ECE1513H - Winter 2020 **Assignment 3**
University of Toronto - LEC0101 (Prof. Papernot) **Due Feb 24**

This assignment contains 11 questions worth a total of 16 points. Because assignments will help you learn things that are complementary to what we covered in class, they should be completed on your own. Otherwise, you will not learn from taking this course and you are harming yourself.

Problem 1 The goal of this problem is to derive the parameters of a support vector machine with a hard margin. We will consider the following toy dataset of 4 points: $x^{(1)} = (1,1)^\top$, $x^{(2)} = (-1,-1)^\top$, $x^{(3)} = (1,0)^\top$, and $x^{(4)} = (0,1)^\top$. $x^{(1)}$ is the only point from the positive class, whereas $x^{(2)}$, $x^{(3)}$, and $x^{(4)}$ are from the negative class. Put another way, $x^{(1)} = 1$ and $x^{(2)} = x^{(3)} = x^{(4)} = -1$.

1. (2 points) Plot the data and draw the maximum-margin hyperplane. Give the equation of this hyperplane.

2. (1 point) Which of the points are support vectors?

3. (1 point) Recall the optimization problem we defined for the SVM with a hard margin:

$$\min_{\vec{w},b} \frac{1}{2}\|\vec{w}\|^2 \ \texttt{s.t.} \quad \forall j(\vec{w} \cdot x^{(j)} + b)y^{(j)} \geq 1 \tag{1}$$

   Write down the Lagrangian $L(\vec{w}, b, \vec{\alpha})$ corresponding to this problem, where $\vec{\alpha}$ is the vector of dual variables.

4. (2 points) We can therefore rewrite Equation 1 as follows:

$$\min_{\vec{w},b} \max_{\vec{\alpha}} L(\vec{w}, b, \vec{\alpha})$$

   Because Slater's condition holds, this is equivalent to solving:

$$\max_{\vec{\alpha}} \min_{\vec{w},b} L(\vec{w}, b, \vec{\alpha}) \tag{2}$$

   Given a fixed $\vec{\alpha}$, solve the inner (minimization) problem. You should obtain two relationships taking the form of

$$A = \sum_j B_j C_j \vec{D}_j$$

$$\sum_j E_j F_j = G$$

5. (2 points) Simplify Equation 2 with the result of question 4. Show that it takes the form of

$$\max_{\vec{\alpha}} \sum_j A_j - \frac{1}{2} \sum_j \sum_k B_j C_k D_j E_k \vec{F} \cdot \vec{G}$$

6. (1 point) We now come back to our toy dataset. Because it has 4 trainings points, we have $j \in 1..4$, i.e., four dual variables. For points that are not support vectors, recall from class that the corresponding constraints in Equation 1 can be removed. Deduce the value of one of the dual variable α_j corresponding to one value of j (i.e., one constraint) which we will write j^* to not reveal the answer. Follow indexes introduced in the problem statement above when describing the dataset.

7. (1 point) Deduce from the second expression found in question 4 (i.e., $\sum_j E_j F_j = G$), a relationship between all remaining dual variables α_j for $j \neq j^*$
8. (3 points) Solving the optimization problem from question 5, find values for all dual variables α_j for $j \in 1..4$.
9. (1 point) Deduce the numerical values of the two components of $\vec{w}$.
10. (2 points) Through an analysis of constraints in Equation 1 which are tight, deduce the numerical value of b. Hint: recall the role of support vectors here.
11. (0 points) Check that the hyperplane you found analytically is the same than the one you found geometrically in the first question.

*
* *