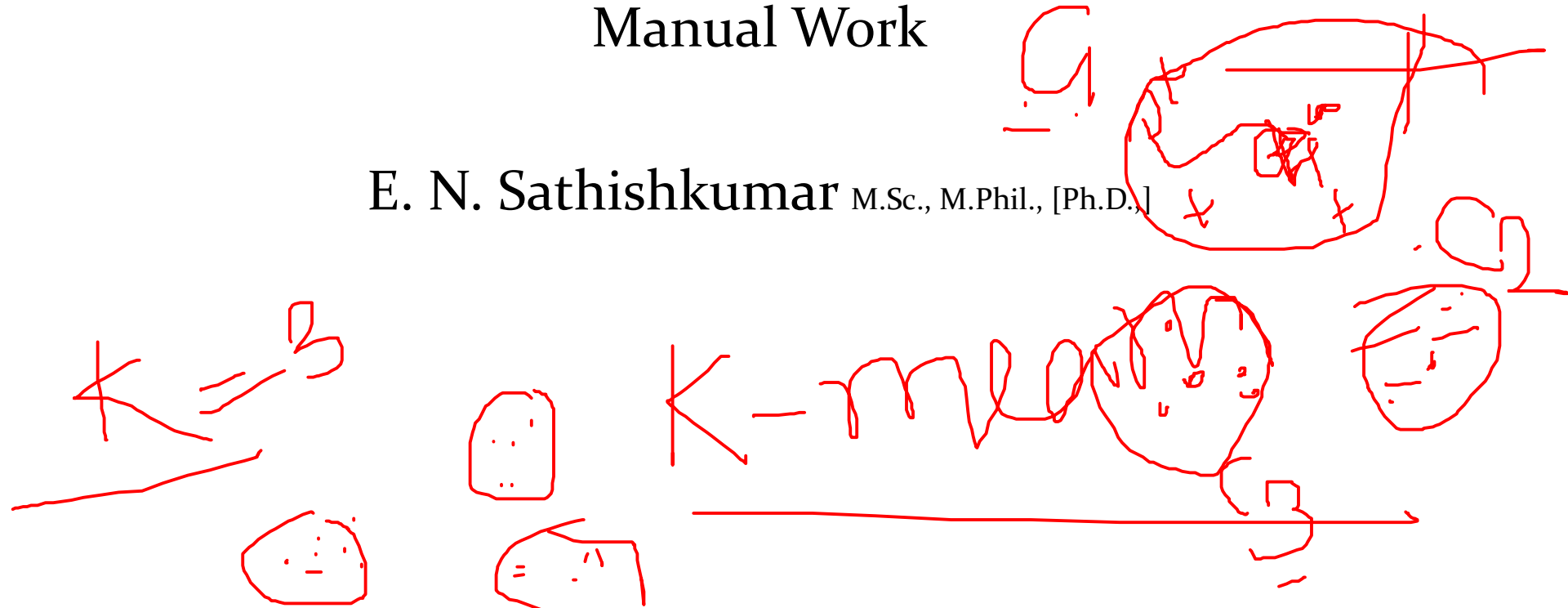
Fuzzy C Means

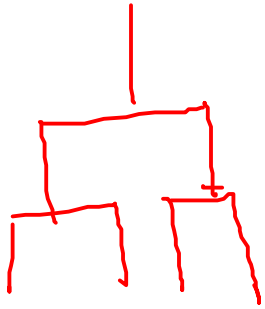
Manual Work

E. N. Sathishkumar M.Sc., M.Phil., [Ph.D.,]



Fuzzy C-Means

- An extension of k-means
- Hierarchical, k-means generates partitions
  - each data point can only be assigned in one cluster
- Fuzzy c-means allows data points to be assigned into more than one cluster
  - each data point has a degree of membership (or probability) of belonging to each cluster



Fuzzy C Means Algorithm

Step-1: Randomly initialize the membership matrix using this equation.

$$\sum_{j=1}^{c} \mu_j(x_i) = 1 \qquad i = 1, 2 \ldots k$$

Step-2: Calculate the Centroid using equation.

$$Cj = \frac{\sum_i [\mu_j(x_i)]^m x_i}{\sum_i [\mu_j(x_i)]^m}$$

Step-3: Calculate dissimilarly between the data points and Centroid using the Euclidean distance.

$$D_i = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Step-4: Update the New membership matrix using the equation.

$$\mu_j(x_i) = \frac{[\frac{1}{d_{ji}}]^{1/m-1}}{\sum_{k=1}^{c} [\frac{1}{d_{ki}}]^{1/m-1}}$$

Here **m** is a fuzzification parameter.
The range **m** is always [1.25, 2]

Step -5: Go back to Step 2, unless the centroids are not changing.

Worked out Example

- ***Input:*** Number of Objects = 6 Number of clusters = 2

X	Y	C1	C2
1	6	0.8	0.2
2	5	0.9	0.1
3	8	0.7	0.3
4	4	0.3	0.7
5	7	0.5	0.5
6	9	0.2	0.8

Step-1: Initialize the membership matrix.

Step-2: Find the constraint using the equation

$$Cj = \left[\frac{\sum_i [\mu_j(x_i)]^m x_i}{\sum_i [\mu_j(x_i)]^m}, \frac{\sum_i [\mu_j(y_i)]^m y_i}{\sum_i [\mu_j(y_i)]^m}\right]$$

$$C1 = \left[\frac{1*0.8^2+2*0.9^2+3*0.7^2+4*0.3^2+5*0.5^2+6*0.2^2}{0.8^2+0.9^2+0.7^2+0.3^2+0.5^2+0.2^2}, \frac{6*0.8^2+5*0.9^2+8*0.7^2+4*0.3^2+7*0.5^2+9*0.2^2}{0.8^2+0.9^2+0.7^2+0.3^2+0.5^2+0.2^2}\right]$$

$$C1 = \frac{5.58}{2.32}, \frac{14.28}{2.32}$$

$$C1 = (2.4, 6.1)$$

$$C2 = \left[\frac{1*0.2^2 + 2*0.1^2 + 3*0.3^2 + 4*0.7^2 + 5*0.5^2 + 6*0.8^2}{0.2^2 + 0.1^2 + 0.3^2 + 0.7^2 + 0.5^2 + 0.8^2}, \right.$$

$$\left. \frac{6*0.2^2 + 5*0.1^2 + 8*0.3^2 + 4*0.7^2 + 7*0.5^2 + 9*0.8^2}{0.2^2 + 0.1^2 + 0.3^2 + 0.7^2 + 0.5^2 + 0.8^2}\right]$$

$$C2 = \frac{7.38}{1.52}, \frac{10.48}{1.52}$$

$$C_2 = (4.8, 6.8)$$

Step- 3 : Find Distance

$$D_i = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Centroid 1 :

$$(1,6)(2.4,6.1) = \sqrt{(1.4)^2 + (0.1)^2} = \sqrt{1.96 + 0.01} = \sqrt{1.97} = 1.40$$

$$(2,5)(2.4,6.1) = \sqrt{0.16 + 1.21} = \sqrt{1.37} = 1.17$$

$$(3,8)(2.4,6.1) = \sqrt{0.36 + 3.61} = \sqrt{3.97} = 1.99$$

$$(4,4)(2.4,6.1) = \sqrt{2.56 + 4.41} = \sqrt{6.97} = 2.64$$

$$(5,7)(2.4,6.1) = \sqrt{6.76 + 0.81} = \sqrt{7.57} = 2.75$$

$$(6,9)(2.4,6.1) = \sqrt{12.96 + 8.41} = \sqrt{21.37} = 4.62$$

Centroid 2:

$(1,6)(4.8,6.8) = \sqrt{14.44 + 0.64} = \sqrt{15.08} = 3.88$

$(2,5)(4.8,6.8) = \sqrt{7.84 + 3.24} = \sqrt{11.08} = 3.32$

$(3,8)(4.8,6.8) = \sqrt{3.24 + 1.44} = \sqrt{4.68} = 2.16$

$(4,4)(4.8,6.8) = \sqrt{0.64 + 7.84} = \sqrt{8.48} = 2.91$

$(5,7)(4.8,6.8) = \sqrt{0.04 + 0.04} = \sqrt{0.08} = 0.28$

$(6,9)(4.8,6.8) = \sqrt{1.44 + 4.84} = \sqrt{6.28} = 2.50$

Cluster 1		Cluster 2	
Datapoint	**Distance**	**Datapoint**	**Distance**
(1,6)	1.40	(1,6)	3.88
(2,5)	1.17	(2,5)	3.32
(3,8)	1.99	(3,8)	2.16
(4,4)	2.64	(4,4)	2.91
(5,7)	2.75	(5,7)	0.28
(6,9)	4.62	(6,9)	2.50

Step-4 : Update the membership value

$$\mu_j(x_i) = \frac{[\frac{1}{d_{ji}}]^{1/m-1}}{\sum_{k=1}^{c}[\frac{1}{d_{ki}}]^{1/m-1}}$$

here m = 2, i – first data point, j - first cluster

Cluster 1

μ_{11} = $(1 / d_{11})^{1/2-1} / (1 / d_{11})^{1/2-1} + (1 / d_{21})^{1/2-1}$

= $(1/1.40)^{1} / (1/1.40)^{1} + (1/3.88)^{1}$ = 0.71 / 0.71+0.25

= 0.71/ 0.96 = 0.7

$\mu_{12} = (1 / d_{12}) / (1 / d_{12}) + (1 / d_{22})$

$\Rightarrow 1/1.17 / 1/1.17 + 1/3.32 = 0.56 / 0.56+0.30$

$= 0.56 / 0.86 = 0.6$

$\mu_{13} = (1 / d_{13}) / (1 / d_{13}) + (1 / d_{23})$

$\Rightarrow 1/1.99 / 1/1.99 + 1/2.16 = 0.50 / 0.50+0.46$

$= 0.50 / 0.96 = 0.5$

$\mu_{14} = (1 / d_{14}) / (1 / d_{14}) + (1 / d_{24})$

$\Rightarrow 1/2.64 / 1/2.64 + 1/2.91 = 0.37 / 0.37+0.34$

$= 0.37 / 0.71 = 0.5$

$\mu_{15} = (1 / d_{15}) / (1 / d_{15}) + (1 / d_{25})$

$\Rightarrow 1/2.75 / 1/2.75 + 1/0.28 = 0.36 / 0.36+3.57$

$= 0.36 / 3.93 = 0.1$

$\mu_{16} = (1 / d_{16}) / (1 / d_{16}) + (1 / d_{26})$

$\Rightarrow 1/4.62 / 1/4.62 + 1/2.50 = 0.21 / 0.21+0.4$

$= 0.21 / 0.61 = 0.3$

Cluster 2

$$\mu_{21} = (1 / d_{21}) / (1 / d_{12}) + (1 / d_{21})$$

$$\Rightarrow 1/3.88 / 1/1.40 + 1/3.88 = 0.25 / 0.71+0.25$$

$$= 0.25 / 0.96 = 0.3$$

$$\mu_{22} = (1 / d_{22}) / (1 / d_{12}) + (1 / d_{22})$$

$$\Rightarrow 1/3.32 / 1/1.17 + 1/3.32 = 0.30 / 0.56+0.30$$

$$= 0.30 / 0.86 = 0.4$$

$$\mu_{23} = (1 / d_{23}) / (1 / d_{13}) + (1 / d_{23})$$

$$\Rightarrow 1/2.16 / 1/1.99 + 1/2.16 = 0.46 / 0.50+0.46$$

$$= 0.46 / 0.96 = 0.5$$

$\mu_{24} = (1 / d_{24}) / (1 / d_{14}) + (1 / d_{24})$

$\Rightarrow 1/2.19 \; / \; 1/2.64 + 1/2.19 = 0.34 \; / \; 0.37+0.34$

$= 0.34 / \; 0.71 = 0.5$

$\mu_{25} = (1 / d_{25}) / (1 / d_{15}) + (1 / d_{25})$

$\Rightarrow 1/0.28 \; / \; 1/2.75 + 1/0.28 = 3.57 \; / \; 0.36+3.57$

$= 3.57 / \; 3.93 = 0.9$

$\mu_{26} = (1 / d_{26}) / (1 / d_{16}) + (1 / d_{26})$

$= 0.4 / \; 0.21+0.4 = 0.4/0.61$

$= 0.7$

- Now the New Membership value is

X	Y	C1	C2
1	6	0.7	0.3
2	5	0.6	0.4
3	8	0.5	0.5
4	4	0.5	0.5
5	7	0.1	0.9
6	9	0.3	0.7

Step 5 : Now continue this process until get the same centroids.