CLASIFICACIÓN DE OBJETOS

Regresión Logística – Aprendizaje (II)

Antonio M. López

DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

Universitat Autònoma de Barcelona

$\boldsymbol{w}$ se aprende a partir de un conjunto de muestras $\mathcal{S} = \{(\boldsymbol{x}^1, y^1), \ldots, (\boldsymbol{x}^M, y^M)\}$

$(\boldsymbol{x}^j, y^j)$ j-ésima muestra de entrenamiento

$\boldsymbol{x}^j \in \mathbb{R}^{n+1}$, descriptor de la ventana j

$y^j \in \{0, 1\}$, etiqueta de la ventana j

$h_{\boldsymbol{w}}(\boldsymbol{x}^j) = Logistic(\boldsymbol{w}^\top \boldsymbol{x}^j)$

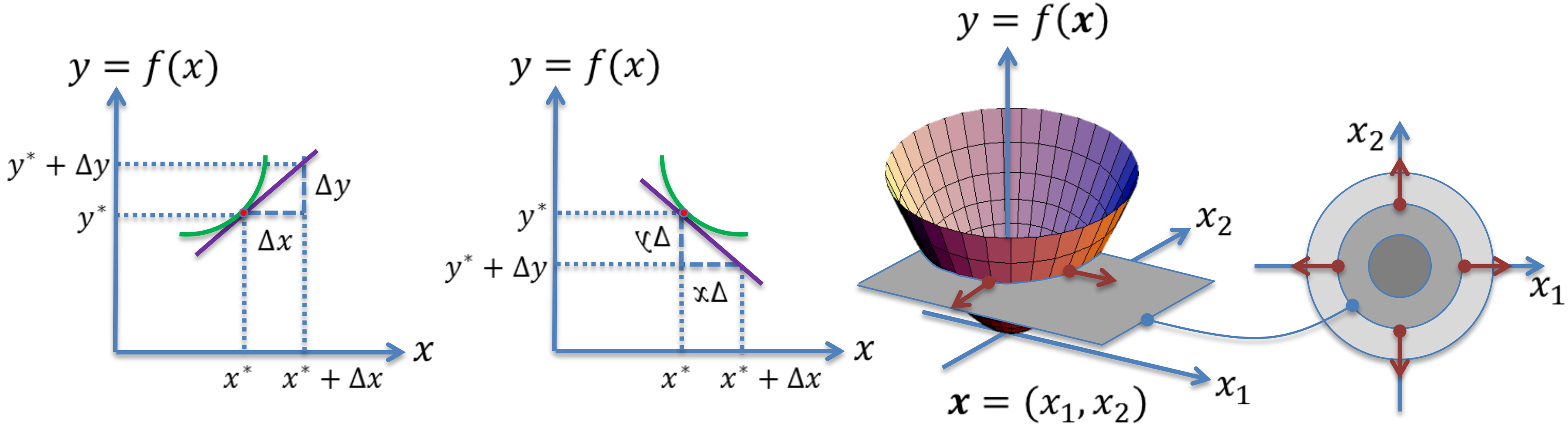
$$\mathcal{J}(\boldsymbol{w}) = -\frac{1}{M}\sum_{j=1}^{M} y^j \, ln\left(h_{\boldsymbol{w}}(\boldsymbol{x}^j)\right) + (1 - y^j)\, ln\left(1 - h_{\boldsymbol{w}}(\boldsymbol{x}^j)\right)$$

$$\boldsymbol{w}^* \longleftarrow \underset{\boldsymbol{w} \in \mathbb{R}^{n+1}}{argmin}\, \mathcal{J}(\boldsymbol{w})$$

$\mathcal{J}(\boldsymbol{w})$ es convexa respecto $\boldsymbol{w}$

Regresión Logística

- Al ser $\mathcal{J}(\boldsymbol{w})$ convexa se puede utilizar algún algoritmo tipo "descenso del gradiente ".

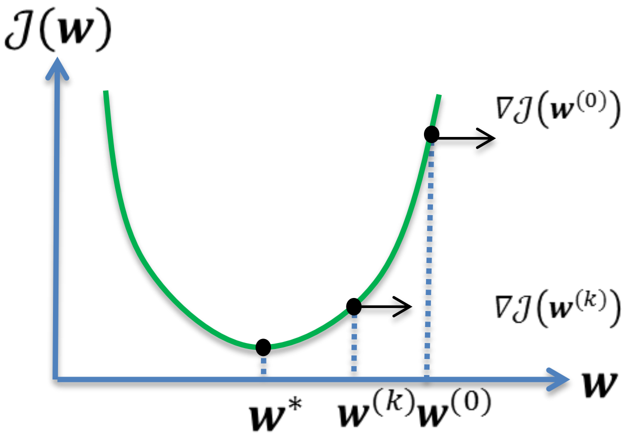


$$\frac{df(x)}{dx} \sim \frac{\Delta y}{\Delta x}$$

$$\nabla f(\boldsymbol{x}) = \left(\frac{\partial f(\boldsymbol{x})}{\partial x_1}, \frac{\partial f(\boldsymbol{x})}{\partial x_2}\right)^{\mathsf{T}}$$

UAB Universitat Autònoma de Barcelona

- Descenso del gradiente:



$$\nabla\mathcal{J}(\boldsymbol{w}) = \left(\frac{\partial\mathcal{J}(\boldsymbol{w})}{\partial w_0}, \dots, \frac{\partial\mathcal{J}(\boldsymbol{w})}{\partial w_n}\right)^{\mathrm{T}}$$

<u>Repetir</u> (en paralelo)

$$w_0{}^{(k)} \longleftarrow w_0{}^{(k-1)} - \frac{\partial}{\partial w_0}\mathcal{J}(\boldsymbol{w}^{(k-1)})$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$w_n{}^{(k)} \longleftarrow w_n{}^{(k-1)} - \frac{\partial}{\partial w_n}\mathcal{J}(\boldsymbol{w}^{(k-1)})$$

<u>Hasta</u> converger /* p.e. monitorizar $\mathcal{J}(\boldsymbol{w})$ */

UAB
Universitat Autònoma de Barcelona

- Descenso del gradiente:

<u>Repetir</u> (en paralelo)

$$w_0^{(k)} \longleftarrow w_0^{(k-1)} - \alpha \frac{\partial}{\partial w_0} \mathcal{J}(\boldsymbol{w}^{(k-1)})$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$w_n^{(k)} \longleftarrow w_n^{(k-1)} - \alpha \frac{\partial}{\partial w_n} \mathcal{J}(\boldsymbol{w}^{(k-1)})$$

<u>Hasta</u> converger /* p.e. monitorizar $\mathcal{J}(\boldsymbol{w})$ */

$\alpha \in \mathbb{R}^+$: velocidad (ratio) de aprendizaje

$$\frac{\partial \mathcal{J}(\boldsymbol{w})}{\partial w_i} = \frac{\partial}{\partial w_i}\left[-\frac{1}{M}\sum_{j=1}^{M} y^j \, ln\left(h_{\boldsymbol{w}}(\boldsymbol{x}^j)\right) \; + \; (1-y^j)\, ln\left(1-h_{\boldsymbol{w}}(\boldsymbol{x}^j)\right)\right]$$

$$= \frac{1}{M}\sum_{j=1}^{M}\left(h_{\boldsymbol{w}}(\boldsymbol{x}^j) - y^j\right) {x_i}^j$$

<u>Repetir</u> (en paralelo)

$$w_0^{(k)} \leftarrow w_0^{(k-1)} - \alpha \frac{1}{M}\sum_{j=1}^{M}\left(h_{\boldsymbol{w}}(\boldsymbol{x}^j) - y^j\right){x_0}^j$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$w_n^{(k)} \leftarrow w_n^{(k-1)} - \alpha \frac{1}{M}\sum_{j=1}^{M}\left(h_{\boldsymbol{w}}(\boldsymbol{x}^j) - y^j\right){x_n}^j$$

<u>Hasta</u> converger /* p.e. monitorizar $\mathcal{J}(\boldsymbol{w})$ */

- Conceptos clave de este vídeo:
    - Descenso del gradiente
    - Regresión logística por descenso del gradiente