**Introduction to Artificial Intelligence - Homework 5**

NE6114011 人工智慧所碩一 楊雲翔

**1 程式執⾏環境及說明**

* Python版本：3.8.15
* 相依套件與版本

1. numpy== 1.23.5
2. matplotlib== 3.6.2
3. scikit-learn ==1.1.3

**2 GMM**

本次作業的目標是要實現高斯混合模型(Gaussian Mixture Model, GMM)，並利用GMM進行資料的分群。我採用了sklearn(scikit-learn)套件所提供的GaussianMixture進行實作，並根據作業檔案提供的input.json來進行實驗，在該檔案中，共有三組不同的資料集，分別為(x1, y1)、(x2, y2)、(x3. y3)，其中x均為二維資料x=()。

**實驗**

GMM中有著許多的不同參數可以調整，我藉由計算不同參數組合下分群的準確度，來探討參數對於結果的好壞

* **Covariance type and Init params**
* Covariance type代表GMM中共變異數的種類，共有四種選項：

1. full: 每個群都有一個一般的共變異數矩陣
2. tied：各個群共享一個共變異數矩陣
3. diag:每個群都有一個對稱共變異數矩陣
4. spherical: 每個群只有一個共變異數

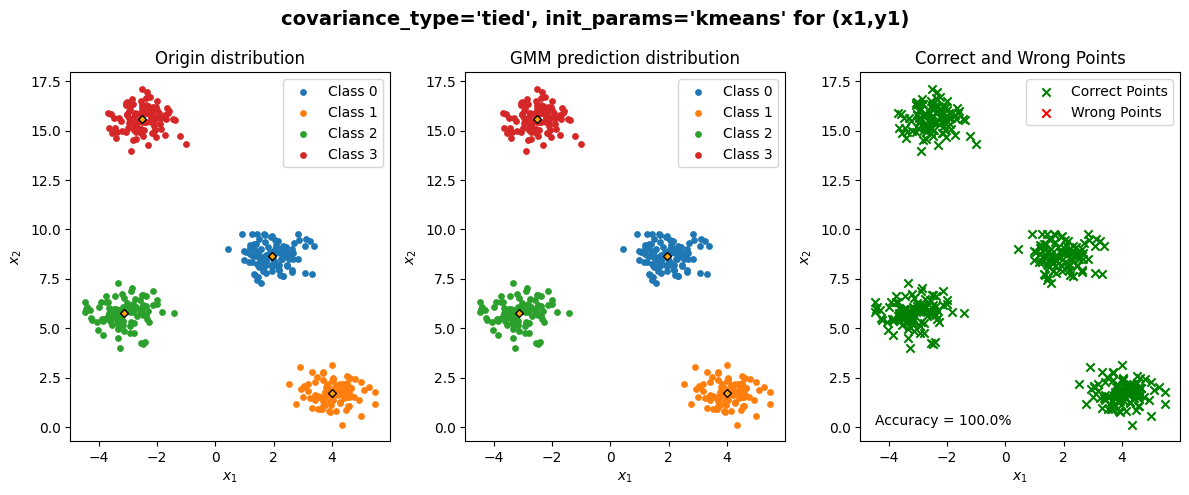
* Init params表初始化weight、mean與precision的方法，共有四種選項：

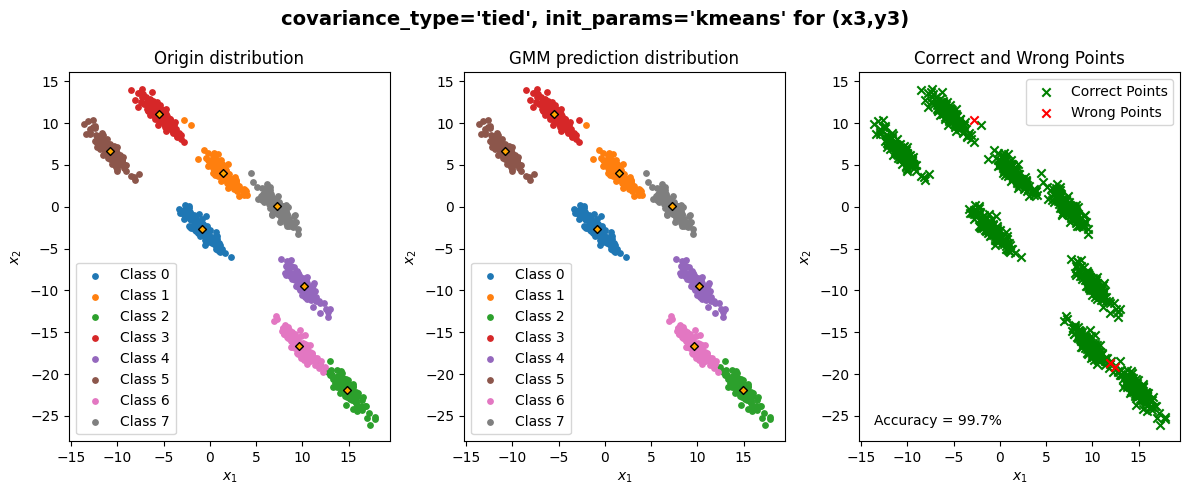
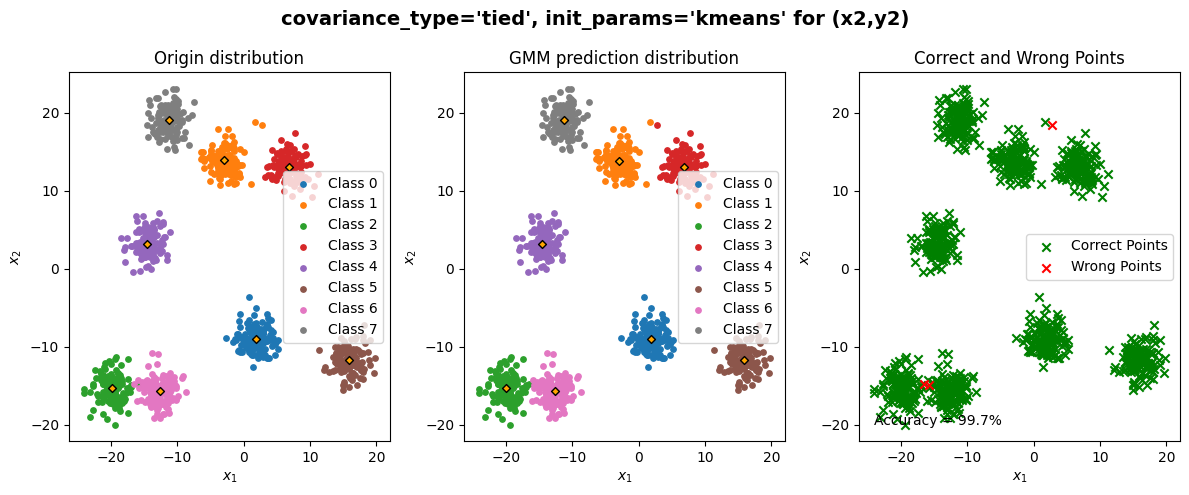
1. Kmeans: 採用kmeans
2. Kmeans++: 採用kmeans++
3. Random: 隨機初始化
4. Random\_from\_data: 隨機從資料中選出

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Covariance Type | Init Params | (x1,y1) Accuracy(%) | (x2,y2) Accuracy(%) | (x3,y3) Accuracy(%) | Average Accuracy(%) |
| full | kmeans | 100 | 99.6 | 99.6 | 99.7 |
| k-means++ | 100 | 99.6 | 99.6 | 99.7 |
| random | 100 | 12.5 | 12.5 | 41.7 |
| random\_from\_data | 54.5 | 71.2 | 87.5 | 71.1 |
| tied | kmeans | 100 | 99.7 | 99.7 | 99.8 |
| k-means++ | 25.0 | 12.5 | 12.5 | 16.7 |
| random | 25.0 | 25.0 | 12.5 | 20.8 |
| random\_from\_data | 25.0 | 12.5 | 12.5 | 16.7 |
| diag | kmeans | 100 | 99.7 | 99.3 | 99.7 |
| k-means++ | 100 | 99.7 | 99.3 | 99.7 |
| random | 25.0 | 12.5 | 12.5 | 16.7 |
| random\_from\_data | 73.0 | 79.1 | 62 | 71.4 |
| spherical | kmeans | 100 | 82.2 | 99.2 | 93.8 |
| k-means++ | 100 | 99.7 | 99.2 | 99.6 |
| random | 25.0 | 25.0 | 12.5 | 20.8 |
| random\_from\_data | 51.2 | 62.9 | 55.4 | 56.5 |

根據上述實驗結果，當covariance type為tied且init params為kmeans時，在三個資料集都有著最高的準確度

* covariance type = ‘tied’, init params=’kmeans’分群結果：

****

****

透過觀察(x2, y2)與(x3, y3)的分群分類錯誤的結果可以發現，GMM分類錯誤的樣本可分為兩類：

1. 樣本點位於兩群交界處的樣本點，舉例來說，在(x2, y2)左下方及(x3, y3)右下方的分類錯誤樣本
2. 樣本點屬於離群值或異常值，舉例來說，在(x3, y3)左上方，觀察原始分布，可以發現距離紅色群非常相近的兩點被歸類為橘色群，