

Math 5364 Homework 14

1. Use gradient descent to minimize the function $f(x, y) = x^2 + y^2 + y$ on the open disk $x^2 + y^2 < 1$. This is essentially the example on p. 7 of the Lagrange multipliers slides, except that we are only interested in interior points on this problem.

(Hints: Start by randomly selecting a vector (x, y) in the disk, and then apply gradient descent. If the vector escapes the disk during the algorithm, randomly select a new vector. Store vectors that converge to a minimum in a matrix, and repeat the entire process a large number of times to ensure that there is only one local minimum inside the disk. You may need to take some care with the learning rate to prevent vectors from escaping the disk too often.)

2. Compare the merits of using a 2-layer neural network (with 1 hidden layer) and a multi-layer neural network to the `wdbc` data set. Let's say you only have 10 minutes to train a network with `nnet` or `mlp`. Which approach produces the model with the higher classification accuracy? It may be interesting to compare the two packages for a variety of time frames and to consider different network topologies for `mlp`, e.g., `size=c(2, 4)` vs. `size=c(2, 2, 2)`.