HW02 STAT W4400

Ethan Grant uni: erg2145

February 23, 2016

1. (a) Based on the formula for classification in the slides x1 is classified as follows:
   $f(x_1) = sgn(< -3, 0 >< \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} > -\frac{1}{2*\sqrt{2}})$
   $f(x_1) = sgn(\frac{-3}{\sqrt{2}} - \frac{1}{2*\sqrt{2}})$
   $f(x_1) = -1$
   now x2:
   $f(x_2) = sgn(< 1/2, 1/2 >< \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} > -\frac{1}{2*\sqrt{2}})$
   $f(x_2) = sgn(\frac{2}{2*\sqrt{2}} - \frac{1}{2*\sqrt{2}})$
   $f(x_2) = sgn(\frac{1}{2*\sqrt{2}})$
   $f(x_2) = 1$

   (b) Since the returned values are the same following the training(v_H and c) the fact that it was trained using an SVM with a margin is irrelevant for classification because the classification rule of seeing which side of the hyperplane defined by v_H and c the point falls on is the same. Thus the classes are the same for both.

   (c) The perceptron cost function approximates the empricial risk function. This must be approximated because the empirical risk function is piece wise constant. This meanss that there is no way to minimize the empirical risk function with respect to z b/c one the derivative of a piecwise constant function is 0 across its whole domain. Setting the derivative of the empirical risk function equal to 0 is how one would determine the optimal point with a non piece wise constant function. Since that is impossible the function must be approximated.