Foundations of Data Science & Machine Learning

Tutotial 04

March 19, 2021

Definition 1. Let X be a subset of $\mathbb{R}^n$. A function $f : X \to \{+1, -1\}$ is called *linearly separable* if there exists a vector $a \in \mathbb{R}^n$ such that for every $x \in X$, $\langle a, x \rangle \geq 1$ if and only if $f(x) = +1$. A 2-partition $(G, X \setminus G)$ of X is called *linear* if there exists a vector $a \in \mathbb{R}^n$ such that for every $x \in G$, $\langle a, x \rangle \geq 1$ and for every $x \in X \setminus G$, $\langle a, x \rangle < 1$.

Remark. The demand that $b = 1$ for a hyperplane defined by $\langle a, x \rangle = b$ restricts the choice of hyperplanes by half. In fact, we are now restricted to those hyperplanes where origin is on its negative side.

Question 1. Let X be a set of N points in $\mathbb{R}^2$. Prove that the number of linearly separable functions $f : X \to \{+1, -1\}$ is at most $\binom{N+1}{2} + 1$.

It is easy to see that a function $f : X \to \{+1, -1\}$ is linearly separable if and only if the partition $(f^{-1}(+1), f^{-1}(-1))$ of X is linear. Hence Question 1 is equivalent to the following claim.

Claim. *The number of linear partitions of a set X of N points in $\mathbb{R}^2$ is at most $\binom{N+1}{2} + 1$.*

Proof. We will prove the claim by induction on N. Base case, $N = 1$, is true since both the 2-partitions (ϕ, X) and (X, ϕ) of X are easily verified to be linear. Now we assume that the claim is true for any set with $N - 1$ points and consider a set $X = \{x_1, \ldots, x_N\}$ of N points and $X' = X \setminus \{x_N\}$.

Let L be the family of linear partitions of X and L' be the family of linear partitions of X'. Taking any linear partition of X and removing x_N (from its part) gives a linear partition of X'. That is, every partition in L is obtained by adding x_N to one of the parts of a partition in L'. Moreover one can see that every partition in L' can be extended to either one or two partitions in L by adding x_N. We split L' into two subfamilies L'_1 and L'_2 based on that. L'_1 consists of those linear partitions of X' which can be extended in only one way to get a linear partition of X and L'_2 consists of those which can be extended in both ways.[1]

A partition $(G, X' \setminus G)$ is in L'_2 if there are two lines l_1 and l_2 that separate G from $X' \setminus G$ such that x_N is on the positive side of l_1 but negative side of l_2. The existence of l_1 and l_2 also means that G can be separated from $X' \setminus G$ by a line passing through x_N. Hence every partition in L'_2 can be obtained by the following process. Start with a horizontal line l passing through x_N. The line l adds one partition to L_2. Now keeping x_N as the pivot, rotate l and each time it crosses a point in X' we get a new partition in L_2. Since there are $N - 1$ points X' we get a total of at most N partitions in L_2. Now

[1]If every partition in L' could be extended in both ways, we will end up with 2^N partitions for X. It is the presence of a large L'_1 that saves our day.

$$
\begin{aligned}
|L| &= |L_1'| + 2\,|L_2'| \\
&= |L'| + |L_2'| && (L' = L_1' \uplus L_2') \\
&\le |L'| + N && (|L_2| \le N) \\
&\le \binom{N}{2} + 1 + N && \text{(induction hypothesis)} \\
&= \binom{N+1}{2} + 1.
\end{aligned}
$$

□

Another Proof (using Tutorial 0)

Let $A = \mathbb{R}^2$ and $B = \mathbb{R}^2$. Imagine a map ϕ which maps every non-zero vector in $x \in A$ to the line in B given by $\{y \in B : \ \langle x, y \rangle = 1\}$. Similarly let ψ be a map which maps every line $L = \{x \in A : \ \langle l, x \rangle = 1\}$ to the point l in B. Let $X = \{x_1, \ldots, x_N\}$ be any set of N points in A. Let $L_1, \ldots, L_N$ be the corresponding lines in B. That is, $L_i = \phi(x_i)$. Our proof will be complete if we establish the following claim.

Claim. *Let $P = \{x \in A : \ \langle p, x \rangle = 1\}$ and $Q = \{x \in A : \ \langle q, x \rangle = 1\}$ be any two lines in A. Then, P and Q define the same partition of X if and only if the points $p = \psi(P)$ and $q = \psi(Q)$ lie inside the same region carved out by the lines $L_1, \ldots, L_N$ in B.*

Proof. Suppose P and Q define the same partition of X. That means

$$\forall i \in [n], \ (\langle p, x_i \rangle \ge 1) \iff (\langle q, x_i \rangle \ge 1),$$

and hence p and q are on the same side of L_i in B for each i.

Suppose p and q are two points in the same region of B cut out by $L_1, \ldots, L_N$. Then

$$\forall i \in [n], \ (\langle x_i, p \rangle \ge 1) \iff (\langle x_i, q \rangle \ge 1),$$

and hence the for each i, the point x_i is on the same side of the lines P and Q. In other words, P and Q result in the same partition of X. □

This second proof is an interesting thought exercise. You can take a line L in A and then wiggle it around a bit and see what happens to the point $\psi(L)$ in B in two cases. First when you do not cross any x_i during the wiggling and second when you cross some point x_i during the wiggling.