Given the following two-dimensional points and their actual labels:

$$\mathbf{x}_A = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad , \quad y_A = 0$$

$$\mathbf{x}_B = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad , \quad y_B = 0$$

$$\mathbf{x}_C = \begin{pmatrix} 3 \\ 3 \end{pmatrix} \quad , \quad y_C = 1$$

$$\mathbf{x}_D = \begin{pmatrix} 4 \\ 3 \end{pmatrix} \quad , \quad y_D = 1$$

If we initial the vector of weights for each dimension (including w_0) as

$\tilde{\mathbf{w}} = \begin{pmatrix} -5 \\ 2 \\ 1 \end{pmatrix}$. What's the vector of weights using Logistic Regression Model after only one iteration by gradient decent (the learning rate η = 0.1)?

Answer:

Based on the initial weights, we have

$$\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{x}}_A = \begin{pmatrix} -5 \\ 2 \\ 1 \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = w_0 \times 1 + w_1 \times x_A^1 + w_2 \times x_A^2 = -5 \times 1 + 2 \times 0 + 1 \times 1 = -4$$

$$\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{x}}_B = -5 \times 1 + 2 \times 1 + 1 \times 1 = -2$$

$$\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{x}}_C = -5 \times 1 + 2 \times 3 + 1 \times 3 = 4$$

$$\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{x}}_D = -5 \times 1 + 2 \times 4 + 1 \times 3 = 6$$

According to the method of gradient decent, the Logistic Regression Model updates weights as follows:

$$
\begin{aligned}
w_0^{new} &= w_0^{old} - \eta \sum \left[\left(\frac{e^{\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{x}}}}{1+e^{\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{x}}}} - y \right) x_0 \right] \\
&= -5 - \eta \left[\left(\frac{e^{-4}}{1+e^{-4}} - 0 \right) \times 1 + \left(\frac{e^{-2}}{1+e^{-2}} - 0 \right) \times 1 + \left(\frac{e^{4}}{1+e^{4}} - 1 \right) \times 1 + \left(\frac{e^{6}}{1+e^{6}} - 1 \right) \times 1 \right] \\
&= -5 - \eta \times 0.1167 \\
&= -5.0117
\end{aligned}
$$

$$
\begin{aligned}
w_1^{new} &= w_1^{old} - \eta \sum \left[\left(\frac{e^{\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{x}}}}{1+e^{\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{x}}}} - y \right) x_1 \right] \\
&= 2 - \eta \left[\left(\frac{e^{-4}}{1+e^{-4}} - 0 \right) \times 0 + \left(\frac{e^{-2}}{1+e^{-2}} - 0 \right) \times 1 + \left(\frac{e^{4}}{1+e^{4}} - 1 \right) \times 3 + \left(\frac{e^{6}}{1+e^{6}} - 1 \right) \times 4 \right] \\
&= 2 - \eta \times 0.0554 \\
&= 1.9945
\end{aligned}
$$

$$
\begin{aligned}
w_2^{new} &= w_2^{old} - \eta \sum \left[\left(\frac{e^{\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{x}}}}{1+e^{\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{x}}}} - y \right) x_2 \right] \\
&= 1 - \eta \left[\left(\frac{e^{-4}}{1+e^{-4}} - 0 \right) \times 1 + \left(\frac{e^{-2}}{1+e^{-2}} - 0 \right) \times 1 + \left(\frac{e^{4}}{1+e^{4}} - 1 \right) \times 3 + \left(\frac{e^{6}}{1+e^{6}} - 1 \right) \times 3 \right] \\
&= 1 - \eta \times 0.0758 \\
&= 0.9924
\end{aligned}
$$

Thus, $\tilde{\mathbf{w}}^{new} = \begin{pmatrix} -5.0117 \\ 1.9945 \\ 0.9924 \end{pmatrix}$.