Homework 2 Solutions

1. **Perceptron Algorithm.**

   (a) Figure 1 shows the problem in weight space. The constraints corresponding to each training examples are:

$$w_1 - 2w_2 > 0$$
$$-w_2 < 0$$

   The orange half-plane shows the first constraint. The blue shows the second one.

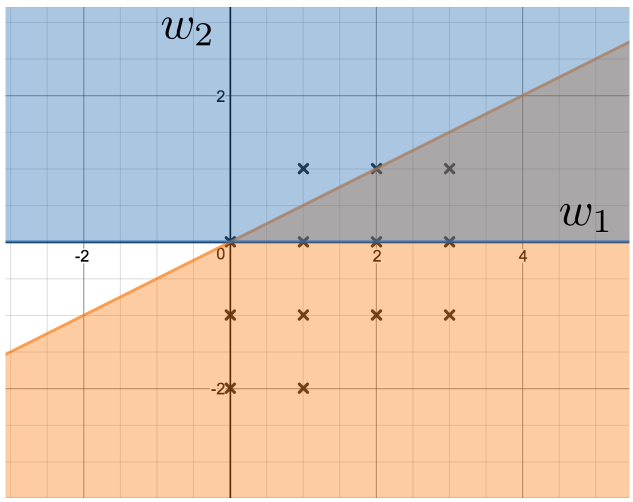

Figure 1: Feature Space

   (b) Figure 1 shows the points in weight space visited by the perceptron algorithm.

2. **Feature Maps.**

   (a) This dataset is not linearly separable because the positive half-space needs to be a positive set. But then, the training point $x = 1$ would have a postiive label since it is between two positive training points $x = -1$ and $x = 3$.

   (b) The constraint on w_1 and w_2 are:

$$-w_1 + w_2 > 0$$
$$w_1 + w_2 < 0$$
$$3w_1 + 9w_2 > 0$$

3. **Loss Functions**

   Let N denote the number of training examples, or the number of columns of the design matrix $\mathbf{X}$. Then:

$$
\begin{aligned}
\mathbf{y} &= \mathbf{X}\mathbf{w} + b\mathbf{1} \\
\frac{\partial \mathcal{E}}{\partial \mathbf{y}} &= \frac{1}{N}\sin(\mathbf{y} - \mathbf{t}) \\
\frac{\partial \mathcal{E}}{\partial \mathbf{w}} &= \mathbf{X}^T \frac{\partial \mathcal{E}}{\partial \mathbf{y}} \\
\frac{\partial \mathcal{E}}{\partial b} &= \mathbf{1}^T \frac{\partial \mathcal{E}}{\partial \mathbf{y}}
\end{aligned}
$$

Where $\mathbf{1}$ is a column vector of 1's with dimension N. (Full credit is also given if you expanded out the formulas explicitly.)