Homework 3

Frederick J. Boehm

10/1/2020

Q1

Q1, part a

We proceed by induction. First, we consider the set $\mathcal{F}_{x_1}$. We need to show that the size of $\mathcal{F}_{x_1}$ is 2. This follows from the observation that the sets $\{0\}$ and $\{1\}$ are in $\mathcal{F}_{x_1}$.

Now, assume that the size of $\mathcal{F}_{x_1,\ldots,x_n}$ is $n+1$. We want to show that the size of $\mathcal{F}_{x_1,\ldots,x_{n+1}}$ is $n+2$.

Suppose, without loss of generality, that the x_i, for $i = 1, \ldots, n$ are ordered such that $x_1 < x_2 < \ldots < x_n$.

Denote the $n+1$ elements of $\mathcal{F}_{x_1,\ldots,x_n}$ by $\{0\ldots0\}$, $\{0\ldots01\}$,$\{0\ldots011\}$, $\ldots$, $\{1\ldots1\}$, where each element contains n digits, all zeros and ones, and the k^{th} element ends with $k-1$ ones.

Observe that each element of $\mathcal{F}_{n+1}$ has $n+1$ binary entries. x_{n+1} lies somewhere on the real line. If $x_{n+1} > x_n$, then we can specify the elements of $\mathcal{F}_{n+1}$ as resulting from the appending of a one to every element of $\mathcal{F}_n$ along with an additional element that results from the $(n+1)$ points all mapping to zero. It's a little trickier when x_{n+1} lies between x_k and x_{k+1}. To illustrate the idea, suppose that $k = 2$, so x_{n+1} is between x_2 and x_3. Then the elements of $\mathcal{F}_{x_1,\ldots,x_{n+1}}$ result from inserting a digit between the second and third elements of each function in $\mathcal{F}_{x_1,\ldots,x_n}$. Whether that digit is a zero or one depends on the other digits. The key is to maintain the runs of consecutive zeros and consecutive ones. You then see that there are $n+2$ functions in $\mathcal{F}_{x_1,\ldots,x_{n+1}}$.

Q1, part b

Note that $\mathcal{A}$ can be written as $1_{[a,b]} + 1_{[c,d]}$. We then use reasoning like that above to see that there's a set of four points that can be shattered, but that no set of five points can be shattered by $\mathcal{A}$. To see that the four-point set $\Lambda = \{-5, -3, 0, 2\}$ can be shattered by $\mathcal{A}$, treat the subsets $\{-5, -3\}$ and $\{0, 2\}$ like we did above. Since the 16 functions all are restrictions of elements of $\mathcal{A}$ to Λ, we know that the vc dimension of $\mathcal{A}$ is at least 4. Now, we want to consider a five-point set, $\{p, q, r, s, t\}$ with $p < q < r < s < t$. Since there are five points, any partitioning of the five points into two groups yields a set with at least three points, we see that there is an "interior" point in one of the intervals. This is problematic because we can't get the "interior" point mapped to zero while the two exterior points map to one.

Q1, part c

We first show that vc dimension of the collection of axis-aligned squares is at least 3; that is, there exists a set of three points that is shattered by the collection of such squares.

Consider the points in $\mathbb{R}^2$: $(0, 0)$, $(2, 0)$, and $(1, -1)$. For each point, we can draw a square that encompasses exactly that one point and neither of the other two. Similarly, we can draw squares that encompass all subsets of size 2 (while excluding exactly one point per square). Finally, it's possible to draw a square that contains all three points.

Now, we want to show that no set of four points can be shattered by the collection of such squares.

Without loss of generality, we assume that the set of four points share no coordinates with each other so that we can identify a unique point with the greatest y value and that with the least y value. Similarly, we identify the point with the greatest and least x values. Suppose, without loss of generality, that the difference between the greatest and least x values is less than the difference between the greatest and least y values. Then, see that there is no square that assigns value 1 to those points with the greatest and least y values while assigning 0 to those points with the greatest and least x values. Thus, a set of four points can't be shattered by the collection of axis-aligned squares.

Q1, part d

Here, we need to show that there's a set of four points that is shattered by the axis aligned rectangles, but that no set of five points is shattered by the same.

Consider the four points: $(0, 2)$, $(1, 4)$, $(3, 3)$, and $(2, 1)$.

That a single function can map all four points to one and that we can draw a rectangle around exactly one point (for each of the four points) is apparent. Similarly, one can draw a single rectangle around any pair of the four points. This is especially apparent with our choice of points. With the four points having distinct x values and distinct y values and all being on the convex hull, we recognize that one can draw a rectangle around any subset of exactly three points.

Now, take a set of any five points. The smallest rectangle that encloses the five points is actually defined by only four points, since there are four edges. Thus, a fifth point is either (also) on an edge or in the interior of the smallest rectangle.

Thus, we can't use rectangles to choose all subsets of size four.

Q1, part e

We first show that a set of three points can be shattered by the set of circles.

Consider the points $(0, 0)$, $(0, 2)$, and $(2, 0)$. It's then apparent that we can draw a circle around any subset of the points. For example, center a circle at (0,1) with radius $1 + \epsilon$ to get only the two points on the y axis. Center a circle at $(2, 2)$ - with radius $2 + \epsilon$ to get only the points not at the origin.

Now, we need to show that no set of four points can be shattered by the collection of circles.

We assume that the four points have a quadrilateral convex hull. If they don't then we're done by the above results.

For the quadrilateral convex hull, let the difference in the extremes of y axis values be less than the difference in extremes of the x axis values.

Then we can't get a circle that contains the two points with the greatest x axis difference while avoiding the two points with the greatest y axis difference.

Thus, a set of four points can't be shattered by circles. So vc of the collection of circles is 3.

Q2

Q2, part a

We recognize that the expected value of $|\{i : X_i \in C\}|$ is $area(C)n$. We thus need a statement about the deviation of $|\{i : X_i \in C\}|$ from its expected value.

$$\mathbb{E}(\sup_{C \in \mathcal{C}} |\frac{|\{i : X_i \in C\}|}{n} - area(C)|) \leq K\sqrt{\frac{vc(\mathcal{C})}{n}} = K\sqrt{\frac{3}{n}}$$

Our bound is achieved by way of Theorem 8.3.23 in Vershynin's book.

Q2, part b

Q3

We follow the hint and start by placing n points on the unit circle. Every subset of the n poitns is a convex polygon that by definition doesn't contain the points that aren't in the subset. This property indicates that the a set of size n can be shattered by the collection of d-gons, with $d \leq n$. Letting n be arbitrarily large, we see that a set of n points on the unit circle can be shattered by the collection of d-gons ($d \leq n$). Thus, such sets of arbitrarily large size can be shattered, so the vc dimension is ∞.

Q4

We first recall that

$\mathcal{F}_{x_1,\ldots,x_n} = \{(f(x_1), \ldots, f(x_n))\}$ and $\mathcal{H}_{x_1,\ldots,x_n} = \{(h(x_1), \ldots, h(x_n))\}$

We know that T maps to $\{0, 1\}$, so $h \in \mathcal{H}$ also maps to $\{0, 1\}$. Suppose that the vc dimension of $\mathcal{F}$ is d. We then ask, is the vc dimension of $\mathcal{H}$ also equal to d?

We note that

$$\mathbb{E}((f(X) - T(X))^2) = \mathbb{E}(|f(X) - T(X)|) = \mathbb{P}(f(X) \neq T(X))$$

Q5

Q5, part a

We use Lemma 1.38 to see that

$$\mathbb{P}\left(\sup |\frac{1}{n}\sum f(X_i) - \mathbb{E}f(X)| > t\right) \leq 2\mathbb{P}\left(\sup |\frac{1}{n}\sum f(X_i) - f(Y_i)| > \frac{t}{2}\right)$$

The above inequality implies that

$$\int \mathbb{P}\left(\sup |\frac{1}{n}\sum f(X_i) - \mathbb{E}f(X)| > t\right) dt \leq \int \mathbb{P}\left(\sup |\frac{1}{n}\sum f(X_i) - f(Y_i)| > \frac{t}{2}\right) dt$$

From this, we get the desired result because

$$\int \mathbb{P}\left(\sup |\frac{1}{n}\sum f(X_i) - \mathbb{E}f(X)| > t\right) dt = \mathbb{E}\left(\sup |\frac{1}{n}\sum f(X_i) - \mathbb{E}f(X)|\right)$$

and

$$\int \mathbb{P}\left(\sup |\frac{1}{n}\sum f(X_i) - f(Y_i)| > \frac{t}{2}\right) dt = \mathbb{E}\left(\sup |\frac{1}{n}\sum f(X_i) - f(Y_i)|\right)$$

Q5, part b