In the exercise one, I create 450 sites, and we need to choose the 3 center.

As for the method of create data, I first set three initial center, the blue one, the green one and the red one. And for each initial center, I used the Gaussian distribution method to generate 150 sites. Meantime, we iterator fuzzy c-means algorithm step by step, until the change of distance of three Centers is less than 0.01.

I have tried set all equal to 1, this will lead the center in the same location, and don’t know how to iterator continue. So I select the initial cluster by greedy. So the left picture are the initial partition. The yellow site means the cluster center. And the color of each site we get by the weight of . After ten steps, we get the right one.

There are four picture of fuzzy c-means algorithm with different m deal with the same data. The left above one the m is 1.1, and we use 10 steps ,the right upper one we also use 10 steps to finish the selection of cluster center. The lower two when m is 2.0, we use 13 steps and when m is 3.0, we use 17 steps.

So we find with m increase, our algorithm needs more iterations. Although I don’t know how to prove this. Another difference is the fuzzy of boundary, the larger m, the more fuzzy of the boundary.

The next question is about the different of k-means and fuzzy c-means, I also calculate the number of steps for iteration. The number of step of k-means close to c-means, when m is not large. And the boundary of k-means is clear, because each site will be divide to a distinct cluster, while fuzzy C-means not.