

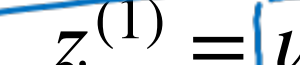


deeplearning.ai

Basics of Neural Network Programming

Vectorizing Logistic Regression

Vectorizing Logistic Regression


$$z^{(1)} = w^T x^{(1)} + b$$
$$a^{(1)} = \sigma(z^{(1)})$$

$$z^{(2)} = w^T x^{(2)} + b$$
$$a^{(2)} = \sigma(z^{(2)})$$

$$z^{(3)} = w^T x^{(3)} + b$$
$$a^{(3)} = \sigma(z^{(3)})$$

$$\underline{\underline{X}} = \begin{bmatrix} x^{(1)} & x^{(2)} & \dots & x^{(n)} \\ | & | & & | \\ | & | & & | \\ | & | & & | \end{bmatrix}$$

$$\frac{(n_x, m)}{\mathbb{R}^{n_x \times m}}$$

$$\vec{1} \rightarrow \omega \left[\begin{array}{c} 1 \\ x^{(1)} \\ 1 \end{array} \quad \begin{array}{c} 1 \\ x^{(2)} \\ 1 \end{array} \quad \dots \quad \begin{array}{c} 1 \\ x^{(m)} \\ 1 \end{array} \right]$$

$$\underline{Z} = \begin{bmatrix} \underline{z}^{(1)} & \underline{z}^{(2)} & \dots & \underline{z}^{(m)} \end{bmatrix} = \underline{w}^T \underline{X} + \underbrace{\begin{bmatrix} b & b & \dots & b \end{bmatrix}}_{1 \times m} = \underbrace{\begin{bmatrix} \underline{w}^T \underline{x}^{(1)} + b \end{bmatrix}}_{1 \times m} \underbrace{\begin{bmatrix} \underline{w}^T \underline{x}^{(2)} + b \end{bmatrix}}_{1 \times m} \dots \underbrace{\begin{bmatrix} \underline{w}^T \underline{x}^{(m)} + b \end{bmatrix}}_{1 \times m}$$

$$\rightarrow \underline{z = np.dot(w.T, x) + \frac{b}{n}}_{(1,1)} \quad \mathbb{R}$$

$$A = [a^{(1)} \ a^{(2)} \ \dots \ a^{(m)}] = \underset{\substack{\uparrow \\ \text{columns}}}{\sigma(z)}$$



deeplearning.ai

Basics of Neural Network

Programming Vectorizing Logistic Regression's Gradient Computation

Vectorizing Logistic Regression

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

$$\underline{dz} = [\underline{dz^{(1)}} \quad \underline{dz^{(2)}} \quad \dots \quad \underline{dz^{(m)}}] \quad \leftarrow$$

$1 \times m$

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

$$\rightarrow \underline{dz} = A - Y = [\underline{a^{(1)} - y^{(1)}} \quad \underline{a^{(2)} - y^{(2)}} \quad \dots]$$

$$\begin{aligned} \rightarrow \underline{dw} &= 0 \\ \underline{dw} &+= \underline{x^{(1)} dz^{(1)}} \\ \underline{dw} &+= \underline{x^{(2)} dz^{(2)}} \\ &\vdots \\ \underline{dw} &= m \end{aligned}$$

~~dw_1~~
 ~~dw_2~~
 ~~\vdots~~

$$\begin{aligned} \underline{db} &= 0 \\ \underline{db} &+= \underline{dz^{(1)}} \\ \underline{db} &+= \underline{dz^{(2)}} \\ &\vdots \\ \underline{db} &+= \underline{dz^{(m)}} \\ \underline{db} &= m. \end{aligned}$$

$$\begin{aligned} \underline{db} &= \frac{1}{m} \sum_{i=1}^m dz^{(i)} \\ &= \frac{1}{m} \text{np.sum}(\underline{dz}) \end{aligned}$$

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

$$\begin{aligned} &= \frac{1}{m} \begin{bmatrix} x^{(1)} & \dots & x^{(m)} \\ 1 & & 1 \end{bmatrix} \begin{bmatrix} dz^{(1)} \\ \vdots \\ dz^{(m)} \end{bmatrix} \\ &= \frac{1}{m} \left[\underline{x^{(1)} dz^{(1)}} + \dots + \underline{x^{(m)} dz^{(m)}} \right] \\ &\quad n \times 1 \end{aligned}$$

Implementing Logistic Regression

$J = 0, d = 0, d = 0, db = 0$

for $i = 1$ to m :

$=$

$=$

$+=$

$d =$

$d+=$

$d+=$

$db += d$

$J = J/m, d = d/m, d = d/m$

$db = db/m$

$\left. \begin{array}{l} d+= \\ d+= \end{array} \right\} dw += x^{(i)} * dz^{(i)}$

for $iter$ in $range(1000)$: \leftarrow

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

$A = \sigma(z)$

$dz = A - Y$

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

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

$w := w - \alpha dw$
 $b := b - \alpha db$