



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Quadern de treball: Regressió logística

Albert Sanchis

DSIC

Departament de Sistemes
Informàtics i Computació

Objectius formatius

- Calcular vectors de logits
- Aplicar la funció softmax als vectors de logits
- Aplicar la regla de decisió d'un classificador basat en la funció softmax
- Calcular el gradient de la funció NLL
- Aplicar descens per gradient per a actualitzar la matriu de pesos

- **Qüestió 1:** La següent taula presenta un conjunt de 2 mostres d'entrenament ($n = \{1, 2\}$) de 2 dimensions (x_{n1}, x_{n2}) procedents de 2 classes ($c_n = \{1, 2\}$):

n	x_{n1}	x_{n2}	c_n
1	1	1	2
2	0	1	1

Addicionalment, la següent taula representa una matriu de pesos inicials \mathbf{W} amb els pesos de cadascuna de les classes per columnes (en notació homogènia):

w_1	w_2
0	0
-0,25	0,25
0	0

Es demana:

1. Calcula el vector de logits associat a cada mostra d'entrenament.

$$\mathbf{a} = f(\mathbf{x}; \mathbf{W}) = \mathbf{W}^t \mathbf{x}$$

$$\mathbf{a}_1 = \mathbf{W}^t \mathbf{x}_1 = \begin{pmatrix} 0 & -0,25 & 0 \\ 0 & 0,25 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -0,25 \\ 0,25 \end{pmatrix}$$

$$\mathbf{a}_2 = \mathbf{W}^t \mathbf{x}_2 = \begin{pmatrix} 0 & -0,25 & 0 \\ 0 & 0,25 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

2. Aplica la funció softmax al vector de logits de cada mostra d'entrenament.

$$\mu_1 = \mathcal{S}(\mathbf{a}_1) = \begin{pmatrix} \frac{e^{-0,25}}{e^{-0,25} + e^{0,25}} \\ \frac{e^{0,25}}{e^{-0,25} + e^{0,25}} \end{pmatrix} = \begin{pmatrix} 0,3775 \\ 0,6225 \end{pmatrix}$$

$$\mu_2 = \mathcal{S}(\mathbf{a}_2) = \begin{pmatrix} \frac{e^0}{e^0 + e^0} \\ \frac{e^0}{e^0 + e^0} \end{pmatrix} = \begin{pmatrix} 0,5 \\ 0,5 \end{pmatrix}$$

3. Classifica cadascuna de les mostres d'entrenament. En cas d'empat, tria qualsevol classe.

$$c(\mathbf{x}) = \arg \max_c \mathcal{S}(\mathbf{a})_c \quad \text{on} \quad \mathbf{a} = \mathbf{W}^t \mathbf{x}$$

$$c(\mathbf{x}_1) = \arg \max_c \mathcal{S}(\mathbf{a}_1)_c = 2$$

$$c(\mathbf{x}_2) = \arg \max_c \mathcal{S}(\mathbf{a}_2)_c = 1 \quad (\text{o } 2)$$

4. Calcula el gradient de la funció NLL en el punt de la matriu de pesos inicials.

$$\begin{pmatrix} \frac{\partial \text{NLL}}{\partial W_{11}} & \cdots & \frac{\partial \text{NLL}}{\partial W_{1C}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \text{NLL}}{\partial W_{D1}} & \cdots & \frac{\partial \text{NLL}}{\partial W_{DC}} \end{pmatrix} = \frac{\partial \text{NLL}}{\partial \mathbf{W}^t} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n (\boldsymbol{\mu}_n - \mathbf{y}_n)^t$$

$$\mathbf{x}_1 (\boldsymbol{\mu}_1 - \mathbf{y}_1)^t = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} ((0,3775 \ 0,6225) - (0 \ 1))$$

$$= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} (0,3775 \ -0,3775)$$

$$= \begin{pmatrix} 0,3775 & -0,3775 \\ 0,3775 & -0,3775 \\ 0,3775 & -0,3775 \end{pmatrix}$$

$$\begin{aligned}
 \mathbf{x}_2(\boldsymbol{\mu}_2 - \mathbf{y}_2)^t &= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} ((0,5 \ 0,5) - (1 \ 0)) \\
 &= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} (-0,5 \ 0,5) \\
 &= \begin{pmatrix} -0,5 & 0,5 \\ 0 & 0 \\ -0,5 & 0,5 \end{pmatrix}
 \end{aligned}$$

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n (\boldsymbol{\mu}_n - \mathbf{y}_n)^t &= \frac{1}{2} \left(\begin{pmatrix} 0,3775 & -0,3775 \\ 0,3775 & -0,3775 \\ 0,3775 & -0,3775 \end{pmatrix} + \begin{pmatrix} -0,5 & 0,5 \\ 0 & 0 \\ -0,5 & 0,5 \end{pmatrix} \right) \\ &= \begin{pmatrix} -0,06 & 0,06 \\ 0,19 & -0,19 \\ -0,06 & 0,06 \end{pmatrix} \end{aligned}$$

5. Actualitza la matriu de pesos inicials aplicant descens per gradient amb factor d'aprenentatge $\eta = 1,0$

$$\mathbf{W}_{i+1} = \mathbf{W}_i - \eta_i \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n (\boldsymbol{\mu}_n - \mathbf{y}_n)^t$$

$$\begin{aligned} \mathbf{W}_1 &= \mathbf{W}_0 - \eta \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n (\boldsymbol{\mu}_n - \mathbf{y}_n)^t \\ &= \begin{pmatrix} 0 & 0 \\ -0,25 & 0,25 \\ 0 & 0 \end{pmatrix} - 1,0 \begin{pmatrix} -0,06 & 0,06 \\ 0,19 & -0,19 \\ -0,06 & 0,06 \end{pmatrix} \\ &= \begin{pmatrix} 0,06 & -0,06 \\ -0,44 & 0,44 \\ 0,06 & -0,06 \end{pmatrix} \end{aligned}$$