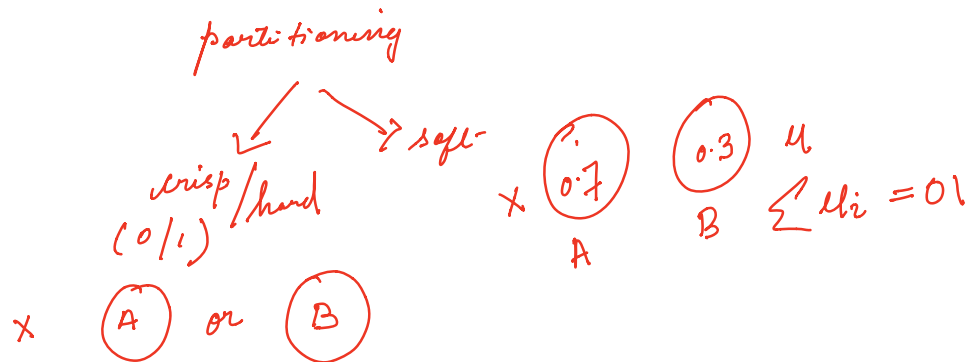


dec-29
20/03/21

Fuzzy C-means clustering algo:

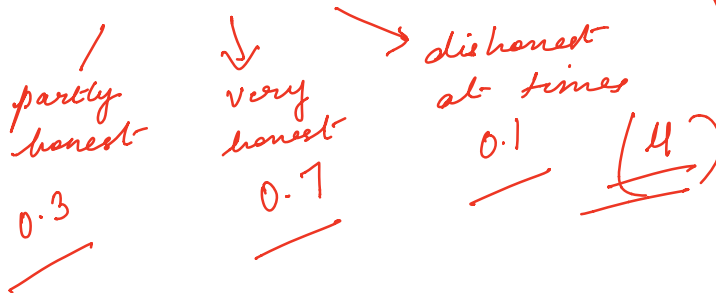
- Basic concept of fuzzy logic
- Algo
- Numerical



Fuzzy Logic :

S1: Sun rises in east- → 01
Sun " " west- → 00

S2: Ram is honest



X → C1 C2
0.6 0.4
Σ = 01

Fuzzy C-mean algo :

This algo works by assigning membership values to the data-point corresponding to each cluster center on the basis of distance between the cluster center and the data point.

→ More the the data point is near to the cluster center, more is the membership value towards the particular cluster center.

- The summation of membership values w.r.t the cluster point should be equal to 1
- After each iteration, membership values and the cluster centers are updated.

Steps of FCM algo :

- 1) Assume the no. of clusters to be made i.e C where $2 \leq C \leq N$.
- 2) Choose an appropriate level of cluster fuzziness i.e $g > 1$.
- 3) Initialize $N \times C$ membership matrix, $[U]$ at random s.t. $U_{ij} \in [0, 1]$ and $\sum_{j=1}^C U_{ij} = 1$ for each i .
- 4) Calculate k th dimension of j th cluster center CC_{jk} using the expression given below:

$$CC_{jk} = \frac{\sum_{i=1}^N U_{ij}^g x_{ik}}{\sum_{i=1}^N U_{ij}^g}$$

- 5) Calculate the Euclidean distance between i -th data point and j th cluster center as

$$d_{ij} = ||(CC_j - x_i)||$$

- 6) Update the fuzzy membership matrix $[U]$ acc. to d_{ij} .

If $d_{ij} \neq 0$, then $U_{ij} = \frac{1}{\sum_{m=1}^C \left(\frac{d_{ij}}{d_{im}} \right)^{2/g-1}}$

If $d_{ij} = 0$, then the data point coincides with j th cluster center cc_j and it will have full membership value i.e. $u_{ij} = 1$

7) Repeat steps (4) to (6) until the changes in U is less than a prespecified value.

Exit.

$$U = \begin{pmatrix} 0.1 & 0.3 \end{pmatrix} \text{ in } U$$

$$U = \begin{pmatrix} 0.699 & 0.33 \end{pmatrix}^{(n+1)}$$

$$\Delta U < \epsilon$$

Numerical :

Assume 10 points in 3-D space.

Assume $q = 1.25$ and termination criteria $\epsilon = 0.01$

	x	y	z
1	0.2	0.4	0.6
2	0.4	0.3	0.8
3	0.8	0.2	0.5
4	0.9	0.5	0.4
5	0.6	0.6	0.6
6	0.3	0.4	0.5
7	0.7	0.6	0.5
8	0.2	0.5	0.3
9	0.3	0.6	0.8
10	0.8	0.3	0.1

10x3

	<u>C1</u>	<u>C2</u>
1	0.6805	0.3194
2	0.4951	0.5048
3	0.8218	0.1781
4	0.3037	0.6962
5	0.339	0.6660
6	0.4315	0.5684
7	0.4153	0.5846
8	0.5096	0.4903
9	0.4698	0.5301
10	0.1891	0.8108

$\sum 4.6509$ 10x2

$$CC_{11} = \frac{\sum_{i=1}^N \mu_{ij}^g x_{i1}}{\sum_{i=1}^N \mu_{i1}} = \frac{A}{B} = 0.4111 \quad \left. \begin{array}{l} CC_{21} = 0.4829 \\ CC_{22} = 0.3995 \\ CC_{23} = 0.4157 \end{array} \right\}$$

$$\underline{CC_{12}} = 0.3481$$

$$CC_{13} = 0.4582$$

$$A = (0.6805^{1.25} \times 0.2) + (0.4951)^{1.25} \times 0.4 + (0.8218)^{1.25} \times 0.8 + (6.3037)^{1.25} \times 0.9 + (0.3339)^{1.25} \times 0.6 + (0.4315)^{1.25} \times 0.3 + (0.4153)^{1.25} \times 0.7 + (0.5096)^{1.25} \times 0.2 + (0.4698)^{1.25} \times 0.3 + (0.1891)^{1.25} \times 0.8 = 2.2668$$

$$B = 4.6509$$

$$\frac{A}{B} = \frac{2.2668}{4.6509} = 0.4111$$

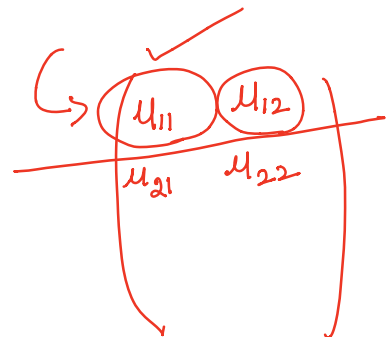
$$\therefore C1 = (0.4111, 0.3481, 0.4582)$$

$$C2 = (0.4829, 0.3995, 0.4157)$$

update the membership values :

$$\mu_{ij} = \frac{1}{\sum_{m=1}^c \left(\frac{d_{ij}}{d_{im}} \right)^{2/g-1}}$$

$$\begin{array}{l} m=1 \\ m=2 \end{array} \quad \mu_{11} = \frac{1}{\left(\frac{d_{11}}{d_{11}} \right)^{2/g-1} + \left(\frac{d_{11}}{d_{12}} \right)^{2/g-1}}$$



$$= \frac{1}{1 + \left(\frac{d_{11}}{d_{12}} \right)^{2/9-1}}$$

$$\mu_{12} = \frac{1}{\left(\frac{d_{12}}{d_{11}} \right)^{2/9-1} + \left(\frac{d_{12}}{d_{12}} \right)^{2/9-1}}$$

$$= \frac{1}{\left(\frac{d_{12}}{d_{11}} \right)^{2/9-1} + 1}$$

$$d_{11} = \sqrt{(0.4111 - 0.2)^2 + (0.3481 - 0.4)^2 + (0.4582 - 0.6)^2}$$

$$= 0.3243$$

$$d_{12} = \sqrt{(0.4829 - 0.2)^2 + (0.3995 - 0.4)^2 + (0.4157 - 0.6)^2}$$

$$= 0.313$$

Now applying d_{11} and d_{12} in μ_{11} & μ_{12}

$$\mu_{11} = 0.8885$$

$$\mu_{12} = 0.1114$$

$$\underline{\underline{\underline{Z = 0.9999}}}$$

After 1st iteration:

$\mu =$

0.8885	0.1114
0.9092	0.0907
0.3765	0.6234
0.1429	0.8570
0.2173	0.7826
0.9222	0.0777
0.0609	0.9390
0.5620	0.4379
0.7883	0.2116
0.2324	0.7675

10 X 2

After few iterations, the cluster centroids:

$$CC1 = (0.2161, 0.4479, 0.6003)$$

$$CC2 = (0.7739, 0.4257, 0.4040)$$