Classification

Machine Learning Techniques

Karthik Thiagarajan

Classification

$$D = \{(\mathbf{x}_1, y_1),\ \cdots, (\mathbf{x}_n, y_n)\}$$

Classification

$$D = \{(\mathbf{x}_1, y_1),\ \cdots, (\mathbf{x}_n, y_n)\}$$

$y_i \in \{0, 1\}$ OR $y_i \in \{-1, 1\}$

Classification

$$D = \{(\mathbf{x}_1, y_1),\ \cdots, (\mathbf{x}_n, y_n)\}$$

$y_i \in \{0, 1\}$ OR $y_i \in \{-1, 1\}$

$$h : \mathbb{R}^d \rightarrow \{0, 1\}$$

$$h(\mathbf{x}) = y$$

Classification

$$D = \{(\mathbf{x}_1, y_1),\ \cdots, (\mathbf{x}_n, y_n)\}$$

$y_i \in \{0, 1\}$ OR $y_i \in \{-1, 1\}$

$$h : \mathbb{R}^d \rightarrow \{0, 1\}$$

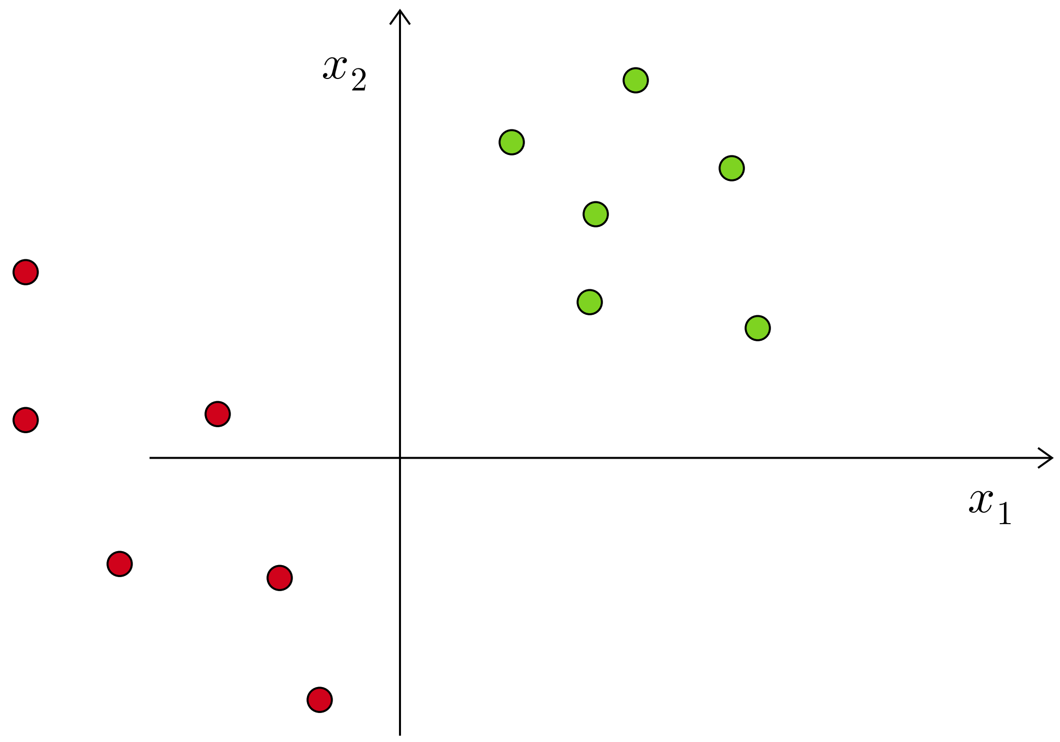
$$h(\mathbf{x}) = y$$

0-1 Loss

$$\text{Loss}(h) = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{1}[h(\mathbf{x}_i) \neq y_i]$$

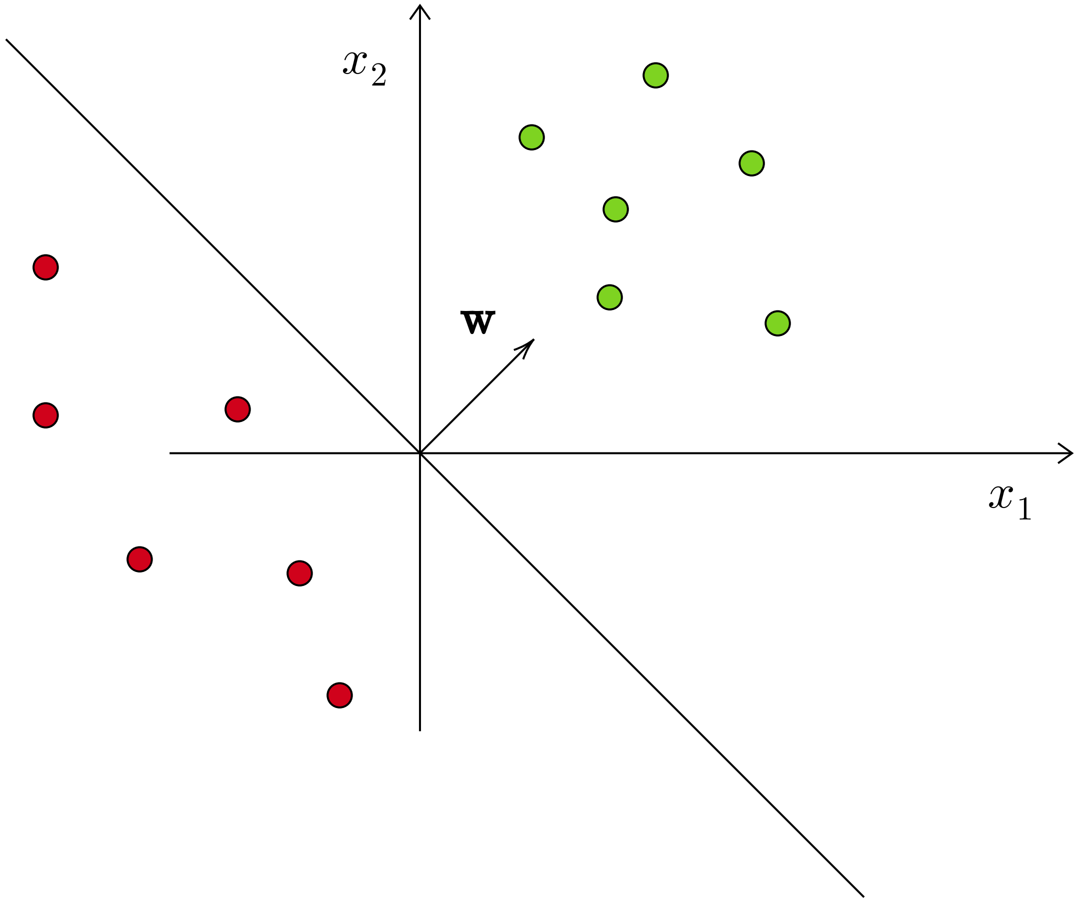
Linear Classifiers

class-1
class-0



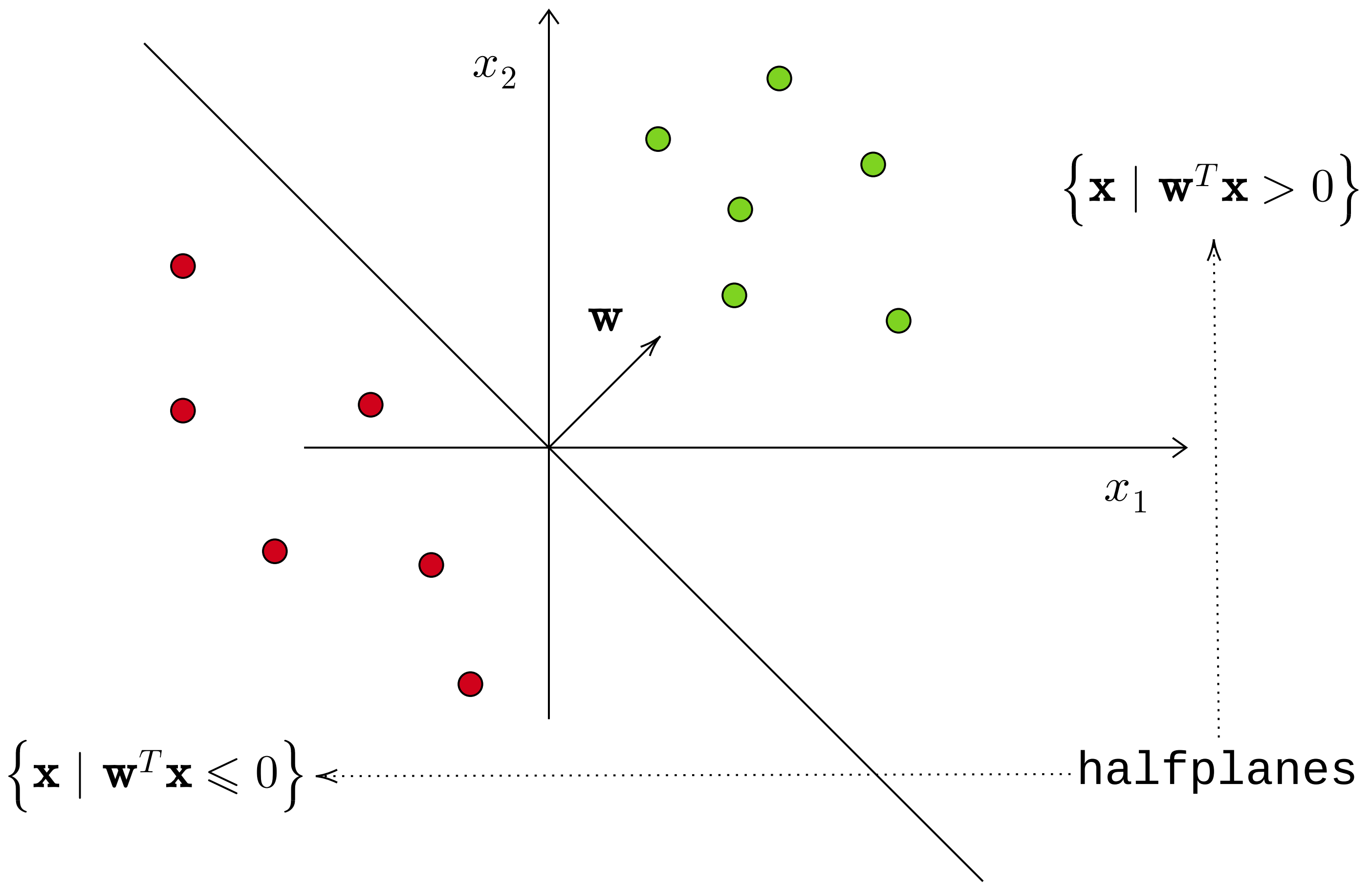
Linear Classifiers

class-1
class-0



Linear Classifiers

class-1
class-0



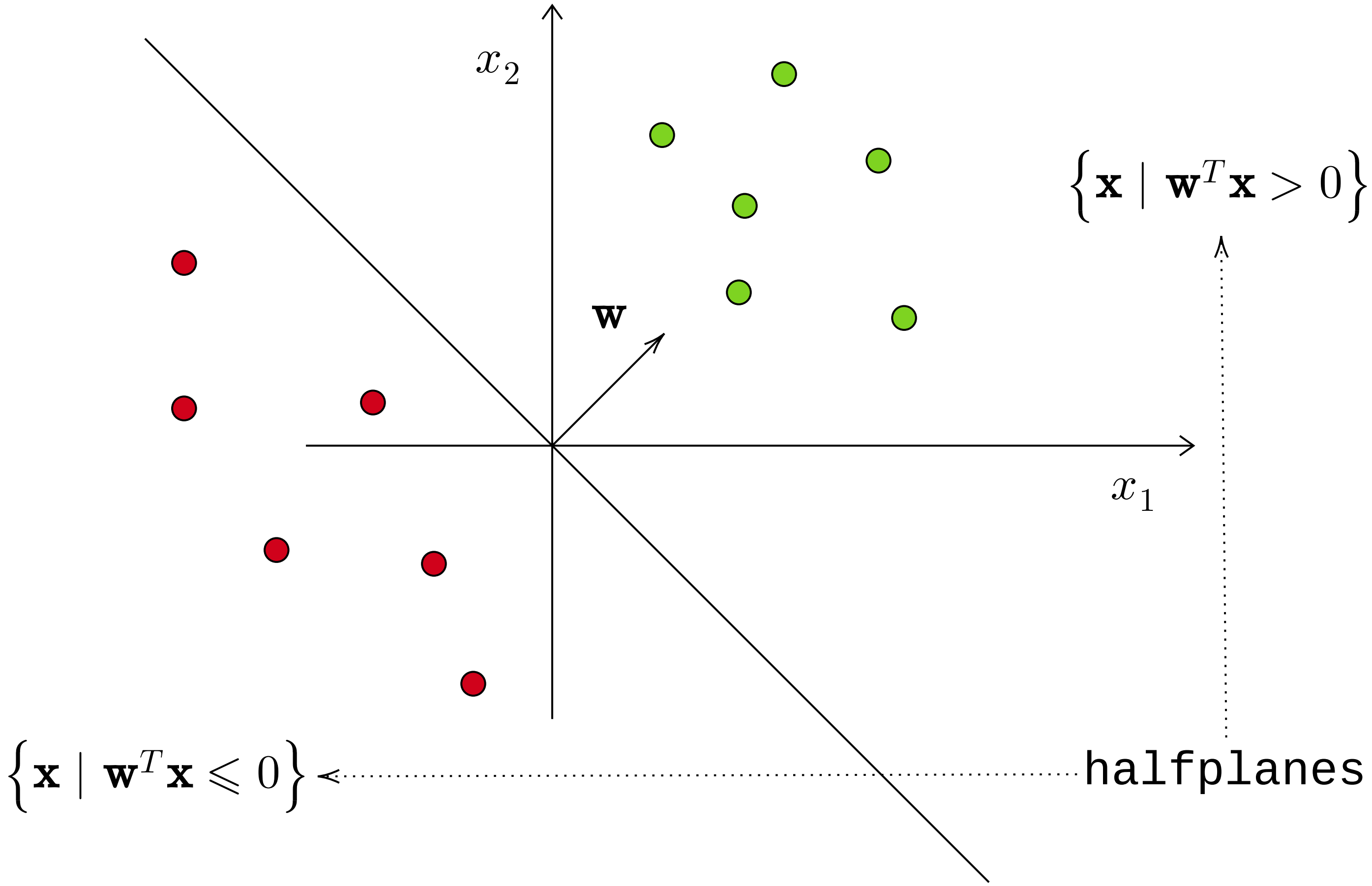
Linear Classifiers

$$h_{\mathbf{w}}(\mathbf{x}) = \mathtt{sign}\left(\mathbf{w}^T\mathbf{x}\right)$$

$$= \begin{cases} 1, & \mathbf{w}^T\mathbf{x} > 0 \\ 0, & \mathbf{w}^T\mathbf{x} \leqslant 0 \end{cases}$$

class-1

class-0



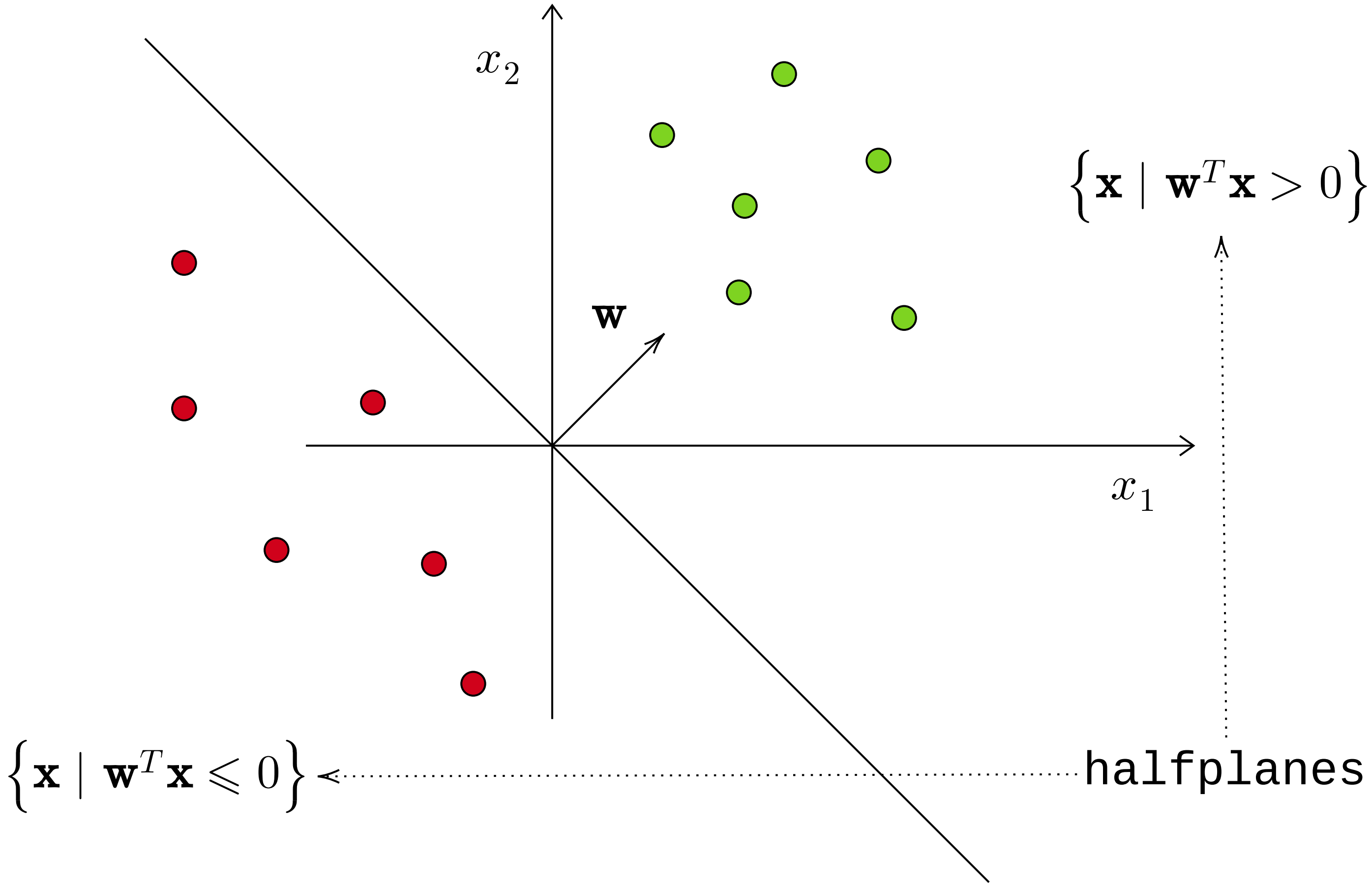
Linear Classifiers

$$h_{\mathbf{w}}(\mathbf{x}) = \mathtt{sign}\left(\mathbf{w}^T\mathbf{x}\right)$$

$$= \begin{cases} 1, & \mathbf{w}^T\mathbf{x} > 0 \\ 0, & \mathbf{w}^T\mathbf{x} \leqslant 0 \end{cases}$$

class-1

class-0

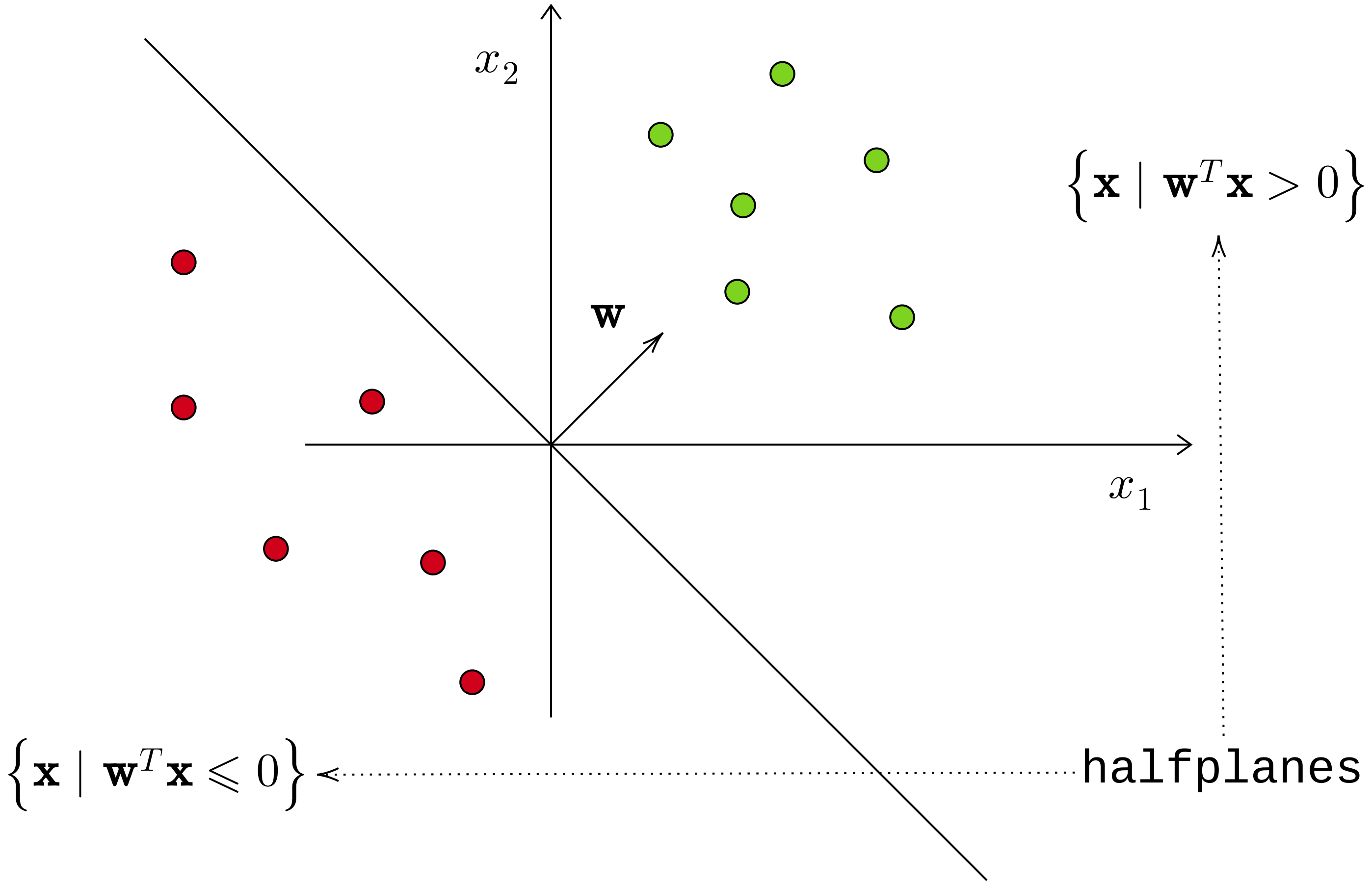


$$\mathcal{H}_{\mathtt{linear}} = \left\{h_{\mathbf{w}} \mid h_{\mathbf{w}}(\mathbf{x}) = \mathtt{sign}\left(\mathbf{w}^T\mathbf{x}\right)\right\}$$

Linear Classifiers

Linear Classifiers

class-1
class-0



$$h_{\mathbf{w}}(\mathbf{x}) = \mathtt{sign}(\mathbf{w}^T\mathbf{x})$$

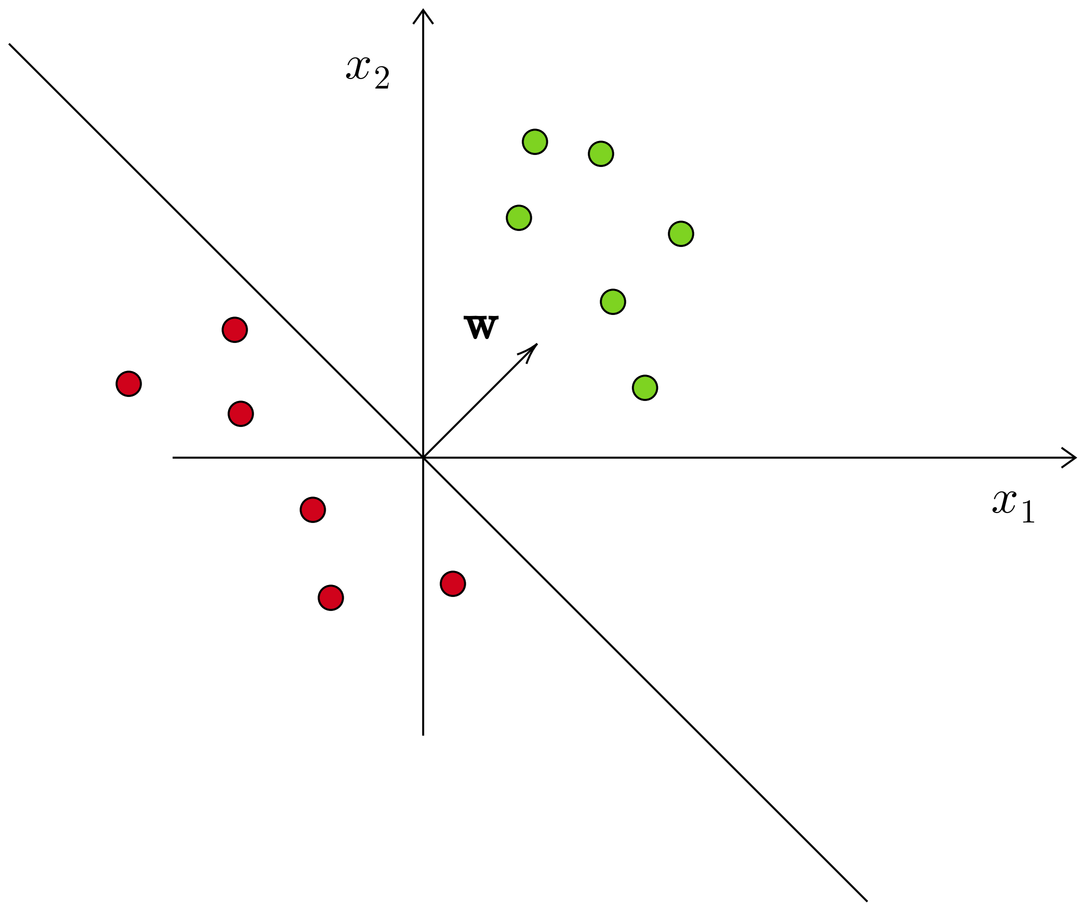
$$= \begin{cases} 1, & \mathbf{w}^T\mathbf{x} > 0 \\ 0, & \mathbf{w}^T\mathbf{x} \leqslant 0 \end{cases}$$

$$\mathcal{H}_{\mathtt{linear}} = \left\{ h_{\mathbf{w}} \mid h_{\mathbf{w}}(\mathbf{x}) = \mathtt{sign}(\mathbf{w}^T\mathbf{x}) \right\}$$

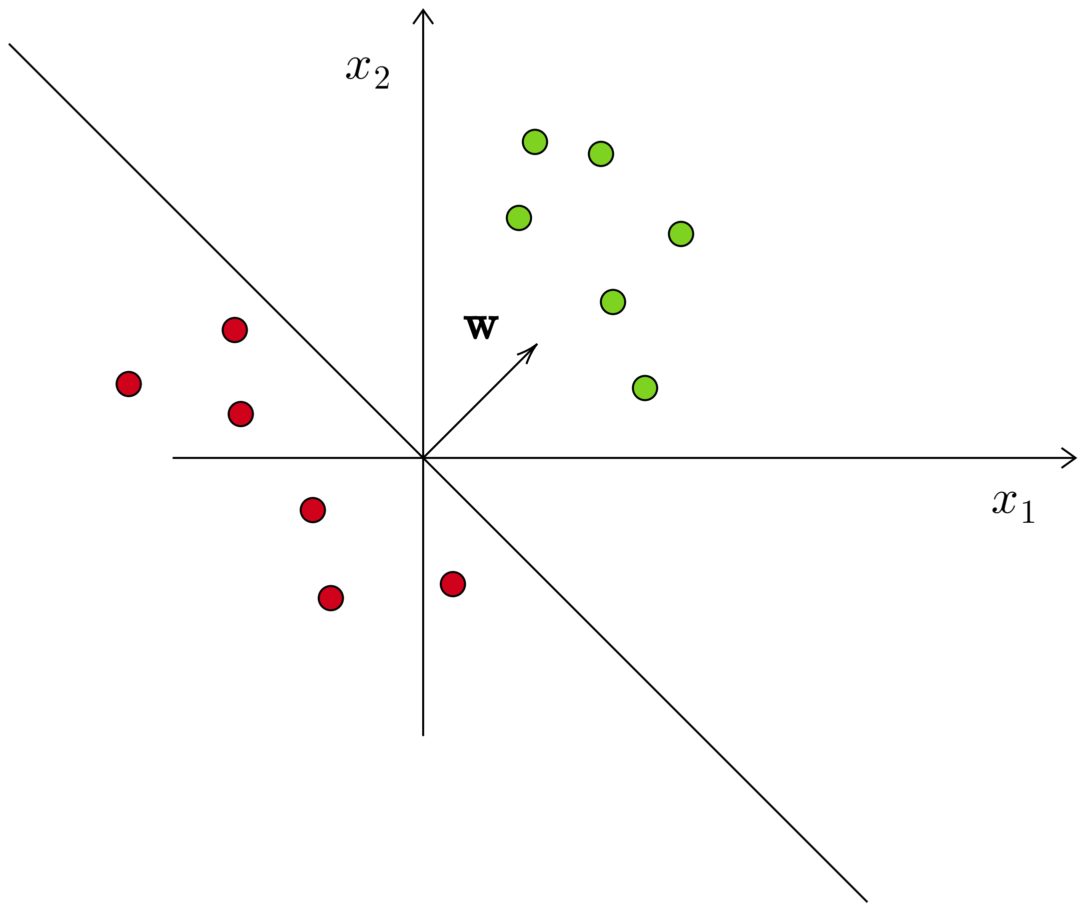
Linear Classifiers

$$\min_{h \in \mathcal{H}_{\mathtt{linear}}} \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{1}[h(\mathbf{x}_i) \neq y_i]$$

Classification as Regression



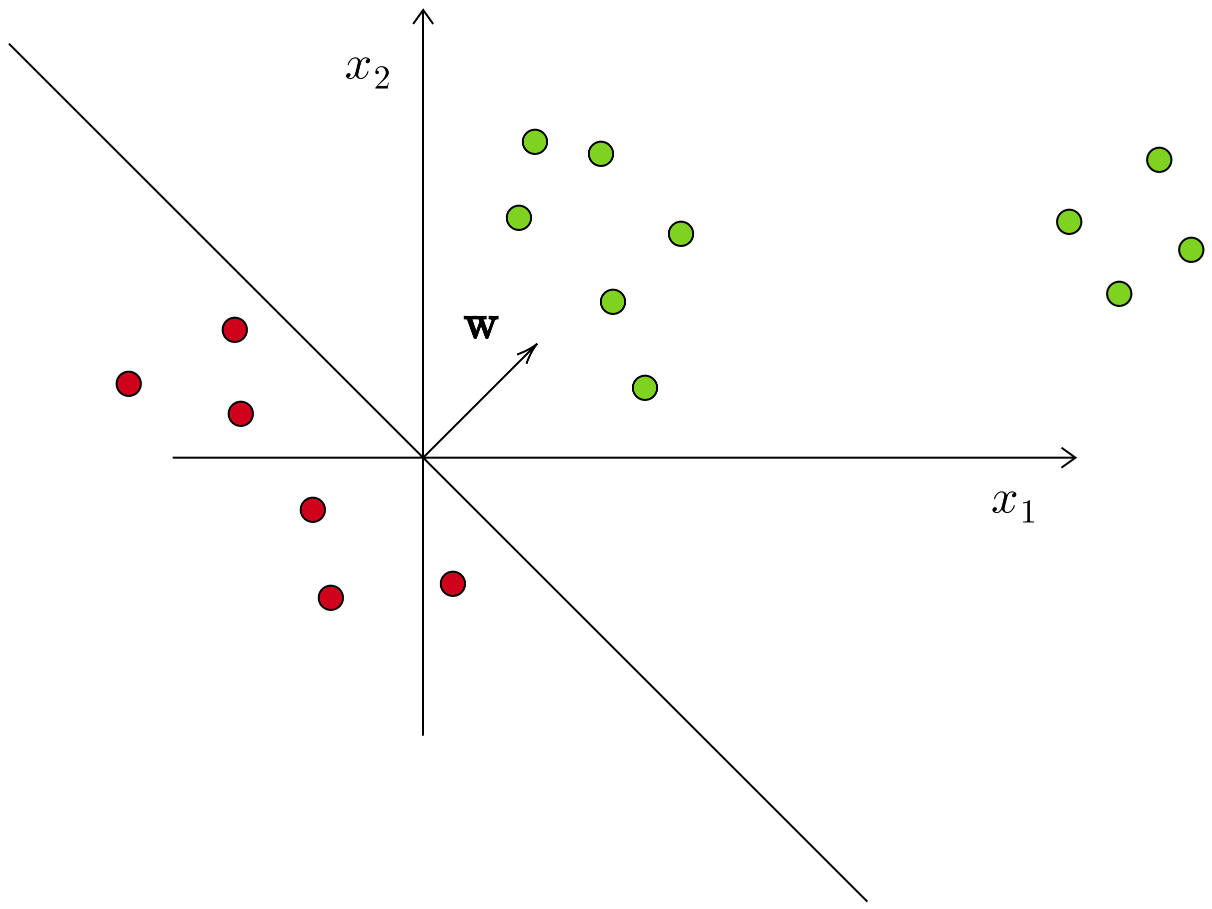
Classification as Regression



$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \ \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{w}^T \mathbf{x}_i - y_i\right)^2$$

$$h_{\mathbf{w}^*}(\mathbf{x}) = \mathtt{sign}\left(\mathbf{w}^{*T} \mathbf{x}\right)$$

Classification as Regression



$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{w}^T \mathbf{x}_i - y_i\right)^2$$

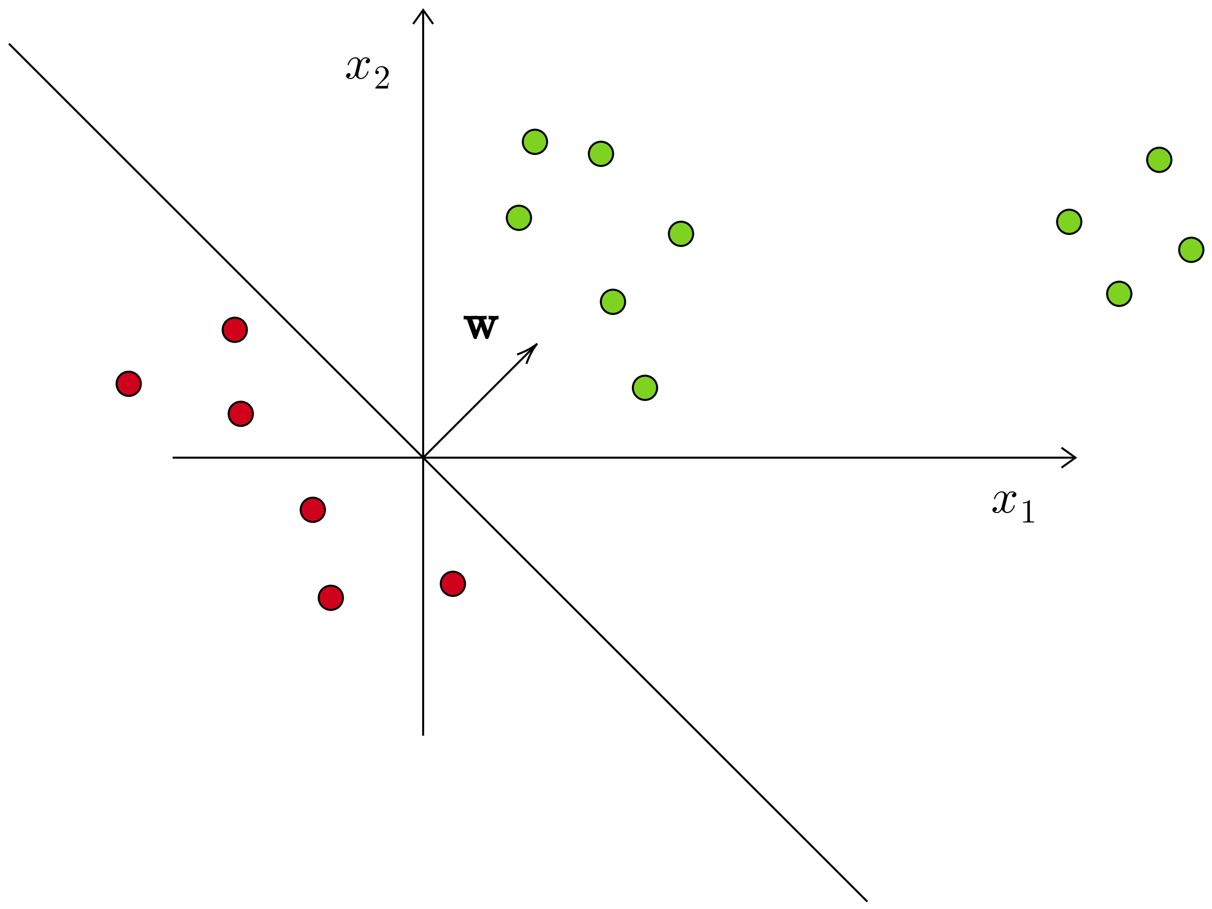
$$h_{\mathbf{w}^*}(\mathbf{x}) = \mathtt{sign}\left(\mathbf{w}^{*T} \mathbf{x}\right)$$

- Sensitive to outliers

Classification as Regression

$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{w}^T \mathbf{x}_i - y_i\right)^2$$

$$h_{\mathbf{w}^*}(\mathbf{x}) = \mathtt{sign}\left({\mathbf{w}^*}^T \mathbf{x}\right)$$



- Sensitive to outliers
- Natural ordering is absent
  - example: spam, work, family