Logistic regression on m examples

$J=0;\ dw_1=0;\ dw_2=0;\ db=0$

For $i=1$ to m

$z^{(i)}=w^Tx^{(i)}+b$

$a^{(i)}=\sigma(z^{(i)})$

$J \mathrel{+}= -[y^{(i)}\log a^{(i)}+(1-y^{(i)})\log(1-a^{(i)})]$

$dz^{(i)}=a^{(i)}-y^{(i)}$

$dw_1 \mathrel{+}= x_1^{(i)}dz^{(i)}$

$dw_2 \mathrel{+}= x_2^{(i)}dz^{(i)}$

$db \mathrel{+}= dz^{(i)}$

dw_3
$\vdots$
dw_n

$n=2$

$J \mathrel{/}= m$

$dw_1 \mathrel{/}= m;\ dw_2 \mathrel{/}= m;\ db \mathrel{/}= m.$

$$dw_1=\frac{\partial J}{\partial w_1}$$

$w_1 := w_1-\alpha\, dw_1$

$w_2 := w_2-\alpha\, dw_2$

$b := b-\alpha\, db$

Vectorization

Logistic regression derivatives

J = 0, ~~dw1 = 0, dw2 = 0~~, db = 0 dw = np.zeros((n_x, 1))

for i = 1 to n:

$z^{(i)} = w^T x^{(i)} + b$

$a^{(i)} = \sigma(z^{(i)})$

$J \mathrel{+}= -\left[y^{(i)} \log \hat{y}^{(i)} + (1 - y^{(i)}) \log(1 - \hat{y}^{(i)})\right]$

$dz^{(i)} = a^{(i)}(1 - a^{(i)})$

for j = 1 … n_x

$dw_j \mathrel{+}= \ldots$

$\cancel{dw_1 \mathrel{+}= x_1^{(i)} dz^{(i)}}$

$\cancel{dw_2 \mathrel{+}= x_2^{(i)} dz^{(i)}}$ $n_x = 2$ $dw \mathrel{+}= x^{(i)} dz^{(i)}$

$db \mathrel{+}= dz^{(i)}$

J = J/m, ~~$dw_1 = dw_1/m$, $dw_2 = dw_2/m$~~, db = db/m

dw /= m.

Vectorizing Logistic Regression

$z^{(1)} = \boxed{w^T x^{(1)} + b}$ $\quad z^{(2)} = \boxed{w^T x^{(2)} + b}$ $\quad z^{(3)} = w^T x^{(3)} + b$

$a^{(1)} = \sigma(z^{(1)})$ $\quad a^{(2)} = \sigma(z^{(2)})$ $\quad a^{(3)} = \sigma(z^{(3)})$

$X = \begin{bmatrix} | & | & & | \\ x^{(1)} & x^{(2)} & \cdots & x^{(m)} \\ | & | & & | \end{bmatrix}$ $\quad (n_x, m) \quad \mathbb{R}^{n_x \times m}$

$[\text{———}] \quad w^T \begin{bmatrix} | & | & & | \\ x^{(1)} & x^{(2)} & \cdots & x^{(m)} \\ | & | & & | \end{bmatrix}$

$Z = \boxed{[z^{(1)} \; z^{(2)} \cdots z^{(m)}]} = \underbrace{w^T X + \underbrace{[b \; b \cdots b]}_{1\times m}} = [\boxed{w^T x^{(1)} + b} \; \boxed{w^T x^{(2)} + b} \cdots w^T x^{(m)} + b]$

$1 \times m$

$\rightarrow Z = np.dot(w.T, X) + b$ $\quad (1,1) \quad \mathbb{R}$ $\quad$ "Broadcasting"

$A = [a^{(1)} \; a^{(2)} \cdots a^{(m)}] = \sigma(Z)$

Vectorizing Logistic Regression

$dz^{(1)} = a^{(1)} - y^{(1)}$ $\quad$ $dz^{(2)} = a^{(2)} - y^{(2)}$ $\quad \ldots$

$$\boxed{dZ = [dz^{(1)} \; dz^{(2)} \cdots dz^{(m)}]} \leftarrow$$

$1 \times m$

$A = [a^{(1)} \cdots a^{(m)}]$. $\quad Y = [y^{(1)} \cdots y^{(m)}]$

$\rightarrow dZ = A - Y = [a^{(1)} - y^{(1)} \quad a^{(2)} - y^{(2)} \quad \ldots]$

$\rightarrow dw = 0$

$dw \mathrel{+}= x^{(1)} dz^{(1)}$

$dw \mathrel{+}= x^{(2)} dz^{(2)}$

$\vdots$

$dw /= m$

$\cancel{dw_1}$

$\cancel{dw_2}$

$\vdots$

$db = 0$

$db \mathrel{+}= dz^{(1)}$

$db \mathrel{+}= dz^{(2)}$

$\vdots$

$db \mathrel{+}= dz^{(m)}$

$db /= m$.

$$\boxed{db = \frac{1}{m} \sum_{i=1}^{m} dz^{(i)} = \frac{1}{m} \, np.sum(dZ)}$$

$$\boxed{dw = \frac{1}{m} X \, dZ^T}$$

$$= \frac{1}{m} \begin{bmatrix} | & & | \\ x^{(1)} & \cdots & x^{(m)} \\ | & & | \end{bmatrix} \begin{bmatrix} dz^{(1)} \\ \vdots \\ dz^{(m)} \end{bmatrix}$$

$$= \frac{1}{m} \left[x^{(1)} dz^{(1)} + \ldots + x^{(m)} dz^{(m)} \right]$$

$n \times 1$

Implementing Logistic Regression

J = 0, dw_1 = 0, dw_2 = 0, db = 0

$\boxed{\texttt{for i = 1 to m:}}$

$z^{(i)} = w^T x^{(i)} + b$

$a^{(i)} = \sigma(z^{(i)})$

$J \mathrel{+}= -[y^{(i)} \log a^{(i)} + (1 - y^{(i)}) \log(1 - a^{(i)})]$

$dz^{(i)} = a^{(i)} - y^{(i)}$

$\left.\begin{array}{l} dw_1 \mathrel{+}= x_1^{(i)} dz^{(i)} \\ dw_2 \mathrel{+}= x_2^{(i)} dz^{(i)} \end{array}\right\}$ $dw \mathrel{+}= x^{(i)} * dz^{(i)}$

db += $dz^{(i)}$

J = J/m, $dw_1 = dw_1$/m, $dw_2 = dw_2$/m

db = db/m

for iter in range(1000):

$Z = w^T X + b$

$= np.dot(w.T, X) + b$

$A = \sigma(Z)$

$dZ = A - Y$

$dw = \frac{1}{m} X dZ^T$

$db = \frac{1}{m}$ np.sum(dZ)

$w := w - \alpha dw$

$b := b - \alpha db$