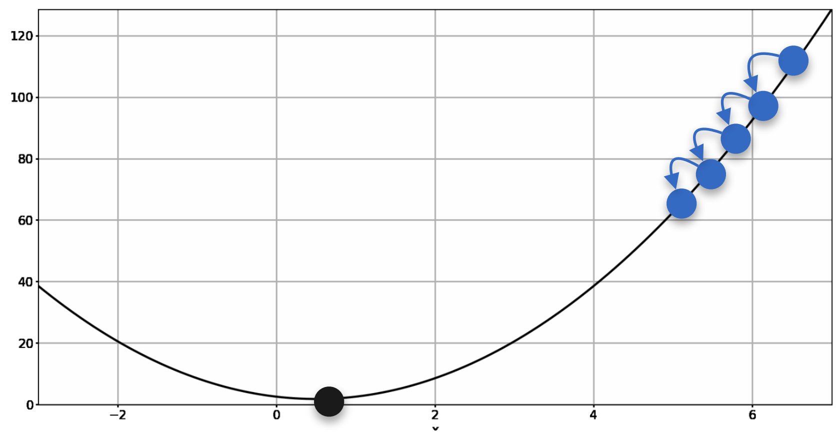


# What Is a Good Learning Rate?

Too large

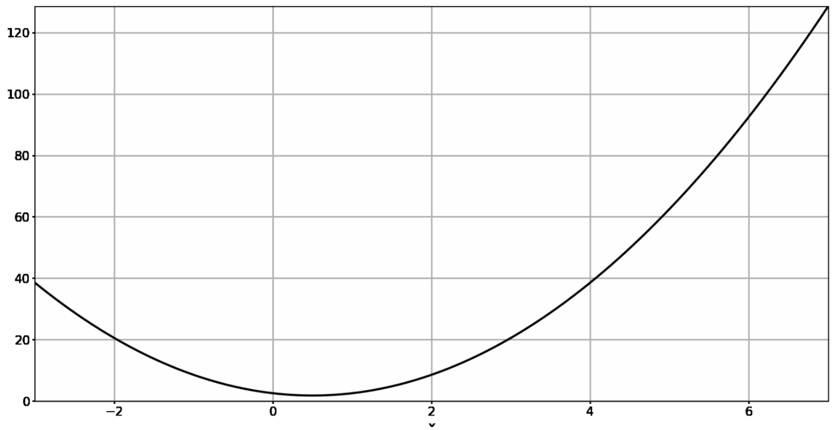



Too small



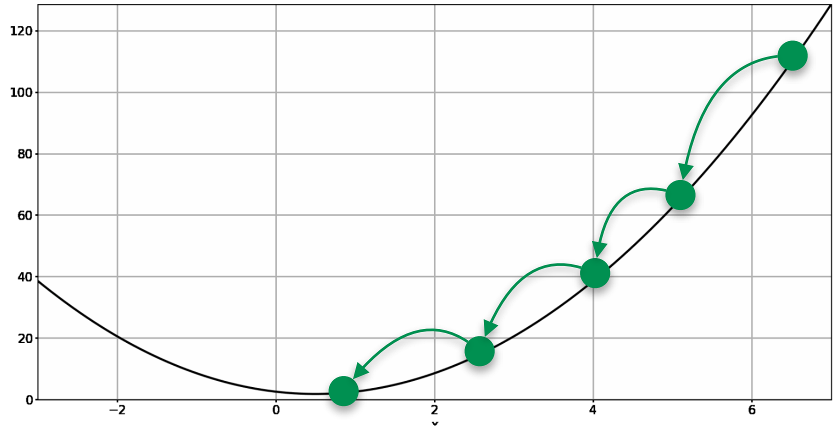
# What Is a Good Learning Rate?

Just right



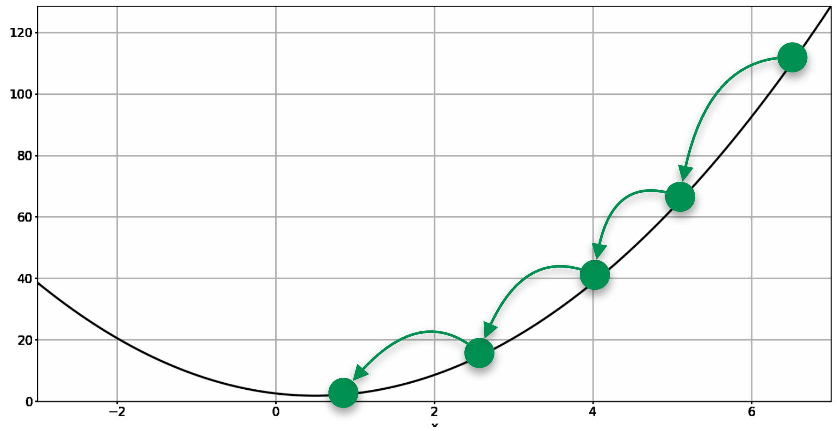
# What Is a Good Learning Rate?

Just right

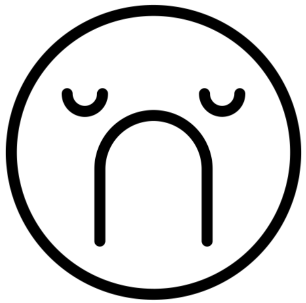


# What Is a Good Learning Rate?

Just right

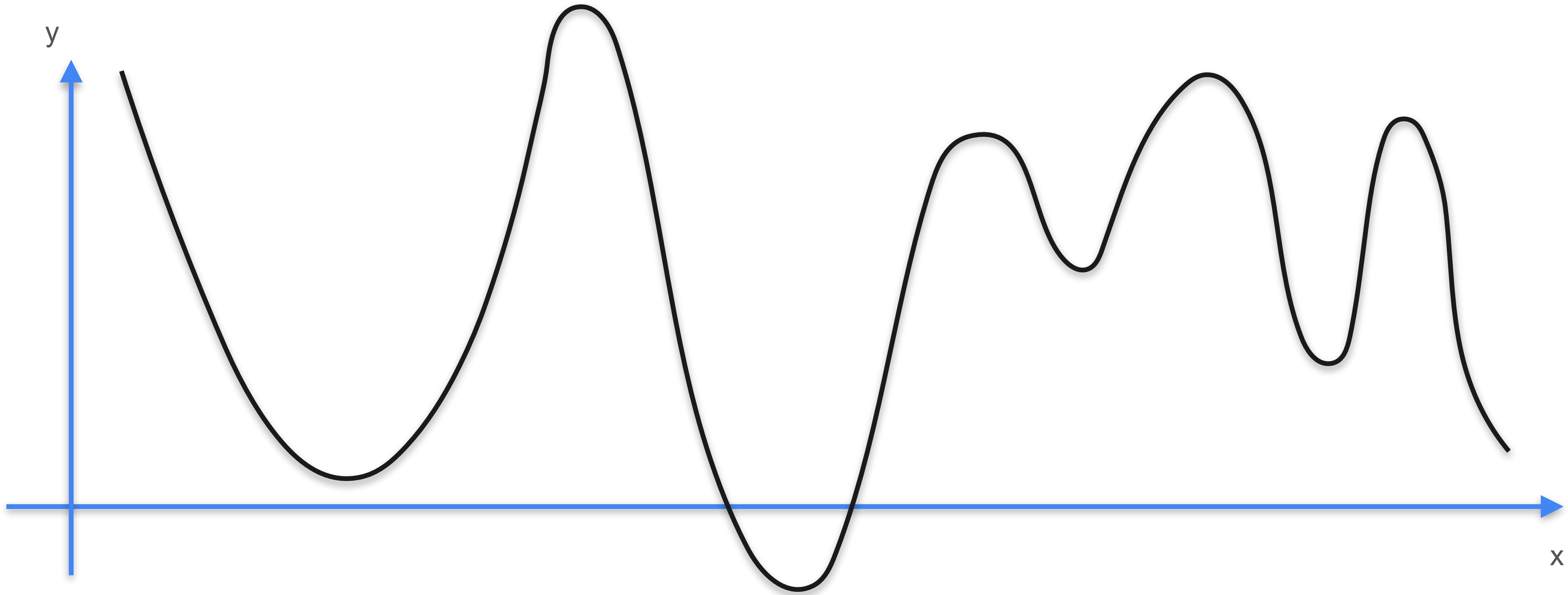



Unfortunately, there is no rule to give the best learning rate

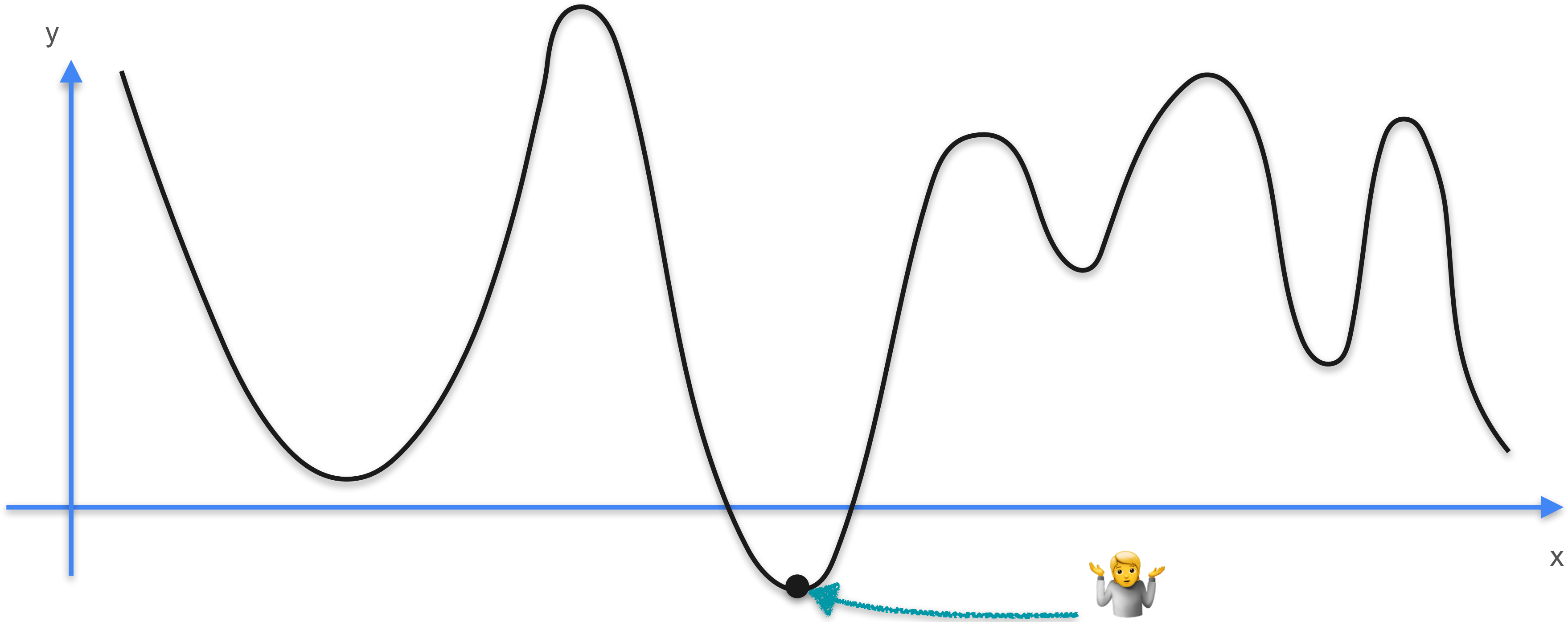


# Drawbacks of Gradient Descent

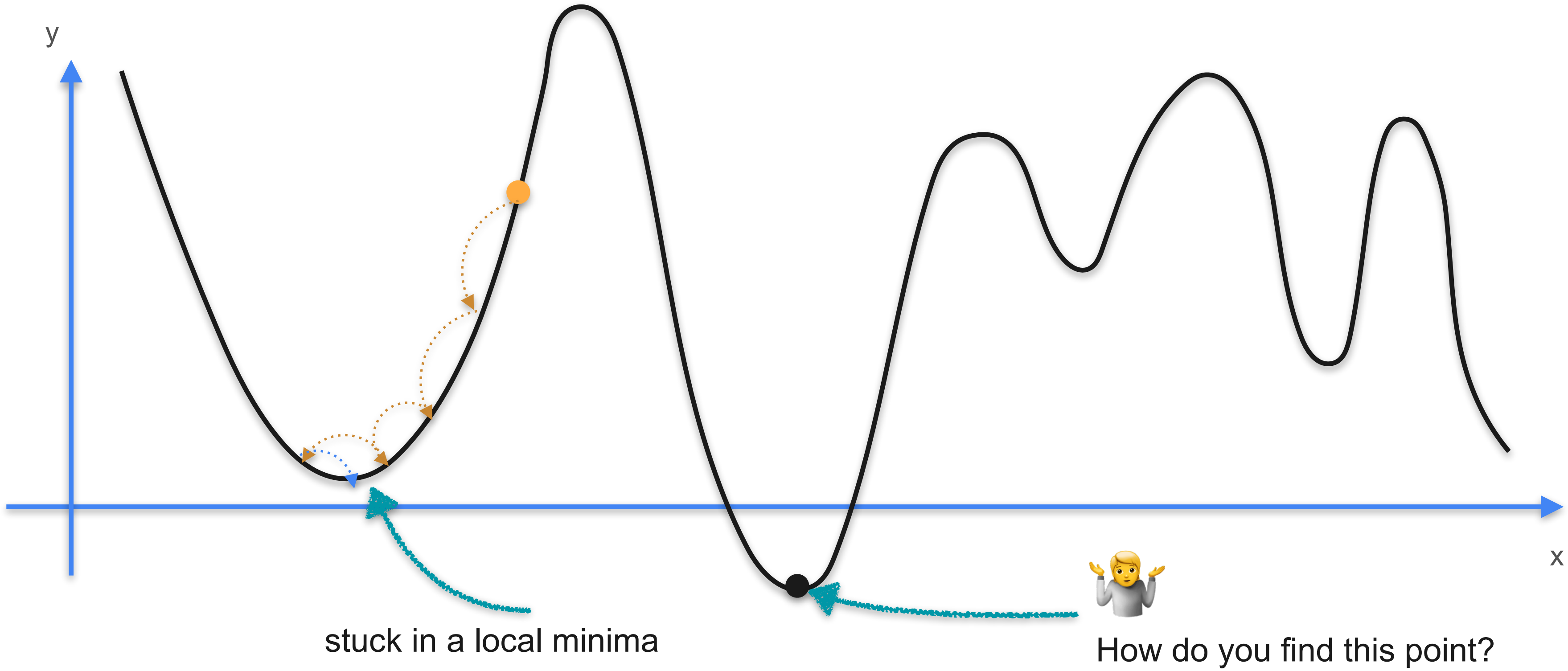
# Drawbacks of Gradient Descent



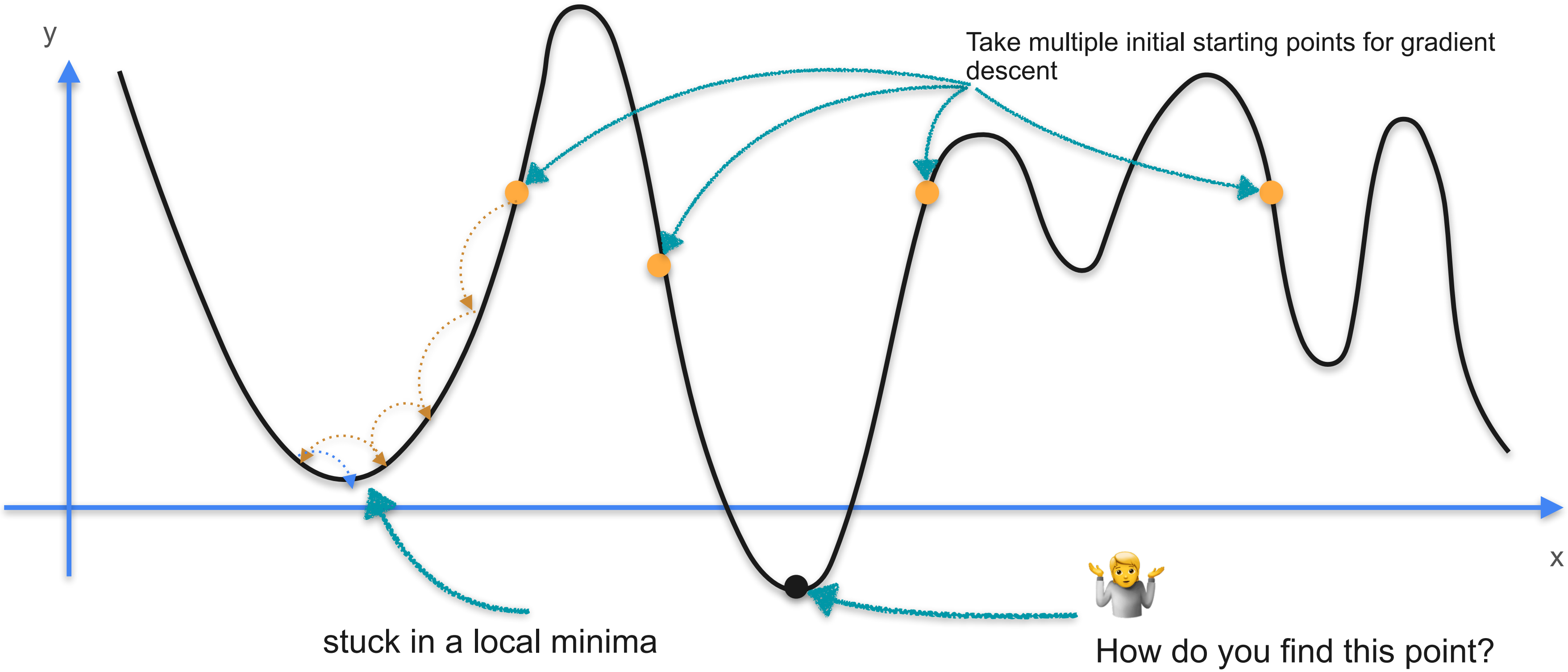
# Drawbacks of Gradient Descent



# Drawbacks of Gradient Descent



# Drawbacks of Gradient Descent

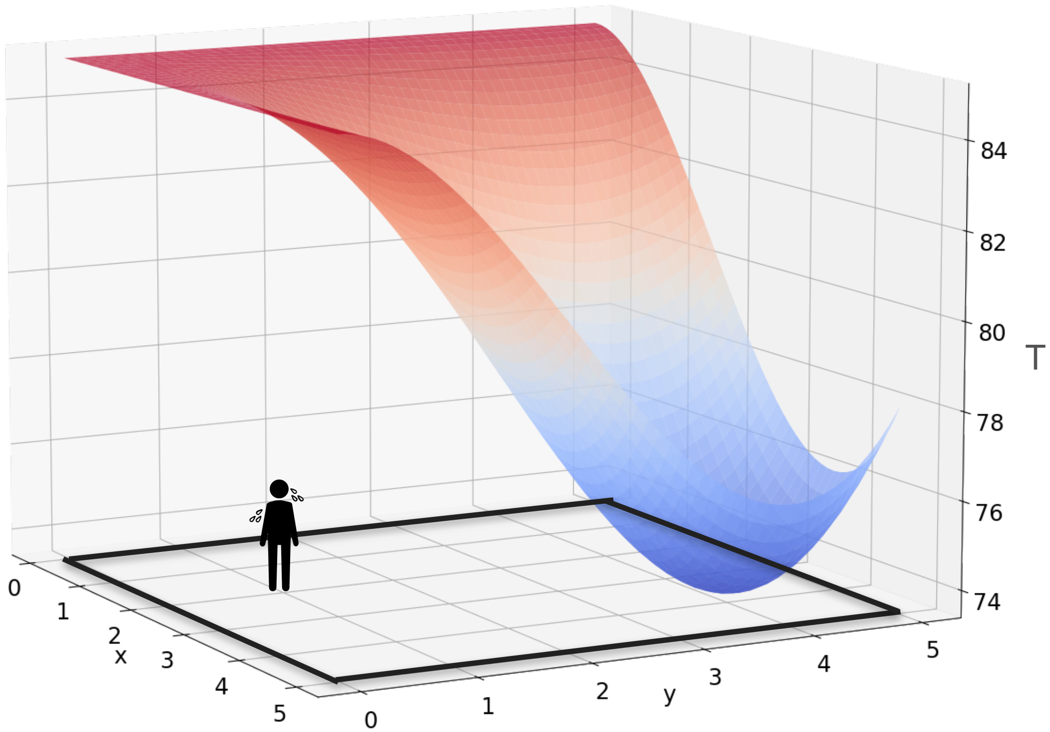


DeepLearning.AI

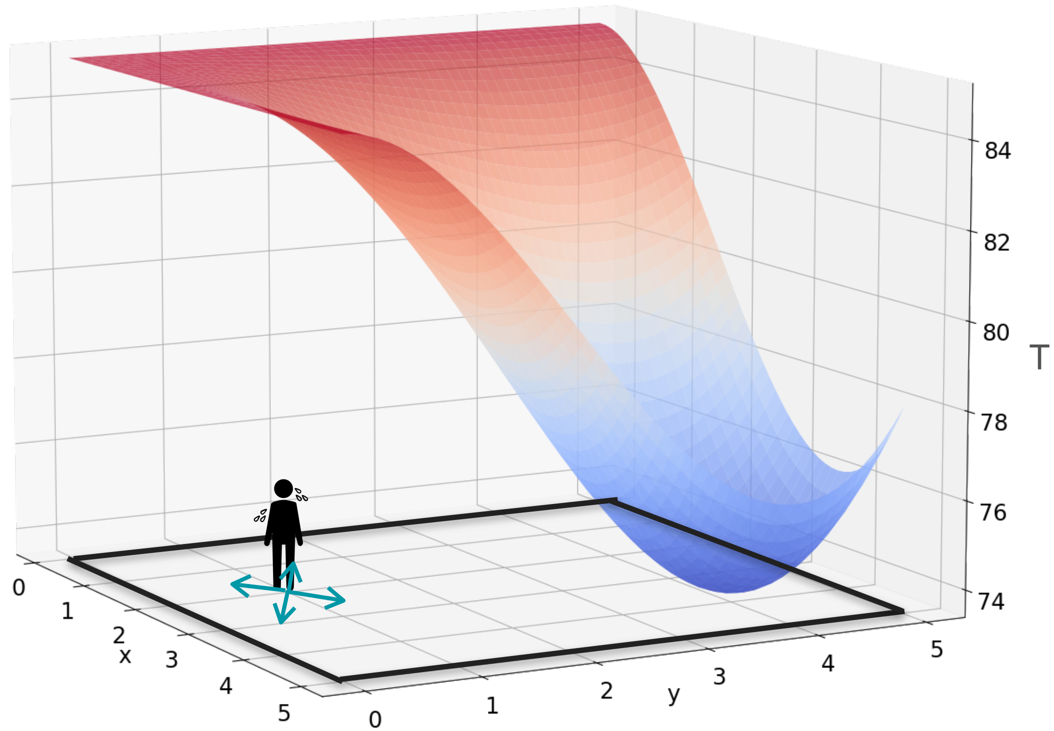
# Gradients and Gradient Descent

## Optimization using Gradient Descent in two variables - Part 1

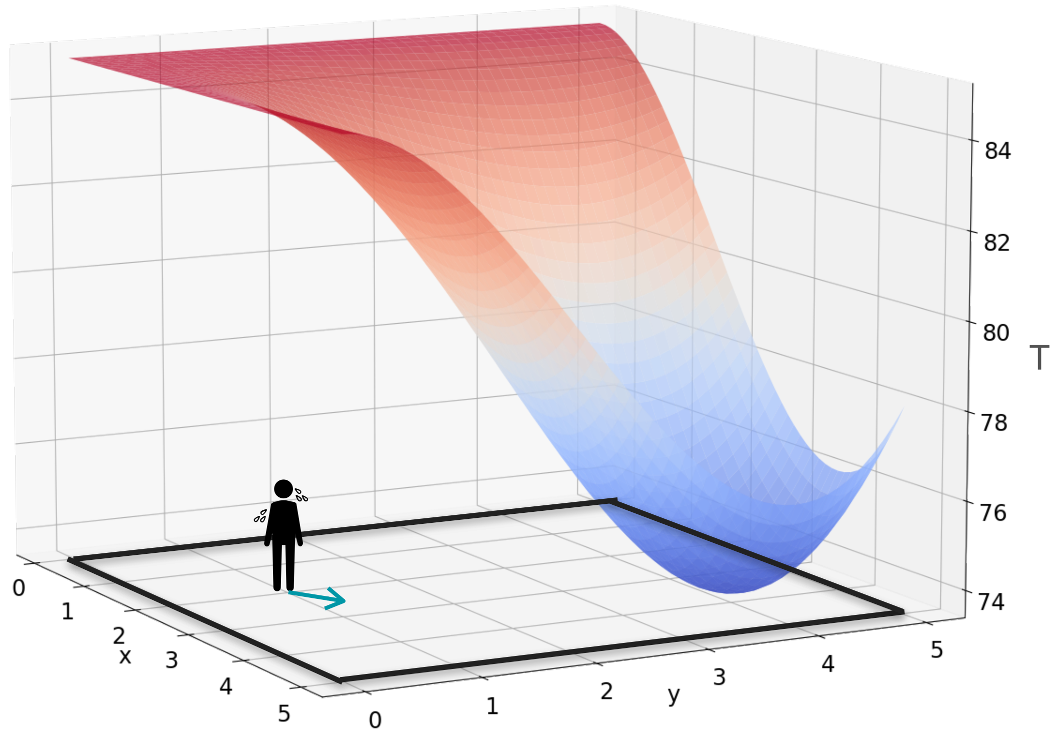
# Gradient Descent With Heat Example



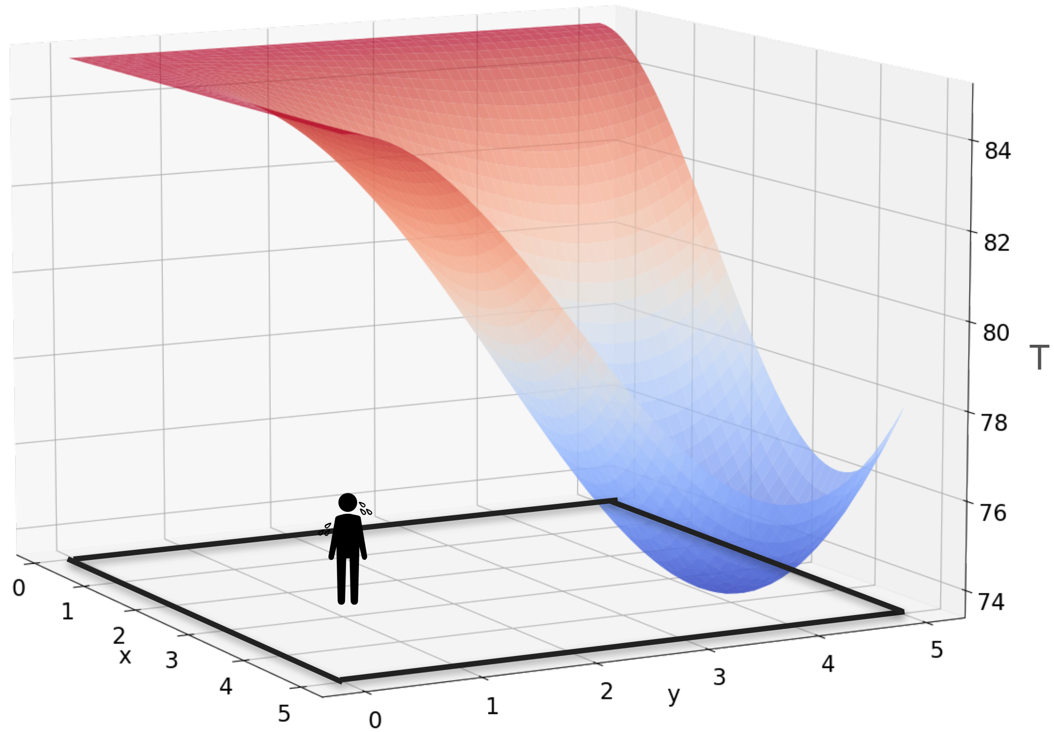
# Gradient Descent With Heat Example



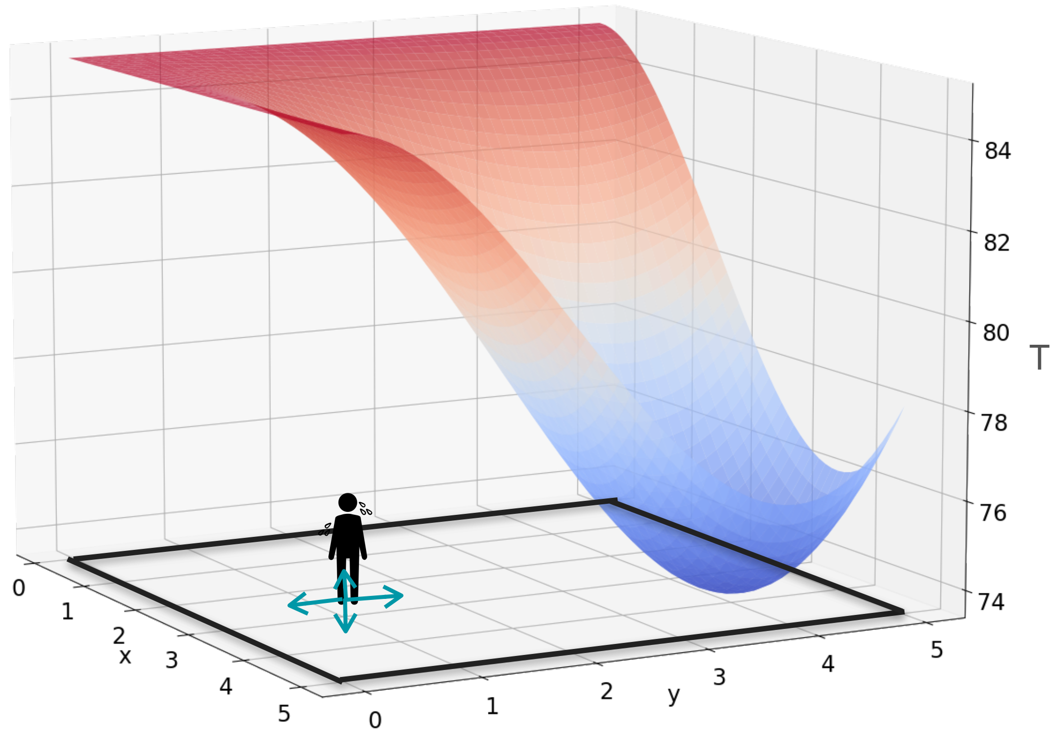
# Gradient Descent With Heat Example



# Gradient Descent With Heat Example

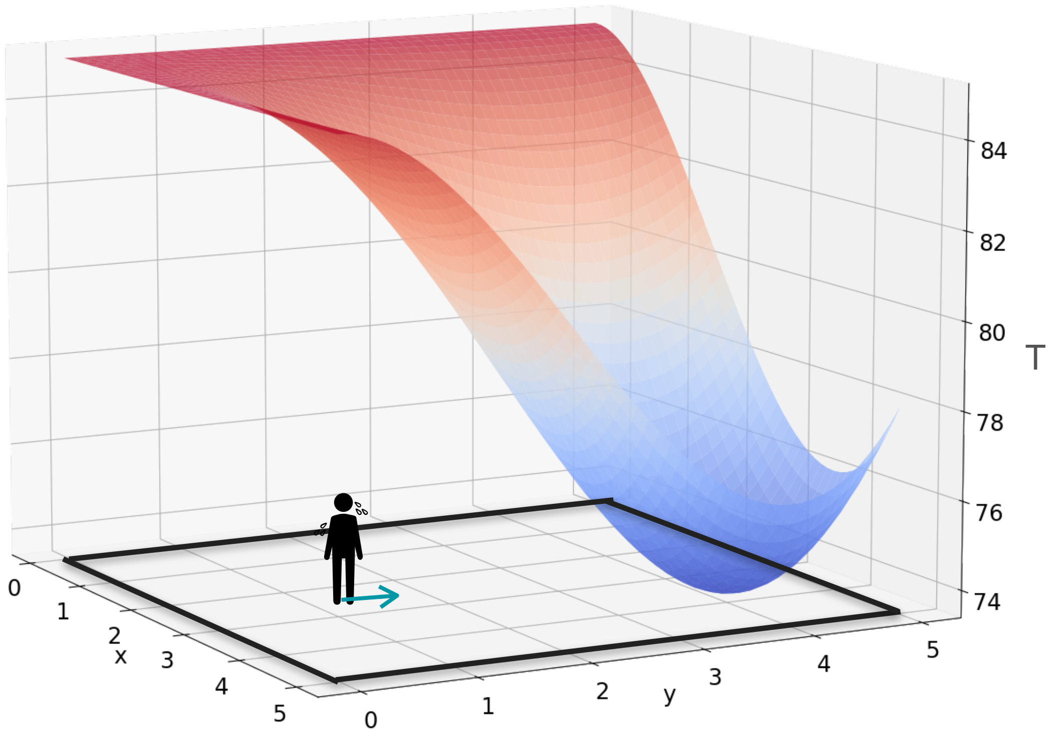


# Gradient Descent With Heat Example

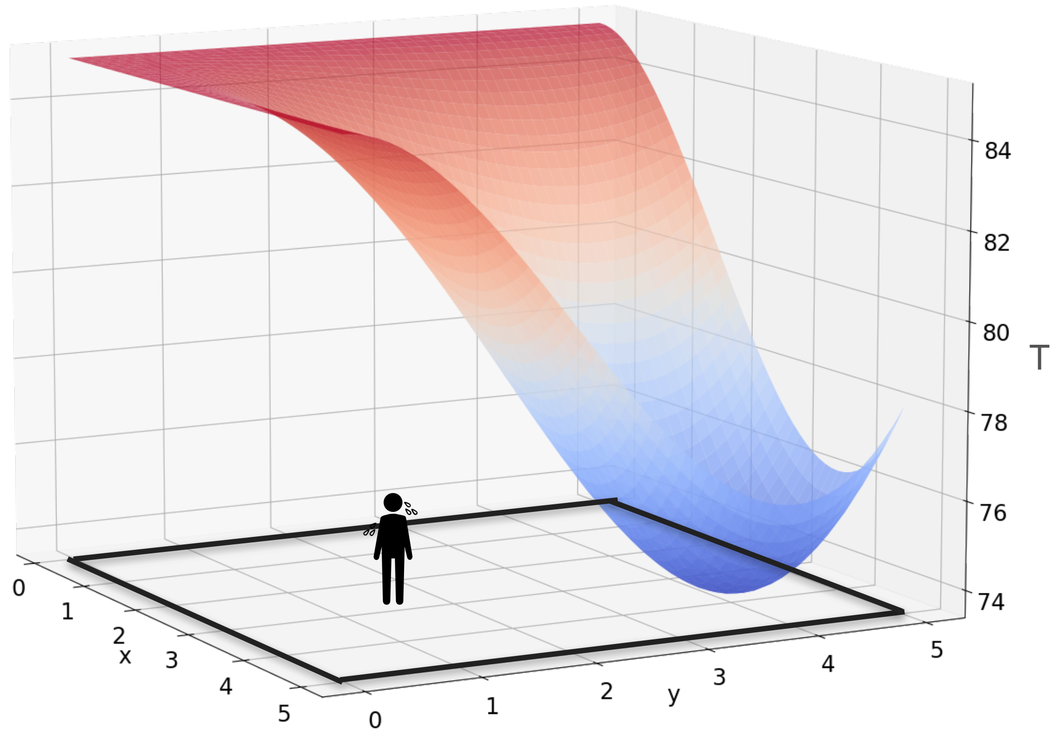


# Gradient Descent With Heat Example

Repeat!

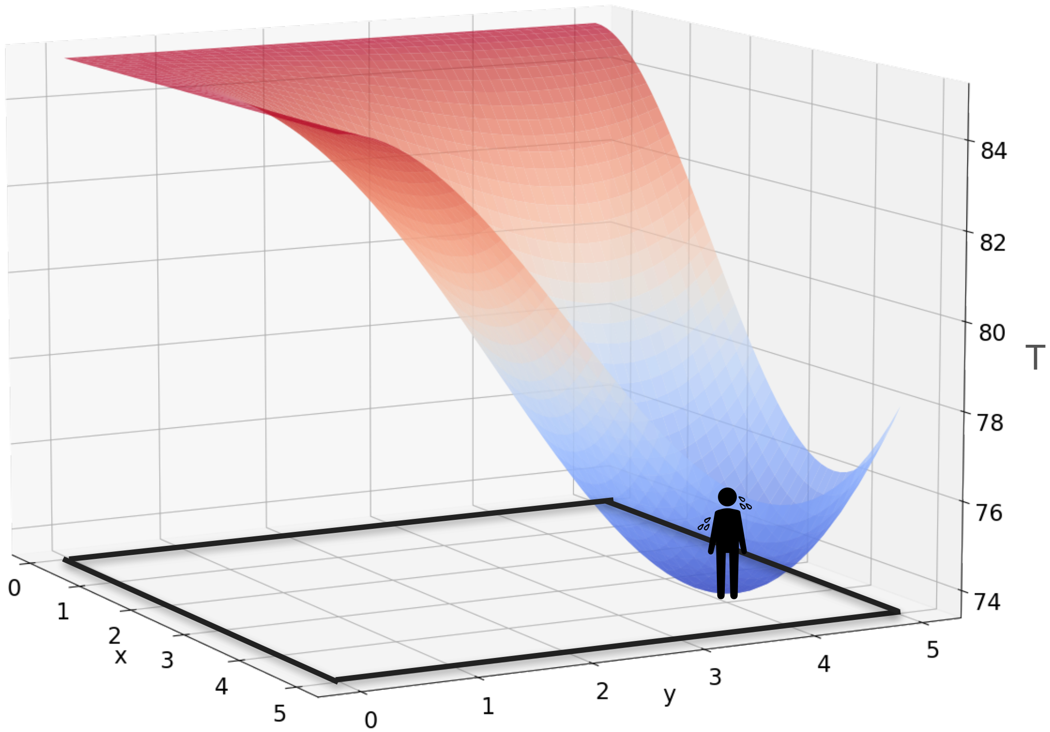


# Gradient Descent With Heat Example



# Gradient Descent With Heat Example

Repeat!

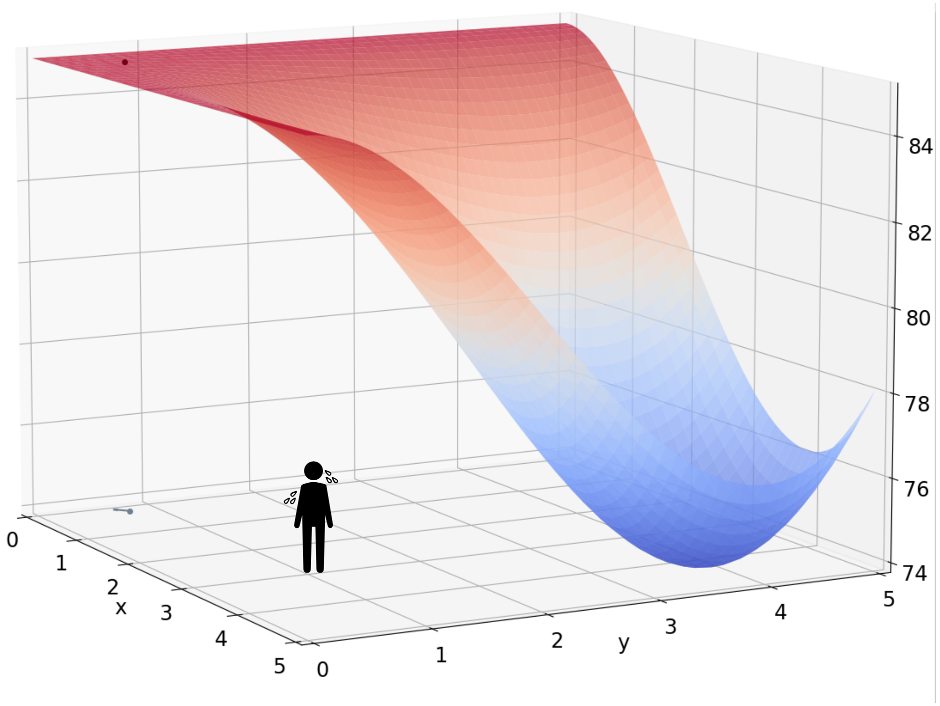


DeepLearning.AI

# Gradients and Gradient Descent

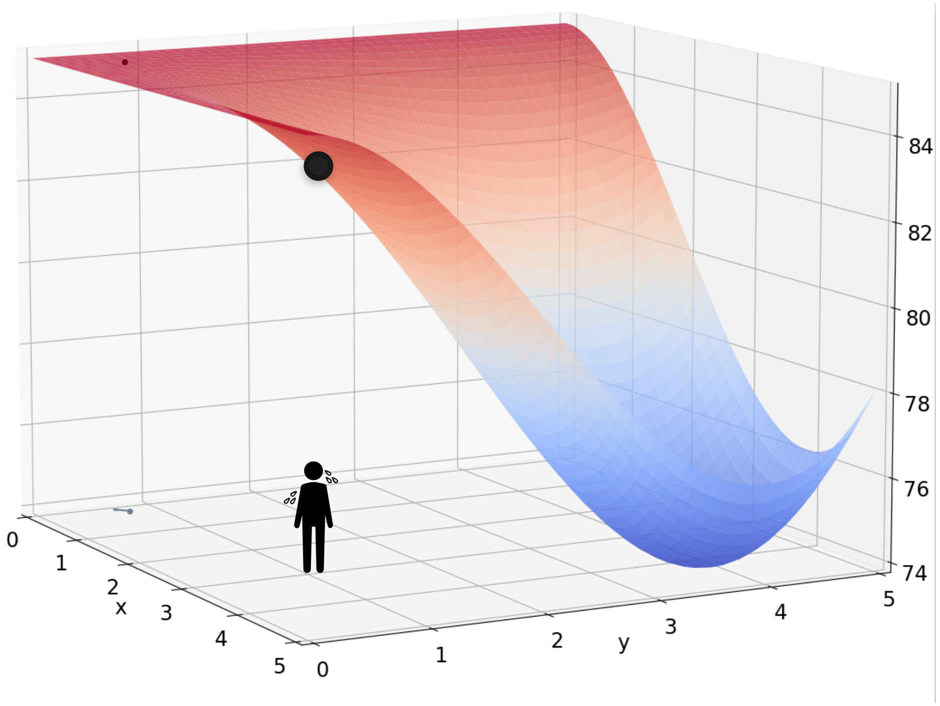
## Optimization using Gradient Descent in two variables - Part 2

# Idea for Gradient Descent



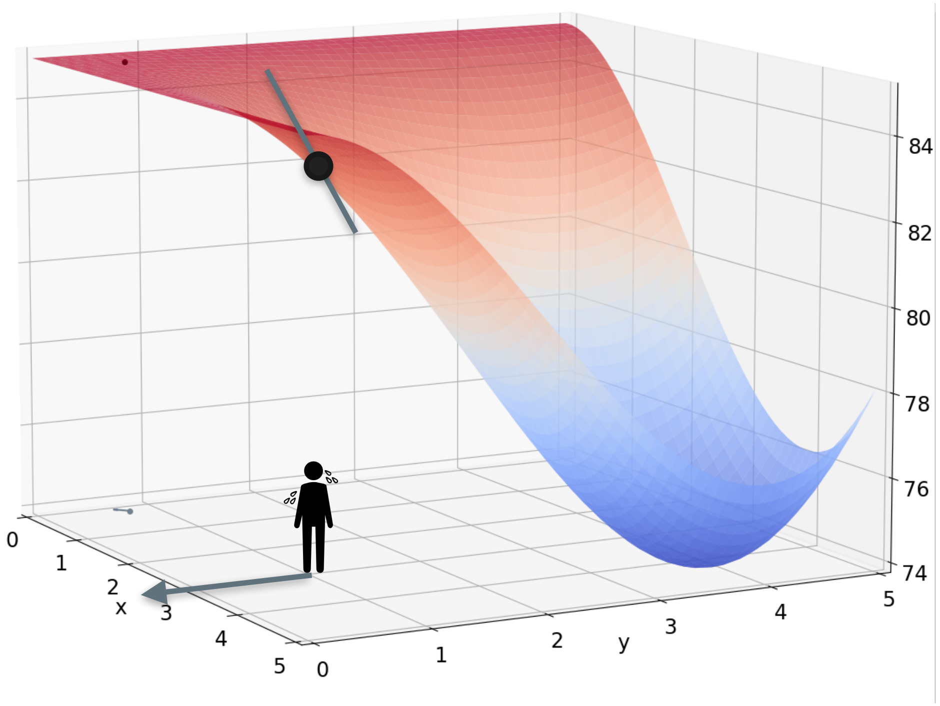
# Idea for Gradient Descent

Initial position: $(x_0, y_0)$



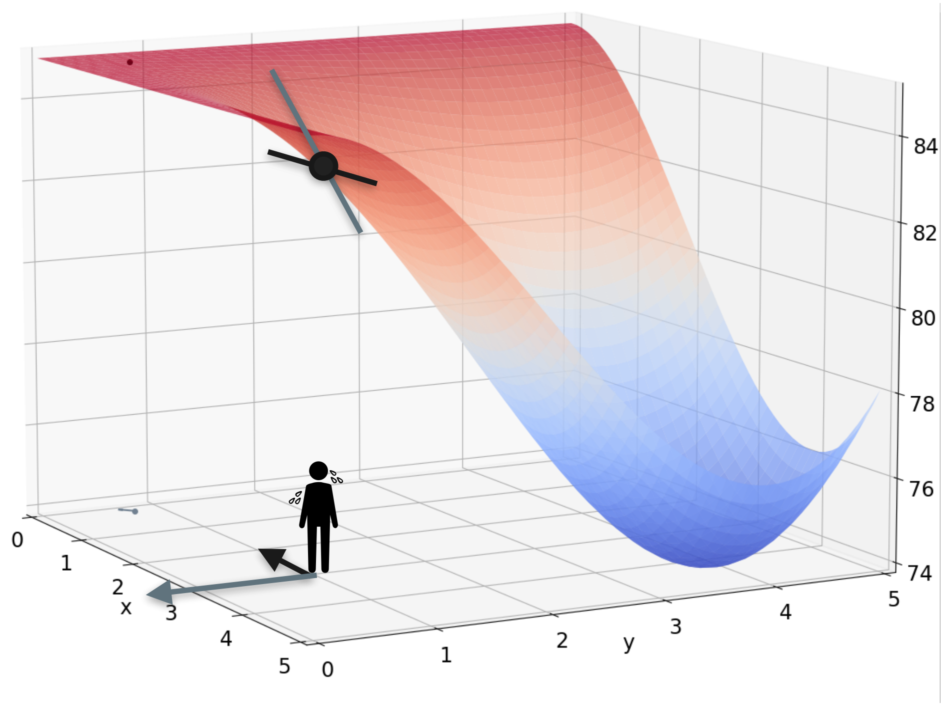
# Idea for Gradient Descent

Initial position: $(x_0, y_0)$



# Idea for Gradient Descent

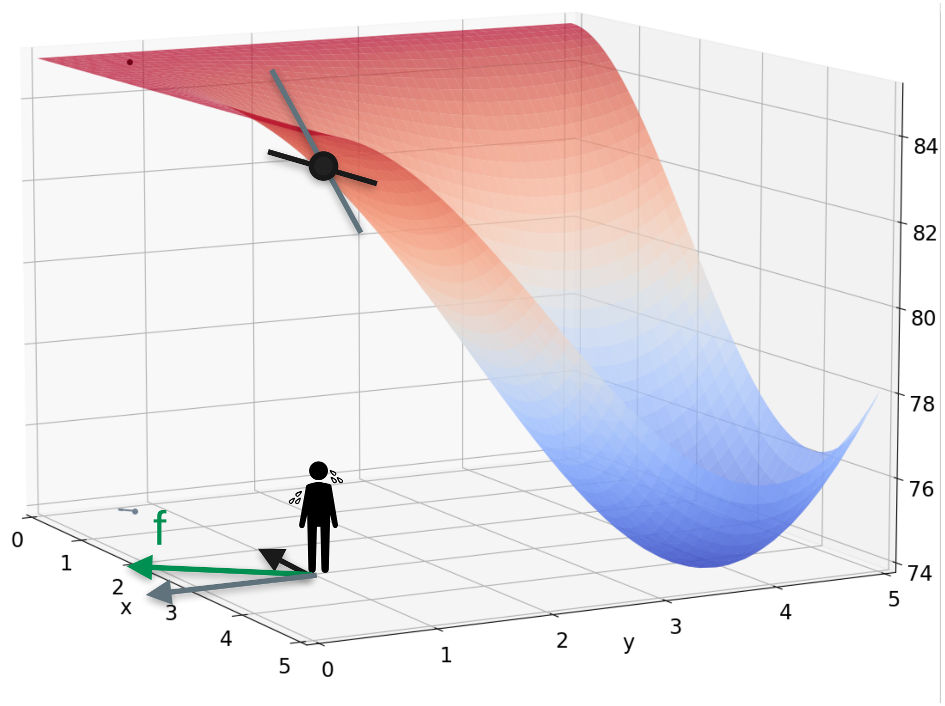
Initial position: $(x_0, y_0)$



# Idea for Gradient Descent

Initial position: $(x_0, y_0)$

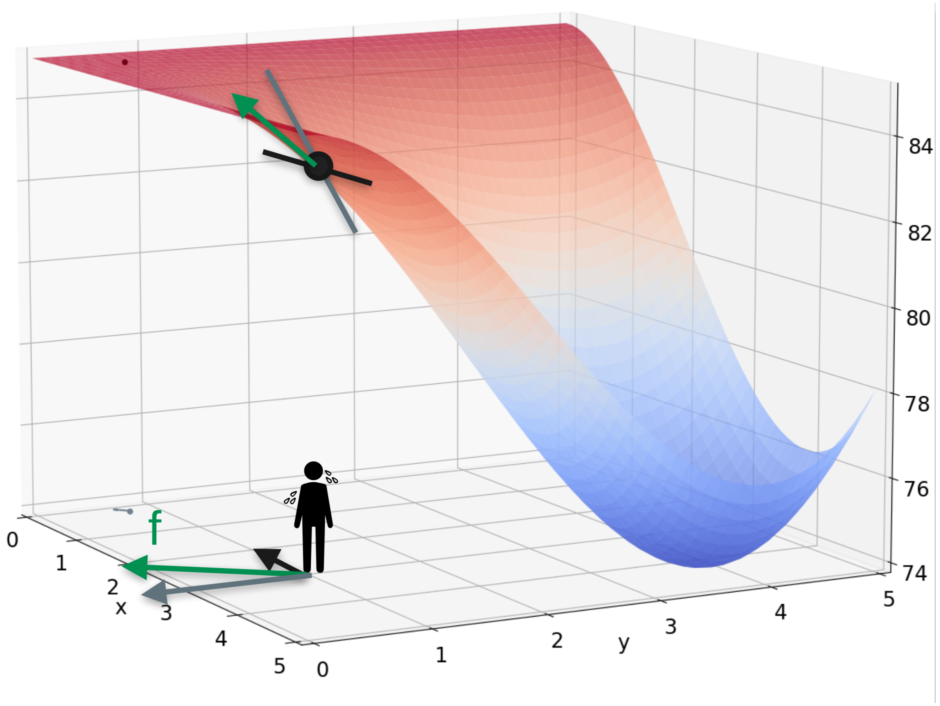
Direction of greatest ascent: f



# Idea for Gradient Descent

Initial position: $(x_0, y_0)$

Direction of greatest ascent: $\nabla f$

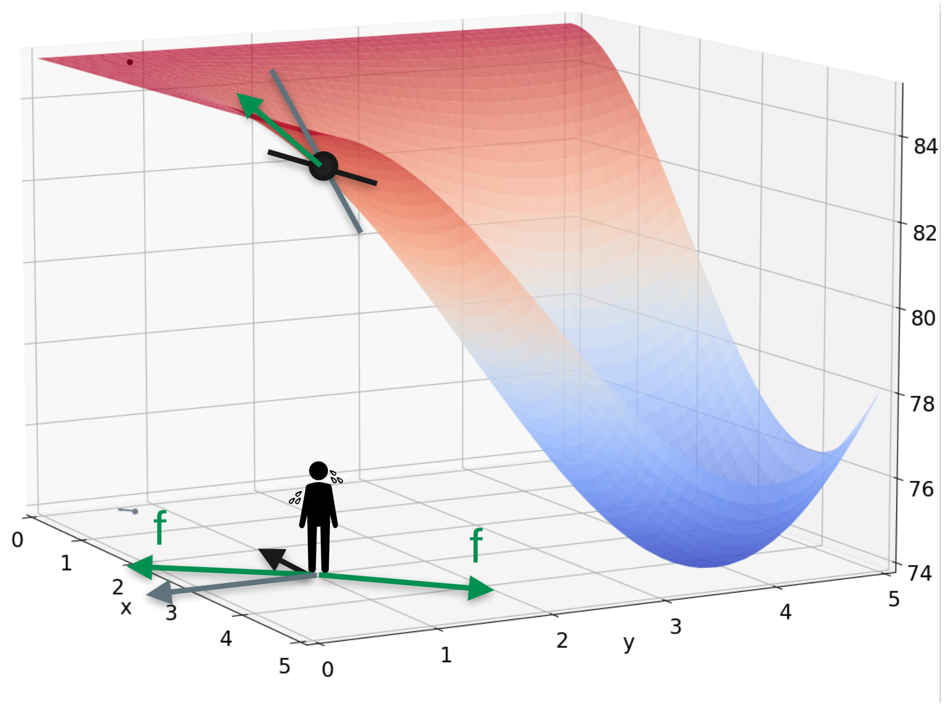


# Idea for Gradient Descent

Initial position: $(x_0, y_0)$

Direction of greatest ascent: f

Direction of greatest descent: f

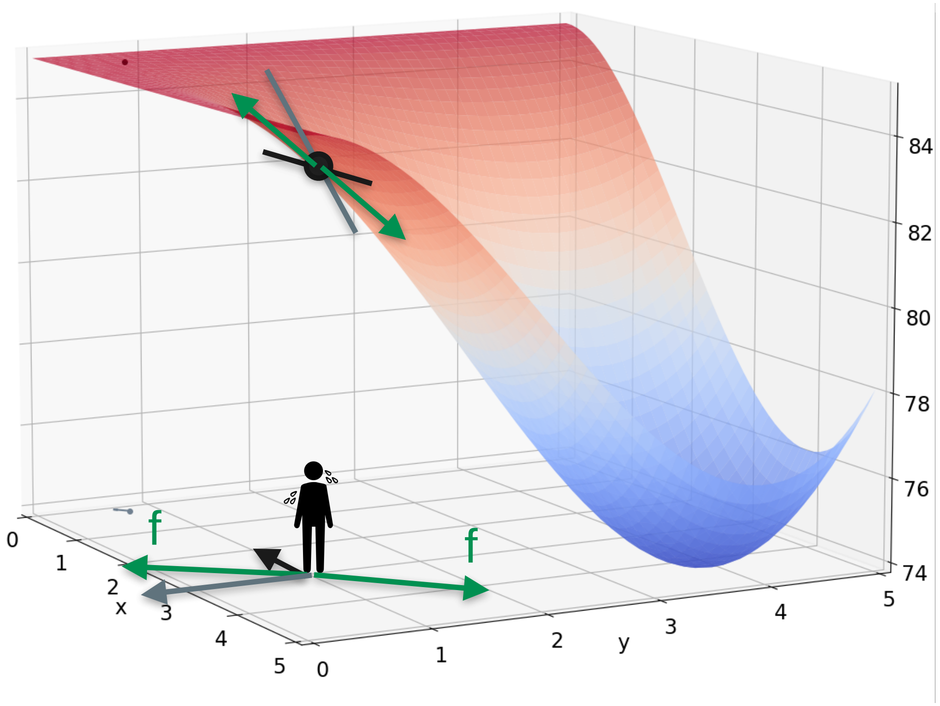


# Idea for Gradient Descent

Initial position: $(x_0, y_0)$

Direction of greatest ascent: f

Direction of greatest descent: f

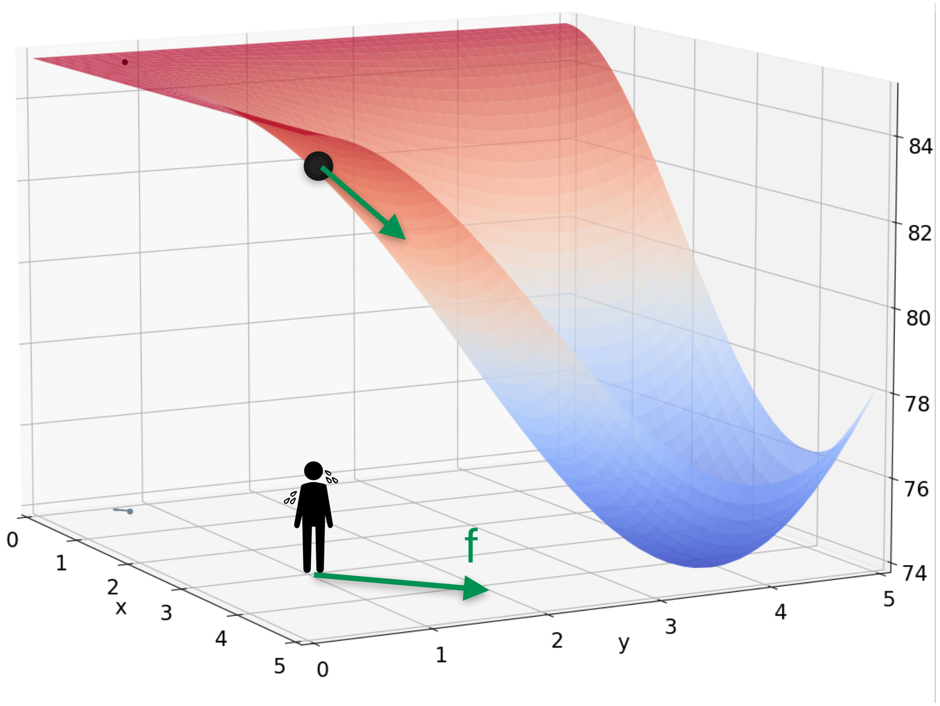


# Idea for Gradient Descent

Initial position: $(x_0, y_0)$

Direction of greatest ascent: f

Direction of greatest descent: f



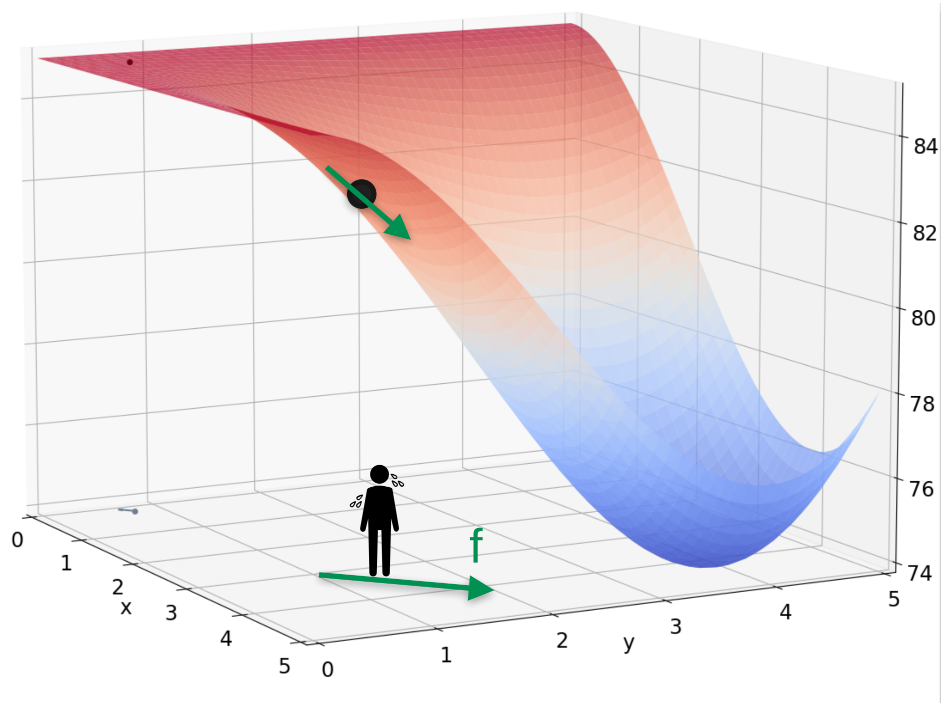
# Idea for Gradient Descent

Initial position: $(x_0, y_0)$

Direction of greatest ascent: f

Direction of greatest descent: f

Updated position: $(x_0, y_0)$ f



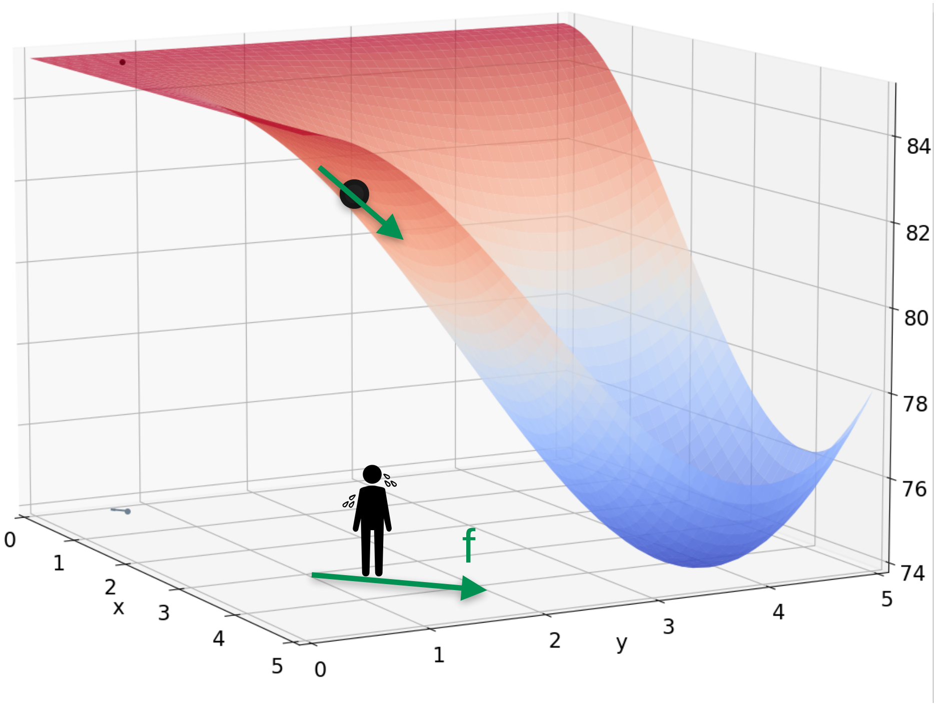
# Idea for Gradient Descent

Initial position: $(x_0, y_0)$

Direction of greatest ascent: $f$

Direction of greatest descent: $f$

Updated position: $\underbrace{(x_0, y_0) \qquad f}_{(x_1, y_1)}$



# Idea for Gradient Descent

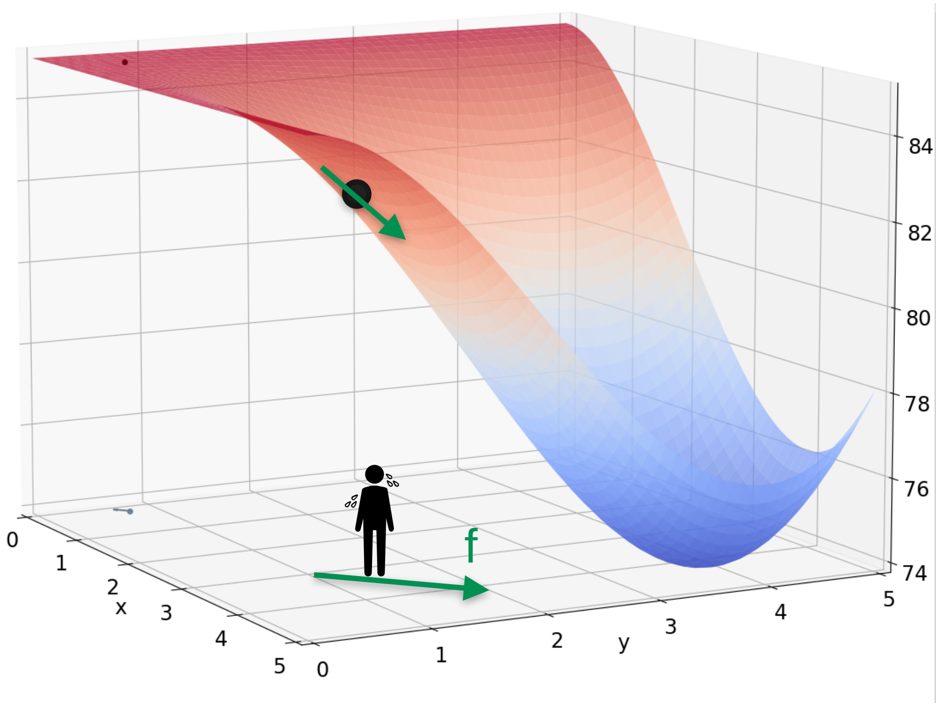
Initial position: $(x_0, y_0)$

Direction of greatest ascent: $f$

Direction of greatest descent: $f$

Updated position: $\underbrace{(x_0, y_0) \qquad f}_{(x_1, y_1)}$

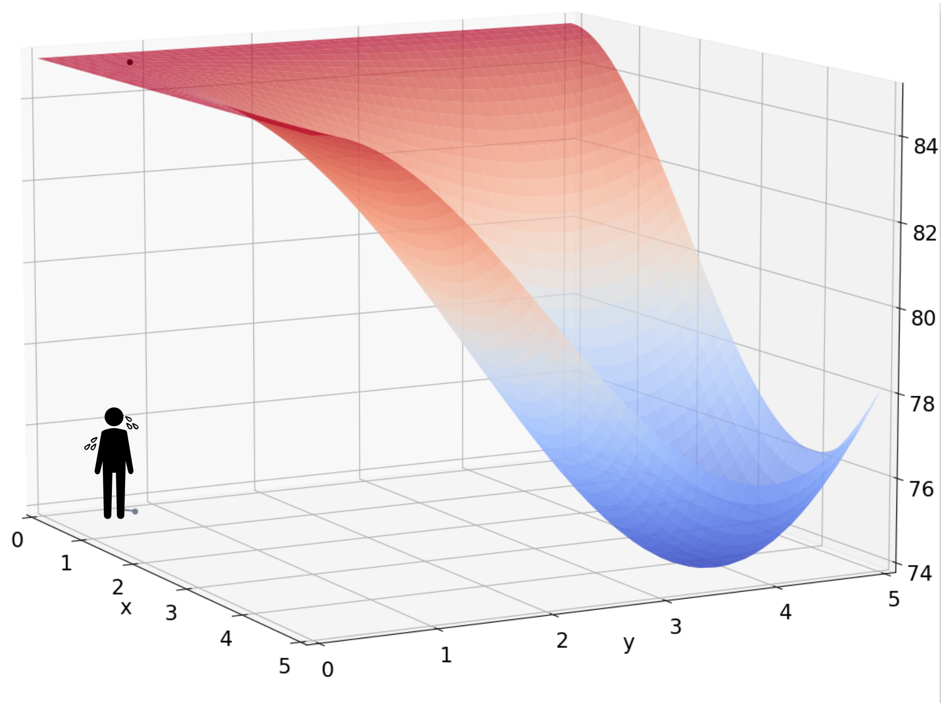
Better point!



# Method 2: Gradient Descent

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

Start: $x = 0.5$ $y = 0.6$

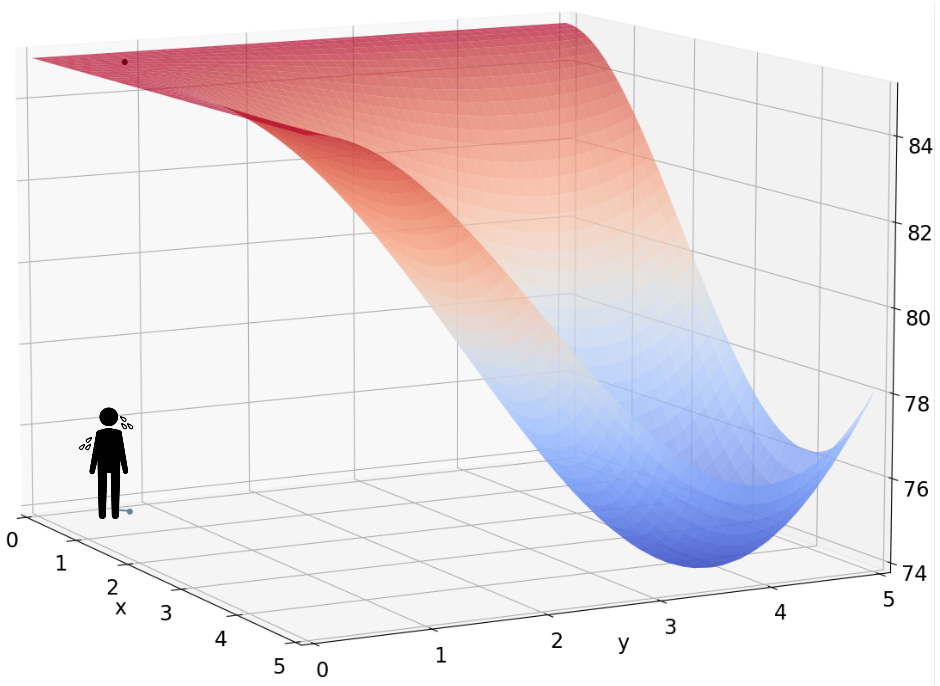


# Method 2: Gradient Descent

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

Start: $x = 0.5$ $y = 0.6$

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



# Method 2: Gradient Descent

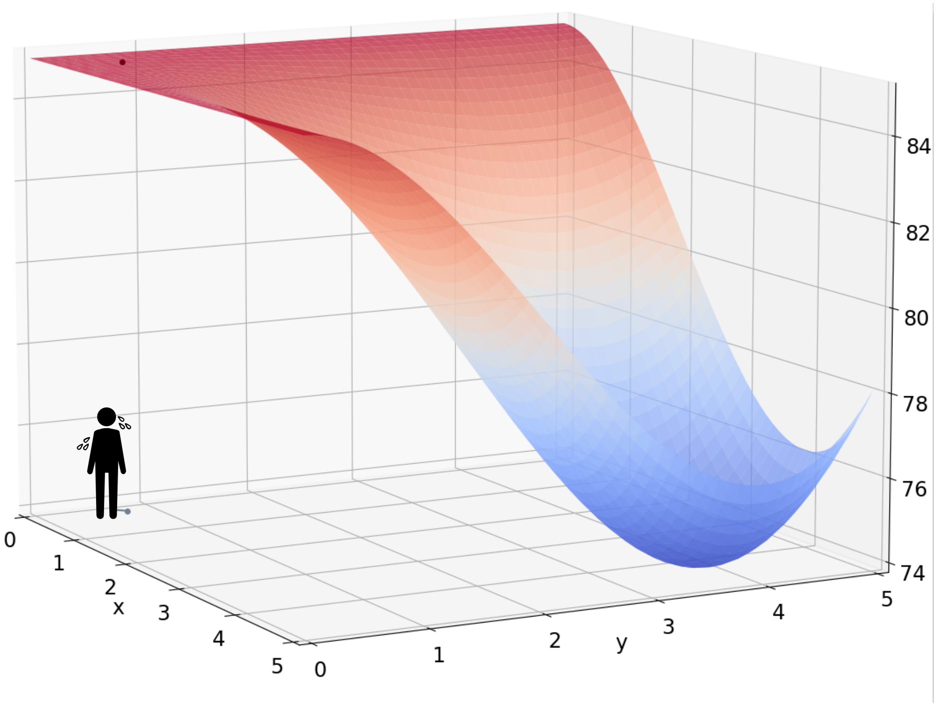
$$T = f(x, y) = 85 \quad \frac{1}{90}x^2(x \quad 6)y^2(y \quad 6)$$

Start: $x = 0.5 \quad y = 0.6$

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

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

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



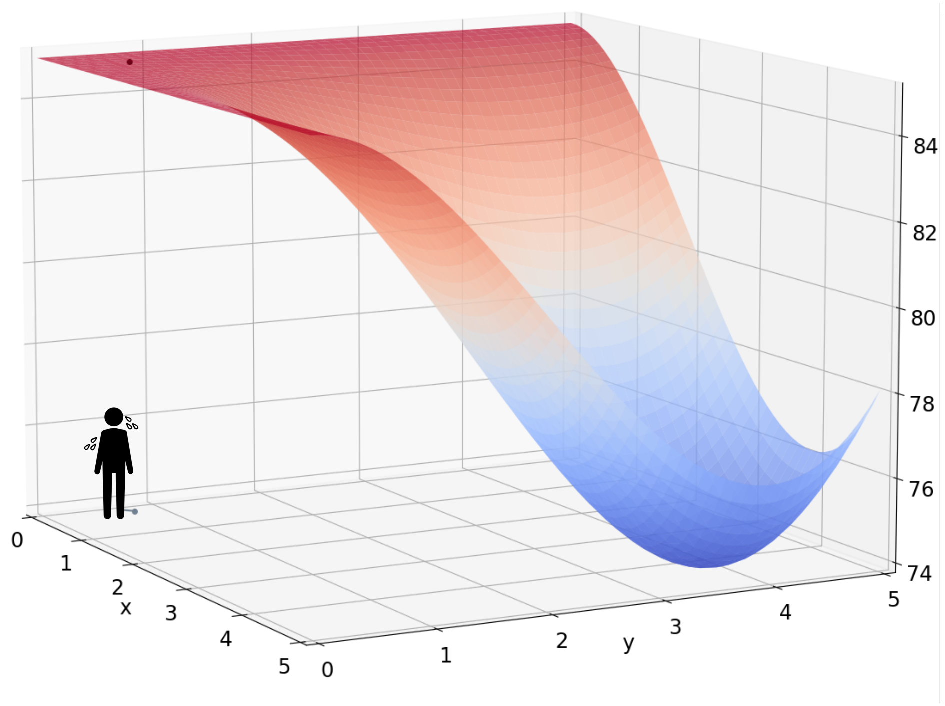
# Method 2: Gradient Descent

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

Start: $x = 0.5 \;\; y = 0.6$

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

$$f = \begin{bmatrix} \frac{1}{90}x(3x \quad 12)y^2(y \quad 6) \\ \frac{1}{90}x^2(x \quad 6)y(3y \quad 12) \end{bmatrix}$$



# Method 2: Gradient Descent

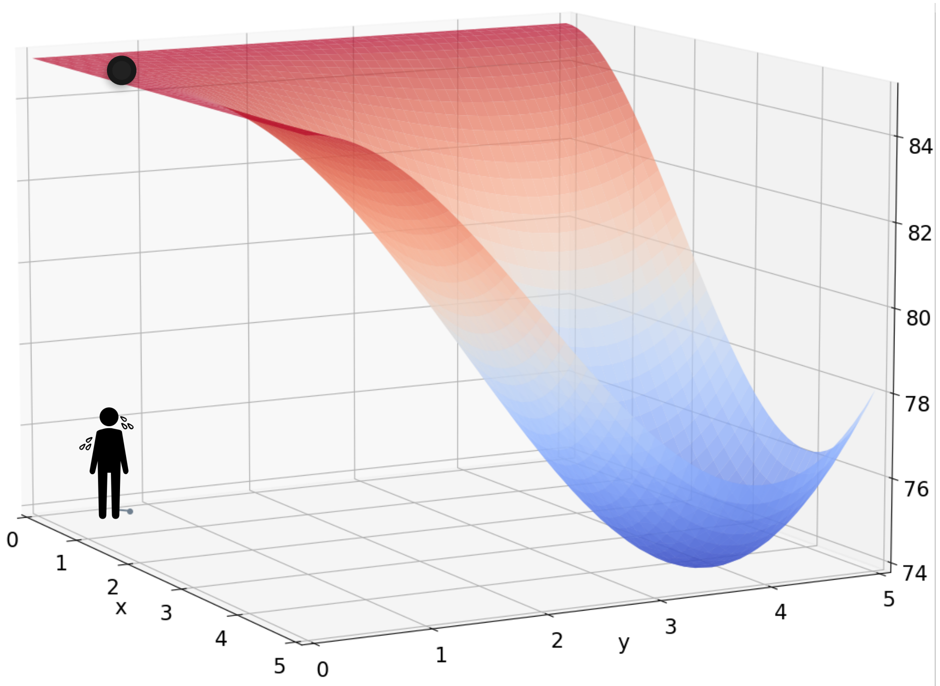
$T = f(x, y) = 85 \quad \frac{1}{90} x^2 (x \quad 6) y^2 (y \quad 6)$

Start: $x = 0.5 \quad y = 0.6$

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

$$f = \begin{bmatrix} \frac{1}{90} x(3x \quad 12) y^2 (y \quad 6) \\ \frac{1}{90} x^2 (x \quad 6) y (3y \quad 12) \end{bmatrix}$$

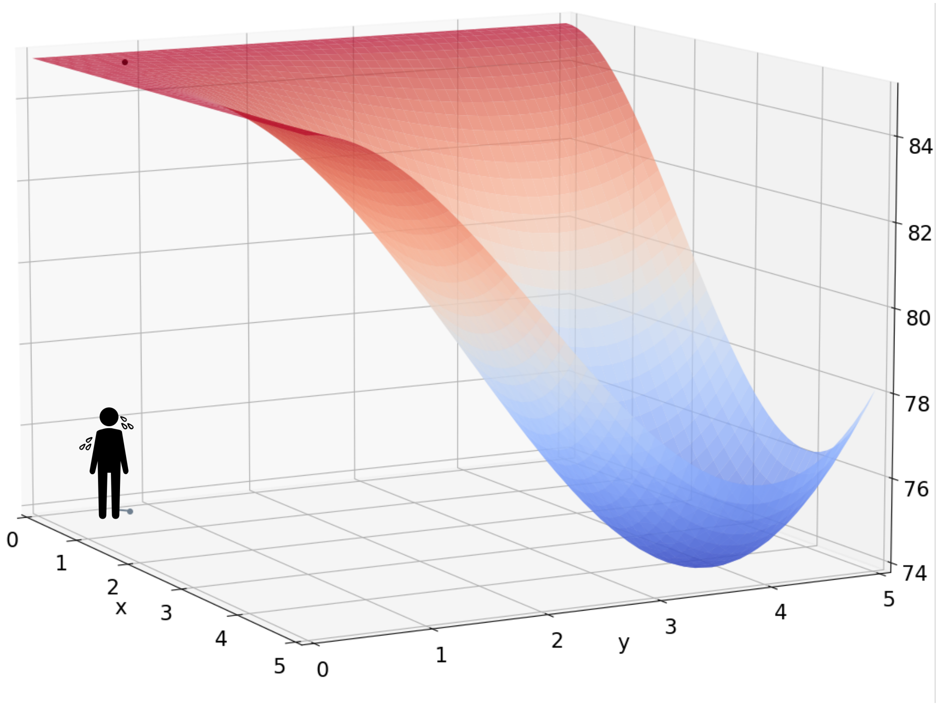
$$f(0.5, 0.6) = \begin{bmatrix} 0.1134 \\ 0.0935 \end{bmatrix}$$



# Method 2: Gradient Descent

Start: $x = 0.5$ $y = 0.6$

$$f(0.5, 0.6) = \begin{bmatrix} 0.1134 \\ 0.0935 \end{bmatrix}$$

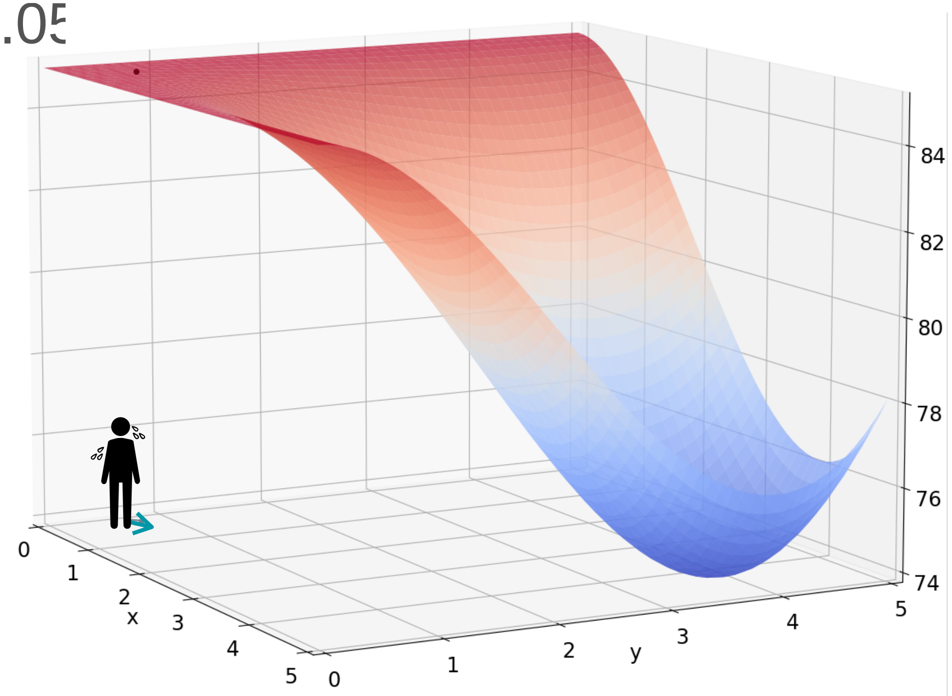


# Method 2: Gradient Descent

Start: x = 0.5 y = 0.6 Rate: = 0.05

f(0.5,0.6) = [ 0.1134 0.0935 ]

Move by
0.05 f(0.5,0.6)



# Method 2: Gradient Descent

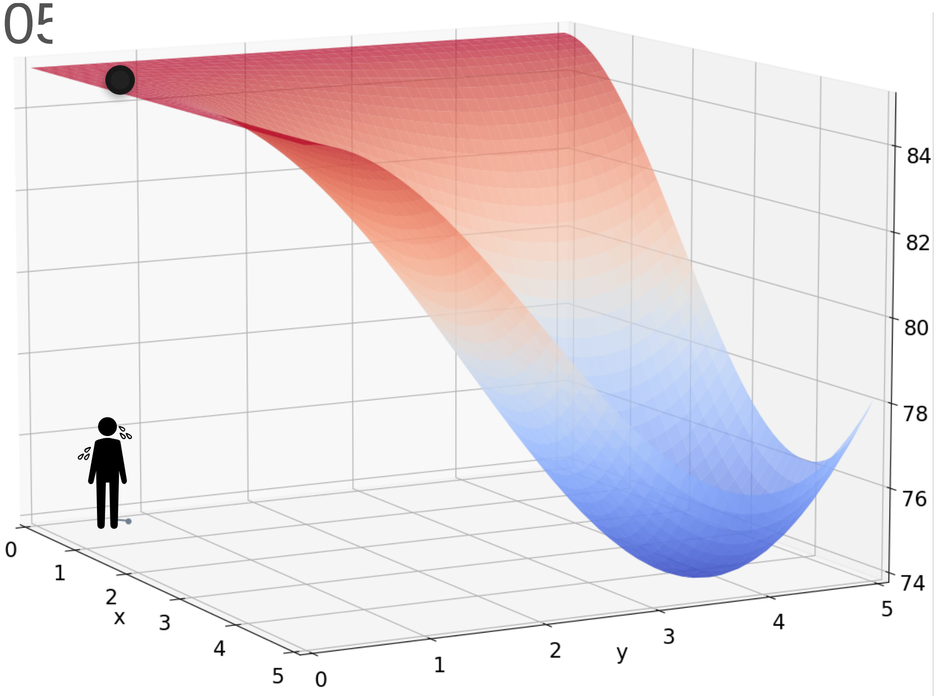
Start: x = 0.5 y = 0.6 Rate: = 0.05

f(0.5,0.6) = [ 0.1134 0.0935]

Move by

0.05 f(0.5,0.6)

x 0.5057
y 0.6047



# Method 2: Gradient Descent

Start: x = 0.5 y = 0.6 Rate: = 0.05

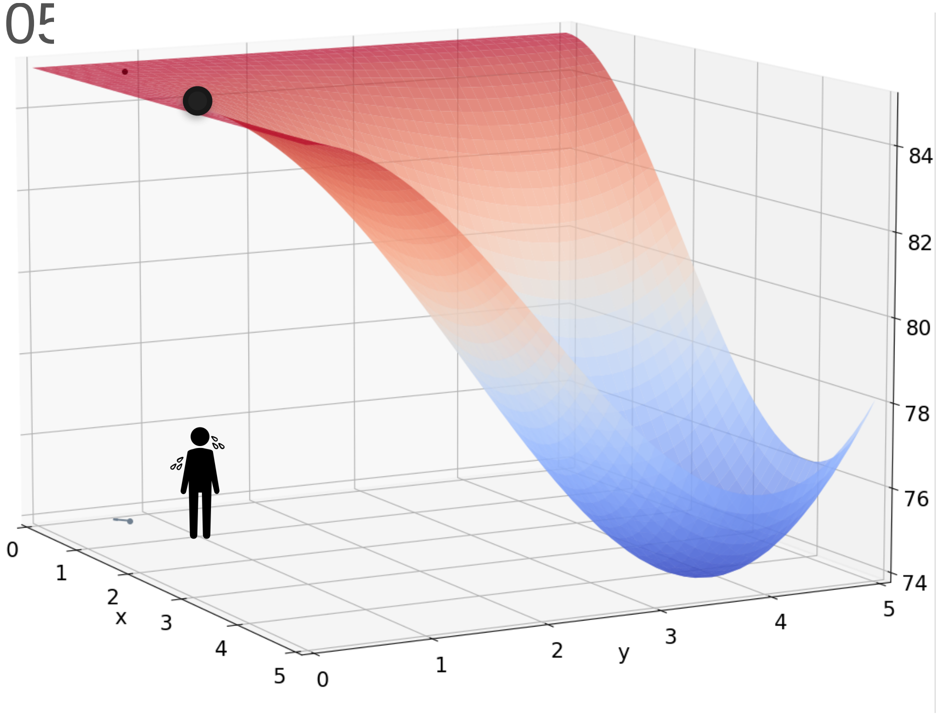
f(0.5,0.6) = [ 0.1134 0.0935 ]

Move by

0.05 f(0.5,0.6)

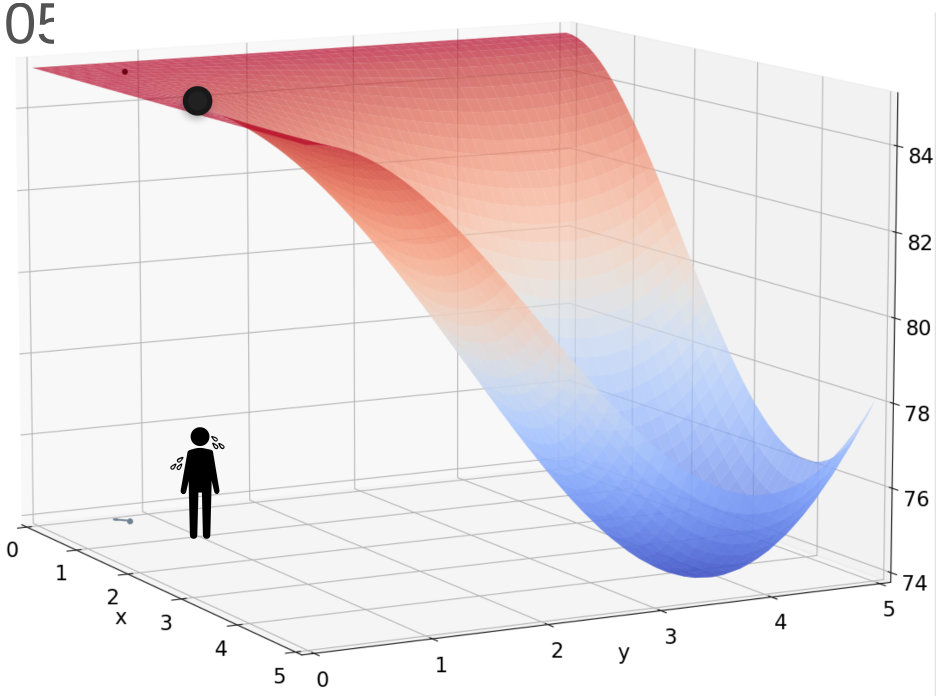
x 0.5057

y 0.6047



# Method 2

Start: x = 0.5 y = 0.6 Rate: = 0.05



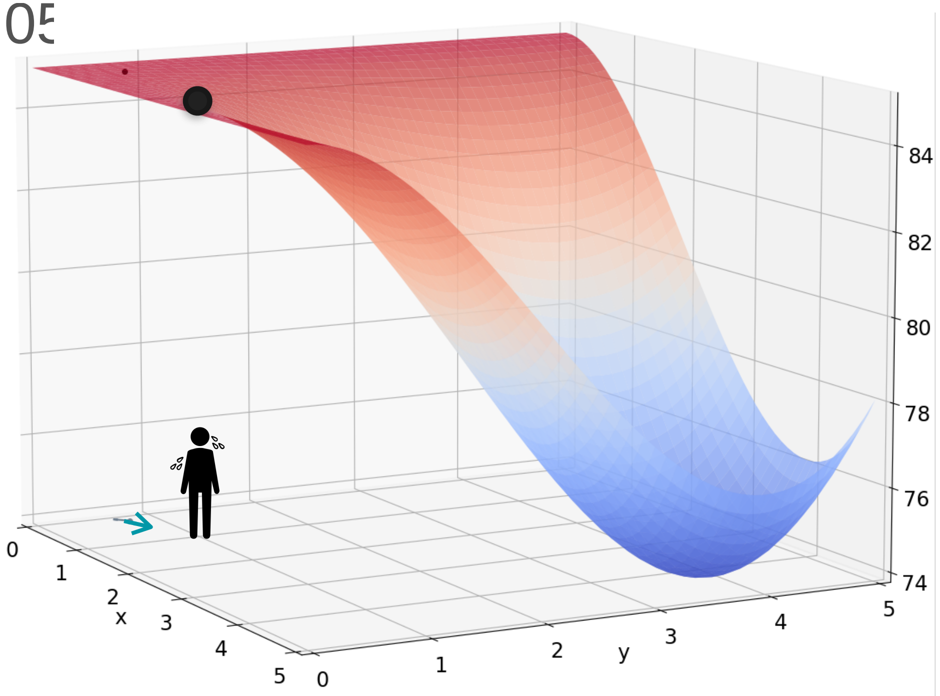
# Method 2

Start: x = 0.5 y = 0.6 Rate: = 0.05

Find:

f(0.5057,0.6047) = 0.1162
0.0961]

Move by
0.05 f(0.5057,0.604)



# Method 2

Start: x = 0.5 y = 0.6 Rate: = 0.05

Find:

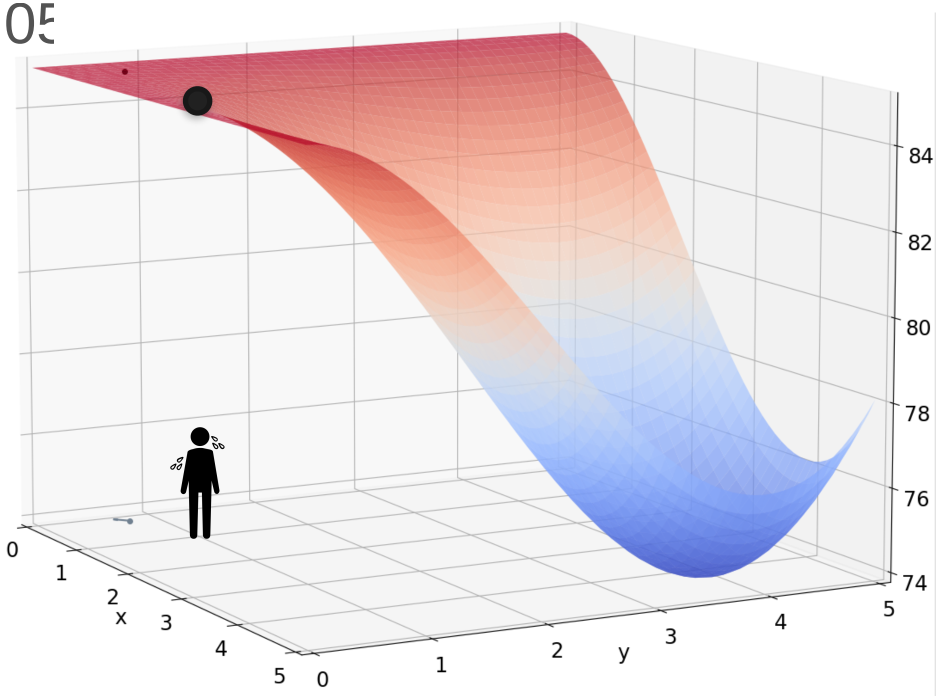
f(0.5057,0.6047) = 0.1162 0.0961]

Move by

0.05 f(0.5057,0.604)

x 0.5115
y 0.6095

Repeat!



# Gradient Descent

Start: x = 0.5 y = 0.6 Rate: = 0.05

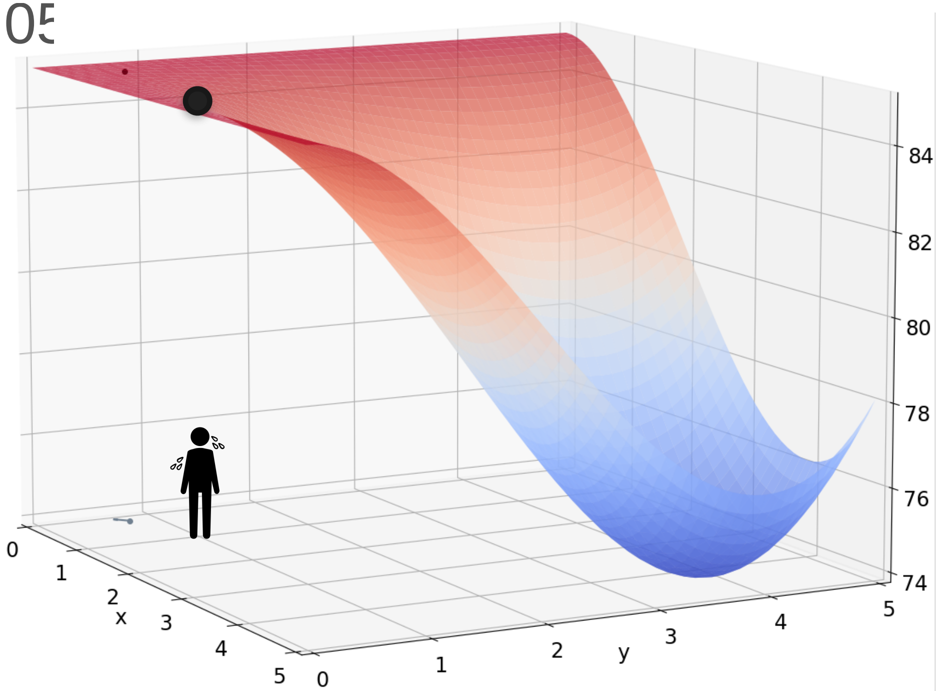
Find:

f(0.5057,0.6047) = 0.1162
0.0961]

Move by
0.05 f(0.5057,0.604)

x 0.5115
y 0.6095

Repeat!



# Gradient Descent

Start: x = 0.5 y = 0.6 Rate: = 0.05

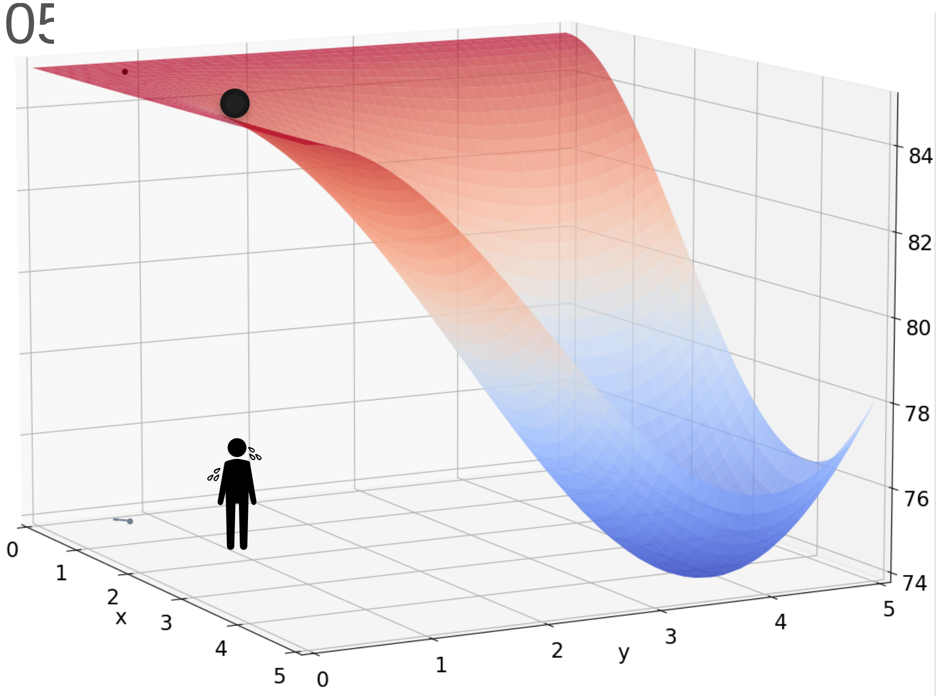
Find:

f(0.5057,0.6047) = [0.1162
0.0961]

Move by
0.05 f(0.5057,0.604)

x 0.5115
y 0.6095

Repeat!



# Gradient Descent

Start: x = 0.5 y = 0.6 Rate: = 0.05

Find:

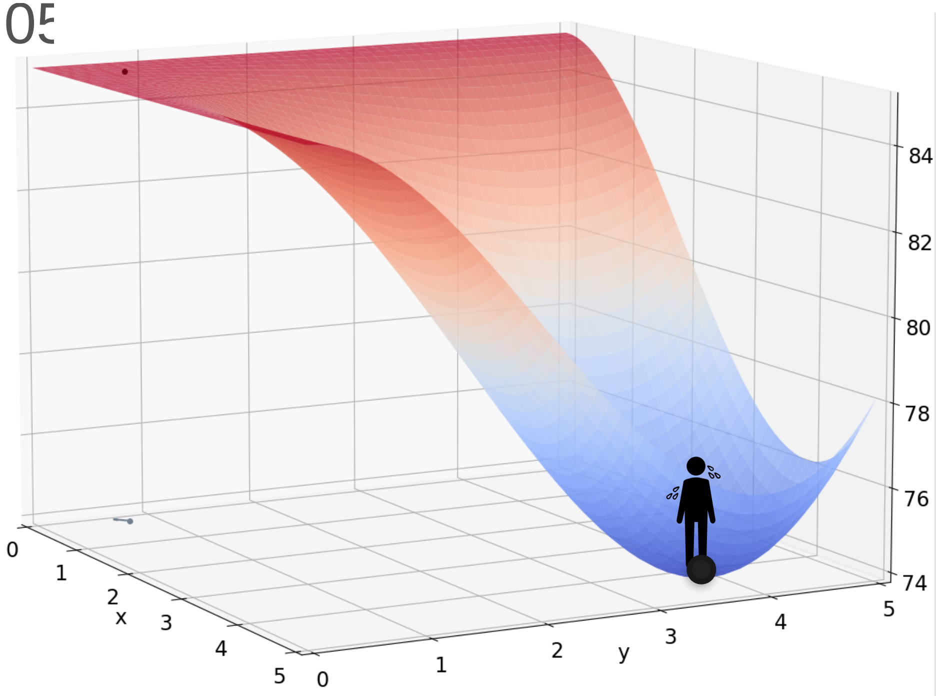
f(0.5057,0.6047) = [0.1162 0.0961]

Move by

0.05 f(0.5057,0.604)

x 0.5115
y 0.6095

Repeat!



# Gradient Descent

# Gradient Descent

Function: $f(x, y)$

# Gradient Descent

Function: $f(x, y)$ Goal: find minimum of $f(x, y)$

# Gradient Descent

Function: $f(x, y)$ Goal: find minimum of $f(x, y)$

Step 1:

Define a learning rate

Choose a starting point $(x_0, y_0)$

# Gradient Descent

Function: $f(x, y)$ Goal: find minimum of $f(x, y)$

Step 1:

Define a learning rate

Choose a starting point $(x_0, y_0)$

Step 2:

Update: $\begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} x_{k\ 1} \\ y_{k\ 1} \end{bmatrix} \qquad f(x_{k\ 1}, y_{k\ 1})$

# Gradient Descent

Function: $f(x, y)$ Goal: find minimum of $f(x, y)$

Step 1:

Define a learning rate

Choose a starting point $(x_0, y_0)$

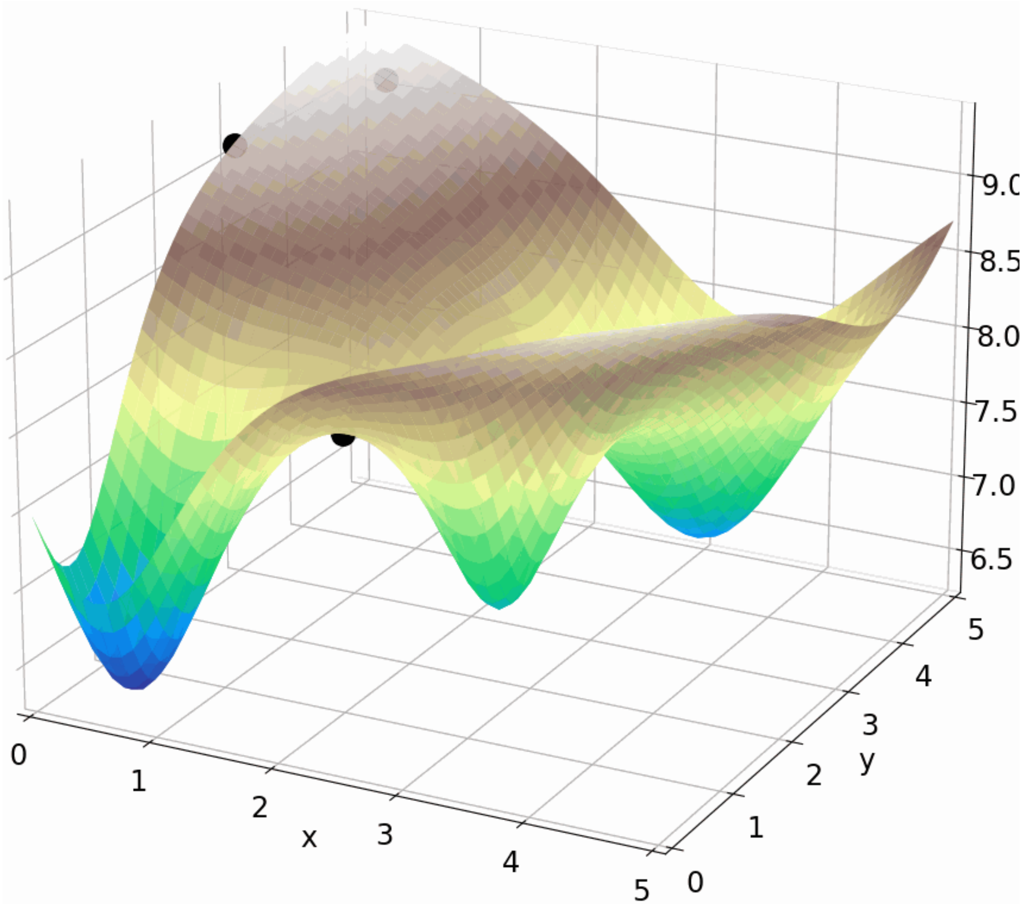
Step 2:

Update: $\begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} x_{k\ 1} \\ y_{k\ 1} \end{bmatrix} \qquad f(x_{k\ 1}, y_{k\ 1})$

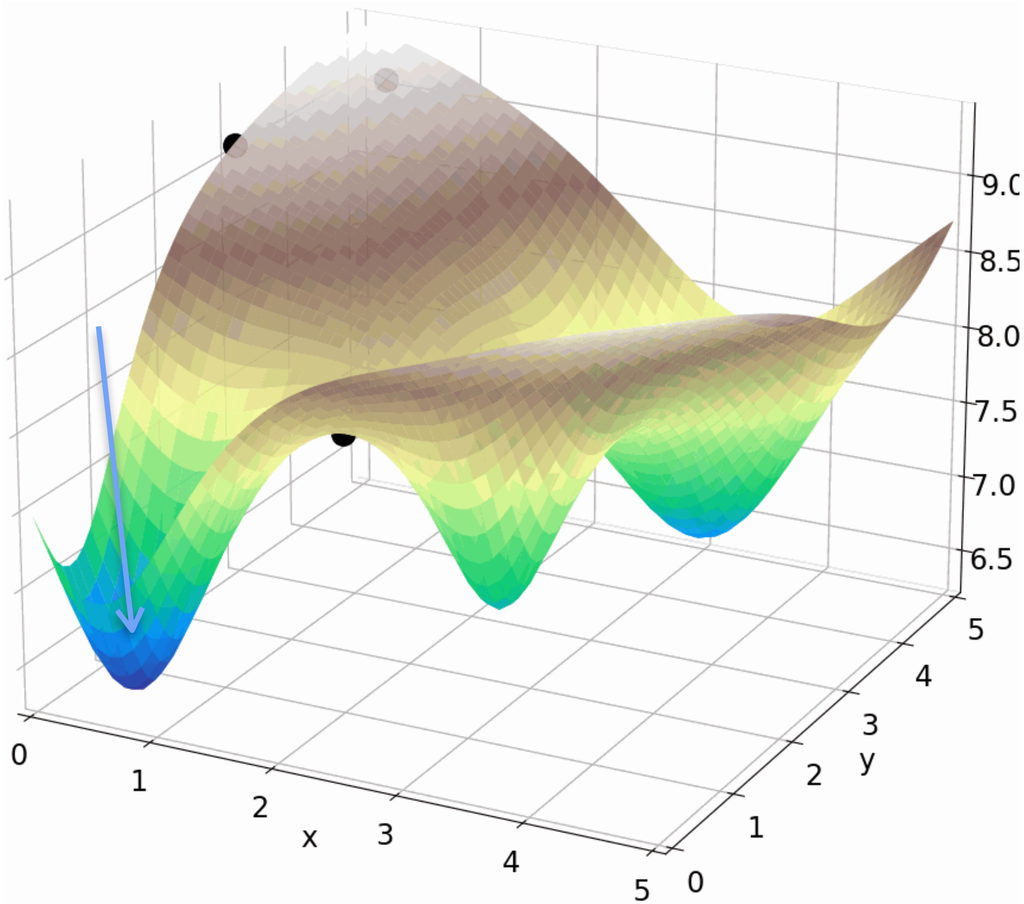
Step 3:

Repeat Step 2 until you are close enough to the true minimum $(x^*, y^*)$

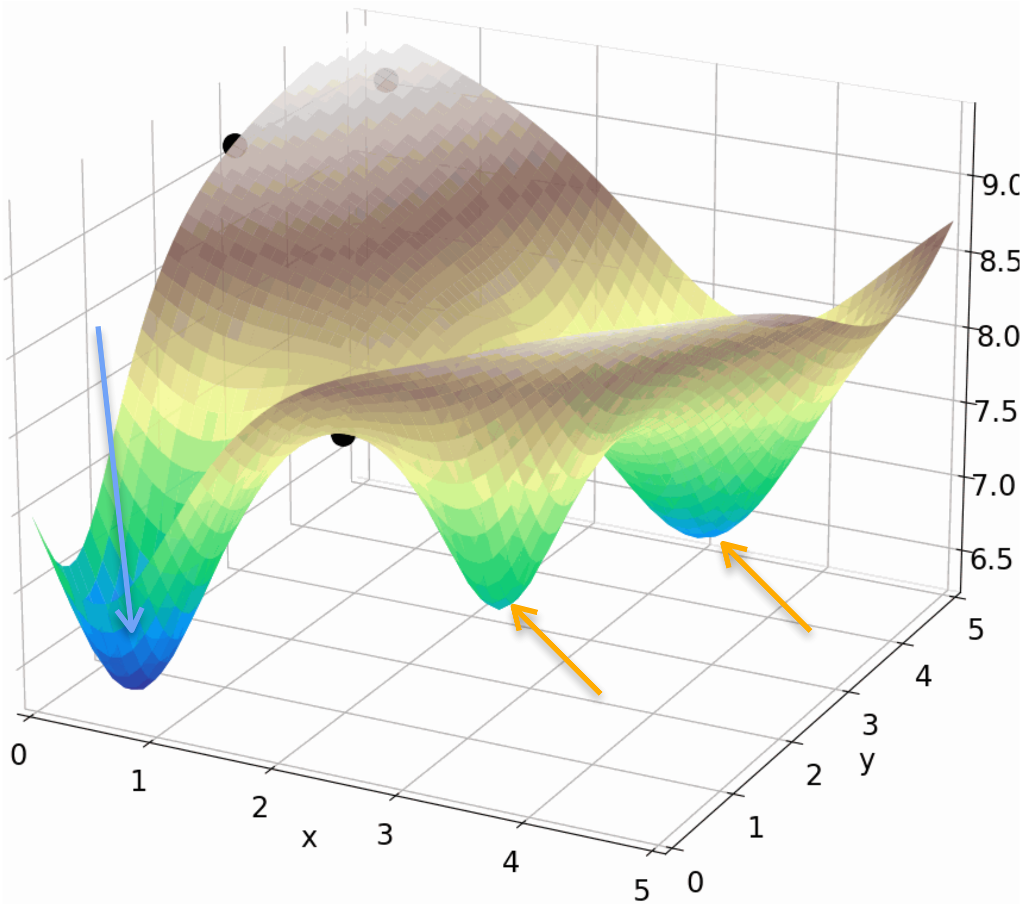
# Gradient Descent With Local Minima



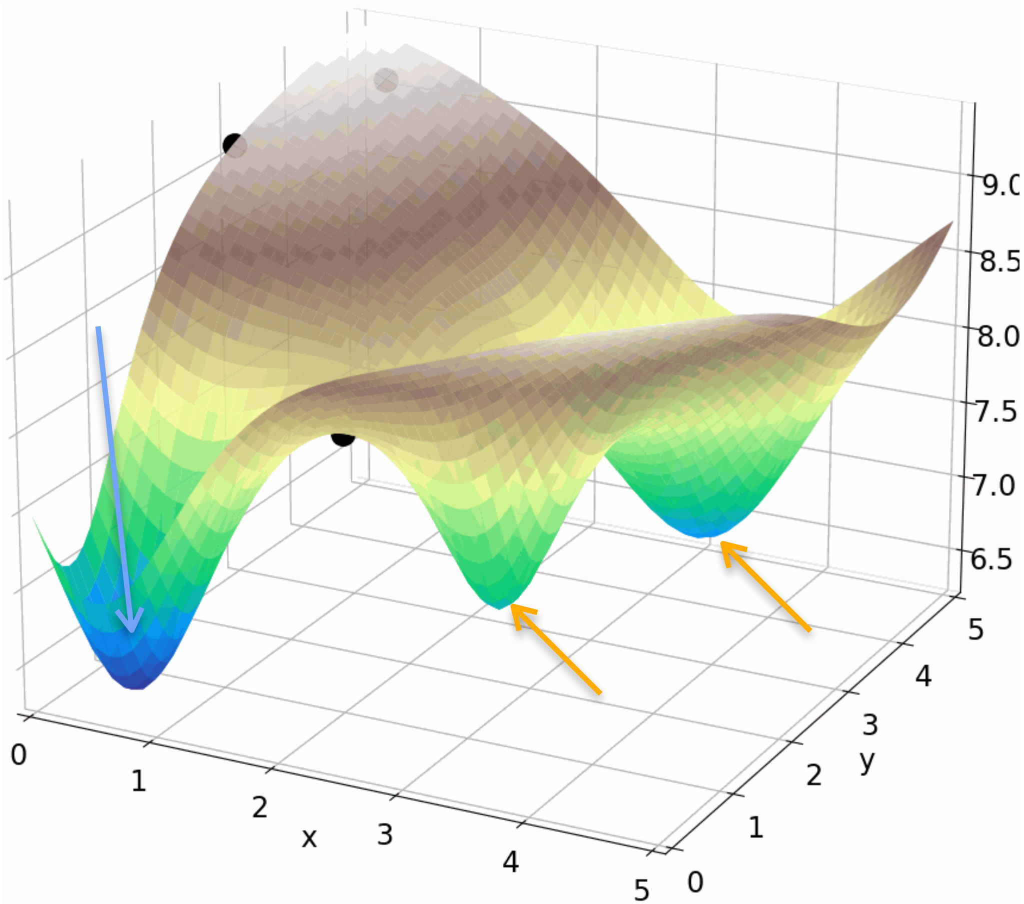
# Gradient Descent With Local Minima



# Gradient Descent With Local Minima



# Gradient Descent With Local Minima

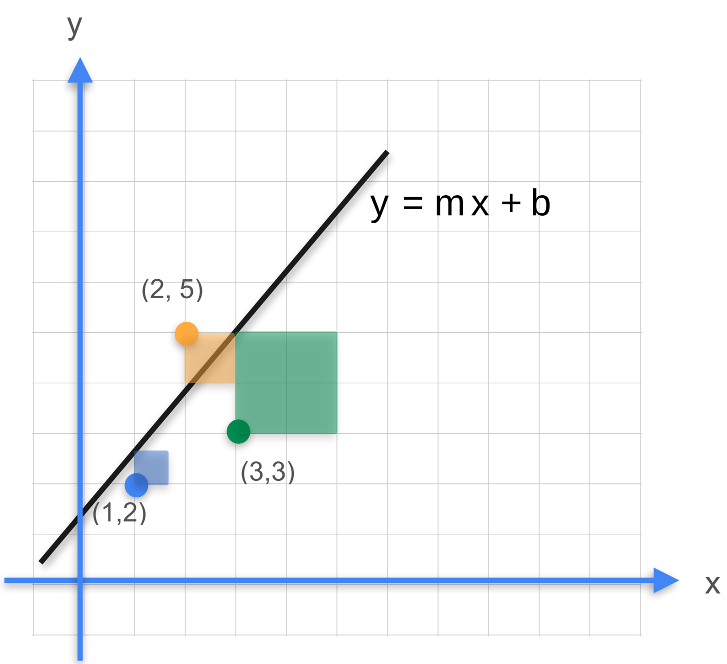


DeepLearning.AI

# Gradients and Gradient Descent

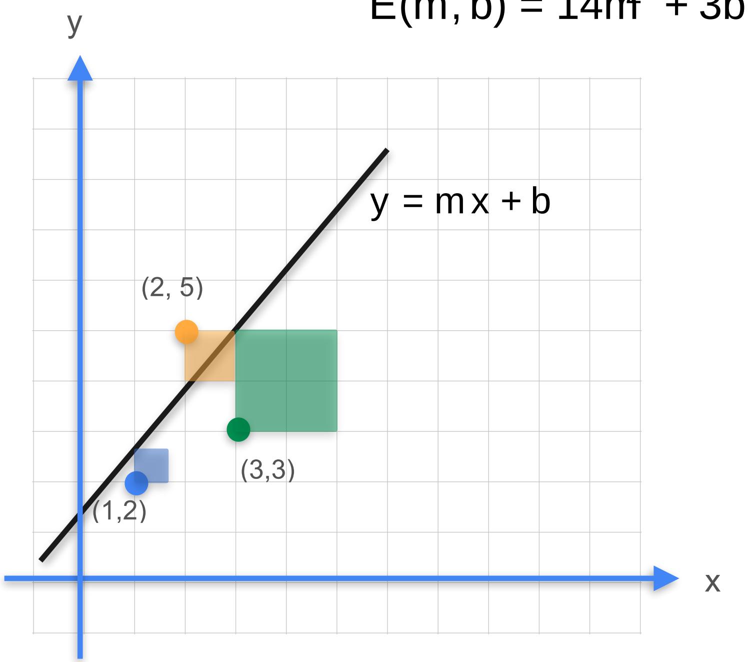
## Optimization using Gradient Descent - Least squares

# Gradient Descent With Power Lines Example



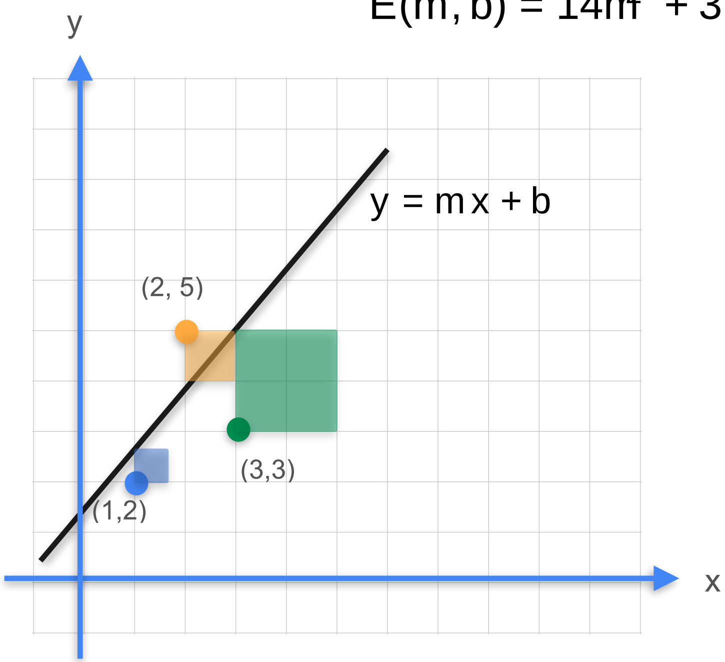
# Gradient Descent With Power Lines Example

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



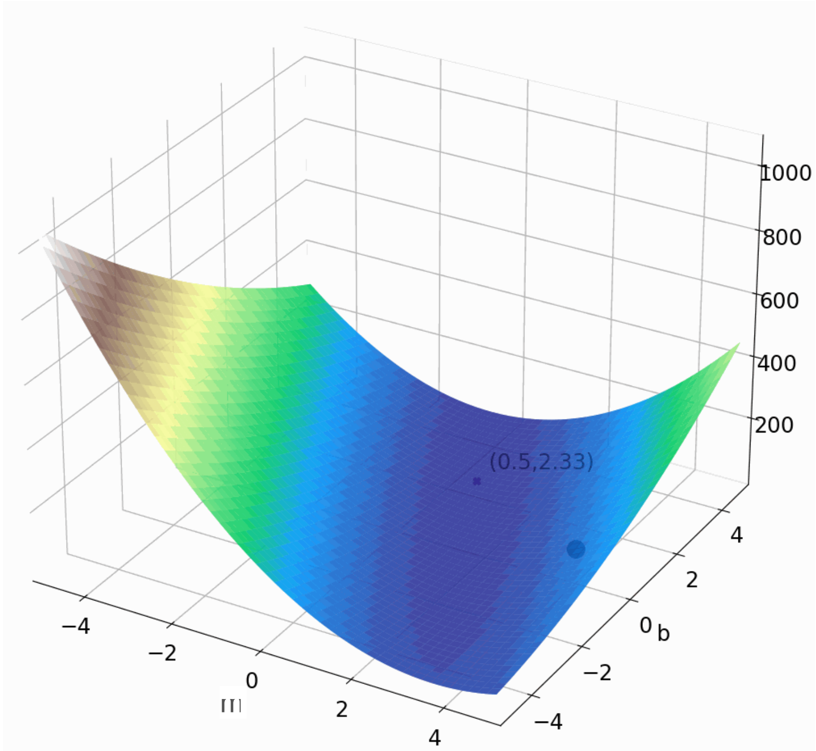
# Gradient Descent With Power Lines Example

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



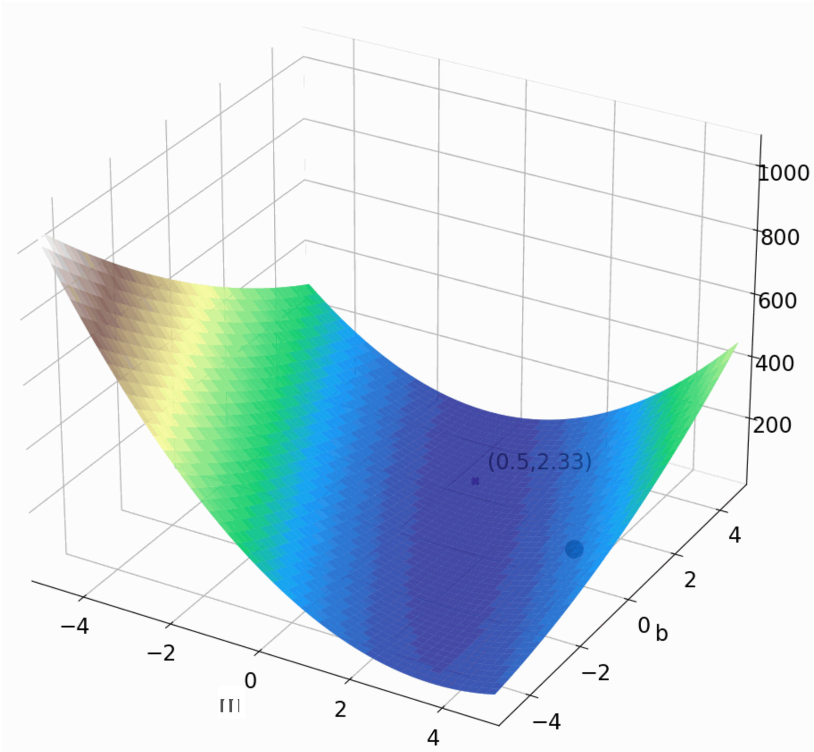

$$E(m = \frac{1}{2}, b = \frac{7}{3}) \quad 4.167$$

# Linear Regression: Gradient Descent



# Linear Regression: Gradient Descent

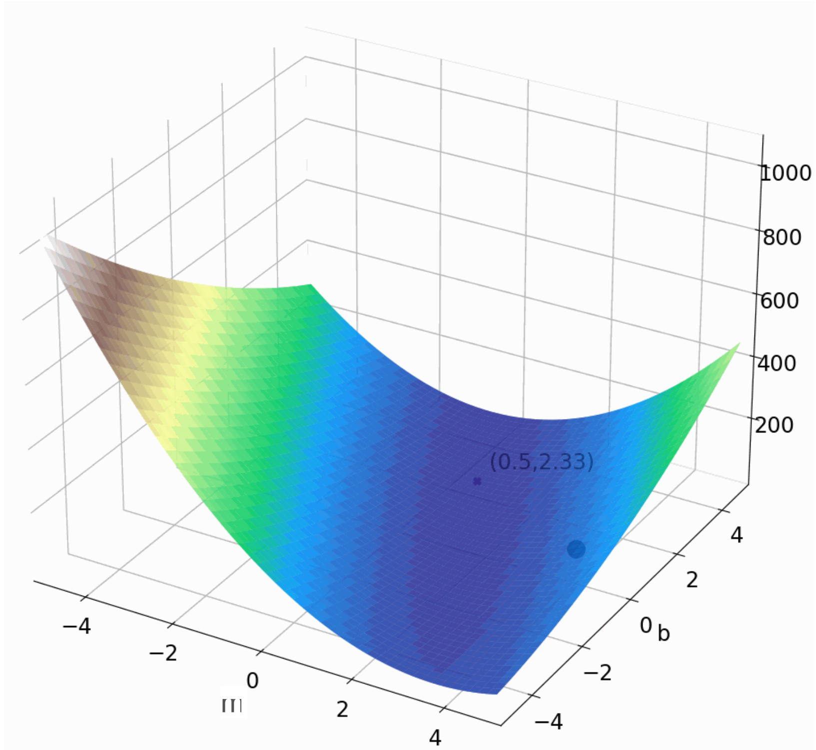
**Goal: Minimize sum of squares cost**



# Linear Regression: Gradient Descent

**Goal: Minimize sum of squares cost**

$$E = [28m + 12b \quad 42, \;\; 6b + 12m \quad 20]$$



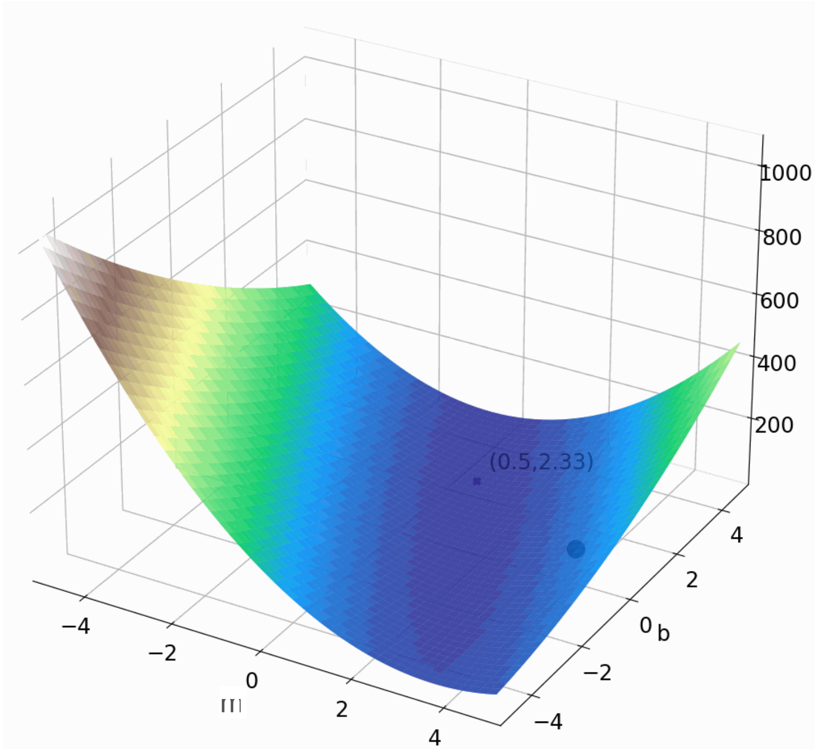
# Linear Regression: Gradient Descent

**Goal: Minimize sum of squares cost**

$$E = [28m + 12b \quad 42, \;\; 6b + 12m \quad 20]$$

m =

b =



# Linear Regression: Gradient Descent

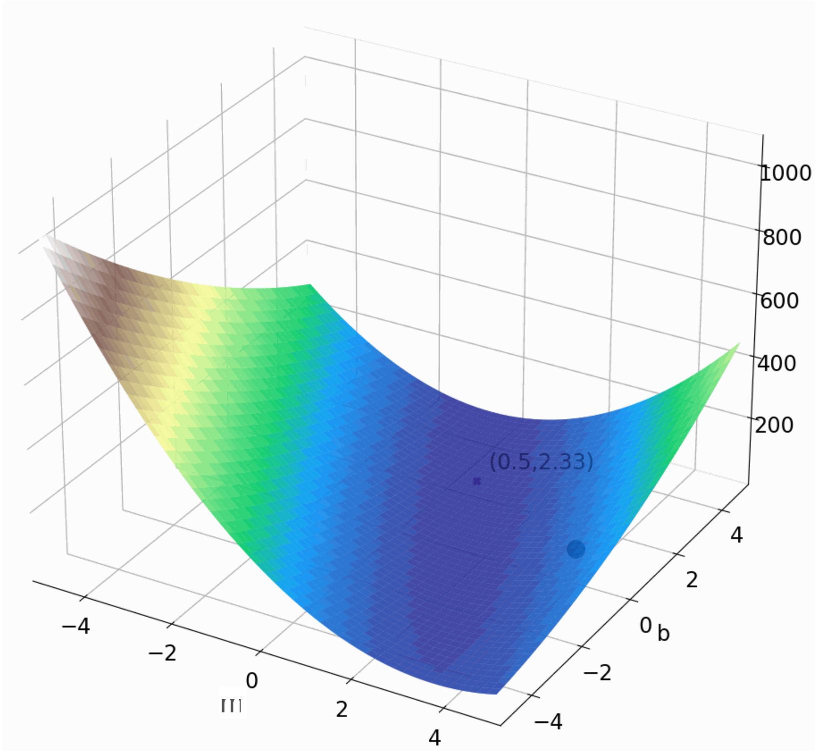
**Goal: Minimize sum of squares cost**

$E = [28m + 12b \quad 42, \; 6b + 12m \quad 20]$

m =

b =

**?**



# Linear Regression: Gradient Descent

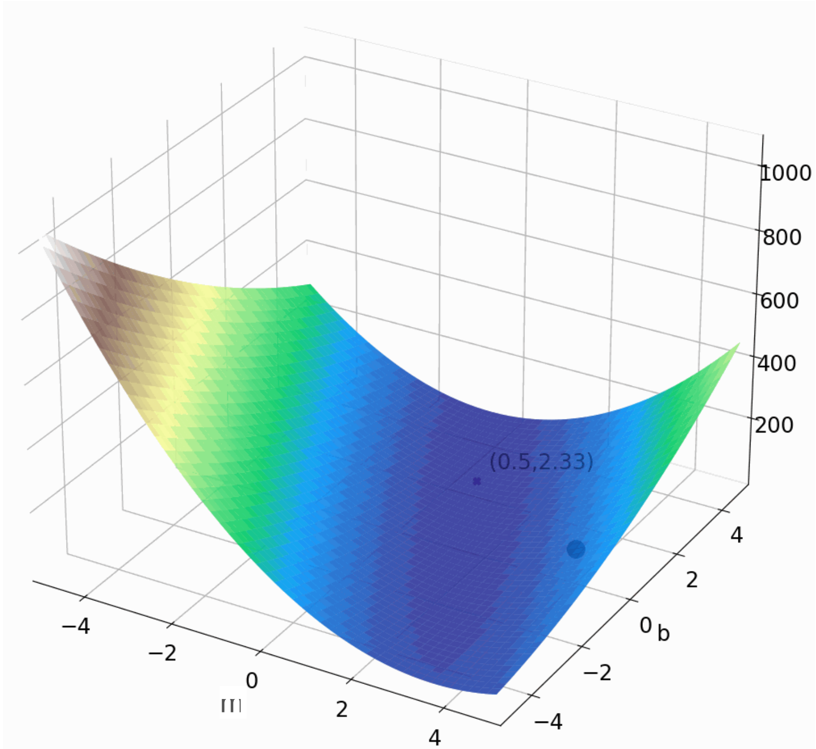
**Goal: Minimize sum of squares cost**

$$E = [28m + 12b \quad 42, \;\; 6b + 12m \quad 20]$$

m =

b =

?



# Linear Regression: Gradient Descent

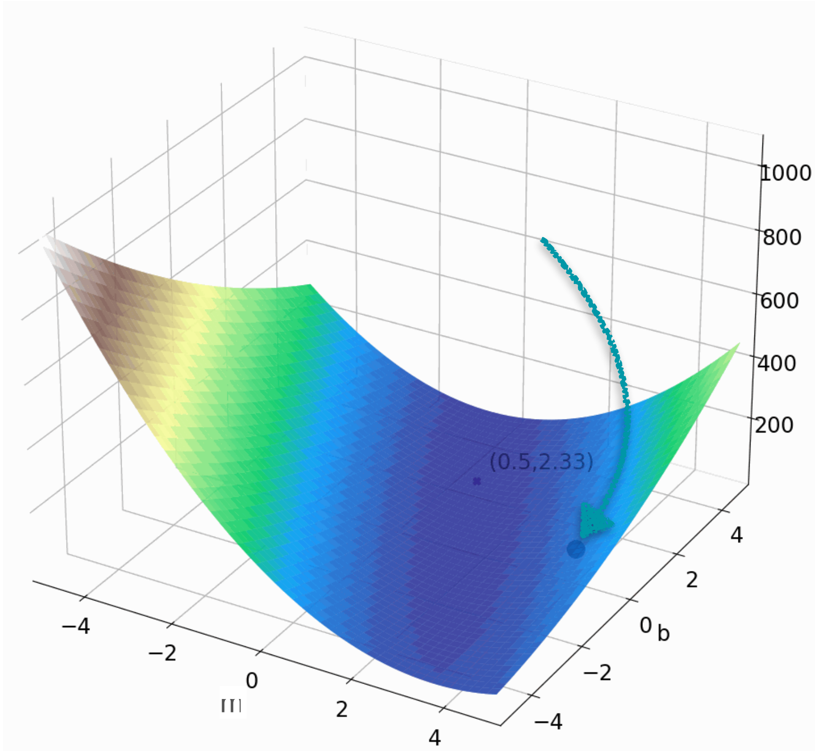
**Goal: Minimize sum of squares cost**

$$E = [28m + 12b \quad 42, \; 6b + 12m \quad 20]$$

$m =$

$b =$

**?**



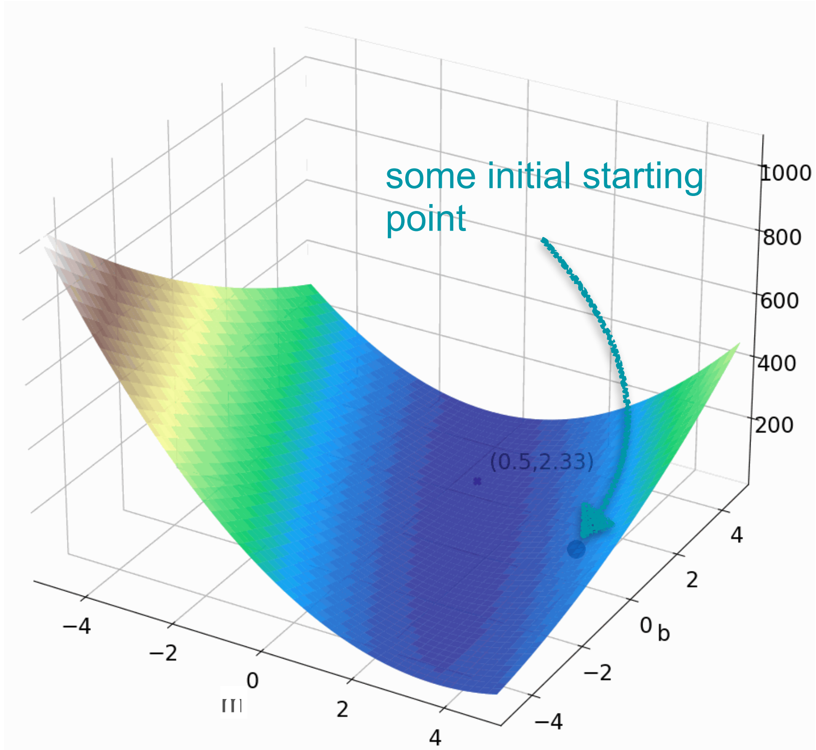
# Linear Regression: Gradient Descent

**Goal: Minimize sum of squares cost**

$$E = [28m + 12b \quad 42, \;\; 6b + 12m \quad 20]$$

m =

b = **?**



# Linear Regression: Gradient Descent

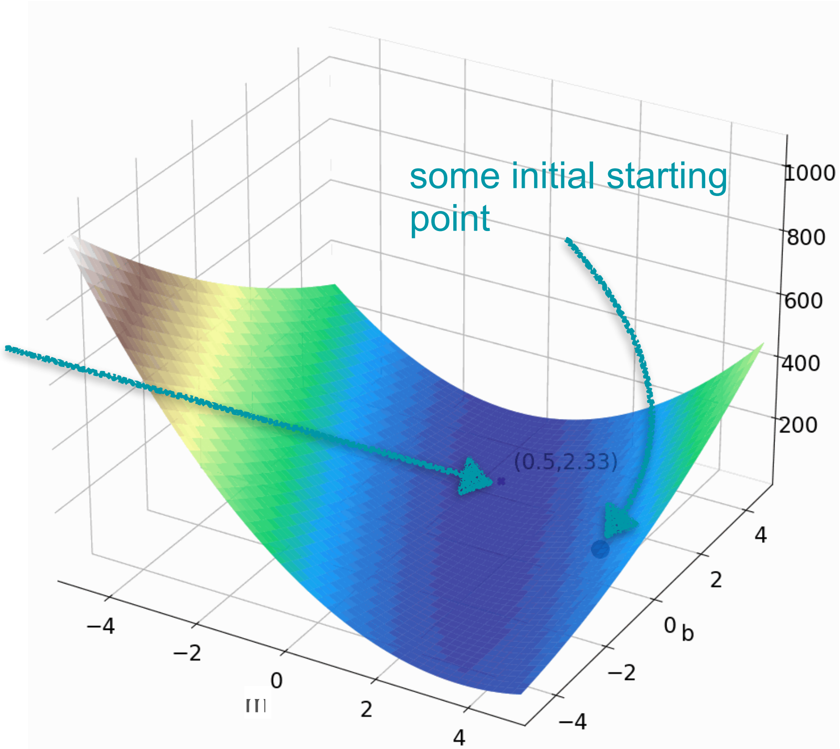
**Goal: Minimize sum of squares cost**

$$E = [28m + 12b \quad 42,\ \ 6b + 12m \quad 20]$$

m =

b =

?



# Linear Regression: Gradient Descent

**Goal: Minimize sum of squares cost**

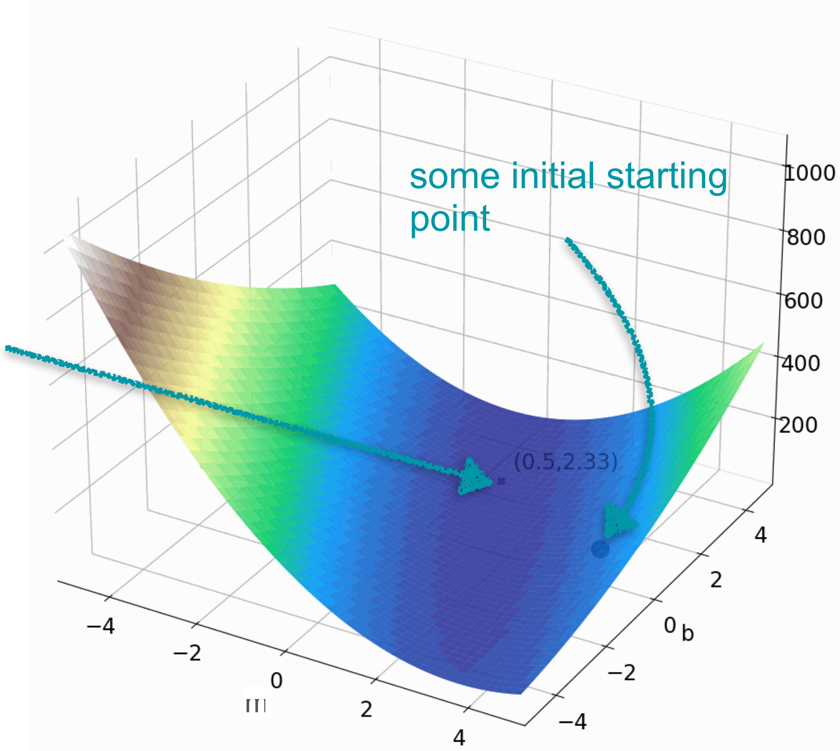
$$E = [28m + 12b \quad 42,\ 6b + 12m \quad 20]$$

m =

b =

?

The points m,b such that the cost is minimum



# Linear Regression: Gradient Descent

**Goal: Minimize sum of squares cost**

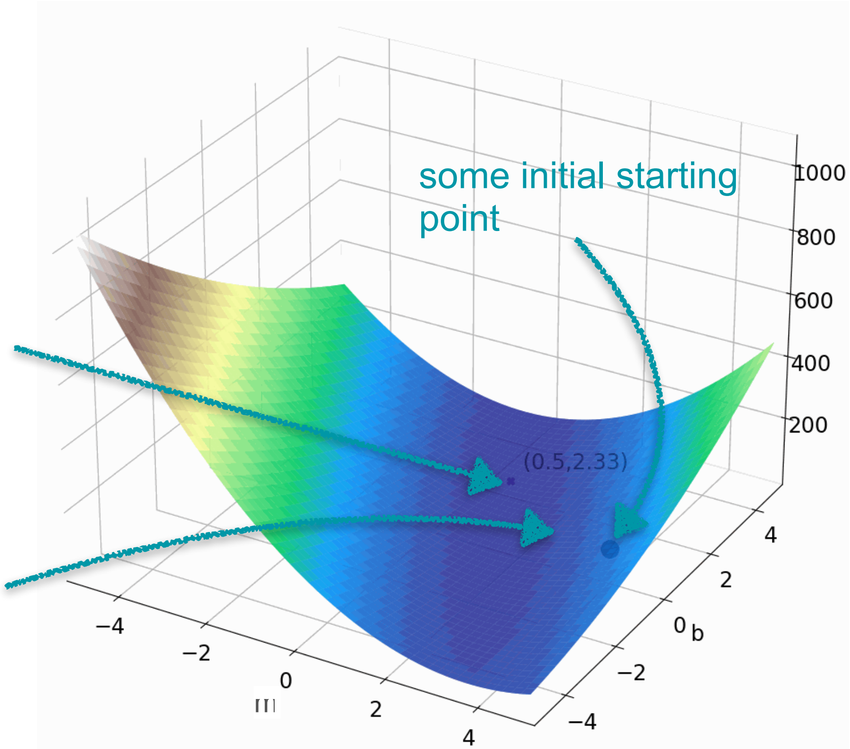
E = [28m + 12b  42,  6b + 12m  20]

m =

b =

?

The points m,b such that the cost is minimum



# Linear Regression: Gradient Descent

**Goal: Minimize sum of squares cost**

$E = [28m + 12b \quad 42, \; 6b + 12m \quad 20]$

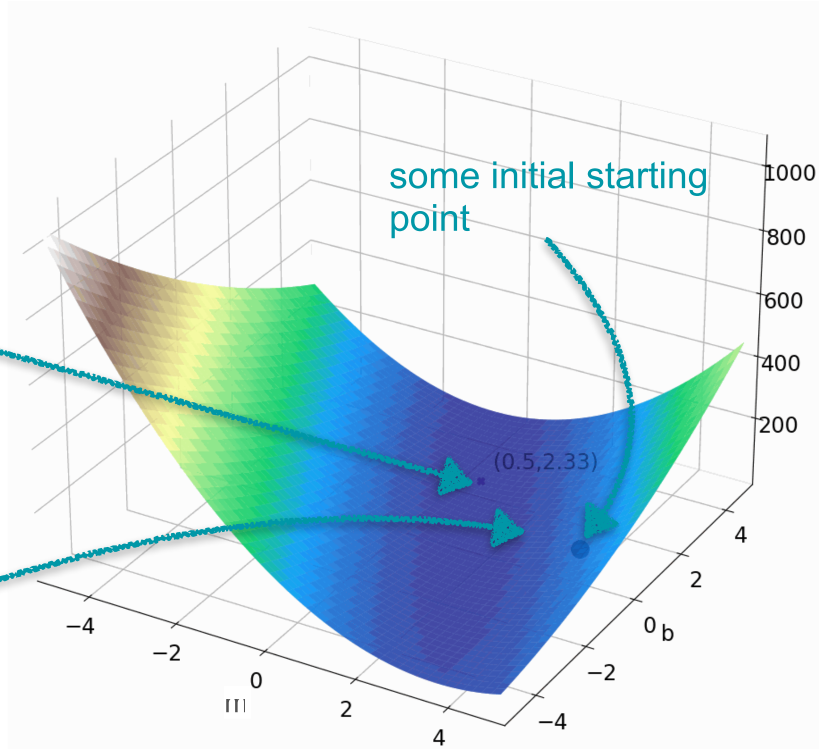
m =

b =

?

The points m,b such that
the cost is minimum

descend until you
find the minimum



# Linear Regression: Gradient Descent

**Goal: Minimize sum of squares cost**

$$E = [28m + 12b \quad 42, \;\; 6b + 12m \quad 20]$$

m =

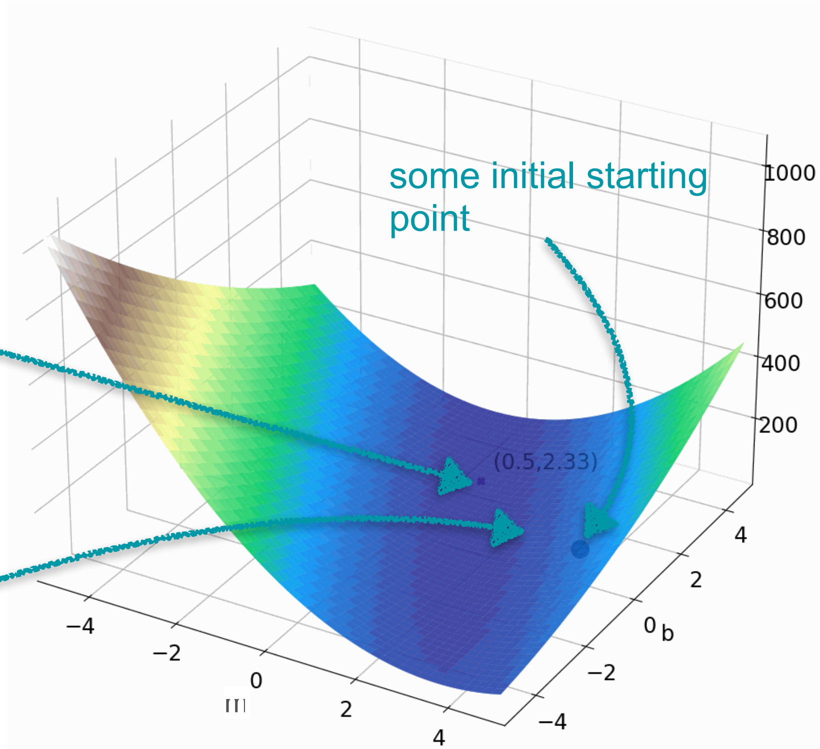
b =

?

Steps:

The points m,b such that the cost is minimum

descend until you find the minimum



# Linear Regression: Gradient Descent

**Goal: Minimize sum of squares cost**

$E = [28m + 12b \quad 42, \;\; 6b + 12m \quad 20]$

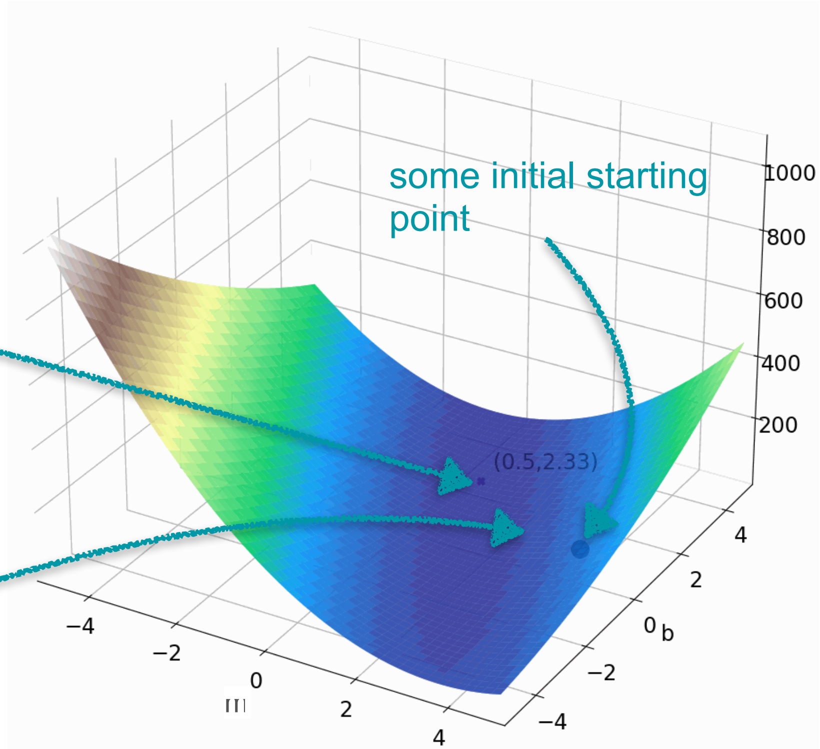
m =
b = **?**

The points m,b such that the cost is minimum

Steps:

Start with $(m_0, b_0)$

descend until you find the minimum



# Linear Regression: Gradient Descent

**Goal: Minimize sum of squares cost**

$$E = [28m + 12b \quad 42, \; 6b + 12m \quad 20]$$

$m =$

$b =$ **?**

The points m,b such that the cost is minimum

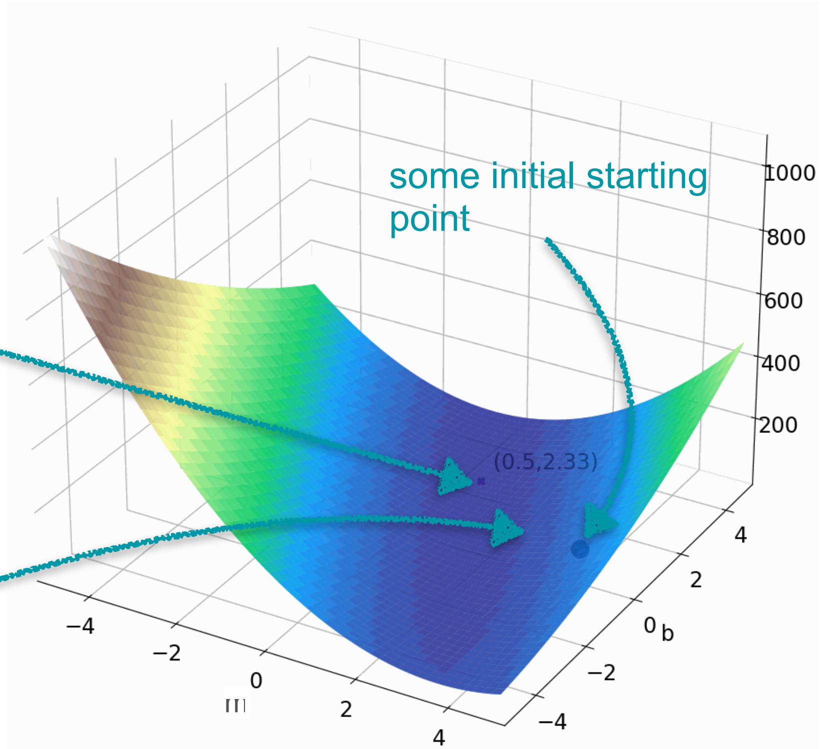
Steps:

Start with $(m_0, b_0)$

descend until you find the minimum

Iterate

$$(m_{k+1}, b_{k+1}) = (m_k, b_k) \qquad E(m_k, b_k)$$

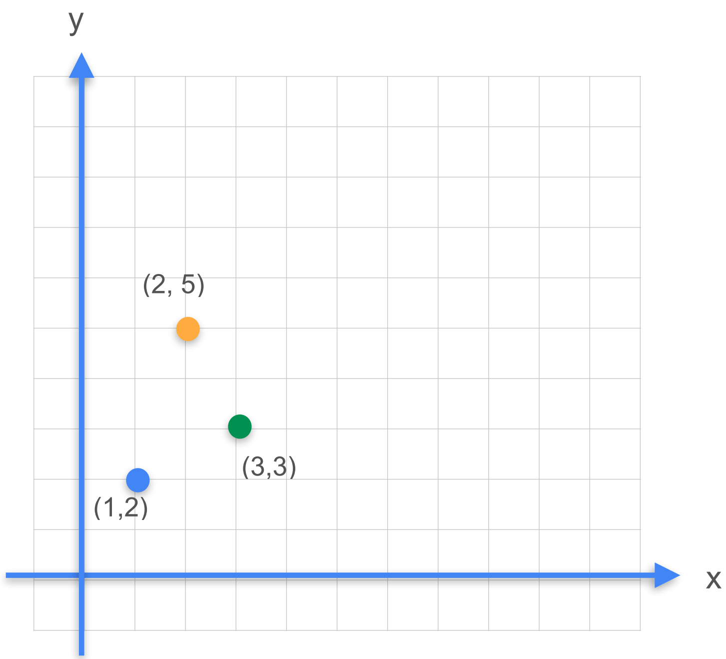


DeepLearning.AI

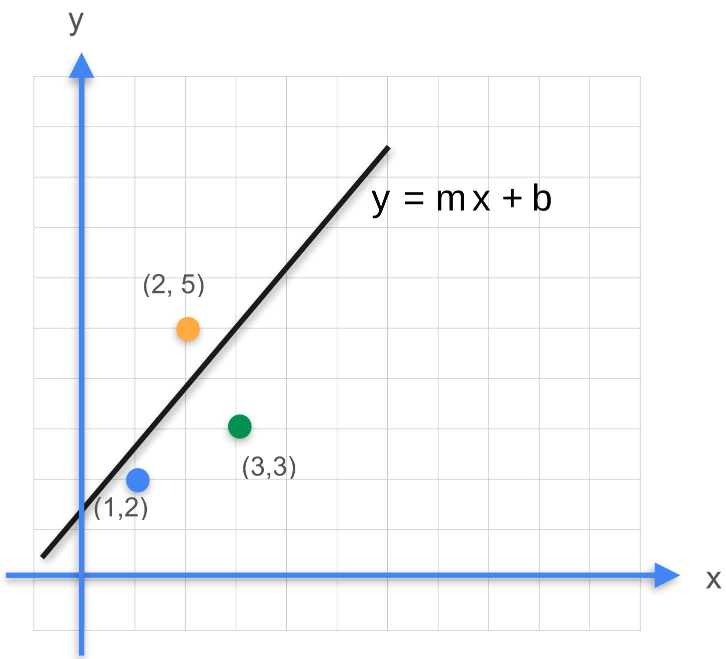
# Gradients and Gradient Descent

## Optimization using Gradient Descent - Least squares with multiple observations

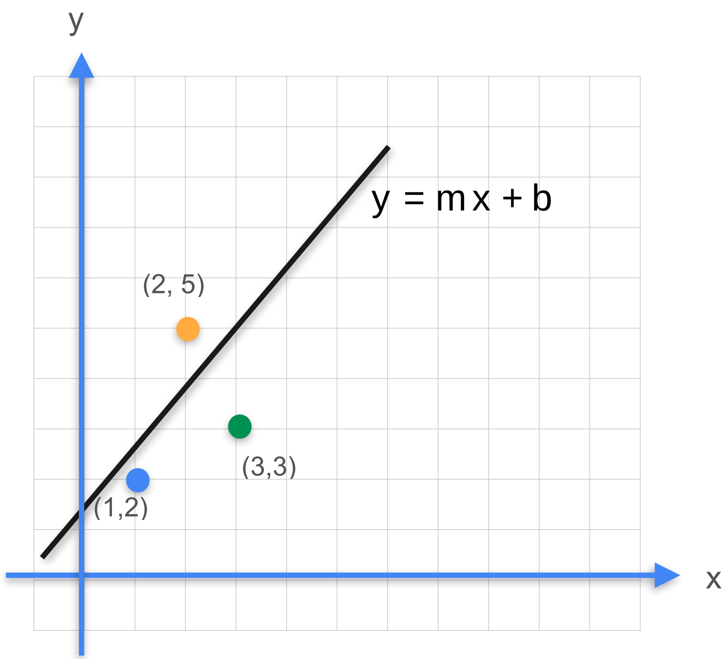
# Gradient Descent

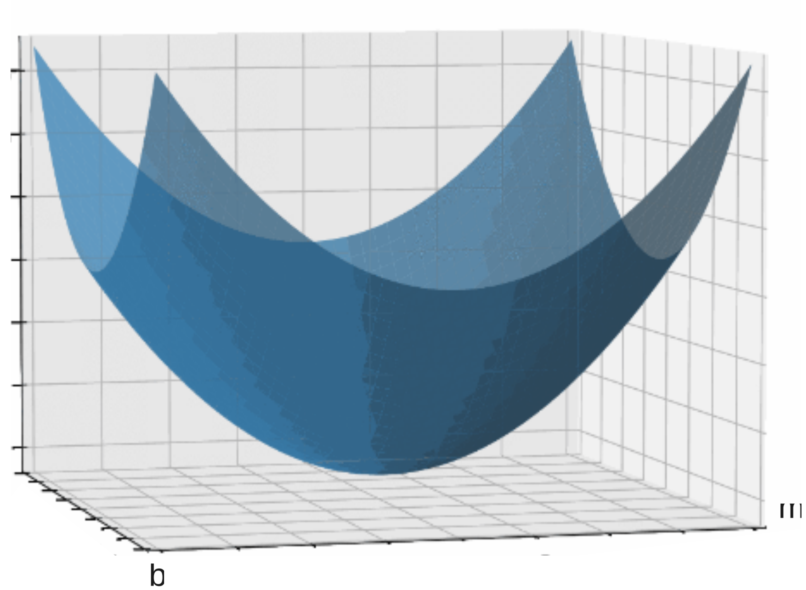


# Gradient Descent

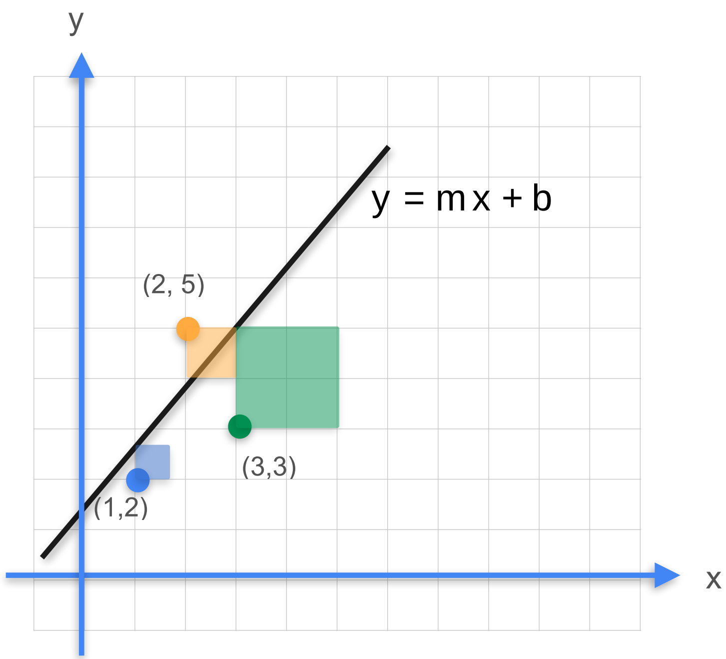


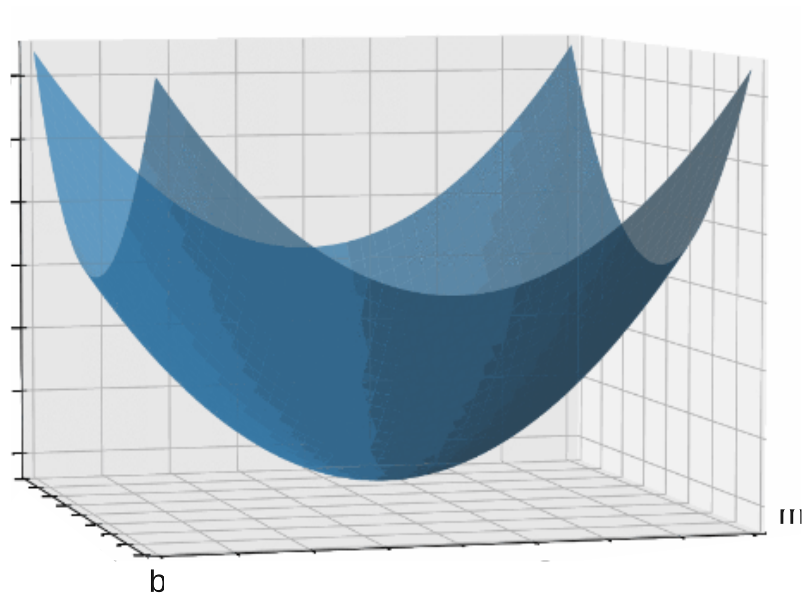
# Gradient Descent



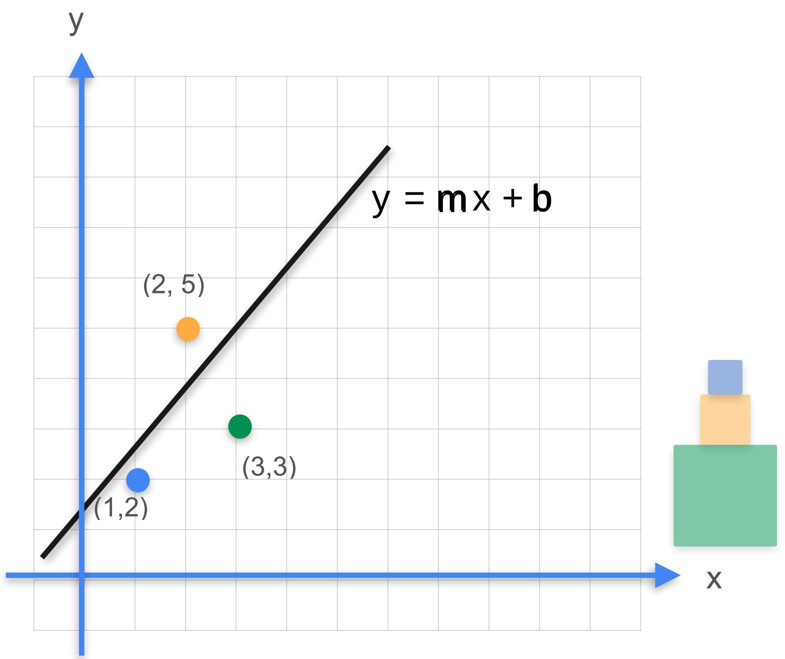



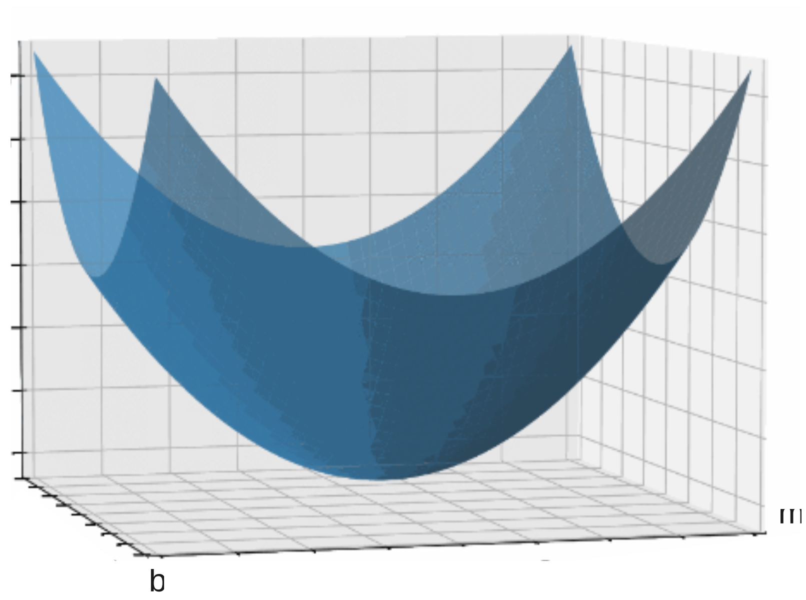
# Gradient Descent



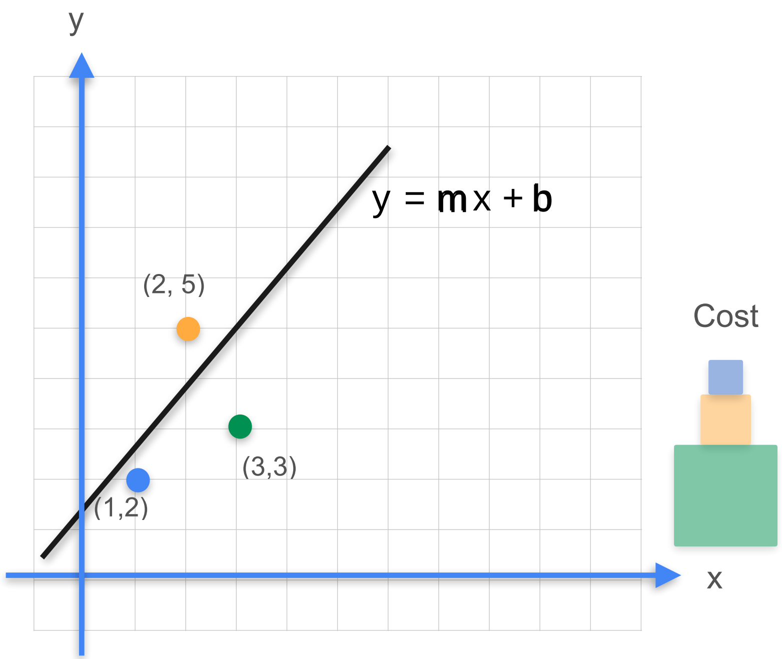



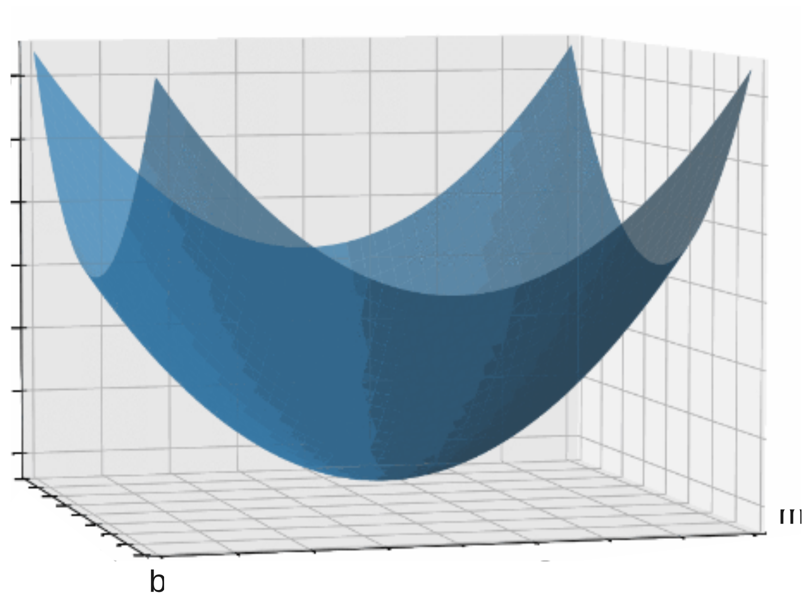
# Gradient Descent



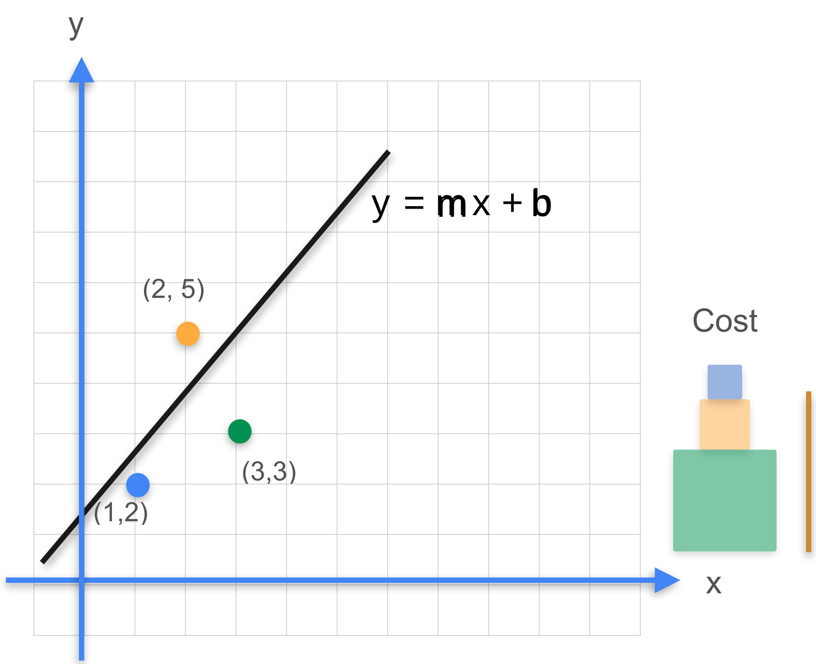



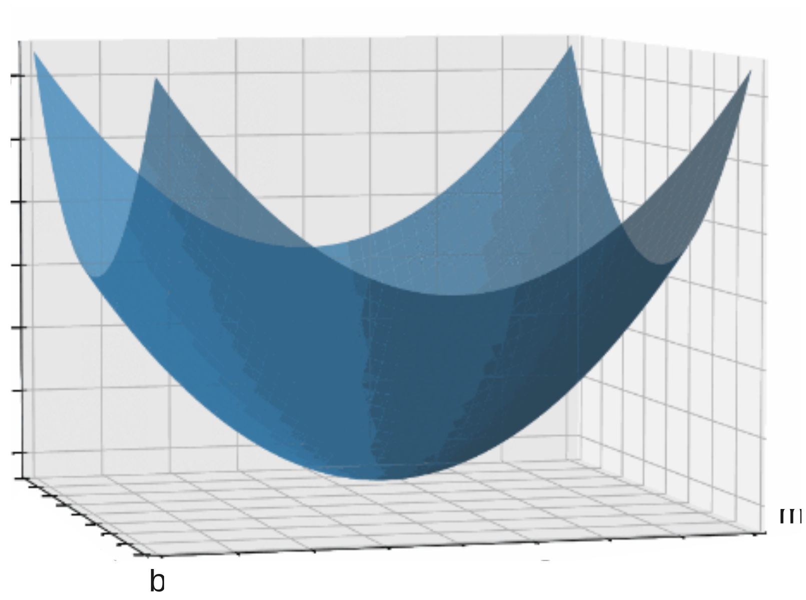
# Gradient Descent



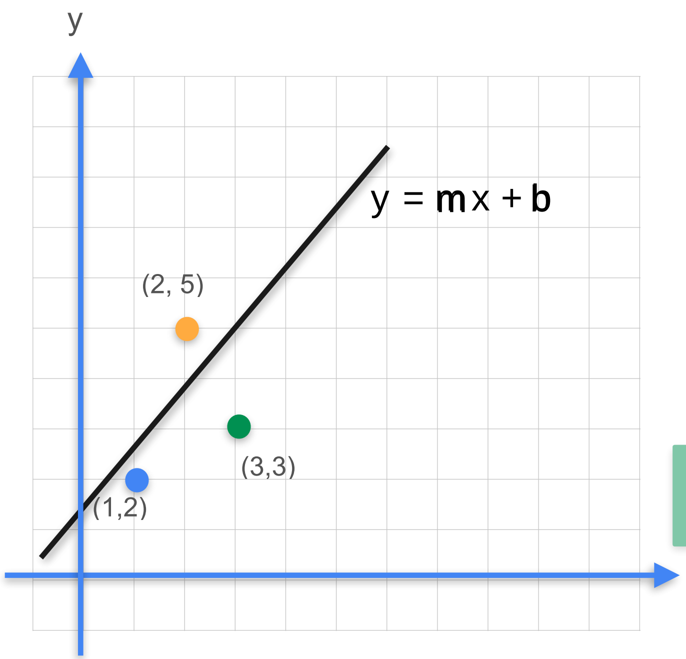



# Gradient Descent

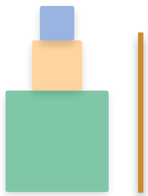


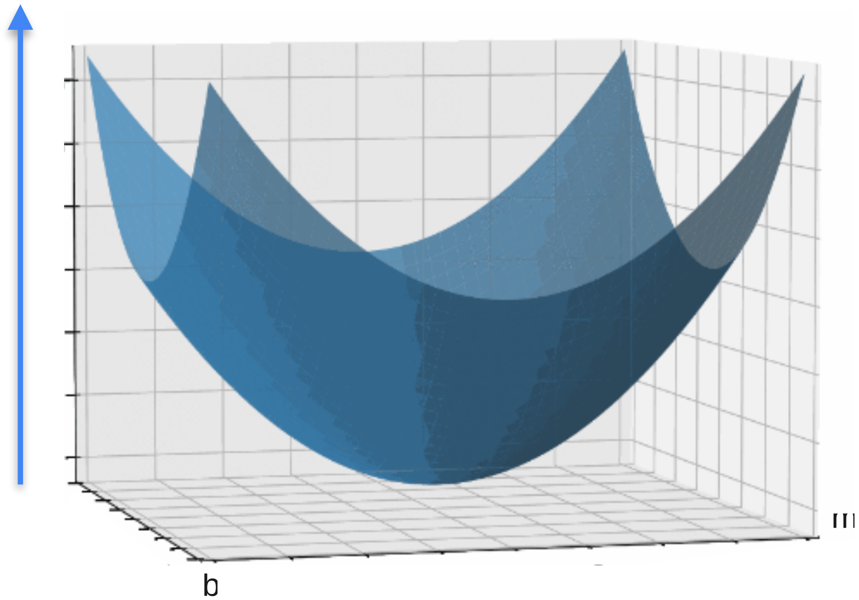
# Gradient Descent



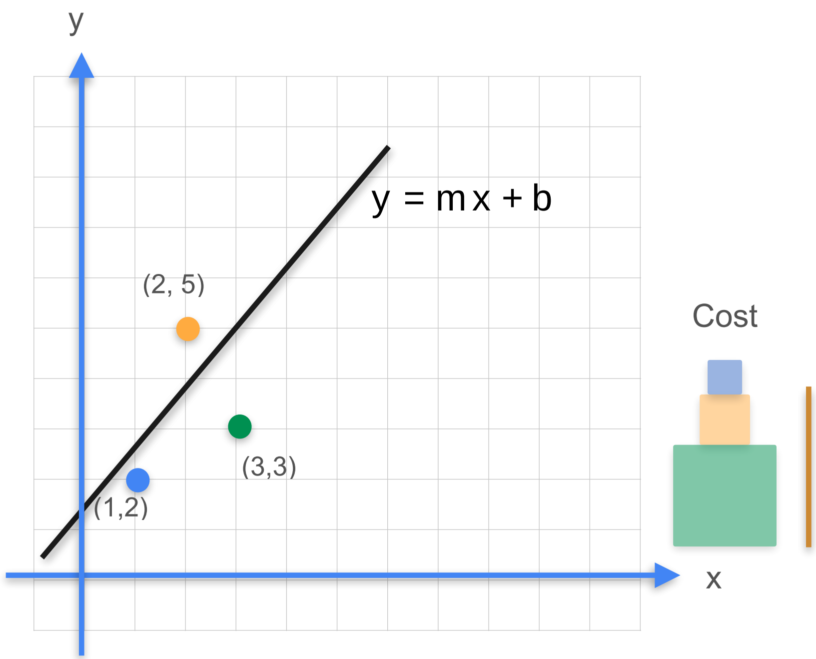

Cost

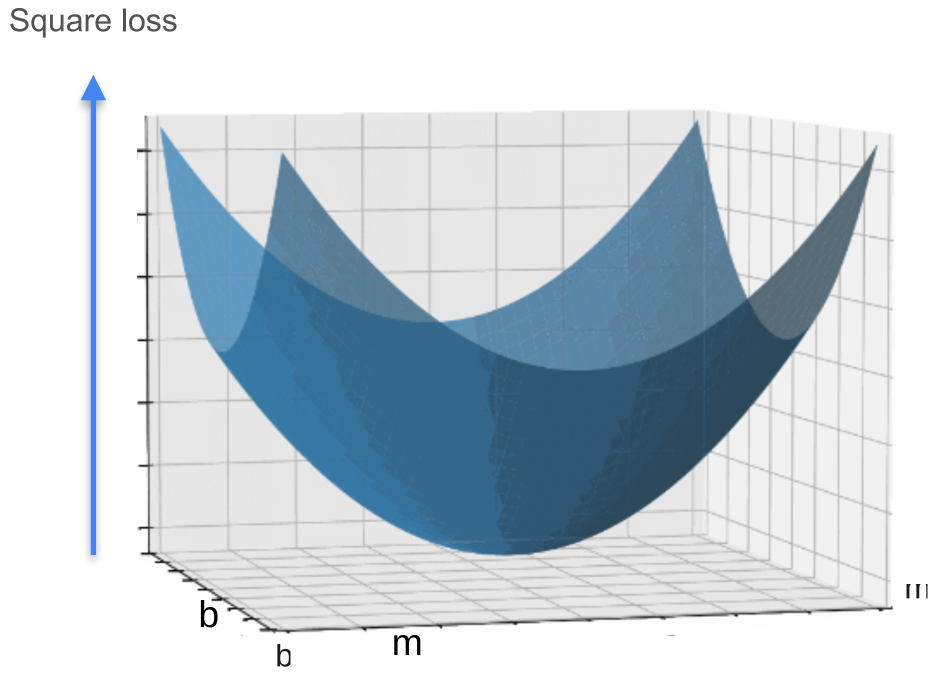


Square loss

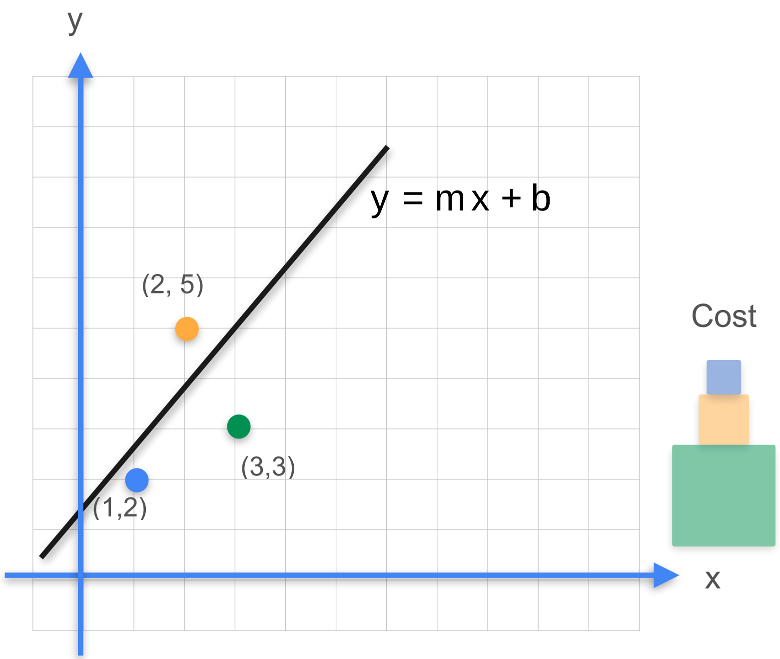


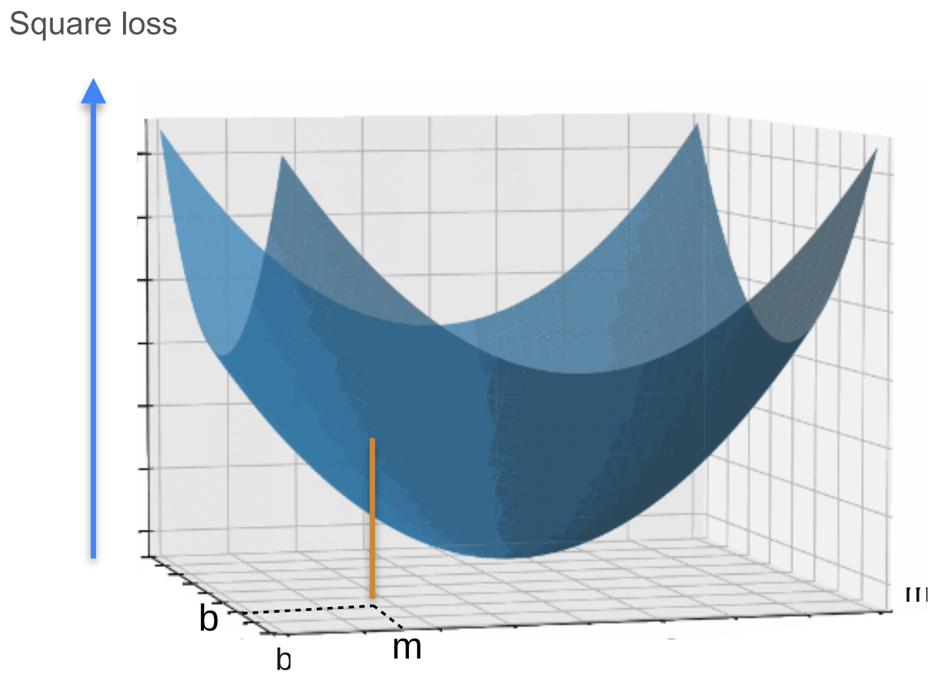
# Gradient Descent



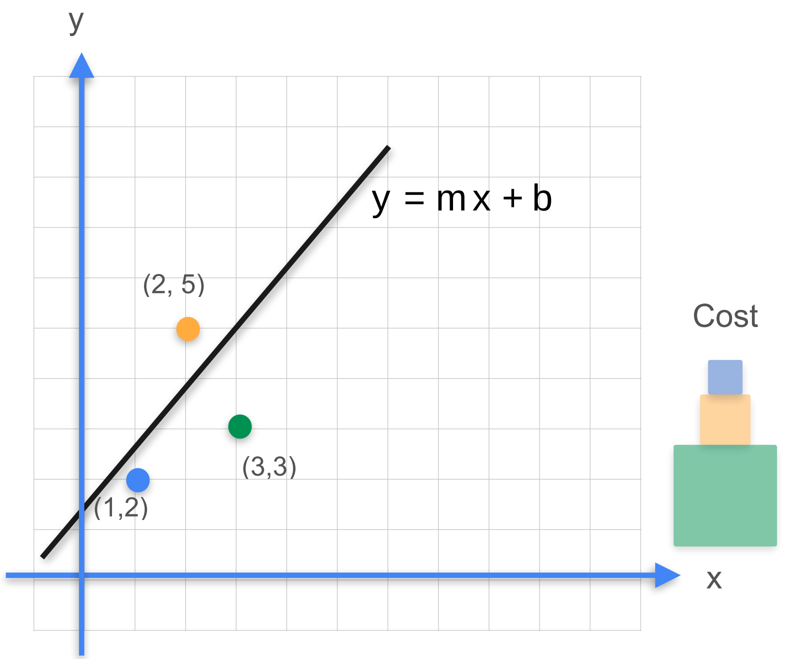



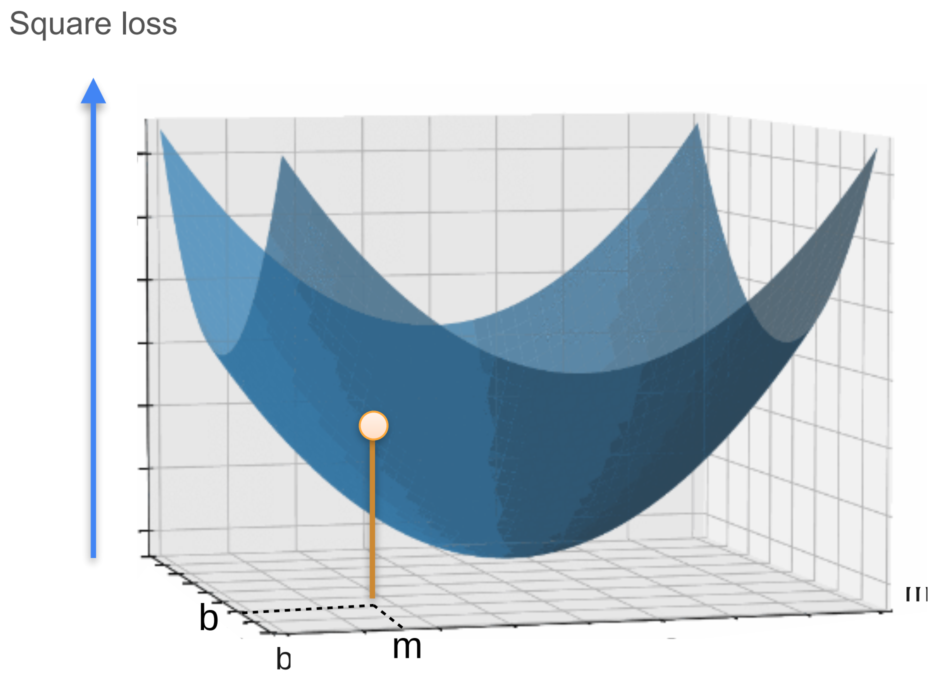
# Gradient Descent



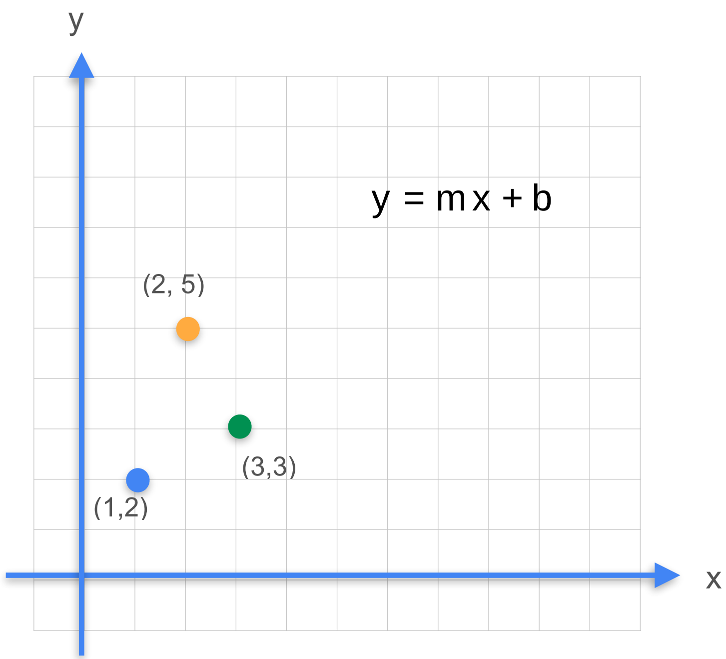



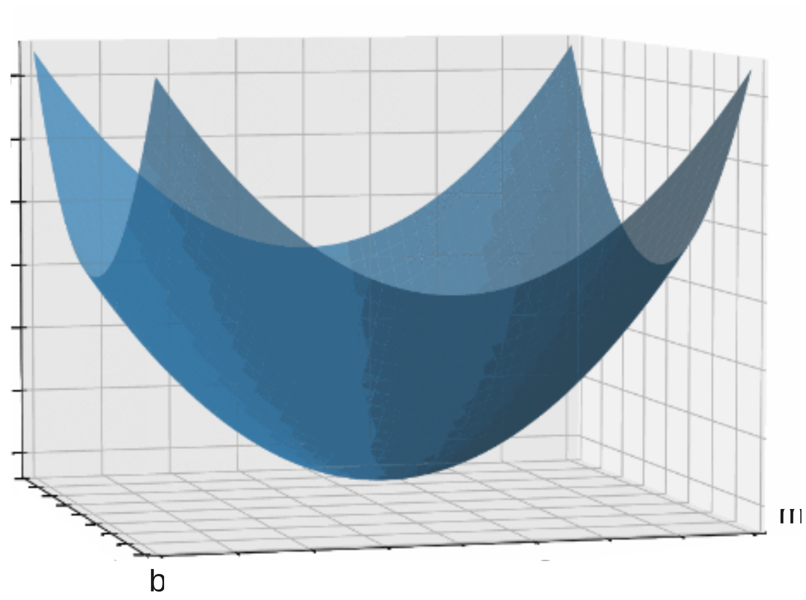
# Gradient Descent



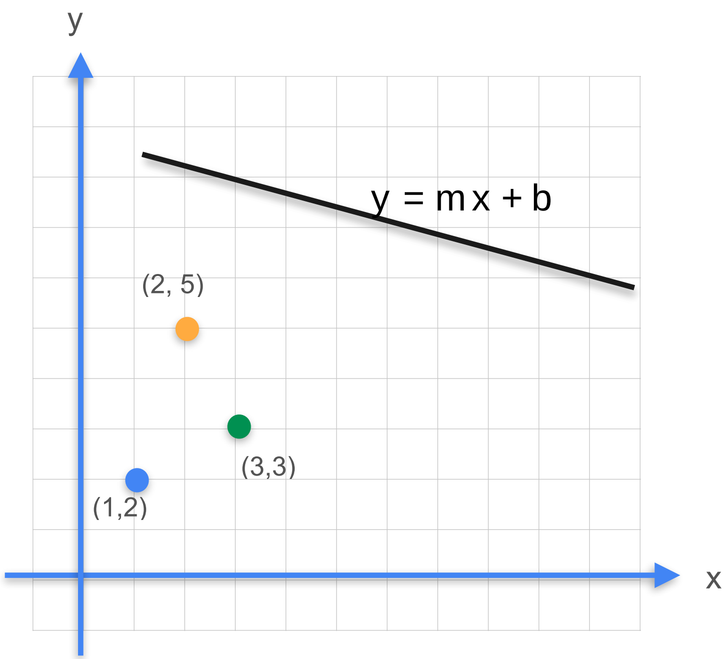



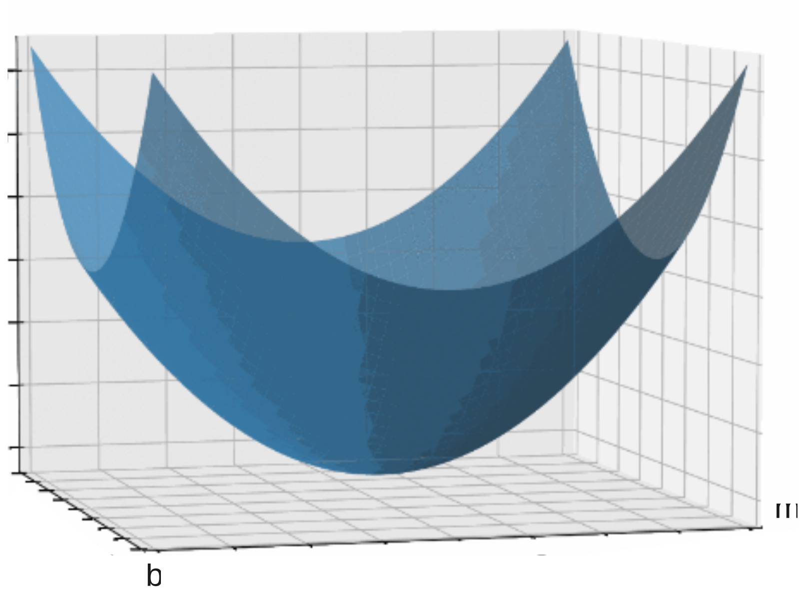
# Gradient Descent



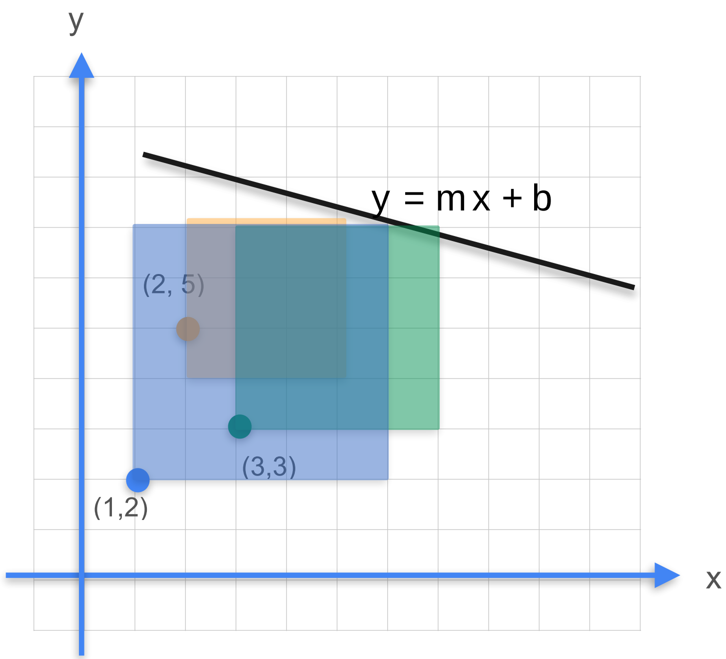



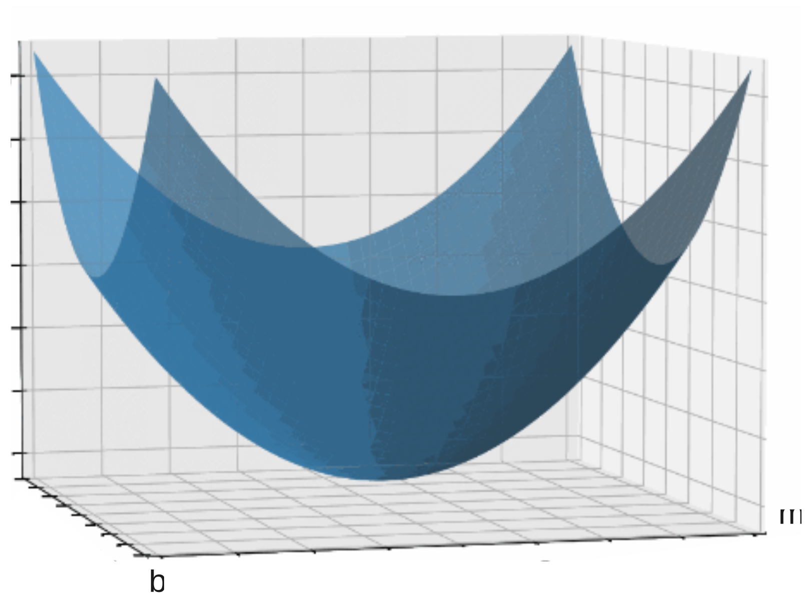
# Gradient Descent



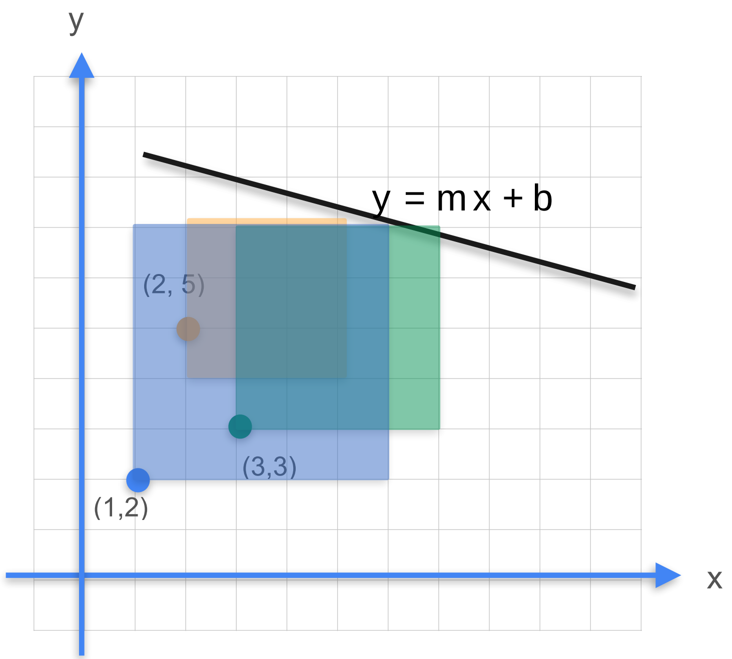



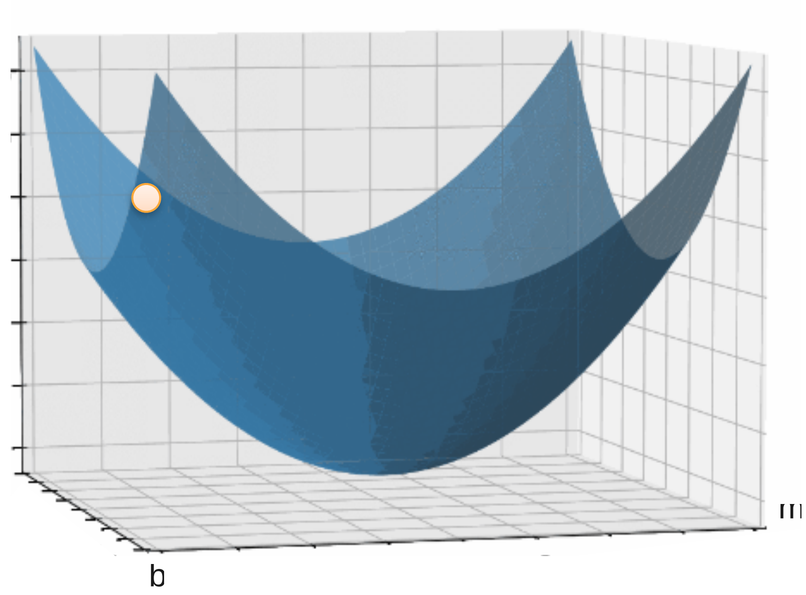
# Gradient Descent



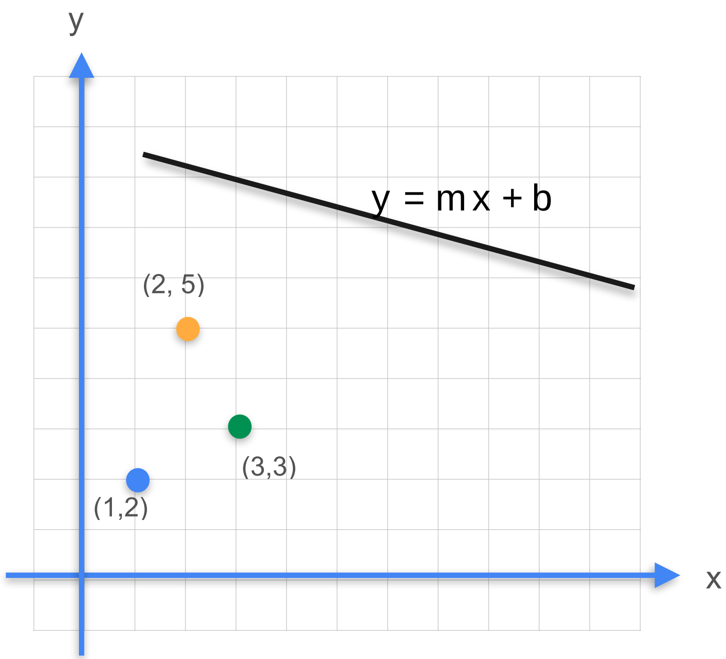



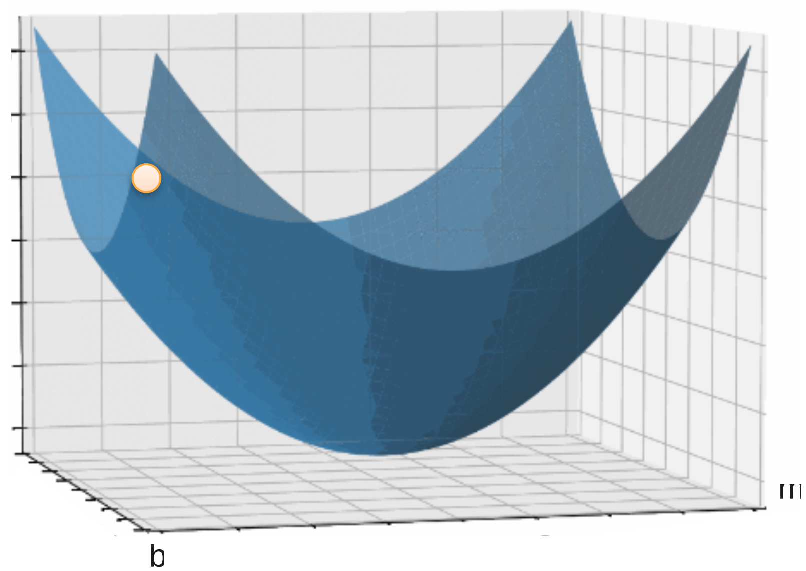
# Gradient Descent



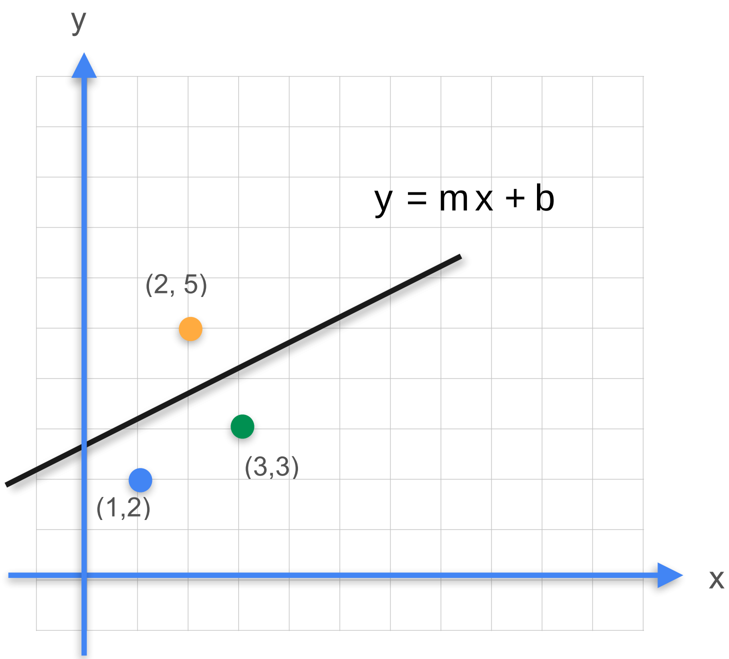



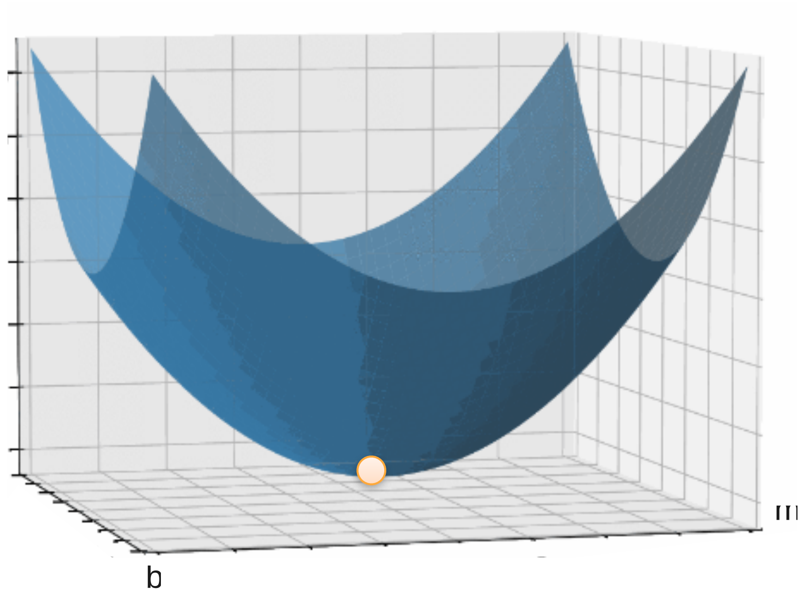
# Gradient Descent



# Gradient Descent



# Another Example

# Another Example

# Another Example

TV advertisement budget

# Another Example

TV advertisement
budget

# Another Example

TV advertisement budget

Number of sales

# Another Example

# Another Example

| TV budget | Sales |
| --- | --- |

# Another Example

| TV budget | Sales |
|---|---|
| 230.1 | 22.1 |

# Another Example

| TV budget | Sales |
|---|---|
| 230.1 | 22.1 |
| 44.5 | 10.4 |

# Another Example

| TV budget | Sales |
|---|---|
| 230.1 | 22.1 |
| 44.5 | 10.4 |
| 17.2 | 9.3 |

# Another Example

| TV budget | Sales |
|---|---|
| 230.1 | 22.1 |
| 44.5 | 10.4 |
| 17.2 | 9.3 |

Goal: Predict sales in terms of TV budget

# Another Example

| TV budget | Sales |
| --- | --- |
| 230.1 | 22.1 |
| 44.5 | 10.4 |
| 17.2 | 9.3 |

Goal: Predict sales in terms of TV budget

Tool: Linear regression

# Another Example

| TV budget | Sales |
|---|---|
| 230.1 | 22.1 |
| 44.5 | 10.4 |
| 17.2 | 9.3 |

Goal: Predict sales in terms of TV budget

Tool: Linear regression

$y = mx + b$

# Another Example

| TV budget | Sales |
|---|---|
| 230.1 | 22.1 |
| 44.5 | 10.4 |
| 17.2 | 9.3 |

Goal: Predict sales in terms of TV budget

Tool: Linear regression

$y = mx + b$

# Another Example

| TV budget | Sales |
|---|---|
| 230.1 | 22.1 |
| 44.5 | 10.4 |
| 17.2 | 9.3 |

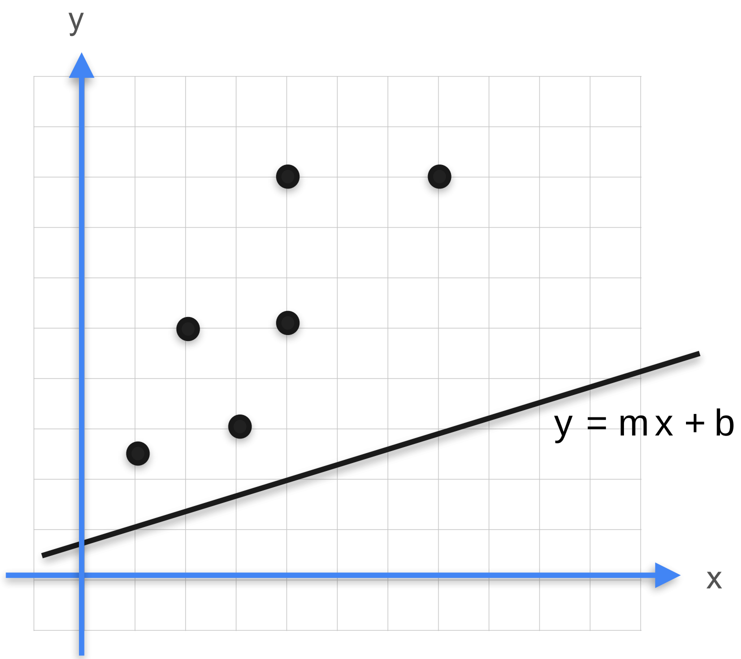
Multiple observations

Goal: Predict sales in terms of TV budget

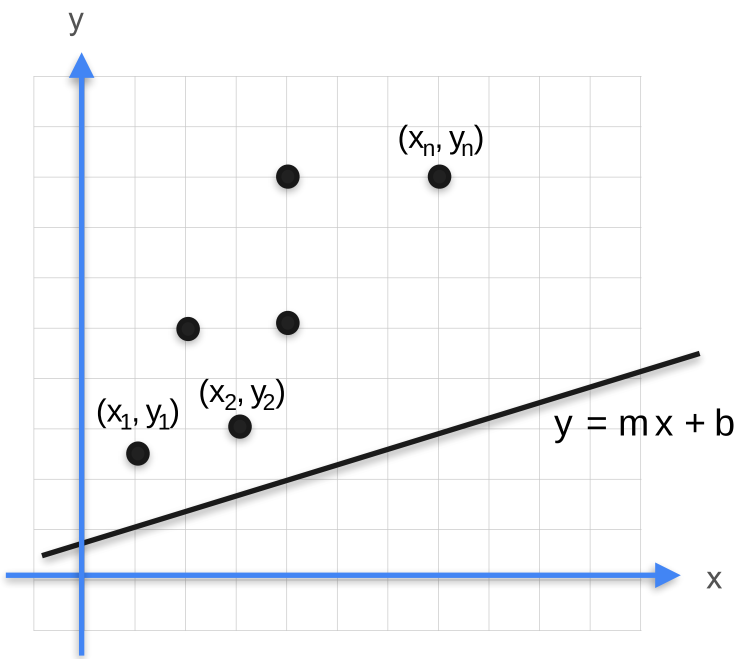
Tool: Linear regression

$y = mx + b$

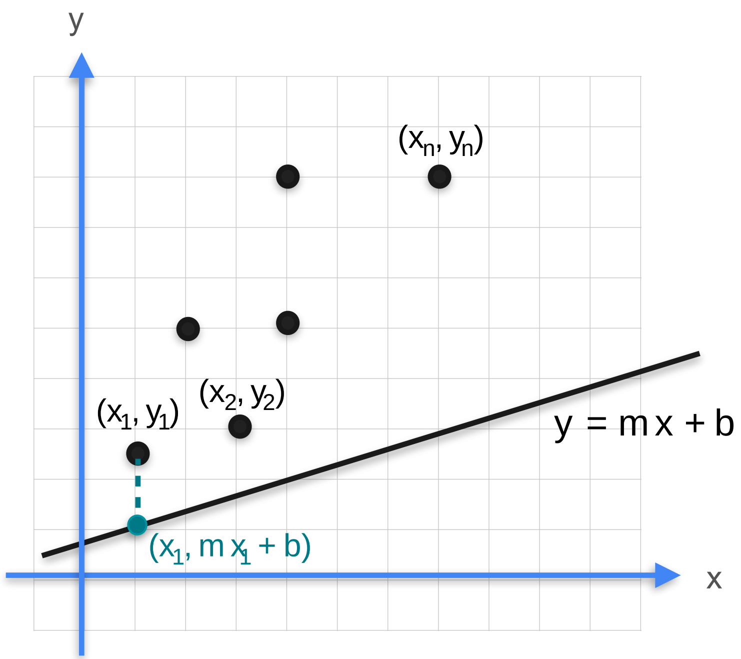
# Gradient Descent



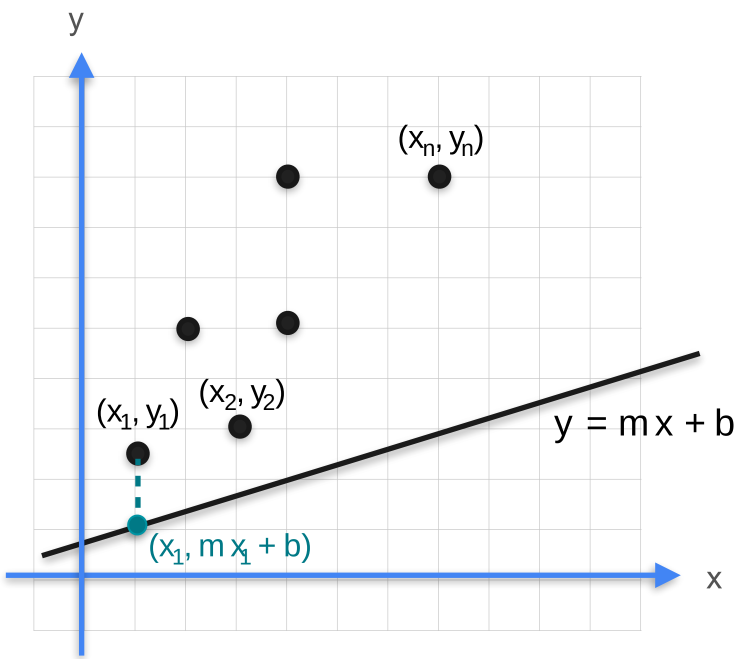
# Gradient Descent



# Gradient Descent



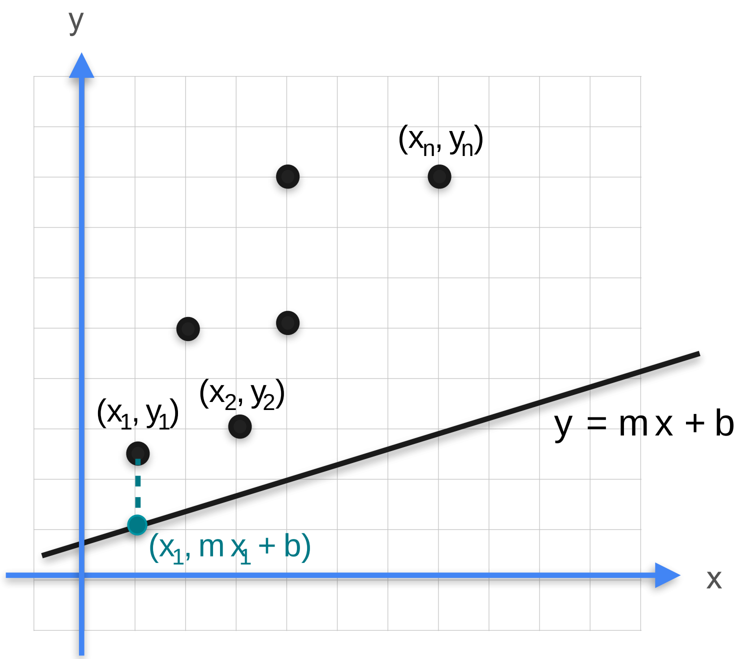
# Gradient Descent




Loss

$$m x_1 + b \quad y_1$$

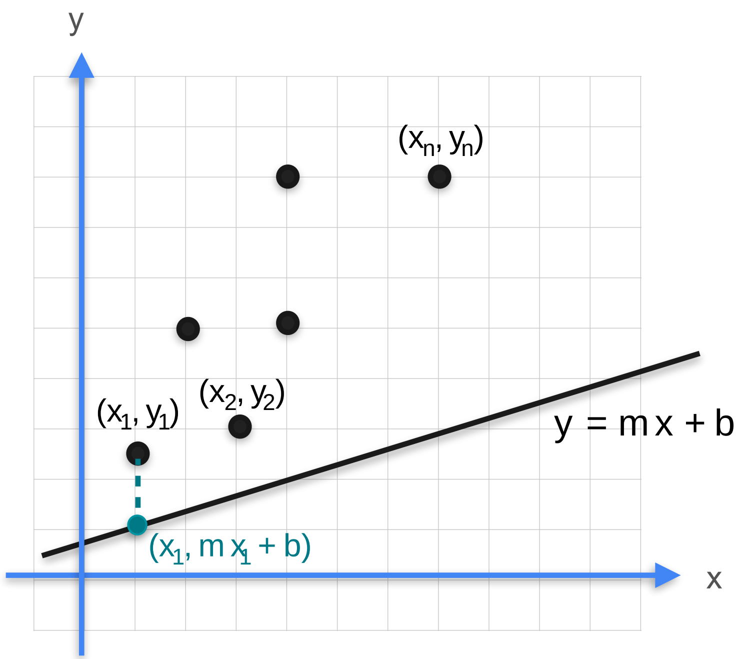
# Gradient Descent




Loss

$$(mx_1 + b \quad y_1)^2$$

# Gradient Descent

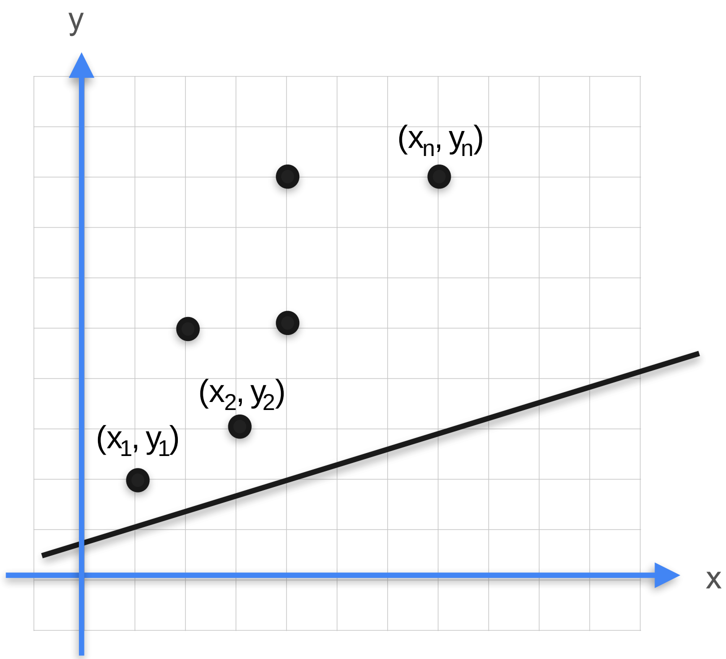



Cost

Loss

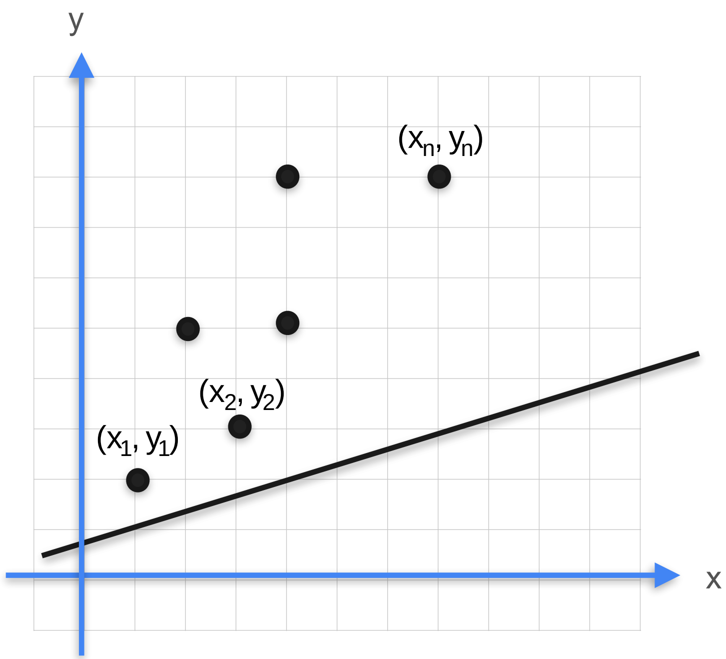
$$(m, b) = \frac{1}{2m}\left[(mx_1 + b \quad y_1)^2 + \quad + (mx_n \quad y_n)^2\right]$$

# Gradient Descent




$$(m, b) = \frac{1}{2m}[(mx_1 + b \quad y_1)^2 + \quad + (mx_n \quad y_n)^2]$$

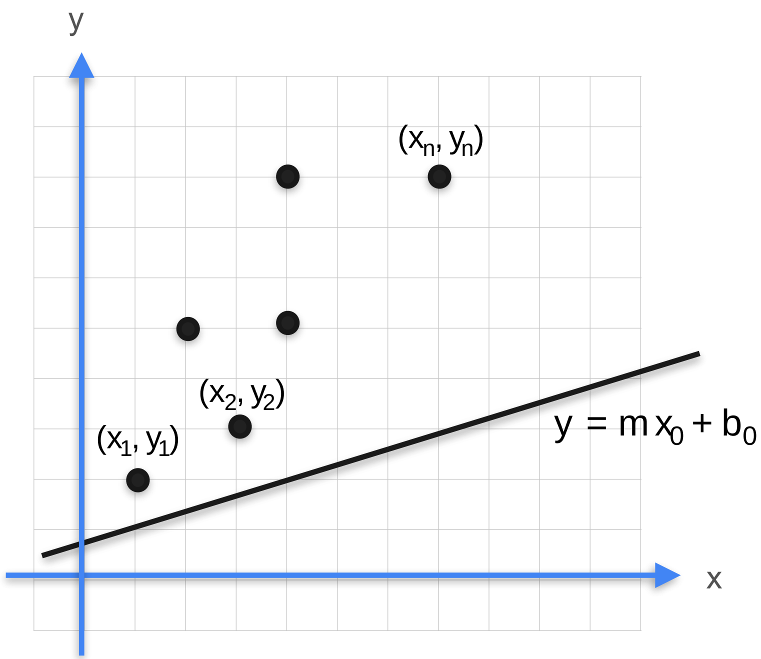
# Gradient Descent




$$(m, b) = \frac{1}{2m}\left[(m x_1 + b \quad y_1)^2 + \quad + (m x_n \quad y_n)^2\right]$$

$$\begin{bmatrix} m_0 \\ b_0 \end{bmatrix}$$

# Gradient Descent

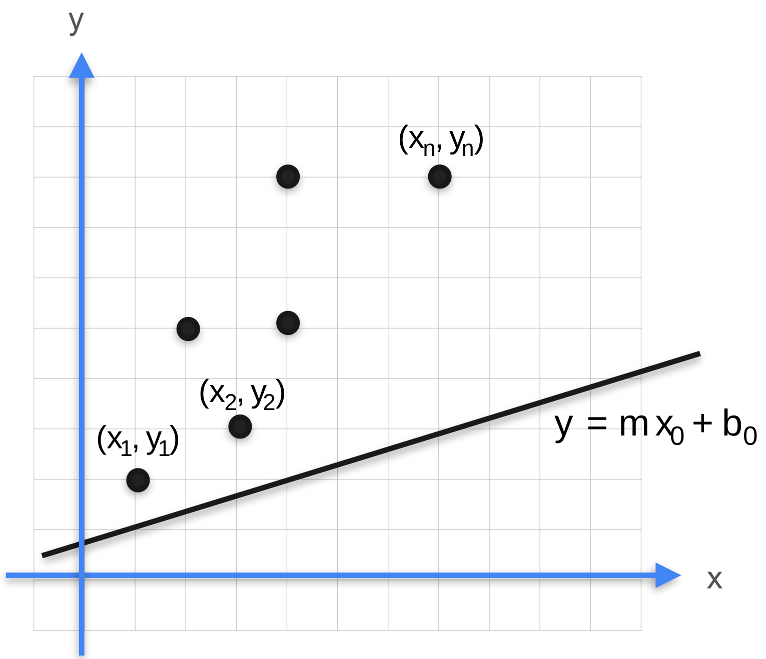



$$(m, b) = \frac{1}{2m}[(mx_1 + b \quad y_1)^2 + \quad + (mx_{\ n} \quad y_n)^2]$$

$m_0$

$[\ b_0]$

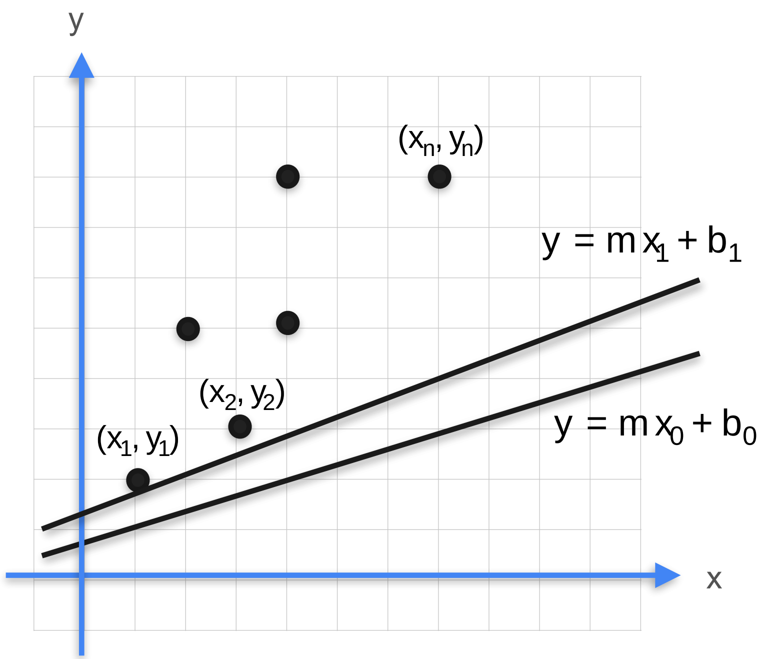
# Gradient Descent




$$(m, b) = \frac{1}{2m}\left[(m x_1 + b \quad y_1)^2 + \quad + (m x_n \quad y_n)^2\right]$$

$$\begin{bmatrix} m_0 \\ b_0 \end{bmatrix} \rightarrow \begin{bmatrix} m_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} m_0 \\ b_0 \end{bmatrix} \quad {}_1(m_0, b_0)$$

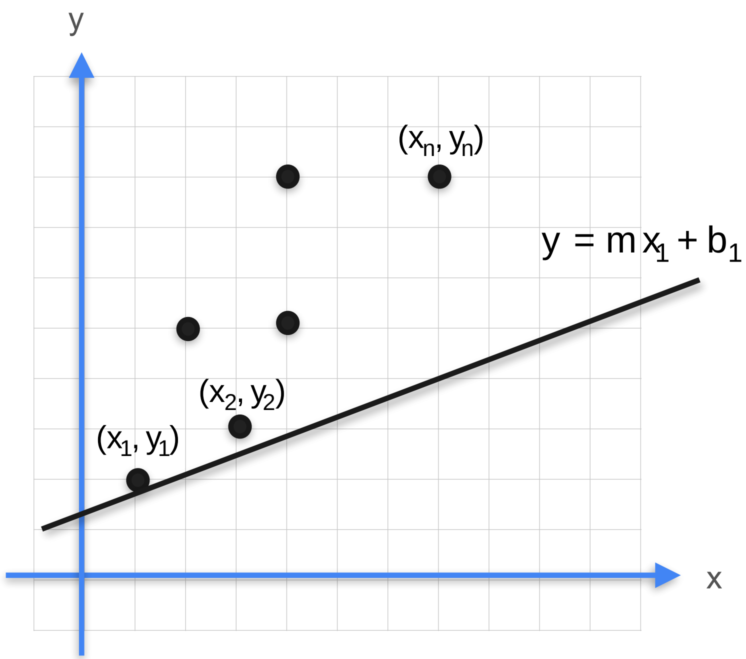
# Gradient Descent




$$(m, b) = \frac{1}{2m}[(m x_1 + b \quad y_1)^2 + \quad + (m x_{\ n} \quad y_n)^2]$$

$$\begin{bmatrix} m_0 \\ b_0 \end{bmatrix} \rightarrow \begin{bmatrix} m_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} m_0 \\ b_0 \end{bmatrix} \qquad {}_1(m_0, b_0)$$

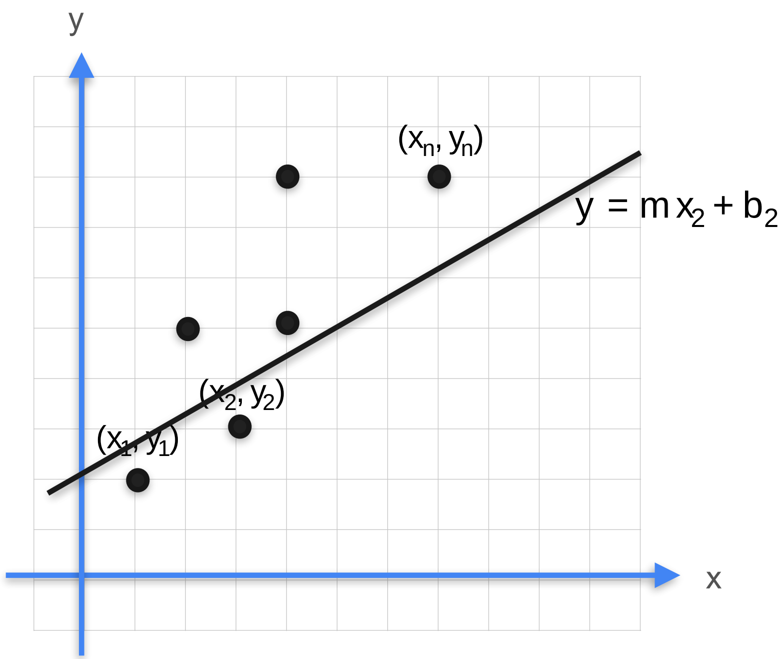
# Gradient Descent




$$(m, b) = \frac{1}{2m}\left[(m x_1 + b \quad y_1)^2 + \quad + (m x \;_n \quad y_n)^2\right]$$

$$\begin{bmatrix} m_0 \\ b_0 \end{bmatrix} \rightarrow \begin{bmatrix} m_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} m_0 \\ b_0 \end{bmatrix} \qquad {}_1(m_0, b_0)$$

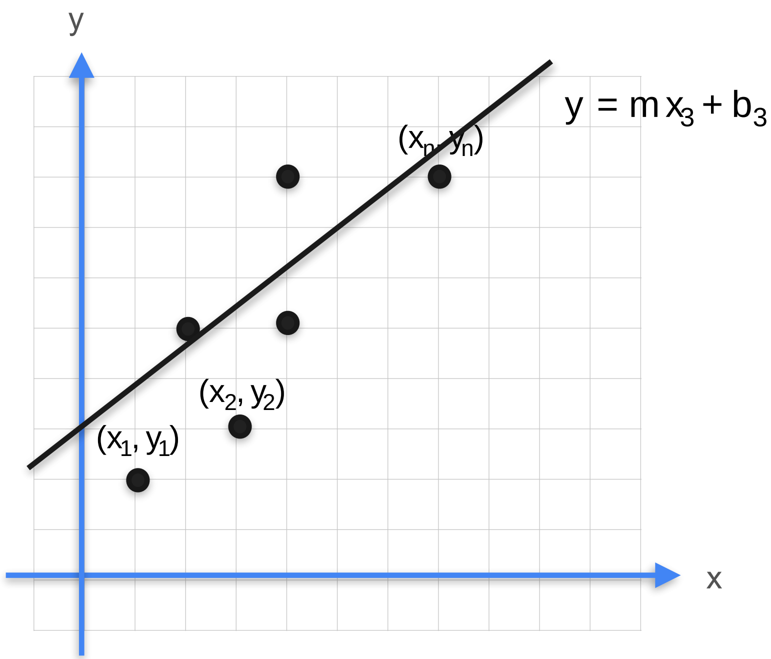
# Gradient Descent




$$(m, b) = \frac{1}{2m}\left[(mx_1 + b \quad y_1)^2 + \quad + (mx_n \quad y_n)^2\right]$$

$$\begin{bmatrix} m_1 \\ b_1 \end{bmatrix} \rightarrow \begin{bmatrix} m_2 \\ b_2 \end{bmatrix} = \begin{bmatrix} m_1 \\ b_1 \end{bmatrix} \qquad {}_1(m_1, b_1)$$

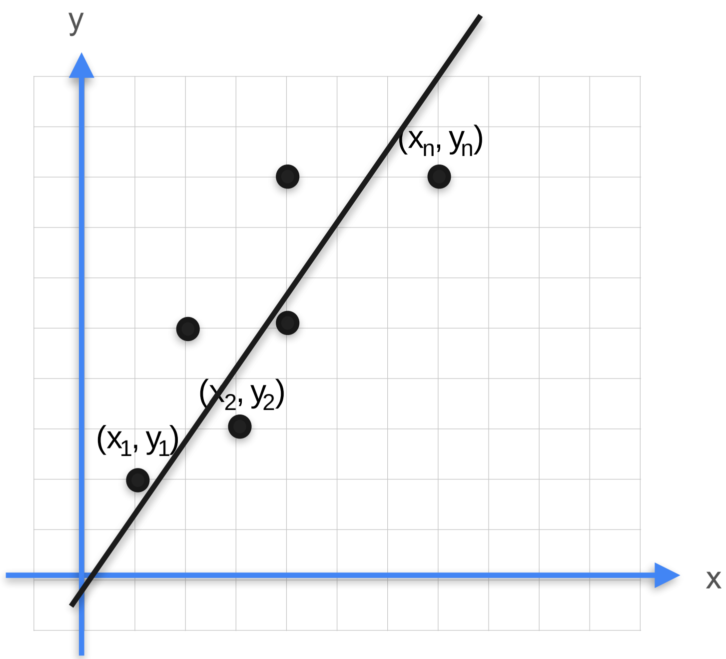
# Gradient Descent




$$(m, b) = \frac{1}{2m} [(m x_1 + b \quad y_1)^2 + \quad + (m x_n \quad y_n)^2]$$

$$\begin{bmatrix} m_2 \\ b_2 \end{bmatrix} \rightarrow \begin{bmatrix} m_3 \\ b_3 \end{bmatrix} = \begin{bmatrix} m_2 \\ b_2 \end{bmatrix} \qquad {}_1(m_2, b_2)$$

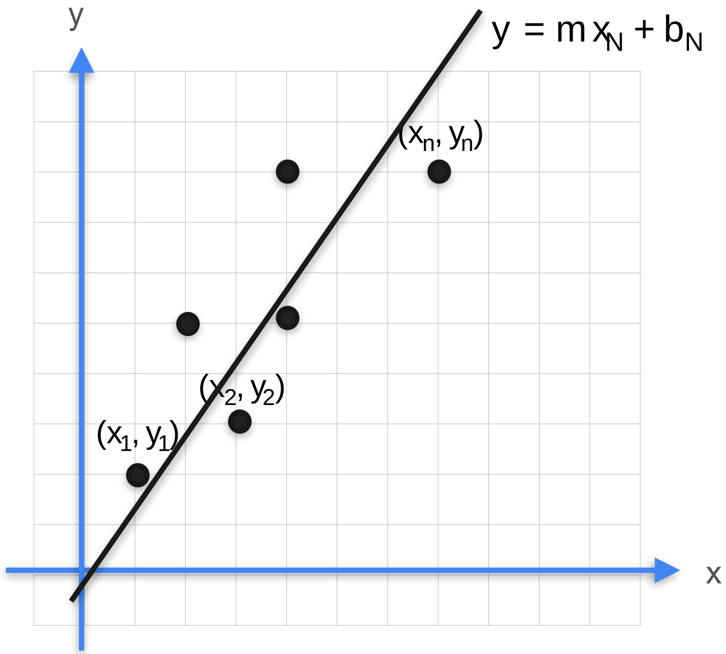
# Gradient Descent




$$(m, b) = \frac{1}{2m}\left[(m x_1 + b \quad y_1)^2 + \quad + (m x_n \quad y_n)^2\right]$$

$$\begin{bmatrix} m_N \\ b_N \end{bmatrix} \rightarrow \begin{bmatrix} m_N \\ b_N \end{bmatrix} = \begin{bmatrix} m_{N \; 1} \\ b_{N \; 1} \end{bmatrix} \qquad {}_1(m_{N \; 1}, b_{N \; 1})$$

# Gradient Descent




$$(m, b) = \frac{1}{2m}\left[(m x_1 + b \quad y_1)^2 + \quad + (m x_n \quad y_n)^2\right]$$

$$\begin{bmatrix} m_N \\ b_N \end{bmatrix} \rightarrow \begin{bmatrix} m_N \\ b_N \end{bmatrix} = \begin{bmatrix} m_{N\ 1} \\ b_{N\ 1} \end{bmatrix} \qquad {}_1(m_{N\ 1}, b_{N\ 1})$$

DeepLearning.AI

# Gradients and Gradient Descent

## Conclusion