Gradient Descent APL

Lakshmiram S
EE22B117

October 2023

Introduction

We are given a cost function and our goal is to locate the global extremum. We employ the gradient descent approach to tackle this. A few points to remember are:

1. The cost function(the function, that needs to be optimised) should be continuous and differentiable.
2. For some cost functions, finding the derivative analytically might be difficult and numerical methods have to be employed in such cases.
3. The starting point plays a crucial role. If a function has multiple extrema, choosing the right starting point is important, since gradient descent always looks for local extrema(while we want global).
4. Making the step size proportional to the gradient is another key strategy, which ensures that the process converges to the extrema.
5. We have to find the sweet spot for the value of the learning rate- too less a value will take a lot of iteratinon to generate a meaningful optimum and too large a value may not result in a convergent value.

Question 1

I have written a function that takes in a bunch of parameters and returns the optimum value along with an image that denotes the same.

here are the results:

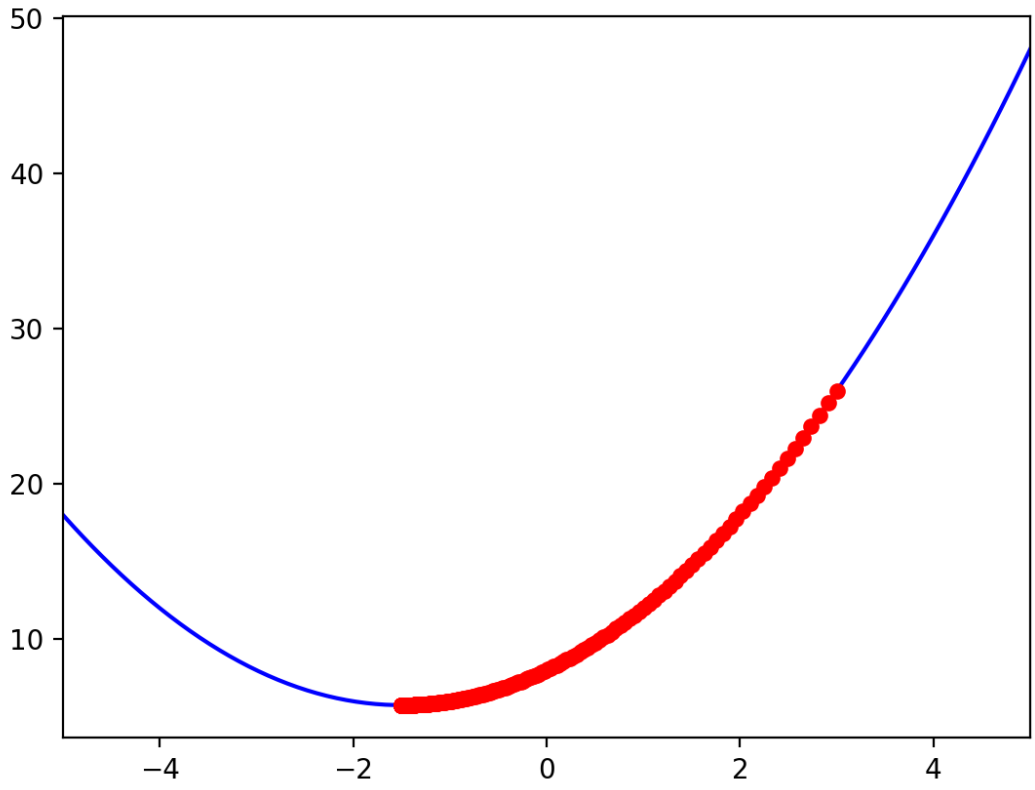

Figure 1: problem 1 gradient descent

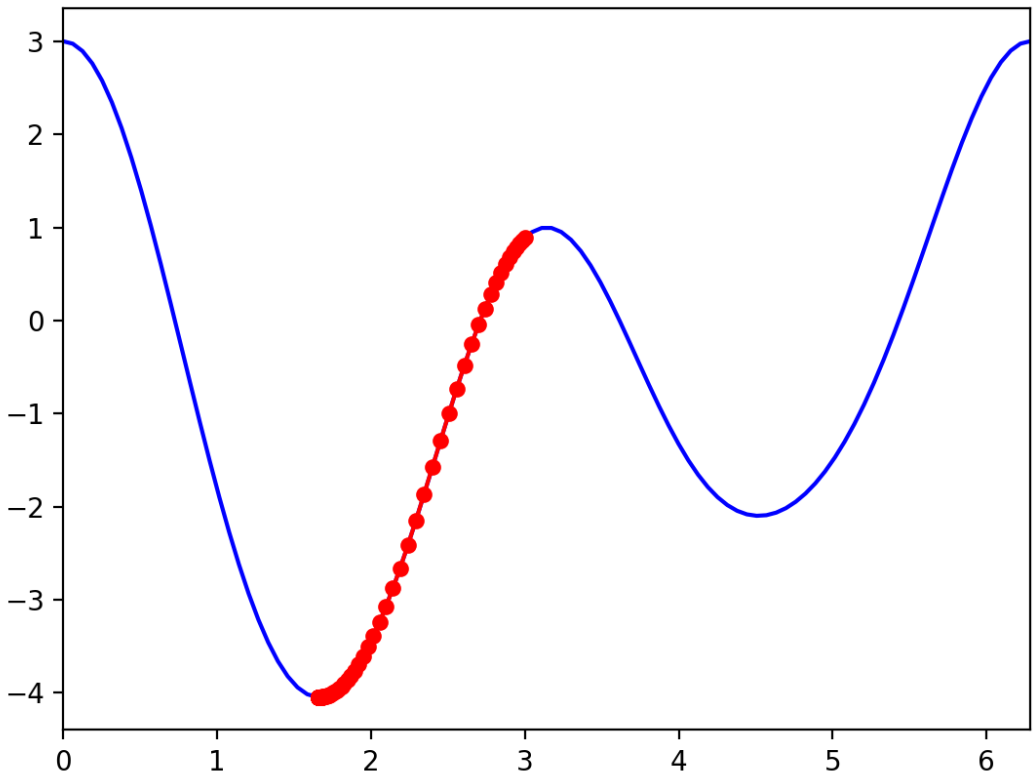

Figure 2: problem 4 results

Question 4

The same gradient descent function is used upon the new cost function to generate results.

Question 2

The domain here is a region in the 2-d plane. The only difference between the 1d and the 2d domains is the fact that both the x-coordinate and the y-coordinate have to be updated simultaneously to take a step towards the optimum.

Question 3

The approach is the same as in question 2, just a different cost function. Here it becomes imperative to choose the right starting point, as not all starting points lead to the global extremum.

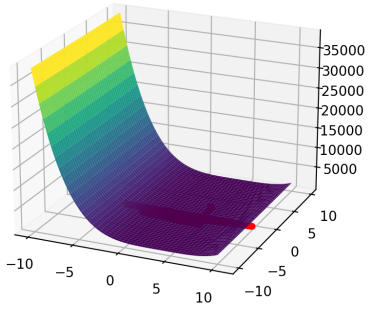

Figure 3: problem 2 results

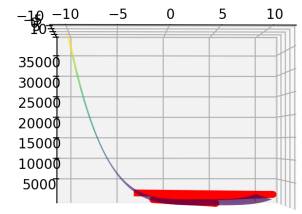

Figure 4: problem 2 results

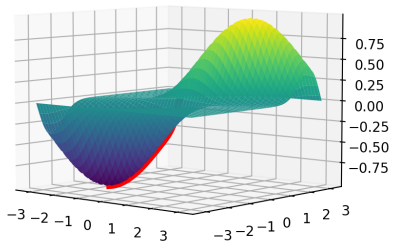

Figure 5: problem 3 results

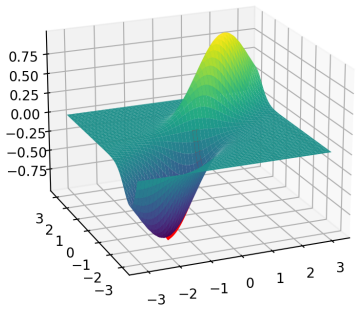

Figure 6: problem 3 results