1 Kernel K-means clustering

First, let us consider following objective function (to minimize) of K-means clustering.

$$\sum_{i \in K} \sum_{x_j \in X_i} ||x_j - \mathcal{C}_i||_2^2$$

Where K is the number of clusters, and X_i is group of data point in i's cluster. Also note that C_i is centroid of i's cluster.

Then we can modify the above objective function as follow.

$$\sum_{i \in K} \sum_{x_j \in X_i} \sum_{l} (x_{jl} - \mathcal{C}_{il})^2$$

And

$$\sum_{i \in K} \sum_{x_j \in X_i} \sum_{l} \left(x_{jl}^2 - 2x_{jl}\mathcal{C}_{il} + \mathcal{C}_{il}^2 \right)$$

Where x_{jl} is l's element of x_j.

Therefore

$$\sum_{i \in K} \sum_{x_j \in X_i} x_j \cdot x_j - 2x_j \cdot \mathcal{C}_i + \mathcal{C}_i \cdot \mathcal{C}_i$$

Now, centroid C_i is average of i's group. For this reason

$$\sum_{i \in K} \sum_{x_j \in X_i} \left(x_j \cdot x_j - \frac{2}{n} \sum_{x_k \in X_i} x_k \cdot x_j + \mathcal{C}_i \cdot \mathcal{C}_i \right)$$

Finary, we obtain new objective function.

Note that we can not calculate the centroid in new objective. So we do not devide into two steps, E and M.

$$\sum_{i \in K} \sum_{x_j \in X_i} \left(-\frac{2}{n} \sum_{x_k \in X_i} K(x_k, x_j) + \frac{1}{n^2} \sum_{x_k \in X_i} \sum_{x_l \in X_i} K(x_k, x_l) \right)$$

In the end, we execute simple one dimensional iteration which make a decision about each data point should join in which cluster. Note that $K(a, b)$ is a kernel function.

I am sorry for my poor english skils.

2018.2.22

参考文献

[1] Toshihiro Kamishima, "カーネル k-means 法",
`http://ibisforest.org/index.php?%E3%82%AB%E3%83%BC%E3%83%8D%E3%83%ABk-means%E6%B3%95`

[2] Inderjit S.Dhillon, Yuqiang Guan, Brian Kulis, "Kernel k-means, Spectral Clustering and Normalized Cuts",
`http://www.cs.utexas.edu/users/inderjit/public_papers/kdd_spectral_kernelkmeans.pdf`