CSCI E-181 Spring 2014 Practical 1

David Wihl

January 31, 2014

Initially I used a KMeans implementation in Octave I had written for a previous course. While this implementation was sufficient for the prior course's Dataset, when I tested it with five random clusters of data, K=5 and random initial centroids, one of the centroids would frequently not converge on any points.
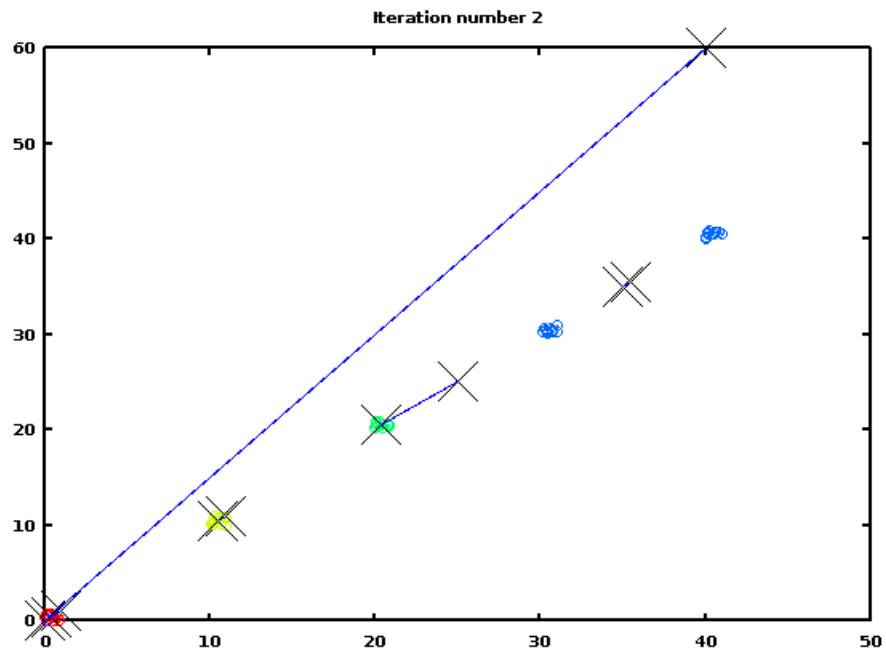

Figure 1: Random Initial Centroids After 1 Iteration

I subsequently modified the code to use a variation of KMeans++, specifically choosing randomly one of the sample data points as an initial centroid. This worked much better.