27/01/26

$$\theta_{k+1} \leftarrow \theta_k - h \nabla f(\theta_k)$$

$$h_\theta(x) = w_1 x_1 + w_2 x_2 + \dots + w_n x_n + b = w^T + b = [x^T, 1] \begin{bmatrix} w \\ b \end{bmatrix} = [x^T, 1]\,\theta$$

$$\theta^* = w^*, b^* = \arg \min_{\theta \in \mathbb{R}^{n+1}}$$

$$\boxed{\frac{1}{2m} \sum_{i=1}^{m} \left(y^{(i)} - [x^T, 1]\,\theta\right)^2}$$

$$\frac{\partial E_{in}(\theta)}{\partial n_i} = \frac{1}{2m} \sum_{i=1}^{m} -2\left(y^{(i)} - \hat{y}^{(i)}\right) x_j^{(i)}$$

$$\frac{\partial E_{in}(\theta)}{\partial b} = \frac{1}{2m} \sum_{i=1}^{m} -2\left(y^{(i)} - y^{(i)}\right)$$

$$\nabla E_{in}(\theta_k) = \begin{bmatrix} \sum_{i=1}^{m} \left(y^{(i)} - y^{(i)}\right) x_1^{(i)} \\ \sum_{i=1}^{m} \left(y^{(i)} - \hat{y}^{(i)}\right) x_2^{(i)} \\ \vdots \\ \sum_{i=1}^{M} \left(y^{(i)} - \hat{y}^{(i)}\right) x_n^{(i)} \end{bmatrix}$$

$$-\frac{1}{m} \begin{bmatrix} x_1^{(1)} & x_1^{(2)} & \cdots & x_1^{(m)} \\ x_1^{(1)} & x_2^{(2)} & \cdots & x_2^{(m)} \\ \vdots & & & \\ x_n^{(1)} & x_n^{(2)} & \cdots & x_n^{(m)} \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y^{(1)} - \hat{y}^{(2)} \\ y^{(1)} - \hat{y}^{(2)} \\ \vdots \\ y^{(m)} - \hat{y}^{(m)} \\ (m, 1) \end{bmatrix}$$

42

```
def    dg-ln(X, Y, wØ, bØ, lr, max-epochs, e-tol):
   m=X.shape [0]
   w = wØ.copy()
    b = bØ.copy()
 hist = []

for _ in range(max_epochs)
Y_est = X@ w+b
Err  = Y- Y_est
hist.append(np.square(Err).mean())
grad_w=-(1/Y.shape) X.T@ Err
d_b=Err.mean()
w-=lr*grad_w
b==lr*d-b
If np.abs(grad_w).max()< e_tol:
return w,b,hist
```

```
X, Y
X.shape=[m,n], Y.shape=[n]
W=np.zeros (X.shape[-1])
b=0

W-n, b-n, hist=dg-lin(X,Y, w, b, 0.1, 50, 1e-4)
```

29/01/26

$$a = Pr(y=1 \mid X=x; \theta)$$

$$\hat{y} = \begin{cases} 1 & \text{si } a > h \\ -1 & \text{en otro caso} \end{cases}$$

$h = 0.5$ MAP

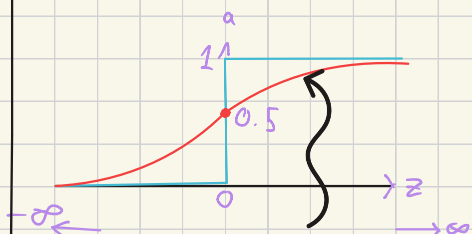
Regresión logística

$$w \leftarrow w - lr\, \nabla_w E_{in}(w,b)$$
$$b \leftarrow b - lr\, \frac{\partial}{\partial b} E_{in}(w,b)$$

$$\frac{\partial}{\partial w_j} E_{in}(w,b) = \frac{\partial}{\partial w_j} \frac{1}{m}\sum_{i=1}^{m} -a^{(i)}\log(\hat{a}^{(i)}) - (1-a^{(i)})\log(1-\hat{a}^{(i)})$$

donde $\hat{a}^{(i)} = \frac{1}{1+e^{-z^{(i)}}}$, y $z^{(i)} = w_1 x_1^{(i)} + \dots + w_n x_n^{(i)} + b$

$$a = f(w_1 x_1 + w_2 x_2 + \dots + w_n x_n + b)$$
$$z = w_1 x_1 + w_2 x_2 + \dots + w_n x_n + b = X^T w + b = X_e^T \theta$$



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

Sigmoide logística

$$\begin{bmatrix} x_1^{(1)} & \dots & x_n^{(1)} \\ x_1^{(2)} & \dots & x_n^{(2)} \\ \vdots & & \vdots \\ x_1^{(m)} & \dots & x_n^{(m)} \end{bmatrix} \quad \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(m)} \end{bmatrix} \quad \begin{bmatrix} a^{(1)} \\ a^{(2)} \\ \vdots \\ a^{(m)} \end{bmatrix}$$

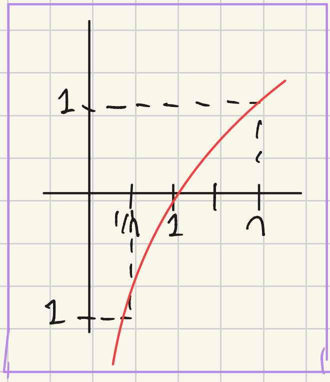
(X) (y) (A)

$$y^{(i)} \in (-1, 1)$$

$$E_{in}(w,b) = \frac{1}{m}\sum_{i=1}^{m} Loss(a^{(i)}, \hat{a}^{(i)})$$

$$Loss(a^{(i)}, \hat{a}^{(i)}) = \begin{cases} -\log(\hat{a}^{(i)}) & \text{si } a^{(i)} = 1 \\ -\log(1-\hat{a}^{(i)}) & \text{si } a^{(i)} = 0 \end{cases}$$

$$Loss(a^{(i)}, \hat{a}^{(i)}) = -a^{(i)}\log(\hat{a}^{(i)}) - (1-a^{(i)})\log(1-\hat{a}^{(i)})$$



03/02/26

Regresión logística

$$\frac{d\,\epsilon in(w,b)}{d\,w_j} = \frac{1}{M}\sum_{i=1}^{M}\left(\frac{a^{(i)}}{\hat{a}^{(i)}} - \frac{1-a^{(i)}}{1-\hat{a}^{(i)}}\right)\hat{a}^{(i)}(1-\hat{a}^{(i)})\,x_j^{(i)}$$

$$\frac{d\,\epsilon in(w,b)}{d\,w_j} = \frac{d}{d\,w_j}\frac{1}{m}\sum_{i=1}^{m}\left[-a^{(i)}\ln(\hat{a}^{(i)}) - (1-a^{(i)})\ln(1-\hat{a}^{(i)})\right]$$

$$= \frac{1}{m}\sum_{i=1}^{m}\frac{-a^{(i)}}{\hat{a}^{(i)}}\frac{d\,\hat{a}^{(i)}}{d\,w_j} + \frac{1-a^{(i)}}{1-\hat{a}^{(i)}}\quad\frac{d\,\hat{a}^{(i)}}{d\,w_j}$$

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

$$z^{(i)} = \sum_{j=1}^{n} w_j x_j^{(i)} + b$$

$$\frac{d\,\hat{a}^{(i)}}{d\,w_j} = \frac{d\,\sigma(z^{(i)})}{d\,w_j} = \frac{d\,\sigma(z^{(i)})}{d\,z}\quad\frac{d\,z^{(i)}}{d\,w_j}$$

$$= \sigma(z^{(i)})(1-\sigma(z^{(i)}))\,x_j^{(i)}$$

$$= \hat{a}^{(i)}(1-\hat{a}^{(i)})\,x_j^{(i)}$$

$$= \frac{1}{m}\sum_{i=1}^{m}\left[-a^{(i)} + (1-a^{(i)})\,\hat{a}^{(i)}\right]x_j^{(i)}$$

$$= \frac{1}{m}\sum_{i=1}^{m}\left[-a^{(i)} + a^{(i)}\hat{a}^{(i)} + \hat{a}^{(i)} - a^{(i)}\hat{a}^{(i)}\right]x_j^{(i)} =$$

$$W \leftarrow W - \frac{lr}{m}\ldots X^{T}(A-\hat{A})$$

$$b \leftarrow b - \frac{lr}{m}\sum_{i=1}^{m}(a^{(i)} - \hat{a}^{(i)})$$

03/02/26

$w^*, b^* = \arg\min \; E_{in}(w, b)$

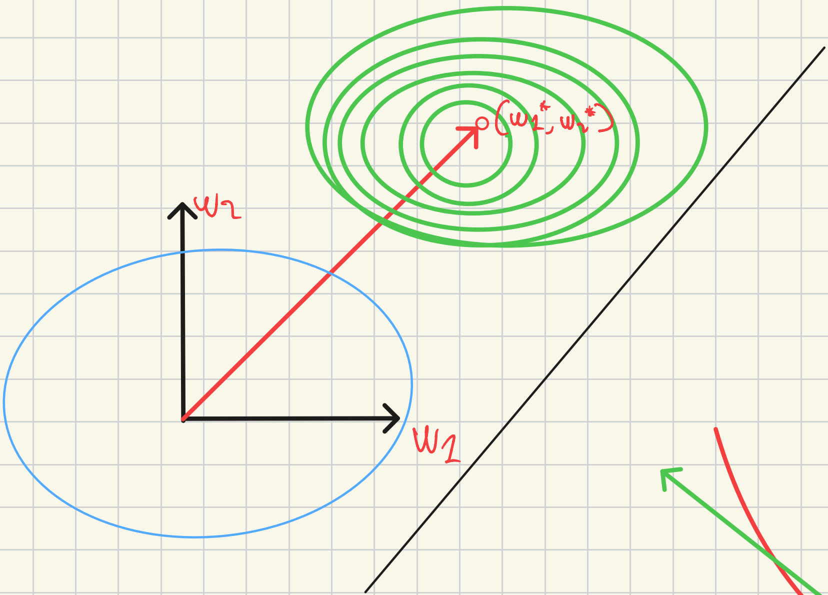
bajo

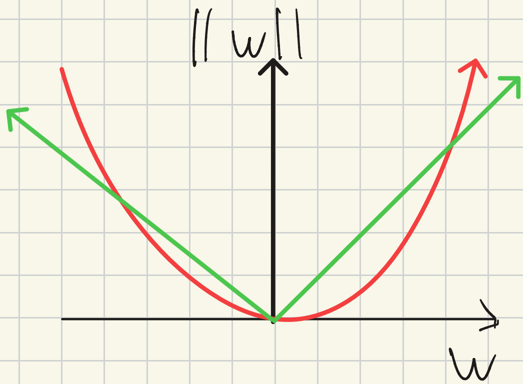
$\sum_{j=1}^{n} w_j^2 \leq C$

$w^*_{res}, b^*_{res} = \arg\min \left[E_n(w,b) + \frac{\lambda}{m} \sum_{j=1}^{n} w_j^2 \right]$

$\frac{d}{dw_j} \left[E_{in}(w,b) + \frac{\lambda}{m} \right] = \frac{-1}{m} \sum_{i=1}^{n} \left(a^{(i)} - \hat{a}^{(i)} \right) x^{(i)}_j + \frac{d \cdot \lambda}{m} w_j$

$w^*_{ras}, b^*_{ras} = \arg\min \left[E_{in}(w,b) + \frac{\lambda}{m} \, regu(w) \right]$





05/02/26

Hypotheses:
decision trees

$f: X \to Y$

Y quiero ajustar

$h_\theta: X \to Y$ a partir de $D = \{(x^{(1)}, y^{(1)}), \ldots, (x^{(m)}, y^{(m)})\}$

$$
\begin{array}{c c c}
 & \begin{array}{cccc} f_1 & f_2 & \cdots & f_n \end{array} & y \\
\begin{array}{c} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(m)} \end{array} &
\begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \cdots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \cdots & x_n^{(2)} \\ \vdots & \vdots & & \vdots \\ x_1^{(m)} & x_2^{(m)} & & x_n^{(m)} \end{bmatrix} &
\begin{vmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(m)} \end{vmatrix}
\end{array}
$$

```
def genera_arbol(features, x, y, nodo):
    Si todos los datos misma clase ó
    no hay datos ó no hay feature

        Regresa nodo

    Var = escoger feature(features, x, y)
    Quitar var de features

    Par valor en valores(var):
        xh, yh = Separa_datos(x, y, var, valor)
        nh = crea_hijo(n, yh)
        nh = Genera_arbol(features, xh, yh, nh)
```

Entropy

$$H(Y) = -\sum_{i=1}^{k} P(Y=y_i)\log_2 P(Y=y_i)$$