Minimizing Cost Function

Our objective is make the error as tiny as possible /ˈpɒs.ə.bəl/, to reach that we have to use some useful concepts:

1. Partial derivatives
2. Gradient
3. Dot product
4. Directional derivative

Partial derivatives

Remember what a derivative means in one single variable, it's no more than the rate of change in every function point. A partial derivative is not really useful on its own, you'll see why. The definition is simple, make one of the variables constant and keep the other one as a variable. But, is this information really useful?

$$\frac{\partial}{\partial x} f(x, y) \, ; \, \frac{\partial}{\partial y} f(x, y)$$

Gradient

This is another definition that's not really handy on its own. It is just a vector composed by the partial derivatives from a function.

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

Directional derivative

The directional derivative is the real analogous /əˈnæləgəs/ to the derivative in one variable. It means I know the rate of change in x and y axis (partial derivatives) of every point, now just state the direction you wanna go (unit vector "v"). If want to know the direction of the maximum rate of change just we must make the cos function equal to 1. Why have we done that? because our objective is to find the minimum rate of change, so, just multiply the maximum by -1.

$$v = (v_x, v_y)$$

$$\left(\frac{\partial}{\partial s} f \right)_{v,P} = \frac{\partial}{\partial x} f * v_x + \frac{\partial}{\partial y} f * v_y$$

$$\left(\frac{\partial}{\partial s} f \right)_{u,P} = \nabla f \, . \, v$$

$$\nabla f \, . \, v = \|\nabla f\| * \|v\| * \cos(\theta)$$

$$\nabla f \, . \, v_{\max} = \|\nabla f\|$$

$$v_{\max} = \frac{\nabla f}{\|\nabla f\|}$$

$$v_{\min} = -n * \frac{\nabla f}{\|\nabla f\|}$$

$$x_{\text{final}} = x_{\text{inicial}} - n * \frac{\nabla f}{\|\nabla f\|}$$