ML Homework 4

Liheng Cao

March 21, 2021

1

When w_0 increases, then the probability of predicting class 1 becomes higher, and vice versa.

$$h(x) = \frac{1}{1 + e^{-(w_0 + \cdots)}}$$

This is because as w_0 gets bigger, the result of exponentiation gets smaller. As the denominator gets smaller, the quotient gets larger. $h(x)$ is used as the probability we will predict a certain class. The higher $h(x)$ is, the more likely we will predict class 1 (as opposed to class 0).

2

The logistic function is

$$\frac{1}{1+e^{-\mathbf{w}^T\mathbf{x}}}$$

If we were to double $\mathbf{w}$, the result wouldn't actually change, as long as the threshold is 0.5.

The geometric interpretation is because $\mathbf{w}^T\mathbf{x}$ represents a (hyper)plane. Multiplying the equation of a plane doesn't affect any points that are on the plane (threshold $= 0.5 \implies$ on hyperplane), but it makes the points that are off it seem further away, increasing the certainty with which we pick a class. For example, if $\mathbf{w}^T\mathbf{x} = 0$, then there is no change ($0 \cdot 2 = 0$). But if it's positive, it becomes a larger positive number, and vice versa.

3

(a)

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

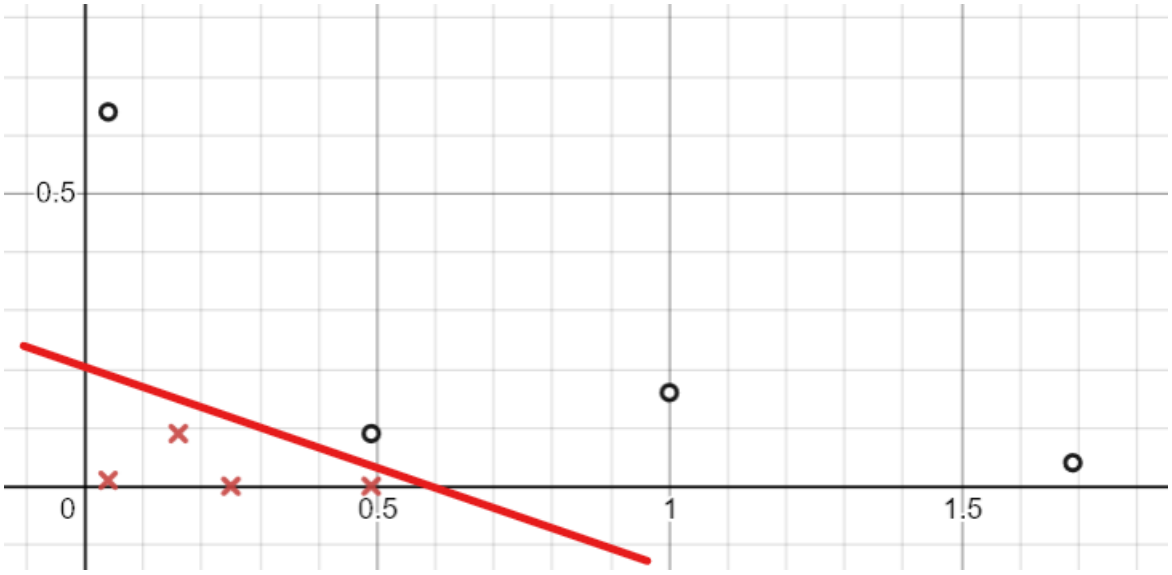

Figure 1: o: class 0, x: class 1

(b)