Homework 8
MATH 123 - Spring 2023
Tufts University, Department of Mathematics
Due: April 5, 2023

Question 1

Download the data 'HW8_TwoClass.mat' to reveal data $\{x_i\}_{i=1}^n \subset \mathbb{R}^2$ together with labels $\{y_i\}_{i=1}^n$.

(a) Let $F(w,b) = \|w\|_2^2 + \lambda \sum_{i=1}^n \max(0, 1 - y_i(w^T x_i + b))$ be the hinge loss. Use MATLAB's black box optimization function 'fminunc.m' to estimate the hyperplane that minimizes the hinge loss for different choices of λ. Plot the results on the data and interpret.

(b) Let $G(w,b) = \|w\|_2^2 + \lambda \sum_{i=1}^n \max(0, 1 - y_i(w^T x_i + b))^2$ be the squared loss. Use MATLAB's black box optimization function 'fminunc.m to estimate the hyperplane that minimizes the hinge loss for different choices of λ. Plot the results on the data and interpret.

Question 2

Suppose the data $\{(x_i, y_i)\}_{i=1}^n$, $x_i \in \mathbb{R}^d, y_i \in \{-1, 1\}$ are linearly separable, i.e. there is a hyperplane such that all points with label -1 are on one side, all point with label -1 on the other side.

(a) Phrase the linear separability condition mathematically in terms of the parameters of the hyperplane.

(b) Prove that there are infinitely many separating hyperplanes, as soon as there is one.

(c) Let $F(w,b) = \|w\|_2^2 + \lambda \sum_{i=1}^n \max(0, 1 - y_i(w^T x_i + b))$ be the soft-margin hinge loss with regularization parameter λ. Describe what kind of hyperplane is learned by minimizing F for $\lambda \to 0$? For $\lambda \to +\infty$?

(d) Suppose (w^*, b^*) are the minimizing hyperplane parameters for a fixed choice of λ. How can we use w^*, b^* to classify a new, unlabeled test point x_{test}?