4 Design of Fuzzy Sets

Fuzzy Systems Engineering
Toward Human-Centric Computing

Contents

4.8 Fuzzy sets as granular representatives of numeric data

4.9 From numeric data to fuzzy sets

4.10 Fuzzy equalization

4.11 Linguistic approximation

4.12 Design guidelines for the construction of fuzzy sets

4.1 Semantics of fuzzy sets: General observations

Semantics of fuzzy sets

- Generic constructs/building conceptual blocks to describe systems in a meaningful way

- Each fuzzy set comes with a well-delineated semantics (meaning)

  - Example: *small – medium – large* error

- Limited number of fuzzy sets

  - "magic" number of 7 +/- 2 (*Miller, 1956*) (short –term memory)

- Fuzzy sets require calibration
  - determination of their membership functions
- Two main approaches to the problem:
  - Expert –driven (designer, user, decision-maker…)
  - Data driven (from data to fuzzy sets)

4.2 Fuzzy sets as a descriptor of feasible solutions

Fuzzy sets as descriptor of feasible solutions (1)

Consider some function $f(x)$ defined in $\mathbf{\Omega}$,

$$f: \mathbf{\Omega} \rightarrow \mathbf{R}. \text{ where } \mathbf{\Omega} \subset \mathbf{R}$$

Determine its maximum

$$x^{\text{opt}} = \arg \max_x f(x).$$

Fuzzy set A of *optimal* solutions $\equiv$ a collection of feasible solutions that could be labeled as *optimal* with some degrees of membership.

$$A(x) = \frac{f(x) - f_{\min}}{f_{\max} - f_{\min}}$$

Fuzzy sets as descriptor of feasible solutions (2)

Consider some function $f(x)$ defined in $\mathbf{\Omega}$,

$$f: \mathbf{\Omega} \rightarrow \mathbf{R}. \text{ where } \mathbf{\Omega} \subset \mathbf{R}$$

Determine its minimum

$$x^{opt} = \arg\max_x f(x).$$

Fuzzy set *A* of *optimal* solutions ≡ a collection of feasible solutions that could be labeled as *optimal* with some degrees of membership.

$$A(x) = 1 - \frac{f(x) - f_{\min}}{f_{\max} - f_{\min}}$$

Fuzzy sets as descriptors of feasible solutions

Example

Linearization error

Linearize function $y = g(x) = \exp(-x)$ around $x_0=1$ and assess the quality of this linearization in the range [–1, 7].

Linearization formula: $y - y_0 = g'(x_0)(x - x_0)$

$y_0 = g(x_0)$ and $g'(x_0)$ is the derivative of $g(x)$ at x_0.

Linearized version of the function $\exp(-1)(2 - x)$.

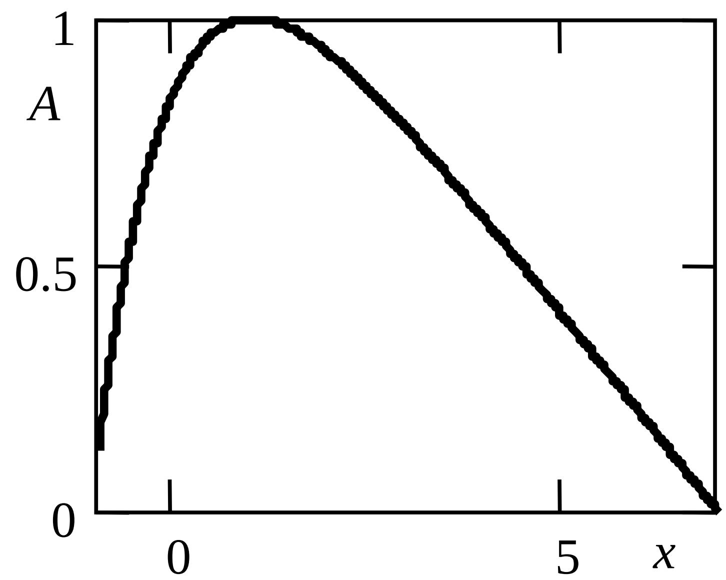
Quality of linearization $f(x) = |g(x) - \exp(-1)(2 - x)|$.

$$A(x) = 1 - \frac{f(x) - f_{\min}}{f_{\max} - f_{\min}}$$

$f_{\max} = f(7) = 1.84$ and $f_{\min} = 0.0$

Fuzzy sets as descriptors of feasible solutions

Example



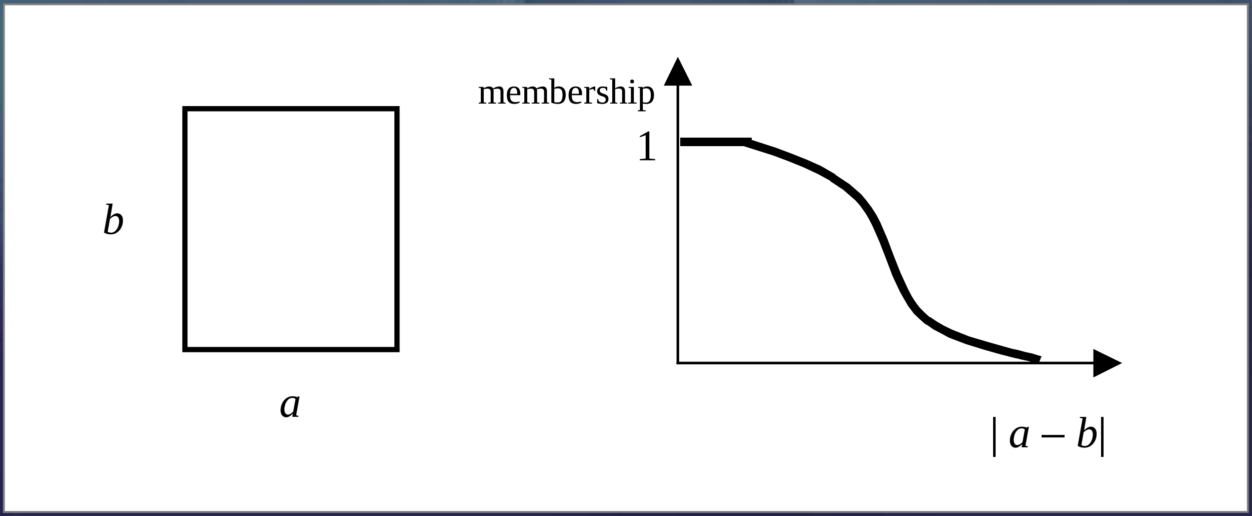
$$A(x) = 1 - \frac{f(x) - f_{\min}}{f_{\max} - f_{\min}}$$

4.3 Fuzzy sets as a descriptor of the notion of typicality

Fuzzy sets as notions of typicality

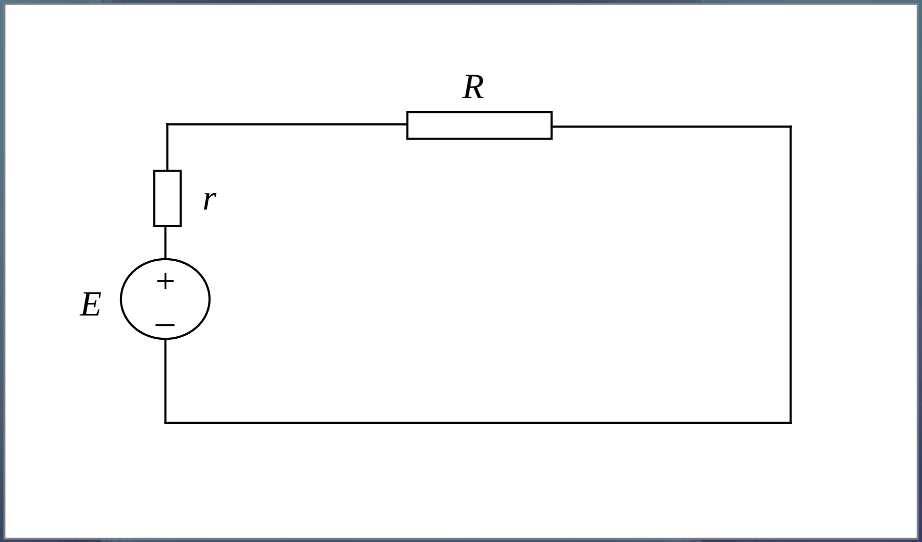
- Fuzzy set as collection of elements of varying degrees of typicality

- Geometric figures : squares, circles….

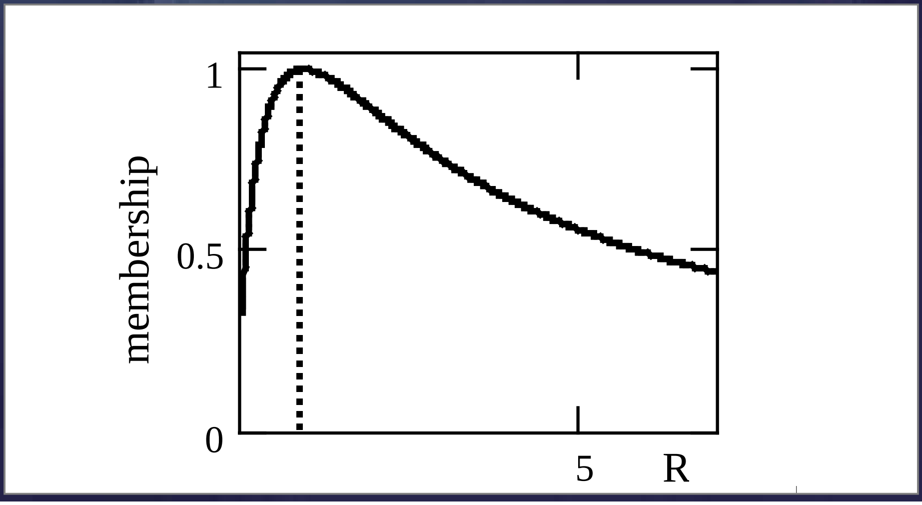


4.4 Membership functions in the visualization of preferences solutions

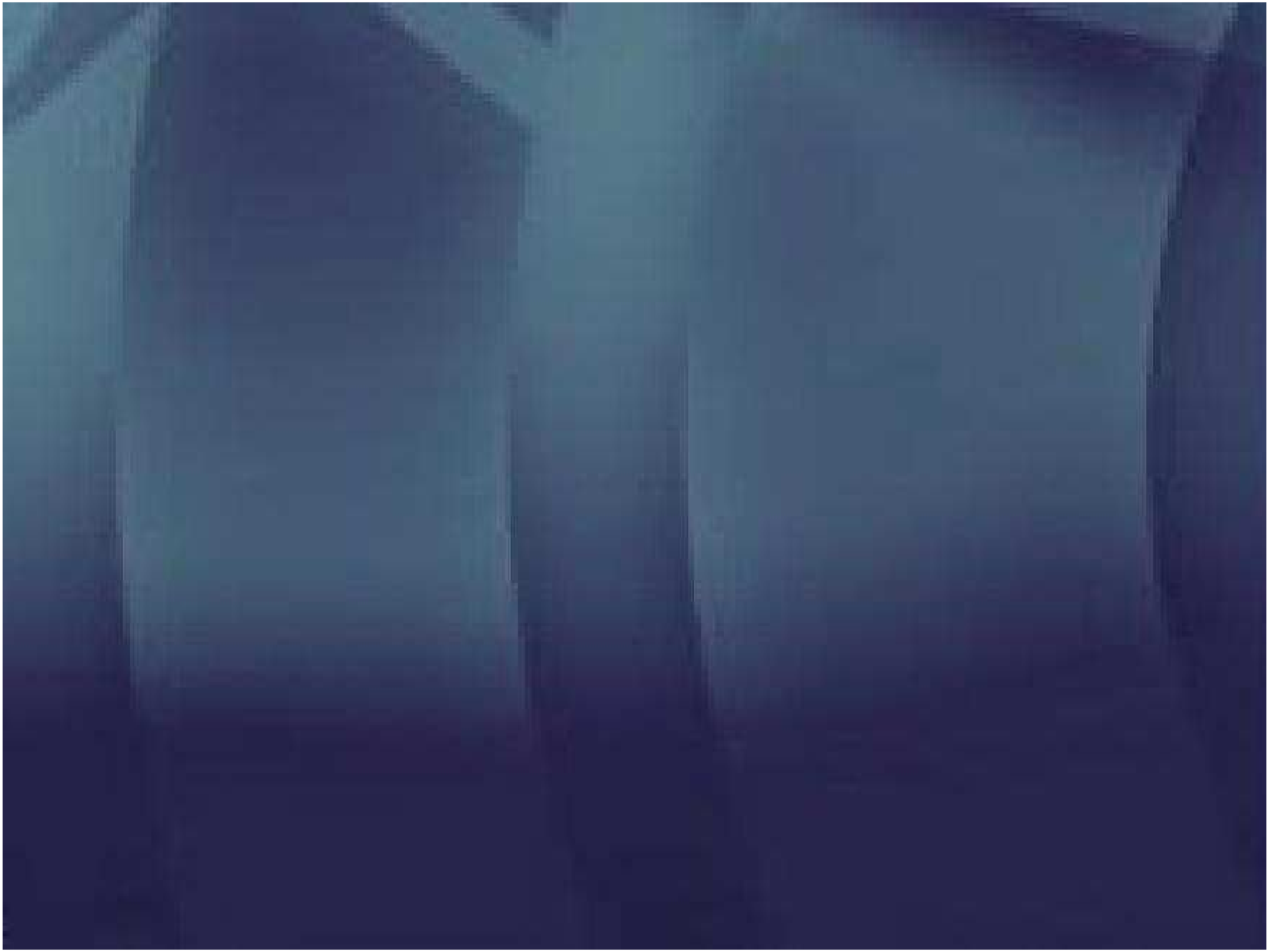
Fuzzy sets in visualization of preferences of solutions



$$P = i^2 R = \left(\frac{E}{R+r} \right)^2 R$$



4.5 Nonlinear transformations of fuzzy sets



4.6 Vertical and horizontal schemes of membership estimation

Horizontal scheme of membership estimation

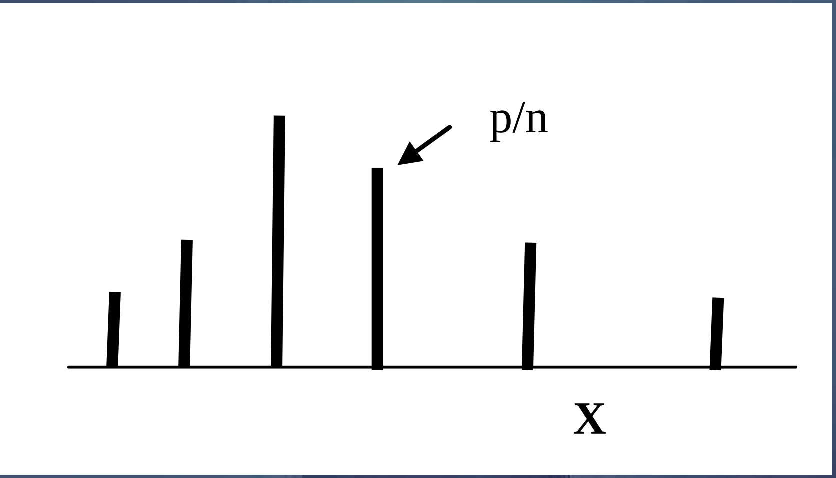
Finite elements of the universe of discourse X
Question of the form

-does x belong to concept A?

Accepted are binary answers (yes-no)

"n" experts – count of positive (yes) answers: p/n

Horizontal scheme of membership estimation



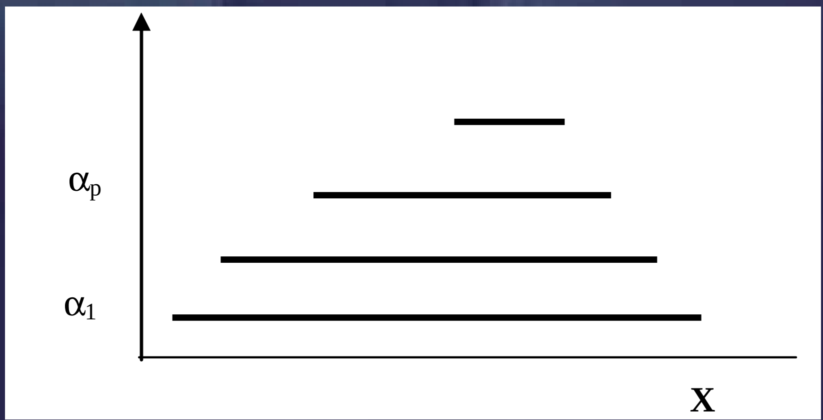
binary replies follow binomial distribution; we can determine confidence interval

$$\sigma = \sqrt{\frac{p(1-p)}{n}}$$

Vertical scheme of membership estimation

Estimation of membership function by determining α-cuts and aggregating them (see representation theorem)

-What are the elements of **X** which belong to fuzzy set A at degree not lower than α?



Pedrycz and Gomide, FSE 2007

Horizontal and vertical schemes of membership estimation

Simple and intuitively appealing

Reflective of domain knowledge

Lack of continuity – elements of **X** considered independently

Saaty's priority method of pairwise comparison

Collection of elements $x_1, x_2, \ldots, x_n$

Membership degrees are given $A(x_1), A(x_2) \ldots. A(x_n)$

Reciprocal matrix R

$$R = [r_{ij}] = \begin{bmatrix} \frac{A(x_1)}{A(x_1)} & \frac{A(x_1)}{A(x_2)} & \ldots & \frac{A(x_1)}{A(x_n)} \\ \frac{A(x_2)}{A(x_1)} & \frac{A(x_2)}{A(x_2)} & \ldots & \frac{A(x_2)}{A(x_1)} \\ & \ldots & \ldots & \\ \frac{A(x_n)}{A(x_1)} & \frac{A(x_2)}{A(x_n)} & \ldots & \frac{A(x_n)}{A(x_n)} \end{bmatrix} = \begin{bmatrix} 1 & \frac{A(x_1)}{A(x_2)} & \ldots & \frac{A(x_1)}{A(x_n)} \\ \frac{A(x_2)}{A(x_1)} & 1 & \ldots & \frac{A(x_2)}{A(x_1)} \\ & & \ldots & \\ \frac{A(x_n)}{A(x_1)} & \frac{A(x_2)}{A(x_n)} & \ldots & 1 \end{bmatrix}$$

Saaty's priority method of pairwise comparison

$$R = [r_{ij}] = \begin{bmatrix} \frac{A(x_1)}{A(x_1)} & \frac{A(x_1)}{A(x_2)} & \dots & \frac{A(x_1)}{A(x_n)} \\ \frac{A(x_2)}{A(x_1)} & \frac{A(x_2)}{A(x_2)} & \dots & \frac{A(x_2)}{A(x_1)} \\ & \dots & \dots & \\ \frac{A(x_n)}{A(x_1)} & \frac{A(x_2)}{A(x_n)} & \dots & \frac{A(x_n)}{A(x_n)} \end{bmatrix} = \begin{bmatrix} 1 & \frac{A(x_1)}{A(x_2)} & \dots & \frac{A(x_1)}{A(x_n)} \\ \frac{A(x_2)}{A(x_1)} & 1 & \dots & \frac{A(x_2)}{A(x_1)} \\ & & \dots & \\ \frac{A(x_n)}{A(x_1)} & \frac{A(x_2)}{A(x_n)} & \dots & 1 \end{bmatrix}$$

Reciprocal matrix R –main properties:

(a) reflexivity

(b) reciprocality

(c) transitivity

Saaty's priority method of pairwise comparison: computing

$$\left[\frac{A(x_i)}{A(x_1)}\ \frac{A(x_i)}{A(x_2)} \cdots \frac{A(x_i)}{A(x_n)}\right]\begin{bmatrix} A(x_1) \\ A(x_2) \\ \cdots \\ A(x_n) \end{bmatrix}$$

i-th row of R

RA = nA

n-the largest eigenvalue of R

Saaty's priority method of pairwise comparison

Estimation of reciprocal matrix:

Scale (typically 1-7 range, could be larger, 1-9)

- strong preference: high values on the scale (7-9)
- preference: 4-7
- weak preference or no preference 1-3

Solving the eigenvalue problem for R, max eigenvalue, λ_{max}

Saaty's priority method : consistency of results

$\nu = (\lambda_{max} - n)/(n-1)$

lack of consistency $\nu > 0.1$

Saaty's priority method : Example

high temperature

Universe of discourse: 10, 20, 30, 40, 45

Scale 1-5

$$R = \begin{bmatrix} 1 & 1/2 & 1/4 & 1/5 \\ 2 & 1 & 1/3 & 1/4 \\ 4 & 3 & 1 & 1/3 \\ 5 & 4 & 3 & 1 \end{bmatrix}$$

max eigenvalue = 4.114 eigenvector [0.122 0.195 0.438 0.869]

after normalization [0.14 0.22 0.50 1.00].

Fuzzy sets as granular representation of numeric data

The principle of justifiable granularity

experiment-driven and intuitively appealing rationale:

(a) we expect that A reflects (or matches) the available experimental data to the highest extent, and

(b) second, the fuzzy set is kept specific enough so that it comes with a well-defined semantics.

The principle of justifiable granularity

(a) we expect that A reflects (or matches) the available experimental data to the highest extent, and

Maximize "coverage" of data

(b) second, the fuzzy set is kept specific enough so that it comes with a well-defined semantics.

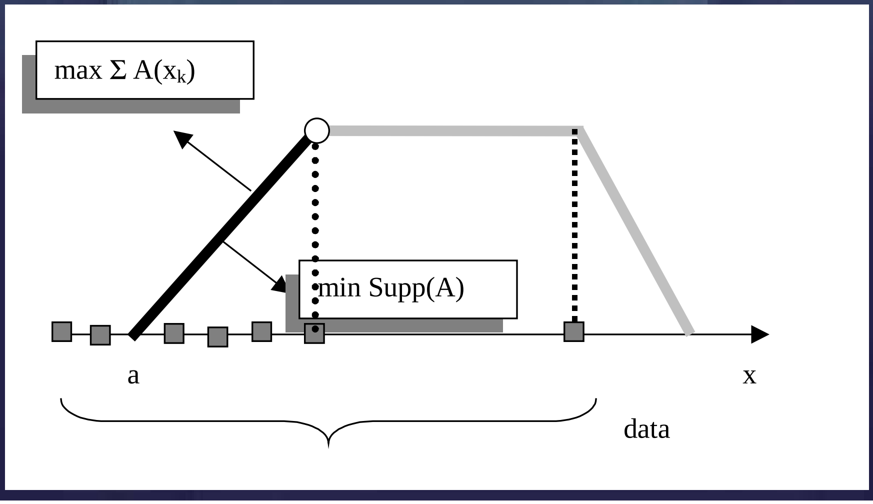
Minimize spread of fuzzy set

The principle of justifiable granularity: unimodal fuzzy set

Numeric data $x_1, x_2, \ldots, x_n$

Determine its "modified" median

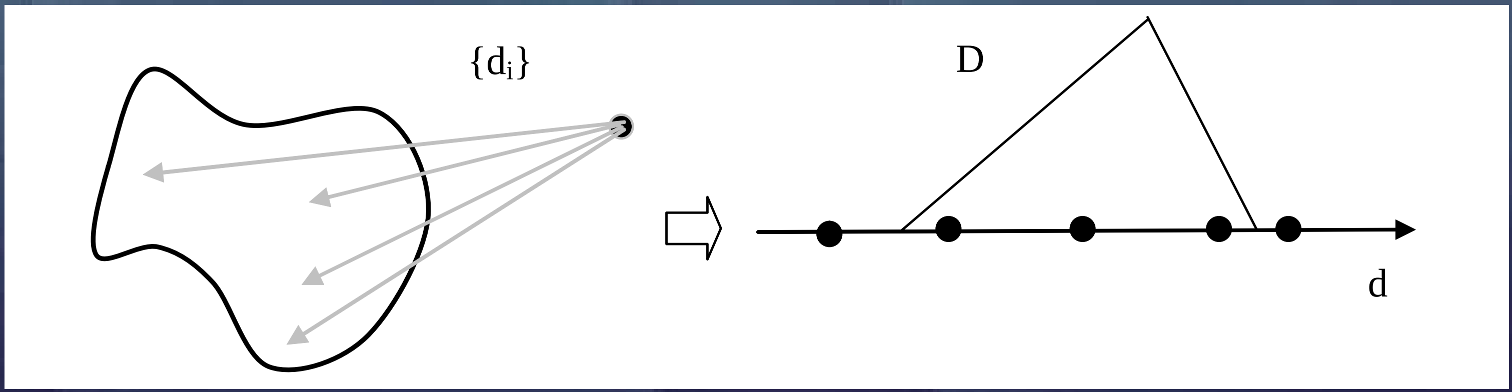
Consider separately data to the left and right from the median



$$\max_{a \neq m} \frac{\sum_k A(x_k)}{|m-a|}$$

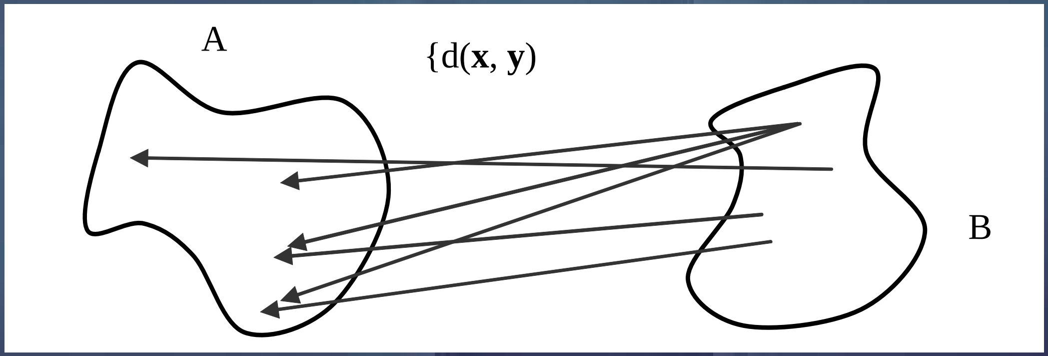
The principle of justifiable granularity: examples

Distance of point from geometric figure



The principle of justifiable granularity: examples

Distance between two geometric figures A and B



$$d_H(A, B) = \max\{\sup_{\mathbf{x}\in A}[\min_{\mathbf{y}\in B} d(\mathbf{x}, \mathbf{y})], \sup_{\mathbf{y}\in B}[\min_{\mathbf{x}\in A} d(\mathbf{x}, \mathbf{y})]\}$$

Clustering: Fuzzy C-Means (FCM)

Given a collection of n-dimensional data set $\{\mathbf{x}_k\}$, k=1,2,…,N,

determine its structure – a collection of “c” clusters.

Minimize the following objective function (performance index) Q

$$Q = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^{m} \| \mathbf{x}_k - \mathbf{v}_i \|^2$$

Fuzzy clustering: structure representation

Partition matrix U

Prototypes $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_c$

$$\sum_{i=1}^{c} u_{ik} = 1, \quad k = 1,2,\ldots, N$$

$$0 < \sum_{k=1}^{N} u_{ik} < N, \quad i = 1,2,\ldots, c$$

FCM – optimization procedure

Optimization with respect to

- Partition matrix U, and
- Prototypes $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_c$

Optimization: partition matrix

use of Lagrange multipliers

$$V = \sum_{i=1}^{c} u_{ik}^{m} d_{ik}^{2} + \lambda\left(\sum_{i=1}^{c} u_{ik} - 1\right)$$

$$\frac{\partial V}{\partial u_{st}} = 0 \qquad \frac{\partial V}{\partial \lambda} = 0$$

Optimization: partition matrix

$$V = \sum_{i=1}^{c} u_{ik}^{m} d_{ik}^{2} + \lambda\left(\sum_{i=1}^{c} u_{ik} - 1\right)$$

$$\frac{\partial V}{\partial u_{st}} = 0 \qquad \frac{\partial V}{\partial \lambda} = 0$$

$$\frac{\partial V}{\partial u_{st}} = m u_{st}^{m-1} d_{st}^{2} + \lambda$$

$$u_{st} = -\left(\frac{\lambda}{m}\right)^{\frac{1}{m-1}} d_{st}^{\frac{2}{m-1}}$$

$$-\left(\frac{\lambda}{m}\right)^{\frac{1}{m-1}} \sum_{j=1}^{c} d_{jt}^{\frac{2}{m-1}} = 1$$

$$-\left(\frac{\lambda}{m}\right)^{\frac{1}{m-1}} = \frac{1}{\sum_{j=1}^{c} d_{jt}^{\frac{2}{m-1}}}$$

$$u_{st} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{st}^{2}}{d_{jt}^{2}}\right)^{\frac{1}{m-1}}}$$

Optimization: prototypes

$$Q = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^{m} \sum_{j=1}^{n} (x_{kj} - v_{ij})^2$$

Gradient of Q w.r.t. prototype $\mathbf{v}_s$

$$\sum_{k=1}^{N} u_{ik}^{m} (x_{kt} - v_{st}) = 0$$

$$v_{st} = \frac{\sum_{k=1}^{N} u_{ik}^{m} x_{kt}}{\sum_{k=1}^{N} u_{ik}^{m}}$$

FCM: an overview of the algorithm

procedure FCM-CLUSTERING ($\mathbf{x}$) **returns** prototypes and partition matrix

input : data $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k\}$

local: fuzzification parameter: m

threshold: ε

norm: ||.||

INITIALIZE-PARTITION-MATRIX

$t \leftarrow 0$

repeat

for i=1:c **do**

$$\mathbf{v}_i(t) \leftarrow \frac{\sum_{k=1}^{N} u_{ik}^m(t)\mathbf{x}_k}{\sum_{k=1}^{N} u_{ik}^m(t)} \quad \text{compute prototypes}$$

for i = 1:c **do**

for k = 1:N **do**

update partition matrix

$$u_{ik}(t+1) = \frac{1}{\sum_{j=1}^{c}\left(\frac{\|\mathbf{x}_k - \mathbf{v}_i(t)\|}{\|\mathbf{x}_k - \mathbf{v}_j(t)\|}\right)^{2/(m-1)}} \quad \text{update partition matrix}$$

$t \leftarrow t + 1$

until $\|U(t+1)-U(t)\| \leq \varepsilon$

return U, V

FCM and its parameters

Number of clusters (c)

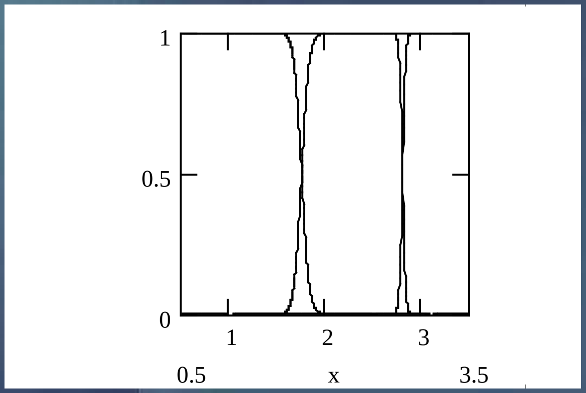
Objective function Q

Distance function ||.||

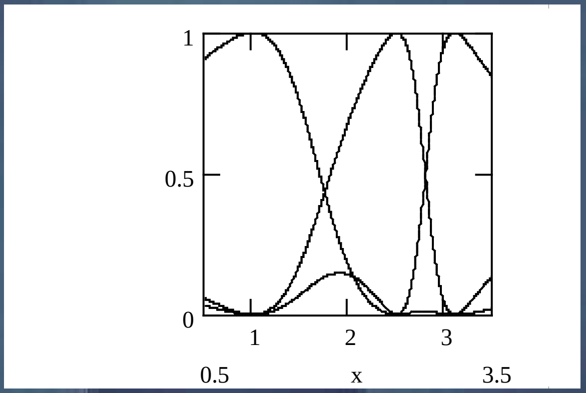
Fuzzification coefficient (m)

Termination criterion

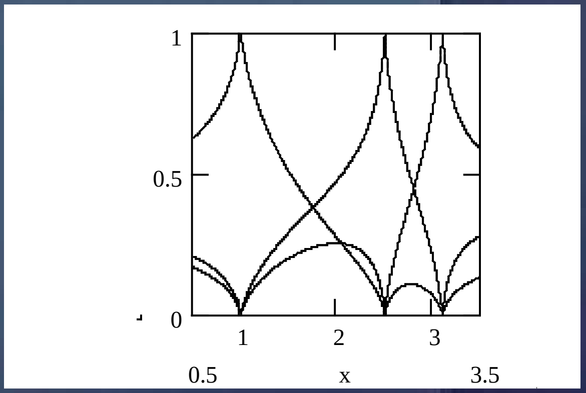
Geometry of clusters and fuzzification coefficient (m)



m =1.2



m =2.0



m =3.5

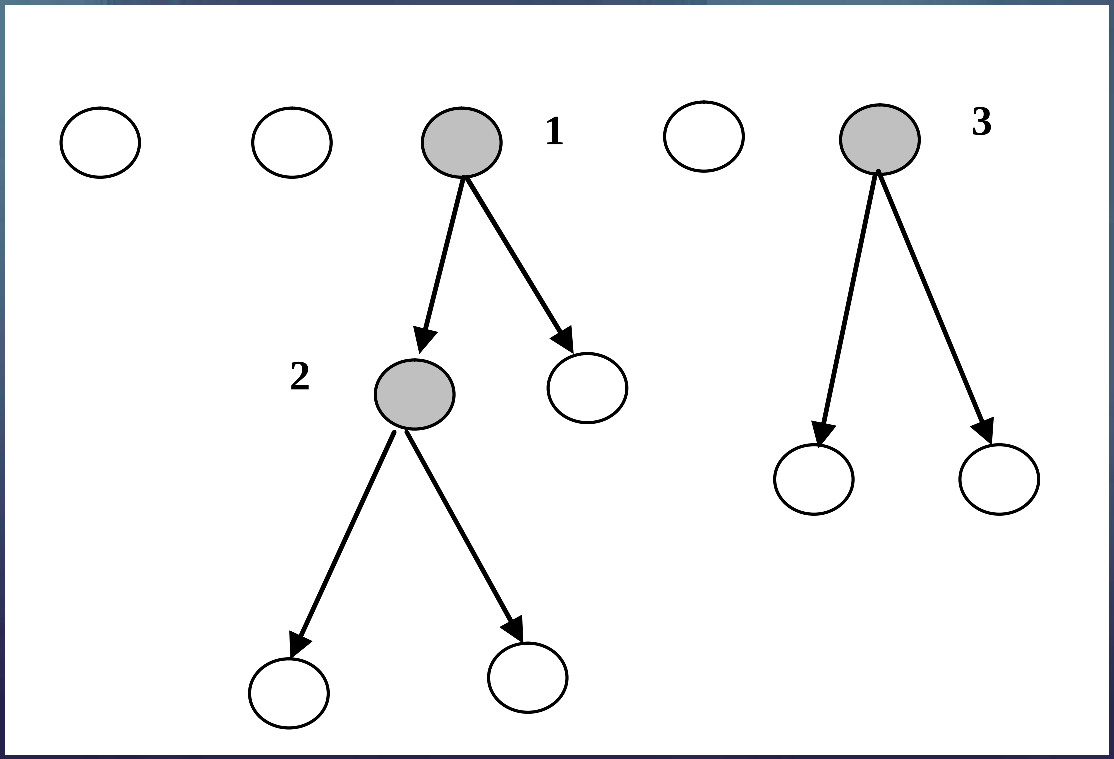
Cluster sharing: a separation measure

$$\varphi(u_1, u_2, \ldots, u_c) = 1 - c^c \prod_{i=1}^{c} u_i$$

Data fully belongs to a single cluster (1- 0)

Data belongs to all clusters at the same level (1/c)

Hierarchical format of FCM: Successive refinements of clusters

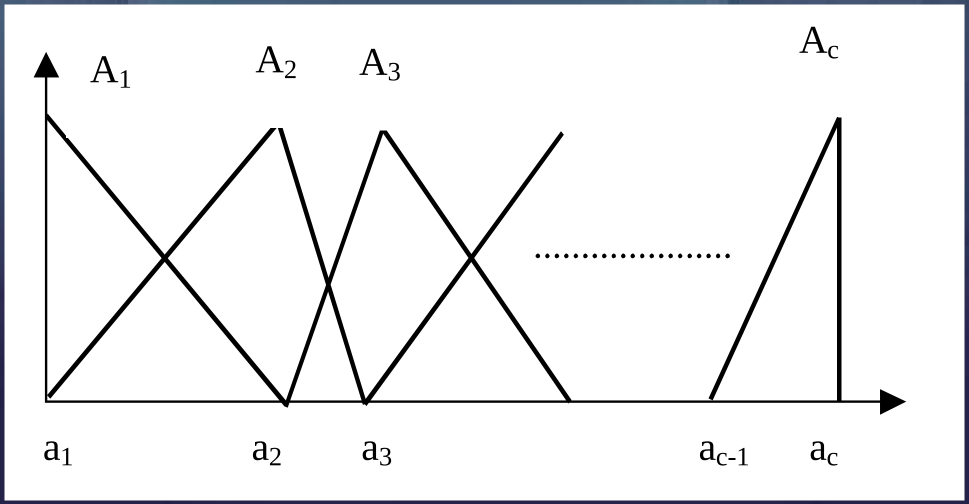


$$V_i = \sum_{k=1}^{N} u_{ik}^{m} \| \mathbf{x}_k - \mathbf{v}_i \|^2$$

$$\mathbf{X}(i_0) = \{ x_k \in \mathbf{X} | \; u_{i_0 k} = \max_i u_{ik} \}$$

Fuzzy equalization

Construct triangular fuzzy sets $A_1, A_2, \ldots, A_c$ defined in **R** such that they come with the same level of experimental evidence (support)



$$\sum_{k=1_k}^{N} A_1(x_k) = \frac{N}{2(c-1)}$$

$$\sum_{k=1_k}^{N} A_2(x_k) = \frac{N}{(c-1)}$$

$$\sum_{k=1_k}^{N} A_{c-1}(x_k) = \frac{N}{(c-1)}$$

$$\sum_{k=1_k}^{N} A_c(x_k) = \frac{N}{2(c-1)}$$

Linguistic approximation

Given is a family of reference fuzzy sets $\{A_i\}$ defined in some space **X**

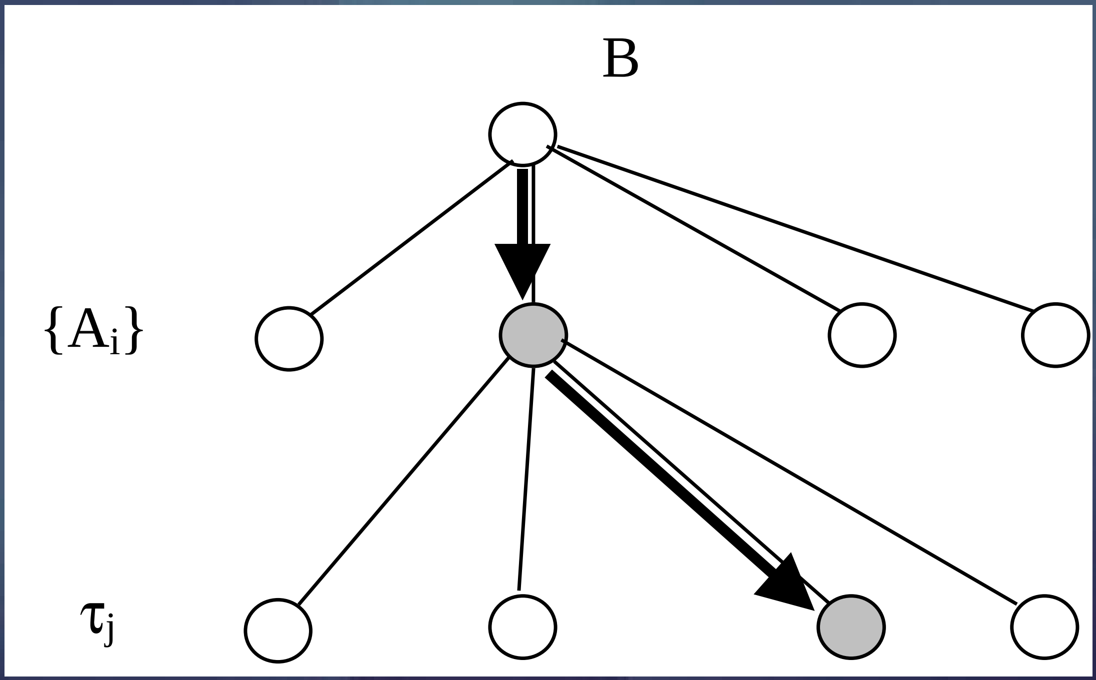
We have at disposal is a family of linguistic modifiers τ_j, say

more or less (dilution),

very (concentration)

Represent (approximate) B in **X** with the use of reference fuzzy sets and linguistic modifiers == linguistic approximation

Linguistic approximation: optimization



$$B \approx \tau_i(A_j)$$

Construction of fuzzy sets: Design guidelines (1)

Strive for highly visible and well-defined semantics of information granules. Keep the number of information granules quite low (7 $\pm$ 2 fuzzy sets).

There are several fundamental views at fuzzy sets and depending upon them, consider the use of various estimation techniques.

Fuzzy sets are context-sensitive constructs and require careful calibration. This The calibration mechanisms being used in the design of the membership function are reflective of human-centricity of fuzzy sets.

Construction of fuzzy sets: Design guidelines (2)

two major categories of approaches supporting the design of membership functions:

data-driven and expert (user)-driven.

The user-driven membership estimation uses the statistics of data yet in an *implicit* manner.

The granular term – fuzzy sets come into existence once there is some experimental evidence behind them

the development of fuzzy sets can be carried out in an stepwise manner.