





I N S E A







Régression logistique

$$\min_{\alpha \in \mathbb{R}, \beta \in \mathbb{R}^d} \sum_{i=1}^n \left( \mathbb{1}_{\{\alpha + \beta^\top \mathbf{x}_i \geq 0\}} - y_i \right)^2$$

Fonction différentiable (discontinue) difficile à ptiniser



Au lieu de prendre le signe, transformer les scores  $\alpha + \beta^T \mathbf{x}_i$  vers  $[0, 1]$  et modéliser des probabilités

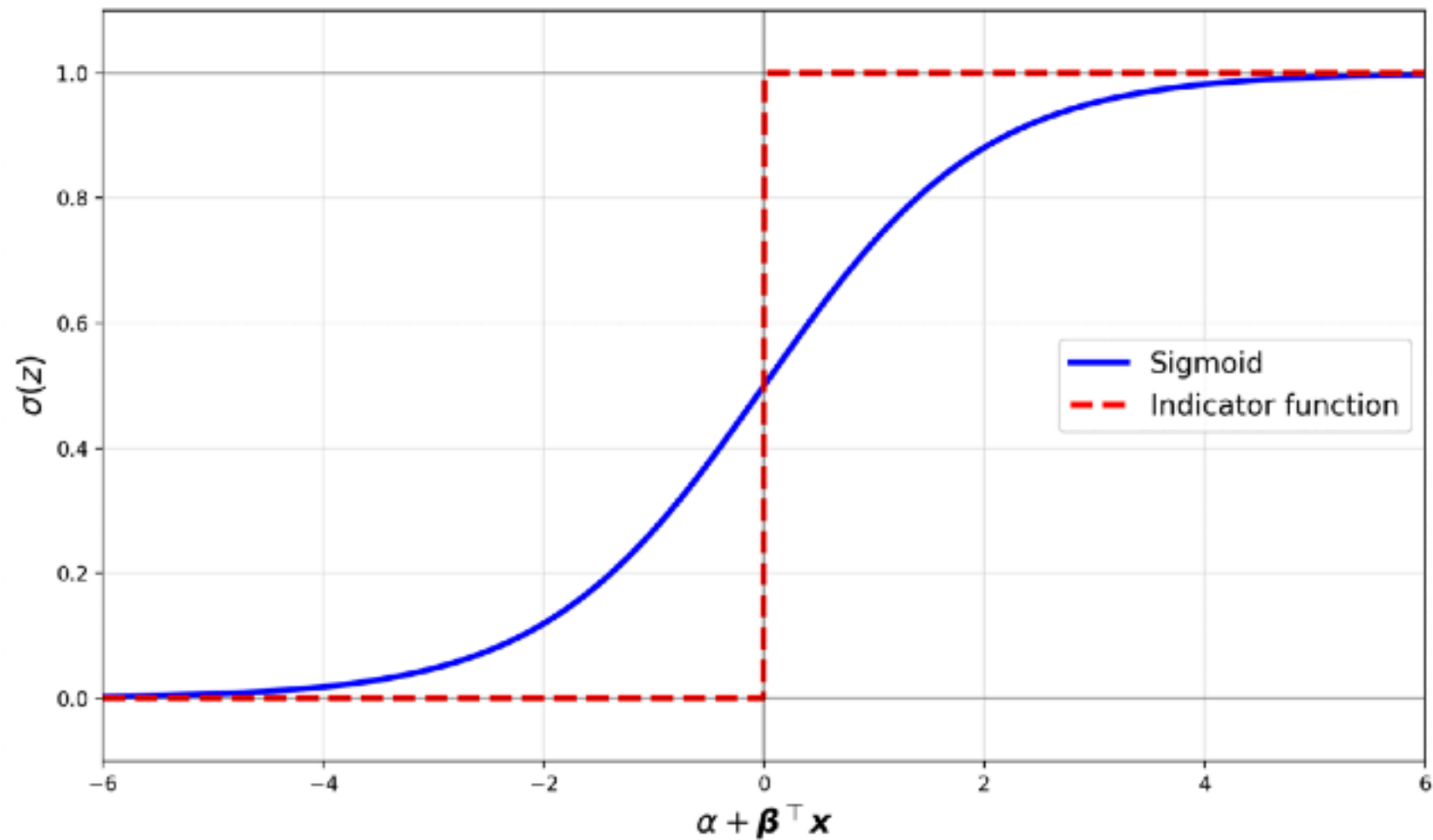
$$p_i \stackrel{\text{def}}{=} P(y_i = 1 | \mathbf{x}_i) = \text{sigmoid}(\alpha + \beta^\top \mathbf{x}_i)$$

On peut comparer les  $p_i$  avec les  $y_i$  avec la *cross-entropy*:

$$\min_{\alpha \in \mathbb{R}, \beta \in \mathbb{R}^d} - \sum_{i=1}^n y_i \log(p_i) + (1 - y_i) \log(1 - p_i)$$

Modèle de régression logistique

sigmoid:  $t \mapsto \frac{1}{1+e^{-t}}$  (logistique)



On a donc une fonction de prédiction:  $f^*(\mathbf{x}_i) = 1 \Leftrightarrow \text{sigmoid}(\alpha^* + \beta^{*\top} \mathbf{x}_i) \geq \frac{1}{2}$

Machine learning classic: zero-to-hero



$$\min_{\alpha \in \mathbb{R}, \beta \in \mathbb{R}^d} \sum_{i=1}^n (\mathbb{1}_{\{\alpha + \beta^\top \mathbf{x}_i \geq 0\}} - y_i)^2 \quad \text{Fonction non différentiable (discontinue même) difficile à optimiser}$$

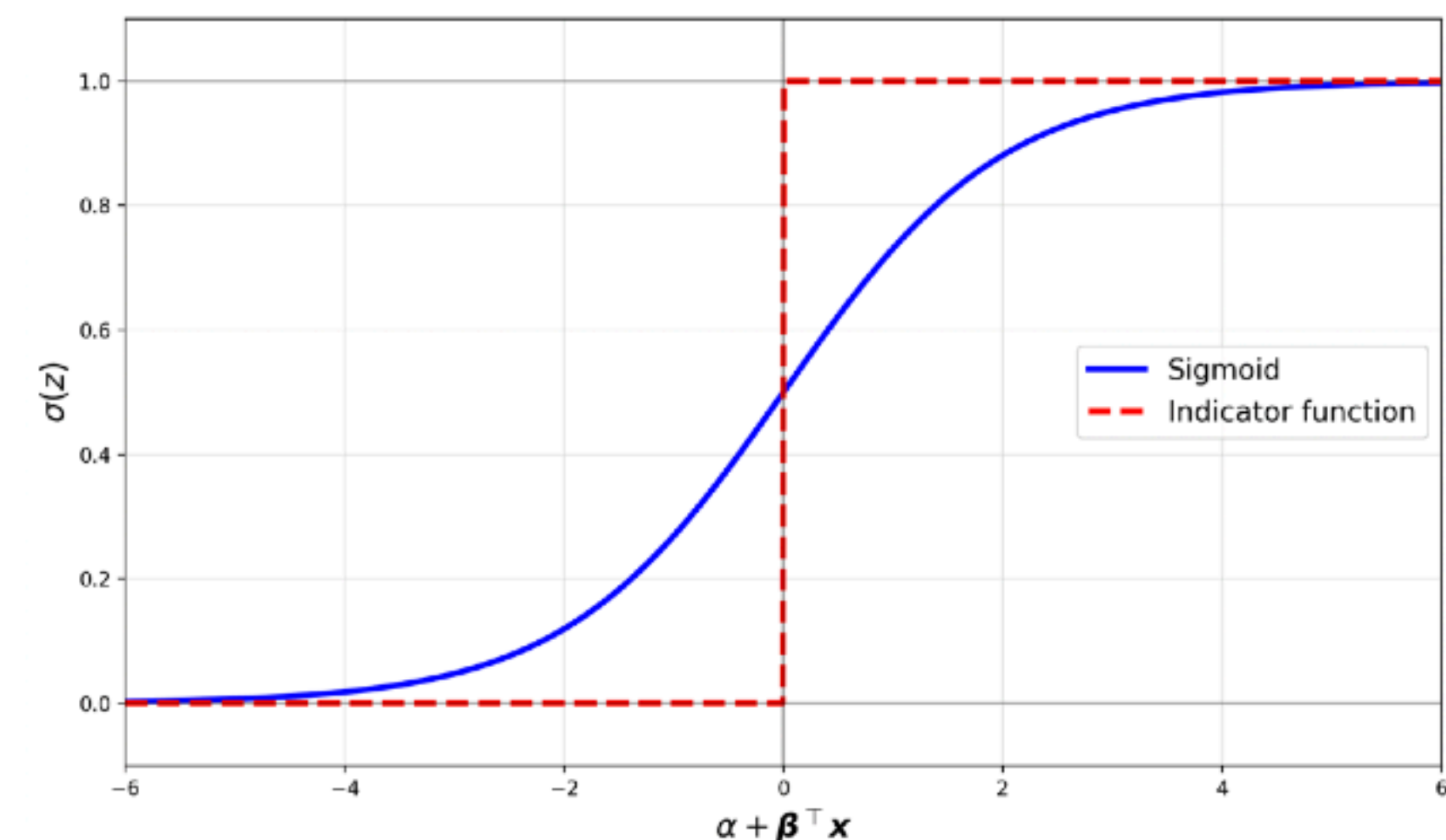
Au lieu de prendre le signe, transformer les scores  $\alpha + \beta^\top \mathbf{x}_i$  vers  $[0, 1]$  et modéliser des probabilités

sigmoid:  $t \mapsto \frac{1}{1+e^{-t}}$  (logistique)

$$p_i \stackrel{\text{def}}{=} \mathbb{P}(y_i = 1 | \mathbf{x}_i) = \text{sigmoid}(\alpha + \beta^\top \mathbf{x}_i)$$

On peut comparer les  $p_i$  avec les  $y_i$  avec la *cross-entropy*:

$$\min_{\alpha \in \mathbb{R}, \beta \in \mathbb{R}^d} - \sum_{i=1}^n y_i \log(p_i) + (1 - y_i) \log(1 - p_i)$$



On a donc une fonction de prédiction:  $f^*(\mathbf{x}_i) = 1 \Leftrightarrow \text{sigmoid}(\alpha^* + \beta^{*\top} \mathbf{x}_i) \geq \frac{1}{2}$

Modèle de régression logistique

$$p_i \stackrel{\text{def}}{=} \mathbb{P}(y_i = 1 | \mathbf{x}_i) = \text{sigmoid}(\alpha + \beta^\top \mathbf{x}_i)$$

$$\min_{\alpha \in \mathbb{R}, \beta \in \mathbb{R}^d} - \sum_{i=1}^n y_i \log(p_i) + (1 - y_i) \log(1 - p_i)$$

