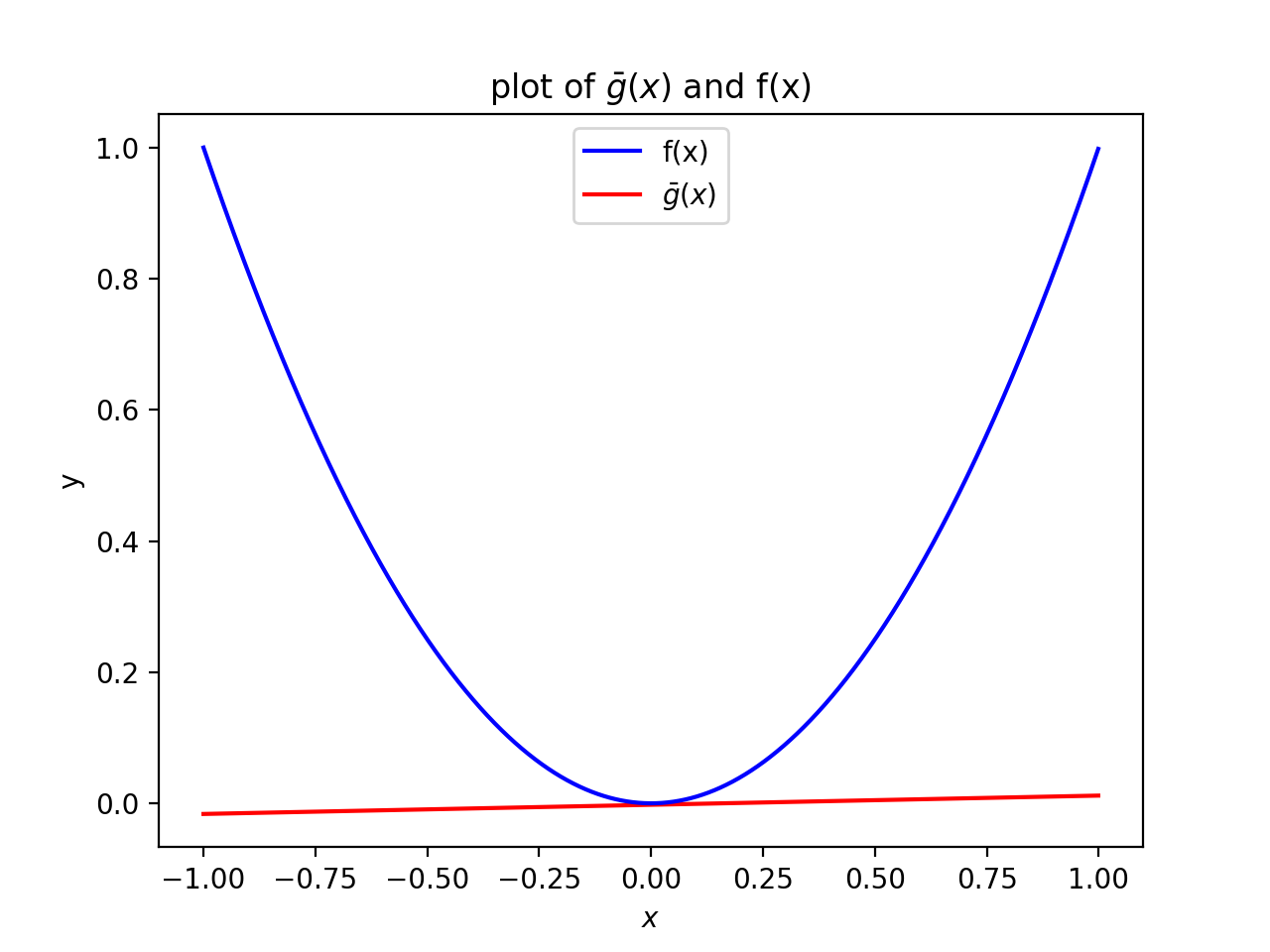
CSCI 4100 Assignment 5

Boliang Yang 661541863

1. Exercise 2.8 in LFD
2. By definition, , then is a linear combination of . Since H is closed under linear combination, then .
3. The simple model is a binary classification. If the model’s hypothesis set contains only two hypotheses: one that always returns +1 and one that always returns -1. Then for dataset that generated +1 and -1 with equal probability, is 0 which clearly will not be in .
4. No, will most likely to be somewhere between -1 and +1, and thus cannot be a binary function.
5. Problem 2.14 in LFD
6. From problem we know, each has break point , then for all where . Then for , . Since , we have .
7. Since for all where , then . Since , then . We have , and it follows that .
8. Assume be the value , and we get:

The inequality holds since the left-hand-side grows much faster than the right-hand-side: exponent in the left-hand-side is , which is much greater than exponent in the right-hand-side. Then satisfies From (a)&(b), , where . Thus, .

1. Problem 2.15 in LFD
2. One example of monotonic classifier in two dimensions is : when returns +1, when returns -1. Since for any two points , , if , we have then the classifier is monotonic.
3. We first choose one point, and then generate the next point by increasing the first component (x) and decreasing the second component (y) in order to keep x+y unchanged until N points are obtained. In this case, all the points can always be shattered and all dichotomies exist. Then , .
4. Problem 2.24 in LFD
5. Since and hypothesis is of the form , then we have slope , and intercept . Since input variable is uniformly distributed in , then expected value . Thus, .
6. First generate 1000 points data set with target function where . Then compute hypothesis based on each point. Now we can compute the average function using . We can compute bias as compute var as , compute as .
7. After running on 1000 points data set, the result is shown as following:

, , ,

1. Since , then

.

.