

0303 Vectorización de la Regresión Logística

La vectorización

es una técnica que nos permitía implementar nuestros algoritmos de forma eficiente

Pseudocódigo con m instancias

```
logistic_regression(x1[m], x2[m], y[m]):  
    J, dw1, dw2, db := 0;  
    z, y_hat, dz := array[m];  
    for i = 1 to m  
        z[i] := w1*x1[i] + w2*x2[i] + b;  
        y_hat[i] := sigmoid(z[i]);  
        J += L(y[i], y_hat[i]);  
        dz[i] := y_hat[i] - y[i];  
        dw1 += x1[i]*dz[i]; dw2 += x2[i]*dz[i];  
        db += dz[i];  
    J /= m, dw1 /=m, dw2 /=m, db /=m;
```

regresion logistica

El código anterior es ineficiente, vamos a vectorizarlo

calcular los m z de forma vectorizada

Vectorización Prop. Adelante

$$\begin{aligned}
 z^{(1)} &= w^T x^{(1)} + b, z^{(2)} = w^T x^{(2)} + b, \dots, z^{(m)} = w^T x^{(m)} + b \\
 w^T X &= [w_1 \quad w_2 \quad \dots \quad w_n]_{(1,n)} \begin{bmatrix} x^{(1)} & x^{(2)} & \dots & x^{(m)} \\ | & | & \dots & | \\ | & | & \dots & | \end{bmatrix}_{(n,m)} \\
 w^T X &= [w^T x^{(1)} \quad w^T x^{(2)} \quad \dots \quad w^T x^{(m)}]_{(1,m)} \\
 w^T X + b &= [w^T x^{(1)} + b \quad w^T x^{(2)} + b \quad \dots \quad w^T x^{(m)} + b]_{(1,m)} \\
 w^T X + b &= [z^{(1)} \quad z^{(2)} \quad \dots \quad z^{(m)}]_{(1,m)} = Z
 \end{aligned}$$

Vectorización Prop. Adelante

$$\begin{aligned}
 Z &= w^T X + b \\
 \hat{Y} &= \sigma(Z) = [\sigma(z^{(1)}) \quad \sigma(z^{(2)}) \quad \dots \quad \sigma(z^{(m)})] \\
 A &= \hat{Y} \\
 J &= ?
 \end{aligned}$$

Vectorización Prop. Atrás

$$\begin{aligned}
 dw &= 1/m \cdot X dZ^T \\
 &= 1/m \cdot \begin{bmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(m)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(m)} \\ \vdots & \vdots & \ddots & \vdots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(m)} \end{bmatrix} \cdot \begin{bmatrix} dz^{(1)} \\ dz^{(2)} \\ \vdots \\ dz^{(m)} \end{bmatrix} = 1/m \cdot \begin{bmatrix} \sum_{i=1}^m x_1^{(i)} dz^{(i)} \\ \vdots \\ \sum_{i=1}^m x_n^{(i)} dz^{(i)} \end{bmatrix}
 \end{aligned}$$

vectorizacion prop atras:

$$dw = \begin{bmatrix} dw_1 \\ dw_2 \\ \vdots \\ dw_n \end{bmatrix} = 1/m \cdot \begin{bmatrix} \sum_{i=1}^m x_1^{(i)} dz^{(i)} \\ \sum_{i=1}^m x_2^{(i)} dz^{(i)} \\ \vdots \\ \sum_{i=1}^m x_n^{(i)} dz^{(i)} \end{bmatrix}$$

$$db = 1/m \cdot \sum_{i=1}^m dz^{(i)} \\ = 1/m * np.sum(dZ)$$

Propagación hacia adelante

$$Z = w^T X + b$$

$$A = \sigma(Z)$$

Propagación hacia atrás

$$dZ = A - Y$$

$$dw = 1/m \cdot X dZ^T$$

$$db = 1/m * np.sum(dZ)$$