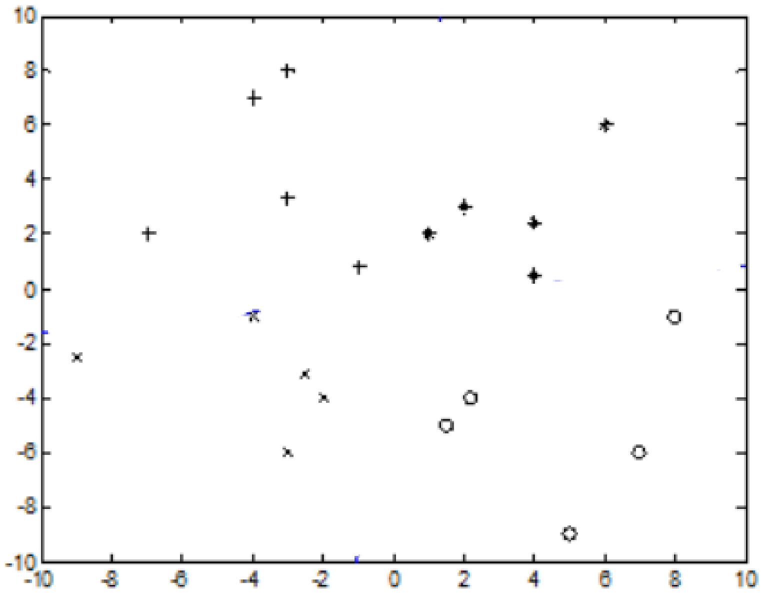
Fuzzy Clustering

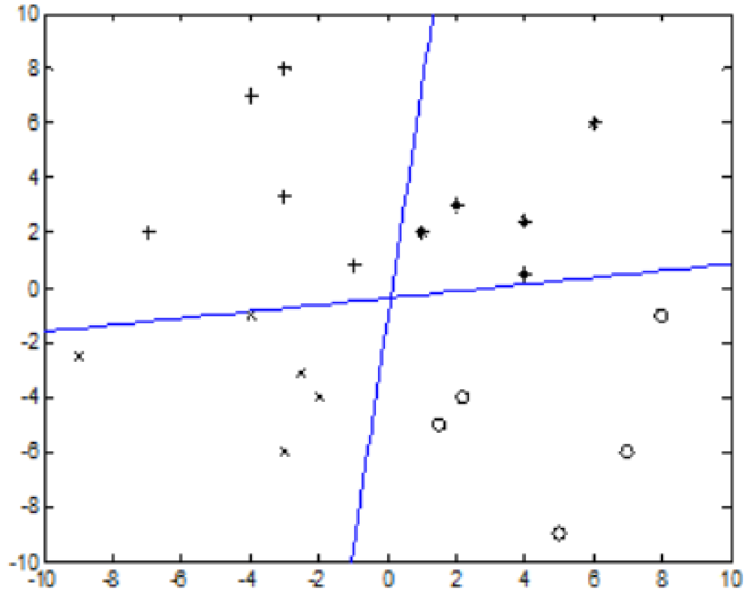
- **What is Clustering?**

  - It deals with finding a structure in a collection of unlabeled data

  - This is an unsupervised study where data of similar types are put into one cluster while data of different types are put into different clusters

  - Hence an unsupervised classification problem

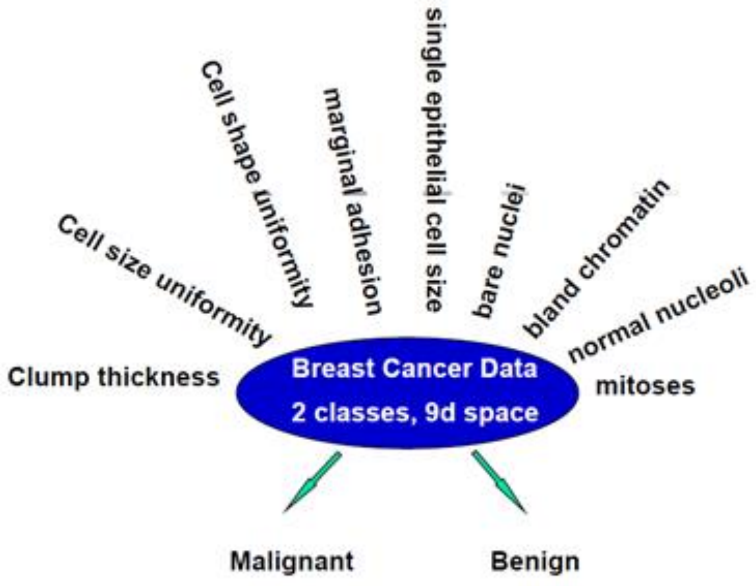


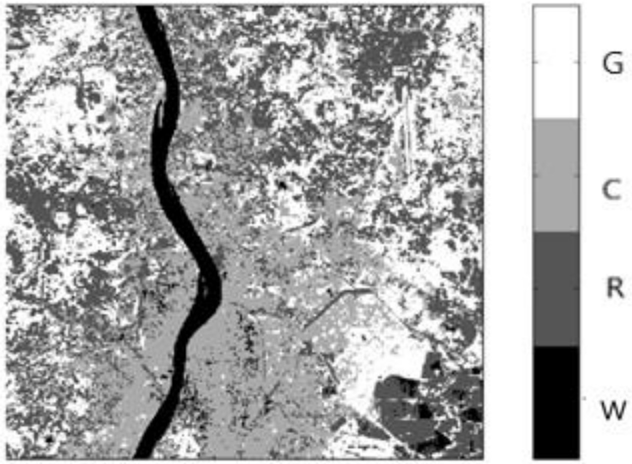


- ✓ Search for appropriate decision boundaries/separating hyper-planes
- ✓ Hence can be cast as an optimization problem

Some Applications:

- ✓ Medical diagnosis
- ✓ Image segmentation
- ✓ Structural health monitoring of structures





- The quality of a clustering method is also measured by its ability to discover some or all of the hidden patterns

- A good clustering method will produce high quality clusters with
  - HIGH INTRA-CLASS SIMILARITY
    (Similar to one another within the same cluster)
  - LOW INTER-CLASS SIMILARITY
    (Dissimilar to the objects in other clusters)

Clustering can be classified as:

- Hard Clustering (or Exclusive Clustering)
- Soft Clustering (Overlapping Clustering)

- K-means is a popular hard clustering technique
- Fuzzy C Means (FCM) is a popular soft clustering technique

K-Means Clustering

$$
U = \begin{array}{cc}
 & \begin{array}{cc} C1 & C2 \end{array} \\
\begin{array}{c} \underline{x}1 \\ \underline{x}2 \\ \\ \cdot \\ \cdot \\ \cdot \\ \underline{x}N \end{array} &
\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ \cdot & \cdot \\ \cdot & \cdot \\ 0 & 1 \end{bmatrix}
\end{array}
$$

- ✓ K means clustering clusters the entire dataset into K number of clusters where a data should belong to only one cluster
- ✓ Membership or partition matrix U is of the form:
- ✓ One of the simplest Similarity Measure is “Euclidean distance” between pairs of feature vectors in the feature space

  (Distance between the points in the same cluster will be considerably less than distance between points in different clusters)

- ✓ The first step is normalization of the data. This step is very important when dealing with parameters of different units and scales.

$$X_{normalized[0,1]} = \frac{X_{current} - X_{\min}}{X_{\max} - X_{\min}}$$

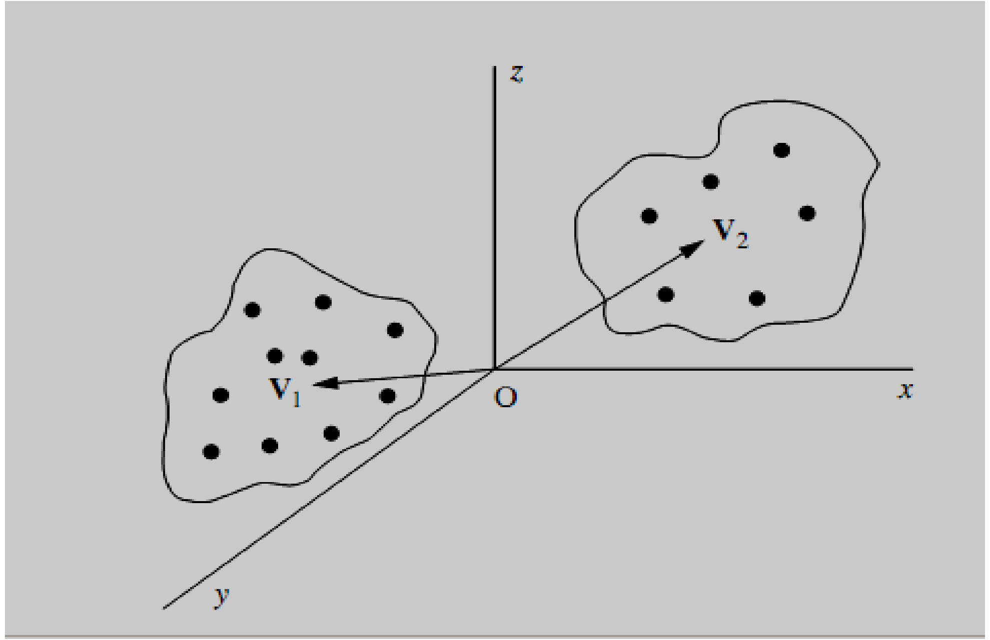
$$X_{normalized[-1,\,1]} = \frac{X_{current} - (X_{\max} + X_{\min})/2}{(X_{\max} - X_{\min})/2}$$

- ✓ In case there are outliers in the data, we prefer standardization $x_{new} = \dfrac{x - \mu}{\sigma}$

Objective function is developed to

1. Minimize Euclidian distance between each data points in a cluster and its cluster centre

2. Maximize the Euclidian distance between cluster centers (centroids)



Steps for K-Means Clustering:

1) Set K– choose a number of desired clusters, K

2) Initialization – choose 'K' starting points which are used as initial estimates of the cluster centroids

3) Classification – Examine each point in the dataset and assign it to the cluster whose centroid is nearest to it

4) Centroid calculation – When each point in the data set is assigned to a cluster,

recalculate the new 'K' centroids.

This is done by taking average of the members of the particular cluster

5) Convergence criteria – Steps 3 & 4 to be repeated until no point changes its cluster

assignment or until the centroids no longer move

A Numerical Example: two dimensional feature space

(using K = No. of clusters = 2)

Individual	Variable 1	Variable 2
1.	1	1
2.	1.5	2
3.	3	4
4.	5	7
5.	3.5	5
6.	4.5	5
7.	3.5	4.5

$$\underline{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\underline{x}_2 = \begin{bmatrix} 1.5 \\ 2 \end{bmatrix}$$

$$\underline{x}_3 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

$$\underline{x}_4 = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

$$\underline{x}_5 = \begin{bmatrix} 3.5 \\ 5 \end{bmatrix}$$

$$\underline{x}_6 = \begin{bmatrix} 4.5 \\ 5 \end{bmatrix}$$

$$\underline{x}_7 = \begin{bmatrix} 3.5 \\ 4.5 \end{bmatrix}$$

Step 1:

Initialization: Randomly choose following two centroids for the two clusters

m1=(1.0,1.0) and m2=(5.0,7.0)

Individual	Variable 1	Variable 2
1	1	1
2	1.5	2
3	3	4
4	5	7
5	3.5	5
6	4.5	5
7	3.5	4.5

$\underline{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad : \underline{m}_1$

$\underline{x}_2 = \begin{bmatrix} 1.5 \\ 2 \end{bmatrix}$

$\underline{x}_3 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

$\underline{x}_4 = \begin{bmatrix} 5 \\ 7 \end{bmatrix} \quad : \underline{m}_2$

$\underline{x}_5 = \begin{bmatrix} 3.5 \\ 5 \end{bmatrix}$

$\underline{x}_6 = \begin{bmatrix} 4.5 \\ 5 \end{bmatrix}$

$\underline{x}_7 = \begin{bmatrix} 3.5 \\ 4.5 \end{bmatrix}$

Individual	Distance from Centroid 1 (1,1)	Distance from Centroid 2 (5,7)
1(1.0,1.0)	0	7.21
2(1.5,2)	1.12	6.1
3(3,4)	3.51	3.51
4(5,7)	7.21	0
5(3.5,5)	4.72	2.5
6(4.5,5)	5.31	2.06
7(3.5,4.5)	4.3	2.92

$$d(m_1, 1) = \|\underline{x}_1 - \underline{m}_1\| = 0$$

$$d(m_2, 1) = \|\underline{x}_1 - \underline{m}_2\|$$

$$= \left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 5 \\ 7 \end{bmatrix} \right\|$$

$$= \left\| \begin{bmatrix} -4 \\ -6 \end{bmatrix} \right\|$$

$$= \sqrt{(-4)^2 + (-6)^2}$$

$$= 7.21$$

Distance of point/individual 3 is same from both centroids

Individual	Distance from Centroid 1 (1,1)	Distance from Centroid 2 (5,7)
1(1.0,1.0)	0	7.21
2(1.5,2)	1.12	6.1
3(3,4)	3.51	3.51
4(5,7)	7.21	0
5(3.5,5)	4.72	2.5
6(4.5,5)	5.31	2.06
7(3.5,4.5)	4.3	2.92

Step 2: Thus, we obtain two clusters containing {1,2,3} and {4,5,6,7}. Their new centroids are:

$$\underline{m}_1 = \frac{1}{3}\left\{ \begin{bmatrix}1\\1\end{bmatrix}+\begin{bmatrix}1.5\\2\end{bmatrix}+\begin{bmatrix}3\\4\end{bmatrix} \right\} = \begin{bmatrix}1.83\\2.33\end{bmatrix}$$

$$\underline{m}_2 = \frac{1}{4}\left\{ \begin{bmatrix}5\\7\end{bmatrix}+\begin{bmatrix}3.5\\5\end{bmatrix}+\begin{bmatrix}4.5\\5\end{bmatrix}+\begin{bmatrix}3.5\\4.5\end{bmatrix} \right\} = \begin{bmatrix}4.12\\5.38\end{bmatrix}$$

Individual	Distance from Centroid 1 (1.83,2.33)	Distance from Centroid 2 (4.12,5.38)
1(1.0,1.0)	1.57	5.38
2(1.5,2)	0.47	4.28
3(3,4)	2.04	1.78
4(5,7)	5.64	1.84
5(3.5,5)	3.15	0.73
6(4.5,5)	3.78	0.54
7(3.5,4.5)	2.74	1.08

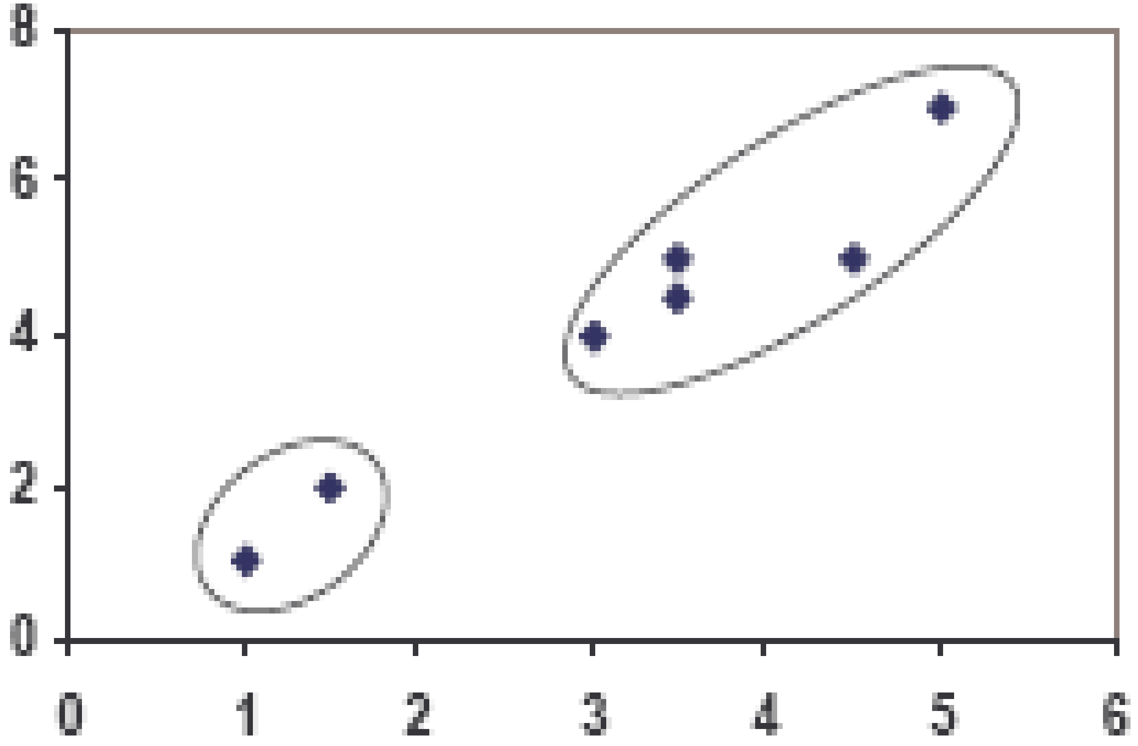
Individual	Distance from Centroid 1 (1.83,2.33)	Distance from Centroid 2 (4.12,5.38)
1(1.0,1.0)	1.57	5.38
2(1.5,2)	0.47	4.28
3(3,4)	2.04	1.78
4(5,7)	5.64	1.84
5(3.5,5)	3.15	0.73
6(4.5,5)	3.78	0.54
7(3.5,4.5)	2.74	1.08

Therefore, the new clusters are: {1,2} and {3,4,5,6,7}

And new centroids are:

$$\underline{m}_1 = \frac{1}{2}\left\{ \begin{bmatrix}1\\1\end{bmatrix}+\begin{bmatrix}1.5\\2\end{bmatrix} \right\} = \begin{bmatrix}1.25\\1.5\end{bmatrix}$$

$$\underline{m}_2 = \frac{1}{5}\left\{ \begin{bmatrix}3\\4\end{bmatrix}+\begin{bmatrix}5\\7\end{bmatrix}+\begin{bmatrix}3.5\\5\end{bmatrix}+\begin{bmatrix}4.5\\5\end{bmatrix}+\begin{bmatrix}3.5\\4.5\end{bmatrix} \right\} = \begin{bmatrix}3.9\\5.1\end{bmatrix}$$



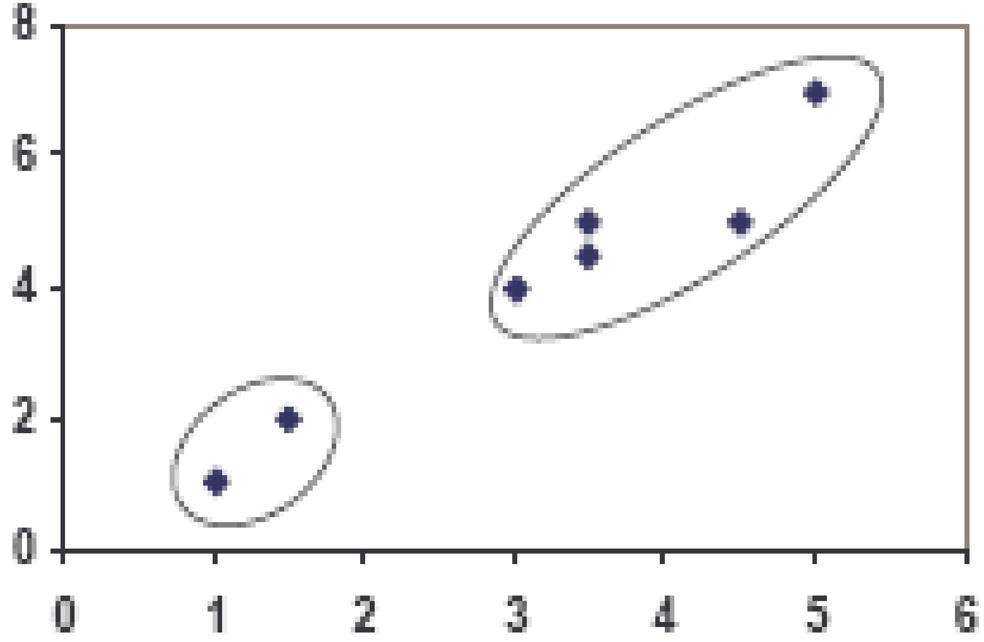
8
6
4
2
0
0
1
2
3
4
5
6

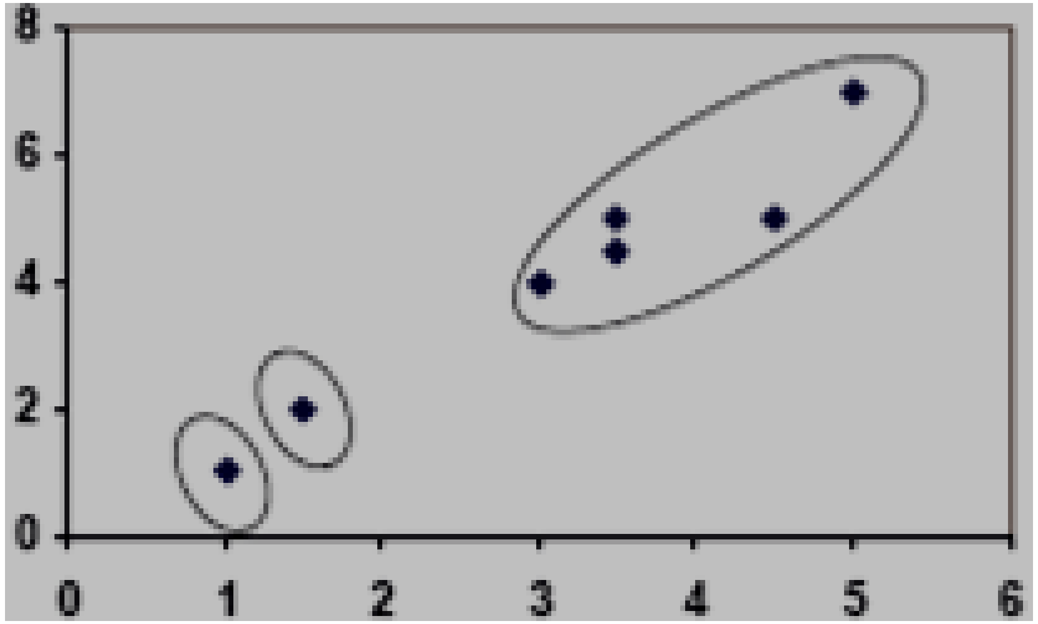
(with K=3)

Individual	Variable 1	Variable 2
1	1	1
2	1.5	2
3	3	4
4	5	7
5	3.5	5
6	4.5	5
7	3.5	4.5

individual	Distance from centroid 1(1,1)	Distance from centroid 2(1.5,2)	Distance from centroid 3(3,4)	Cluster
1(1,1)	0	1.11	3.61	1
2(1.5,2)	1.12	0	2.5	2
3(3,4)	3.61	2.5	0	3
4(5,7)	7.21	6.1	3.61	3
5(3.5,5)	4.72	3.61	1.12	3
6(4.5,5)	5.31	4.24	1.8	3
7(3.5,4.5)	4.3	3.2	0.71	3

K=2





K=3

Some Features:

1) Apriori knowledge of number of clusters

    - may be problem specific

    - otherwise may be determined by trial and error

2) Sensitivity to initialization

    - suffers from local minima problem

    - initial centroids should be well spread out

GA based K-Means Clustering:

Population	Chromosome	Fitness
1.	[1.0 1.0 5.0 7.0]	$\frac{1}{\mathrm{M}^{(1)}}$
2.	[3.0 4.0 3.5 4.5]	$\frac{1}{\mathrm{M}^{(2)}}$
.	.	.
.	.	.
P	.	.

Computational Steps :

1) Set K (no. of clusters)

2) Randomly pick 'K' points from the data set and form a chromosome

Repeat 'P' times (P=population size)

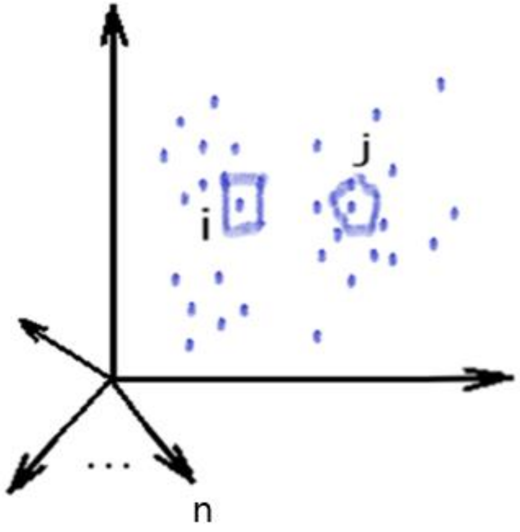
3) Assign each point in the dataset to the cluster whose centroid is nearest

4) After assignment of each point update the chromosomes

5) Compute the clustering metric for each chromosome as

$$M = \sum_{j=1}^{K} M_j$$

$$M_j = \sum_{\underline{x}_i \in C_j} \left\| \underline{x}_i - \underline{m}_j \right\|$$



6) Fitness of each chromosome is 1/M

Numerical Example:

Consider the following 4 data points in a 2D feature space. The data points are to be grouped in 2 hard clusters using a PSO based K-means clustering algorithm.

Considering randomly chosen initial cluster centres to be

$\begin{bmatrix}2.8\\1.7\end{bmatrix}$ & $\begin{bmatrix}7.0\\3.0\end{bmatrix}$; and $\begin{bmatrix}2.3\\5.1\end{bmatrix}$ & $\begin{bmatrix}2.8\\1.7\end{bmatrix}$ for a PSO population size of 2, update the population ONCE.

Assume the PSO parameters to be

w=1.0, c1=c2=2.0, r1=0.3 and r2=0.6

and initial velocities of the particles to be zero.

S. No.	Data point
1.	$\begin{bmatrix}2.3\\5.1\end{bmatrix}$
2.	$\begin{bmatrix}2.8\\1.7\end{bmatrix}$
3.	$\begin{bmatrix}6.7\\4.1\end{bmatrix}$
4.	$\begin{bmatrix}7.0\\3.0\end{bmatrix}$

Particle	Data Points	Dist. from $\underline{c}_1$ (i.e. $\begin{bmatrix} 2.8 \\ 1.7 \end{bmatrix}$)	Dist. from $\underline{c}_2$ (i.e. $\begin{bmatrix} 7 \\ 3 \end{bmatrix}$)	Cluster 1.	Cluster 2
1.> $\underline{X}_1$ (i.e. $\begin{bmatrix} 2.8 \\ 1.7 \\ 7.0 \\ 3.0 \end{bmatrix}$)	1.> $\underline{x}_1$	3.44	5.15	Data pt. 1.> & 2.>	Data pt. 3.> & 4.>
	2.> $\underline{x}_2$	0	4.40		
	3.> $\underline{x}_3$	4.58	1.14		
	4.> $\underline{x}_4$	4.40	0		

updating the 1st particle $\rightarrow$ $\begin{bmatrix} \frac{2.3+2.8}{2} \\ \frac{5.1+1.7}{2} \\ \frac{6.7+7}{2} \\ \frac{4.1+3}{2} \end{bmatrix} = \begin{bmatrix} 2.55 \\ 3.4 \\ 6.85 \\ 3.55 \end{bmatrix}$

Fitness of 1st particle', $F_1 = \frac{1}{M_1 + M_2}$ where

$$M_1 = \| \underline{x}_1 - \underline{c}_1 \| + \| \underline{x}_2 - \underline{c}_1 \|$$; $\underline{x}_1, \underline{x}_2$ being data points belonging to cluster 1.

$$= \left\| \begin{bmatrix} 2.3 \\ 5.1 \end{bmatrix} - \begin{bmatrix} 2.55 \\ 3.4 \end{bmatrix} \right\| + \left\| \begin{bmatrix} 2.8 \\ 1.7 \end{bmatrix} - \begin{bmatrix} 2.55 \\ 3.4 \end{bmatrix} \right\|$$

$$= 1.72 + 1.72 = 3.44.$$

& $$M_2 = \| \underline{x}_3 - \underline{c}_2 \| + \| \underline{x}_4 - \underline{c}_2 \|$$; $\underline{x}_3, \underline{x}_4$ being data points belonging to cluster 2.

$$= \left\| \begin{bmatrix} 6.7 \\ 4.1 \end{bmatrix} - \begin{bmatrix} 6.85 \\ 3.55 \end{bmatrix} \right\| + \left\| \begin{bmatrix} 7.0 \\ 3.0 \end{bmatrix} - \begin{bmatrix} 6.85 \\ 3.55 \end{bmatrix} \right\|$$

$$= 0.57 + 0.57 = 1.14.$$

$$\therefore F_1 = \frac{1}{3.44 + 1.14} = 0.2183.$$

Particle 1 is G_{best} i.e. $G_{best} = \begin{bmatrix} 2.55 \\ 3.4 \\ 6.85 \\ 3.55 \end{bmatrix}$

$$\underline{X}_1 \text{ (new)} = \begin{bmatrix} 2.55 \\ 3.4 \\ 6.85 \\ 3.55 \end{bmatrix} \qquad \underline{X}_2 \text{ (new)} = \begin{bmatrix} 2.16 \\ 3.16 \\ 7.24 \\ 3.79 \end{bmatrix}$$

Fuzzy C-Means Clustering

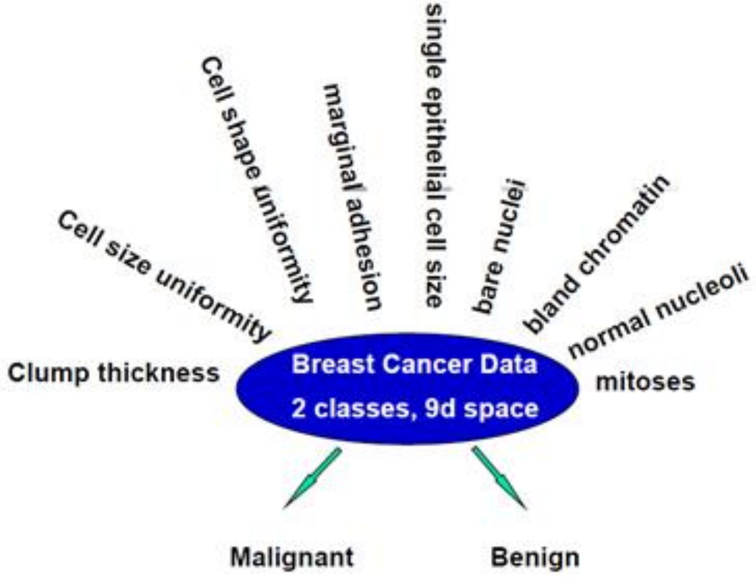
- ✓ Every point in the feature space belongs to all C clusters with varied membership grades
- ✓ Membership or partition matrix U takes the form:
- ✓ By Bezdek in late 1970s
- ✓ Natural clustering may not be unique

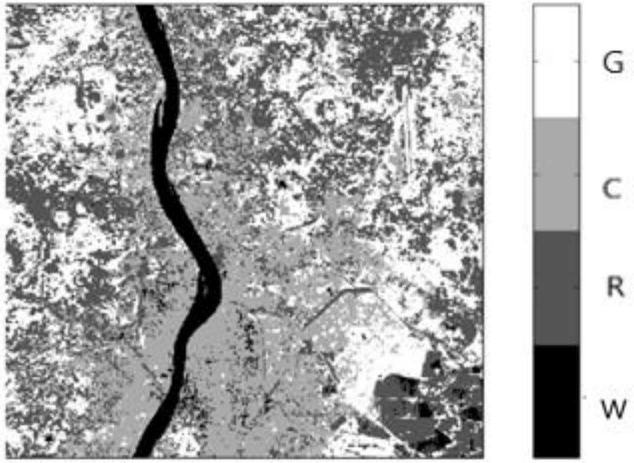
$$U = \begin{array}{c} \underline{x}_1 \\ \underline{x}_2 \\ \\ \vdots \\ \\ \underline{x}_N \end{array} \begin{array}{c} \begin{array}{cc} C1 & C2 \end{array} \\ \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \\ 0.6 & 0.4 \\ \cdot & \cdot \\ \cdot & \cdot \\ 0.1 & 0.9 \end{bmatrix} \end{array}$$

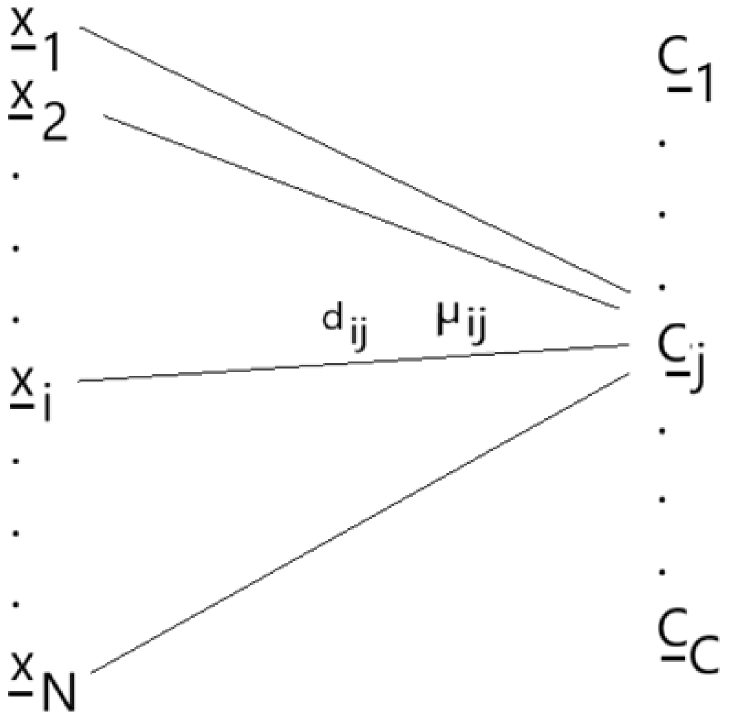
✓ Medical diagnosis

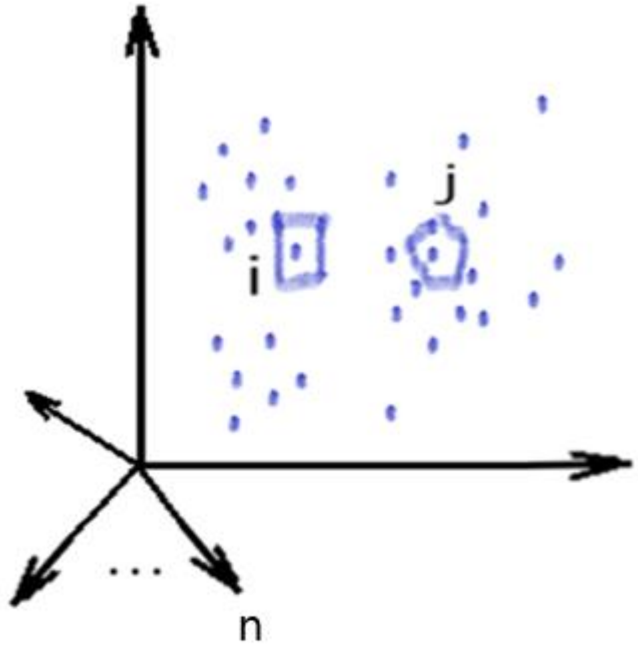
✓ Image segmentation

✓ Structural health monitoring of structures



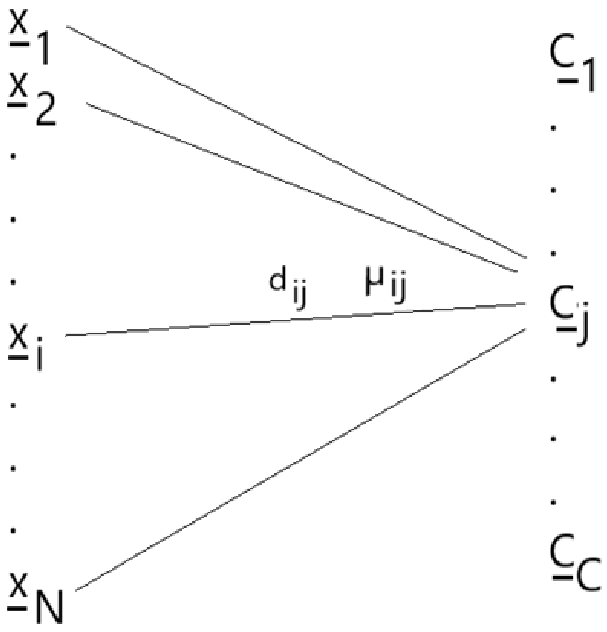






Dissimilarity, $F = \sum_{j=1}^{C} \sum_{i=1}^{N} \mu_{ij}^2 \, d_{ij}^2$ where $\sum_{j=1}^{C} \mu_{ij} = 1.0; \quad \forall\, i = 1, 2, \ldots N$

$$\mathcal{L} = \sum_{j=1}^{C} \sum_{i=1}^{N} \mu_{ij}^2 \left\| \underline{C}_j - \underline{x}_i \right\|^2 + \sum_{i=1}^{N} \lambda_i \left\{ \sum_{j=1}^{C} \mu_{ij} - 1.0 \right\}$$



$$\underline{C}_j = \frac{\sum_N \ \mu_{ij}^2 \ \underline{x}_i}{\sum_N \ \mu_{ij}^2}$$

$$\mu_{ij} = \frac{1}{\sum_{m=1}^{C} \left(\frac{d_{ij}}{d_{im}} \right)^2} \quad \text{where,} \quad d_{ij} = \left\| \underline{C}_j - \underline{x}_i \right\|$$

$$\mu_{ij} = \frac{1}{\left(\frac{d_{ij}^2}{d_{i1}^2} + \frac{d_{ij}^2}{d_{i2}^2} + \ \dots + \frac{d_{ij}^2}{d_{ij}^2} + \ \dots + \frac{d_{ij}^2}{d_{iC}^2} \right)}$$

Steps for FCM Algorithm:

1) Choose the number of clusters C

$$F = \sum_{j=1}^{C} \sum_{i=1}^{N} \mu_{ij}^{g} \, d_{ij}^{2}$$

2) Choose a cluster fuzziness level (g > 1)

3) Initialize the membership or partition matrix U at random

(satisfying the constraint that row sum = 1)

4) Compute the cluster centers as

$$\underline{C}_j = \frac{\sum_N \mu_{ij}^{g} \, \underline{x}_i}{\sum_N \mu_{ij}^{g}}$$

5) Compute the Euclidean distances, $d_{ij} = \left\| \underline{C}_j - \underline{x}_i \right\|$

6) Update the membership matrix as

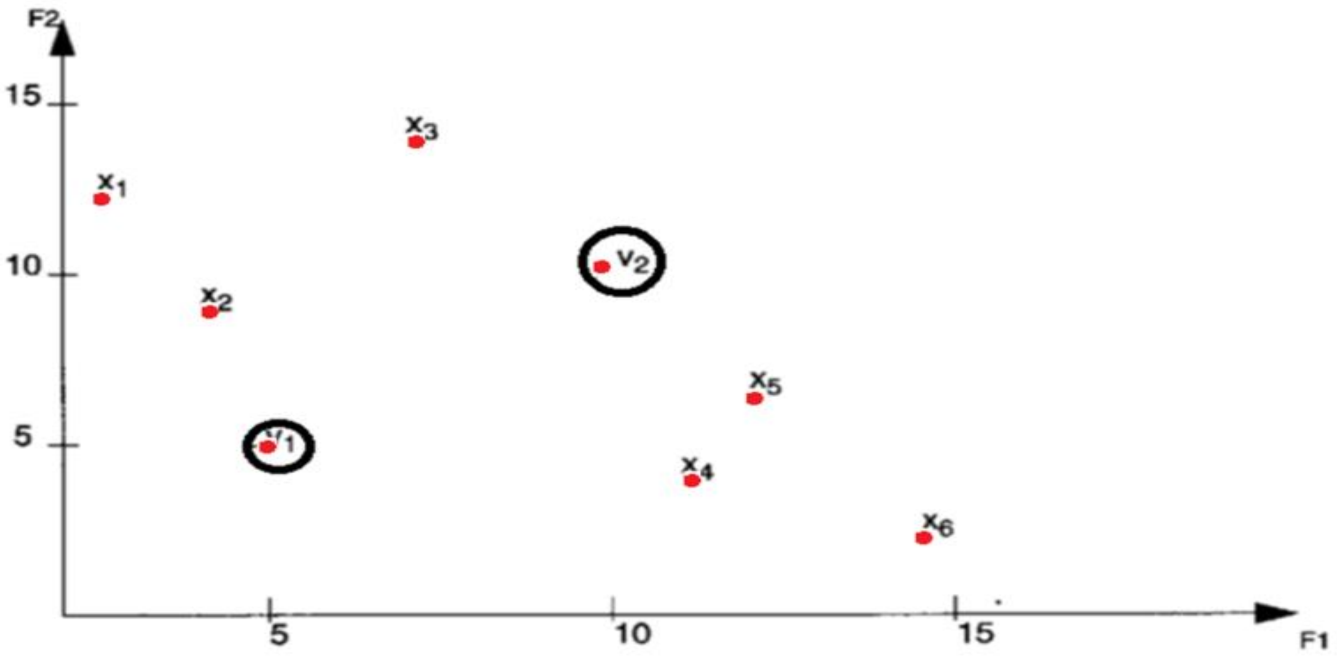
$$\mu_{ij} = \frac{1}{\sum_{m=1}^{C} \left(\frac{d_{ij}}{d_{im}} \right)^{2/(g-1)}}$$

if $d_{ij} = 0$, then $\mu_{ij} = 1$

7) Repeat steps 4, 5 and 6 till change in μ_{ij} values fall below some small threshold

Numerical Example: (C=2, n=2, N=6)

Membership value	X_1 (2,12)	X_2 (4,9)	X_3 (7,13)	X_4 (11,5)	X_5 (12,7)	X_6 (14,4)
C_1 (5,5)	0.5397	0.6852	0.2093	0.4194	0.197	0.3881
C_2 (10,10)	0.4603	0.3148	0.7907	0.5806	0.803	0.6119



$$
\begin{array}{cc}
C1 & C2 \\
\end{array}
$$

$$
U = \begin{bmatrix} 0.5397 & 0.4603 \\ 0.6852 & 0.3148 \\ 0.2093 & 0.7907 \\ 0.4194 & 0.5806 \\ 0.1970 & 0.8030 \\ 0.3881 & 0.6119 \end{bmatrix}
$$

	X_1 (2,12)	X_2 (4,9)	X_3 (7,13)	X_4 (11,5)	X_5 (12,7)	X_6 (14,4)
C_1	**0.5397**	**0.6852**	**0.2093**	**0.4194**	**0.197**	**0.3881**
C_2	**0.4603**	**0.3148**	**0.7907**	**0.5806**	**0.803**	**0.6119**

Calculate coordinates of new centre of cluster 1

$$C_1 = \frac{(0.5397)^2 \times (2,12) + (0.6852)^2 \times (4,9) + (0.2093)^2 \times (7,13) + (0.4194)^2 \times (11,5) + (0.197)^2 \times (12,7) + (0.3881)^2 \times (14,4)}{(0.5397)^2 + (0.6852)^2 + (0.2093)^2 + (0.4194)^2 + (0.197)^2 + (0.3881)^2}$$

$$= (6.6273, 9.1484)$$

	X_1 (2,12)	X_2 (4,9)	X_3 (7,13)	X_4 (11,5)	X_5 (12,7)	X_6 (14,4)
C_1	0.5397	0.6852	0.2093	0.4194	0.197	0.3881
C_2	0.4603	0.3148	0.7907	0.5806	0.803	0.6119

Calculate coordinates of new centre of cluster 2

$$C_2 = \frac{\begin{gathered}(0.4603)^2 \times (2,12) + (0.3148)^2 \times (4,9) + (0.7907)^2 \times (7,13) + \\ (0.5806)^2 \times (11,5) + (0.803)^2 \times (12,7) + (0.6119)^2 \times (14,4)\end{gathered}}{(0.4603)^2 + (0.3148)^2 + (0.7907)^2 + (0.5806)^2 + (0.803)^2 + (0.6119)^2}$$

$$= (9.7374,\ 8.4887)$$

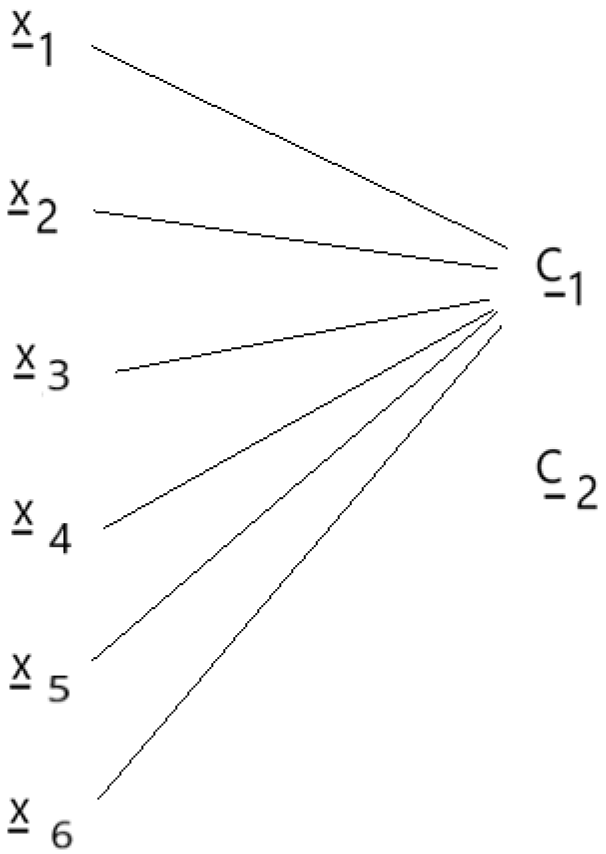
X_1 (2,12)	X_2 (4,9)	X_3 (7,13)	X_4 (11,5)	X_5 (12,7)	X_6 (14,4)

$$U = \begin{array}{c} \begin{array}{cc} C1 & C2 \end{array} \\ \begin{bmatrix} 0.5397 & 0.4603 \\ 0.6852 & 0.3148 \\ 0.2093 & 0.7907 \\ 0.4194 & 0.5806 \\ 0.1970 & 0.8030 \\ 0.3881 & 0.6119 \end{bmatrix} \end{array}$$

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

$$\mu_{21} = \frac{1}{\sum_C \left(\frac{d_{ij}}{d_{im}}\right)^{2/(g-1)}} = \frac{1}{1 + \left(\frac{d_{21}}{d_{22}}\right)^2}$$

$$\mu_{31} = \frac{1}{\sum_C \left(\frac{d_{ij}}{d_{im}}\right)^{2/(g-1)}} = \frac{1}{1 + \left(\frac{d_{31}}{d_{32}}\right)^2}$$



$$\mu_{ij} = \frac{1}{\left(\frac{d_{ij}^2}{d_{i1}^2} + \frac{d_{ij}^2}{d_{i2}^2} + \ \dots + \frac{d_{ij}^2}{d_{ij}^2} + \ \dots + \frac{d_{ij}^2}{d_{iC}^2} \right)}$$

X_1 (2,12)	X_2 (4,9)	X_3 (7,13)	X_4 (11,5)	X_5 (12,7)	X_6 (14,4)

$$U = \begin{array}{cc} C1 & C2 \\ \end{array}$$
$$U = \begin{bmatrix} 0.5397 & 0.4603 \\ 0.6852 & 0.3148 \\ 0.2093 & 0.7907 \\ 0.4194 & 0.5806 \\ 0.1970 & 0.8030 \\ 0.3881 & 0.6119 \end{bmatrix}$$

$$\mu_{41} = \frac{1}{\sum_C \left(\frac{d_{ij}}{d_{im}}\right)^{2/(g-1)}} = \frac{1}{1 + \left(\frac{d_{41}}{d_{42}}\right)^2}$$

$$\mu_{51} = \frac{1}{\sum_C \left(\frac{d_{ij}}{d_{im}}\right)^{2/(g-1)}} = \frac{1}{1 + \left(\frac{d_{51}}{d_{52}}\right)^2}$$

$$\mu_{61} = \frac{1}{\sum_C \left(\frac{d_{ij}}{d_{im}}\right)^{2/(g-1)}} = \frac{1}{1 + \left(\frac{d_{61}}{d_{62}}\right)^2}$$

X_1 (2,12)	X_2 (4,9)	X_3 (7,13)	X_4 (11,5)	X_5 (12,7)	X_6 (14,4)

New centers :

c_1 = (6.6273,9.1484)
c_2 = (9.7344, 8.4887)

$$d_{ij} = \|\underline{C}_j - \underline{x}_i\|$$

$$d_{11}^2 = \|\underline{C}_1 - \underline{x}_1\|^2 = \left\|\begin{bmatrix} 6.6273 - 2 \\ 9.1484 - 12 \end{bmatrix}\right\|^2 = 29.5435$$

$$d_{12}^2 = \|\underline{C}_2 - \underline{x}_1\|^2 = \left\|\begin{bmatrix} 9.7344 - 2 \\ 8.4887 - 12 \end{bmatrix}\right\|^2 = 72.1502$$

$$\mu_{11} = \frac{1}{1 + \left(\frac{d_{11}}{d_{12}}\right)^2} = 0.7095$$

$$d_{21}^2 = \|\underline{C}_1 - \underline{x}_2\|^2 = \left\|\begin{bmatrix} 6.6273 - 4 \\ 9.1484 - 9 \end{bmatrix}\right\|^2 = 6.9247$$

$$d_{22}^2 = \|\underline{C}_2 - \underline{x}_2\|^2 = \left\|\begin{bmatrix} 9.7344 - 4 \\ 8.4887 - 9 \end{bmatrix}\right\|^2 = 33.1448$$

$$\mu_{21} = 0.8272$$

X_1 (2,12)	X_2 (4,9)	X_3 (7,13)	X_4 (11,5)	X_5 (12,7)	X_6 (14,4)

New centers :

c_1 = (6.6273,9.1484)
c_2 = (9.7344, 8.4887)

$$d_{ij} = \left\| \underline{C}_j - \underline{x}_i \right\|$$

$$d_{31}^2 = \left\| \underline{C}_1 - \underline{x}_3 \right\|^2 = \left\| \begin{bmatrix} 6.6273 - 7 \\ 9.1484 - 13 \end{bmatrix} \right\|^2 = 14.9737$$

$$d_{32}^2 = \left\| \underline{C}_2 - \underline{x}_3 \right\|^2 = \left\| \begin{bmatrix} 9.7344 - 7 \\ 8.4887 - 13 \end{bmatrix} \right\|^2 = 27.8288$$

$$\mu_{31} = 0.6502$$

$$d_{41}^2 = \left\| \underline{C}_1 - \underline{x}_4 \right\|^2 = \left\| \begin{bmatrix} 6.6273 - 11 \\ 9.1484 - 5 \end{bmatrix} \right\|^2 = 36.3297$$

$$d_{42}^2 = \left\| \underline{C}_2 - \underline{x}_4 \right\|^2 = \left\| \begin{bmatrix} 9.7344 - 11 \\ 8.4887 - 5 \end{bmatrix} \right\|^2 = 13.7728$$

$$\mu_{41} = 0.2749$$

X_1 (2,12)	X_2 (4,9)	X_3 (7,13)	X_4 (11,5)	X_5 (12,7)	X_6 (14,4)

New centers :

c_1 = (6.6273,9.1484)
c_2 = (9.7344, 8.4887)

$$d_{ij} = \left\| \underline{C}_j - \underline{x}_i \right\|$$

$$d_{51}^2 = \left\| \underline{C}_1 - \underline{x}_5 \right\|^2 = \left\| \begin{bmatrix} 6.6273 - 12 \\ 9.1484 - 7 \end{bmatrix} \right\|^2 = 33.4815$$

$$d_{52}^2 = \left\| \underline{C}_2 - \underline{x}_5 \right\|^2 = \left\| \begin{bmatrix} 9.7344 - 12 \\ 8.4887 - 7 \end{bmatrix} \right\|^2 = 7.3492$$

$$\mu_{51} = 0.1800$$

$$d_{61}^2 = \left\| \underline{C}_1 - \underline{x}_6 \right\|^2 = \left\| \begin{bmatrix} 6.6273 - 14 \\ 9.1484 - 4 \end{bmatrix} \right\|^2 = 80.8627$$

$$d_{62}^2 = \left\| \underline{C}_2 - \underline{x}_6 \right\|^2 = \left\| \begin{bmatrix} 9.7344 - 14 \\ 8.4887 - 4 \end{bmatrix} \right\|^2 = 38.3438$$

$$\mu_{61} = 0.3217$$

C1 (6.6273,9.1484)   C2 (9.7344, 8.4887)

$$U=\begin{bmatrix} 0.7095 & 0.2905 \\ 0.8272 & 0.1728 \\ 0.6502 & 0.3498 \\ 0.2749 & 0.7251 \\ 0.1800 & 0.8200 \\ 0.3217 & 0.6783 \end{bmatrix}$$

$$U=\begin{bmatrix} 0.5397 & 0.4603 \\ 0.6852 & 0.3148 \\ 0.2093 & 0.7907 \\ 0.4194 & 0.5806 \\ 0.1970 & 0.8030 \\ 0.3881 & 0.6119 \end{bmatrix}$$

Some Features:

1) Apriori knowledge of number of clusters

   - may be problem specific

   - may be determined by trial and error

   - some other simpler clustering algorithm may be used to have a rough idea about the number of clusters

2) Sensitivity to initialization

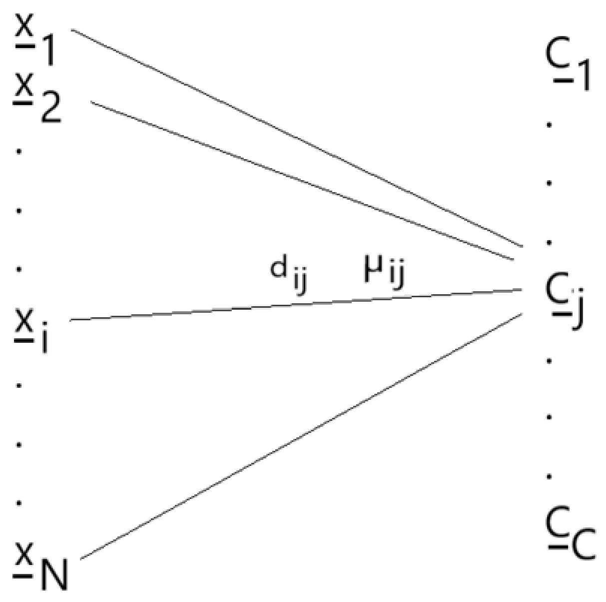
   - suffers from local minima problem

3) Higher ‘g’ value implies higher fuzziness; results in slower convergence

   (g=1 implies hard clustering)

GA based FCM Algorithm:

- ✓ Initialization problem is overcome
- ✓ Bezdek, 1999



Dissimilarity, $$F = \sum_{j=1}^{C} \sum_{i=1}^{N} \mu_{ij}^{2}\, d_{ij}^{2}$$

$$\mu_{ij} = \frac{1}{\sum_{m=1}^{C} \left(\frac{d_{ij}}{d_{im}} \right)^{2}}$$

$$F = \sum_{i=1}^{N} \left\{ \sum_{j=1}^{C} d_{ij}^{-1} \right\}^{-1}$$

Computational Steps:

1) Choose the number of clusters C

2) Choose a population of chromosomes where each chromosome represents a set of cluster centres i.e. $\{ \underline{C}_1 \ \underline{C}_2 \ \ldots \ \underline{C}_j \ \ldots \ \underline{C}_C \}$

3) Compute the distances and therefore the dissimilarity and fitness $\frac{1}{F}$

$$F = \sum_{i=1}^{N} \left\{ \sum_{j=1}^{C} d_{ij}^{-1} \right\}^{-1}$$

4) After GA converges, assign membership values as

$$\mu_{ij} = \frac{1}{\sum_{m=1}^{C} \left(\frac{d_{ij}}{d_{im}} \right)^2}$$

Numerical Example:

Consider the following 4 data points in a 2D feature space. The data points are to be grouped in 2 soft clusters using a PSO based FCM clustering algorithm.

Considering randomly chosen initial cluster centres to be

$\begin{bmatrix} 2.8 \\ 1.7 \end{bmatrix}$ & $\begin{bmatrix} 7.0 \\ 3.0 \end{bmatrix}$; and $\begin{bmatrix} 2.3 \\ 5.1 \end{bmatrix}$ & $\begin{bmatrix} 2.8 \\ 1.7 \end{bmatrix}$ for a PSO population size of 2, update the population ONCE.

Also compute the initial and updated partition matrix.

Assume the PSO parameters to be

w=1.0, c1=c2=2.0, r1=0.3 and r2=0.6

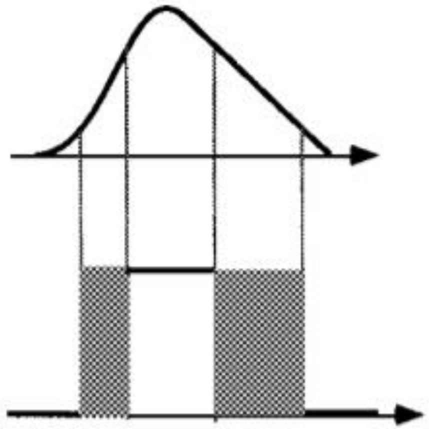
and initial velocities of the particles to be zero.

S. No.	Data point
1.	$\begin{bmatrix} 2.3 \\ 5.1 \end{bmatrix}$
2.	$\begin{bmatrix} 2.8 \\ 1.7 \end{bmatrix}$
3.	$\begin{bmatrix} 6.7 \\ 4.1 \end{bmatrix}$
4.	$\begin{bmatrix} 7.0 \\ 3.0 \end{bmatrix}$

Variable String Length GA for Automatic Clustering (2001)

- ✓ GA to evolve the number of clusters
- ✓ Requires GA string length to vary
- ✓ Applicable to K-means algorithm also

Some Further Topics:



- ✓ General and Interval Type-2 Fuzzy Logic
- ✓ Shadowed Fuzzy Sets
- ✓ Intuitionistic Fuzzy Logic
- ✓ Mediative Fuzzy Logic
- ✓ Rough Sets
- ✓ Fuzzy- Rough Set Hybridization