assignment4 handwriting problems

20190846 정희종

problem 1e: Clustering 2D pointsChart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

problem 1f: Trying the Algorithm on MNIST

Each centroid is a point in R^784, and it is mean of each cluster. By applying implemented algorithm on given data set, we attempt to classify given data set into 10 classes. But as we can see from the result, cluster 4(centroid 4) is omitted. This is due to the behavior of function update\_assignment. update\_assignment's return value does not include centroid which does not have any data point closest to it. So as a result we obtain 9 clusters(centroids), and again, since centroids are mean value of each cluster, the resulting centroid image looks like actual digits.

problem 2d: Clustering 2D pointsChart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

plots above are new 7 plots from soft\_kmean with original beta value(3)

result plot with beta = 50Chart, scatter chart

Description automatically generated

We can see that data points near the cluster boundary(decision boundary) became more certain about which cluster to belong. This is due to increasing beta. As we increase the value of beta, the softness of the cluster reduces.