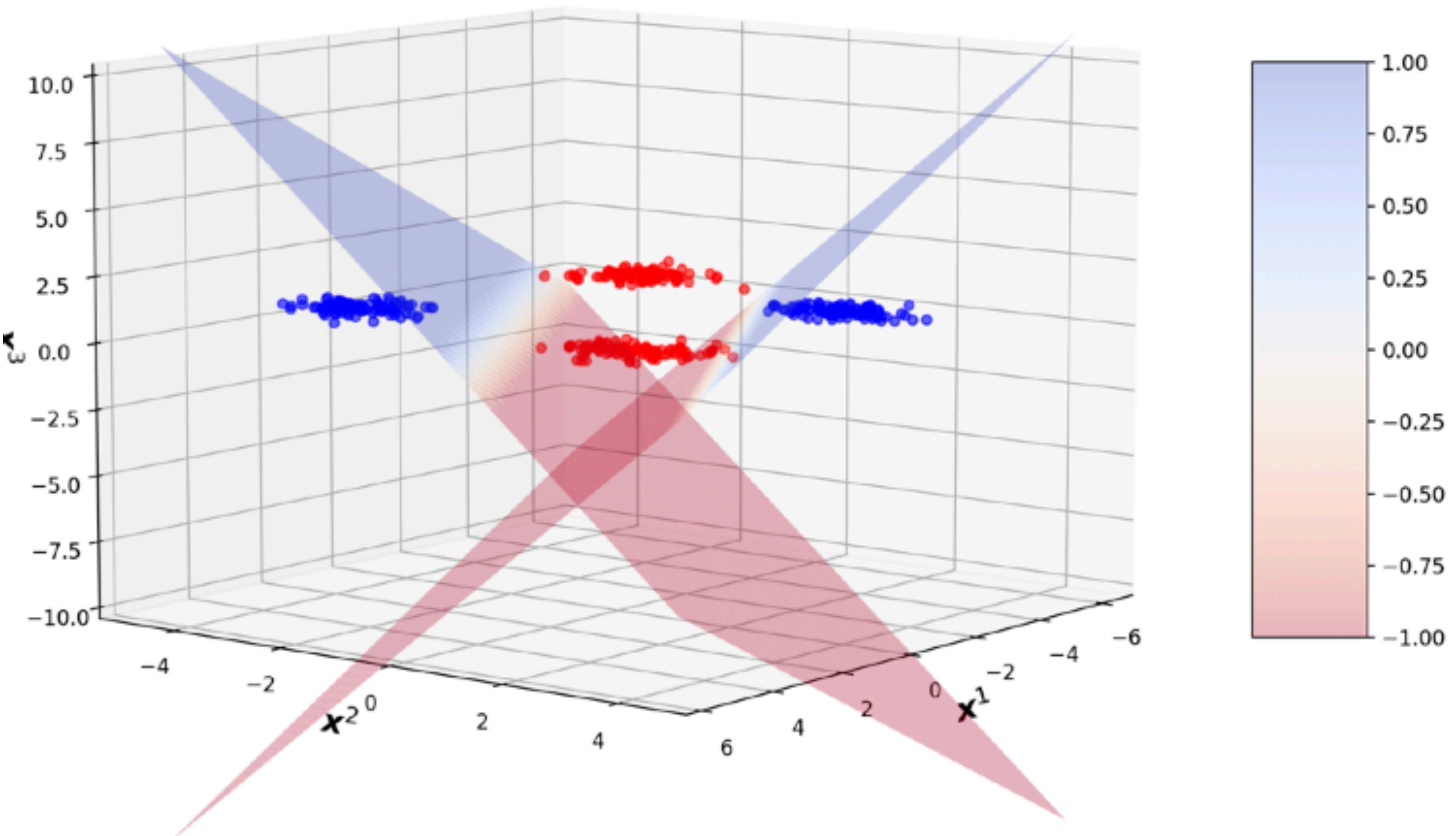
INSEA

30

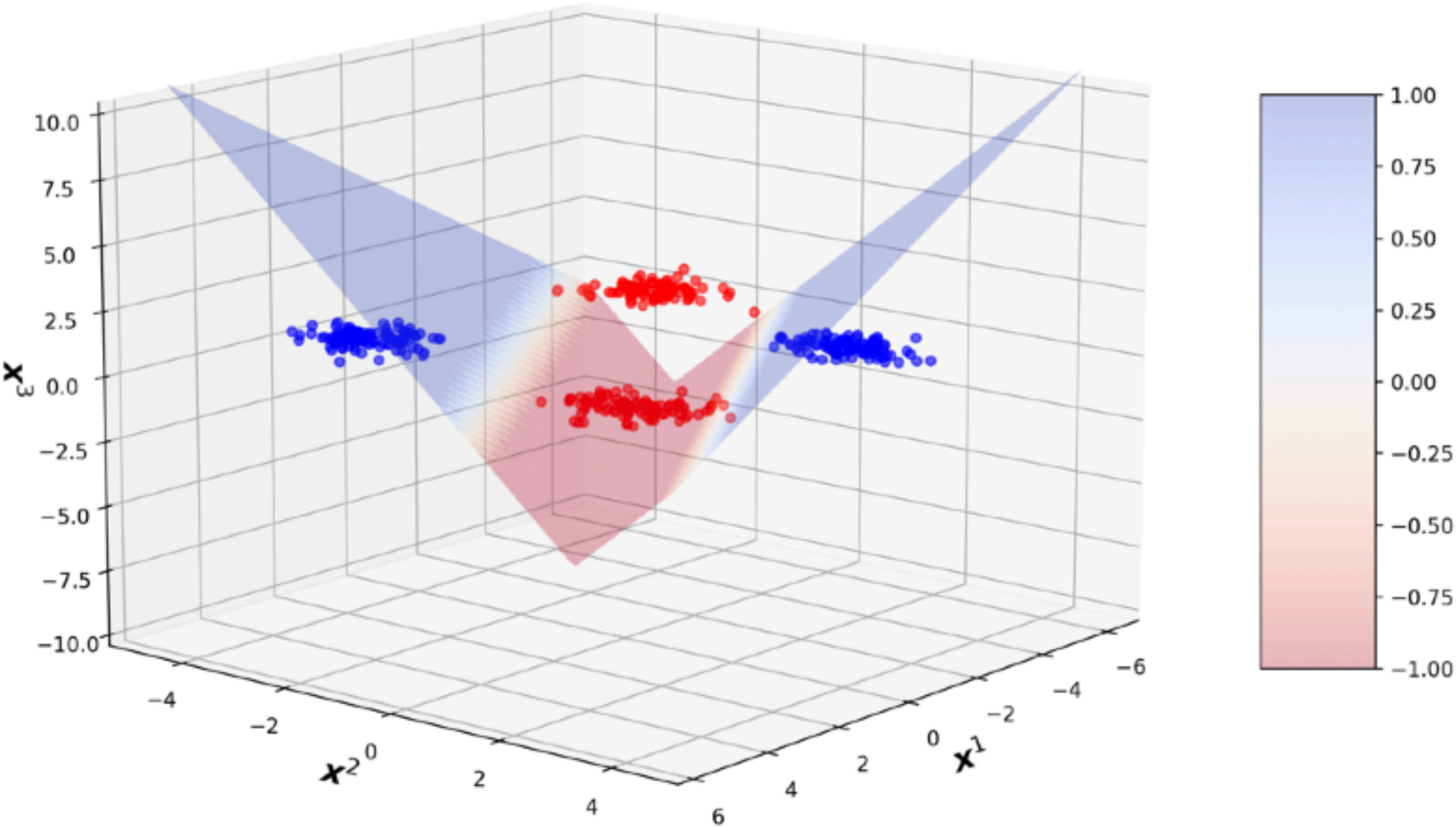
Machine learning classique

Intro to neural nets


10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
−10.0
x3
−4
−2
0
2
4
x2
−6
−4
−2
0
2
4
6
x1
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Surfaces des hyperplans z_1, z_2

Surface de $\max(z_1, z_2)$



Prédire $Y = 1$ si l'un des z_i est positif $\Leftrightarrow \max(z_1, z_2) > 0$

$$f_{\alpha,\beta}(\mathbf{x}) = \mathbb{1}_{\{\mathrm{max}(\mathbf{x}^\top \boldsymbol{\beta}^{\mathbf{1}} + \alpha_1, \mathbf{x}^\top \boldsymbol{\beta}^{\mathbf{2}} + \alpha_2) > 0\}}$$

Comment entraîner ce modèle, c-à-d optimiser α, β ?

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

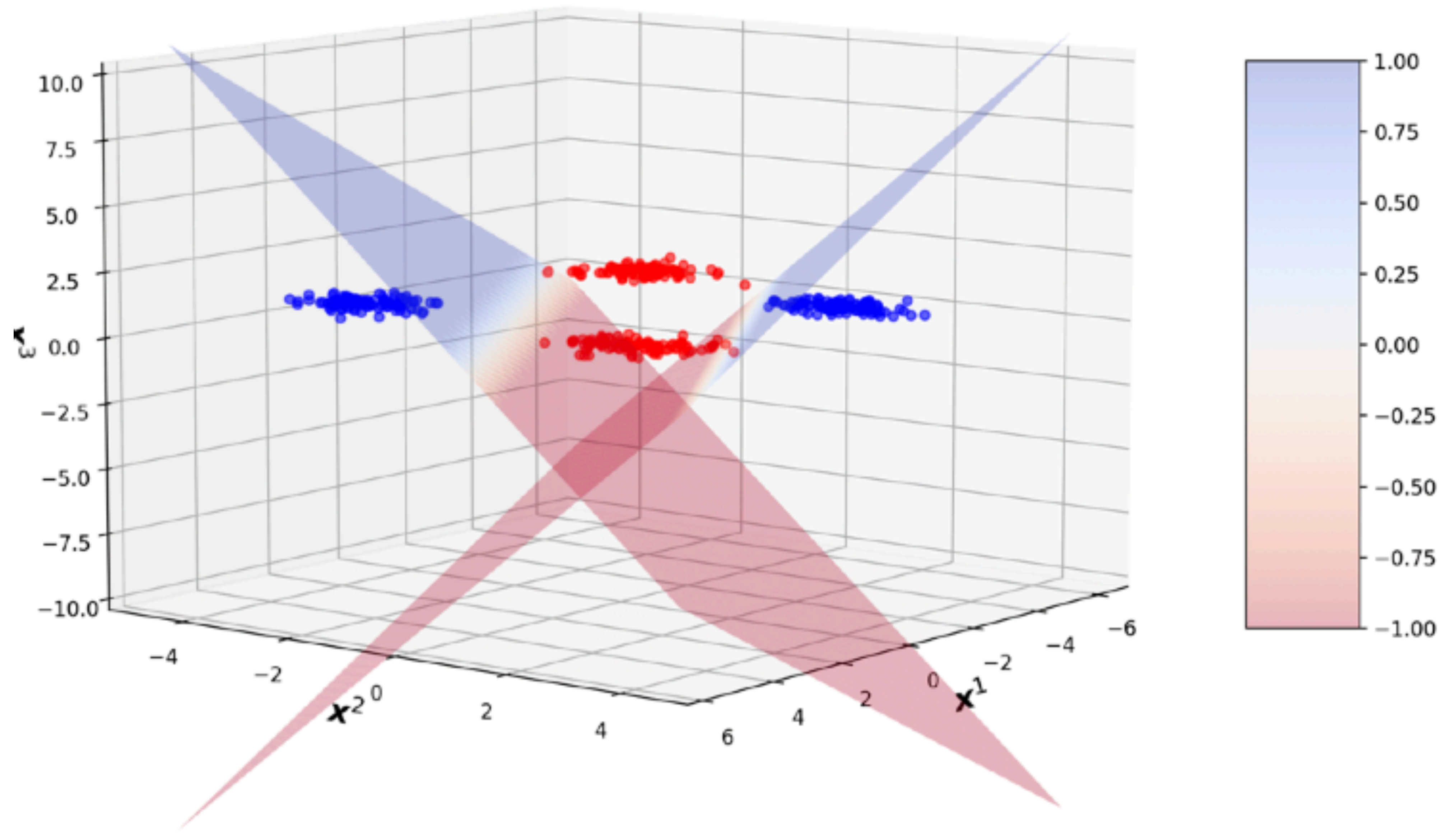
$$p_i \stackrel{\text{def}}{=} \mathbb{P}_{\alpha,\beta}(Y = 1|\mathbf{x}_i) = \text{sigmoid}(\max(\mathbf{x}_i^\top \boldsymbol{\beta}^{\mathbf{1}} + \alpha_1, \mathbf{x}_i^\top \boldsymbol{\beta}^{\mathbf{2}} + \alpha_2))$$

+

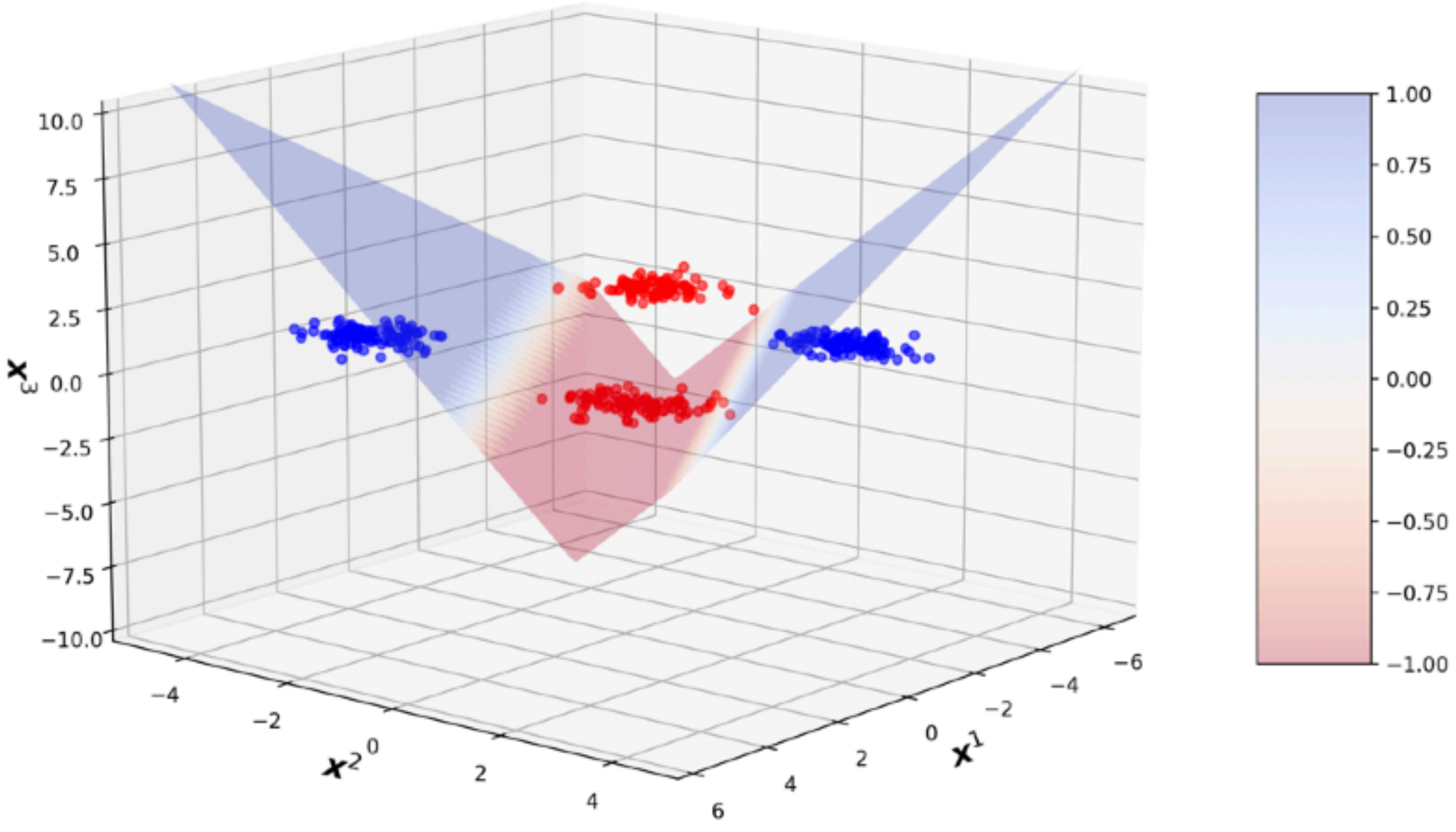
Linear functions

Non-linearity

Surfaces des hyperplans z_1, z_2



Surface de $\max(z_1, z_2)$



Prédire $Y = 1$ si l'un des z_i est positif $\Leftrightarrow \max(z_1, z_2) > 0$

$$f_{\alpha,\beta}(\mathbf{x}) = \mathbb{1}_{\{\max(\mathbf{x}^\top \boldsymbol{\beta}^{\mathbf{1}} + \alpha_1, \mathbf{x}^\top \boldsymbol{\beta}^{\mathbf{2}} + \alpha_2) > 0\}}$$

Linear functions

Comment entraîner ce modèle, c-à-d optimiser α, β ?

+

$$p_i \stackrel{\text{def}}{=} \mathbb{P}_{\alpha,\beta}(Y = 1 | \mathbf{x}_i) = \text{sigmoid}(\max(\mathbf{x}_i^\top \boldsymbol{\beta}^{\mathbf{1}} + \alpha_1, \mathbf{x}_i^\top \boldsymbol{\beta}^{\mathbf{2}} + \alpha_2))$$

Non-linearity

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

Comment adapter ce modèle à des données plus complexes ?