Basics of Neural Network Programming

Vectorizing Logistic Regression

deeplearning.ai

Vectorizing Logistic Regression

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

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

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

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

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

$1 \times m$

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

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

Andrew Ng

Basics of Neural Network Programming

Vectorizing Logistic Regression's Gradient Computation

deeplearning.ai

Vectorizing Logistic Regression

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

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

$1 \times m$

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

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

$\rightarrow dw = 0$

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

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

$\vdots$

$dw /= m$

$\cancel{dw_1}$

$\cancel{dw_2}$

$\vdots$

$db = 0$

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

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

$\vdots$

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

$db /= m$.

$$db = \frac{1}{m} \sum_{i=1}^{m} dz^{(i)}$$

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

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

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

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

$n \times 1$

Andrew Ng

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
```

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

for iter in range(1000):

$Z = w^T X + b$

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

$A = \sigma(Z)$

$dZ = A - Y$

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

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

$w := w - \alpha dw$

$b := b - \alpha db$