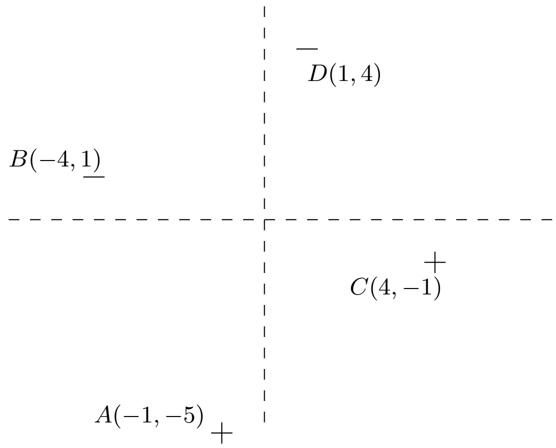
CMSC5724: Exercise List 5

Answer all the problems below based on the following set P of points A, B, C and D:



where "+" represents label 1 and "-" represents label -1.

Problem 1. What is the margin of the separation line $\ell : -x - 5y = 0$?

Answer: The distance between ℓ and the points in P are as follows:

- A: $\frac{|-1\times(-1)-5\times(-5)|}{\sqrt{(-1)^2+(-5)^2}} = \sqrt{26}$.
- B: $\frac{|-4\times(-1)+1\times(-5)|}{\sqrt{(-1)^2+(-5)^2}} = 1/\sqrt{26}$.
- C: $\frac{|4\times(-1)-1\times(-5)|}{\sqrt{(-1)^2+(-5)^2}} = 1/\sqrt{26}$.
- D: $\frac{|1\times(-1)+4\times(-5)|}{\sqrt{(-1)^2+(-5)^2}} = 21/\sqrt{26}$.

Therefore, the margin of ℓ is $1/\sqrt{26}$.

Problem 2. Run Margin Perceptron on P with $\gamma_{guess} = 0.1$, and give the equation of the line that is maintained by the algorithm at the end of each iteration.

Answer: Let us represent the line maintained by Margin Perceptron as $c_1 x + c_2 y = 0$. Define $\vec{c} = [c_1, c_2]$. At the beginning of Margin Perceptron, $\vec{c} = [0, 0]$. We use $\vec{A}$ to denote the vector $[-1, -5]$, obtained by listing the coordinates of A. Define $\vec{B}, \vec{C}, \vec{D}$ similarly.

Iteration 1. A does not satisfy $\vec{A} \cdot \vec{c} > 0$. So we update $\vec{c}$ to $\vec{c} + \vec{A} = [0, 0] + [-1, -5] = [-1, -5]$.

Iteration 2. No more violation. So we have found a separation line $-x - 5y = 0$.

Problem 3. Same as the previous problem but with $\gamma_{guess} = 4/\sqrt{26}$.

Answer: Starting with $\vec{c} = [0, 0]$, Margin Perceptron runs as follows:

Iteration 1. A does not satisfy $\vec{A} \cdot \vec{c} > 0$. So we update $\vec{c}$ to $\vec{c} + \vec{A} = [0, 0] + [-1, -5] = [-1, -5]$.

Iteration 2. The distance between B and the line determined by $\vec{c}$ is $1/\sqrt{26}$, which is smaller than $\gamma_1/2$. So we update $\vec{c}$ to $\vec{c} - \vec{B} = [-1, -5] - [-4, 1] = [3, -6]$.

Iteration 3. No more violation. So we have found a separation line $3x - 6y = 0$.

Problem 4. Give an instance of quadratic programming to find an origin-passing separation plane with the maximum margin.

Answer: Minimize $w_1^2 + w_2^2$ subject to the following constraints:

- $(-1)w_1 + (-5)w_2 \geq 1$
- $4w_1 + (-1)w_2 \geq 1$
- $(-4)w_1 + w_2 \leq -1$
- $w_1 + 4w_2 \leq -1$

Problem 5. Consider the following instance of quadratic programming in $\mathbb{R}^d$:

$$\text{minimize } |\boldsymbol{w}| \text{ subject to}$$
$$\boldsymbol{w} \cdot \boldsymbol{p}_i \geq 1 \text{ for each } i \in [1, n]$$

where $\boldsymbol{p}_1, ..., \boldsymbol{p}_n$ are n given points in $\mathbb{R}^d$. Prove: if an optimal $\boldsymbol{w}$ exists, there must exist at least one $i \in [1, n]$ such that $\boldsymbol{w} \cdot \boldsymbol{p}_i = 1$.

Answer: We will give a proof by contradiction. Suppose that $\boldsymbol{w}$ is an optimal solution and $\boldsymbol{w} \cdot \boldsymbol{p_i} > 1$ for every $i \in [1, n]$. Define $\tau = \min_i \boldsymbol{w} \cdot \boldsymbol{p_i}$ and $\boldsymbol{w'} = \boldsymbol{w}/\tau$. We know $\tau > 1$ (otherwise, there exists an i such that $\boldsymbol{w} \cdot \boldsymbol{p}_i = 1$):

- $\boldsymbol{w} \cdot \boldsymbol{p_i} \geq \tau$ for each $i \in [1, n]$

which implies

- $\boldsymbol{w'} \cdot \boldsymbol{p_i} \geq 1$ for each $i \in [1, n]$.

Hence, $\boldsymbol{w'}$ is a feasible solution of the quadratic programming. However, the fact $|\boldsymbol{w'}| < \boldsymbol{w}$ contradicts the optimality of $\boldsymbol{w}$.

Problem 6. Let γ_{opt} be the maximum margin of an origin-passing separation plane on a set P of points. Denote by R the largest distance from a point in P to the origin.

Suppose that, given a value γ, margin Perceptron ensures the following:

- if it terminates, it definitely returns a separation plane with margin at least $\alpha \cdot \gamma$, where α is an arbitrary constant less than 1;
- if $\gamma \leq \gamma_{opt}$, it definitely terminates after at most $c \cdot R^2/\gamma^2$ corrections, for some constant (which depends on α).

Design an algorithm to find a separation plane with margin at least $\alpha \cdot \beta \cdot \gamma_{opt}$ after $O(R^2/\gamma_{opt}^2)$ corrections in total, where β can be any constant less than 1.

Answer: Use exactly the same algorithm taught in the class that repeatedly runs margin Perceptron with an increasingly smaller γ, except that we set γ to $\beta^{i-1} \cdot R$ in the i-th run.

Suppose that the value of γ in the final run is $\gamma_{final} = \beta^x \cdot R$. Since we did not stop at the previous run, we know that $\gamma_{final}/\beta > \gamma_{opt}$, namely, $\gamma_{final} > \beta \cdot \gamma_{opt}$.

In the final run, the separation plane returned must have a margin at least $\alpha \cdot \gamma \geq \alpha \cdot \beta \cdot \gamma_{opt}$.

The total number of corrections is no more than

$$cR^2 \left(\frac{1}{\gamma_{final}^2} + \frac{\beta^2}{\gamma_{final}^2} + \frac{\beta^4}{\gamma_{final}^2} ... \right) = O(R^2/\gamma_{final}^2) = O(R^2/\gamma_{opt}^2).$$