

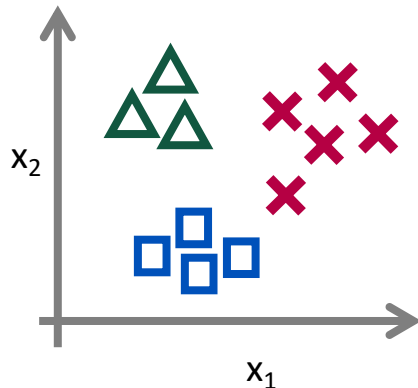
Multi-Class Classification



Three Classes as Examples

Multi-Class Classification: Part I

One-versus-all
(one-versus-rest)



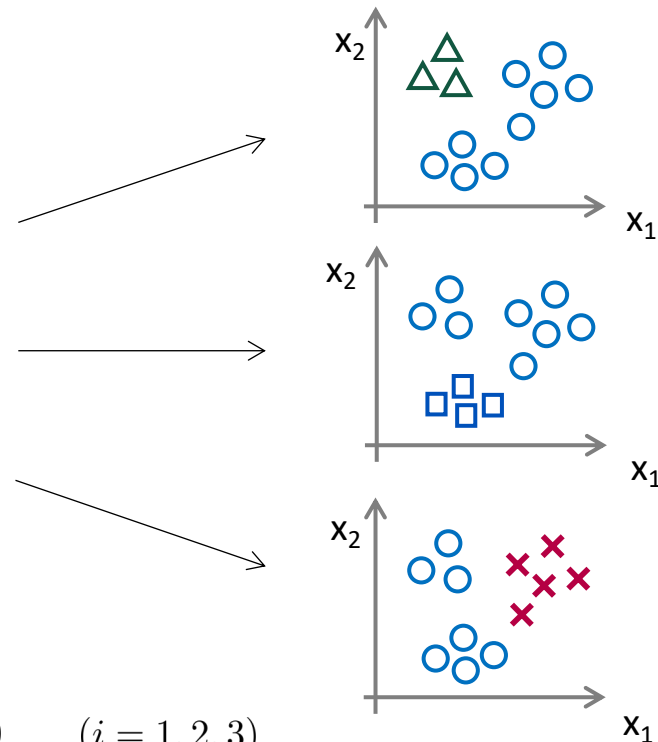
Class 1: 

Class 2: 

Class 3: 

$$h_{\theta}^{(i)}(x) = P(y = i|x; \theta) \quad (i = 1, 2, 3)$$

$$\max_i h_{\theta}^{(i)}(x)$$



Three Classes as Examples

Multi-Class Classification: Part II

$$C_1: w^1, b_1 \quad z_1 = w^1 \cdot x + b_1$$

$$C_2: w^2, b_2 \quad z_2 = w^2 \cdot x + b_2$$

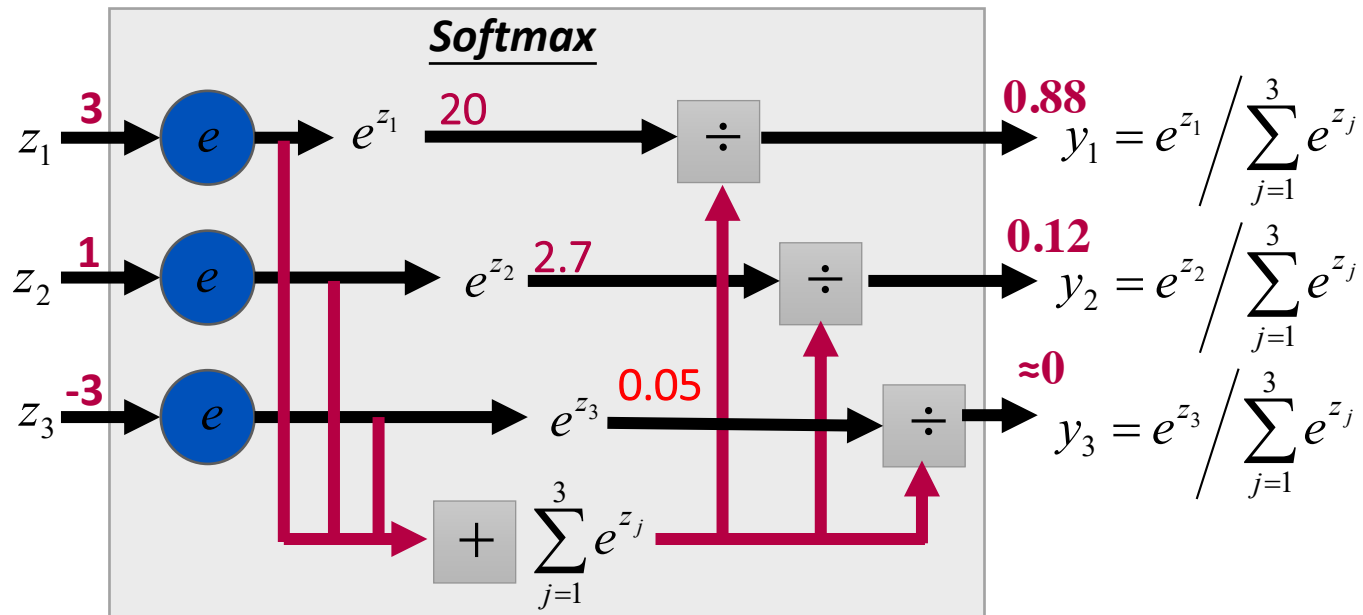
$$C_3: w^3, b_3 \quad z_3 = w^3 \cdot x + b_3$$

Probability:

$$\blacksquare 1 > y_i > 0$$

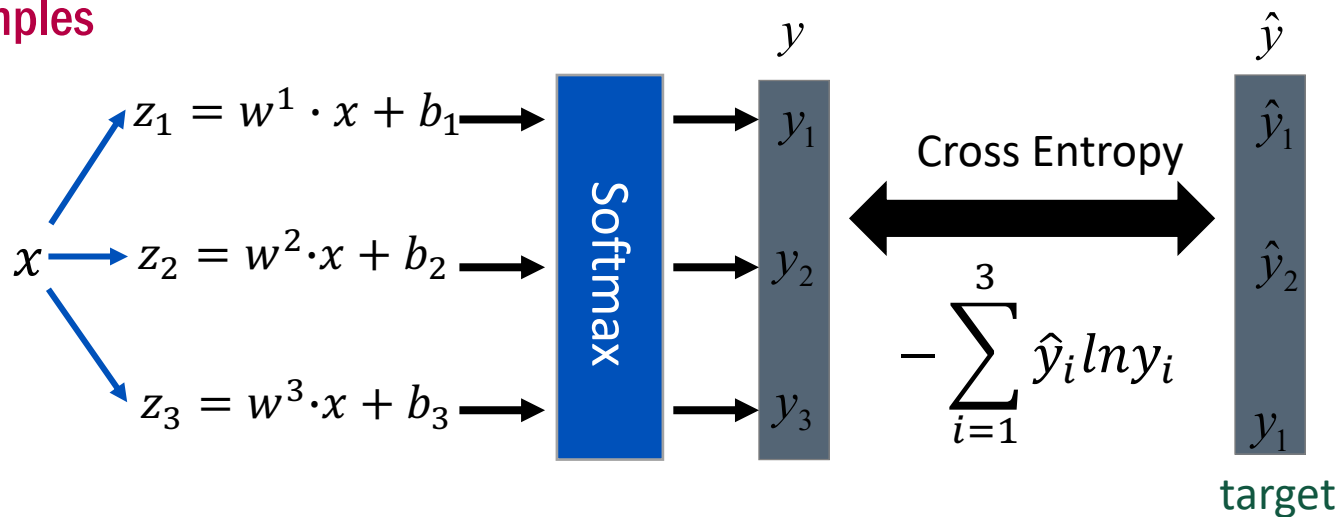
$$\blacksquare \sum_i y_i = 1$$

$$y_i = P(C_i | x)$$



Three Classes as Examples

Multi-Class Classification: Part III



If $x \in \text{class 1}$

$$\hat{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$-\ln y_1$$

If $x \in \text{class 2}$

$$\hat{y} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$-\ln y_2$$

If $x \in \text{class 3}$

$$\hat{y} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$-\ln y_3$$

Bishop, C.M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer.

