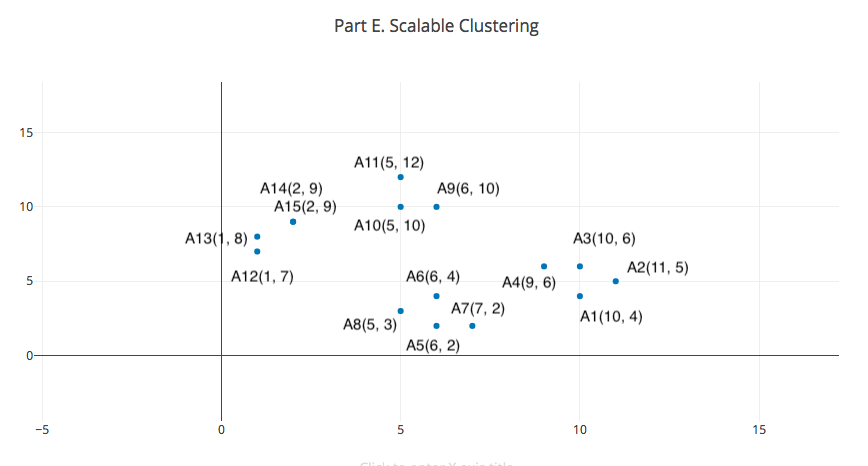
The X-Y plot of 15-data points:



Assume we cluster these 15 data points into 4 clusters. The initial seeds are A2, A7, A9, A12.

1, (3 points) The new clusters (i.e., the examples belonging to each cluster)

Cluster1: A2, A1, A3, A4

Cluster2: A7, A5, A8, A6

Cluster3: A9, A10, A11

Cluster4: A12, A13, A14, A15

2, (3 points) The centers of the new clusters

|  |  |
| --- | --- |
| ClusterId | Cluster Centroid |
| 1 | 10, 5.25 |
| 2 | 6, 2.75 |
| 3 | 5.33, 10.67 |
| 4 | 1.5, 8.25 |

3, (4 points) How many more iterations are needed to converge? Please write

down the final centers.

Iteration1:

|  |  |  |
| --- | --- | --- |
| ClusterId | Items | Cluster Centroid |
| 1 | A1, A2, A3, A4 | 10, 5.25 |
| 2 | A5, A6, A7, A8 | 6, 2.75 |
| 3 | A9, A10, A11 | 5.33, 10.67 |
| 4 | A12, A13, A14, A15 | 1.5, 8.25 |

Iteration2:

|  |  |  |
| --- | --- | --- |
| ClusterId | Items | Cluster Centroid |
| 1 | A1, A2, A3, A4 | 10, 5.25 |
| 2 | A5, A6, A7, A8 | 6, 2.75 |
| 3 | A9, A10, A11 | 5.33, 10.67 |
| 4 | A12, A13, A14, A15 | 1.5, 8.25 |

The cluster centroid has no changes in this iteration, so the iteration will stop at clusters:

Cluster1: A1, A2, A3, A4

Cluster2: A5, A6, A7, A8

Cluster3: A9, A10, A11

Cluster4: A12, A13, A14, A15