Homework #4

Friday, 13 December 2024 20:24

I

1.1

First iter :

$a_1 = w^T x + b = 0,1 \cdot 1 + 0,1 \cdot 0 + 0,1 = 0,2$

$\hookrightarrow h_1 = \frac{1}{1+e^{-a_1}} = 0,5498$

$o = w'_1 h_1 + b' = 0,1 \cdot 0,5498 + 0,1 = 0,15498$

$\hookrightarrow \hat{y} = \frac{1}{1+e^{-o}} = 0,5387$

$L = -y \log(\hat{y}) - (1-y)\log(1-\hat{y})$

Gradients :

Out put layers:

$\frac{\partial L}{\partial w'_1} = (\hat{y} - y) h_1 = (0,5387 - 1) \cdot 0,5498 = -0,2536$

$\Rightarrow w'_1 \leftarrow w'_1 - \eta \cdot (-0,2536) = 0,176$

$\frac{\partial L}{\partial b'} = \hat{y} - y = -0,4613$

$\Rightarrow b' \leftarrow 0,238$

Hidden layers:

$\frac{\partial L}{\partial w_1} = (\hat{y} - y) \cdot w'_1 \cdot h_1(1-h_1) \cdot x_1 = -0,0114$

$\Rightarrow w_1 \leftarrow 0,1034$

$\frac{\partial L}{\partial w_2} = (\hat{y} - y) \cdot w'_2 \, h(1-h_1) \cdot x_2 = 0$

$\Rightarrow w_2 \leftarrow 0,1$

$\frac{\partial L}{\partial b} = (\hat{y} - y) \cdot w'_1 \cdot h_1(1-h_1) = -0,0114$

$\Rightarrow b \leftarrow 0,1034$

Do the same with second iteration :

$w'_1 \leftarrow 0,0798$

$b' \leftarrow 0,0622$

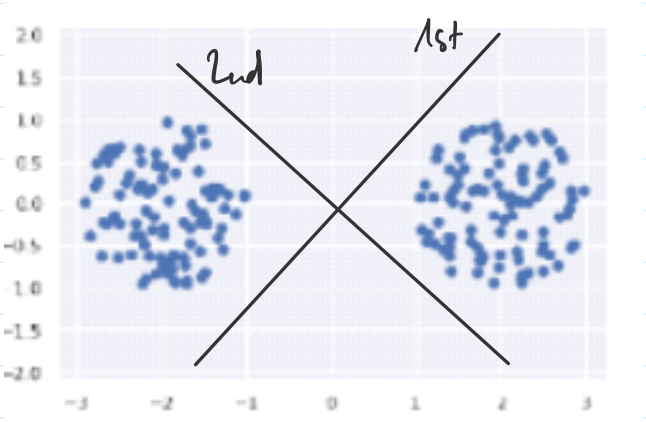
$w_1 \leftarrow 0,1034$

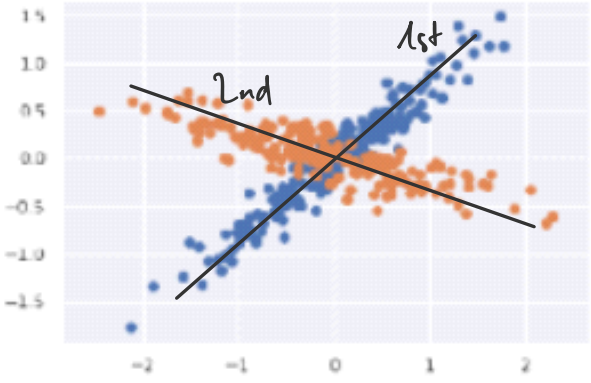
$w_2 \leftarrow 0,0924$

$b \leftarrow 0,0954$

1.2

a/





b/

Mean: $\mu_1 = 3,5$ $\mu_2 = 3,5$ $\Rightarrow X_c = \begin{bmatrix} x_1-\mu_1 & x_2-\mu_2 \end{bmatrix}$

$$\Rightarrow X_{centered} = \begin{bmatrix} -1,5 & -1,5 \\ 0,5 & -0,5 \\ 1,5 & 1,5 \\ 1,5 & 0,5 \\ -0,5 & 0,5 \\ -1,5 & -0,5 \end{bmatrix} \Rightarrow X_c^T X_c = \begin{bmatrix} 9,5 & 5,5 \\ 5,5 & 5,5 \end{bmatrix}$$

$$Cov(X) = \Sigma = \frac{1}{n-1} X_c^T X_c = \begin{bmatrix} 1,9 & 1,1 \\ 1,1 & 1,1 \end{bmatrix}$$

$$\det(\Sigma - \lambda I) = 0 \Rightarrow \begin{bmatrix} 1,9-\lambda & 1,1 \\ \cdots & \cdots \end{bmatrix} = 0$$

$$\det(\Sigma - \lambda I) = 0 \Rightarrow \begin{bmatrix} 1,9-\lambda & 1,1 \\ 1,1 & 1,1-\lambda \end{bmatrix} = 0$$

$$\Leftrightarrow \lambda^2 - 3\lambda + 0,88 = 0$$

Eigenvalue $\in \begin{cases} \lambda_1 = 2,67 \\ \lambda_2 = 0,33 \end{cases}$

$\Rightarrow$ Eigenvector $v_1 \Rightarrow \begin{bmatrix} -0,77 & 1,1 \\ 1,1 & -1,57 \end{bmatrix} v = 0$

$$\Rightarrow v_1 = \begin{bmatrix} 1,4276 \\ 1 \end{bmatrix}$$

$\Rightarrow$ Project X_c onto $v_1 \rightarrow \hat{X} = X_c \cdot v_1 = \begin{bmatrix} -3,6414 \\ 0,2138 \\ 3,6414 \\ 2,6414 \\ -0,2138 \\ -2,6414 \end{bmatrix}$

1.3

a/ Yes it would because:
- k-mean minimizes within cluster sum of squares.
- Each data point can only belong to one of k clusters → finite point
- the algo will terminate when L doesn't change

b/

Using point #1 and #2 as cluster centers and Euclidean dist:

Cluster 1: Points 1, 3, 4, 5, 10
Cluster 2: Points 2, 6, 7, 8, 9

Update cluster centers:

new Center 1 $= 1/5 \sum_{i \in C_1} (x_i, y_i) = (3,39 ; 3,03)$

new Center 2 $= (5,68 ; 5,08)$