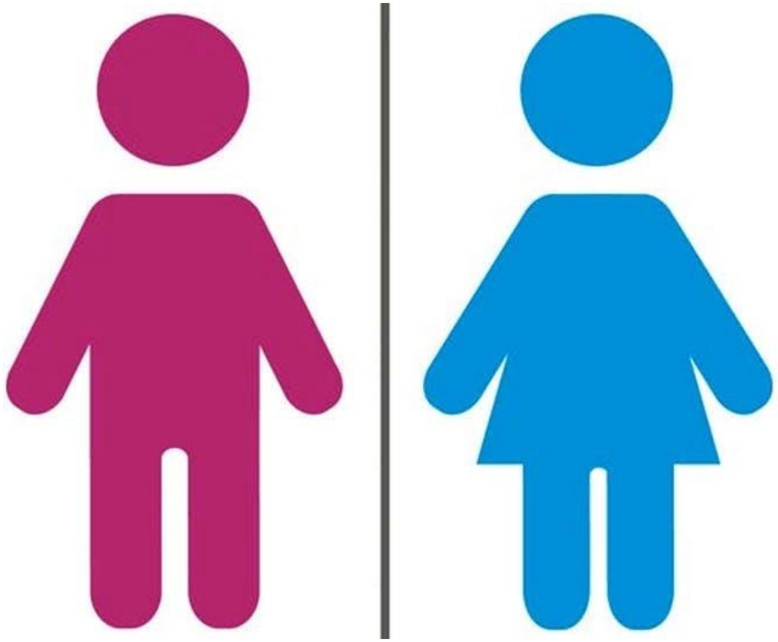
DATA MINING

Pertemuan XIII

Fuzzy Logic

Introduction

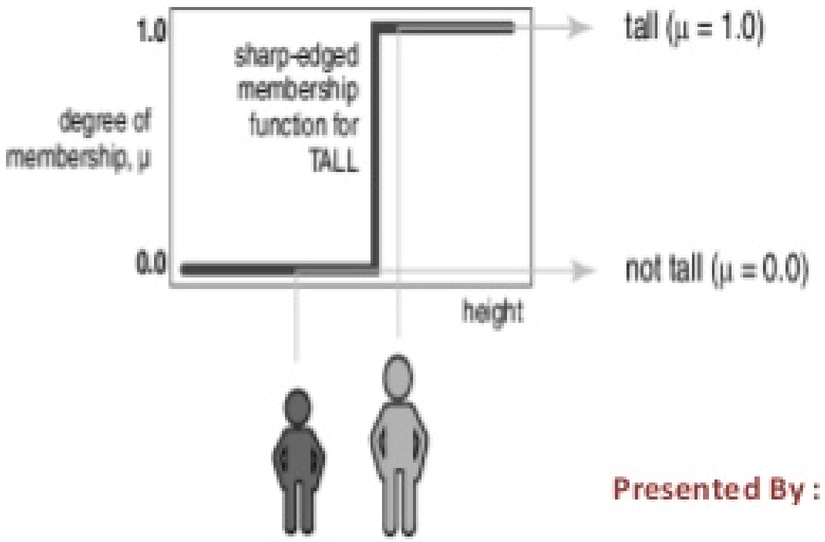
- The word **fuzzy** refers to things which are not clear or are vague.
- Any event, process, or function that is changing continuously cannot always be defined as either true or false, which means that we need to define such activities in a Fuzzy manner.

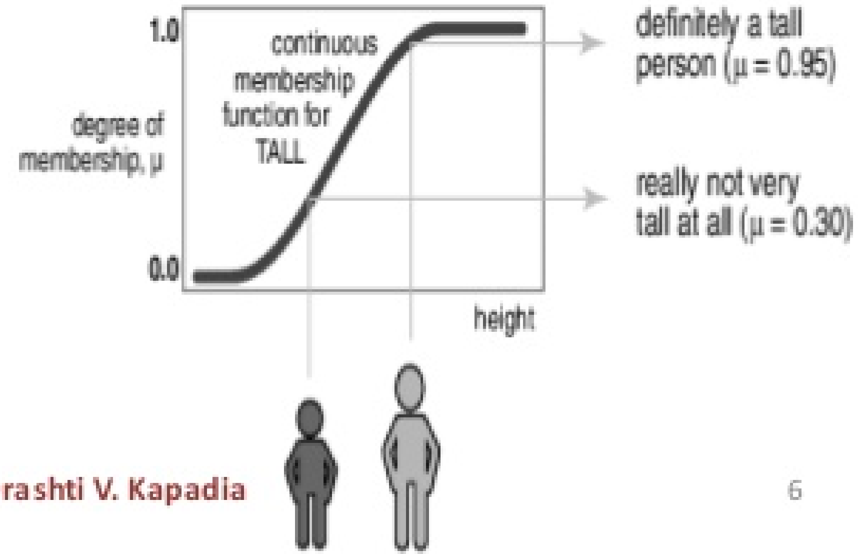


Male or female?

Crisp Set vs Fuzzy Set

- The x-axis represents the universe of discourse – the range of all possible values applicable to a chosen variable. The variable is the man height. The universe of men's heights consists of all tall men
- The y-axis represents the membership value of the fuzzy set. The fuzzy set of "tall men" maps height values into corresponding membership values.





More on Crisp Set

- Everything is either true or false
- No uncertainty is allowed
- An item either is
  - entirely within a set, or
  - entirely not in a set
- The Law of the Excluded Middle
  - X must be either in set A or in set not-A
  - no middle ground is allowed
- Opposite sets (A and not-A) must between them contain everything

Fuzzy Set

- Items can belong to a fuzzy set to different degrees
  - ➔ In fuzzy set, each item has degrees of membership to each cluster
- Completely within a set is a membership degree of 1
- Completely outside a set is a membership degree of 0

Fuzzy Sets

Example

Let the values of temperature in °C under consideration be

$T = \{0, 5, 10, 15, 20, 25, 30, 35, 40\}$.

Then, the term *hot* can be defined by a fuzzy set as follows

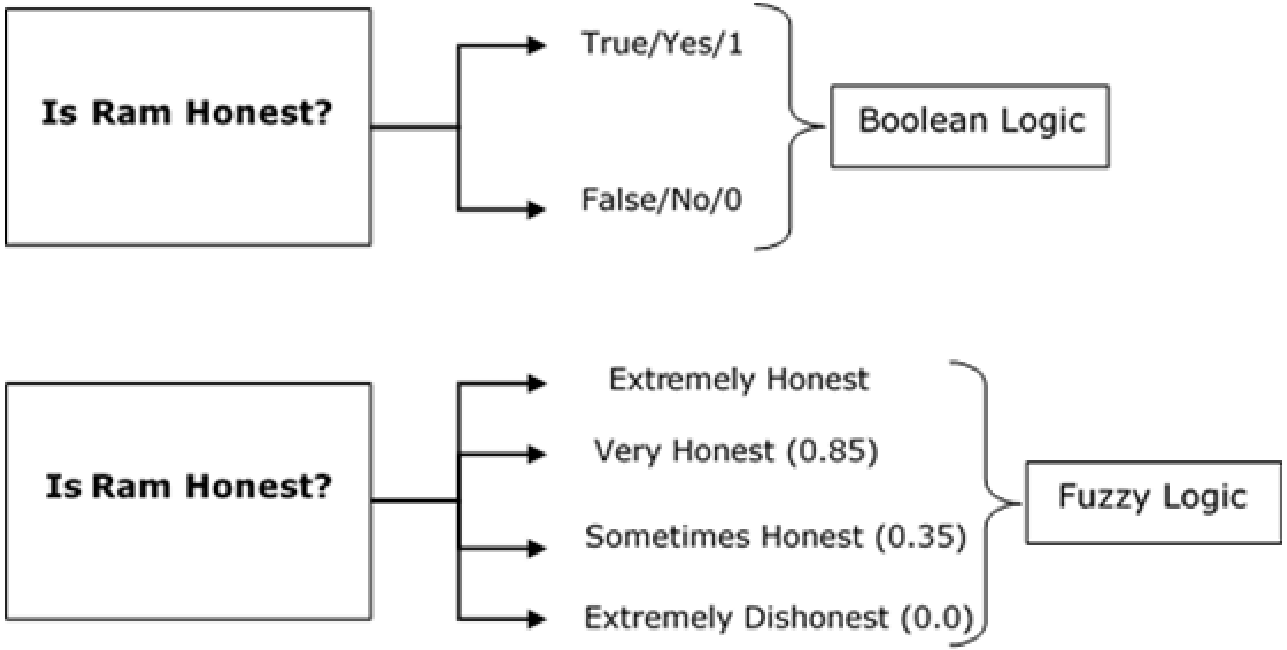
HOT = {(0,0), (5,0.1), (10,0.3), (15,0.5), (20,0.6), (25,0.7), (30,0.8), (35,0.9), (40,1.0)}.

This fuzzy set reflects the point of view that 0 °C is not hot at all, 5, 10, and 15 °C are somewhat hot, and 40 °C is indeed hot. Another person could have defined the set differently.

Fuzzy Logic

What is Fuzzy Logic?

- In fuzzy mathematics, **fuzzy logic** is a form of many-valued logic in which the truth values of variables may be any real number between 0 and 1 both inclusive. It is employed to handle the concept of partial truth, where the truth value may range between completely true and completely false. By contrast, in Boolean logic, the truth values of variables may only be the integer values 0 or 1.



fuzzy logic is not logic that is fuzzy, but logic that is used to describe fuzziness

Fuzzy Logic – Set Theory

- Same operations and function as in crisp logic:
  - AND, OR, NOT
- Crisp logical functions
  - AND true is both parameters are true
  - OR true if either parameter is true
  - NOT reverses truth of argument

- Fuzzy Logic: Must deal with degrees of truth rather than absolute truths
- Fuzzy logic is a superset of crisp (Boolean) logic

Crisp Logic vs Fuzzy Logic

- AND function - crisp version

AND		
A	B	A AND B
0	0	0
0	1	0
1	0	0
1	1	1

- AND function - fuzzy version
  - take the minimum of the two arguments

A = 0,2 / TRUE + 0,8 / FALSE
B = 0.6 / TRUE + 0,4 / FALSE

A and B = 0,2 / TRUE + 0,4 / FALSE

AND		
A	B	min(A,B)
0	0	0
0	1	0
1	0	0
1	1	1

Crisp Logic vs Fuzzy Logic

- OR function - crisp version

OR		
A	B	A OR B
0	0	0
0	1	1
1	0	1
1	1	1

- OR function - fuzzy version
  - take the maximum of the two arguments

A = 0,2 / TRUE + 0,8 / FALSE
B = 0.6 / TRUE + 0,4 / FALSE

A or B = 0,6 / TRUE + 0,8 / FALSE

OR		
A	B	max(A,C)
0	0	0
0	1	1
1	0	1
1	1	1

Crisp Logic vs Fuzzy Logic

- NOT function - crisp version

A	NOT A
0	1
1	0

- NOT function - fuzzy version
  - subtract the truth value from one

A	1-A
0	1
1	0

A = 0,2 / TRUE + 0,8 / FALSE

Not A = 0,8 / TRUE + 0,2 / FALSE

- Output of fuzzy logical functions are the same as crisp functions
- It just calculated differently
- Fuzzy logic handle *degrees* of truth, rather than *absolute* truths
- The basis of fuzzy rule based systems

Fuzzy Clustering

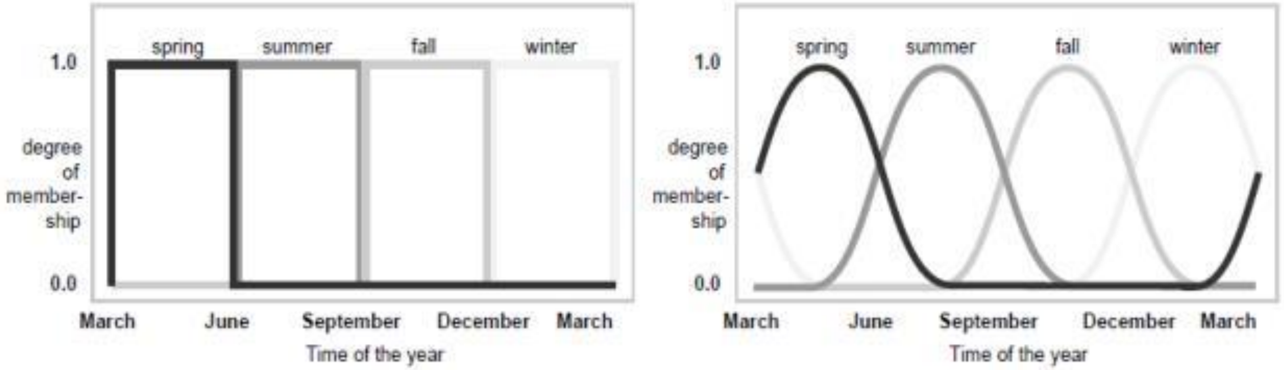
What is Fuzzy Clustering??

- Clustering divides data points into groups based in similarity between items and looks to find patterns or similarity between items in a set; Items in clusters should be as similar as possible to each other and as dissimilar as possible to items in other groups.
- Computationally, **it's much easier to create fuzzy boundaries than it is to settle on one cluster for one point.**
- Fuzzy clustering is a clustering method where **data points can belong in more than one group ("cluster")**.

Hard (Crisp) Cluster vs Soft (Fuzzy) Cluster

- In “hard” clustering, each data point can only be in one cluster.
- In “soft” or “fuzzy” clustering, data points can belong to more than one group.

Crisp Set vs Fuzzy Set

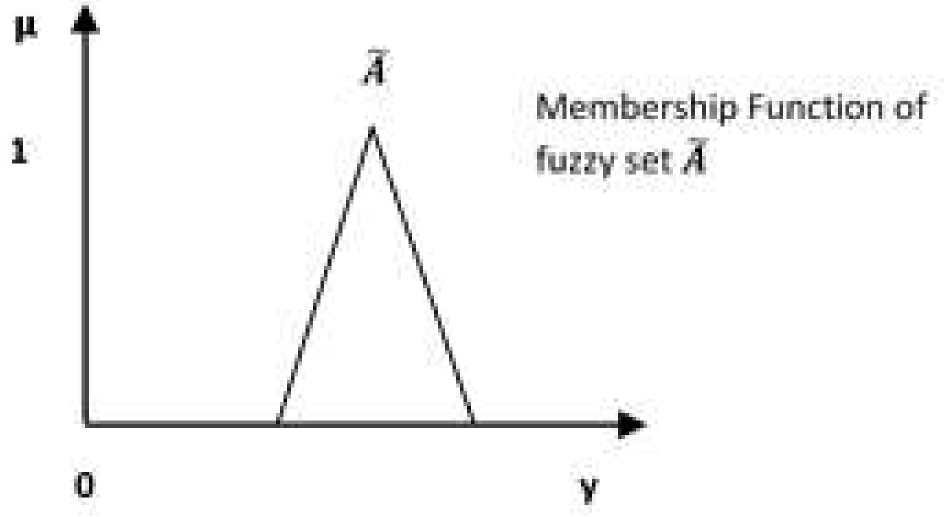


Membership

- Membership grades are assigned to each of the data points (tags).
- Membership grades indicate the degree to which data points belong to each cluster. Thus, points on the edge of a cluster, with lower membership grades, may be *in the cluster* to a lesser degree than points in the center of cluster.
- Degrees of membership must sum to 1
- An item can be both A and not-A to different degrees
- e.g. A to a degree of 0.8, not-A 0.2
- Degrees of membership are expressed with membership functions

Membership Function

- We already know that fuzzy logic is not logic that is fuzzy but logic that is used to describe fuzziness.
- This fuzziness is best characterized by its membership function.
- In other words, we can say that membership function represents the degree of truth in fuzzy logic.



Membership Function

Important Points of MF

- Membership functions were first introduced in 1965 by Lofti A. Zadeh in his first research paper “fuzzy sets”.
- Membership functions characterize fuzziness (i.e., all the information in fuzzy set), whether the elements in fuzzy sets are discrete or continuous.
- Membership functions can be defined as a technique to solve practical problems by experience rather than knowledge.
- Membership functions are represented by graphical forms.
- Rules for defining fuzziness are fuzzy too.

Features of Membership Functions

- **Core**

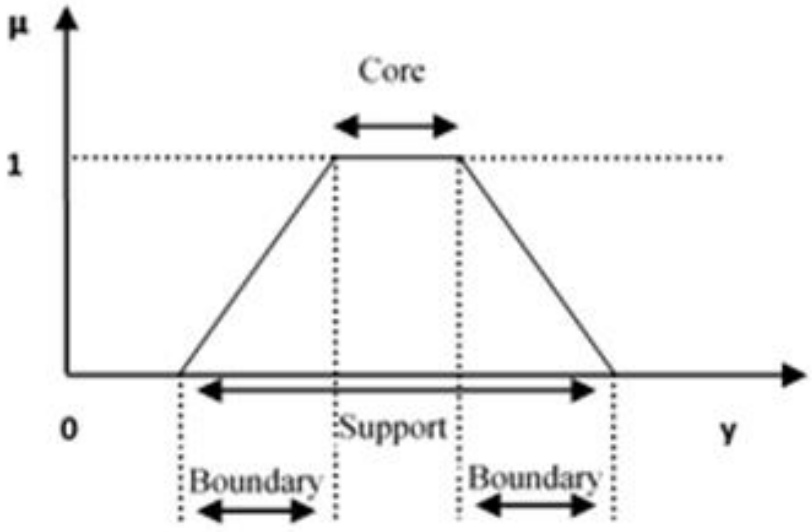
For any fuzzy set A˜, the core of a membership function is that region of universe that is characterize by full membership in the set.

- **Support**

For any fuzzy set A˜, the support of a membership function is the region of universe that is characterize by a nonzero membership in the set.

- **Boundary**

For any fuzzy set A˜, the boundary of a membership function is the region of universe that is characterized by a nonzero but incomplete membership in the set.



Features of Membership Function

Membership Functions

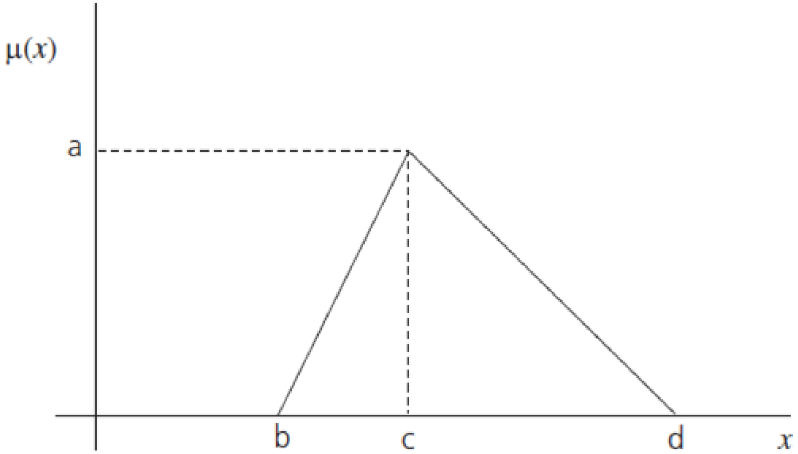
Fuzzy sets can also be defined by assigning a continuous function to describe the membership either analytically or graphically.

The triangular membership function

$$\mu(x) = a\,(b\text{-}x)/(b\text{-}c)\ ; \qquad b \geq x \leq c$$

$$= a\,(d\text{-}x)/(d\text{-}c)\ ; \qquad c \geq x \leq d$$

$$= 0\ ; \qquad \text{otherwise}$$



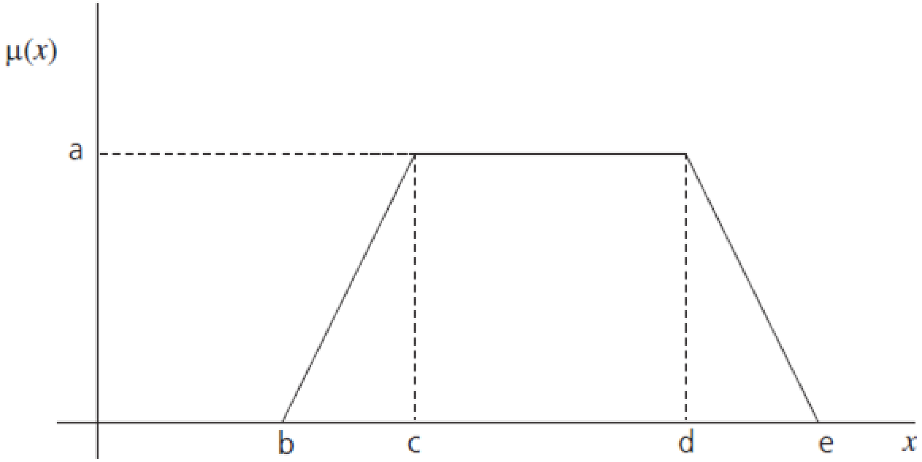
The trapezoidal function s

$$\mu(x) = a\,(b\text{-}x)/(b\text{-}c)\ ; \qquad b \geq x \leq c$$

$$= a\ ; \qquad c \geq x \leq d$$

$$= a\,(e\text{-}x)/(e\text{-}d)\ ; \qquad d \geq x \leq e$$

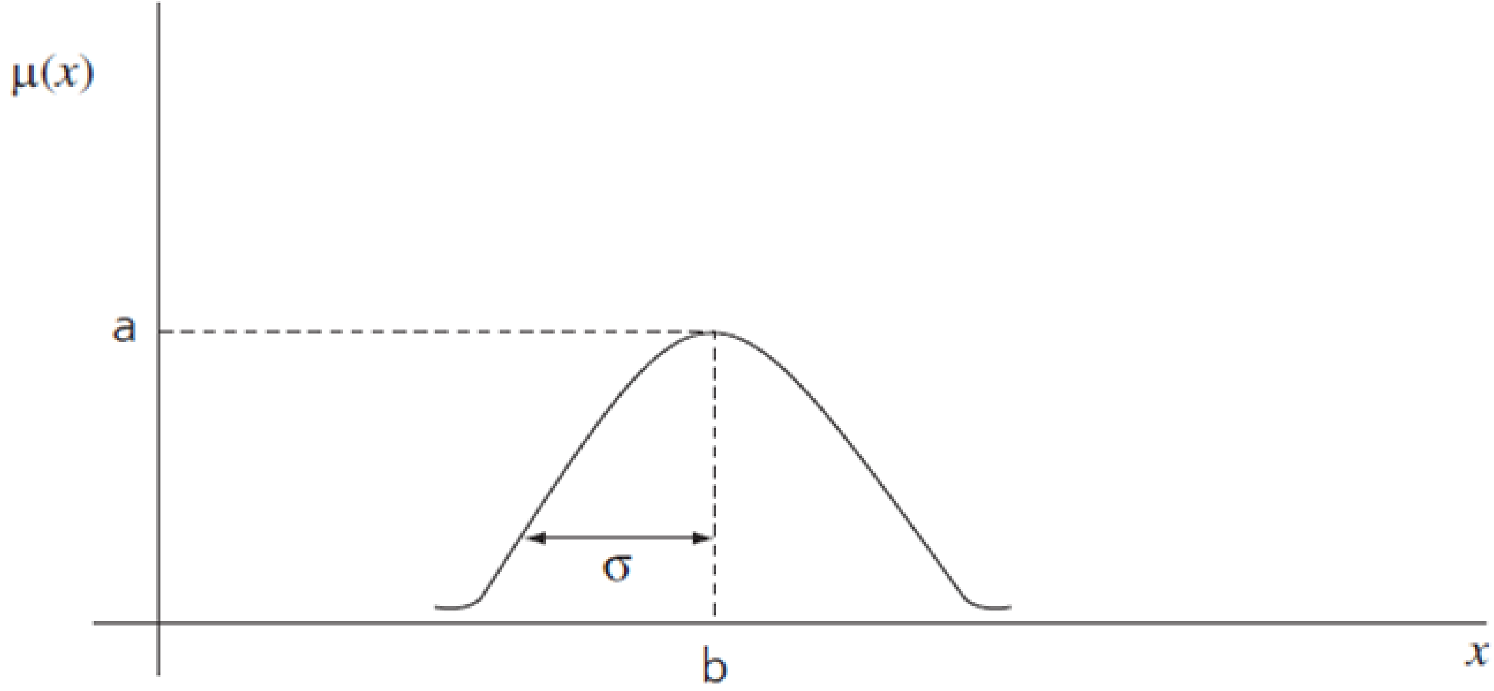
$$= 0\ ; \qquad \text{otherwise}$$



Membership Functions

The Gaussian function

$$\mu(x) = a \exp(- x\text{-}b)^2/2\sigma^2)$$

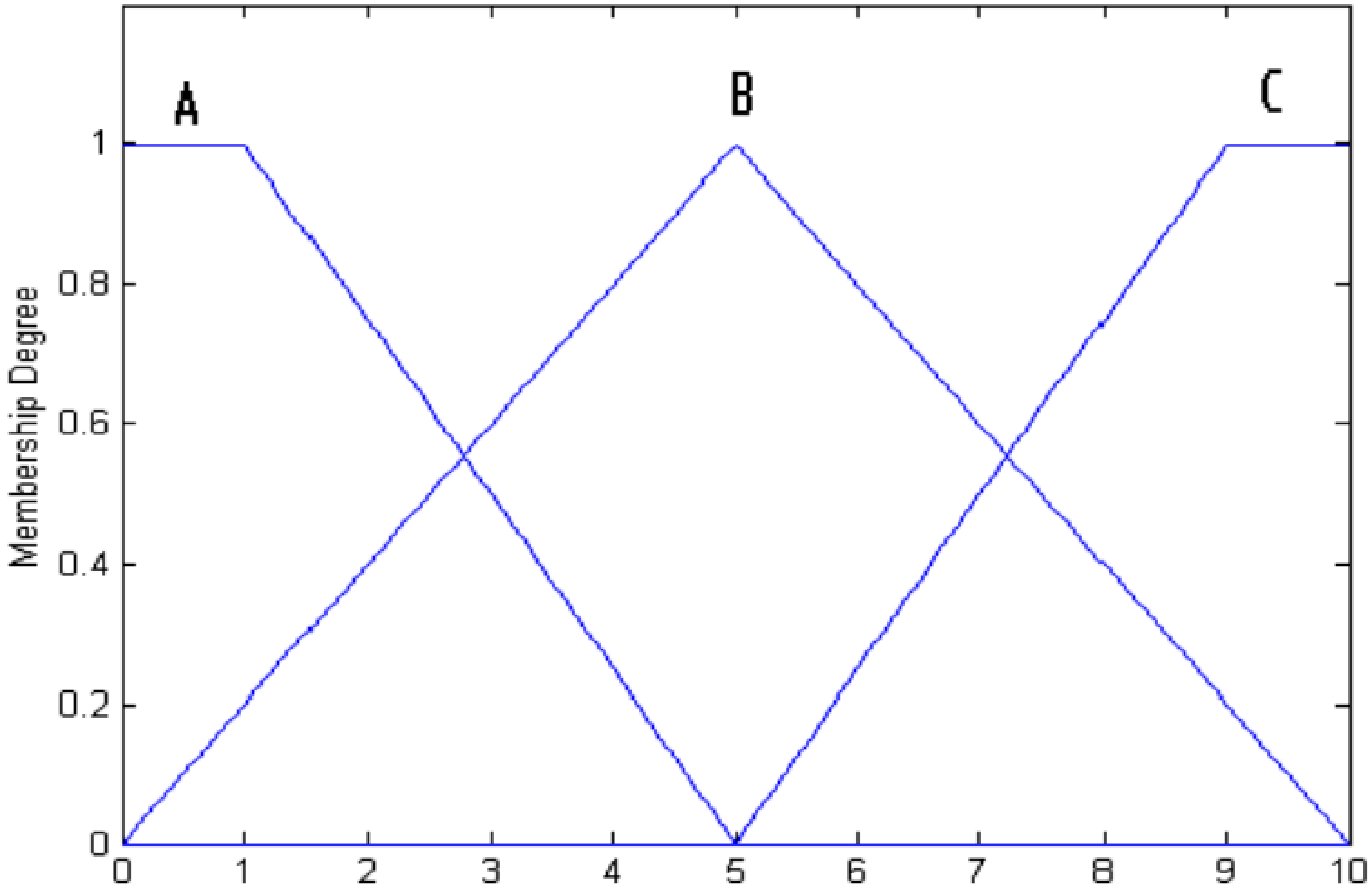


Membership Functions

- MF can also be represented by a set of ordered pairs

- Pairs are crisp-fuzzy values
  - A={(0,1.0),(1,1.0),(2,0.75),(3,0.5),(4,0.25),(5,0.0),(6,0.0),(7,0.0),(8,0.0),(9,0.0),(10,0.0)}
  - B={(0,0.0),(1,0.2),(2,0.4),(3,0.6),(4,0.8),(5,1.0),(6,0.8),(7,0.6),(8,0.4),(9,0.2),(10,0.0)}
  - C={(0,0.0),(1,0.0),(2,0.0),(3,0.0),(4,0.0),(5,0.0)(6,0.25),(7,0.5),(8,0.75),(9,1.0),(10,1.0)}

Membership Functions



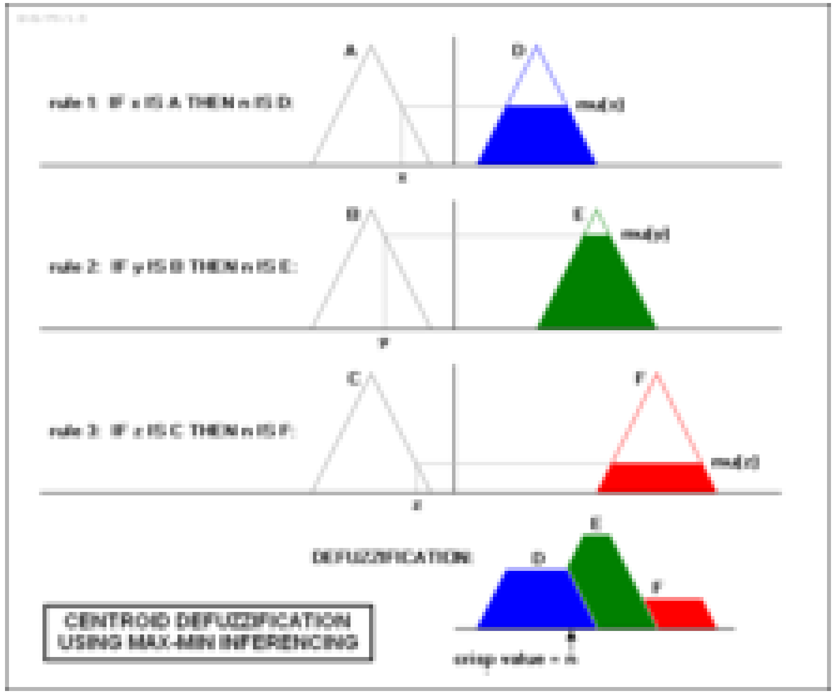
Fuzzification

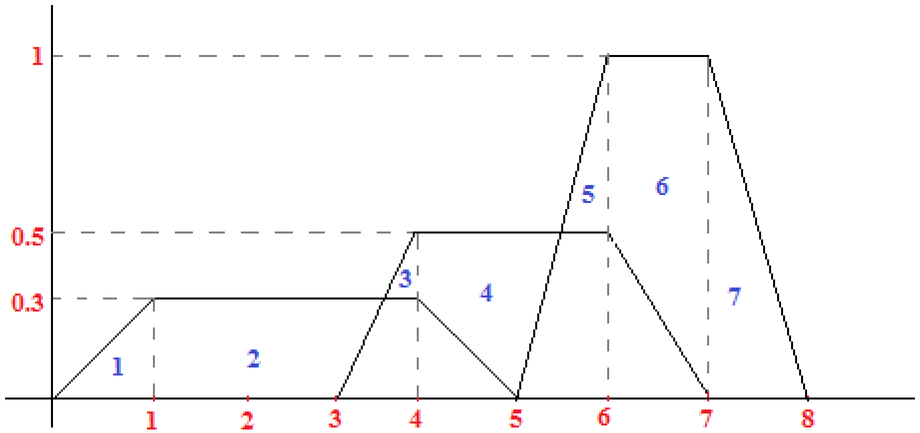
- Fuzzification is the process of assigning the numerical input of a system to fuzzy sets with some degree of membership.
- This degree of membership may be anywhere within the interval [0,1]. If it is 0 then the value does not belong to the given fuzzy set, and if it is 1 then the value completely belongs within the fuzzy set. Any value between 0 and 1 represents the degree of uncertainty that the value belongs in the set ➔ The value returned by a fuzzy MF
- Most variables in a fuzzy system have multiple MF attached to them
- Fuzzifying that variable involves passing the crisp value through each MF attached to that value

Defuzzification

- **Defuzzification** is the process of producing a quantifiable result in Crisp logic, given fuzzy sets and corresponding membership degrees. It is the process that maps a fuzzy set to a crisp set.
- For example, rules designed to decide how much pressure to apply might result in "Decrease Pressure (15%), Maintain Pressure (34%), Increase Pressure (72%)". Defuzzification is interpreting the membership degrees of the fuzzy sets into a specific decision or real value.

- The simplest but least useful defuzzification method is to choose the set with the highest membership, The problem with this approach is that it loses information. The rules that called for decreasing or maintaining pressure might as well have not been there in this case.
- A common and useful defuzzification technique is *center of gravity*. First, the results of the rules must be added together in some way. The most typical fuzzy set membership function has the graph of a triangle. Now, if this triangle were to be cut in a straight horizontal line somewhere between the top and the bottom, and the top portion were to be removed, the remaining portion forms a trapezoid (or other shape). all of these trapezoids are then superimposed one upon another, forming a single geometric shape. Then, the centroid of this shape, called the *fuzzy centroid*, is calculated. The *x* coordinate of the centroid is the defuzzified value.





Sub Area No.	Area	$\bar{x}$	Area × $\bar{x}$
1.	$\frac{1\times0.3}{2} = 0.150$	0.67	0.100
2.	$3\times0.3 = 0.90$	2.50	2.250
3.	$\frac{0.4\times0.2}{2} = 0.04$	3.73	0.149
4.	$2\times0.5 = 1.00$	1.00	5.000
5.	$\frac{0.5\times0.5}{2} = 0.125$	5.87	7.330
6.	$1\times1 = 1.00$	6.50	6.500
7.	$\frac{1\times1}{2} = 0.50$	7.33	3.660
	$\sum \text{Area}=3.715$		$\sum \text{Area} \times \bar{x}=24.989$

$x^*=\sum \text{Area} . \bar{x} \div \sum \text{Area}$

$x^*=24.989 \div 3.715$

$x^*=6.72$

Algorithm

- Fuzzy clustering algorithms are divided into two areas:

  classical fuzzy clustering

  shape-based fuzzy clustering.

-

Classical fuzzy clustering algorithms.

- **Fuzzy C-Means algorithm (FCM)**. This widely-used algorithm is practically identical to the K-Means algorithm. A data point can theoretically belong to all groups, with a membership function (also called a membership grade) between 0 and 1, where: 0 is where the data point is at the farthest possible point from a cluster's center and 1 is where the data point is closest to the center. Subtypes include Possibilistic C-Means (PCM), Fuzzy Possibilistic C-Means (FPCM) and Possibilistic Fuzzy C-Means (PFCM).
- **Gustafson-Kessel (GK) algorithm**: associates a data point with a cluster and a matrix. While C-means assumes the clusters are spherical, GK has elliptical-shaped clusters.
- **Gath-Geva algorithm** (also called Gaussian Mixture Decomposition): similar to FCM, but clusters can have *any* shape.

Fuzzy C-Means Clustering Preliminary

Given a set S composed of pattern vectors which we wish to cluster

$$S = \{ \boldsymbol{x}_1, \boldsymbol{x}_{2,} \dots , \boldsymbol{x}_N \}$$

Define C Cluster Membership Functions

$$\boldsymbol{\mu}_1 = [\, \mu_1(\boldsymbol{x}_1),\ \mu_1(\boldsymbol{x}_2), \dots ,\ \mu_1(\boldsymbol{x}_n) \,]$$

$$\boldsymbol{\mu}_2 = [\, \mu_2(\boldsymbol{x}_1),\ \mu_2(\boldsymbol{x}_2), \dots ,\ \mu_2(\boldsymbol{x}_n) \,]$$

$$\vdots \qquad\qquad\qquad \vdots$$

$$\boldsymbol{\mu}_K = [\, \mu_C(\boldsymbol{x}_1),\ \mu_C(\boldsymbol{x}_2), \dots ,\ \mu_C(\boldsymbol{x}_n) \,]$$

Define C Cluster Centroids as follows

Let V_j be the Cluster Centroid for Fuzzy Cluster Cl_j, j = 1, 2, …, C

Define a Performance Objective J as

$$J(U,V)=\sum_{i=1}^{n}\sum_{j=1}^{c}(\mu_{ij})^m\left\|\mathbf{x}_i-\mathbf{v}_j\right\|^2$$

where

'$||x_i - v_j||$' is the Euclidean distance between i^{th} data and j^{th} cluster center.

$$\mu_{ij}=1/\sum_{k=1}^{c}(d_{ij}/d_{ik})^{(2/m-1)}$$

$$\mathbf{v}_j=\left(\sum_{i=1}^{n}(\mu_{ij})^m x_i\right)/\left(\sum_{i=1}^{n}(\mu_{ij})^m\right),\forall j=1,2,\ldots..c$$

Definitions

n' is the number of data points.

'vj' represents the j^{th} cluster center. ‘

m' is the fuzziness index (m >1) → Higher numbers being more fuzzy

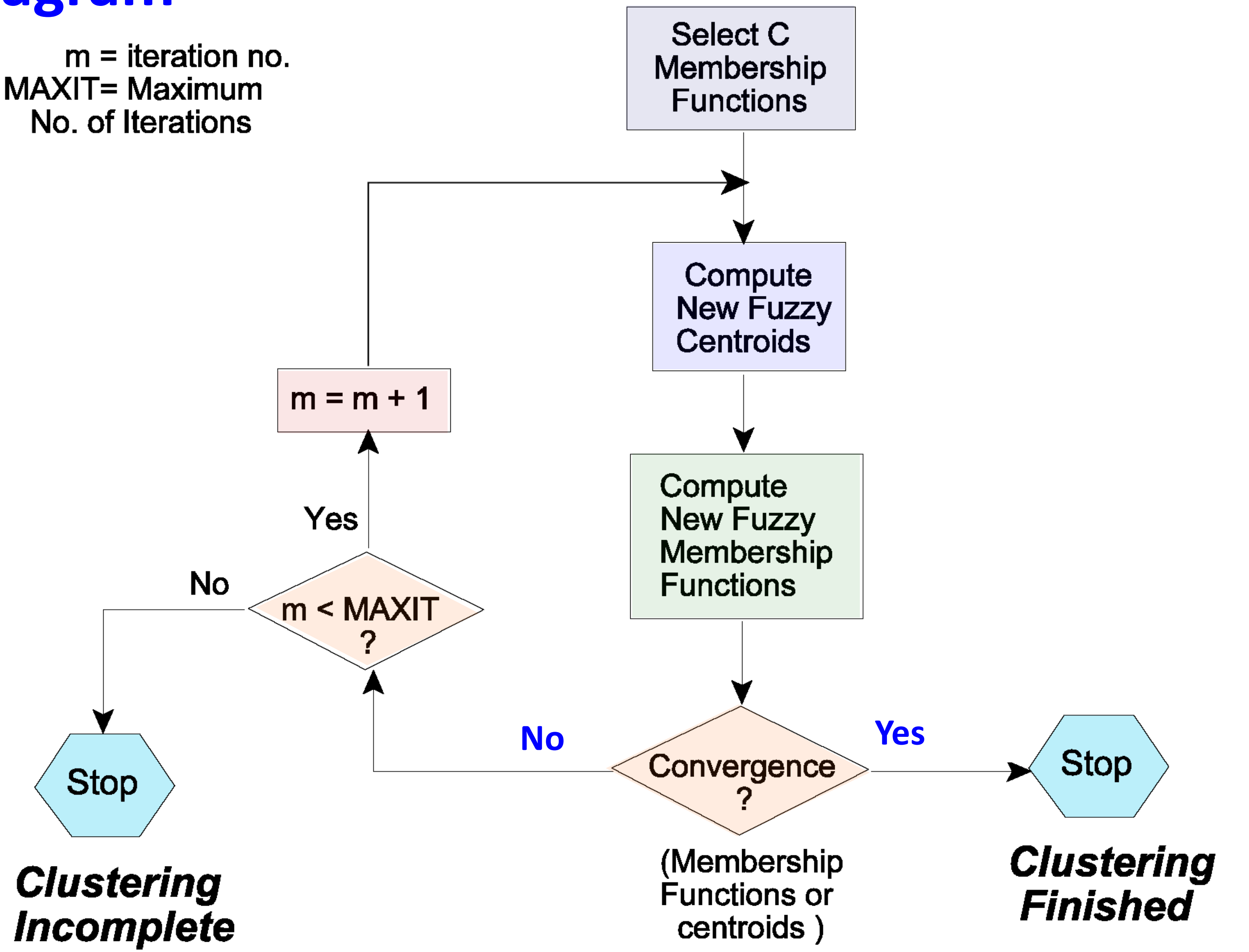
'c' represents the number of cluster center.

'μ_{ij}' represents the membership of i^{th} data to j^{th} cluster center.

'dij' represents the Euclidean distance between i^{th} data and j^{th} cluster center.

The **Fuzzy C-Means Algorithm** minimizes J_m by selecting V_j and μ_j , j =1, 2, … , C by an alternating iterative procedure as described in the algorithm’s details

Fuzzy C-Means Clustering Algorithm (a) Flow Diagram



Advantages and disadvantages

- **<u>Advantages</u>**

1) Gives best result for overlapped data set and comparatively better then k-means algorithm.
2) Unlike k-means where data point must exclusively belong to one cluster center here data point is assigned
3) membership to each cluster center as a result of which data point may belong to more then one cluster center.

- **<u>Disadvantages</u>**

1) Apriori specification of the number of clusters.
2) With lower value of β we get the better result but at the expense of more number of iteration.
3) Euclidean distance measures can unequally weight underlying factors.

Fuzzy C-means using R

Compute **fuzzy clustering** using the function **cmeans**() [in *e1071* R package]:

- The simplified format of the function **cmeans**() is as follow:

cmeans(x, centers, iter.max = 100, dist = "euclidean", m = 2)

  - x: a data matrix where columns are variables and rows are observations
  - centers: Number of clusters or initial values for cluster centers
  - iter.max: Maximum number of iterations
  - dist: Possible values are “euclidean” or “manhattan”
  - m: A number greater than 1 giving the degree of fuzzification.

terima kasih