

① One example  $x$

Feature vector:  $\phi(x) = [\phi_1(x), \dots, \phi_k(x)]$

$$x = [x_1, x_2, \dots, x_k]^T$$

Multiple ( $N$ ) examples  $x^{(1)}, x^{(2)}, \dots, x^{(N)}$

$$\hookrightarrow [x_1^{(2)}, x_2^{(2)}, \dots, x_k^{(2)}]$$

Params (weights)

$$w = [w_1, w_2, \dots, w_k]$$

Dot product  $w^T x = \sum_{k=1}^K w_k x_k$   $\langle w, x \rangle$  or  $w \cdot x$  or  $w * x$

$k \times 1$

$k \times 1$

$1 \times k$

$$w^T x \geq \tau ? \Leftrightarrow w^T x - \tau \geq 0 \Leftrightarrow w^T x + (-\tau) \geq 0$$

$\hookrightarrow$  threshold  $\Leftrightarrow w^T x + b \geq 0$  Logistic Regression

$$\Leftrightarrow \underbrace{w_1 x_1 + \dots + w_k x_k}_{w^T x} + \underbrace{1}_{x_0} \cdot \underbrace{b}_{w_0} = \sum_{j=0}^k w_j x_j$$

Redefined  $w = [w_0, w_1, \dots, w_k]$

$$x = [1, x_1, \dots, x_k]$$

$= w^T x \rightarrow$  Perceptron

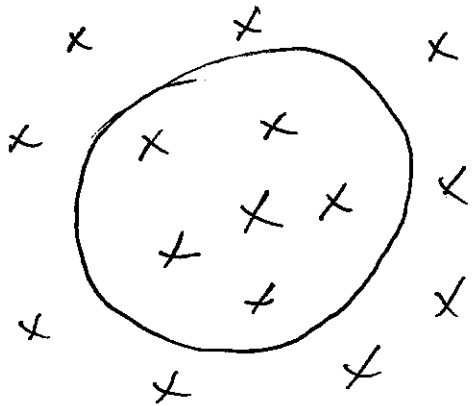
② Decision boundary  $D = \{x \mid w^T x + w_0 = 0\}$

$w = [w_1, w_2]$ ,  $w_0$  are given, fixed, known

$$\begin{aligned} w^T x + w_0 = 0 &\Rightarrow w_1 x_1 + w_2 x_2 + w_0 = 0 \Rightarrow x_2 = -\frac{w_1 x_1 + w_0}{w_2} \\ &= \left(-\frac{w_1}{w_2}\right) x_1 + \left(-\frac{w_0}{w_2}\right) \end{aligned}$$

$$\|w\|_1 = \|w\|_2 = \sqrt{w_1^2 + w_2^2}$$

$\hookrightarrow L_2$  norm  $\equiv$  Euclidean norm  $\equiv$  length.



not linearly separable.