Statistical Learning
Homework 1

Fall 2024

Chenchuan He

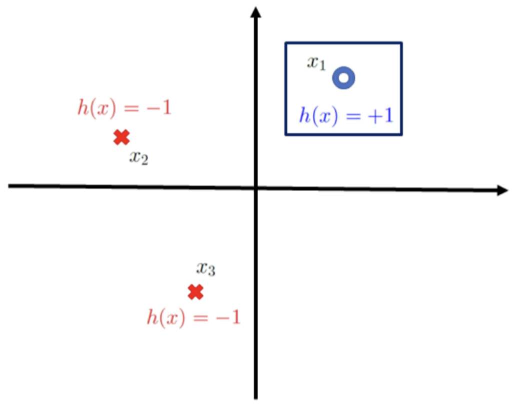

Figure 1: $\mathcal{H}$ - Positive rectangles, $N = 3$.

2. Consider the learning model of positive rectangles i.e. $\mathcal{H}$ contains rectangles that are aligned horizontally or vertically and are positive in the inside and negative elsewhere as it is shown in Figure 1. Determine:

   (a) The breaking point k and the VC dimension, d_{VC}. Note: Demonstrate your results, showing a set of points that cannot be shattered by $\mathcal{H}$.

   (b) A bound for the growth function:

$$m_{\mathcal{H}}(N) \leq \sum_{i=0}^{d_{VC}} \binom{N}{i}$$

Solution:
(a). Intuitively, the break point $k = 5$, and $d_{VC} = 4$.
To prove $k = 5$, firstly show 4 points can be classified. Suppose 4 points are in (-2,0), (0,2), (2,0), (0,-2), Then, there exists $\mathcal{H}$ such that any types of labels can be applied, so that $\mathcal{H}$ shatters all the 4 points.
However, if there are 5 points, all the points that cannot be shattered by H. Suppose we randomly select 5 points in the axis, and we randomly select the 4 of them and draw a smallest rectangle for the four points.
1. If the last point is in the middle of the rectangle, then we cannot classify it as negative while all other points as positive,
2. If the last point fells outside the rectangle, then we can always find three other points and form a new rectangle, such that the remaining poing would fall in the middle, making the shatterings infeasible.
To conclude: the break point $k = 5$, and $d_{VC} = 4$

(b). Since $d_{VC} = 4$, we need to prove

$$m_{\mathcal{H}}(N) \leq \sum_{i=0}^{4} \binom{N}{i}$$

To prove this, we only need to prove $m_{\mathcal{H}}(5) < 2^k$ according to the theorem.

We know $m_{\mathcal{H}}(5) < \sum_{i=0}^{5} \binom{5}{i} = 32$, as proved in (a) that all shatterings are infeasible when $k = 5$, while this equation simply captures all the possible labelling combinations of the 5 points.

Thus, we conclude that

$$m_{\mathcal{H}}(N) \leq \sum_{i=0}^{d_{VC}} \binom{N}{i}$$

3. Remember the inequality for multiple hypotheses:

$$\Pr\left[|E_{\text{in}}(g) - E_{\text{out}}(g)| \geq \epsilon\right] \leq 2Me^{-2N\epsilon}.$$

If we replace M by $m_{\mathcal{H}}(N)$ which can be bounded by a polynomial, the generation error will go to zero as $N \to \infty$ which implies learning is feasible. To prove this, assume $m_{\mathcal{H}}(N)$ can be bounded by the polynomial N^{k-1} and compute the following simplified limit for $\epsilon > 0$ and k being a finite positive integer (i.e. $k \in \mathbb{Z}^+, 0 < k < \infty$):

$$\lim_{N\to\infty} N^{k-1}e^{-N}.$$

Solution:
To prove

$$\lim_{N\to\infty} N^{k-1}e^{-N} = \lim_{N\to\infty} \frac{N^{k-1}}{e^N}$$

simply use L'Hopital's rule as N is differentiable when approaching infinity and,

$$\lim_{N\to\infty} N^{k-1} \to \infty$$

,

$$\lim_{N\to\infty} e^N \to \infty$$

$$\lim_{N\to\infty} \frac{N^{k-1}}{e^N} = \lim_{N\to\infty} \frac{\frac{dN^{k-1}}{dN}}{\frac{de^N}{dN}} = \lim_{N\to\infty} \frac{(k-1)N^{k-2}}{e^N}$$

After applying the L'Hopital's rule for $k-1$ times, we have

$$\lim_{N\to\infty} \frac{N^{k-1}}{e^N} = \lim_{N\to\infty} \frac{(k-1)!}{e^N} = 0$$

Thus, the bound has been proved.