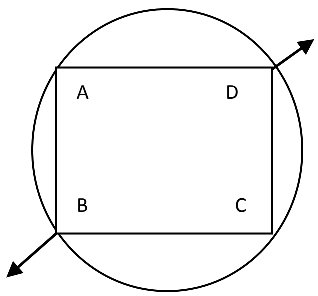
CS 4850: Mathematical Foundations for the Information Age April 10, 2009

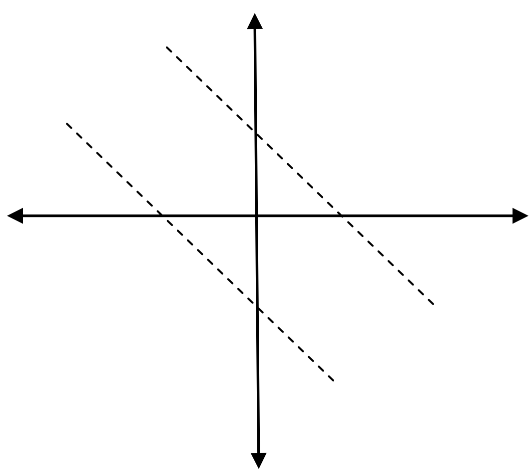
Lecture Notes 33

Professor: *John Hopcroft* Scribe: *Hari Sreedharan(hs465)*
Hari Sankar (hss78)

Correction from previous lecture: In the proof of shattering 4 points for a circle, the points A and C should necessarily be fixed and only B and D should be allowed to move outside the circle.

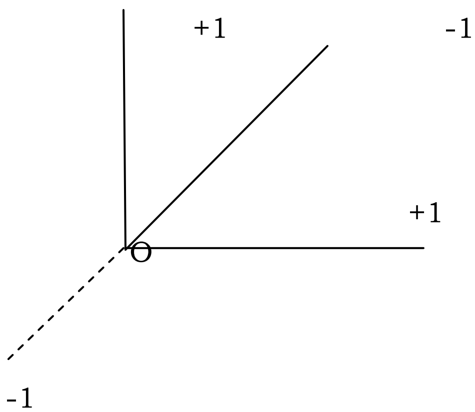


VC-Half Space

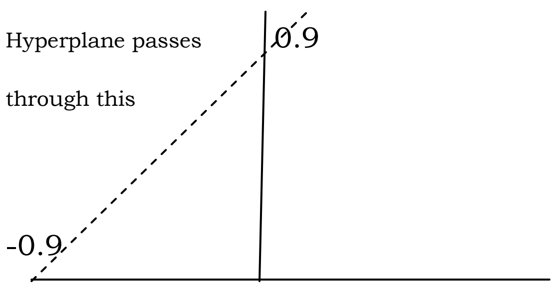


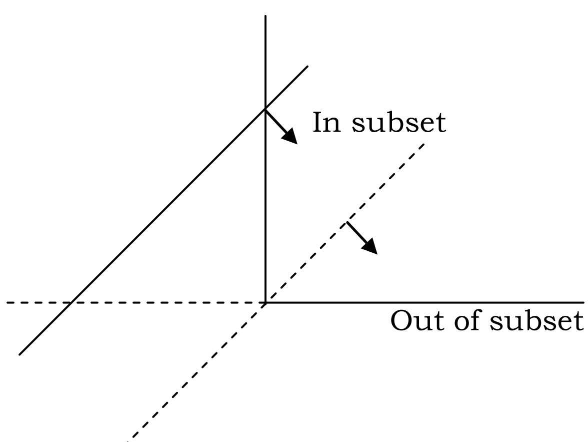
VC-Dimension of Half space is n+1. I.e., some set of n+1 points can be shattered in VC-Half Space.

Suppose n unit coordinate vectors and the origin is selected.



Finding the hyperplane that satisfies n equations and finding whether a point is on the right or left side of the plane.





To prove that a point is on the right or left side of the plane, we need to take the dot product.

Taking dot product, x.a1 = +1 (within)

x.a2=-1(outside)

Taking the Threshold = +1/2 or -1/2

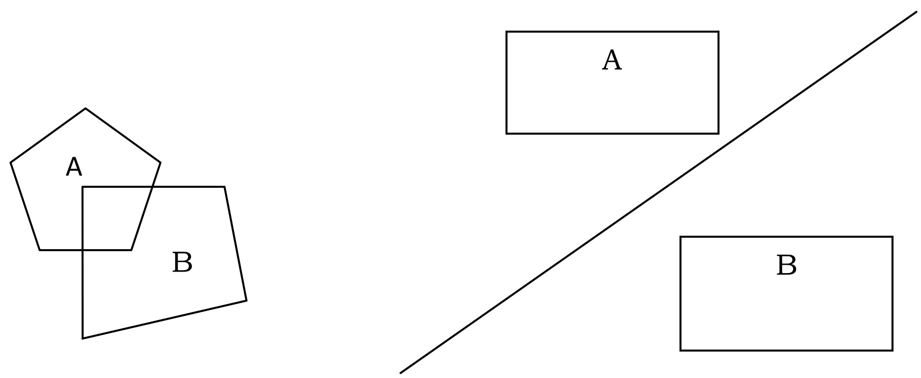
This proves that VC-Dimension is n+1

VC – Dimension is atmost n+1 points.

To prove this, we need to prove that we cannot shatter any set of n+2 points.

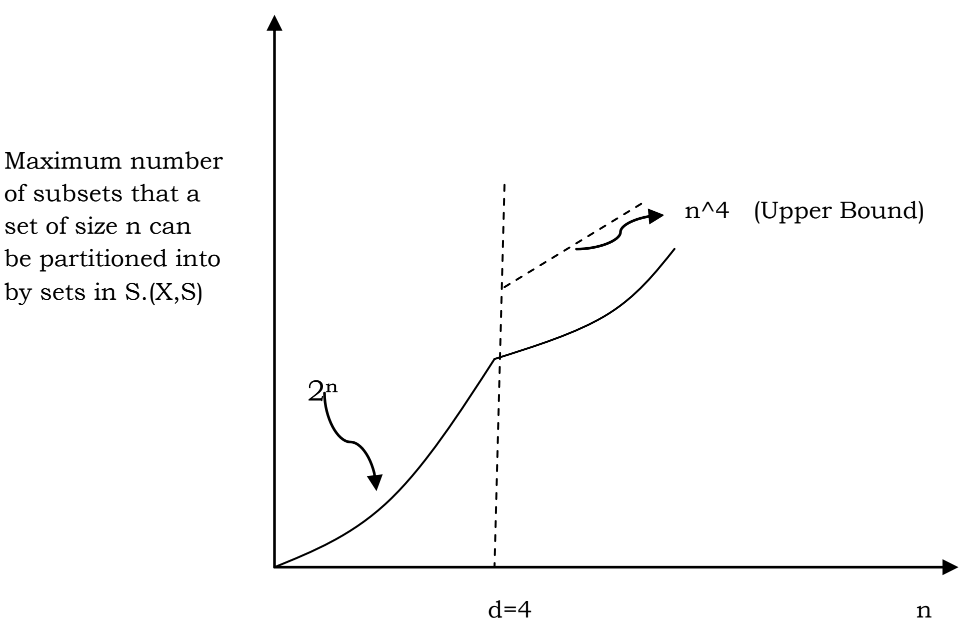
Proof by Contradiction: Partition the n+2 points into 2 sets A and B.The convex hull of A and convex hull of B overlap or convex(A) convex(B) $\neq \emptyset$

Intuitively it can be seen that if there are n+2 points, the convex hull overlap.(Shown in the diagram below).If there are n+1 points there exists a plane which can divide these n+1 points into 2 subspaces. Hence the VC-Dimension of half space is n+1.



Primal Shatter Function

A primal shatter function of an integer n is the maximum number of subsets that a set of size n can be partitioned into by sets in S.



It can be shown that a partial shatter function of a figure grows exponentially from n = 0 to the VC dimension of the figure, and then grows only at the rate of a polynomial. I.e., if the set system is finite dimension, the nature of the graph goes as 2^n until that dimension. After that the nature of the graph becomes polynomial. For rectangle, the polynomial is function, n^4.

Define 2 rectangles equivalent if they define the same set of subset. How many equivalence classes of rectangles are there?

Shrink the rectangles to get canonical equivalent rectangle. How many canonical rectangles can I get?

- Number the points.
- Assign the lowest numbered point to an edge.
- There are at most n^4 equivalent rectangles.
- Since two equivalent triangles define the same points, at most n^4 subsets can be defined.

For a set system of finite VC Dimensions, growth of the # of subsets that can be obtained is polynomial in n.

For VC Dimension d, upper bound is the summation:

$$\sum_{i=0}^{d} \binom{n}{i} = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{d} \leq n^d$$

Proof:
By induction on n and for fixed n, induction on d.

Base case: For n, use n ≤ d.

$$\prod_{s}(n) = 2^n \; for \;\; d \geq n, \sum_{i=0}^{d} \binom{n}{i} = 2^n$$

Base Case for d: d=0

A set of size 0 can have VC Dimension of at most 1.