Problem 1. Let $P$ be a set of 4 points: $A = (1,2), B = (2,1), C = (0,1)$ and $D = (1,0)$ where $A, B$ have label 1, while $C, D$ have label $-1$. We want to find a plane that separates the points of the two labels.

1. Convert the problem to $\mathbb{R}^3$ so that it can be solved by the Perceptron algorithm. Give the resulting dataset $P'$.

2. Execute Perceptron on $P'$. Give the equation of the plane that is maintained by the algorithm at the end of each iteration.

3. Convert the plane output by Perceptron back to the original 2d space to obtain a separation plane on $P$.

Answer.

1. $P'$ includes the following points: $A' = (1,2,1), B' = (2,1,1), C' = (0,1,1), D' = (1,0,1)$.

2. Let us represent the plane maintained by Perceptron as $c_1 x_1 + c_2 x_2 + c_3 x_3 = 0$. Denote by $\vec{c} = (c_1, c_2, c_3)$. At the beginning of Perceptron, $\vec{c} = (0,0,0)$.

   We use $\vec{A}'$ to denote the vector $(1,2,1)$, obtained by listing the coordinates of $A'$. Define $\vec{B}', \vec{C}', \vec{D}'$ similarly.

   Iteration 1. Since $\vec{A}'$ does not satisfy $\vec{A}' \cdot \vec{c} > 0$, we update $\vec{c}$ to $\vec{c} + \vec{A}' = (0,0,0) + (1,2,1) = (1,2,1)$.

   Iteration 2. Since $\vec{C}'$ does not satisfy $\vec{C}' \cdot \vec{c} < 0$, we update $\vec{c}$ to $\vec{c} - \vec{C}' = (1,2,1) - (0,1,1) = (1,1,0)$.

   Iteration 3. Since $\vec{C}'$ does not satisfy $\vec{C}' \cdot \vec{c} < 0$, we update $\vec{c}$ to $\vec{c} - \vec{C}' = (1,1,0) - (0,1,1) = (1,0,-1)$.

   Iteration 4. Since $\vec{A}'$ does not satisfy $\vec{A}' \cdot \vec{c} > 0$, we update $\vec{c}$ to $\vec{c} + \vec{A}' = (1,0,-1) + (1,2,1) = (2,2,0)$.

   Iteration 5. Since $\vec{C}'$ does not satisfy $\vec{C}' \cdot \vec{c} < 0$, we update $\vec{c}$ to $\vec{c} - \vec{C}' = (2,2,0) - (0,1,1) = (2,1,-1)$.

   Iteration 6. Since $\vec{C}'$ does not satisfy $\vec{C}' \cdot \vec{c} < 0$, we update $\vec{c}$ to $\vec{c} - \vec{C}' = (2,1,-1) - (0,1,1) = (2,0,-2)$.

   Iteration 7. Since $\vec{A}'$ does not satisfy $\vec{A}' \cdot \vec{c} > 0$, we update $\vec{c}$ to $\vec{c} + \vec{A}' = (2,0,-2) + (1,2,1) = (3,2,-1)$.

   Iteration 8. Since $\vec{C}'$ does not satisfy $\vec{C}' \cdot \vec{c} < 0$, we update $\vec{c}$ to $\vec{c} - \vec{C}' = (3,2,-1) - (0,1,1) = (3,1,-2)$.
**Iteration 9.** Since $\vec{D}'$ does not satisfy $\vec{D}' \cdot \vec{c} < 0$, we update $\vec{c}$ to $\vec{c} - \vec{D}' = (3, 1, -2) - (1, 0, 1) = (2, 1, -3)$.

**Iteration 10.** No more violation. So we have found a separation plane $2x_1 + x_2 - 3x_3 = 0$ for $P'$.

3. The corresponding separation plane in the original 2d space is therefore $2x_1 + x_2 - 3 = 0$.

**Problem 2.** Let $P$ be a set of multidimensional points where each point has a label equal to 1 or $-1$. We want to design an algorithm to achieve the following purpose:

- Either return a separation plane;
- Or declare that $P$ has no separation planes with a margin at least $\gamma$ (recall that the margin of a separation plane $\pi$ equals the minimum of the distances from the points of $P$ to $\pi$).

Note that your algorithm must still work even if $P$ is not linearly separable.

**Answer.** Simply run Perceptron, and return whatever plane found by the algorithm. If the algorithm still has not finished after $R^2/\gamma^2$ vector corrections, manually force it to stop, and declare that no separation plane has a margin at least $\gamma$.

**Problem 3.** Consider the set of points below. The coordinates of two points are given as shown. Suppose that we apply Perceptron to find a plane that passes the origin, and separates the black points from the white points. Prove: Perceptron finishes after at most 5 iterations.

![Diagram](image)

**Answer.** The $y$-axis is a separation plane with margin $\gamma = 1$. Clearly, the largest distance from a point to the origin is $R = \sqrt{5}$. Hence, Perceptron performs at most $R^2/\gamma^2 = 5$ iterations.

**Problem 4.** Some people prefer the following variant of the Perceptron algorithm:

1. $\vec{c} = \vec{0}$
2. while there is a violating point $\vec{p}$
3. if $\vec{p}$ has label 1
4. $\vec{c} = \vec{c} + \lambda \cdot \vec{p}$
   else
5. $\vec{c} = \vec{c} - \lambda \cdot \vec{p}$
where $\lambda$ is a positive real-value constant that is often referred to as the “step length”. In the version we discussed in the lecture, $\lambda = 1$. Prove: regardless of $\lambda$, Perceptron always terminates in $R^2/\gamma^2$ iterations, where $R$ is the maximum norm of all points, and $\gamma$ the margin of an arbitrary separation plane. In other words, the value of $\lambda$ has no effect in this bound.

**Answer.** Let $\vec{c}_k$ denote the vector $\vec{c}$ at the end of the $k$-th iteration. Following the analysis we discussed in the lecture, we can prove the following two inequalities:

\[
|\vec{c}_k| \geq \lambda \cdot k \cdot \gamma \\
|\vec{c}_k|^2 \leq k \cdot \lambda^2 \cdot R^2
\]

The two inequalities give $\lambda^2 \cdot k^2 \cdot \gamma^2 \leq k \cdot \lambda^2 \cdot R^2$, which indicates $k \leq R^2/\gamma^2$. 

3