Separating Two Sets of Points via Feasibility of an LP

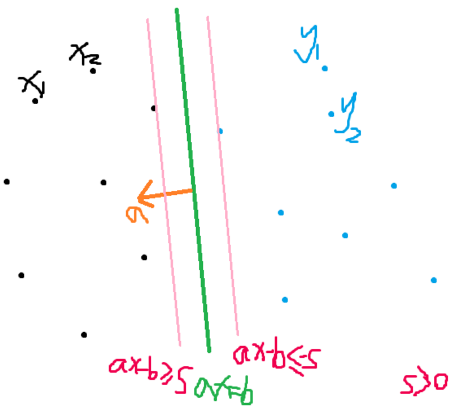
Consider two sets of points
$\{x_1, x_2, \dots, x_N\}$ and $\{y_1, y_2, \dots, y_M\}$, where each point is two dimensional (so we have N+M points on the plane). We say that these sets are linearly separable if there is a line such that all the x_s are on one side of the line, and all the y_s are on the other side. Express this as a linear feasibility problem (see eqn 8.21). Develop an algorithm from scratch to solve this problem using the approach in (11.19) where $f_i(x)$ are affine functions of x. Express the gradients and the Hessians needed for the algorithm explicitly in your report and describe the algorithm. Show that for two simple examples with two variables (one feasible, and one infeasible) that your algorithm can identify feasibility.

Analysis of the problem:

If two sets of points $\{x_1, x_2, \dots, x_N\}$ and $\{y_1, y_2, \dots, y_M\}$ are separable, the separating line is *ax* =*b*, then:

$$\begin{cases} {x_i}^T a > b, & \mathrm{i} = 1 \dots \dots \mathrm{N} \\ {y_j}^T a < b, & \mathrm{j} = 1 \dots \dots \mathrm{M} \end{cases}, \text{ where } x_i, y_j \in R^2, \mathrm{a} \in R^2, \mathrm{b} \in R$$

We assume that a and b are variables. We want to find feasible a, b satisfying those N+M constraints.



In textbook, feasibility problem (11.19) is:
Minimize: s
Subject to: $f_i(x) < s, \quad i = 1, \dots, m$
$Ax = b$

We form our problem similarly, assume s>0,

Minimize: s

Subject to: $\begin{cases} {x_i}^T a - b \geq s, & \text{i} = 1 \dots\dots \text{N} \\ {y_j}^T a - b \leq -s, & \text{j} = 1 \dots\dots \text{M} \end{cases}$, which can be written

as: $\begin{cases} -{x_i}^T a + b + s \leq 0, & \text{i} = 1 \dots\dots \text{N} \\ {y_j}^T a - b + s \leq 0, & \text{j} = 1 \dots\dots \text{M} \end{cases}$, i.e. $\begin{cases} f_i(s) \leq 0, & \text{i} = 1 \dots\dots \text{N} \\ \text{g}_i(s) \leq 0, & \text{j} = 1 \dots\dots \text{M} \end{cases}$

Where $f_i(s) = -{x_i}^T a + b + s, \quad \text{g}_i(s) = {y_j}^T a - b + s$

Therefore the original feasibility problem boils down to this an inequality constrained ($Fs \leq 0, Gs \leq 0$) minimization problem:

Minimize: s

Subject to: $\begin{cases} f_i(s) \leq 0, & \text{i} = 1 \dots\dots \text{N} \\ \text{g}_i(s) \leq 0, & \text{j} = 1 \dots\dots \text{M} \end{cases}$

Note that although we have variables a and b, we can regard them as determined parameters now. Because what we want to do is "to prove the existence of a separating line", which is equivalent to find a feasible s. there is no constraint on a and b. ??

inequality constrained problem cannot be solved by newton's method. Therefore, we transform it to equality constrained (or unconstrained) problem.

The problem is transformed as an unconstrained problem :

Minimize: $\text{s} + \sum_{i=1}^{N} I\_({x_i}^T a + b + s) + \sum_{i=1}^{M} I\_\left({y_j}^T a - b + s\right),$

$where \quad \text{i} = 1 \dots\dots \text{N}, \text{j} = 1 \dots\dots \text{M}, \; I\_(u) = \begin{cases} 0 & u \leq 0 \\ \infty & u > 0 \end{cases}$, Variable is s.

The indicator function $I\_(u)$ is not differentiable, where newton's method cannot be applied. But it can be approximated by logarithmic barrier function.

$$\widehat{I\_}(u) = -\frac{1}{t} log(-u)$$

Then the problem is transformed into:

Minimize: $\text{s} + \left(-\frac{1}{t}\right) [\, \sum_{i=1}^{N} \log(-{x_i}^T a - b - s) + \sum_{i=1}^{M} \log\left(-{y_j}^T a + b - s\right) \,],$

$where \quad \text{i} = 1 \dots\dots \text{N}, \text{j} = 1 \dots\dots \text{M}$

Variable is s.

Denote:

$$\Phi(\mathrm{s}) = \sum_{i=1}^{N} \log(-x_i{}^T a - b - s) + \sum_{i=1}^{M} \log\left(-y_j{}^T a + b - s\right)$$

$$\operatorname{dom} \Phi = \{x_i{}^T a + b + s < 0\,, y_j{}^T a - b + s < 0\}$$

the problem is:

minimize $\quad \mathrm{ts} + \Phi(\mathrm{s})$, $\quad$ with no constraint

Now the problem can be solved by Newton's method. It has a unique solution for each t>0

Gradient: $\nabla\Phi(\mathrm{s}) = \sum_{i=1}^{N} \frac{-1}{-x_i{}^T a - b - s} + \sum_{i=1}^{M} \frac{-1}{-y_j{}^T a + b - s}$

Hessian: $\nabla^2\Phi(\mathrm{s}) = \sum_{i=1}^{N} \frac{-1}{(-x_i{}^T a - b - s)^2} + \sum_{i=1}^{M} \frac{-1}{\left(-y_j{}^T a + b - s\right)^2}$

For each t>0, we define $s^*(t)$ as the optimal solution. The series of $s^*(t)$ forms central path.

Each of them satisfies: $t + \nabla\Phi\left(s^*(t)\right) = 0$, i.e. $t + \sum_{i=1}^{N} \frac{-1}{-x_i{}^T a - b - s^*(t)} + \sum_{i=1}^{M} \frac{-1}{-y_j{}^T a + b - s^*(t)} = 0$

Implement in Matlab: