Name: Janmejay Mohanty

CS 559-B Homework Assignment 4

Solution 1:

Given:

$$A_1 = (2,10)$$

$$A_2 = (2,5)$$

$$A_3 = (8,4)$$

$$B_1 = (5,8)$$

$$C_1 = (1,2)$$

$$C_2 = (4,9)$$

The three cluster centers after the first round of execution.

The three clusters are A_1, B_1 and C_2, so calculating the Euclidean distance between of each point from all the three clusters.

First Iteration:

	A_1	A_2	A_3	B_1	C_1	C_2
Centroid:1(A_1)	0	5	8.49	3.61	8.06	2.24
Centroid:2(B_1)	3.61	4.24	5	0	7.21	1.41
Centroid:3(C_2)	2.24	4.47	6.40	1.41	7.62	0

seed1 = $A_1 = (2,10)$

seed2 = $B_1 = (5,8)$

seed3 = $C_2 = (4,9)$

A_1:

$d(A_1, seed1) = 0$

$d(A_1, seed2) = 3.61 > 0$

$d(A_1, seed3) = 2.24 > 0$

$=> A_1 \in cluster1$

A_2:

$d(A_2, seed1) = 5$

$d(A_2, seed2) = 4.24$

$d(A_2, seed3) = 4.47$

$=> \ A_2 \in cluster2$

A_3:

$d(A_3, seed1) = 8.49$

$d(A_3, seed2) = 5$

$d(A_3, seed3) = 6.40$

$=> \ A_3 \in cluster2$

B_1:

$d(B_1, seed1) = 3.61$

$d(B_1, seed2) = \ 0$

$d(B_1, seed3) = \ 1.41$

$=> \ B_1 \in cluster2$

C_1:

$d(C_1, seed1) = 8.06$

$d(C_1, seed2) = \ 7.21$

$d(C_1, seed3) = \ 7.62$

$=> \ C_1 \in cluster2$

C_2:

$d(C_2, seed1) = 2.24$

$d(C_2, seed2) = \ 1.41$

$d(C_2, seed3) = \ 0$

$=> \ C_2 \in cluster3$

1)

The 3 clusters with cluster points are:

Cluster 1 = $\{A_1(2{,}10)\}$

Cluster 2 = $\{A_2(2{,}5),\ A_3(8{,}4),\ B_1(5{,}8),\ C_1(1{,}2)\}$

Cluster 3 = $\{C_2(4{,}9)\}$

2)

Calculating the center ($centroids$) after the first round:

Center1 = $(2, 10)$

Center2 = $\left\{\frac{(2+8+5+1)}{4}, \frac{(5+4+8+2)}{4}\right\} = \left(\frac{16}{4}, \frac{19}{4}\right) = (4, 4.75)$

Center3 = $(4, 9)$

Solution 3:

1)

$$X = (Gender = M, Car\ Type = Family, Shirt\ Size = Large)$$

$$P(X|C0) = P(Gender = M|C0) \times P(Car\ Type = Family|C0) \times P(Shirt\ Size = Large|C0)$$

$$= \frac{6}{10} \times \frac{1}{10} \times \frac{2}{10}$$

$$= \frac{12}{1000}$$

$$P(X|C0) = 0.012$$

$$P(X|C1) = P(Gender = M|C1) \times P(Car\ Type = Family|C1) \times P(Shirt\ Size = Large|C1)$$

$$= \frac{4}{10} \times \frac{3}{10} \times \frac{2}{10}$$

$$= \frac{24}{1000}$$

$$P(X|C1) = 0.024$$

Since $P(X|C1) > P(X|C0)$, the new test example is classified as C1

2)

$$x_1 = Gender, x_2 = Car\ Type, x_3 = ShirtSize$$

$$P\left(C|x_{1,}x_2, x_3\right) = \frac{P(x_1|C)P(x_2|C, x_1)P(x_3|C, x_1, x_2)P(C)}{P(x_1, x_2, x_3)}$$

$$P(x_1 = M, x_2 = Family, x_3 = Large|C0)P(C0) = \left[\frac{6}{10}, \frac{1}{6}, 0\right] \times \frac{1}{2} = 0$$

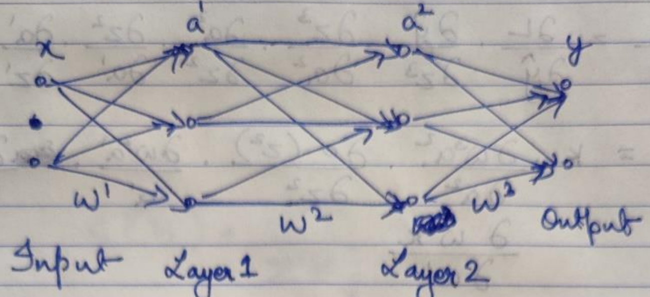
$$P(x_1 = M, x_2 = Family, x_3 = Large|C1)P(C1) = \left[\frac{4}{10}, \frac{3}{4}, \frac{1}{10}\right] \times \frac{1}{2} = 0.06$$

=> C1

Solution 4:

1

Solution4



$$X \to \underbrace{w^1 x}_{z^1}\ \underbrace{\sigma(z^1)}_{a^1} \to \underbrace{w^2 a^1}_{z^2}\ \sigma(z^2) \to \underbrace{w^3 a^2}_{z^3}\ \sigma(z^3)$$

$$\to L(\hat{y}, y)$$

$$\sigma(z) = \frac{1}{1+e^{-z}} \Rightarrow \sigma'(z) = \sigma(z)(1-\sigma(z))$$

(1) $\frac{\partial L}{\partial w^3} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z^3} \cdot \frac{\partial z^3}{\partial w^3}$ as per now take $\frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z^3} = K$

$$= K \cdot \frac{\partial w^3 a^2}{\partial w^3}$$

$$= K a^2$$

(2) $\to \frac{\partial L}{\partial w^2} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z^3} \cdot \frac{\partial z^3}{\partial a^2} \cdot \frac{\partial a^2}{\partial z^2} \cdot \frac{\partial z^2}{\partial w^2}$

$$= K \frac{\partial w^3 a^2}{\partial a^2} \cdot \frac{\partial \sigma(z^2)}{\partial z^2} \cdot \frac{\partial (w^2 a^1)}{\partial w^2}$$

$$= K w^3 \cdot \sigma(w^2 a^1)(1 - \sigma(w^2 a^1)) \cdot a^1$$

②

③ $\frac{\partial L}{\partial w} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z^3} \cdot \frac{\partial z^3}{\partial a^2} \cdot \frac{\partial a^2}{\partial z^2} \cdot \frac{\partial z^2}{\partial a'} \cdot \frac{\partial a'}{\partial z'} \cdot \frac{\partial z'}{\partial w}$

$= K \cdot \frac{\partial w^3 a^2}{\partial a^2} \cdot \frac{\partial \sigma(z^2)}{\partial z^2} \cdot \frac{\partial w^2 a'}{a'} \cdot \frac{\partial \sigma(z')}{\partial z'} \cdot \frac{\partial w' x}{\partial w'}$

$= K \cdot w^3 \cdot \sigma(w^2 a')(1 - \sigma(w^2 a')) \cdot w^2 + (wx)(1 - \sigma(w'x)) \cdot X$

$\frac{\partial L}{\partial w^3} = \frac{\partial L}{\partial w^3} + \frac{\partial L}{\partial w_2^3}$

for $\frac{\partial L_1}{\partial w_1^3} \Rightarrow \frac{\partial L_1}{\partial \hat{y}_1} \cdot \frac{\partial \hat{y}}{\partial z_1^3} \cdot \frac{\partial z_1^3}{\partial w_1}$

$= \frac{\partial (y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)}{\partial \hat{y}_1} \times \frac{\partial s(z_1)^3}{\partial z_1^3} \cdot \frac{\partial w^3 a^2}{\partial w'}$

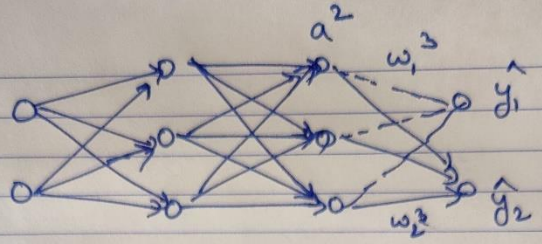
$= 2(\hat{y}_1 - y_1) s(z_1^3)(1 - s(z_1^3)) a^2$

$= 2(\hat{y}_1 - y_1) \cdot \frac{e^{z_1^3}}{e^{z_1^3} + e^{z_2^3}} \cdot \frac{e^{z_2^3}}{e^{z_1^3} + e^{z_2^3}} \cdot a^2$

$\frac{\partial L_1}{\partial w_1} = 2(\hat{y}_1 - y_1) \cdot \frac{e^{z_1^3 + z_2^3}}{(e^{z_1^3} + e^{z_2^3})^2} a^2$

Similarly, $\frac{\partial L_2}{\partial w_2} = 2(\hat{y}_2 - y_2) \cdot \frac{e^{z_1^3 + z_2^3}}{(e^{z_1^3} + e^{z_2^3})} \cdot a^2$

③



$$S(z_i) = \frac{e^{z_i}}{\sum e^{z_j}}$$

$$S'(z_i) = S(z_i)(1 - S(z_i))$$

$$\frac{\partial L}{\partial w} = \frac{\partial L_1}{\partial w_1} + \frac{\partial L_2}{\partial w_2} = \frac{2(\hat{y}_1 - y_1 + \hat{y}_2 - y_1)}{\hat{y} - y} \frac{e^{z_1^3 + z_2^3}}{(e^{z_1^3} + e^{z_2^3})^2} a^2$$

$$\therefore \frac{\partial L}{\partial w^3} = 2(y^2 - y_1) \cdot \frac{e^{z_1^3} \times e^{z_2^3}}{(e^{z_1^3} + e^{z_2^3})} \cdot a^2$$

$$\Rightarrow k = 2(\hat{y} - y_1) \frac{e^{z_1^3 + z_2^3}}{(e^{z_1^3} + e^{z_2^3})}$$

Substitute k in equations ② & ③

$$\frac{\partial L}{\partial w^2} = 2(\hat{y} - y) \cdot \frac{e^{z_1^3} \times e^{z_2^3}}{(e^{z_1^3} + e^{z_2^3})} \cdot w^3 \sigma(w^2 a^1) . (1 - \sigma(w^2 a^1)) a_1$$

$$\frac{\partial L}{\partial w_1} = 2(\hat{y} - y) \cdot \frac{e^{z_1^3} \times e^{z_2^3}}{(e^{z_1^3} + e^{z_2^3})} \cdot w^3 . \sigma(w^2 a^1) . (1 - \sigma(w^2 a^1)) . w^2 \sigma(w^1 x)(1 - \sigma)(w^1 x) .$$

Solution 2:

```
import numpy as np

import pandas as pd

from scipy.stats import multivariate_normal

from scipy.stats import mode

from sklearn.metrics import confusion_matrix

import matplotlib.pyplot as plt

X = pd.read_csv('points.dat.txt', sep = " ", header=None)

#Reference from http://www.oranlooney.com/post/ml-from-scratch-part-5-gmm/

class GMM:

    def __init__(self, k, max_iter=5):

        self.k = k

        self.max_iter = int(max_iter)


    def initialize(self, X):

        self.shape = X.shape

        self.n, self.m = self.shape


        self.phi = np.full(shape=self.k, fill_value=1/self.k)

        self.weights = np.full( shape=self.shape, fill_value=1/self.k)


        random_row = np.random.randint(low=0, high=self.n, size=self.k)

        self.mu = [  X[row_index,:] for row_index in random_row ]

        self.sigma = [ np.cov(X.T) for _ in range(self.k) ]


    def e_step(self, X):

        self.weights = self.predict_proba(X)

        self.phi = self.weights.mean(axis=0)


    def m_step(self, X):

        for i in range(self.k):
```

```
            weight = self.weights[:, [i]]
            total_weight = weight.sum()
            self.mu[i] = (X * weight).sum(axis=0) / total_weight
            self.sigma[i] = np.cov(X.T,
                aweights=(weight/total_weight).flatten(),
                bias=True)

    def fit(self, X):
        self.initialize(X)

        for iteration in range(self.max_iter):
            self.e_step(X)
            self.m_step(X)

    def predict_proba(self, X):
        likelihood = np.zeros( (self.n, self.k) )
        for i in range(self.k):
            distribution = multivariate_normal(
                mean=self.mu[i],
                cov=self.sigma[i])
            likelihood[:,i] = distribution.pdf(X)

        numerator = likelihood * self.phi
        denominator = numerator.sum(axis=1)[:, np.newaxis]
        weights = numerator / denominator
        return weights

    def predict(self, X):
        weights = self.predict_proba(X)
        return np.argmax(weights, axis=1)
gmm = GMM(10,5)
```

gm = gmm.fit(X)

```
[-6.3670924  -0.35240556]
[[2.  0.5]
 [0.5 0.5]
 [0.5 0.5]
 ...
 [0.5 0.5]
 [0.5 0.5]
 [0.5 0.5]]
pi:  [0.5 0.5]
mu:  [[-0.84275947 -1.14251164]
 [-0.79945878 -0.05602496]]
epsilon:  [[[14.92980262  1.07714443]
  [ 1.14705169  7.14493712]]

 [[ 5.55922667  1.58282789]
  [ 1.68555429  6.5471054 ]]]
```