- Classification can be approached from the perspective of building a decision boundary that separates class labels, y, in the input space.

Classification as a decision boundary problem

6

$$y \in \{1,\ldots,K\}$$

$$\hat{y} = g(f(\mathbf{x}; \boldsymbol{\theta}))$$

decision boundary is a hyperplane, H:

In this lecture we'll assume a scalar label:

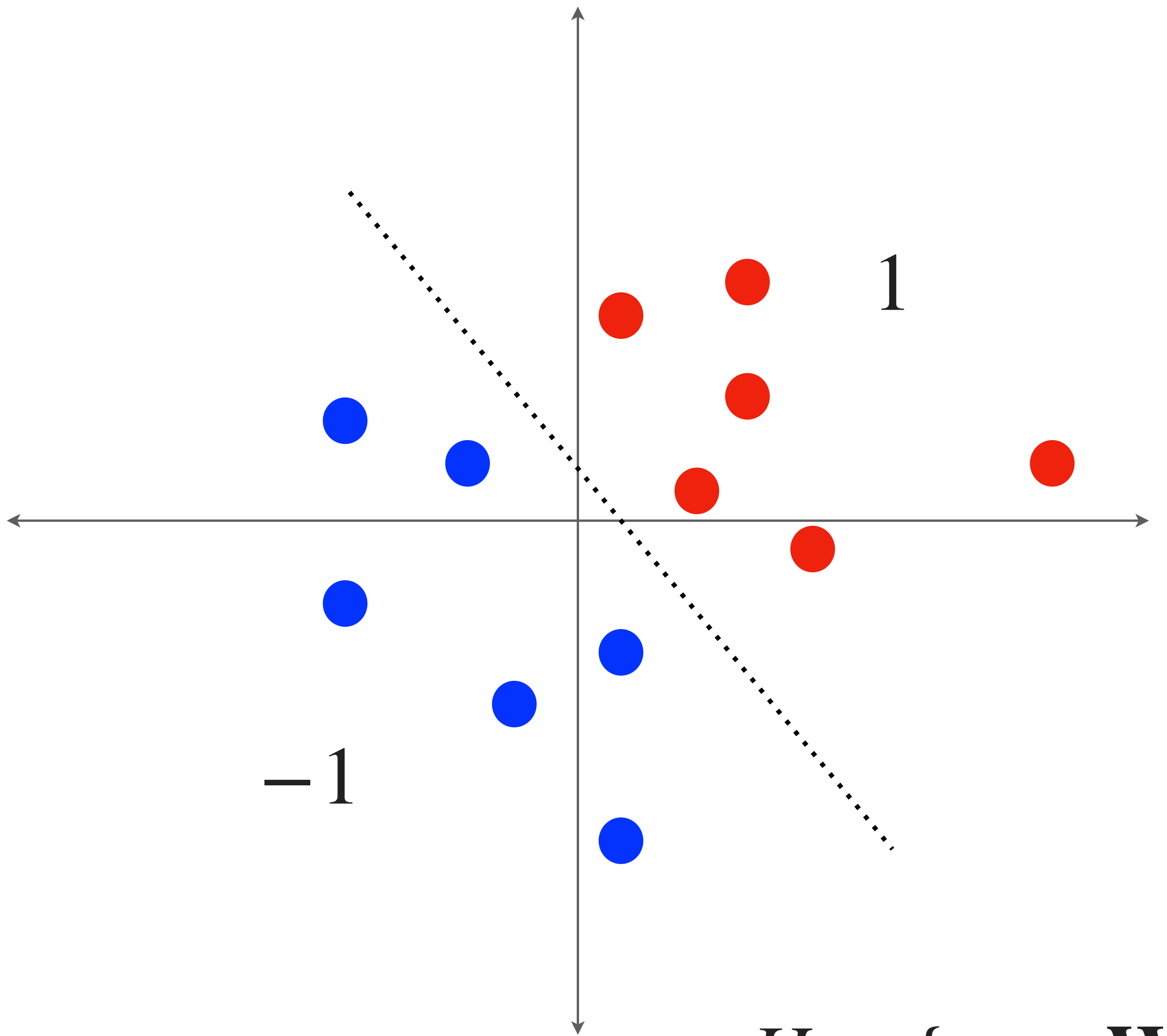
$$H = \{ \ \mathbf{x} : \ f(\mathbf{x}; \boldsymbol{\theta}) = \mathbf{0} \ \}$$

$$f: \mathbb{R}^N \to \mathbb{R}^N$$

decision boundary is represented by $f(\cdot)$:

then our predictor is $g : f(\mathbf{x}; \boldsymbol{\theta}) \rightarrow y :$

binary classification, linear decision boundary



$$H = \{\ \mathbf{x} :\ \ \mathbf{W}\mathbf{x} + \mathbf{b} = \mathbf{0}\ \}$$

$$y \in \{ \bullet\ , \bullet \}$$

Classification as a decision boundary problem

- Classification can be approached from the perspective of building a decision boundary that separates class labels, y, in the input space.

In this lecture we'll assume a scalar label:

$$y \in \{1, \ldots, K\}$$

decision boundary is a hyperplane, H:

$$H = \{\ \mathbf{x} :\ f(\mathbf{x}; \boldsymbol{\theta}) = \mathbf{0}\ \}$$

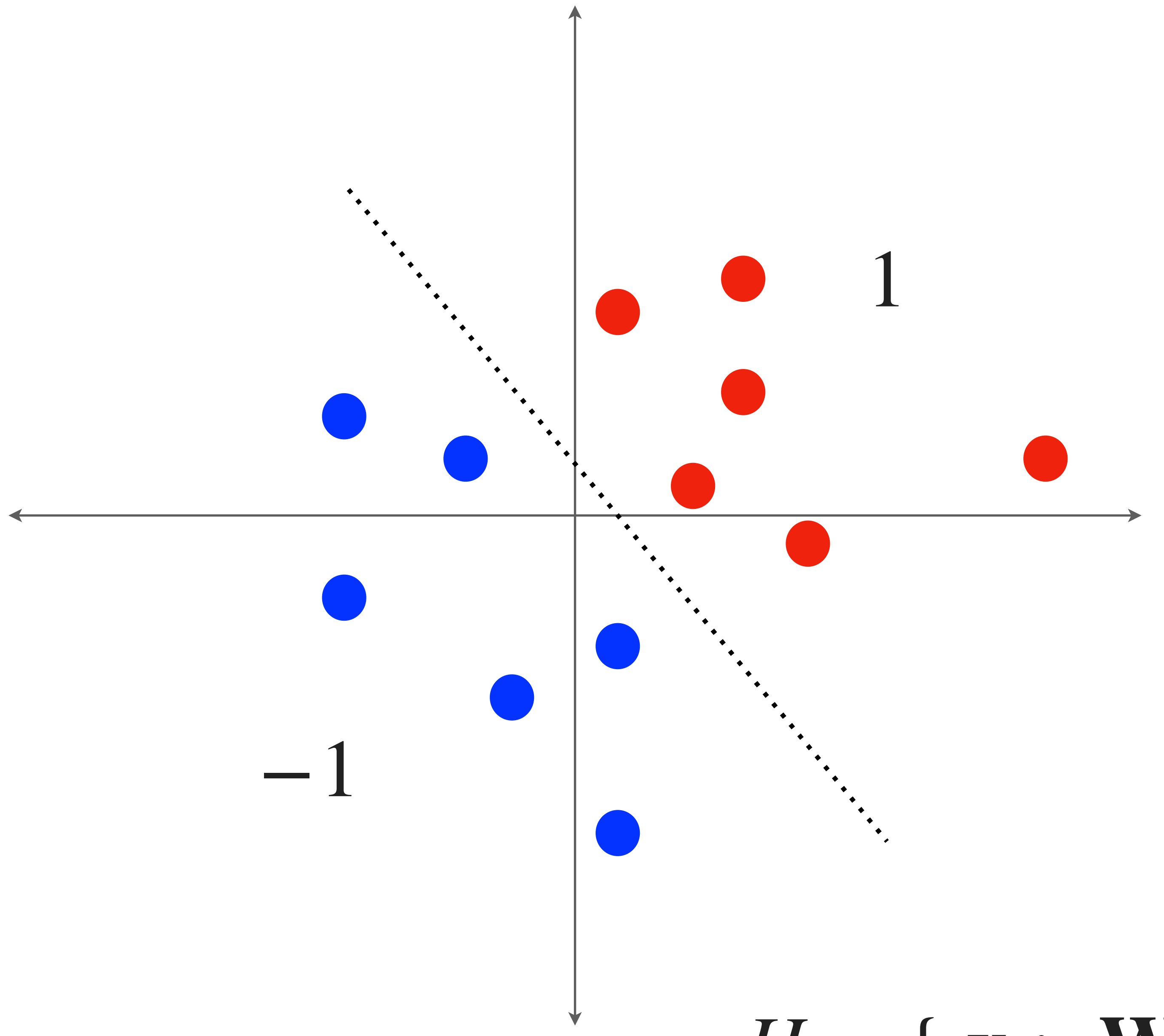
decision boundary is represented by $f(\cdot)$:

$$f : \mathbb{R}^N \rightarrow \mathbb{R}^N$$

then our predictor is $g : f(\mathbf{x}; \boldsymbol{\theta}) \rightarrow y$:

$$\hat{y} = g(f(\mathbf{x}; \boldsymbol{\theta}))$$

binary classification, linear decision boundary



$$H = \{\ \mathbf{x} :\ \mathbf{W}\mathbf{x} + \mathbf{b} = \mathbf{0}\ \}$$

$$y \in \{\bullet\ ,\bullet\}$$

Curse of dimensionality

- Somewhat counter intuitively, in high dimensional space, volume gets concentrated at the boundaries. In the world of decision boundaries, this means that presence of a linearly separating hyperplane increases with the number of dimensions, and in fact is guaranteed to exist at $N \to \infty$ … great news!?!?!