Advanced Machine Learning - Assignment 1

Ragja Palakkadavath - SC18M003

January 2019

1. **Solution:**

   (a) Best No of Clusters obtained for Data 1: 4
   Silhouette Coefficient metric was obtained to get the best number of clusters and to analyse the quality of the clusters

   (b) Plot of the clusters:

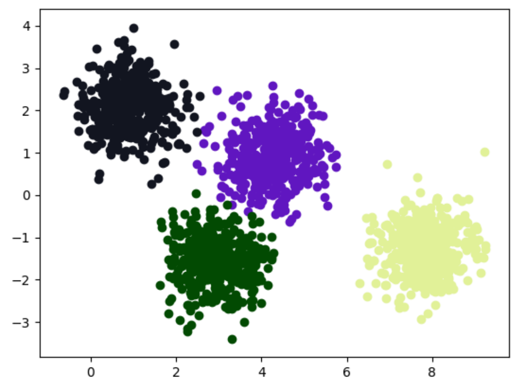

Figure 1: Cluster Plot - Data 1

   (c) Prior of each cluster

$$\begin{bmatrix} 0.24999366 & 0.25031739 & 0.25217938 & 0.24750957 \end{bmatrix}$$

Covariance Matrices of Each Cluster

$$\begin{bmatrix} 0.33181063 & 0.02446567 \\ 0.02446567 & 0.33903328 \end{bmatrix}$$

$$\begin{bmatrix} 0.33065482 & -0.01227766 \\ -0.01227766 & 0.33919031 \end{bmatrix}$$

$$\begin{bmatrix} 0.36177068 & 0.0096626 \\ 0.0096626 & 0.35275575 \end{bmatrix}$$

$$\begin{bmatrix} 0.32779448 & -0.01371615 \\ -0.01371615 & 0.32988095 \end{bmatrix}$$

Mean of each cluster:

$$\begin{bmatrix} 7.80803973 & -1.29762497 \end{bmatrix}$$
$$\begin{bmatrix} 0.92021663 & 2.07854824 \end{bmatrix}$$
$$\begin{bmatrix} 4.26165325 & 0.90687511 \end{bmatrix}$$
$$\begin{bmatrix} 2.91901799 & -1.58180678 \end{bmatrix}$$

(d) Quality of the clusters were assessed using Silhouette Coefficient
Silhouette Coeffecient for Data1: 0.6781.

2. **Solution:**

(a) (i) Logistic Function
Logistic function is a function with an S shaped curve given by $F(x) = \frac{1}{1+e^{-x}}$
The initial stage of growth of function is approximately exponential (geometric); then, as saturation begins, the growth slows to linear (arithmetic), and at maturity, growth stops, which in case of a standard logistic function =1.

(ii) Logistic Distribution
Logistic Distribution is a continuous probability distribution.
Its distribution function is given by, $F(x; \mu, s) = \frac{1}{1+e^{-\frac{(x-\mu)}{s}}}$
Its probability density function is given by, $f(x; \mu, \sigma) = \frac{e^{-\frac{(x-\mu)}{\sigma}}}{\sigma(1+e^{-\frac{(x-\mu)}{\sigma}})^2}, x \in R$
It is symmetrical, unimodal and has a shape similar to normal distribution.

(b) Logistic Function is used as the hypothesis function in Logistic Regression.

(c) $f(x; \mu, \sigma) = \frac{e^{-\frac{(x-\mu)}{\sigma}}}{\sigma(1+e^{-\frac{(x-\mu)}{\sigma}})^2}(-\infty < x < \infty) > 0$ and $\int f(x, \mu, \sigma)dx = 1$. Hence, properties of probability density function are being satisfied here by f. Therefore, (i) is pdf of (ii) as f=F'.

(d) $F(-\infty) = 0$ and $F(\infty) = 1$ and F is a continuous function in the given domain. Hence, (ii) is cdf of (i)

3. **Solution:**

(a) ICA was performed on $Data2.csv$ and the results were noted.

(b) Parameter W of the model is present in the file: $W\_ICA.csv$

(c) The sources where identified for the data for the various time. It is present in the file *sources.csv*

4. **Solution: Central Limit Theorem:**
Let $X_1, X_2, ...X_n$ be independent and identically distributed random variables with expected value $E[X_i] = \mu < \infty$ and variance $0 < Var[X_i] = \sigma^2 < \infty$. Then, the random variable,

$$Z_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + X_2 + ...X_n - n\mu}{\sqrt{n}\sigma}$$

converges in distribution to the standard normal random variables as n goes to infinity, that is

$$lim_{n\to\infty} P(Z_n \leq x) = \phi(x)$$

for all $x \epsilon R$ where $\phi(x)$ is the standard normal cumulative distribution function.

5. **Solution:** Given matrix A, ϵR^{nxn}, we want to find $\nabla_A |A|$

For matrix A ϵR^{nxn}, we define $A_{/i,/j} \epsilon R^{\text{(n-1)x(n-1)}}$to be the matrix that results from deleting ith row and jth column from the matrix A.

From this definition, the recursive formula for $|A|$ can be written as
$|A| = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} A_{\backslash i, \backslash j}$ for any j $\in$ 1,2....n

$\frac{\partial}{\partial_{A_{kl}}} |A| = \frac{\partial}{\partial_{A_{kl}}} (\sum_{i=1}^{n} (-1)^{i+j} a_{ij} A_{\backslash i, \backslash j})$
$=(-1)^{k+l} |A_{\backslash k, \backslash l}|$
$=((adj(A))_{lk} = ((adj(A^T))_{kl}$

We know that $adj(A) = |A| A^{-1}$

Computing the above derivative over all A_{kl}, we get $\nabla_A |A| = (adj(A))^T = |A| A^{-T}$