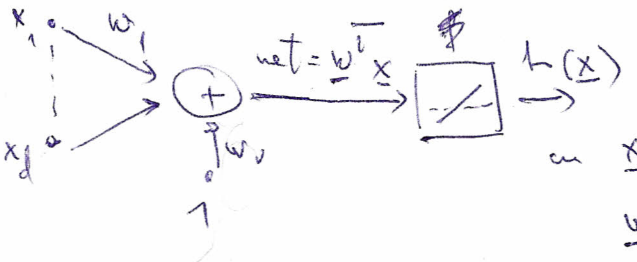
Fie perceptronul A:



cu $\underline{x} = (1, x_1, \ldots, x_d)^T$

$\underline{w} = (w_0, w_1, \ldots, w_d)^T$

a) Scrieți formula lui h [Precizați dom. și codom. lui h

$h: \{1\}\times\mathbb{R}^d \to \mathbb{R}$, $h(\underline{x}) = \underline{w}^T\underline{x} = \sum_{i=1}^{d} w_i x_i + w_0$

b) Precizați sp. de funcții $\mathcal{H}$ implementată de A

$\mathcal{H} = \{ h: \{1\}\times\mathbb{R}^d \to \mathbb{R} \mid h(\underline{x}) = w_0 + \sum_{i=1}^{d} w_i x_i$

c) Fie $S = \{(\underline{x}^i, y^i)\}_{i=1}^{n}$ o mulțime de antrenare. Deduceți formula lui w^* ce minimizează riscul empiric al lui A pe mulțimea S ~~dacă~~ cu funcția ~~de pierdere (loss~~

$$J(\underline{w}) = \sum_{i=1}^{n} (y^i - h(\underline{x}^i))^2 + \lambda \sum_{i=0}^{d} w_i^2, \quad \lambda \geq 0 \text{ dat}$$

$$J(\underline{w}) = (\underline{y} - X\underline{w})^T(\underline{y} - X\underline{w}) + \lambda\, \underline{w}^T\underline{w} \quad \text{cu } y = (y_1, \ldots, y_n)^T$$

$$X = \underset{n \times d}{\begin{bmatrix} x^{1,T} \\ \vdots \\ x^{n,T} \end{bmatrix}}$$

$$\frac{\partial J}{\partial w} = 0 \Rightarrow -2X^T(y - Xw) + 2\lambda w = 0$$

$$(X^TX + \lambda I)w = X^Ty$$

$$\Rightarrow w^* = (X^TX + \lambda I)^{-1}X^Ty$$

evaluat hessianul și arătat că-i poz. def

d) Dacă $\{y^i\}_{i=1}^{n}$ sunt i.i.d. și $D^2(y^i) = \sigma^2$ calculați $D^2(w^*)$

Aceeași probl. pt. $\lambda = 0$.

$$D^2[w^*] = [(X^TX + \lambda I)^{-1}X^T]\, D^2(y)\, [(X^TX + \lambda I)^{-1}X^T]^T$$

$$= \sigma^2 (X^TX + \lambda I)^{-1}(X^TX)(X^TX + \lambda I)^{-1}$$

dacă $\lambda = 0 \Rightarrow D^2[w^*] = \sigma^2 (X^TX)^{-1}$

dacă inv.

$w^* = (X^TX)^{-1}X^Ty$