CS 383 - Machine Learning

Assignment 2 - Clustering
Summer 2017
Amir Omidi

1 Programming Questions

1.1 Basic k-Means Clustering

1.1.1 Initial Setup Visualization

The following figure displays the initial data visualization.

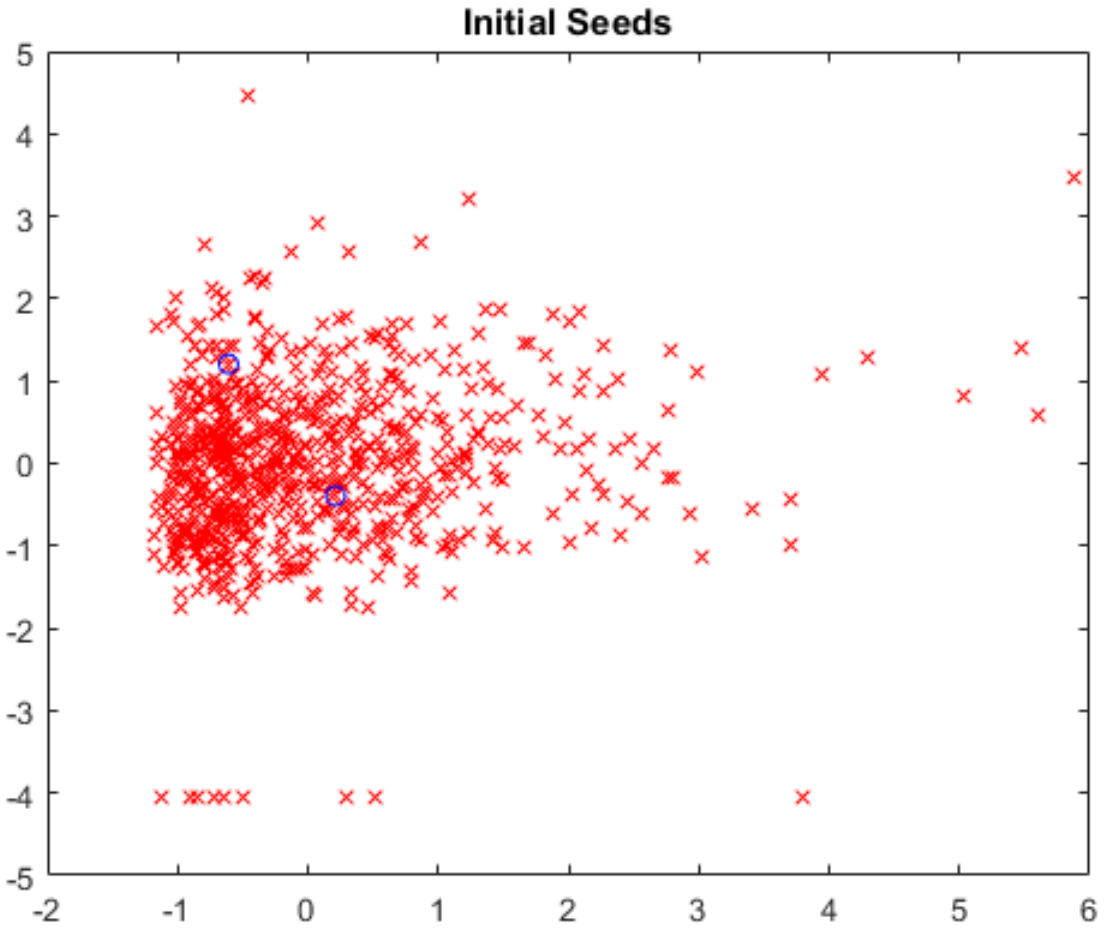

Figure 1: Initial setup visualization

Code: (Yes; the function is called kmeme on purpose :D)

```
otherData = kmeme(newData(:,8:-1:7), 2, 1, 2);
```

1.1.2 Initial Cluster Assignment

The following figure displays the initial clustering of the data.

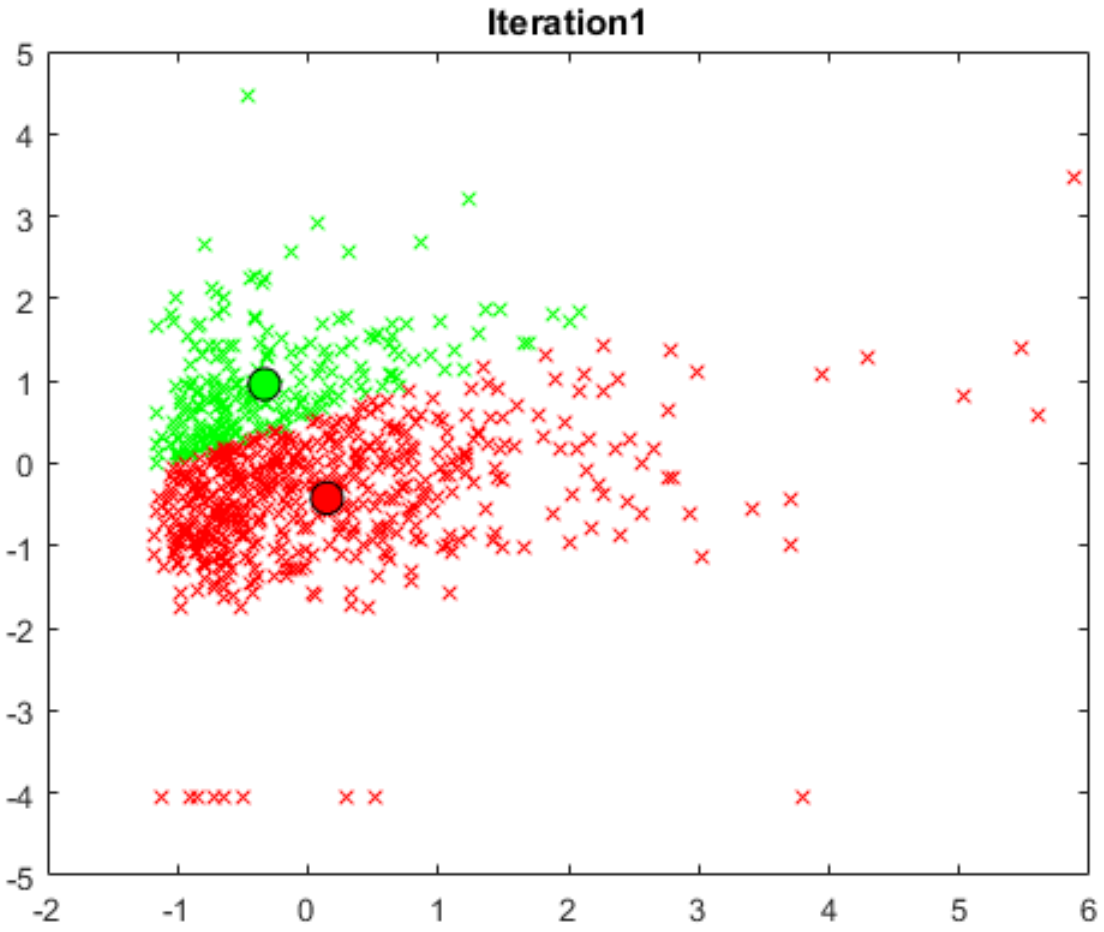

Figure 2: Initial clustering of data

1.1.3 Final Cluster Assignment

The following figure displays the final clustering of the data.

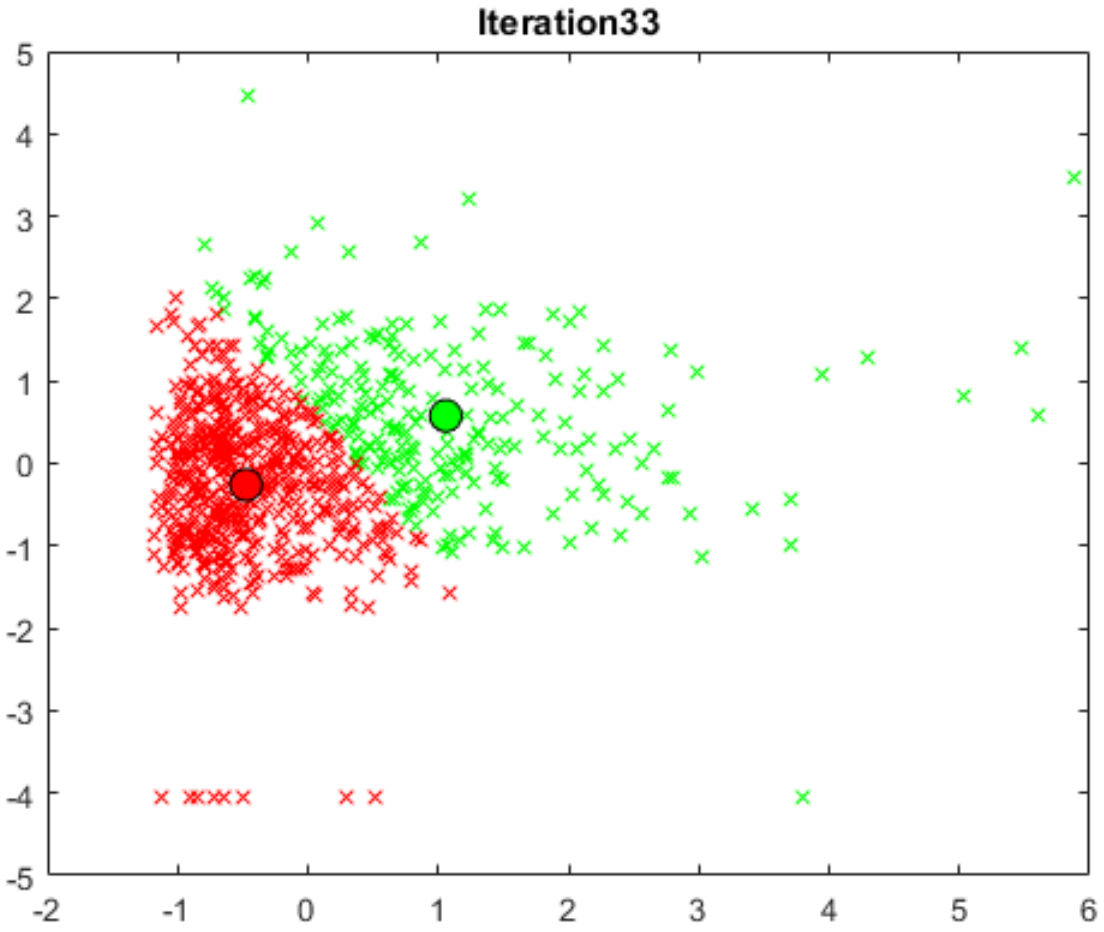

Figure 3: Final clustering of data

1.2 Flexible k-Means Clustering

1.2.1 Sample 1

This sample of data is the clustering of all the data with $k = 2$, but displaying the 6th and 7th features only (The BMI and Pedigree values).

Initial Seeds

The inital seeds are ofcourse going to be the same as section 1.1.1:

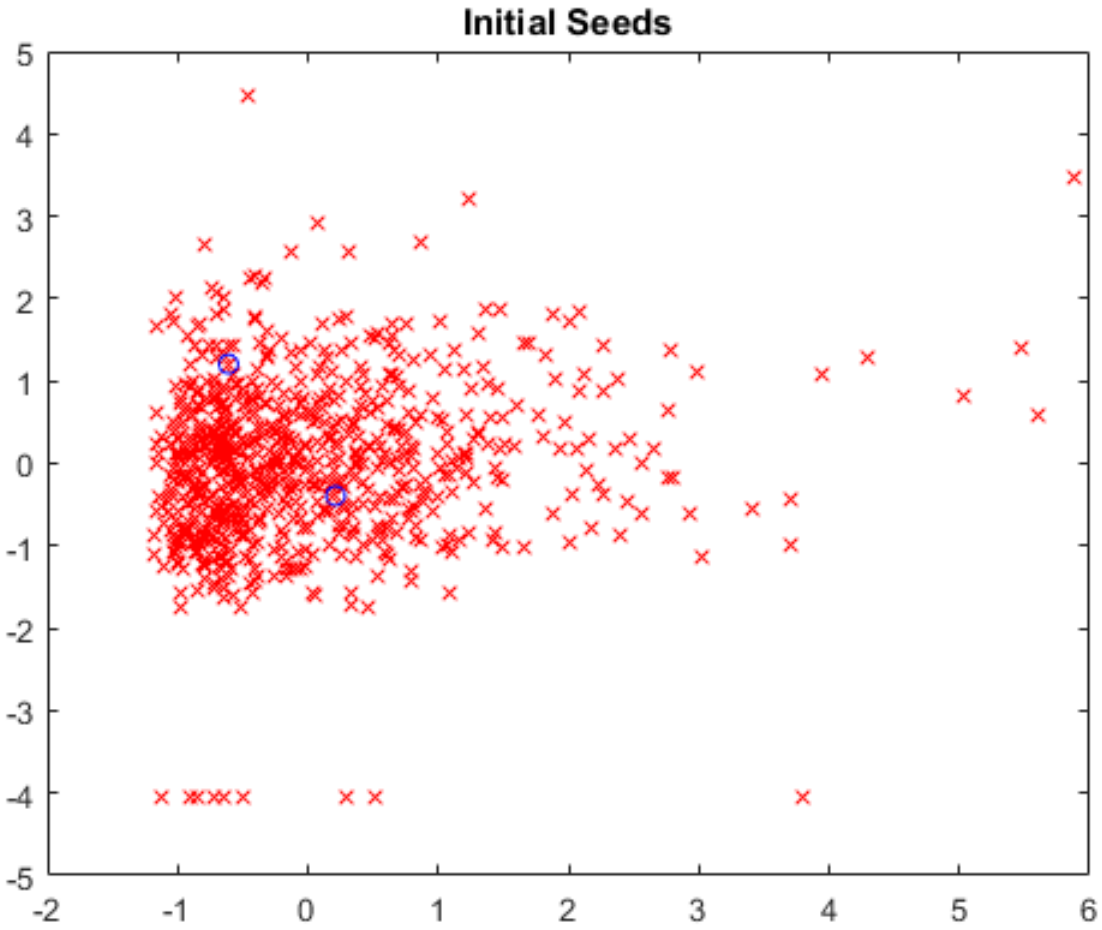

Figure 4: Initial setup visualization

Initial clustering assignments

The inital clustering for this data.

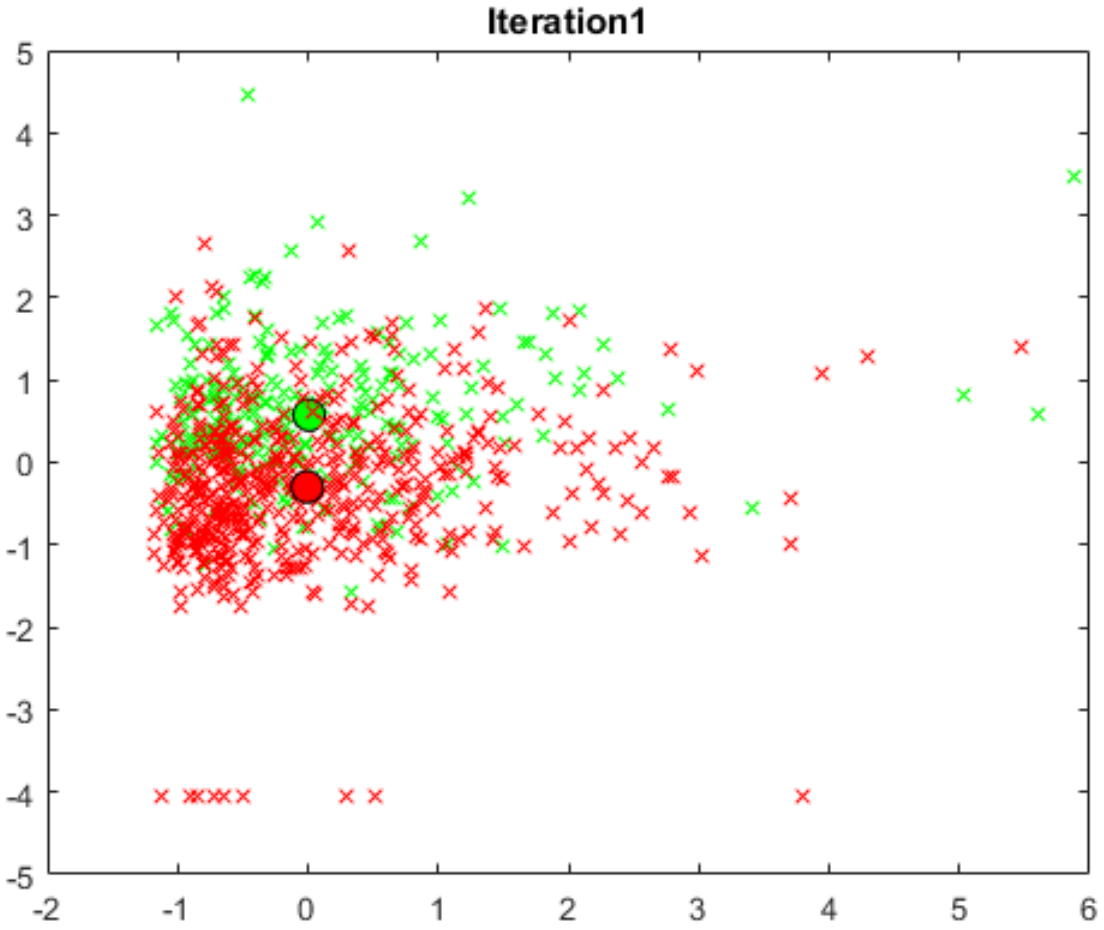

Figure 5: Initial clustering of data

As you can see it looks quite different from 1.1.2.

Final clustering assignments

The final clustering for this data.

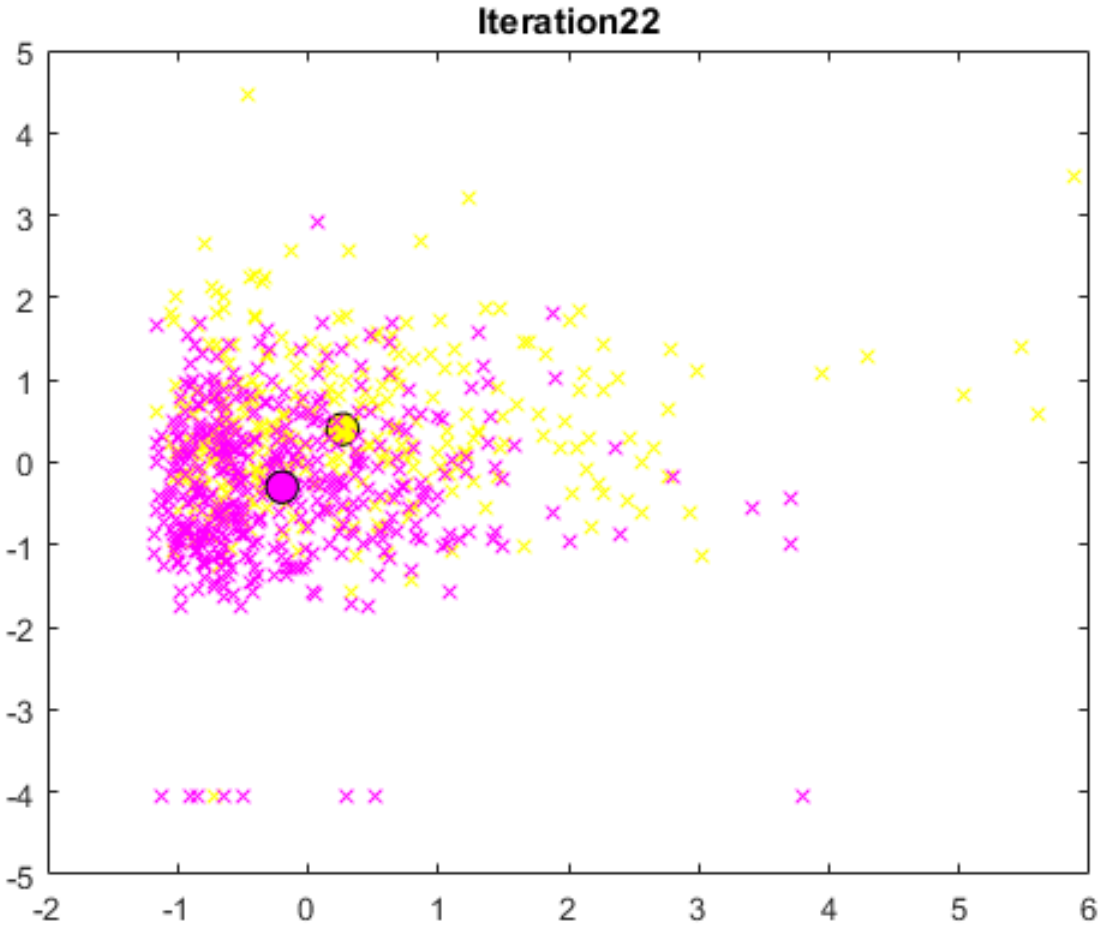

Figure 6: Final clustering of data

1.2.2 Sample 2

This sample of data is the clustering of all the data with $k = 4$, but displaying the 2nd and 3rd features only.
The code to run this:

```
otherData = kmeme(newData, 4, 4, 3)
```

The final clustering for this data.

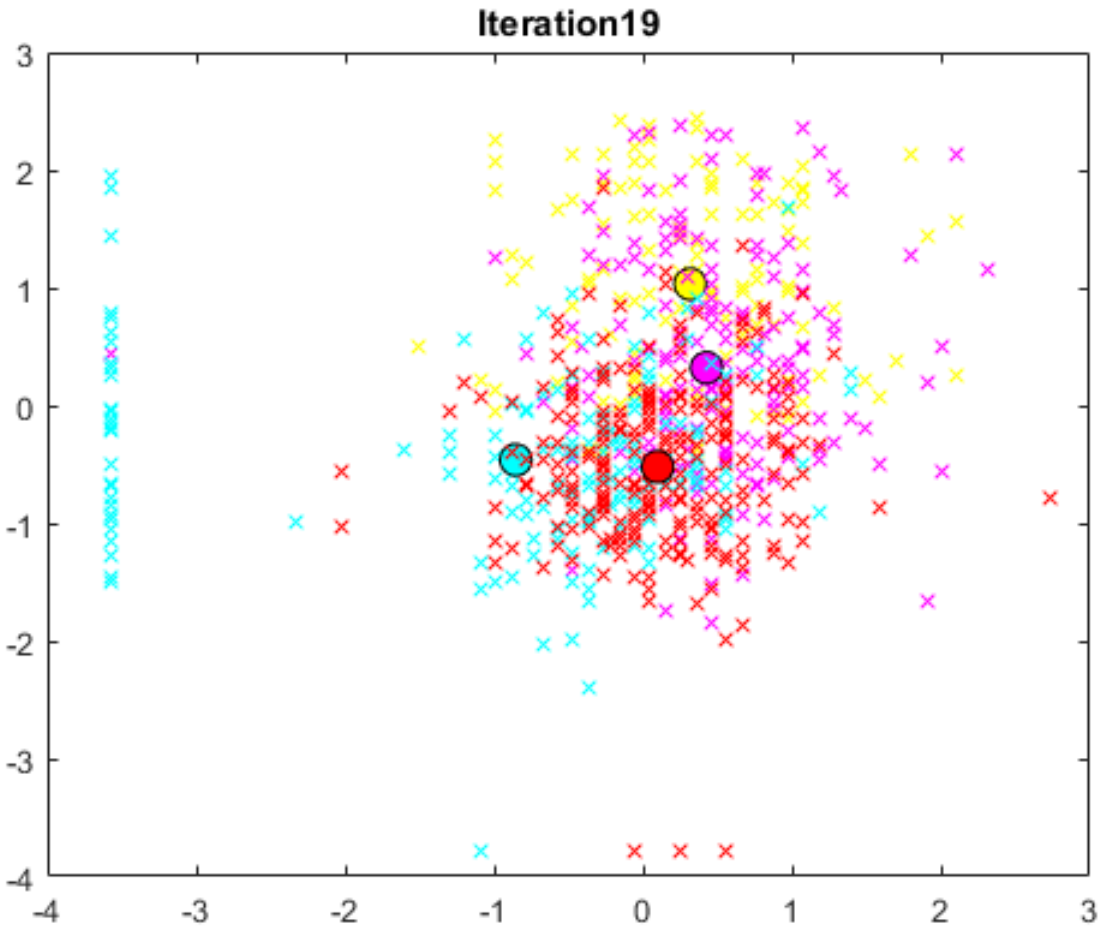

Figure 7: Final clustering of data

1.2.3 Sample 3

This sample of data is the clustering of all the data with $k = 5$, but displaying the 1st and 2nd features only.
The code to run this:

```
otherData = kmeme(newData, 5, 2, 3);
```

The final clustering for this data.

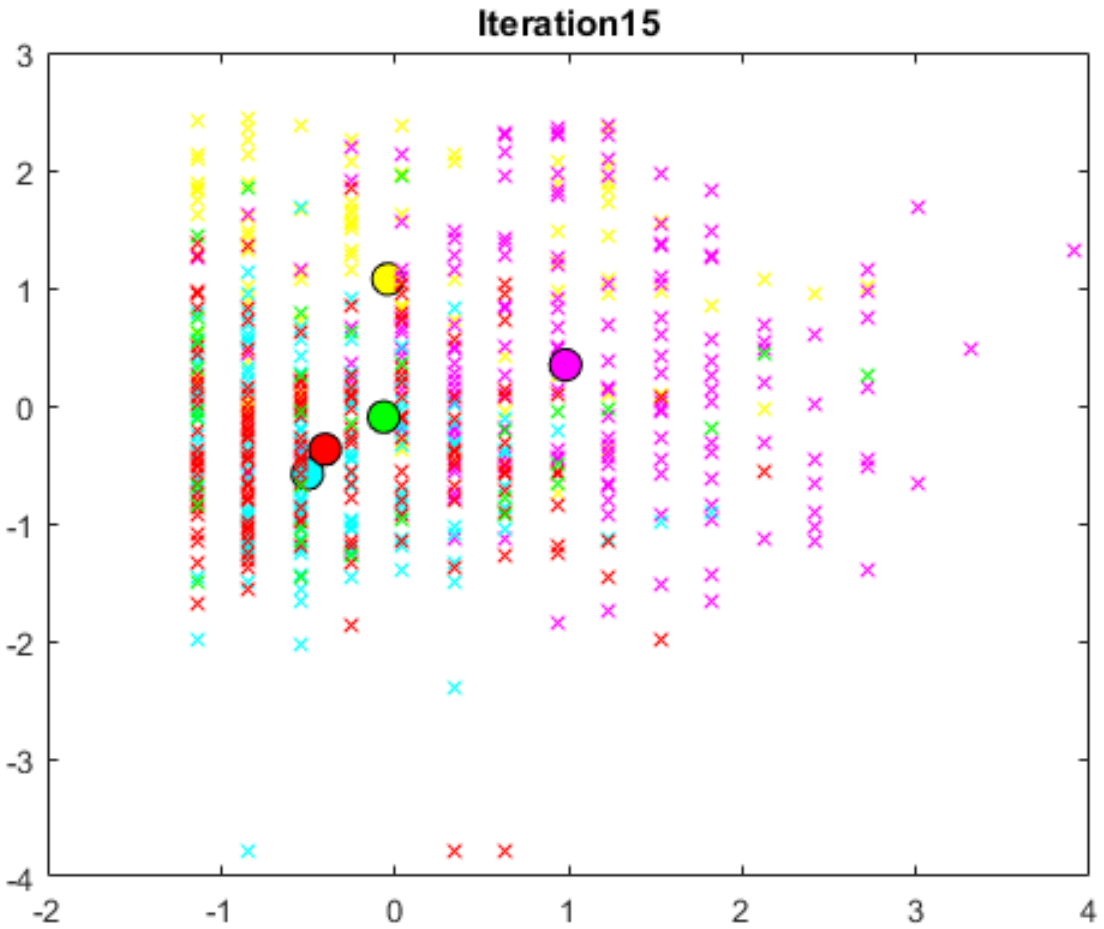

Figure 8: Final clustering of data

1.2.4 Sample 4

This sample of data is the clustering of all the data with $k = 7$, but displaying the 5th and 7th features only.
The code to run this:

```
otherData = kmeme(newData, 7, 8, 6);
```

The final clustering for this data.

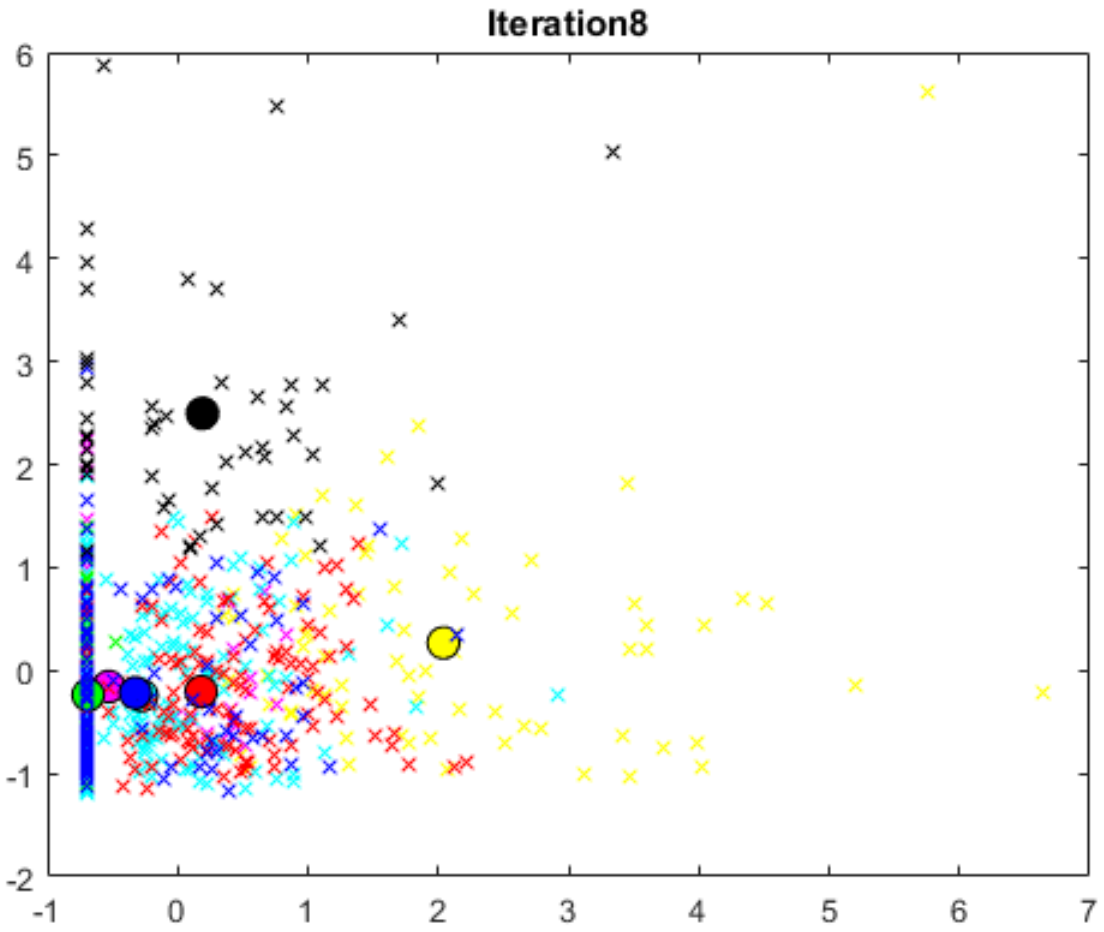

Figure 9: Final clustering of data