Definition of VC dimension

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

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

"the most points $\mathcal{H}$ can shatter"

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

$k > d_{\mathrm{VC}}(\mathcal{H}) \implies$ k is a break point for $\mathcal{H}$

The growth function

In terms of a break point k:

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

In terms of the VC dimension d_{VC}:

$$m_{\mathcal{H}}(N) \leq \underbrace{\sum_{i=0}^{d_{\text{VC}}} \binom{N}{i}}_{\text{maximum power is } N^{d_{\text{VC}}}}$$

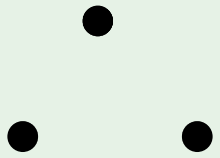
Examples

- $\mathcal{H}$ is positive rays:

$$d_{\text{VC}} = 1$$

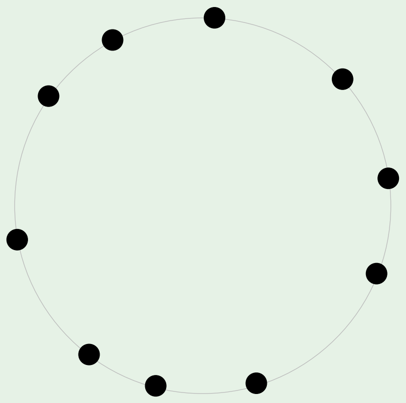
- $\mathcal{H}$ is 2D perceptrons:

$$d_{\text{VC}} = 3$$



- $\mathcal{H}$ is convex sets:

$$d_{\text{VC}} = \infty$$



VC dimension and learning

$d_{\mathrm{VC}}(\mathcal{H})$ is finite $\Longrightarrow$ $g \in \mathcal{H}$ will generalize

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