**Homework# 6 Clustering**

**0756079 陳冠聞**

1. **Clustering Procedure for k-means/kernel k-means for k = 2, 3, 4**

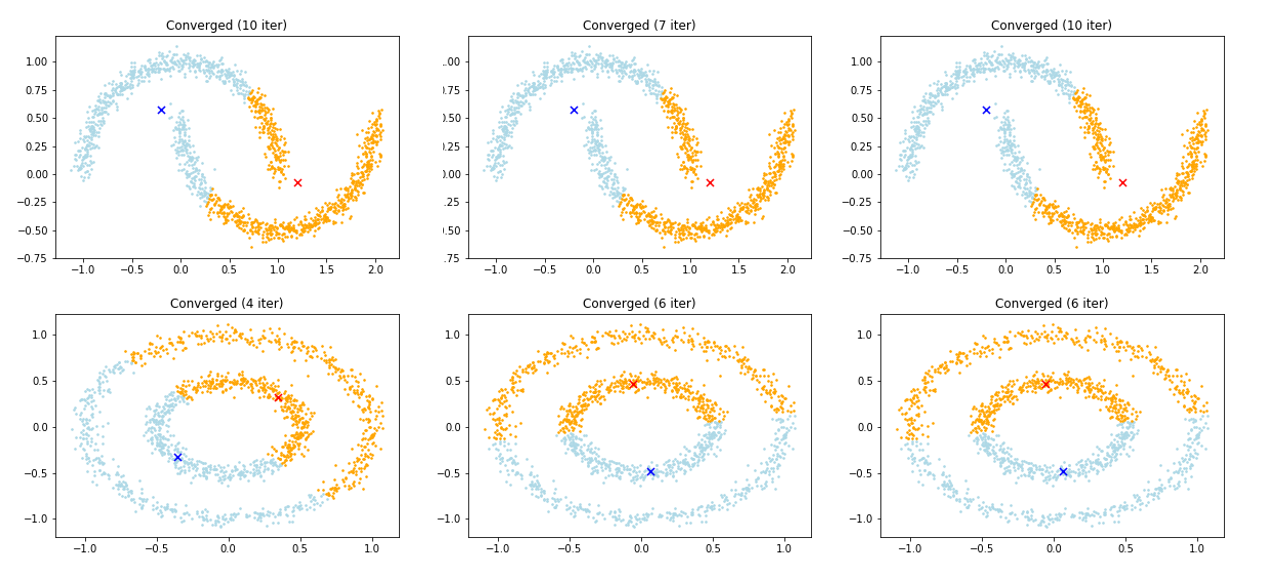
Please see***kmeans.mp4*** & ***kernel-kmeans.mp4*** in detail.

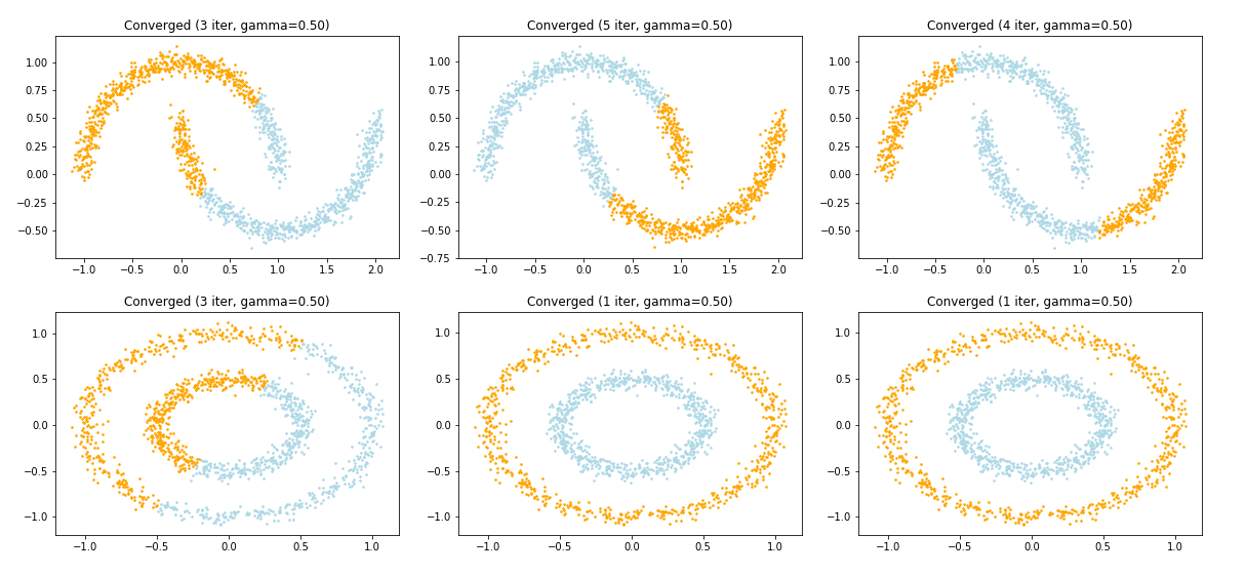
1. **Different initialization of k-means/kernel k-means**

We use three different methods (Random, Distance-To-Origin and Distance-To-Center) for initialization. And we can see that for these two dataset, k-means can not get desired result for any initialization method, and kernel k-means can get desired result of circle dataset if we use Dist2Origin or Dist2Center for initialization.

Following are figures of clustering result for k=2, and from left to right are Random, Distance-To-Origin and Distance-To-Center method respectively.

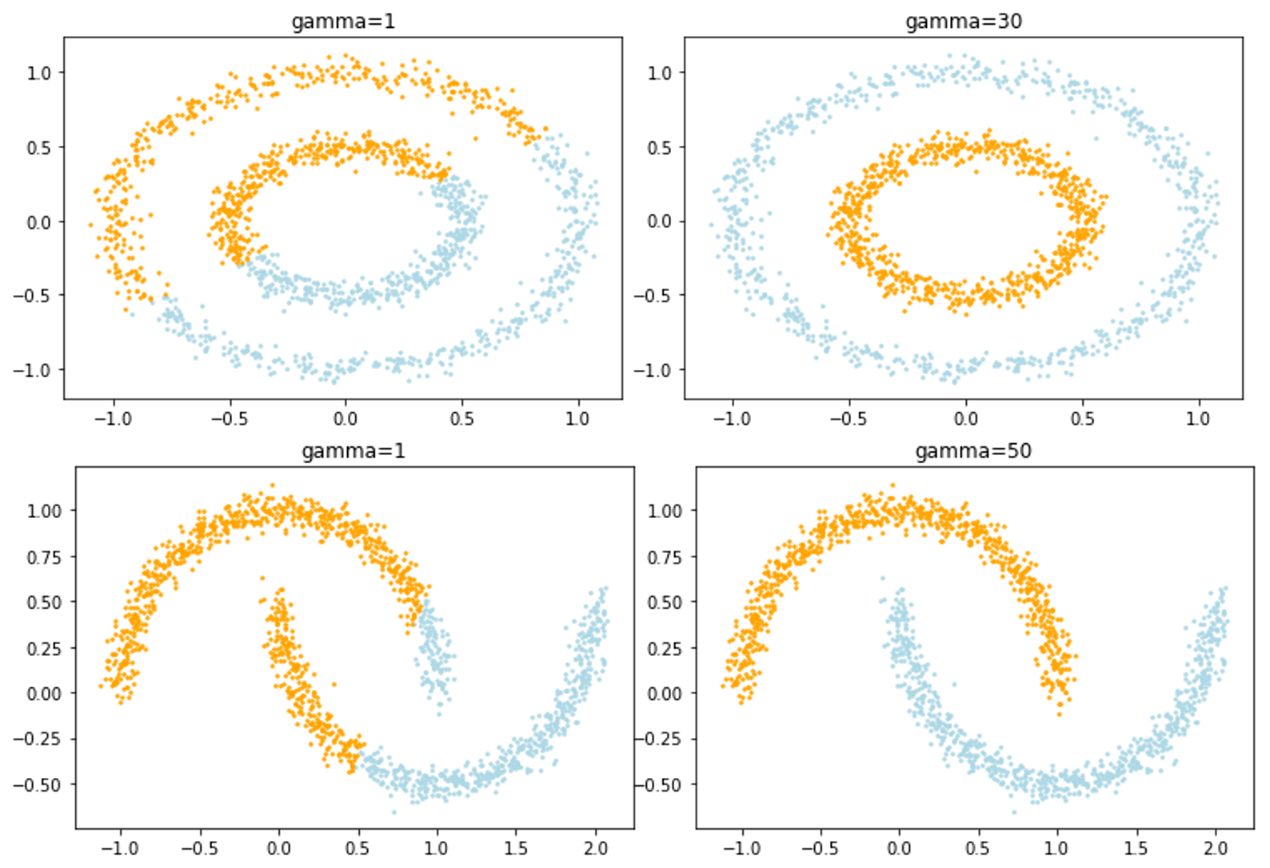
K-means



Kernel K-means

1. **Spectral clustering**

We found that the value of gamma for RBF kernel will effect the result of Spectral clustering. Following are figures of clustering result for Spectral clustering of different gamma.



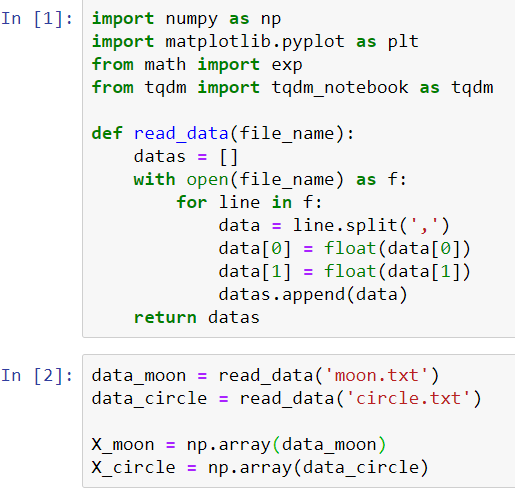
For spectral clustering, you can see if data points within the same cluster do have the same coordinates in the eigenspace of graph Laplacian, discuss in the report.

Because the geometric meaning of spectral clustering is to use Laplacian Eigenmap to do dimension reduction, and then do k-means clustering. Since k-means will assign the membership of a data point by the closest distance to the cluster, the data points which have the same cluster will have similar coordinate.

Reference

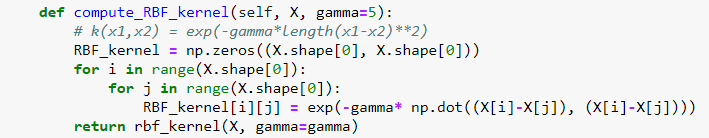
Belkin, M., & Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, *15*(6), 1373-1396.

1. **Code : RBF Kernel K-means**

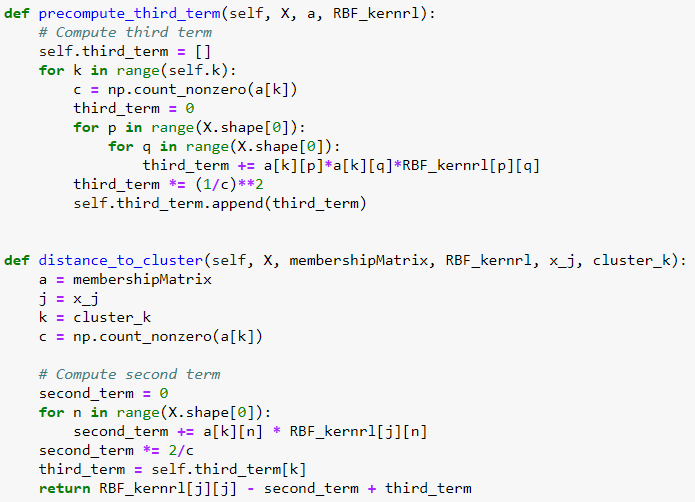
****

Initialization Method

Read Data and convert to matrix form

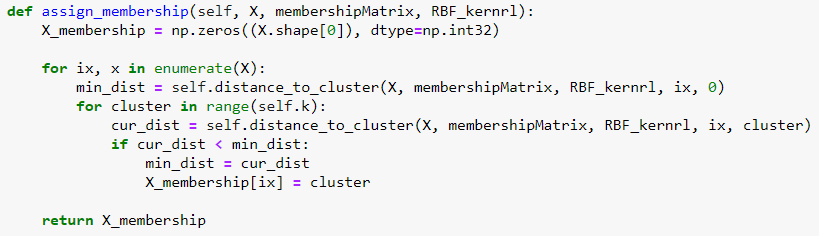
****

Compute RBF kernel

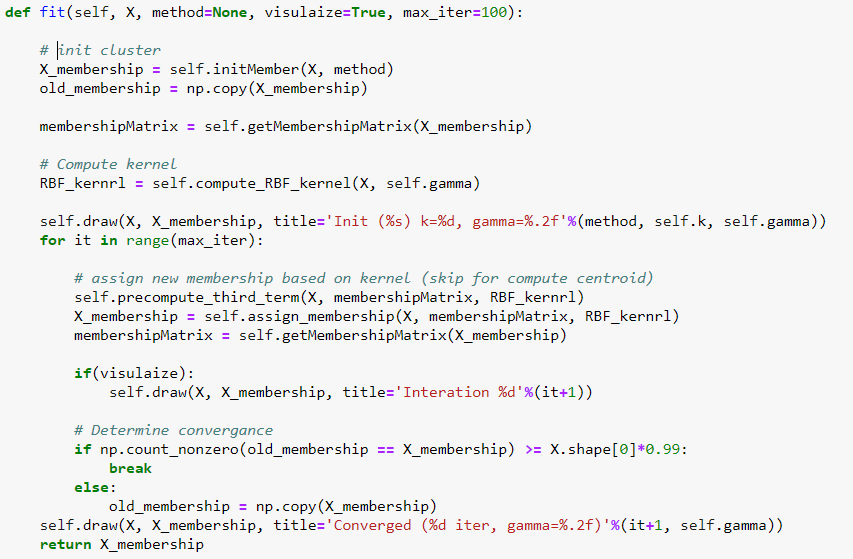
****

Compute Distance of x\_j to cluster\_k

(precompute third term in formula to speed up)

****

Assign membership by min distance to cluster

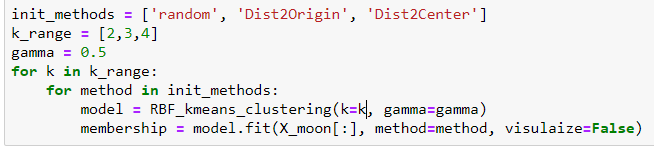
****

Until Converge

Compute new membership

Compute Kernel

Initialize Membership

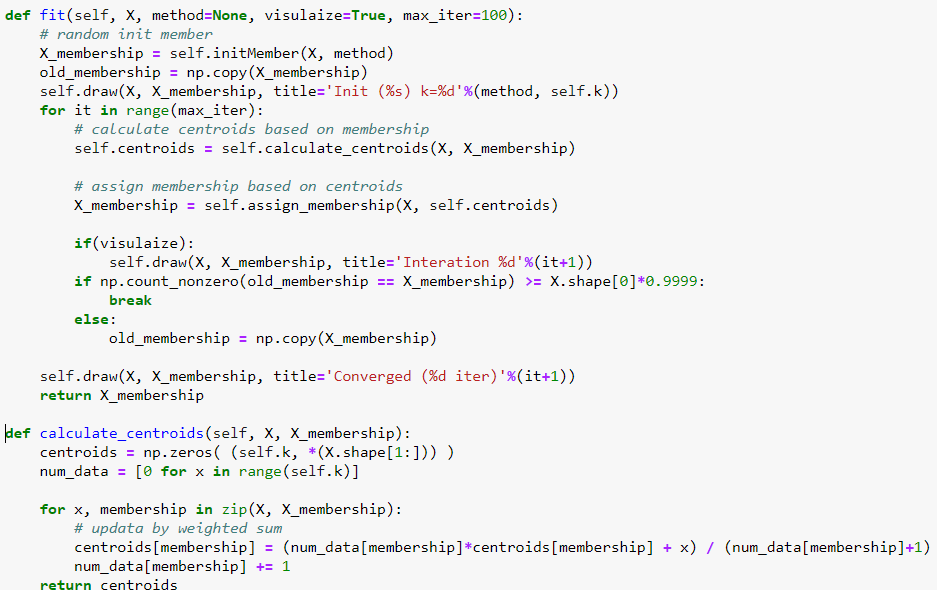
****

Do kernel k-means clustering on

different k and initialization method

1. **Code : K-means**

Only show codes which are different from kernel k-means

****

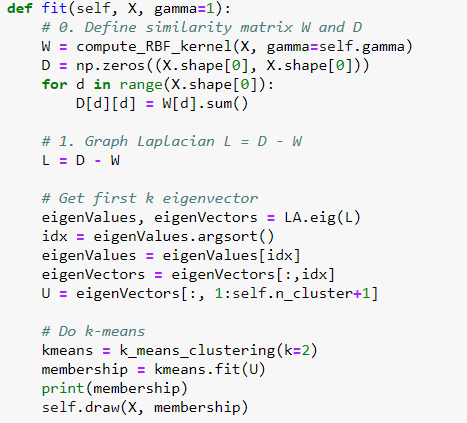
Calculate centroids of a cluster

Update memberships by fixing centroids

Update centroids by fixing memberships

1. **Code : Spectral clustering**

Only show codes which are different from kernel k-means

****

Do k-means

Calculate First K eigenvector

of Graph Laplacian

Calculate Graph Laplacian

For more detail, please see ***k-means\_clustering.py*, *RBF\_k-means\_clustering.py*** and ***spectral\_clustering.py***