Assignment 5

18/05/2020

Name: Yusuf Elnady

Q1.a

First, we have the following objective function $J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk}||x_n - \mu_k||^2$, which is a quadratic function of μ_k.

To minimize, we take the derivative w.r.t each μ_k and set it to zero, then we solve to obtain the formula of μ_k. Because we will have the fixed (kth) cluster to solve for, so we can remove the inner loop which is $\sum_{k=1}^{K} r_{nk}$ as I did in the second step.

$$\begin{aligned}\frac{\partial}{\partial \mu_k} J &= \frac{\partial}{\partial \mu_k} \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk}||x_n - \mu_k||^2 \\ &= \frac{\partial}{\partial \mu_k} \sum_{n=1}^{N} r_{nk}||x_n - \mu_k||^2 \\ &= -2 \sum_{n=1}^{N} r_{nk}(x_n - \mu_k) \\ &= -2 \sum_{n=1}^{N} r_{nk}(x_n) + 2 \sum_{n=1}^{N} r_{nk}(\mu_k) = 0\end{aligned} \tag{1}$$

Therefore, as μ_k doesn't depend on any sum, we can factor it out as following:

$$\begin{aligned}\sum_{n=1}^{N} r_{nk}(\mu_k) &= \sum_{n=1}^{N} r_{nk}(x_n) \\ \mu_k &= \frac{\sum_{n=1}^{N} r_{nk}(x_n)}{\sum_{n=1}^{N} r_{nk}}\end{aligned} \tag{2}$$

Q1.b

$$L2 = ||Y - XW||_2^2$$

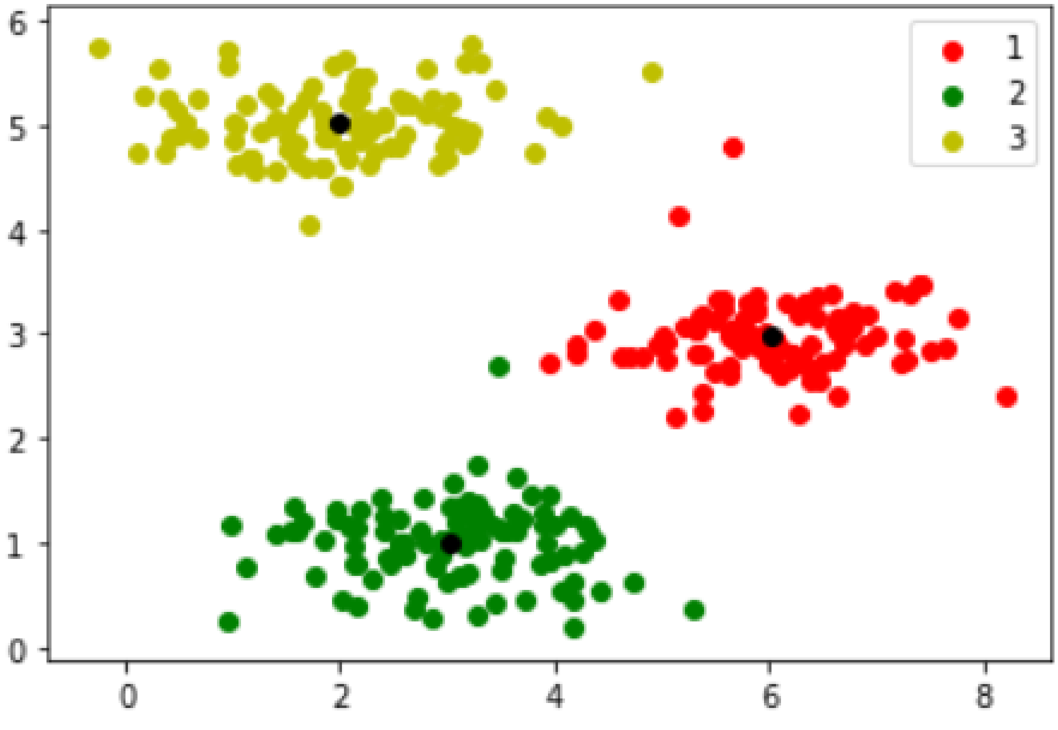
$$L1 = |Y - XW|$$

If we have data corrupted with a portion ϵ of outliers, then we need our cost function to be more robust to outliers, so we should choose L1 norm because it only takes the the absolute value, and considers them linearly. Least absolute errors is robust in that it is resistant to outliers in the data. Otherwise, L2-norm squares the error and the model will see a much larger error.

If we have only 1 outlier, we may choose L2 norm because it squares the values, and in this case we only have 1 outlier, so the cost will not increase exponentially.

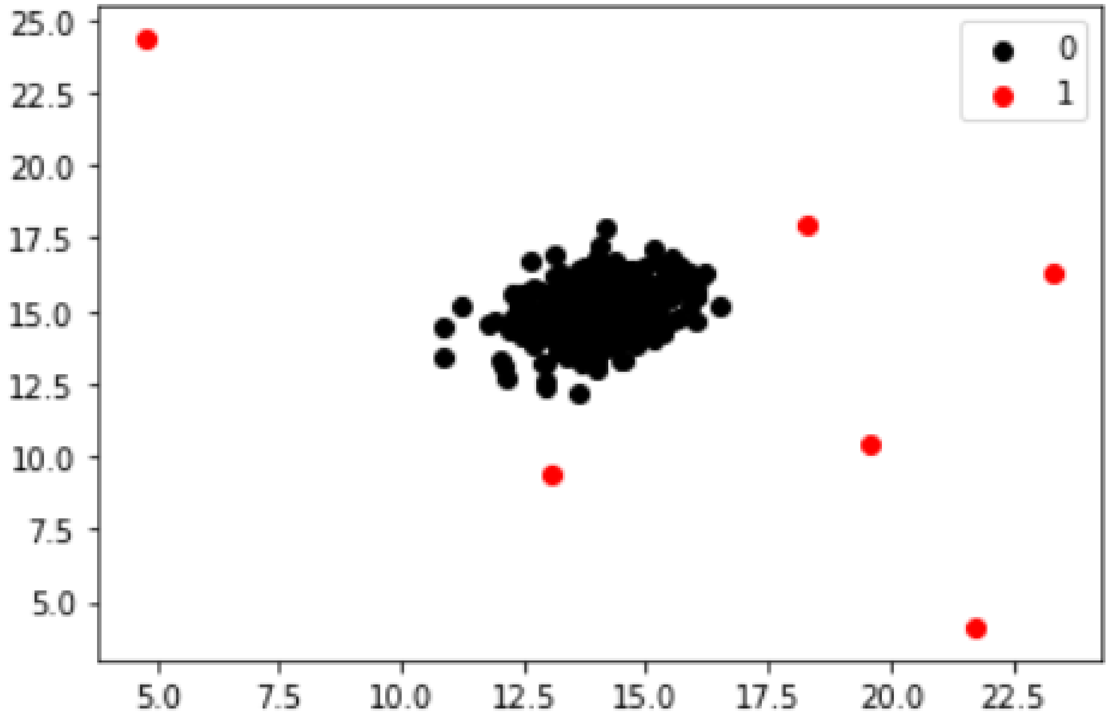
Q2.a

The Centroids after 10 iterations are:
(X1, X2)
(6.024403169393790, 2.9726605369458947)
(3.034582544606674, 0.9985308034402206)
(1.983631519927153, 5.030430038142128)



Q2.b

```
The maximum F1-Score is 0.875
The best threshold is 9.009009009009009e-05
The number of outliers is 6
```


25.0
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
5.0
7.5
10.0
12.5
15.0
17.5
20.0
22.5
● 0
● 1