

EECS E6892 Fall 2015: Homework #4

Daniel Kronovet - dbk2123@columbia.edu

Discussants: mkt2126, jma2215

December 15, 2015

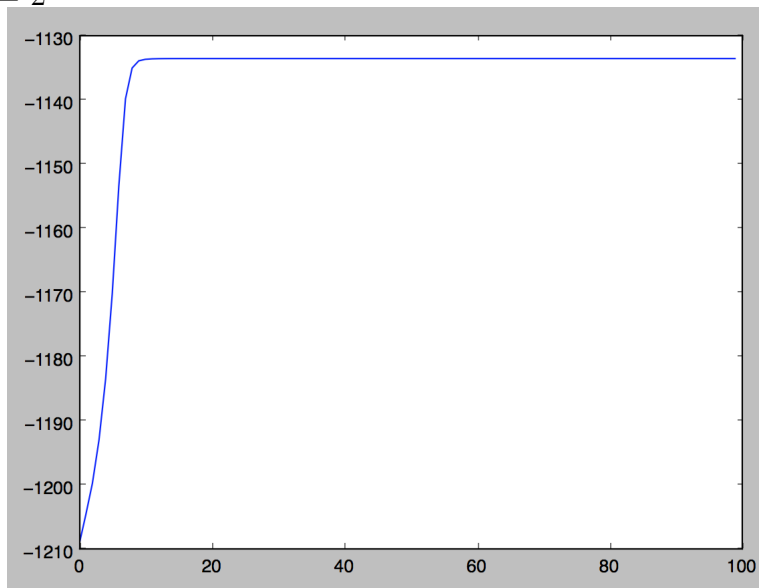
Problem 1

Part A

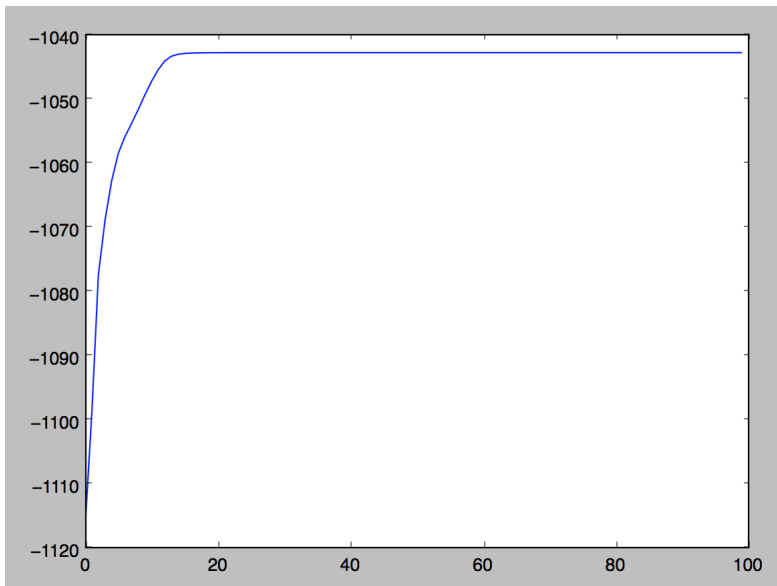
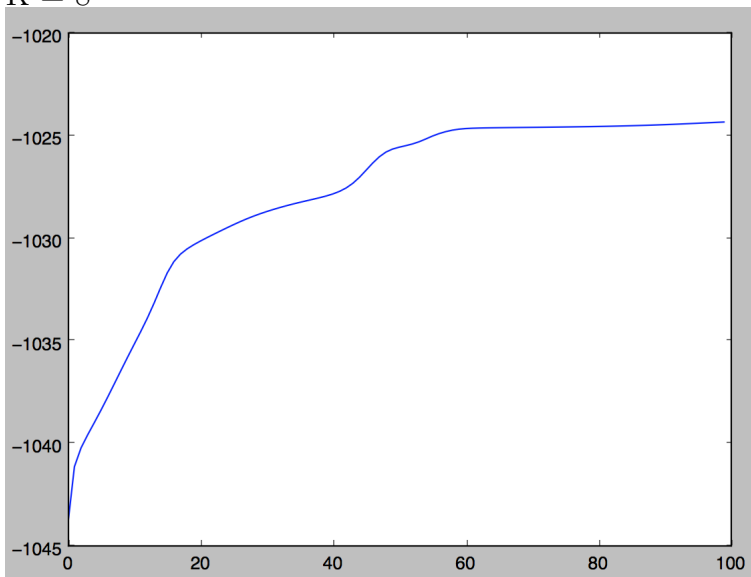
I did this it was fun.

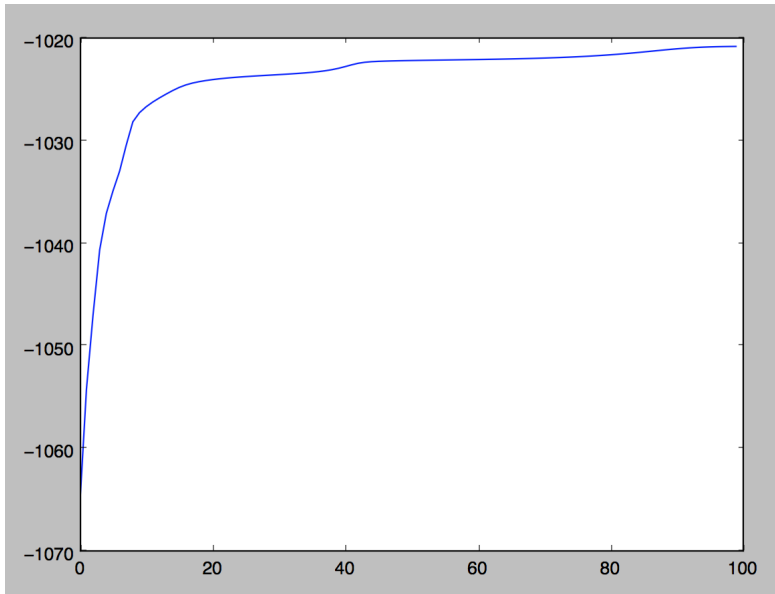
Part B

$K = 2$



$K = 4$

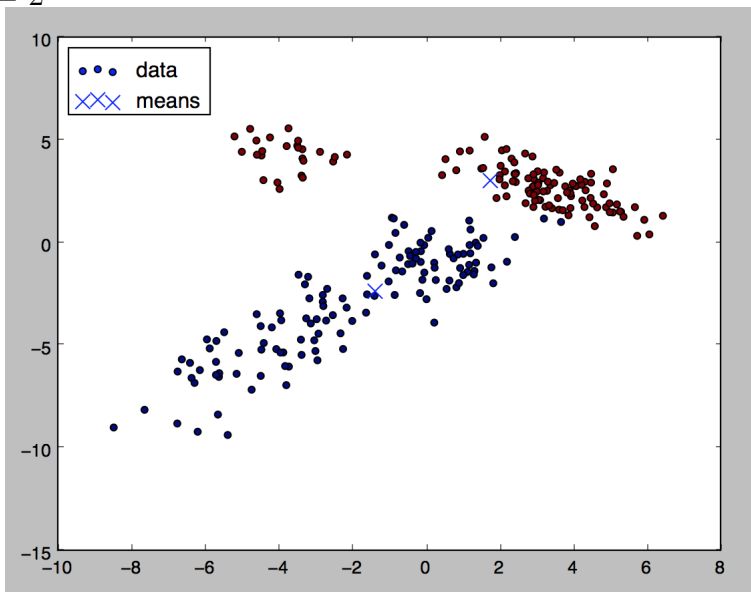
 $K = 8$  $K = 10$



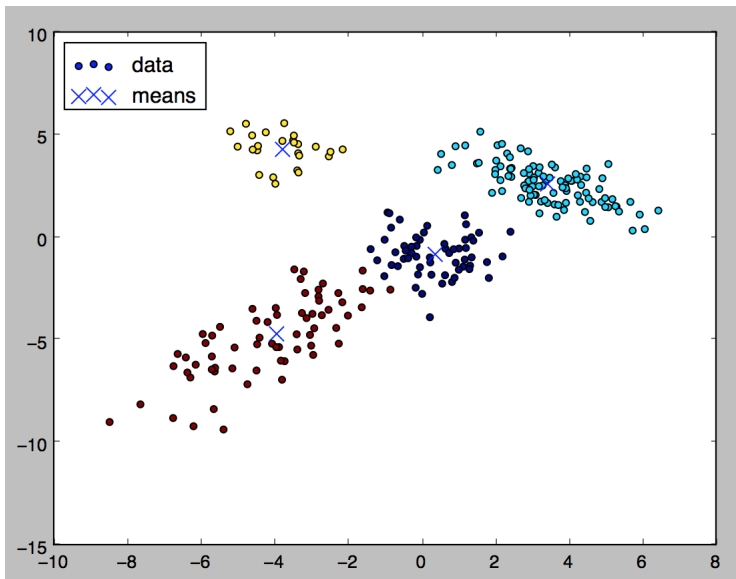
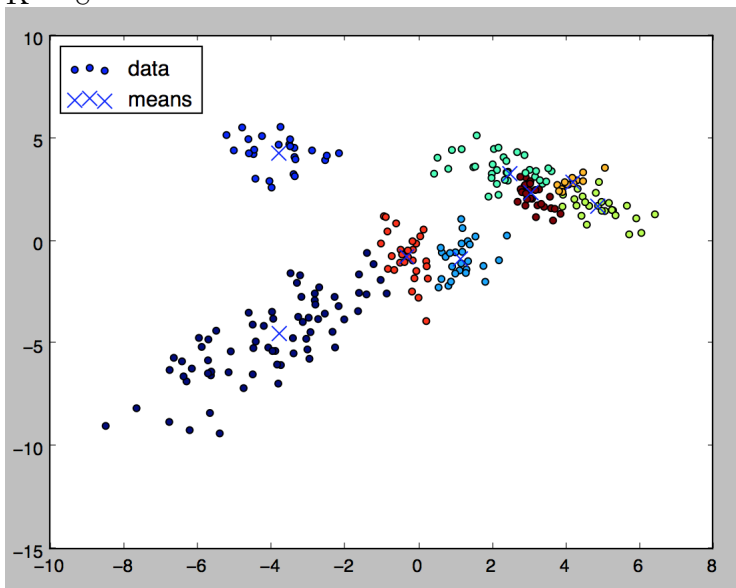
I notice that as K increases, the objective function converges more slowly, but converges to a higher overall likelihood. This makes sense as the algorithm is fitting a larger number of points; having more points minimizes the distance between any point and its cluster, increasing likelihood.

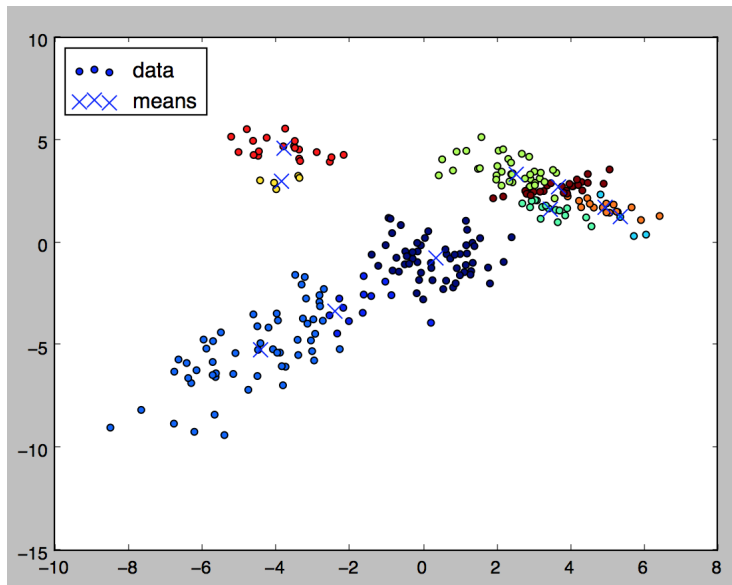
Part C

$K = 2$



$K = 4$

 $K = 8$  $K = 10$



I noticed here that additional clusters tended to spread out over the points, minimizing overall error. When $K=4$, the clusters reflect what appears to be the true distribution. As K increases, the clusters spread out over the data, but reflect less the underlying structure.

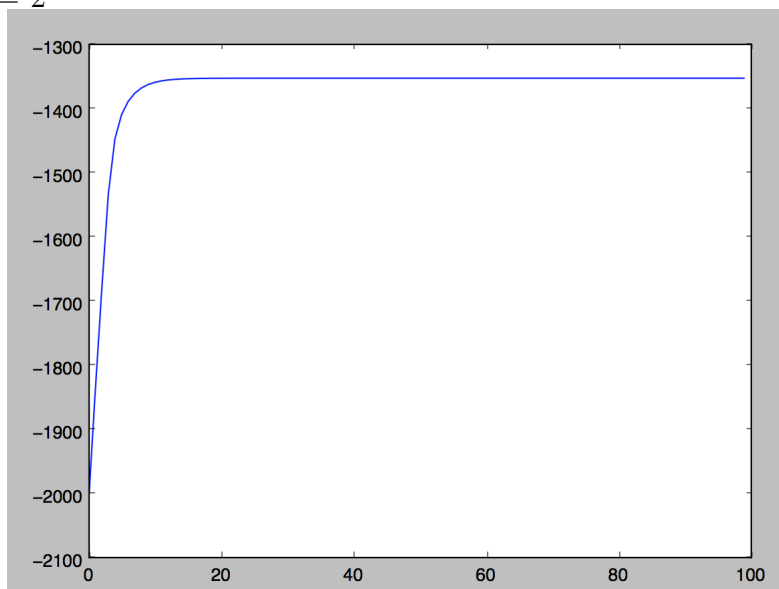
Problem 2

Part A

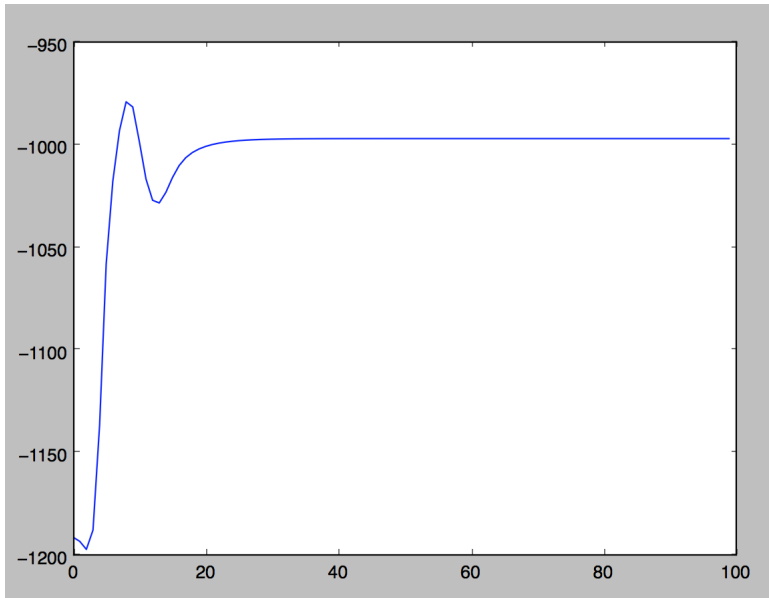
I did this it was challenging.

Part B

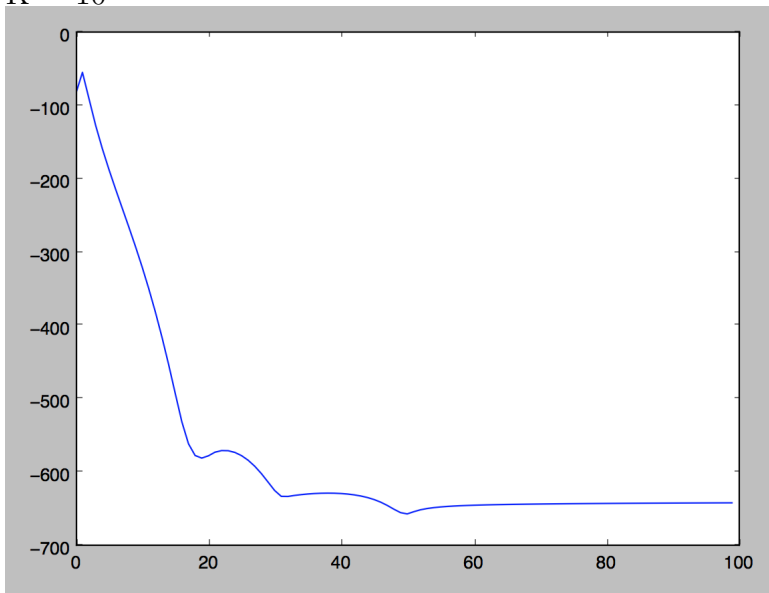
$K = 2$



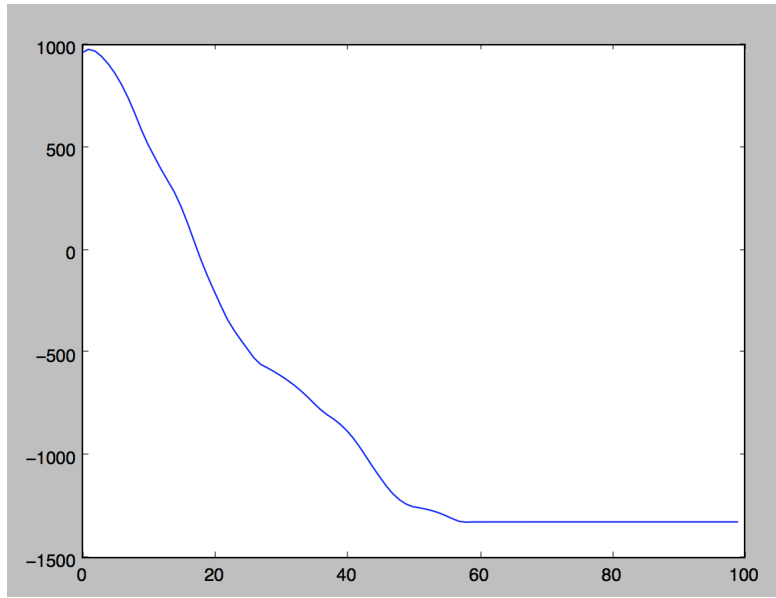
$K = 4$



$K = 10$



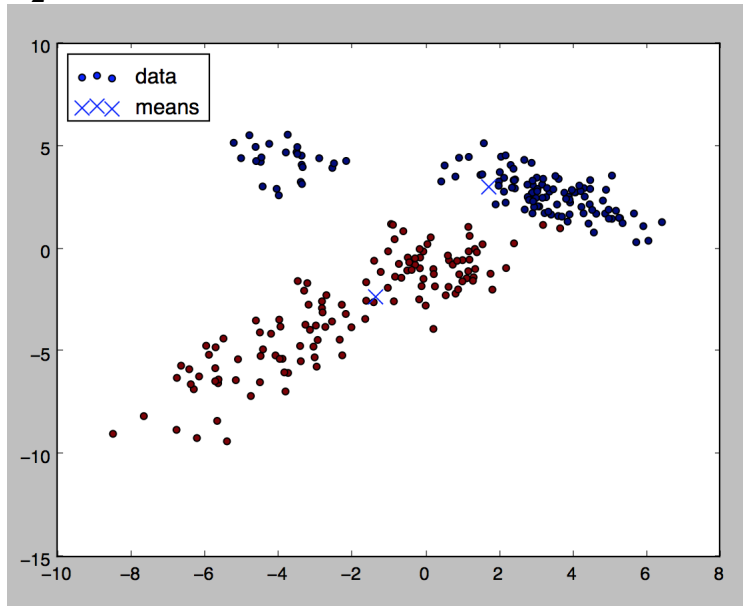
$K = 25$



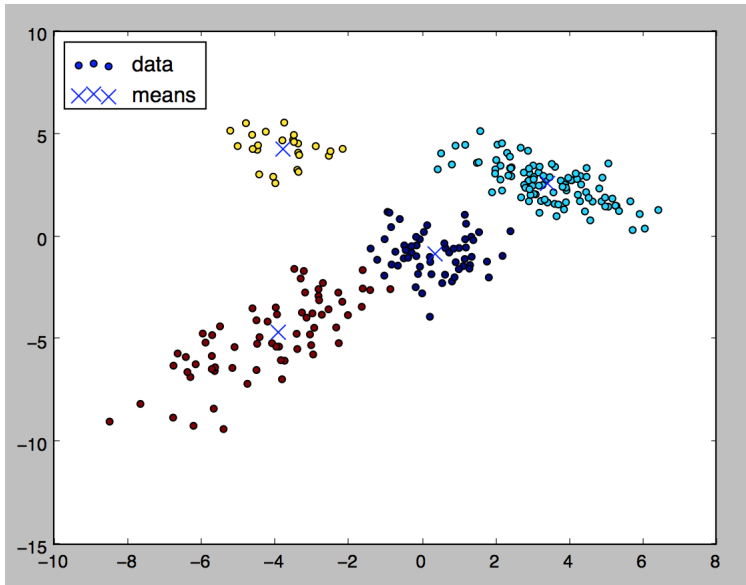
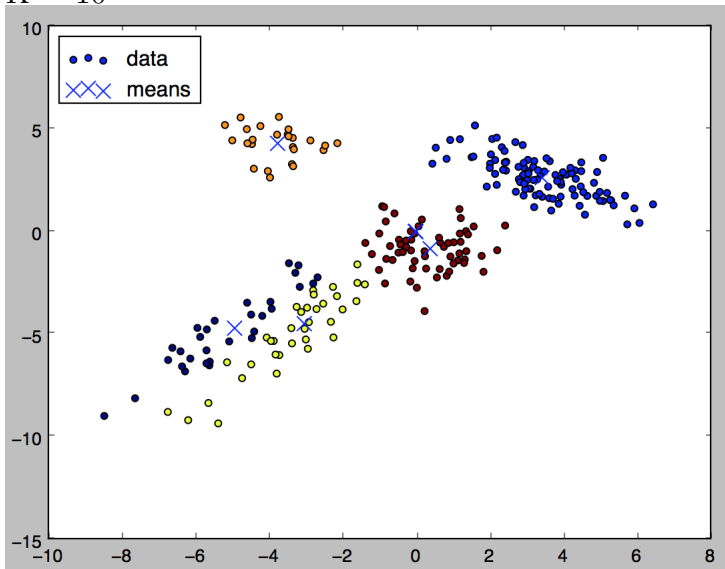
I was unable to fully debug the VI objective function; it is not consistently monotonically increasing. My algorithm is performing well on data, however, so I believe the error is with the objective function itself.

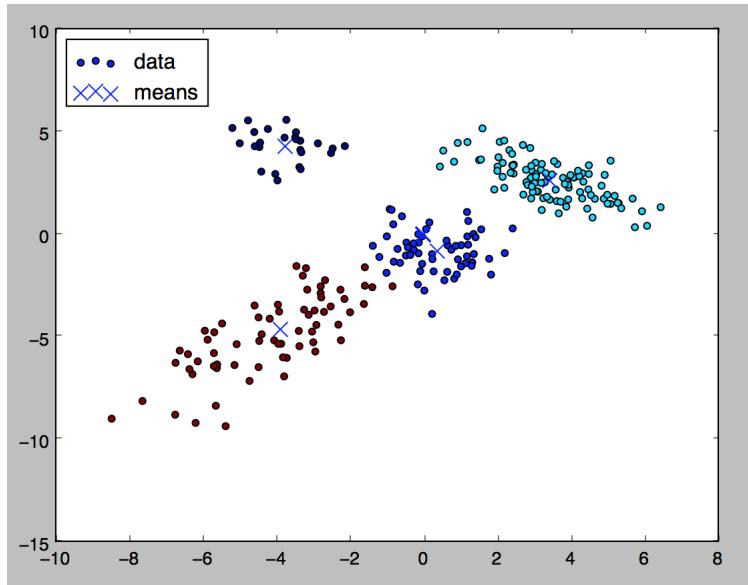
Part C

$K = 2$



$K = 4$

 $K = 10$  $K = 25$



I noticed that the VI algorithm will "learn" four clusters and place all of the density on those clusters; the density on the other points will go to zero, even though they are still being included in the algorithm.

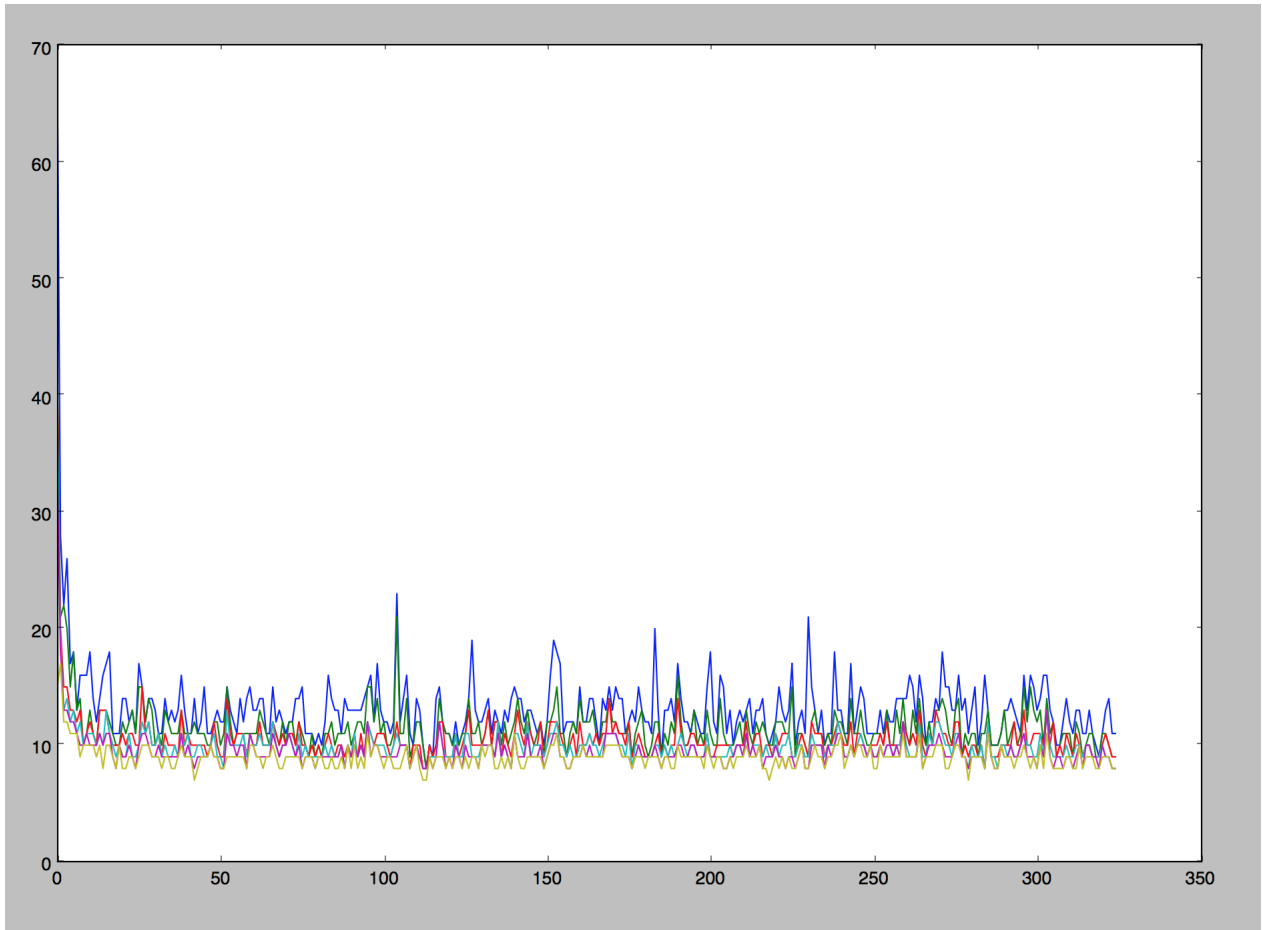
Problem 3

Part A

I did this it was challenging.

Part B

Points in top 6 clusters over 325 iterations (ended before 500 because the values were stable and it was past the HW submission deadline)



Part C

Number of clusters over time – stable in the 50 cluster region.

