Assignment 2

Rahul Krishna
rkrish11@ncsu.edu

February 13, 2014

3.1. To show that the perceptron loss function is a continuous piecewise linear function.

The perceptron loss function is given by,

$$J(w) = \sum_{x \in Y} (\delta_x w^T x) \tag{1}$$

If we use the gradient descent method to minimize the cost function, the weight will change according to following relation,

$$w(n+1) = w(n) - \rho_n \sum_{x \in Y} \delta_x x \tag{2}$$

Here, n is the number of iterations. The variable ρ_n determines the rate at which the algorithm converges to the local minima, the global minima or the saddle point. When we change the weight smoothly the cost function changes in a linear manner. Once the number of misclassified vector changes, the limits of the summation – denoted by x∈Y– changes and so does the cost function. Therefore, as long as the number of misclassified vectors remain the same the cost function changes linearly with the weight; once the number of misclassified vectors changes, the cost function changes altogether. This makes it a piecewise linear function.

3.4 To compute a hyperplane using a perceptron algorithm

We have 4 feature vectors and 2 classes ω_1 and ω_2. We'll first extend the feature vectors to a 2D space as below,

$$\omega_1 = [0,\ 0,\ 1]\ and\ [0,\ 1, 1]$$
$$\omega_2 = [1,\ 0,\ 1]\ and\ [1,\ 1, 1]$$

The initial weight at iteration 0 is given by $w(0) = [0, 0, 0]$. Also, $rho = 1$. To apply the perceptron algorithm using the reward and punishment form, we use the following 3 fundamental equations.

$w(n+1) = w(n) + px_n$ if $x_n \in w_1$ and $w^T(n)x_n \leq 0$
$w(n+1) = w(n) - px_n$ if $x_n \in w_2$ and $w^T(n)x_n \geq 0$
$w(n+1) = w(n)$ elsewise

Using these the following computations can be made,