Math 281C Homework 3 Solutions

1. Let $\mathcal{H}_k$ denote the indicators of all closed half-spaces in $\mathbb{R}^k$, i.e. $\mathcal{H}_k = \{ x \mapsto I((a,x) + b \leq 0) : a \in \mathbb{R}^k, b \in \mathbb{R} \}$. Show that the VC dimension of $\mathcal{H}_k$ is exactly equal to $k + 1$.

**Solution:** First we show that $\mathcal{V}(\mathcal{H}_k) \geq k + 1$. Consider the set $\{ x_1, \ldots, x_{k+1} \}$ with $x_i = e_i$, for $i = 1, \ldots, k$, and $x_{k+1} = 0$. For any binary labels $(y_1, \ldots, y_{k+1}) \in \{0,1\}^{k+1}$, let $a_i = y_{k+1} - y_i$, for $i = 1, \ldots, k$, and $b = -y_{k+1} + 1/2$, then the classifier $I((a,x) + b \leq 0)$ realizes the labels. This shows that the set $\{ x_1, \ldots, x_{k+1} \}$ is shattered by $\mathcal{H}_k$.

Then we show that $\mathcal{V}(\mathcal{H}_k) \leq k + 1$. In the following proof, we denote $f_{a,b}(x) = (a,x) + b$.

**Method 1:** The function class $\{ f_{a,b} : a \in \mathbb{R}^k, b \in \mathbb{R} \}$ is a vector space of dimension $k + 1$, then the result follows by Example 2.2 in Lecture 5.

**Method 2:** Suppose there is a set $\{ x_1, \ldots, x_{k+2} \}$ that is shattered by $\mathcal{H}_k$. By Radon theorem, the set can be partitioned into two disjoint subsets $A$ and $B$ such that $\text{conv}(A) \cap \text{conv}(B) \neq \emptyset$. Since $\{ x_1, \ldots, x_{k+2} \}$ is shattered by $\mathcal{H}_k$, there are $a \in \mathbb{R}^k$ and $b \in \mathbb{R}$ such that for any $x \in A, f_{a,b}(x) \leq 0$ and for any $x \in B, f_{a,b}(x) > 0$. This further means that for any $x \in \text{conv}(A), f_{a,b}(x) \leq 0$, and for any $x \in \text{conv}(B), f_{a,b}(x) > 0$, which is contradictory with $\text{conv}(A) \cap \text{conv}(B) \neq \emptyset$.

2. Consider the sphere $S_{a,b} = \{ x \in \mathbb{R}^k : \| x - a \|_2 \leq b \}$, where $(a,b) \in \mathbb{R}^k \times \mathbb{R}_+$ specify its center and radius, respectively. Define the function $f_{a,b}(x) = \| x \|_2^2 - 2 \sum_{j=1}^k a_j x_j + \| a \|_2^2 - b^2$, so that $S_{a,b} = \{ x \in \mathbb{R}^k : f_{a,b}(x) \leq 0 \}$. Let $S_k = \{ x \mapsto I\{ f_{a,b}(x) \leq 0 \} : a \in \mathbb{R}^k, b \geq 0 \}$. Show that the VC dimension of $S_k$ is at most $k + 2$.

**Solution:** For any $x \in \mathbb{R}^k$, define $\phi(x) = (x_1, \ldots, x_k, \| x \|_2^2)^T : \mathbb{R}^k \rightarrow \mathbb{R}^{k+1}$. Then $f_{a,b}(x) = \langle u, \phi(x) \rangle + v$, with $u = (-2a_1, \ldots, -2a_k, 1)^T$ and $v = \| a \|_2^2 - b$. In the following proof, we denote $g_{u,v}(x) = \langle u, \phi(x) \rangle + v$.

**Method 1:** The function class $\{ g_{u,v} : u \in \mathbb{R}^{k+1}, v \in \mathbb{R} \}$ is a vector space with dimension $k + 2$, and it contains the function class $\{ f_{a,b} : a \in \mathbb{R}^k, b \in \mathbb{R}_+ \}$. Applying Example 2.2 in Lecture 5 to this larger vector space completes the proof.

**Method 2:** Suppose there is a set $\{ x_1, \ldots, x_{k+3} \}$ that is shattered by $\mathcal{S}_k$, then $\{ \phi(x_1), \ldots, \phi(x_{k+3}) \}$ is shattered by $\mathcal{H}_{k+1}$, where $\mathcal{H}_{k+1}$ is defined in Question 1. This is a contradiction with $\mathcal{V}(\mathcal{H}_{k+1}) \leq k + 2$.

3. Consider the class of all spheres in $\mathbb{R}^2$: $S_2$, where $S_k$ is defined in question 2.

(a) Show that $S_2$ can shatter any subset of three points that are not collinear.

**Solution:** There are $2^3 = 8$ possible ways to label 3 points that are not collinear. We omit the graphs but it’s easy to see that circular classifiers can realize these labels.

(b) Conclude that the VC dimension of $S_2$ is 3.

**Solution:** We prove the general result that $\mathcal{V}(S_k) = k + 1$.

First we show that $\mathcal{V}(S_k) \geq k + 1$. Consider the set $\{ x_1, \ldots, x_{k+1} \}$ with $x_i = e_i$, for $i = 1, \ldots, k$, and $x_{k+1} = 0$ as in Question 1. For any binary labels, suppose the subset $A$ contains the points labeled as 1. The result is trivial if $A = \emptyset$. If $A \neq \emptyset$, we consider the sphere with the center $a = \sum_{i \in A} e_i$ and the radius $b = \sqrt{|A| - 1}$. Then the classifier $I\{ f_{a,b}(x) \leq 0 \}$ realizes the labels. This means that the set $\{ x_1, \ldots, x_{k+1} \}$ is shattered by $S_k$.

Then we show that $\mathcal{V}(S_k) \leq k + 1$. If there is a set $V = \{ x_1, \ldots, x_{k+2} \}$ shattered by $S_k$. Then there is a Radon partition $V = A \cup B$, such that we have a sphere $S_A$ that contains $A$ but not $B$, and another sphere $S_B$ that contains $B$ but not $A$. Whether $S_A \cap S_B = \emptyset$ or not, we can find a hyperplane that separates $A$ and $B$. This is a contradiction, with the argument in Question 1.