Homework 1

Problem 1. The PYTHON example for gradient descent had a cost function given by

$$C(x,y) = \frac{1}{2}(\cos^2(x) + \sin^2(y)).$$

On the plot of the gradient descent path *all* possible (local) minimum points in the region $-5\pi/2 \leq x \leq 5\pi/2, -2\pi \leq y \leq 2\pi$ were also shown.

(a) Show that all of these points were local minimums (not local maximums).

Hint: Look at the Hessian of $C(x,y)$.

(b) Modify the PYTHON code for gradient descent to show all of the (local) maximums in the region $-5\pi/2 \leq x \leq 5\pi/2, -2\pi \leq y \leq 2\pi$ (instead of the local minimums). Also modify the code so that it computes the gradient *ascent*. That is, starting at a random point in the region $-5\pi/2 \leq x \leq 5\pi/2, -2\pi \leq y \leq 2\pi$ the path converges to a local *maximum*. Submit your code and your output.

Problem 2. Download from Canvas the code given in the module **Chapter 1 code for MNIST**. Run the program that implements the three level neural network. Then run the code that plots the weights.

(a) First run these files on your PC/laptop to see that they work. Submit your outputs.

(b) Run these files on the Borah server. Go to the module **Borah server cluster** on Canvas to get the procedure to run there. I found the queue was fast if I used the cpu procedure rather than the gpu procedure (apparently a lot jobs request a gpu). I have sent a request to OIT for all of you to get accounts on Borah. The point of this problem is make sure you are able to run on Borah. Submit your outputs.