Outline

- The definition
- VC dimension of perceptrons
- Interpreting the VC dimension
- Generalization bounds

1. Degrees of freedom

Parameters create degrees of freedom

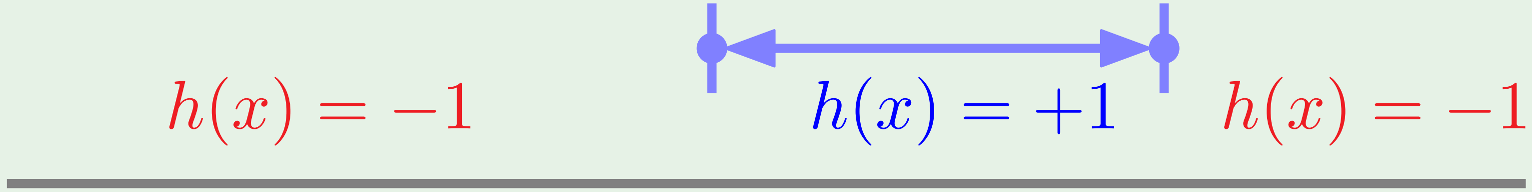
# of parameters: **analog** degrees of freedom

d_{VC}: equivalent **'binary'** degrees of freedom

The usual suspects

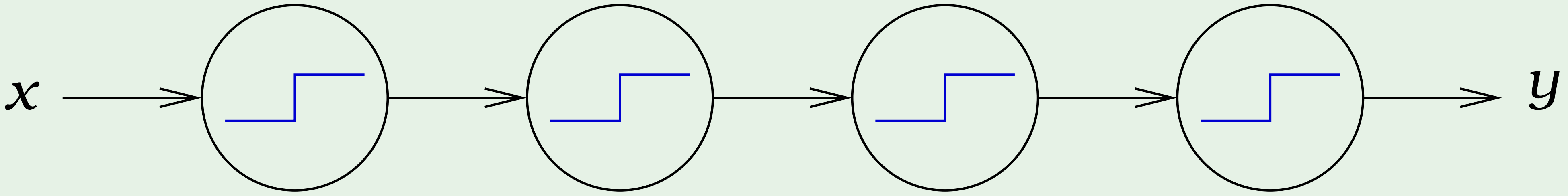
Positive rays ($d_{\text{VC}} = 1$):

Positive intervals ($d_{\text{VC}} = 2$):



Not just parameters

Parameters may not contribute degrees of freedom:



d_{VC} measures the **effective** number of parameters

2. Number of data points needed

Two small quantities in the VC inequality:

$$\mathbb{P}\left[\,|E_{\text{in}}(g) - E_{\text{out}}(g)| > \epsilon\,\right] \;\le\; \underbrace{4 m_{\mathcal{H}}(2N) e^{-\frac{1}{8}\epsilon^2 N}}_{\delta}$$

If we want certain ϵ and δ, how does N depend on d_{VC}?

Let us look at $N^{d} e^{-N}$

$N^{d}e^{-N}$

Fix $N^{d}e^{-N}$ = small value

How does N change with d?

Rule of thumb:

$$N \;\geq\; 10\, d_{\rm VC}$$