

Problem 1

As stated by the problem, we have $\epsilon = .15$ —the error of the hypothesis h_i . The probability of hypothesis h being outputted with error ϵ is the confidence of the hypothesis which is given by the equation $1 - \delta$. In this particular problem, $1 - \delta = .95$. This means $\delta = .05$.

This problem has a finite, consistent hypothesis class H . It is finite because each hypothesis h_i is a triangle with 3 distinct vertices on the interval $[0,99]$ —there is a maximum number of triangles. It is consistent because each h_i determines if a given training example x_i is inside or outside of its boundaries. This is aligned with the concept c of labeling each point positive or negative depending on whether that point is interior and exterior to the triangle boundary. There is an algorithm A that exists that given a point x_i and 3 vertices that comprise a triangle, will tell if that point is inside or outside of the triangle boundary.

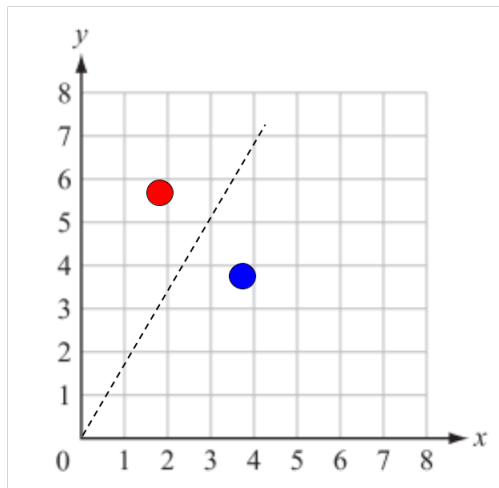
Because this problem has a finite, consistent hypothesis class we can use the following equation to find the bound on training examples m to ensure each hypothesis has a confidence of 95% and error of .15.

$$m \geq \frac{1}{\epsilon} (\ln |H| + \ln \frac{1}{\delta})$$

$|H|$ is the number of hypothesis which is equal to the number of triangles that can be found in the problem. The number of triangles that can be found in a given interval is the total number of combinations of the total number of points in that interval made up of 3 vertices $\binom{n}{3}$. n in the case when the interval for both x and y $[0,99]$ is 200 points. This means there are $\binom{200}{3} = 1313400$ possible triangles. This means $|H| = 1313400$. Plugging $|H|, \epsilon, \delta$ in to the equation for above we get

$$m \geq \frac{1}{.15} (\ln(1313400) + \ln \frac{1}{.05}) \approx 114 \text{ training samples}$$

Problem 2



(a) Lower VC Dim bound example

The figure above is an example of a configuration of 2 points that are shattered by H . Because this is the lowest number of points H can shatter, the $VCdim(H) \geq 2$.