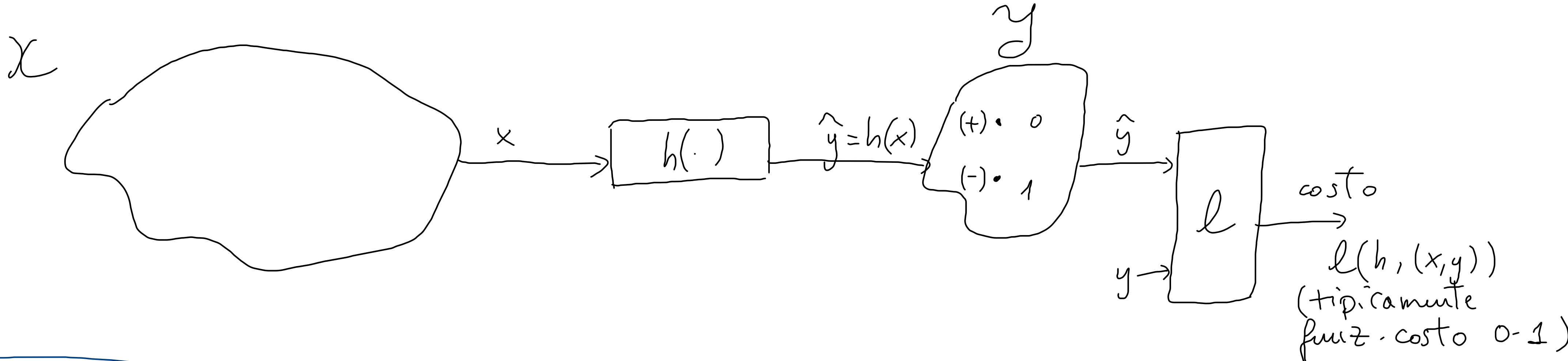
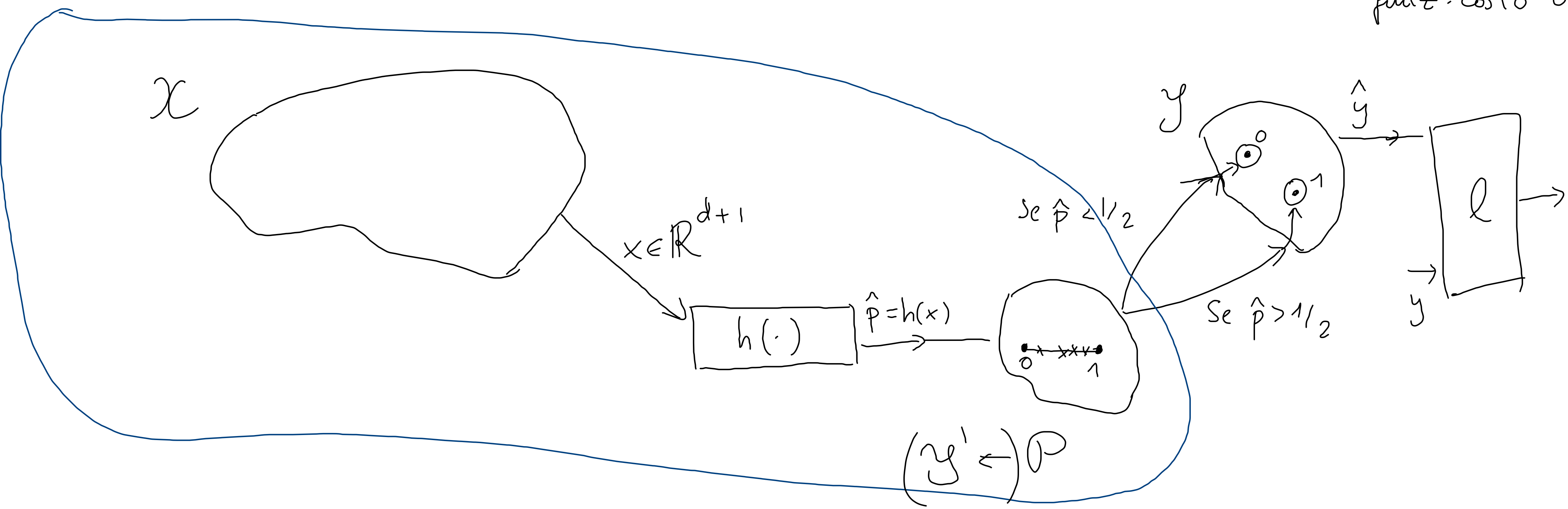

𝒳
x
h(·)
ŷ = h(x)
𝒴
(+) 0
(−) 1
ŷ
ℓ
y
costo
ℓ(h, (x,y))
(tipicamente funz. costo 0-1)


𝒳
x ∈ ℝ^{d+1}
h(·)
p̂ = h(x)
0
1
Se p̂ < 1/2
Se p̂ > 1/2
𝒴
0
1
ŷ
y
ℓ
(𝒴' ←) P

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

$$\boxed{1-\sigma(z)} = 1 - \frac{1}{1+e^{-z}} = \frac{\cancel{1}+e^{-z} - \cancel{1}}{1+e^{-z}}$$

$$= \boxed{\frac{e^{-z}}{1+e^{-z}}} \cdot \frac{e^{z}}{e^{z}} = \frac{e^{0}}{e^{z}+1}$$

$$= \frac{1}{1+e^{z}} = \boxed{\sigma(-z)}$$

$$\sigma'(z) = \frac{d}{dz}\frac{1}{1+e^{-z}} = -\frac{1}{(1+e^{-z})^2} \cdot \frac{d}{dz}(1+e^{-z}) = +\frac{e^{-z}\cdot(+1)}{(1+e^{-z})^2}$$

x

$$\frac{d}{dx}\frac{1}{x} = -\frac{1}{x^2}$$

$$= \frac{1}{(1+e^{-z})} \cdot \boxed{\frac{e^{-z}}{(1+e^{-z})}} = \sigma(z) \cdot (1-\sigma(z))$$

$\sigma(z)$

$$\arg\max_{w \in \mathbb{R}^{d+1}} \log \mathcal{L}(w) = \arg\max_w \sum_{i=1}^m \log \underbrace{\Pr(y^{(i)} | x^{(i)}; w)}_{= \begin{cases} h(x^{(i)}) & \text{se } y^{(i)} = 1 \\ 1 - h(x^{(i)}) & \text{se } y^{(i)} = 0 \end{cases} = h(x^{(i)})^{y^{(i)}} \cdot (1 - h(x^{(i)}))^{1 - y^{(i)}}}$$

$$= \arg\max_w \sum_{i=1}^m \left[y^{(i)} \log h(x^{(i)}) + (1 - y^{(i)}) \log(1 - h(x^{(i)})) \right]$$

$$= \arg\min_w \underbrace{\sum_{i=1}^m \left[-y^{(i)} \log h(x^{(i)}) - (\underbrace{1 - y^{(i)}}_{0}) \log(1 - h(x^{(i)})) \right]}_{\text{cross-entropie (sul training set)}}$$

Se $y = 1$ → 1

Se $y = 0$

$$h(x) = \Pr(y = 1 | x)$$

$$1 - h(x) = \Pr(y = 0 | x)$$

$$\ell(h,(x,y)) = \underbrace{-y\cdot\log h(x) - (1-y)\cdot\log(1-h(x))} \qquad \text{dove } h(x) = \sigma(w^T x)$$

$$\nabla \ell(h,(x,y))$$

$$\frac{\partial}{\partial w_j}\ell(h_w) \ ?$$

$$\frac{\partial}{\partial w_j}\left[\ \cdots\cdots\ \right] = -y\,\frac{\partial}{\partial w_j}\log h(x) - (1-y)\frac{\partial}{\partial w_j}\log(1-h(x))$$

$$= -y\cdot\frac{1}{h(x)}\cdot\frac{\partial}{\partial w_j}h(x) - (1-y)\,\frac{1}{1-h(x)}\,\boxed{\frac{\partial}{\partial w_j}\left[1-h(x)\right]} = -\frac{\partial}{\partial w_j}h(x)$$

$$= -\frac{y}{h(x)}\,\sigma(w^T x)(1-\sigma(w^T x))\cdot\frac{\partial}{\partial w_j}(w^T x) + \frac{1-y}{1-h(x)}\cdot\sigma(w^T x)(1-\sigma(w^T x))\,\frac{\partial}{\partial w_j}(w^T x)$$

$$= -\frac{y}{h(x)}\,\sigma(w^T x)(1-\sigma(w^T x))\cdot \underbrace{\frac{\partial}{\partial w_j}(w^T x)}_{= x_j} + \frac{1-y}{1-h(x)}\cdot \sigma(w^T x)(1-\sigma(w^T x))\underbrace{\frac{\partial}{\partial w_j}(w^T x)}_{= x_j}$$

$$= \sigma(w^T x)(1-\sigma(w^T x))\, x_j \left[-\frac{y}{\sigma(w^T x)} + \frac{1-y}{1-\sigma(w^T x)} \right]$$

$$= x_j \underbrace{\left[-y(1-\sigma(w^T x)) + (1-y)\,\sigma(w^T x) \right]}_{+\cancel{y\sigma(w^T x)} - y + \sigma(w^T x) - \cancel{y\sigma(w^T x)}}$$

$$= \left[\underbrace{\sigma(w^T x)}_{h(x)} - y \right]\cdot x_j \qquad = (h(x) - y)\cdot x_j$$