1 Context

We are trying to find the parameters for a linear function, notably $y = ax+b$, that passes as close to all the points in our dataset, meaning an a and a b that minimize this function:

$$J(a,b) = \frac{1}{2m} \cdot \sum_{i=1}^{m} [a \cdot x_i + b - y_i]^2$$

with the x_is and the y_is being the points of our dataset. This function is our cost function and it is a convex function, meaning it has a single minima over R^2 for a and b. The proof of that is outside the scope of this explanation (and not that i have done it) but one can easily imagine that it is..

That minima happens to also be where the derivative is equal to zero, so the question turns into, what is the root of the derivative of $J(a,b)$. And that "simply", we now just have to find the root of the derivative of $J(a,b)$!

2 Problematic

There are several iterative methods that finds the root of a function, we'll be using Newtons' algorithm. Which goes as this:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

The problem however is that our cost function is multi-variable, it has a and b as input values. Meanwhile Newtons' algorithm is set for uni-variable functions.

3 Solution

We can work around this trough the partial derivative, and it would work in the following manner: We first fix one of the parameters, let say b, with a random initial value. We then apply Newtons' algorithm to just a with that fixed value of b. This would give us an approximate value for the minima of the cost function for that specific value b. And then we switch, with our value of a, which will now be fixed, we'll apply Newtons' algorithm to approximate the minima for the fixed a. And we repeat.

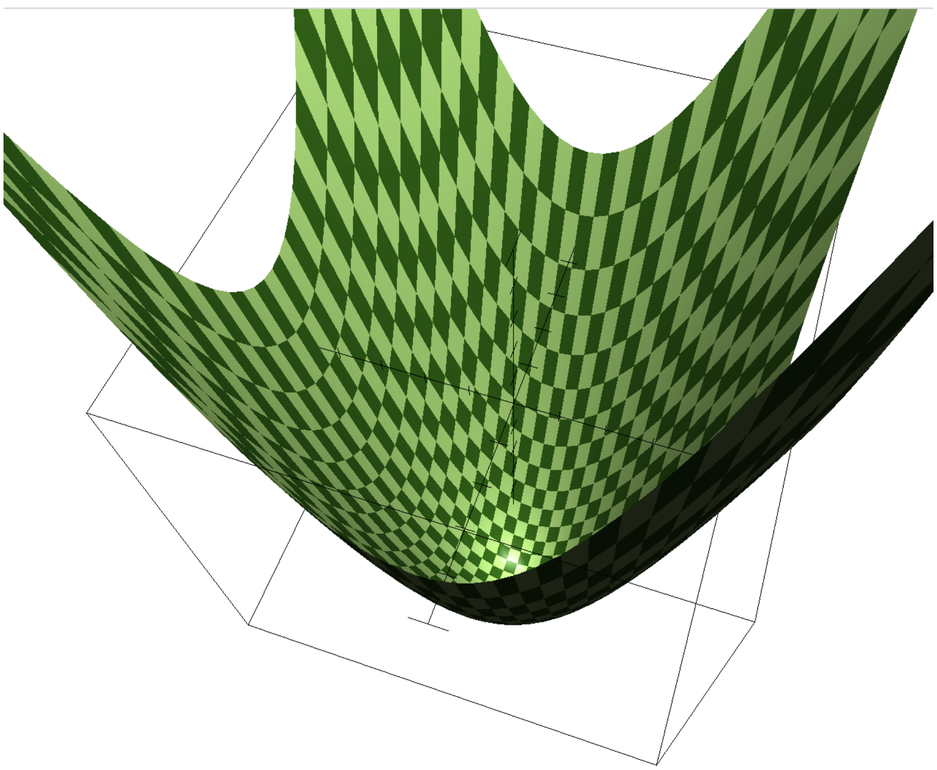

Figure 1: Representation of the cost function

You can maybe see, now that you have a visualisation of what the graph of the cost function, which is in function of a and b, may look like, how with every iteration we are approaching the minima of either a or b on a specific plane of the function. Combining the two together, approximates the minima of the function as a whole.

4 How-to solution

Alright, we have our solution, now let's apply it, first let's get the partial derivative of J, our cost function, relative to a

$$\frac{\partial J}{\partial a} = \frac{1}{m} \cdot \sum_{i=1}^{m} [x_i \cdot (a \cdot x_i + b - y_i)]$$

And the partial derivative relative to b:

$$\frac{\partial J}{\partial b} = \frac{1}{m} \cdot \sum_{i=1}^{m} [a \cdot x_i + b - y_i]$$

Now that we have the function that we need to find the root for, we now need its derivative, which is, for a, the second partial derivative for a:

$$\frac{\partial^2 J}{\partial^2 a} = \frac{1}{m} \cdot \sum_{i=1}^{m} \left[x_i^2\right]$$

And for b, the second partial derivative for b:

$$\frac{\partial^2 J}{\partial^2 b} = \frac{1}{m} \cdot \sum_{i=1}^{m} [1] = 1$$

So for a fixed b, we get:

$$a_{n+1} = a_n - \frac{\frac{\partial J}{\partial a}}{\frac{\partial^2 J}{\partial^2 a}}$$

Which we simplify to:

$$a_{n+1} = a_n - \frac{\frac{1}{m} \cdot \sum_{i=1}^{m} [x_i \cdot (a \cdot x_i + b - y_i)]}{\frac{1}{m} \cdot \sum_{i=1}^{m} [x_i^2]} = a_n - \frac{\sum_{i=1}^{m} [x_i \cdot (a \cdot x_i + b - y_i)]}{\sum_{i=1}^{m} [x_i^2]}$$

And for a fixed a, we get:

$$b_{n+1} = b_n - \frac{\frac{\partial J}{\partial b}}{\frac{\partial^2 J}{\partial^2 b}}$$

Which is:

$$b_{n+1} = b_n - \frac{\frac{1}{m} \cdot \sum_{i=1}^{m} [a \cdot x_i + b - y_i]}{1}$$

Finally, as to gain some times in the computation, we pre-calculate some constants once. So the final form of our equations becomes:

$$a_{n+1} = a_n - \frac{a \cdot X_.^2 + b \cdot X_. - X_.Y_.}{X_.^2}$$

$$b_{n+1} = b_n - \frac{\frac{1}{m} \cdot [a \cdot X_. + b \cdot m - Y_.]}{1}$$

With:

$$X_.^2 = \sum_{i=1}^{m} \left[x_i^2\right]$$

$$X_. = \sum_{i=1}^{m} [x_i]$$

$$Y_. = \sum_{i=1}^{m} [y_i]$$

$$X_.Y_. = \sum_{i=1}^{m} [x_i \cdot y_i]$$

There is still room for optimization, but that's good enough for my needs..

And that's that, we simply pick an arbitrary initial value for a and b, we iterate once over a with that fixed b, then we iterate once over b with the new a, and we repeat!

HEMMOUDA Aymane