Regresión Logística

Rafael Zambrano
rafazamb@gmail.com

Regresión Logística

- Utilizada para predecir probabilidades

- Ejemplo: Probabilidad de sufrir hipertensión

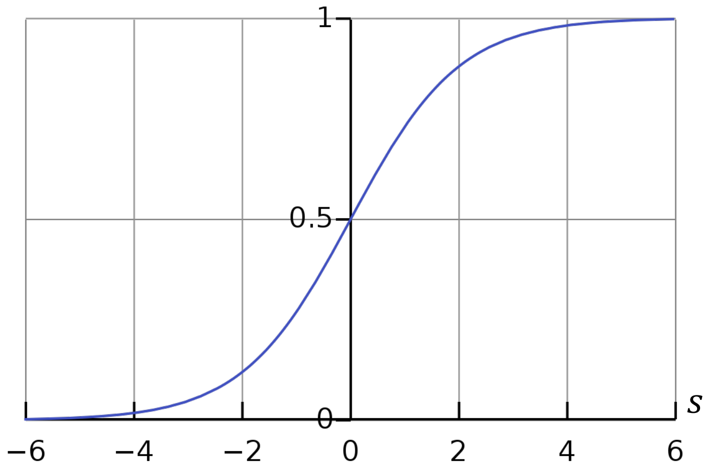
Edad	Altura	Peso	Colesterol	Fumador/a	Padece hipertensión
53	170	70	120	1	SÍ
67	156	85	240	0	NO
21	191	56	100	0	NO
34	182	77	500	1	SÍ
…	…	…	…	…	…

- Para un nuevo paciente, el modelo devolverá un valor de probabilidad de padecer hipertensión (entre 0 y 1)

Regresión Logística

- En la regresión lineal, teníamos $h(x) = \sum w_i x_i = \boldsymbol{w}^T \boldsymbol{x} = \mathrm{w}_0 + \mathrm{w}_1 x_1 + w_2 x_2 + \cdots$
- En regresión logística, tenemos $h(x) = \theta(\sum w_i x_i) = \theta(\boldsymbol{w}^T \boldsymbol{x})$
- Función sigmoide

$$\theta(s) = \frac{e^s}{1+e^s}$$

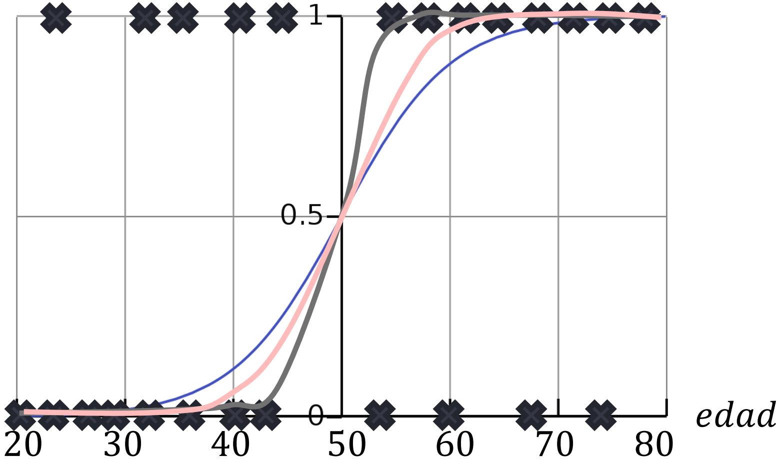


Regresión Logística

- A modo de ejemplo, vamos a usar solamente la edad como variable explicativa de padecer hipertensión

- $h(x) = \theta(w_0 + w_1 \cdot \ edad)$

Edad	Padece hipertensión
53	1
67	0
21	0
34	1
…	…



Regresión Logística

- El hecho de padecer hipertensión se ve afectado por una probabilidad a la cual no tenemos acceso ni podemos medir
- Sólo podemos observar la ocurrencia de un evento e intentar **inferir** esa probabilidad
- $P(y|x) = \begin{cases} \theta(\boldsymbol{w}^T\boldsymbol{x}) & \text{si } y = +1 \\ 1 - \theta(\boldsymbol{w}^T\boldsymbol{x}) & \text{si } y = -1 \end{cases} = \begin{cases} \theta(w_0 + w_1 \cdot \ edad) & \text{si } y = +1 \\ 1 - \theta(w_0 + w_1 \cdot \ edad) & \text{si } y = -1 \end{cases}$

$w_0 = -3.2$
$w_1 = 0.07$

Edad	Padece hipertensión	$\theta(w_0 + w_1 \cdot \ edad)$	P(y|x)
53	1	0.62	0.62
67	-1	0.82	0.18
21	-1	0.15	0.85
34	1	0.31	0.31
…	…	…	…

Hay que escoger los pesos *w* tales que maximicen el producto de las P(y|x)

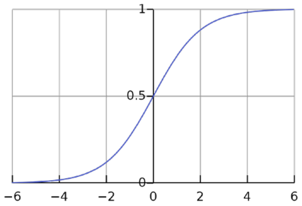
$$\Rightarrow \boldsymbol{max} \prod P(y|x)$$

$$\prod P(y|x) = 0.62 \times 0.18 \times 0.85 \times 0.31 \times \cdots$$

Regresión Logística

- Hay que escoger los pesos w tales que maximicen el producto de las P(y|x)

$$P(y|x) = \begin{cases} \theta(\boldsymbol{w}^T\boldsymbol{x}) & \text{si } y = +1 \\ 1 - \theta(\boldsymbol{w}^T\boldsymbol{x}) & \text{si } y = -1 \end{cases} = \begin{cases} \theta(\boldsymbol{w}^T\boldsymbol{x}) & \text{si } y = +1 \\ \theta(-\boldsymbol{w}^T\boldsymbol{x}) & \text{si } y = -1 \end{cases}$$



$$P(y|x) = \theta(y \cdot \boldsymbol{w}^T\boldsymbol{x})$$

$$1 - \theta(s) = \theta(-s)$$

$\Rightarrow$ maximizar $\prod \theta(y \cdot \boldsymbol{w}^T\boldsymbol{x})$

$\Rightarrow$ maximizar 1/N ln $\prod \theta(y \cdot \boldsymbol{w}^T\boldsymbol{x})$

$\Rightarrow$ minimizar - 1/N ln $\prod \theta(y \cdot \boldsymbol{w}^T\boldsymbol{x})$

$\Rightarrow$ minimizar 1/N $\sum -\ln\, \theta(y \cdot \boldsymbol{w}^T\boldsymbol{x})$

$\Rightarrow$ minimizar 1/N $\sum \ln\ 1/\theta(y \cdot \boldsymbol{w}^T\boldsymbol{x})$

$\Rightarrow$ minimizar 1/N $\sum \ln\,(1 + e^{-y \cdot \boldsymbol{w}^T\boldsymbol{x}})$

$$\theta(s) = \frac{e^s}{1+e^s} = \frac{1}{1+e^{-s}}$$

$$\Rightarrow \frac{1}{\theta(s)} = 1 + e^{-s}$$

Regresión Logística

- Minimizar $E_{in}(\boldsymbol{w}) = \frac{1}{N}\sum \ln(1 + e^{-y\boldsymbol{w}^T\boldsymbol{x}})$

- A diferencia de la regresión lineal, no hay una solución analítica para esta minimización

- La función $E_{in}(\boldsymbol{w})$ sólo tiene un mínimo

- Para minimizar la función se utiliza el método del **Gradiente Descendente**

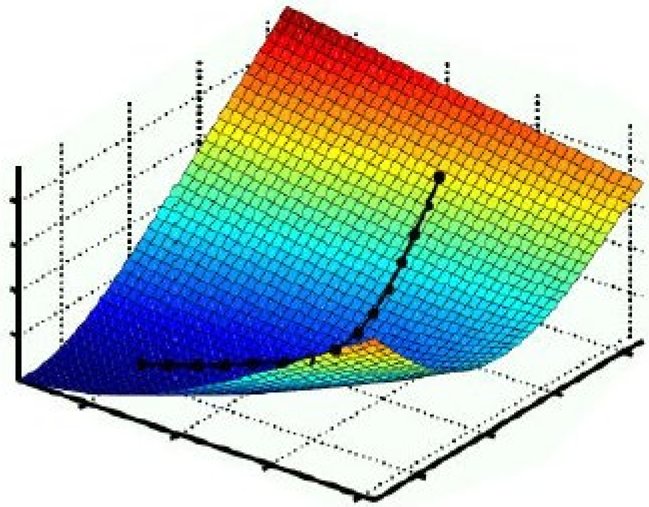
Regresión Logística

- Minimizar $E_{in}(\boldsymbol{w}) = \frac{1}{N}\sum \ln(1 + e^{-y\boldsymbol{w}^T\boldsymbol{x}})$

- La mejor dirección para "moverse" es el negativo del gradiente

$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - \alpha \nabla E_{in}$$

$$\nabla E_{in} = -\frac{1}{N}\sum \frac{y \cdot \boldsymbol{x}}{1 + e^{y\boldsymbol{w}^T\boldsymbol{x}}}$$



$w(0) \to w(1) \to w(2) \to w(3) \ldots$

Regresión Logística

$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - \alpha \nabla E_{in}$$

$$\nabla E_{in} = -\frac{1}{N}\sum \frac{y \cdot \boldsymbol{x}}{1+e^{y\boldsymbol{w}^T\boldsymbol{x}}} = -\frac{1}{2}\left([0.059 \quad 4.16 \quad 3.16] + [-0.89 \quad -50.89 \quad -59.82]\right)$$

$$= [0.42 \quad 23.36 \quad 28.33]$$

Peso	Edad	Padece hipertensión
70	53	1
57	67	-1

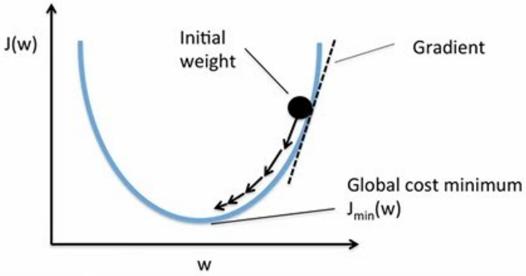
$$w_0 = -3.2$$
$$w_1 = 0.07$$
$$w_2 = 0.02$$

$$w_0 = -3.2 - \alpha \cdot 0.42$$
$$w_1 = 0.07 - \alpha \cdot 23.36$$
$$w_2 = 0.02 - \alpha \cdot 28.33$$

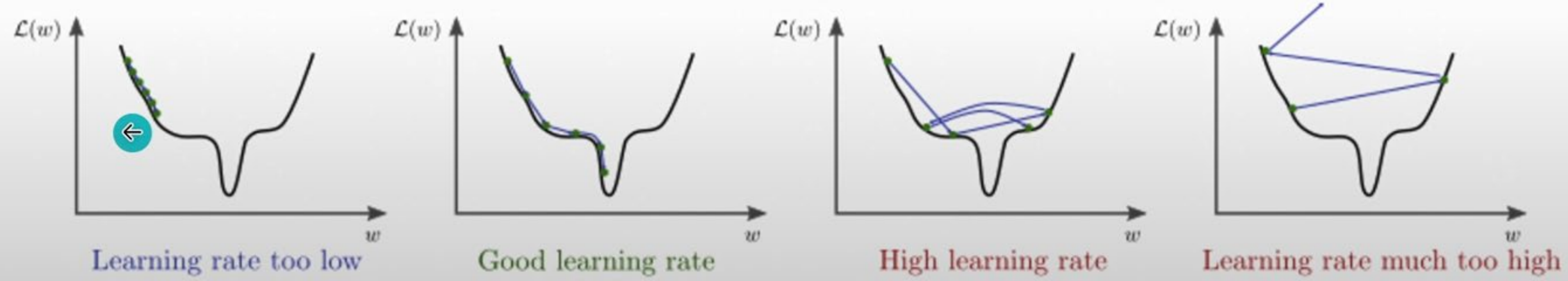
$$\frac{y \cdot \boldsymbol{x}}{1+e^{y\boldsymbol{w}^T\boldsymbol{x}}} = \frac{1 \cdot [1 \quad 70 \quad 53]}{1+e^{1 \cdot [-3.2 \quad 0.07 \quad 0.02]^T [1 \quad 70 \quad 53]}} = \frac{[1 \quad 70 \quad 53]}{1+e^{2.76}} = \frac{[1 \quad 70 \quad 53]}{1+15.79} = [0.059 \quad 4.16 \quad 3.16]$$

$$\frac{y \cdot \boldsymbol{x}}{1+e^{y\boldsymbol{w}^T\boldsymbol{x}}} = \frac{-1 \cdot [1 \quad 57 \quad 67]}{1+e^{-1 \cdot [-3.2 \quad 0.07 \quad 0.02]^T [1 \quad 57 \quad 67]}} = \frac{[-1 \quad -57 \quad -67]}{1+e^{-2.13}} = \frac{[-1 \quad -57 \quad -67]}{1.12} = [-0.89 \quad -50.89 \quad -59.82]$$

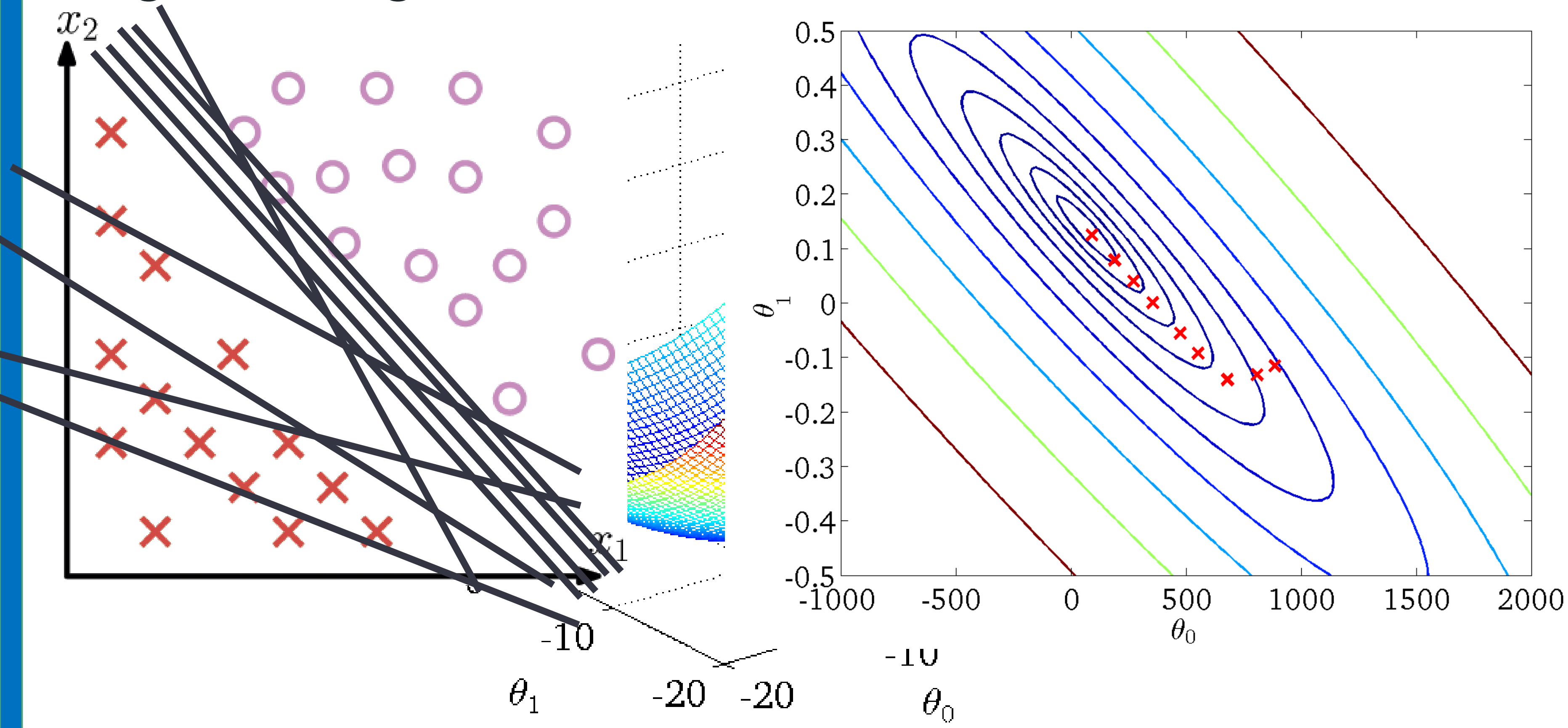
Regresión Logística



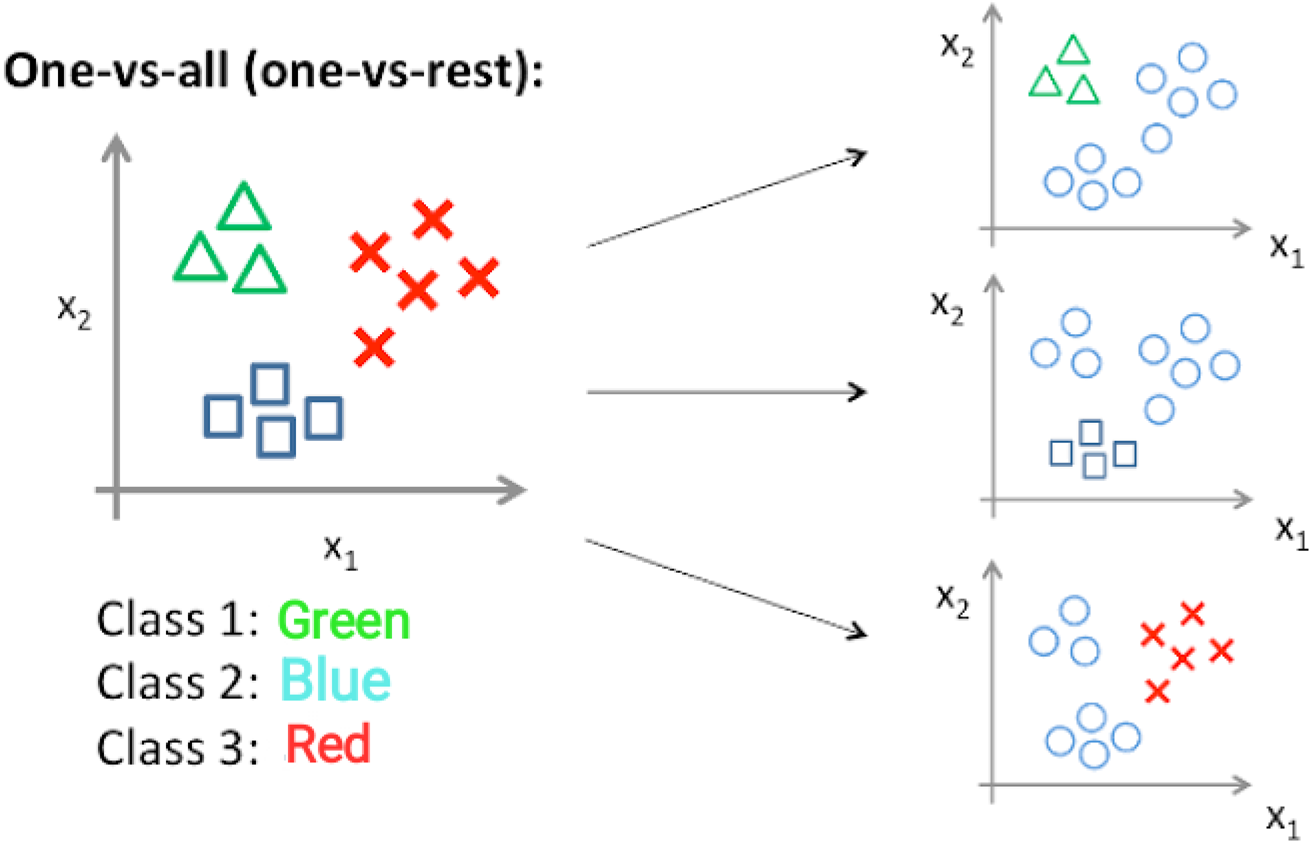
$$w = w - \alpha \cdot dw$$



Regresión Logística



Clasificación multiclase



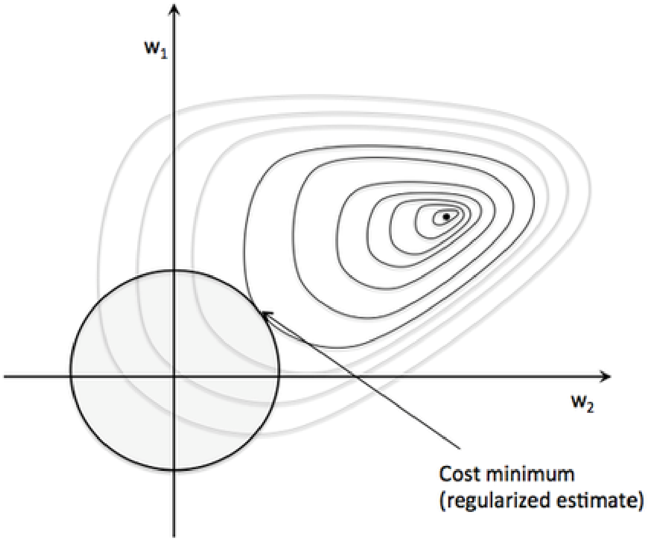
Regularización

- Regularización L2:

$$E_{in}(\boldsymbol{w}) = \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + C\sum \ln(1 + e^{-y\boldsymbol{w}^T\boldsymbol{x}})$$

- Regularización L1:

$$E_{in}(\boldsymbol{w}) = \|\boldsymbol{w}\|_1 + C\sum \ln(1 + e^{-y\boldsymbol{w}^T\boldsymbol{x}})$$



¡Gracias!

Contacto: Rafael Zambrano

rafazamb@gmail.com