

CSE 546 HW #3

Sam Kowash

November 20, 2018

1 Bayesian Inference

1. (a)
- (b)
- (c)
- (d)
- (e)
2. (a)
- (b)

2 Kernel Regression

3. (a)
- (b)
- (c)
- (d)
- (e)

3 k -means clustering

4. (a) Figures below show cluster centers in the full MNIST set (train joined with test) from k -means for a variety of different cluster counts. In general, they all resemble identifiable numerals, but more clusters tends to lead to more distinct features. At $k = 5$ we see centers somewhere in between similar numerals like 3 and 8 or 9 and 4, but at $k = 20$ the algorithm has the freedom to center clusters around more specific forms like taller or rounder or more slanted 9s.

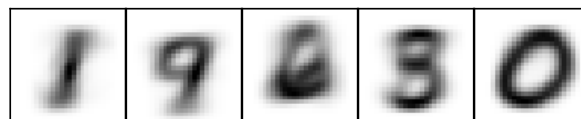


Figure 3.1: Final centers for $k = 5$ with uniform initialization

1	0	1	9	7
3	9	0	8	6

Figure 3.2: Final centers for $k = 10$ with uniform initialization

0	3	9	1	9
0	3	7	6	8
8	1	2	0	5
9	6	7	5	0

Figure 3.3: Final centers for $k = 20$ with uniform initialization

- (b) Figures below show centers obtained with the **k-means++** initialization scheme. They look broadly similar to the uniform case, although perhaps modestly crisper.

Fig. 4b compare the convergence rate to uniform initialization through the mean squared distance from a point to its cluster center. We see that other than for $k = 5$, the new initialization converges faster and to smaller mean distances.

0	1	3	6	9
---	---	---	---	---

Figure 3.4: Final centers for $k = 5$ with **k-means++** initialization

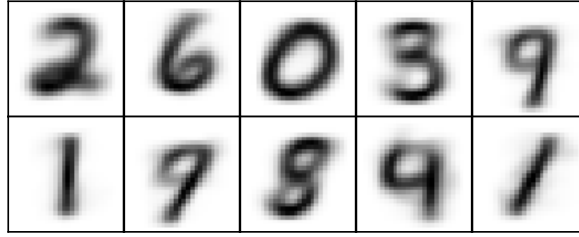


Figure 3.5: Final centers for $k = 10$ with k-means++ initialization

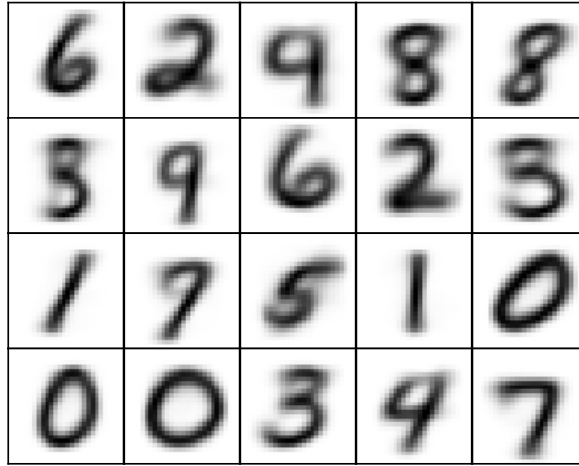


Figure 3.6: Final centers for $k = 20$ with k-means++ initialization

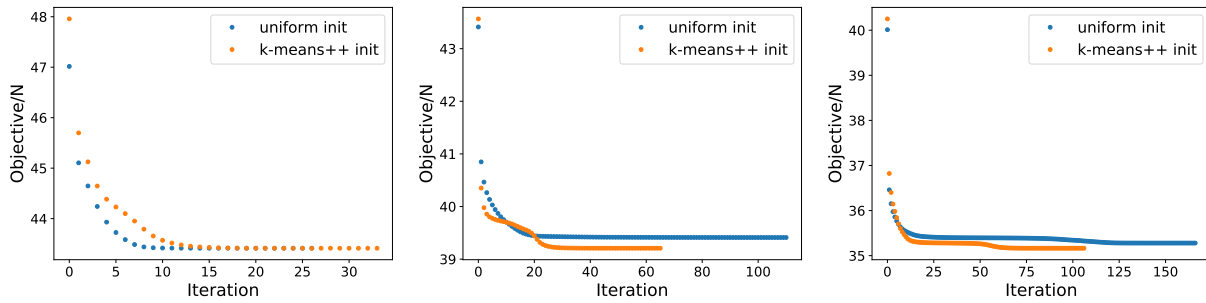


Figure 3.7: Objective function vs. iteration number for $k = 5, 10, 20$ from left to right

4 Joke Recommender System

5. (a)
- (b)
- (c)