

Ryan Farr – rlf238

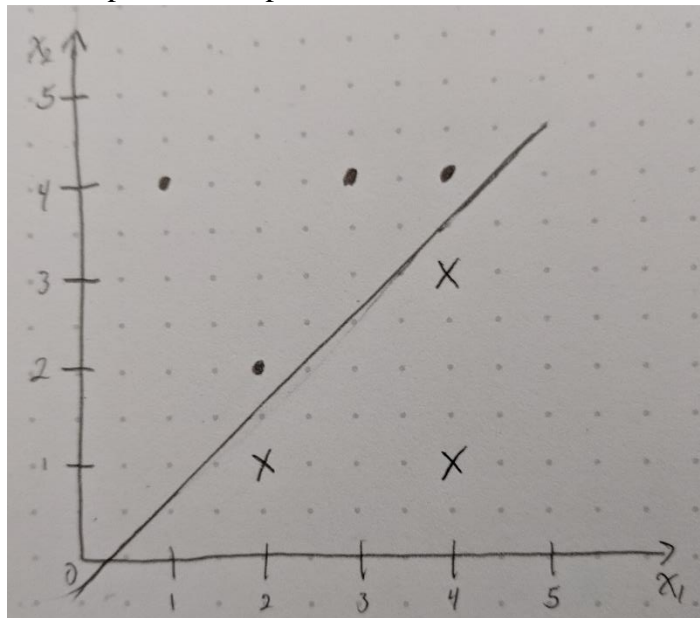
November 25th, 2018

CS 5785 – Applied Machine Learning

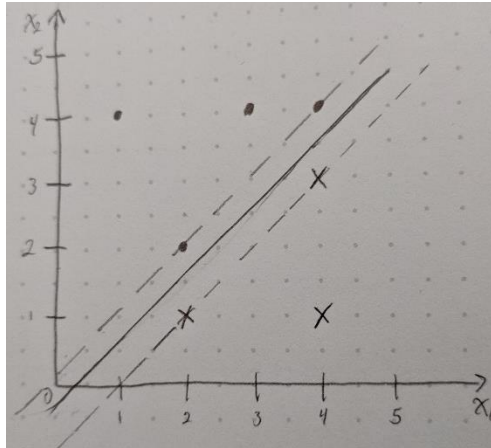
Homework 4 Written Exercises

1. Maximum-margin classifiers

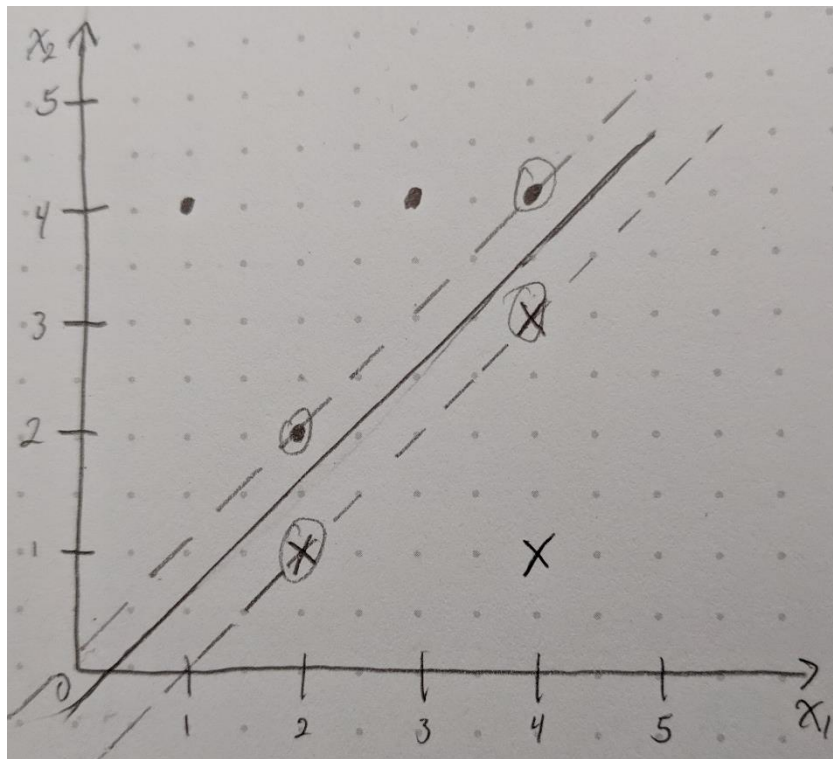
- a. Below is a sketch of the observations and maximum-margin separating hyperplane. The blue points are represented with dots and the red with 'x' marks.



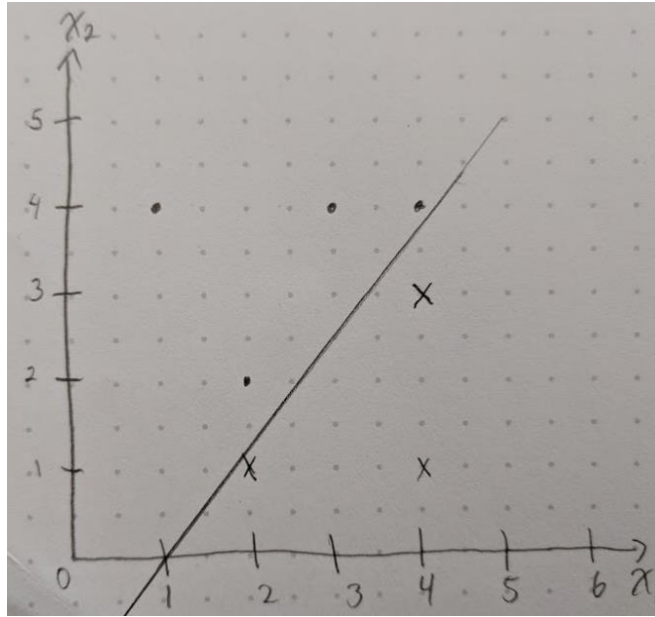
- b. As can be seen, the maximum-margin line should fall between the blue (2,1) and (4,3) points and the red (2,2) and (4,4) points. This implies it should have a slope of one and fall halfway between the lines defined by the two red and two blue points. The classification rule then is red if $0.5 - X_1 + X_2 > 0$ and blue otherwise.
- c. The margin is defined by the red and blue points listed in part b and is shown below.



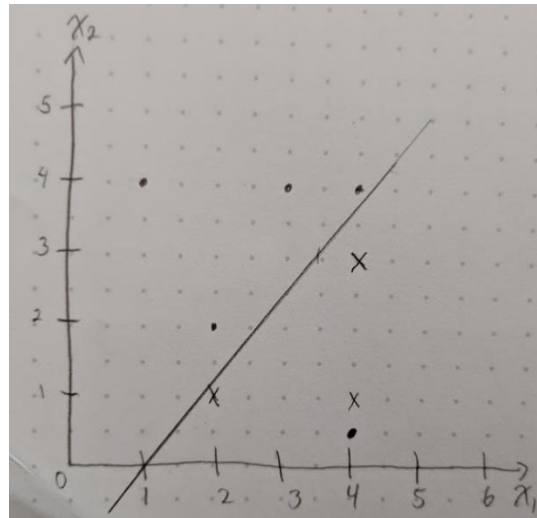
- d. The support vectors are the red and blue points found on the margin and are circled below.



- e. A small change in location of the seventh point (4,1) would not affect the maximal-margin hyperplane because it is not found on or near the margins. Thus, it could move almost anywhere in the lower right without affecting the hyperplane so long as it doesn't become a support vector or cross the current hyperplane.
- f. The hyperplane shown below separates the data but is not the maximum-margin separating hyperplane. The line is defined as $y = \frac{6}{5}x - \frac{6}{5}$.



- g. By drawing a red point at (4, 0.5) the classes can no longer be separated in two dimensions.



2. I've designed this neural network so that it has two hidden layers, one with four nodes and one with one node. A diagram of the network is shown below and inspired by *Neural Networks and Deep Learning* Chapter 4 from hint 2. The values inside each node are the biases and the values on each line are the weights. The first hidden layer takes care of all input values from 0 to 9 and the last layer ensures that from 9 to 10 we give 0 as output. The biases on the first layer of hidden nodes ensure that the weights on the outgoing edges will not affect values of x smaller than or equal to the given bias. The bias was then selected to ensure that the sum of the outgoing positive values equals a given value. For

example, selecting the bias for the first node in the first hidden layer was chosen to ensure that the output goes from 0 to 2 between $x = 1$ and $x = 2$.

