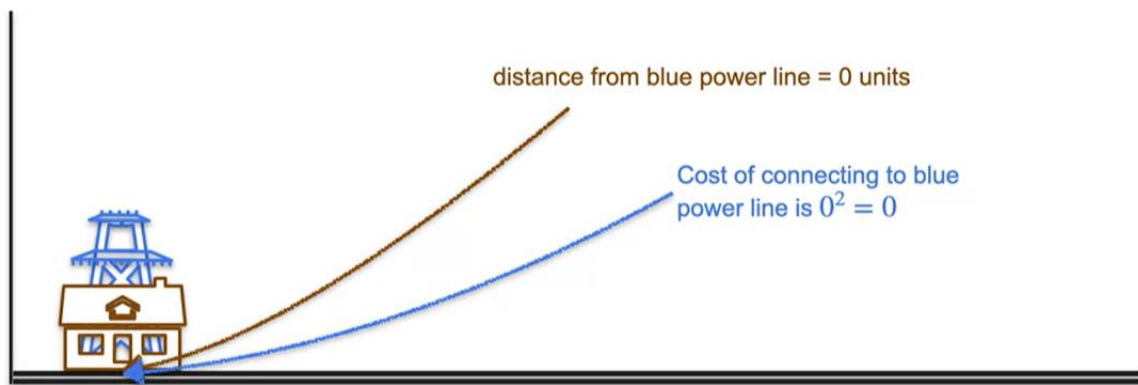


Optimization

Optimization is the process of finding maximum and minimum values given constraints using calculus. For example, you'll be given a situation where you're asked to find: The Maximum Profit. The Minimum Travel Time. Or Possibly the Least Costly Enclosure.

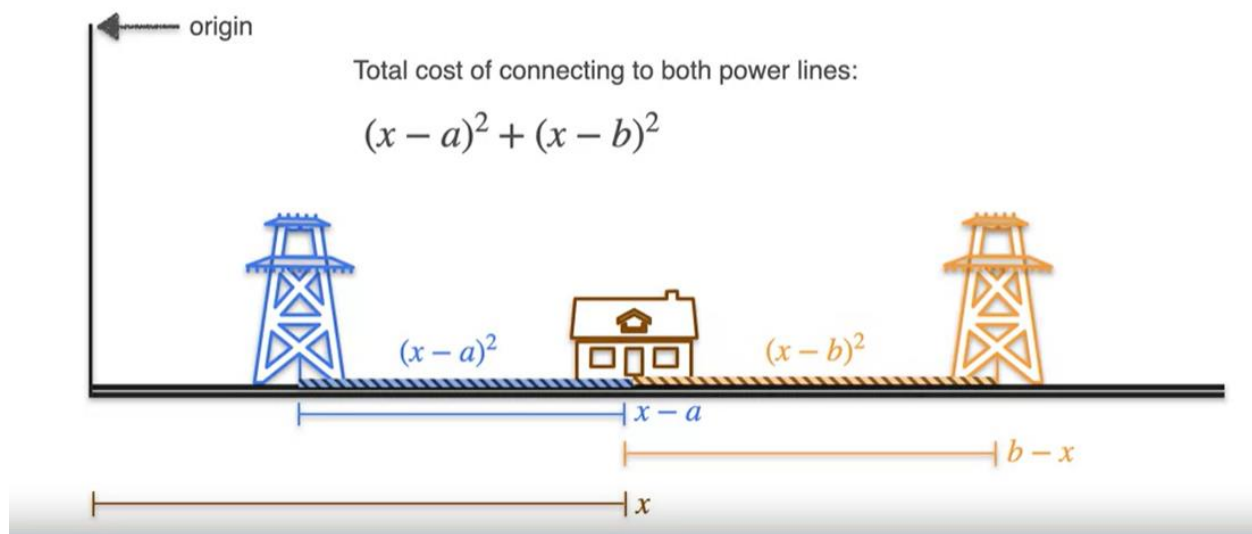
1. Optimization of squared loss - The one powerline problem

One Power Line Problem

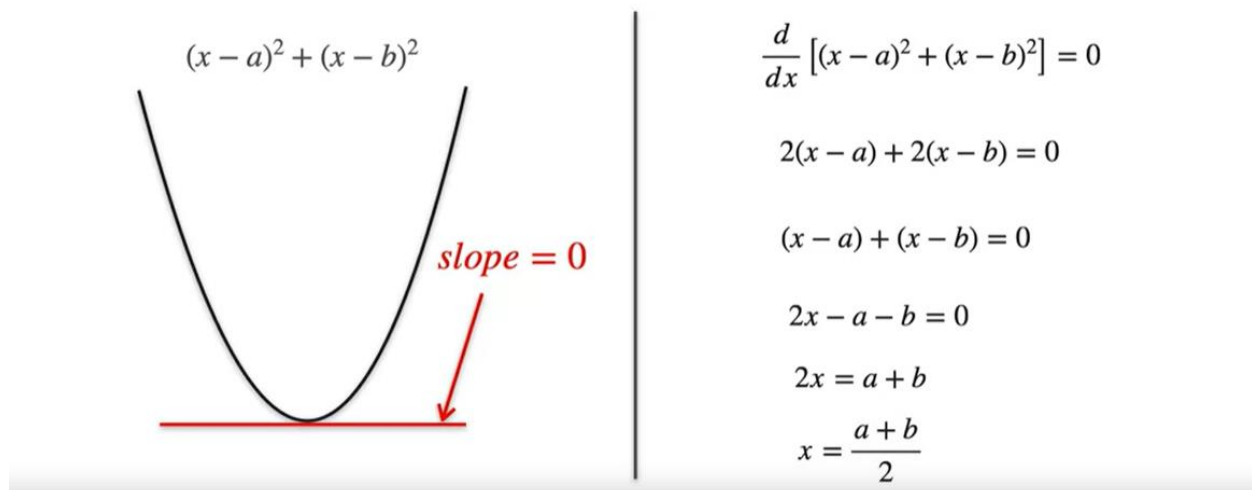


2. Optimization of squared loss - The two-powerline problem

Two Power Line Problem

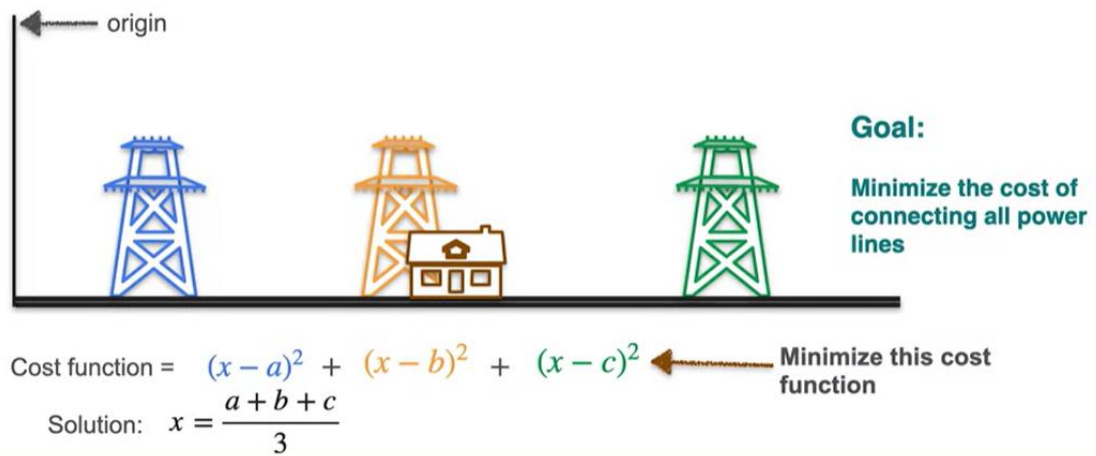


Two Power Line Problem

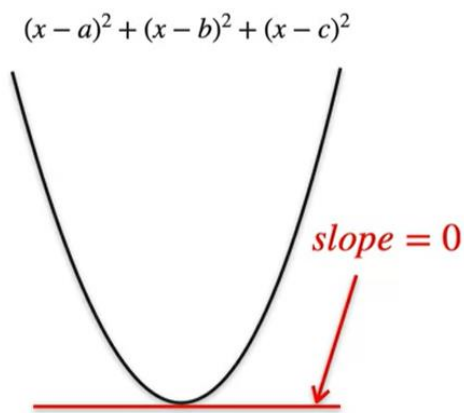


3. Optimization of squared loss - The three-powerline problem

Three Power Line Problem



Three Power Line Problem



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

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

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

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

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

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

4. Optimization of log-loss

Coins Toss



Coin 1 $0.7 \times 0.7 \times 0.7 \times 0.7 \times 0.7 \times 0.7 \times 0.7 \times 0.3 \times 0.3 \times 0.3 = 0.7^7 0.3^3$
 $= 0.00222$

Coin 2 $0.5 \times 0.5 \times 0.5 \times 0.5 \times 0.5 \times 0.5 \times 0.5 \times 0.5 \times 0.5 \times 0.5 = 0.5^7 0.5^3$
 $= 0.00097$

Coin 3 $0.3 \times 0.3 \times 0.3 \times 0.3 \times 0.3 \times 0.3 \times 0.3 \times 0.7 \times 0.7 \times 0.7 = 0.3^7 0.7^3$
 $= 0.00008$

Coin Toss



p $(1 - p)$

Chances of winning: $p^7(1 - p)^3 = g(p)$

Goal: maximize $g(p)$

Coin Toss

Product rule

$$\begin{aligned}
 \frac{dg}{dp} &= \frac{d}{dp}(p^7(1-p)^3) \stackrel{\text{Product rule}}{=} \frac{d(p^7)}{dp}(1-p)^3 + p^7 \frac{d((1-p)^3)}{dp} \\
 &= 7p^6(1-p)^3 + p^7 \boxed{3(1-p)^2(-1)} \quad \text{Chain rule} \\
 &= p^6(1-p)^2[7(1-p) - 3p] \\
 &= \boxed{p^6} \boxed{(1-p)^2} \boxed{7-10p} = 0
 \end{aligned}$$

\swarrow
 $p = 0$ $\rightarrow p = 1$ $\rightarrow p = 0.7$

Coin Toss

$$\begin{aligned}
 \log(g(p)) &= \log(p^7(1-p)^3) = \log(p^7) + \log((1-p)^3) \\
 &= 7\log(p) + 3\log(1-p) = G(p)
 \end{aligned}$$

$$\begin{aligned}
 \frac{dG(p)}{dp} &= \frac{d}{dp}(7\log(p) + 3\log(1-p)) = 7\frac{1}{p} + 3\frac{1}{1-p}(-1) \\
 &= \frac{7(1-p) - 3p}{p(1-p)} = 0
 \end{aligned}$$

$$7(1-p) - 3p = 0 \quad p = 0.7$$



Why the Logarithm?

1. Derivative of products is hard, derivative of sums is easy

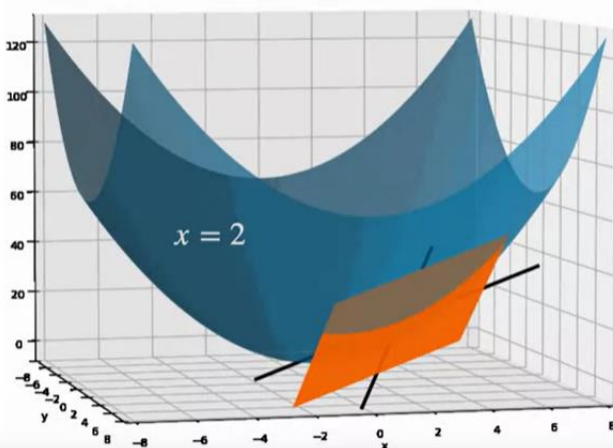
$$f(p) = p^6(1-p)^2(3-p)^9(p-4)^{13}(10-p)^{500}$$
$$\frac{df}{dp} \quad \text{😓} \quad \rightarrow \quad \frac{d}{dp} \log(f) \quad \text{😊}$$

2. Product of lots of tiny things is tiny!

Gradients

1. Introduction to tangent planes

Finding the Tangent Plane



$$\text{Fix } y=4 \quad f(x,4) = x^2 + 4^2$$
$$\frac{d}{dx} (f(x,4)) = 2x$$

$$\text{Fix } x=2 \quad f(2,y) = 2^2 + y^2$$
$$\frac{d}{dy} (f(2,y)) = 2y$$

The tangent plane contains both tangent lines.

2. Partial derivatives

Intro To Partial Derivatives

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

TASK

Find partial derivative of f with respect to x

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

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

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

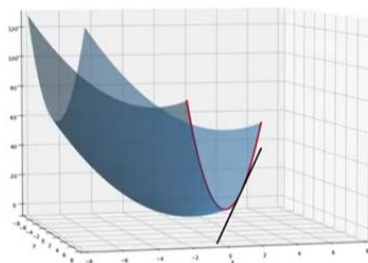
Step 2: Differentiate the function using the normal rules of differentiation.

3. Gradients

Gradient

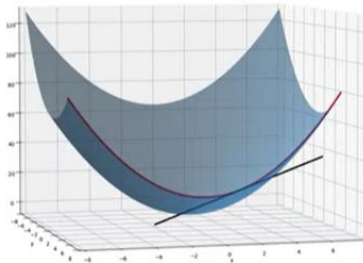
$$f(x, y) = x^2 + y^2$$

Treat y as a constant



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

Treat x as a constant



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

Gradient

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

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

4. Gradients and maxima/minima

Functions of Two Variables

