Minimizing Cost Function

Our objective is making the error as low as possible /ˈpɒs.ə.bəl/. For that we have to use some useful concepts:

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

Partial derivatives

Do you remember what a derivative at one single variable is?. It's no more than the rate of change at each point of a function. A partial derivative is making one variable constant, keep the other one as a variable and find the rate of change respect to that variable. Phisically, this means to make a section (where the variable works as a constant) and find the derivative on that section for each point of that function.

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

Gradient vector

The gradient vector doesn't represent anything by itself. It is just a vector composed of the partial derivatives of the function we are working on.

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

Directional derivative

A directional derivative is giving a direction to the gradient vector, so that we know where we are going to. For that, we will state the vector "v".

$$v = (v_x, v_y)$$

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

It could be written as the dot product of v and the gradient vector.

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

Our objective is to find the minimum rate of change, because that means if we follow that direction we will be closer to find the global or a local minimum of the function. What we're going to do is finding the maximum rate of change and then we will change its direction.

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

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

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

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

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

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