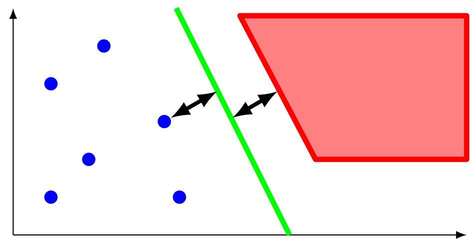
CSCI 5654, Spring 2023: Assignment 3*

Assigned date: Th 3/2/2023
Due date: Th 3/16/2023

Instructions: Please upload your HW as a PDF file.

* updated version: 3/11/2023

We consider the problem of separating two sets of data points. This problem has application for instance in Machine Learning with Support Vector Machines (SVM). To make things more interesting, and to connect with the topic of polyhedra computations of the course, we assume that one data set is a polyhedron. See the figure below for an illustration:



In this homework, we will see, step by step, how to compute a hyperplane separating the two data sets.

Problem 1 (5 points): Halfspace containing a polyhedron

Consider a polyhedron $S = \{x \in \mathbb{R}^n : Ax \leq b\}$, wherein $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Now consider a halfspace $H^{++} = \{x \in \mathbb{R}^n : a^\top x > \beta\}$, wherein $a \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$. We will derive a condition on (a, β), based on (A, b), for S to be contained in H^{++}, i.e., for $S \subseteq H^{++}$. Note that $S \subseteq H^{++}$ if and only if $S \cap H^- = \emptyset$, wherein $H^- = \{x \in \mathbb{R}^n : a^\top x \leq \beta\}$ is the complement of H^{++}.

(a) From the above, show that $S \subseteq H^{++}$ if and only if the following set of linear inequalities is infeasible:

$$\begin{bmatrix} A \\ a^\top \end{bmatrix} x \leq \begin{bmatrix} b \\ \beta \end{bmatrix} \tag{1}$$

(b) Using Farkas' Lemma[1], show that (1) is infeasible if and only if there is $y \in \mathbb{R}^m \geq 0$ and $\gamma \in \mathbb{R} \geq 0$ such that $A^\top y + \gamma a = 0$ and $b^\top y + \beta\gamma < 0$.

(c) Assume that S is nonempty. Using Farkas' Lemma, deduce that there is no $y \in \mathbb{R}^m \geq 0$ such that $A^\top y = 0$ and $b^\top y < 0$.

(d) From (b) and (c), conclude that $S \subseteq H^{++}$ if and only if there is $y \in \mathbb{R}^m \geq 0$ such that $A^\top y + a = 0$ and $b^\top y + \beta < 0$. (Hint: first, deduce from (c) that the solution in (b) satisfies $\gamma > 0$.)

Problem 2 (5 points): Separating hyperplane

Consider a set of p points in $\mathbb{R}^n$: $V = \{x_1, \ldots, x_p\} \subseteq \mathbb{R}^n$. Let S be a polyhedron as in Problem 1. Our goal is to find a hyperplane $H = \{x \in \mathbb{R}^n : a^\top x = \beta\}$ separating V and S. In other words, we want to compute $a \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$ such that $V \subseteq H^{--} = \{x \in \mathbb{R}^n : a^\top x < \beta\}$ and $S \subseteq H^{++} = \{x \in \mathbb{R}^n : a^\top x > \beta\}$.

(a) Show that if $a \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$ satisfy that $V \subseteq \{x \in \mathbb{R}^n : a^\top x < \beta\}$ and $S \subseteq \{x \in \mathbb{R}^n : a^\top x > \beta\}$, then for any multiplier $\lambda > 0$, λa and $\lambda\beta$ also satisfy $V \subseteq \{x \in \mathbb{R}^n : \lambda a^\top x < \lambda\beta\}$ and $S \subseteq \{x \in \mathbb{R}^n : \lambda a^\top x > \lambda\beta\}$.

[1] We have seen different versions of Farkas' Lemma in the class. For a reference, see, e.g., Vanderbei's book, Theorem 10.5, p. 165.

(b) Deduce from (a) that, without loss of generality, we can put bounds on a: namely, show that there is $a \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$ satisfying $V \subseteq \{x \in \mathbb{R}^n : a^\top x < \beta\}$ and $S \subseteq \{x \in \mathbb{R}^n : a^\top x > \beta\}$ if and only if there is $a \in [-1,1]^n$ and $\beta \in \mathbb{R}$ satisfying $V \subseteq \{x \in \mathbb{R}^n : a^\top x < \beta\}$ and $S \subseteq \{x \in \mathbb{R}^n : a^\top x > \beta\}$.

(c) From (b) and from Problem (1), show that the computation of a and β can be formulated as the following Linear Program, with variables $a \in \mathbb{R}^n$, $\beta \in \mathbb{R}$, $y \in \mathbb{R}^m$ and $r \in \mathbb{R}$:

$$\begin{array}{rl} \max & r \\ \text{s.t.} & a^\top x_i \leq \beta - r, \quad \forall\, i = 1, \ldots, p \\ & A^\top y + a = 0 \\ & b^\top y + \beta \leq -r \\ & y \geq 0 \\ & -1 \leq a \leq 1\ . \end{array} \tag{2}$$

In particular, show that there is an optimal solution (a^*, β^*, y^*, r^*) with $r^* > 0$ if and only if there is a separating hyperplane.

Problem 3 (2 bonus points[2]): Implementation

For a nonempty polyhedron $S \subseteq \mathbb{R}^2$ (as in Problem 1) and a nonempty set of points $V \subseteq \mathbb{R}^2$ (as in Problem 2) that you are free to choose, but restricted to the 2-dimensional plane, do the following:

- Plot S and V.
- Implement and solve (2) for S and V using a language and an optimization solver of your choice.
- Plot the separating hyperplane in the plot with S and V.

You do not need to provide your source code.

[2] Up to 2 points added to the final score of this assignment, with a ceil of 10 points for the final score.