Foundations of Data Science & Machine Learning

Tutotial 01 Solutions

February 27, 2021

Definition 1 (Axis-normal hyperplane)**.** An *axis-normal hyperplane* in $\mathbb{R}^n$ is a hyperplane whose normal vector is along one of the axes in $\mathbb{R}^n$. Equivalently, it is a hyperplane defined by $\langle a, x\rangle = b$ $a = (x_1, x_2 \dots, x_n)$ such that exactly one $x_i = 1$ and all others 0.

Definition 2 (Bounding box)**.** An *axis-parallel box* in $\mathbb{R}^n$ is the Cartesian product of n closed intervals on $\mathbb{R}$. The *bounding box* of a set of points P in $\mathbb{R}^n$ is the smallest axis-parallel box containing all points in P.

Theorem 3. *Two sets X and Y of points in $\mathbb{R}^n$ is separable by an axis-normal hyperplane if and only if the bounding boxes of X and Y are disjoint.*

Proof. In this proof, we will denote by e_i the standard basis vector of $\mathbb{R}^n$ which has 1 in the i-th coordinate and 0 everywhere else. The set {1, …, n} will be denoted by the shorthand $[n]$.

Let the bounding boxes B_X and B_Y of X and Y be disjoint. By definition, both B_X and B_Y are Cartesian products of closed intervals, say $I_1, \dots, I_n$ and $J_i, \dots, J_n$ respectively. That is $B_X = I_1 \times \cdots \times I_n$ and $B_Y = J_1 \times \cdots \times J_n$. Moreover for each $x = (x_1, \dots, x_n) \in X$, we have $x_i \in I_i$ and for each $x = (x_1, \dots, x_n) \in Y$, we have $x_i \in J_i$.

Suppose for each $i \in [n]$, the two intervals I_i and J_i are overlapping. Then pick any point $z_i \in I_i \cap J_i$. It is easy to check that $(z_1, \dots, z_n) \in B_X \cap B_Y$. In fact $(I_1 \cap J_1) \times \cdots \times (I_n \cap J_n) \subseteq B_X \cap B_Y$. Since we are under the assumption that B_X and B_Y are disjoint, there exists an i such that I_i and J_i are disjoint. Since I_i and J_i are disjoint intervals in $\mathbb{R}$, there exist a point $b \in \mathbb{R}$ such that all points in I_i is to one side of b and all points in J_i is to the other side. Hence the axis-normal plane defined by $\langle e_i, x\rangle = b$ separates B_X and B_Y and therefore X and Y. Notice that $\langle e_i, x\rangle = x_i$ which is in I_i for all $x \in X$ and J_i for all $x \in Y$.

In the other direction, let X and Y be on the two sides of the axis-normal hyperplane defined by $\langle e_j, x\rangle = b$ for some $j \in [n]$. Without loss of generality let $b_x = \max_{x\in X} \langle e_j, x\rangle < b$ and $b_y = \min_{y\in Y} \langle e_j, y\rangle > b$. (If not swap the roles of X and Y). Let $s \in \mathbb{R}$ be a number that is smaller than the smallest coordinate among all the vectors in $X \cup Y$ and $t \in \mathbb{R}$ be a number that is larger than the largest coordinate among all the vectors in $X \cup Y$. Let $I_i = J_i = [s, t]$ for all $i \neq j$ and let $I_j = [s, b_x]$ and $J_j = [b_y, t]$. Now $B = I_1 \times \cdots \times I_n$ and $C = J_1 \times \cdots \times J_n$ are disjoint axis-parallel boxes which respectively contain X and Y. Hence the bounding boxes of X and Y which are contained in B and C are also disjoint. □

Definition 4. The *convex hull* of a set of points S in $\mathbb{R}^n$ is the smallest convex set containing all points in S.

Theorem 5. *The convex hull of a set of points S in $\mathbb{R}^n$ is the intersection of all convex sets in $\mathbb{R}^n$ which contain S.*

Proof. First let us observe that for any two convex sets A and B, $A \cap B$ is also convex. To see this, pick any two points $x, y \in A \cap B$ and any $\alpha \in [0, 1]$. The point $z = \alpha x + (1 - \alpha)y$ is in A and B by the convexity of the respective sets and hence in $A \cap B$.

Now let H_M be the minimal convex set containing S and H_I be the intersection of all convex sets containing S. We need to show that $H_M = H_I$.

Since H_M is also a convex set containing S, it is part of the intersection which defines H_I. Since $A \cap B \subseteq A$, we have $H_I \subseteq H_M$. In the other direction, first notice that $S \subset H_I$. Moreover, since convexity is preserved by intersection (first para of the proof), H_I is also convex. Hence if H_I is a proper subset of H_M, it violates the minimality of H_M. Hence $H_I = H_M$. $\square$