Definition of VC dimension

The VC dimension of a hypothesis set $\mathcal{H}$, denoted by $d_{\text{VC}}(\mathcal{H})$, is

the largest value of $N$ for which $m_\mathcal{H}(N) = 2^N$

“the most points $\mathcal{H}$ can shatter”

$N \leq d_{\text{VC}}(\mathcal{H}) \implies \mathcal{H}$ can shatter $N$ points

$k > d_{\text{VC}}(\mathcal{H}) \implies k$ is a break point for $\mathcal{H}$
The growth function

In terms of a break point $k$:

$$m_H(N) \leq \sum_{i=0}^{k-1} \binom{N}{i}$$

In terms of the VC dimension $d_{VC}$:

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

maximum power is $N^{d_{VC}}$
Examples

• $\mathcal{H}$ is positive rays:
  
  \[ d_{VC} = 1 \]

• $\mathcal{H}$ is 2D perceptrons:
  
  \[ d_{VC} = 3 \]

• $\mathcal{H}$ is convex sets:
  
  \[ d_{VC} = \infty \]
VC dimension and learning

\[ d_{VC}(\mathcal{H}) \text{ is finite} \implies g \in \mathcal{H} \text{ will generalize} \]

- Independent of the learning algorithm
- Independent of the input distribution
- Independent of the target function