Analysis - HW4

Jianxiang Fan jianxiang.fan@colorado.edu

1. Order of hypothesis classes in terms of complexity:

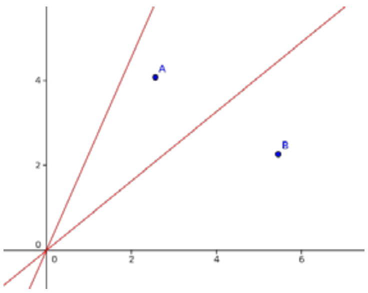
Hyperplanes through the origin < arbitrary hyperplanes < axis-aligned rectangles

Experimental result of Rademacher Complexity (20 random points in [-20, 20] x [-20, 20]):

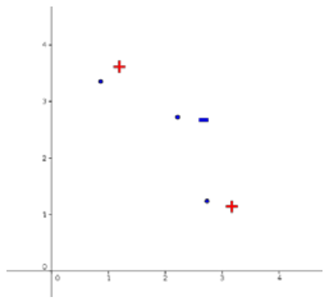
Axis-aligned rectangles: 0.573000 Hyperplanes (origin): 0.400500 Arbitrary hyperplanes: 0.528100

VC-Dimension analysis:

(a) Hyperplanes through the origin: VC-Dim = 2. Two points can be shattered (a). Three points: If we let the left and right ones be positive, and the middle one be negative, (b) no hypothesis can realize this labeling.

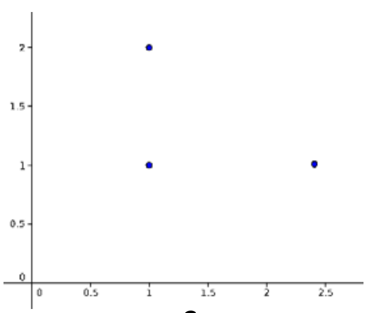


a

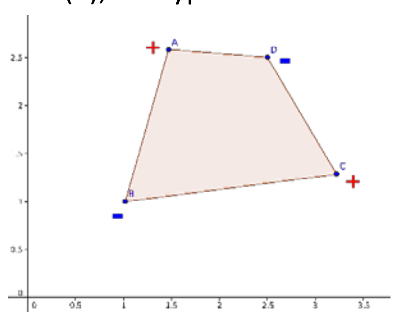


b

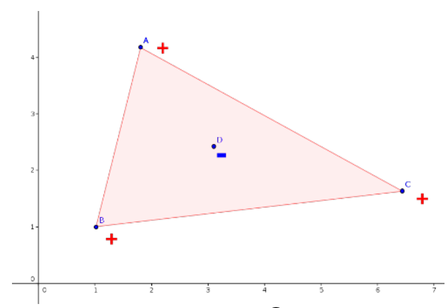
(b) Arbitrary hyperplanes: VC-Dim = 3. Three points can be shattered (c). Four points: (Case 1) The four points can make up a quadrangle (d): let one opposite pair be positive and another pair be negative, no hypothesis can realize this labeling. (Case 2) The four points can't make up a quadrangle: Let the interior point be negative and the others be positive (e), no hypothesis can realize this labeling.



c

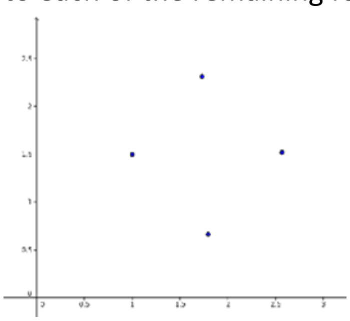


d

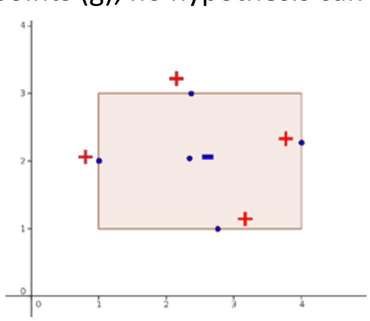


e

Axis-aligned rectangles: VC-Dim = 4. Four points can be shattered (f). Five points: If we construct the minimal axis-aligned rectangle containing these points, assign a negative label to this interior point and a positive label to each of the remaining four points (g), no hypothesis can realize this labeling..



f



g

2. Prove that your frequency correctly classifies any training set

Frequency: $\omega = \pi\left(1 + \sum_{i=1}^{m} \frac{1-y_i}{2} 2^{x_i}\right)$

For any j $(1 \le j \le m)$

$$\omega 2^{-x_j} = \pi\left(2^{-x_j} + \sum_{i=1}^{m} \frac{1-y_i}{2} 2^{x_i - x_j}\right) = \pi\left(2^{-x_j} + \frac{1-y_j}{2} + \sum_{i=1}^{x_j-1} \frac{1-y_i}{2} 2^{x_i - x_j}\right)$$

$$\omega 2^{-x_j} \le \pi\left(2^{-x_j} + \frac{1-y_j}{2} + \sum_{i=1}^{x_j-1} 2^{x_i - x_j}\right) = \pi\left(\frac{1-y_j}{2} + \sum_{i=1}^{x_j} 2^{-i}\right) < \pi\left(\frac{1-y_j}{2} + 1\right)$$

$$\omega 2^{-x_j} \ge \pi\left(2^{-x_j} + \frac{1-y_j}{2}\right) > \pi \frac{1-y_j}{2}$$

If $y_j = 1,\ 0 < \omega 2^{-x_j} < \pi$, then $\sin(\omega 2^{-x_j}) > 0,\ h(x_j) = \text{sign}(\sin(\omega 2^{-x_j})) = 1$

If $y_j = -1,\ \pi < \omega 2^{-x_j} < 2\pi$, then $\sin(\omega 2^{-x_j}) < 0,\ h(x_j) = \text{sign}(\sin(\omega 2^{-x_j})) = -1$

3. Classifying real numbers

Choose four points x = 1, 2, 3, 4. They can't be shattered by this sine classifier.
Proof:
Suppose there is an ω labeling these four points as +, -, -, -, thus:

$$\sin\omega \ge 0, \sin 2\omega < 0, \sin 3\omega < 0, \sin 4\omega < 0$$

Since $\sin 4\omega = 2\sin 2\omega \cos 2\omega = 2 \sin 2\omega\,(1 - 2\sin^2\omega) < 0$ and $\sin 2\omega < 0$,
we have $1 - 2\sin^2\omega > 0$, then $\sin^2\omega < 1/2$.
While $\sin 3\omega = 3\sin\omega - 4\sin^3\omega = \sin\omega\,(3 - 4\sin^2\omega) < 0$ and $\sin\omega \ge 0$,
we have $3 - 4\sin^2\omega < 0$, then $\sin^2\omega > 3/4$.
These two results lead to a contradiction, so this classifier can’t realize this labeling.

Although the sine classifier can’t shatter these four points, this conclusion is not related to its VC dimension. Actually the VC-dimension of the family of sine functions is infinite, as we can see in question 2.