

Ajinkya Kishor Vadane

1001657016

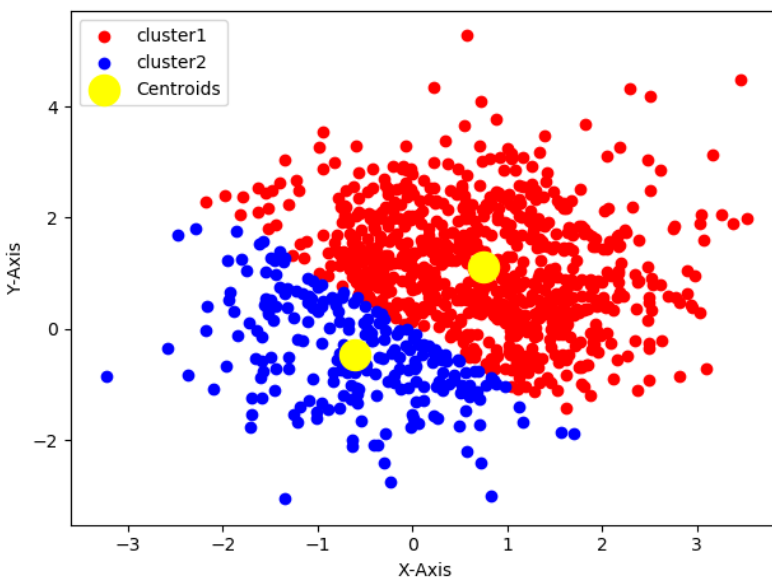
## 1) KMeans

- Firstly, I have used `np.random.multivariate_normal` function for calculating random data
- I am taking input for Number of Clusters from user as Enter the Number of Clusters.
- For each training example compute the Euclidean distance from the centroid and assign the cluster based on the minimal distance.
- Then the mean of the clusters are taken to find new centroids.
- All steps repeated until previous centroids not equal to current centroids.

2) Apply your code to the data generated above with  $k = 2$  and initial centers  $c_1 = (10; 10)$  and

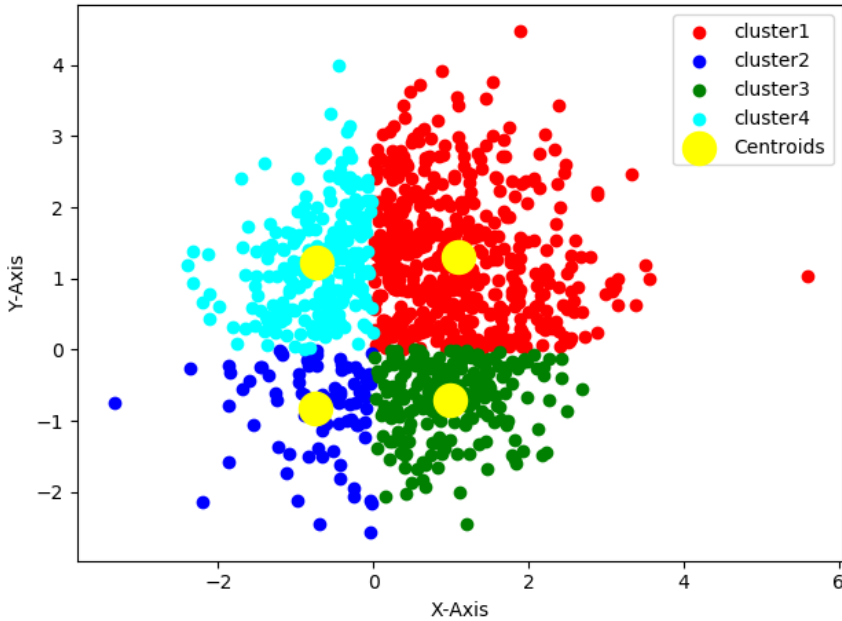
$c_2 = (-10; -10)$ .

Graph :-



Iterations it took to reach final output :-

3) Apply your code to the data generated above with  $k = 4$  and initial centers  $c_1 = (10; 10)$  and  $c_2 = (-10; -10)$ ,  $c_3 = (10; -10)$  and  $c_4 = (-10; 10)$ .



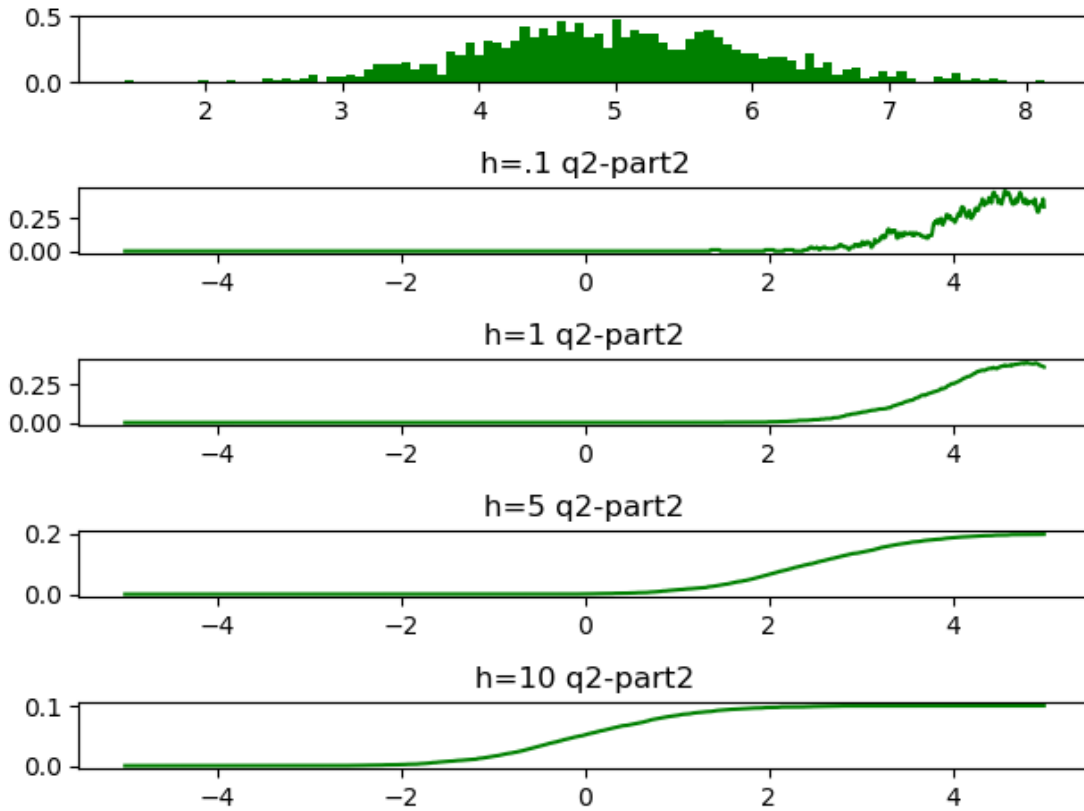
Number of Iteration :-

Question 2)

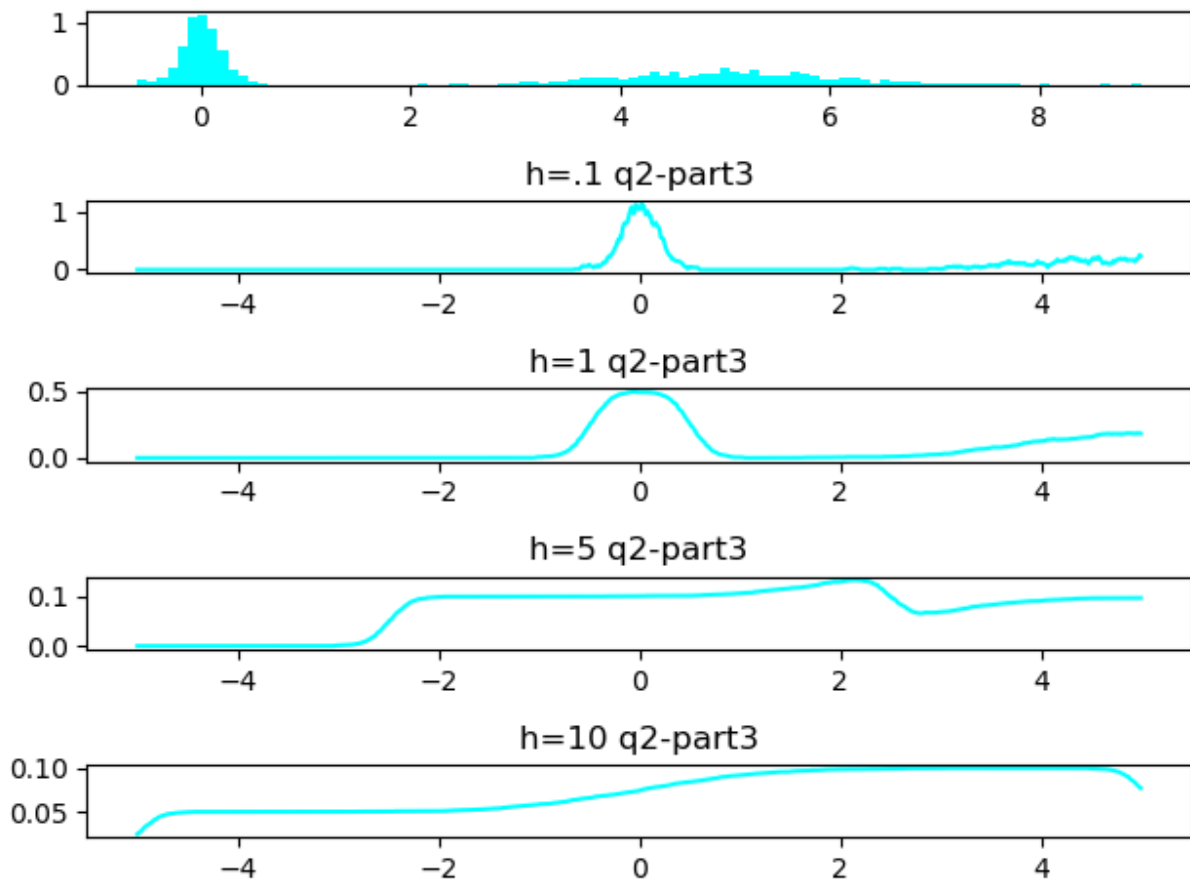
- 1) I have generated the data using function `np.random.multivariate_normal()`
- 2) By referring to the algorithm presented in the sides I am calculating probability for each data point.
- 3) For 1-D I am taking  $d = 1$
- 4) For 2-D I am taking  $d = 2$

Graph :-

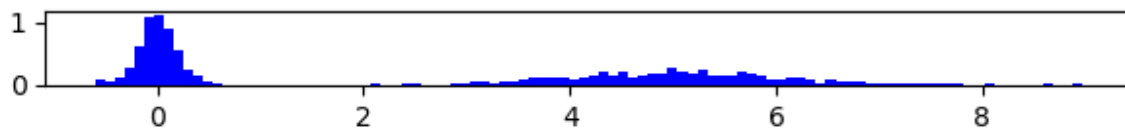
2) Generate  $N = 1000$  Gaussian random data with  $\sigma_1 = 5$  and  $\sigma_2 = 1$ . Test your function `mykde` on this data with  $h = \{.1, 1, 5, 10\}$



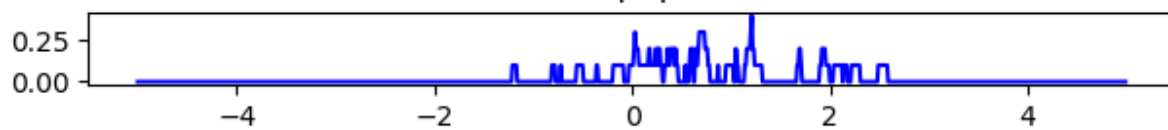
3) Generate  $N = 1000$  Gaussian random data with  $m_1 = 5$  and  $\text{sig}_1 = 1$  and another Gaussian random data with  $m_2 = 0$  and  $\text{sig}_2 = 0.2$ . Test your function `mykde` on this data with  $h = \{0.1, 1, 5, 10\}$



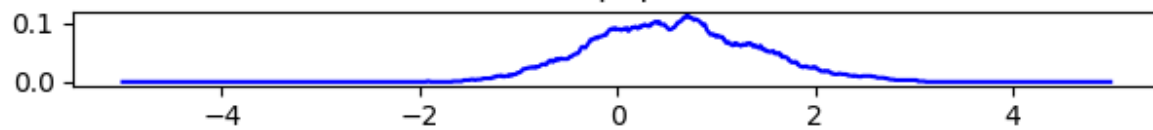
4)For part 4



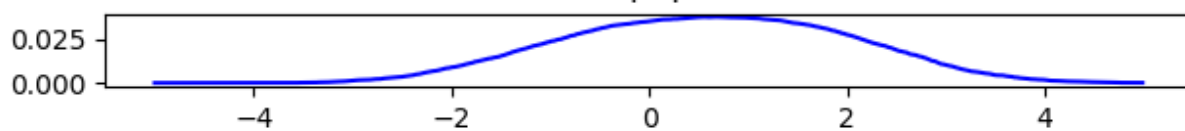
$h=.1$   $q2\text{-part4}$



$h=1$   $q2\text{-part4}$



$h=5$   $q2\text{-part4}$



$h=10$   $q2\text{-part4}$

