dec-29
20|03|21

Fuzzy c-means clustering algo:

→ Basic concept of fuzzy logic
→ Algo
→ Numerical

partitioning → crisp/hard (0/1), soft

x → A or B

x → A 0.7, B 0.3 μ $\sum \mu_i = 01$

Fuzzy logic:

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

S2: Ram is honest → partly honest 0.3, very honest 0.7, dishonest at times 0.1   (μ)

X → C1 0.6, C2 0.4 $\sum = 01$

Fuzzy-Cmean algo:

This algo works by assigning membership values to the data point corresponding to each cluster center on the basis of distance between the clusters 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 01

→ 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 $\underset{\text{(no. of data points)}}{N} \times C$ 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 Kth dimension of jth cluster center CC_{jK} using the expression given below:

$$\boxed{CC_{jK} = \frac{\sum_{i=1}^{N} U_{ij}^{g} x_{iK}}{\sum_{i=1}^{N} U_{ij}^{g}}}$$

5) Calculate the Euclidean distance between ith data point and jth cluster center as

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

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

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

If $d_{ij} = 0$, then the data point coincides with jth cluster center cc_j and it will have full membership value ie $U_{ij} = 1$

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

exit.

$$U = \begin{pmatrix} 0.7 & 0.3 \\ & \end{pmatrix} \text{ nth}$$

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

$$\Delta U < \epsilon$$

Numerical :

Assume 10 points in 3-D space. Assume $q = 1.25$ and termination criteria $\epsilon = 0.01$

$c = 2$

	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

10×3

$U =$

	C1	C2
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

10×2

$\Sigma \; 4.6509$

$$CC_{11} = \frac{\sum_{i=1}^{N} u_{ij}^{g} \, x_{i1}}{\sum_{i=1}^{N} u_{i1}} = \frac{A}{B} = 0.4111$$

$$CC_{12} = 0.3481$$

$$CC_{13} = 0.4582$$

$$CC_{21} = 0.4829$$

$$CC_{22} = 0.39995$$

$$CC_{23} = 0.41577$$

$$A = (0.6805^{1.25} \times 0.2) + (0.4951)^{1.25} \times 0.4 + (0.8218)^{1.25} \times 0.8 + (0.3037^{1.25} \times 0.9) + (0.3339^{1.25} \times 0.6) + (0.4315^{1.25} \times 0.3) + (0.4153^{1.25} \times 0.7) + (0.5096^{1.25} \times 0.2) + (0.4698^{1.25} \times 0.3) + (0.1891^{1.25} \times 0.8) = 2.2668$$

$$B = 4.6509$$

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

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

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

update the membership values :

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

$$\begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix}$$

$m=1$
$m=2$

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

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

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

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

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

$$= 0.3243$$

$$d_{12} = \sqrt{(0.4829 - 0.2)^2 + (0.39995 - 0.4)^2 + (0.4157 - 0.6)^2}$$

$$= 0.313$$

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

$$u_{11} = 0.8885$$

$$u_{12} = 0.1114$$

$$\Sigma = .9999$$

After 1st itern :

$$
U = \begin{bmatrix} 0.8885 & 0.1114 \\ 0.9092 & 0.0907 \\ 0.3765 & 0.6234 \\ 0.1429 & 0.8570 \\ 0.2173 & 0.7826 \\ 0.9222 & 0.0779 \\ 0.0609 & 0.9390 \\ 0.5620 & 0.4379 \\ 0.7883 & 0.2116 \\ 0.2324 & 0.7675 \end{bmatrix}
$$

10×2

After few iterations, the cluster centroids :

$CC1 = (0.2961,\ 0.4479,\ 0.6003)$

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