Problem set 2
Classification algorithms

Excercise 1

In the lecture we saw how Perceptron updates the weights and the threshold if it misclassifies a positive object (i.e. when $y = +1$). Conduct a similar analysis when the algorithm misclassifies a negative instance (i.e. when $y = -1$).

Solution Updated Weight Vector $W' = W + YX$
Updated Activation $a' = W'^T X = (W + YX)X = (W - X)X$, since $Y = -1$
Hence $a' = WX - X^2 = a - X^2$ $[a = WX]$

Hecne $a' < a$.

Excercise 2

Provide a geometric interpretation for the update rule in Perceptron when a negative instance is mistaken to be positive (in the lecture we conidered the case when a positive instance is mistaken to be positive).

Solution

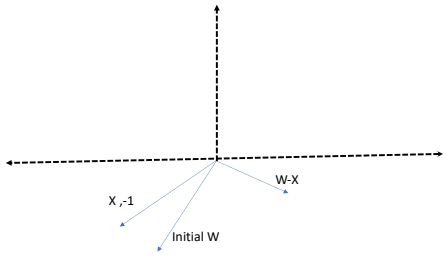

Figure 1: Caption

Excercise 3

Let us consider the loss function $(y - \overline{W}^T \overline{X})^2$. Given a train dataset $D = \{(\overline{X}_i, y_i)\}_{i=1}^N$, we define the loss of classifying an instance $(\overline{X}_n, y_n)$ by $y_n -$

$\overline{W}^T \overline{X}_n)^2$. Here, $y_n \in \{1, -1\}$ is the target label of the instance $\overline{X}_n$. Answer the following questions.

1. Compute the derivative of the loss function w.r.t W.

2. Write the gradient descent rule for obtaining a new vector $\overline{W}^{(t+1)}$ from the current weight vector $\overline{W}^{(t)}$ after observing a train instance $(\overline{X}_n, y_n)$. Assume step-size to be γ.

   Solution

   - Derivative is $\frac{\partial (y - W^T X)^2}{\partial w} = 2(Y - W^T X)(-X)$
   - Updation rule is $W^{t+1} = W^t + 2(Y - W^{t^T} X)(X)$