Week 3 Homework Submission
Optimization

Iuliia Skobleva
Matriculation Number 03723809

November 24, 2019

Optimization / Gradient Descent

Problem 2

The following function is given:

$$f(x_1, x_2) = 0.5x_1^2 + x_2^2 + 2x_1 + x_2 + cos(sin(\sqrt{\pi})) \tag{1}$$

First we compute the minimum of this function. It would be helpful to know is this function is convex since then the found extremum will automatically be the minimum. The Hessian of this function equals:

$$H_{f(x_1,x_2)} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \tag{2}$$

This matrix is positive definite and therefore the function f is convex on $\Re$. To find the extremum, we set the gradient of f to zero and obtain the following system of equations:

$$\begin{aligned} x_1 + 2 &= 0 \\ 2x_2 + 1 &= 0 \end{aligned} \tag{3}$$

we find that the minimum of the function f is obtained at $(-2, -0.5)$. In the next step we perform 2 steps of gradient descent. We have found

$$\nabla f(x_1, x_2) = \begin{pmatrix} x_1 + 2 \\ 2x_2 + 1 \end{pmatrix}$$

and staring with $\mathbf{x_0} = (0, 0)$ in the first step we have:

$$\mathbf{x_1} = \mathbf{x_0} - \tau \nabla f(\mathbf{x_0}) \tag{4}$$

where $\tau = 1$. The results equals $\mathbf{x_1} = (-2, -1)$. In the second step:

$$\mathbf{x_2} = \mathbf{x_1} - \tau \nabla f(\mathbf{x_1}) \tag{5}$$

which results in $\mathbf{x_2} = (-2, 0)$. This result tells us that we cannot find the minimum of the function in two steps. Since the learning rate is too large we will never be able to find it because the solution will oscillate around the minimum. To solve this problem we could set $\tau = 0.5$.

Problem 4

We consider the following covex function:

$$f(x_1, x_2) = e^{x_1 + x_2} - 5 \log x_2 \tag{6}$$

We first consider the shaded region S and see that it is not convex. For example, if we take the points $(3.5, 1.5)$ and $(5.5, 2.5)$, we see that the line connecting them does not lie fully in S.

The next question is to find the maximizer $\mathbf{x}^*$ over S. From the lecture we know that "maximum over a convex function on a convex set is obtained on a vertex". The convex hull of S can be obtained by connecting the 4 vertices $(1, 3.5), (3.5, 6), (6, 3.5), (3.5, 1)$. Now we can evaluate the function at these vertices and the largest value will correspond to the maximum of f. This can be seen in the Jupyter Notebook under 'Problem 4'. We find the maximizer $\mathbf{x}^* = (6, 3.5)$.

In the last question we are given an algorithm $ConvOpt(f, D)$ that takes as input a convex function f and convex region D, and returns the minimum of f over D. If we were to find a minimum of f over S we could do the following:

- Divide S into five convex regions with a square in the middle.
- Apply $ConvOpt(f, D)$, where D is one of the five convex regions within S.
- Compare the minima obtained in the previous step to find the global minimum of f over S.