14 Granular Models and Human-Centric Computing

Fuzzy Systems Engineering
Toward Human-Centric Computing

Contents

14.1 Cluster-Based representation of input-output mappings

Human-Centric systems and computing

- Concerns with
  - – functionality responsive to human user needs
  - – diversity of requirements and user preferences
  - – relevance feedback
- Examples
  - – system modeling within a context chosen by the user
  - – information retrieval depending upon user preferences
  - – context-based learning

Cluster-based representation of I/O mapping

- Fuzzy clustering

  – sound basis to construct fuzzy models
  – clustering in the input$\times$output space
  – collection of prototypes $\rightarrow$ model skeleton/blueprint
  – different ways to use prototypes to develop the model

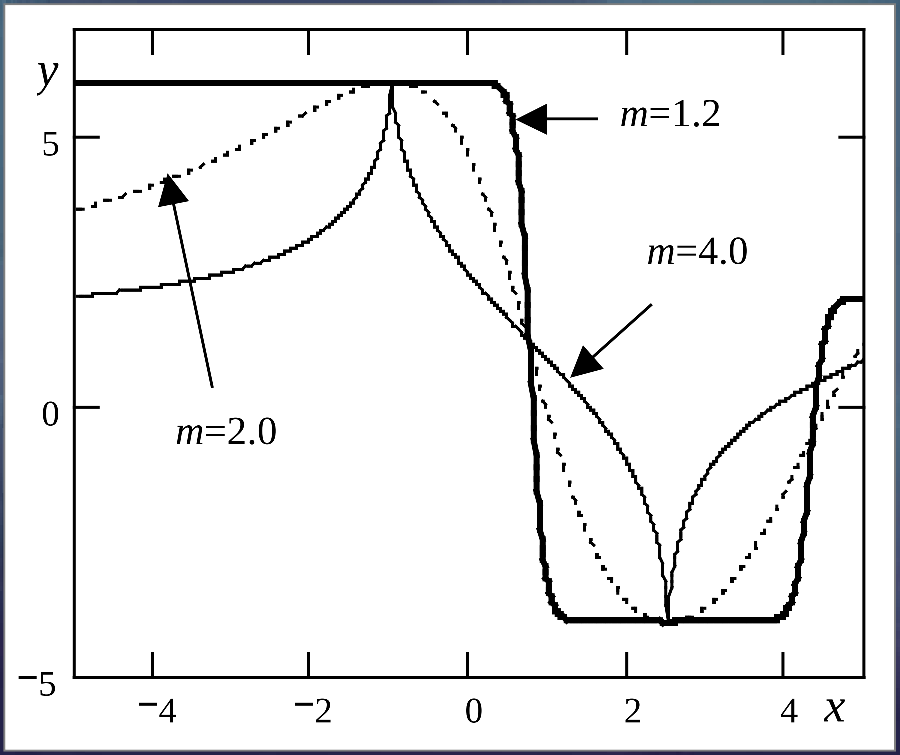
- Example

$z_1, z_2, \ldots., z_c$ prototypes formed at the output space
$\mathbf{v}_1, \mathbf{v}_2, \ldots., \mathbf{v}_c$ prototypes formed at the input space
$u_1(\mathbf{x}), \ldots., u_c(\mathbf{x})$ membership grades

$$y = \sum_{i=1}^{c} z_i u_i(\mathbf{x})$$

$$u_i(x) = \frac{1}{\sum_{i=1}^{c} \left(\frac{\| \mathbf{x} - \mathbf{v}_i \|}{\| \mathbf{x} - \mathbf{v}_i \|} \right)^{2/(m-1)}}$$

Example



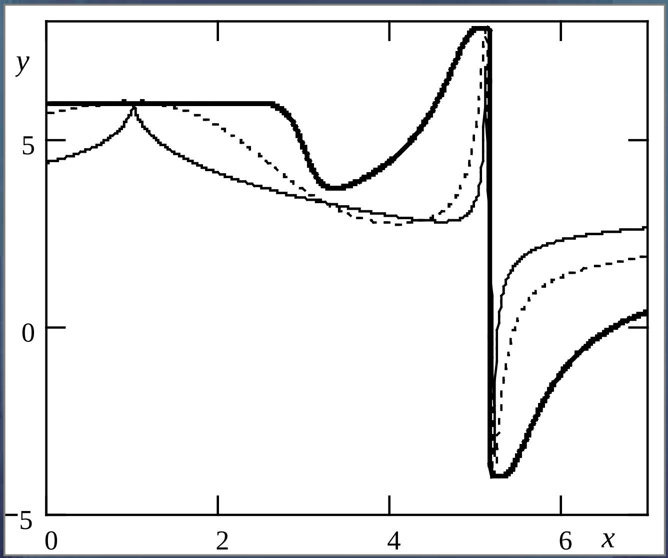
$v_1 = -1$ $z_1 = -6$

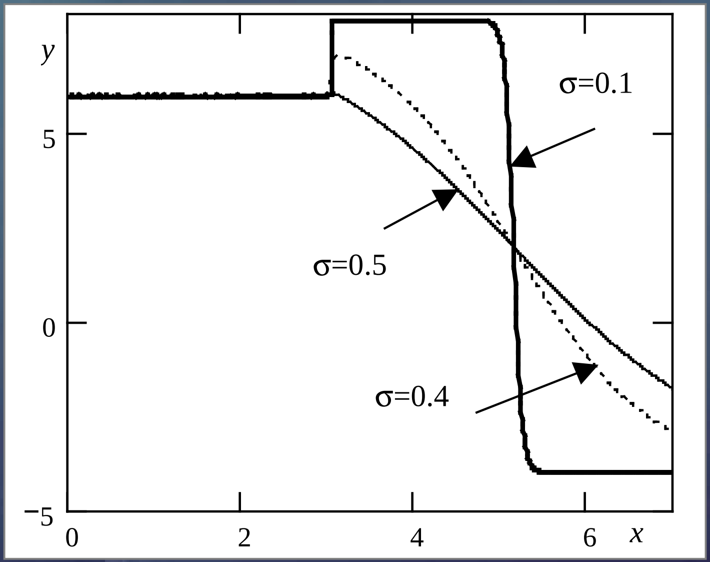
$v_2 = 2.5$ $z_2 = -4$

$v_3 = 6.1$ $z_3 = 2$

Cluster-Based × RBF

$m = 2$





$v_1 = 1$ $\quad z_1 = 6$

$v_2 = 5.2$ $\quad z_2 = -4$

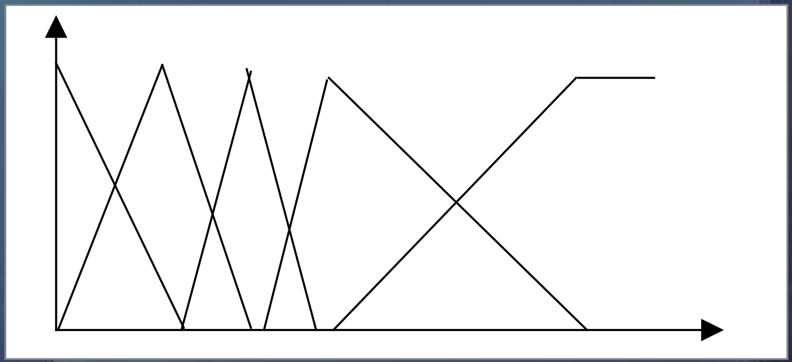
$v_3 = 5.1$ $\quad z_3 = 8$

$$y = \frac{\sum_{i=1}^{c} z_i G(x; v_i, \sigma)}{\sum_{i=1}^{c} G(x; v_i, \sigma)}$$

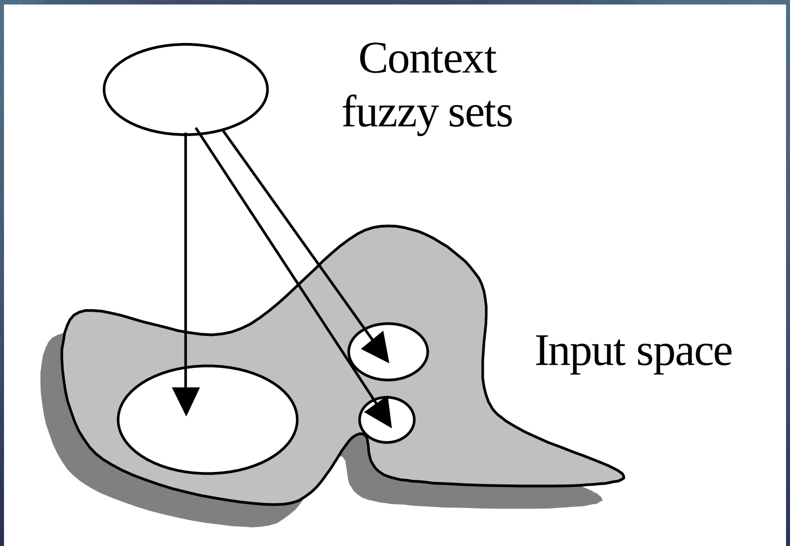
Pedrycz and Gomide, FSE 2007

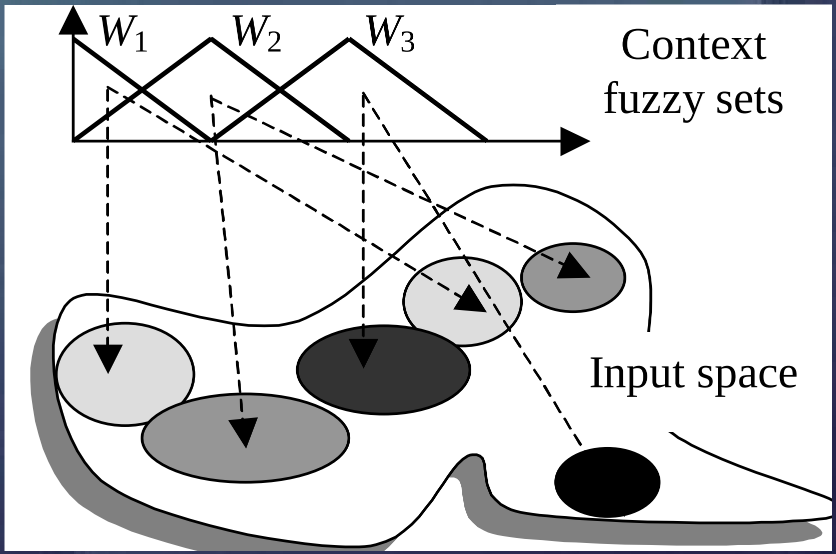
14.2 Context-Based clustering in the development of granular models

- Contexts: $W_1, \ldots., W_p$
- W_j is a fuzzy set
- data point (target_k)
- $w_{jk} = W_j(\text{target}_k)$



Context-Based clustering





Partition matrices induced by the *j*th context

$$U(W_j) = \left\{ u_{ik} \in [0,1] \mid \sum_{i=1}^{c} u_{ik} = w_{ik}, \forall k, 0 < \sum_{k=1}^{N} u_{ik} < N, \forall i \right\}$$

- Context-Based clustering algorithm

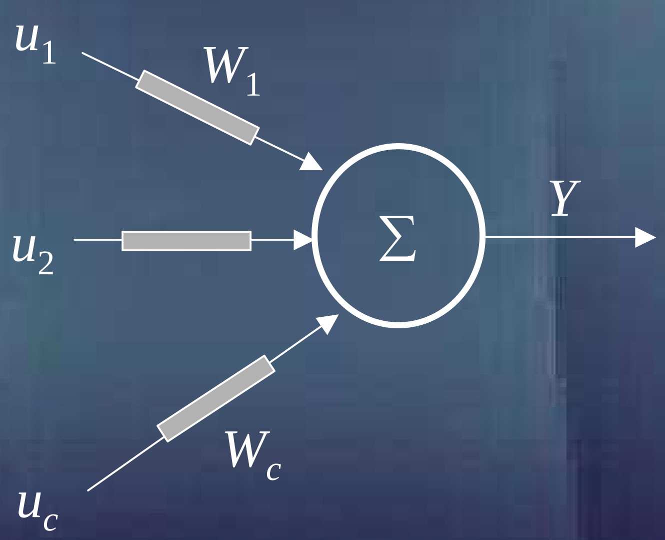
$$Q = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^{m} \| \mathbf{x}_k - \mathbf{v}_i \|^2$$

$$u_{ik} = \frac{w_{jk}}{\sum_{j=1}^{c} \left(\frac{\| \mathbf{x} - \mathbf{v}_j \|}{\| \mathbf{x} - \mathbf{v}_j \|} \right)^{2/(m-1)}}$$

$$\mathbf{v}_i = \frac{\sum_{k=1}^{N} u_{ik}^{m} \mathbf{x}_k}{\sum_{k=1}^{N} u_{ik}^{m}}$$

14.3 Granular neuron as a generic processing element in granular networks

Granular neuron



$$Y = N(u_1,....,u_c,W_1,....,W_c) = \sum_{\oplus}(W_i \otimes u_i)$$

$$W_i = [w_{i-}, w_{i+}]$$

$$W_i \otimes u_i = [w_{i-}u_i, w_{i+}u_i]$$

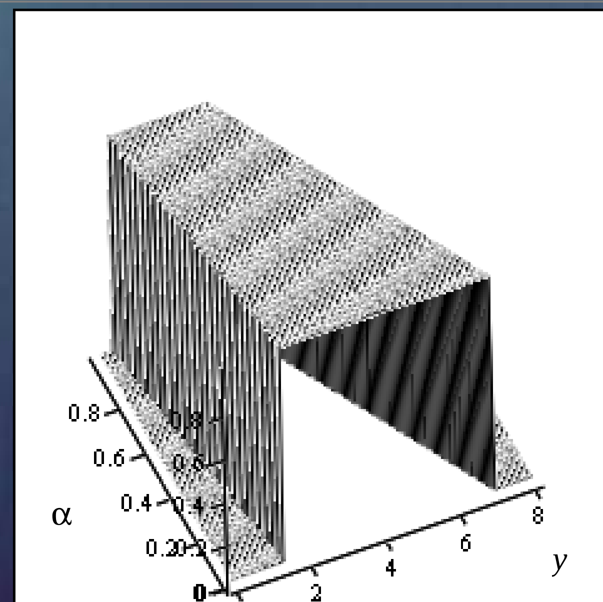
Pedrycz and Gomide, FSE 2007

Interval-valued connections

$$W_i = [w_{i-}, w_{i+}]$$

$$W_i \otimes u_i = [w_{i-}u_i, w_{i+}u_i]$$

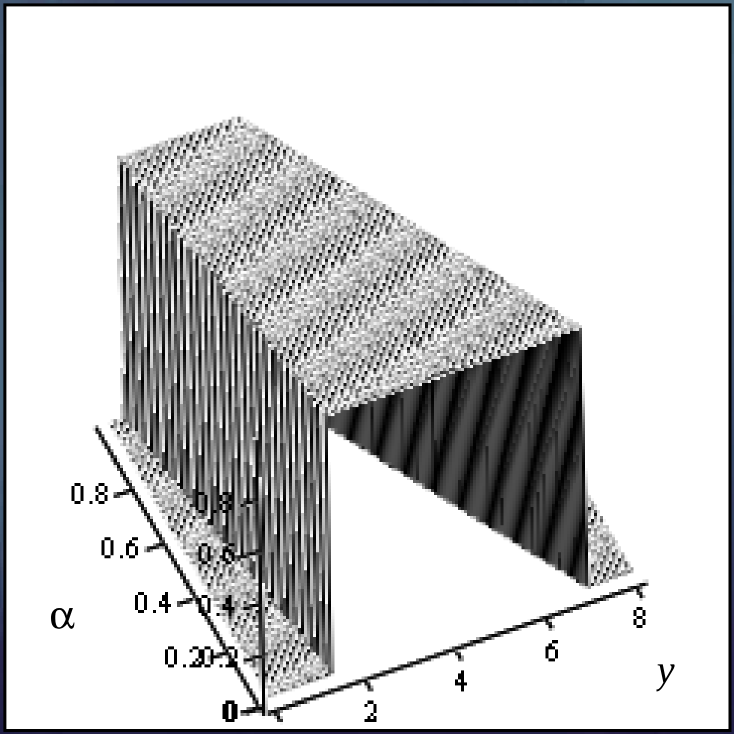
$$Y = \left[\sum_{i=1}^{c} w_{i-}u_i, \sum_{i=1}^{c} w_{i+}u_i, \right]$$



$u_1 = \alpha$ $\qquad$ $W_1 = [0.3, 3]$

$u_2 = 1–\alpha$ $\qquad$ $W_2 = [1.4, 7]$

Example



$u_1 = \alpha$
$u_2 = 1–\alpha$

$W_1 = [0.3, 3]$
$W_2 = [1.4, 7]$

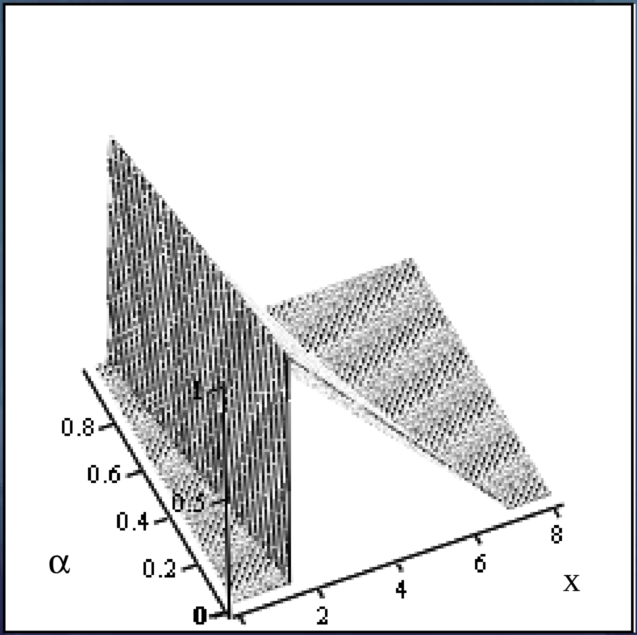
Fuzzy set-valued connections

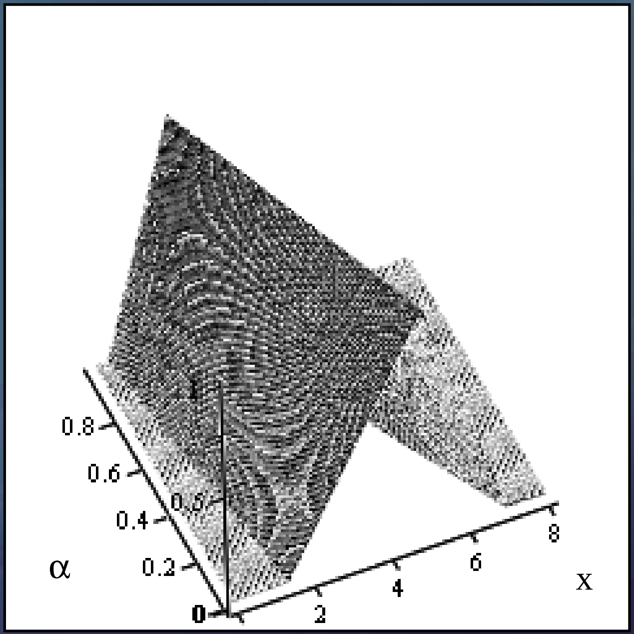
$$W_i \otimes u_i = \sup_{w: y = w u_i} W_i(w) = W_i(y / u_i)$$

$$Y = Z_1 \oplus Z_2 \oplus \cdots \oplus Z_n$$

$$Y(y) = \sup_{y = y_1 + \cdots + y_c} \{\min(Z_1(y_1), \ldots, Z_c(y_c))\}$$

Example



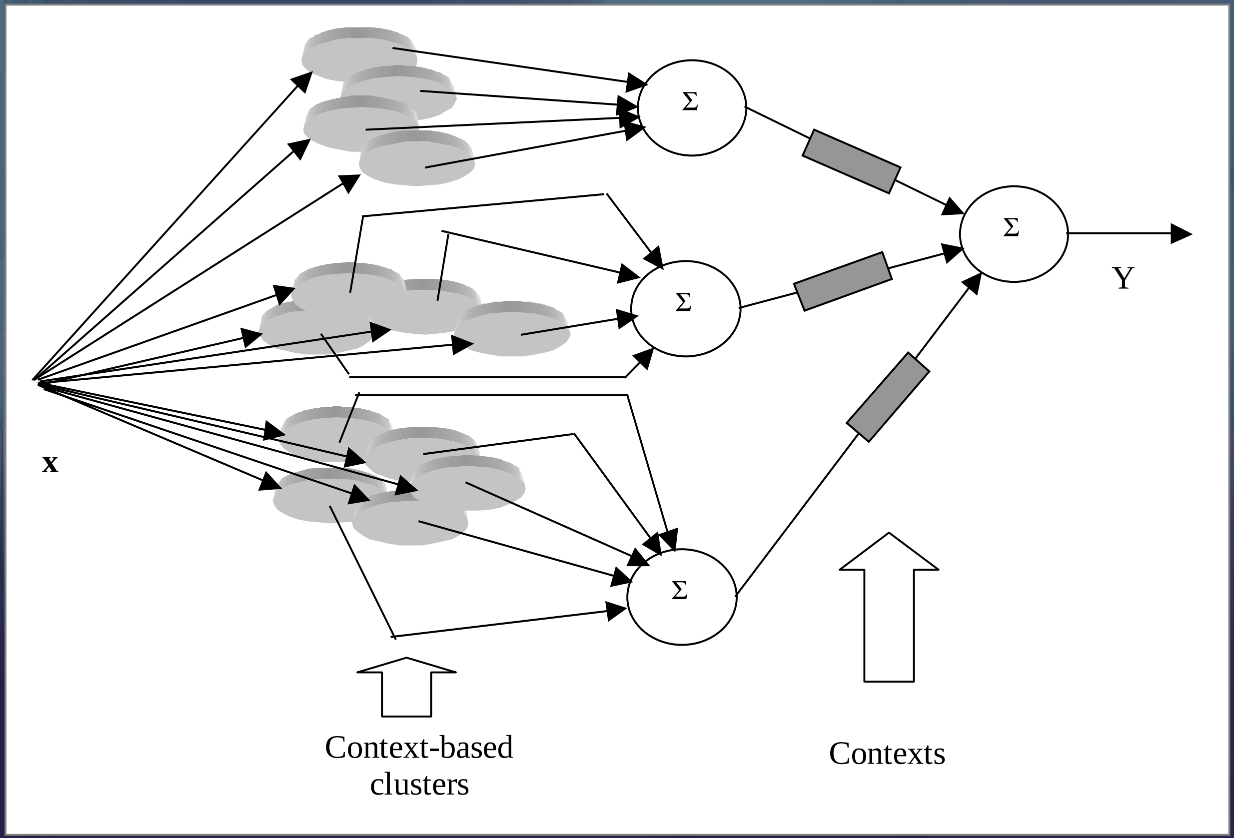


W_1 = <0.3,0.5,3.0>
W_2 = <1.4,1.5,7.0>

W_1 = <0.3,2.0,3.0>
W_2 = <1.4,5.0,7.0>

14.4 Architecture of granular models based on conditional fuzzy clustering

Overall architecture of granular models



- Development phases of granular models

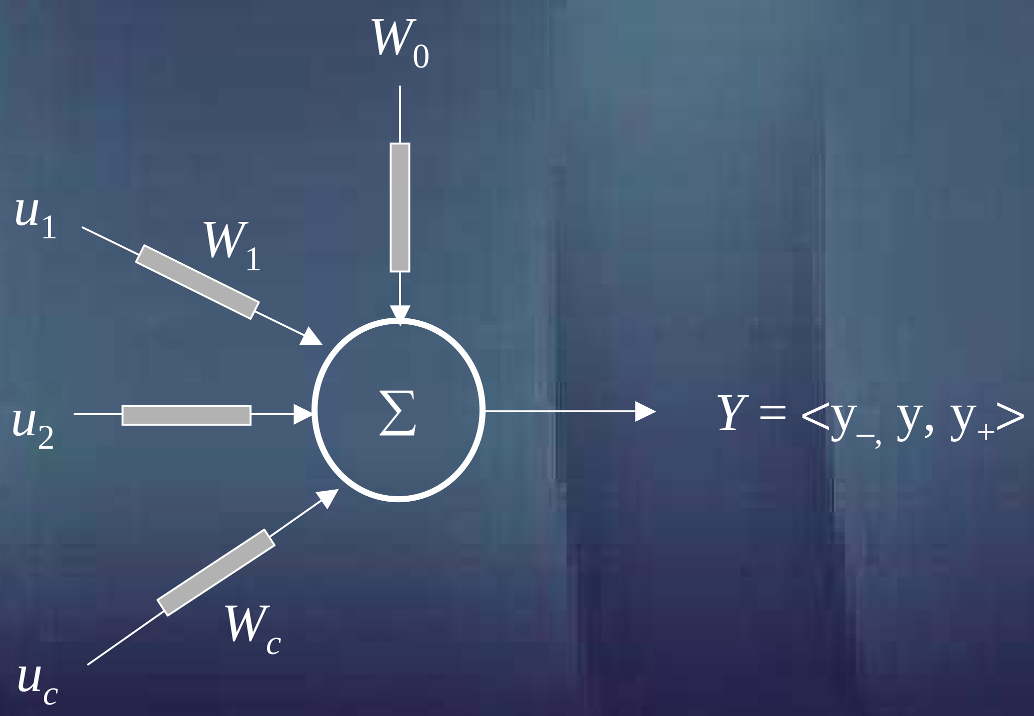
    1– form fuzzy sets of context
    2– conditional clustering based on the contexts

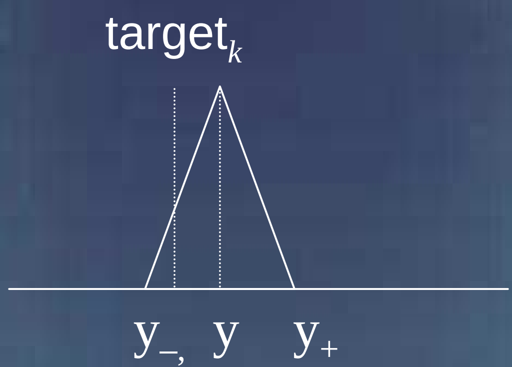
- Features of granular models

    – web of associations between information granules
    – inherently granular models (granular outputs for numeric inputs)
    – design using rapid prototyping scheme

14.5 Refinements of granular models

Bias of granular neurons





$$w_0 = -\frac{1}{N}\sum_{k=1}^{N}(\text{target}_k - y_k)$$

$$\sum_{t=1}^{p} z_t w_{t-} + w_0$$

- lower bound

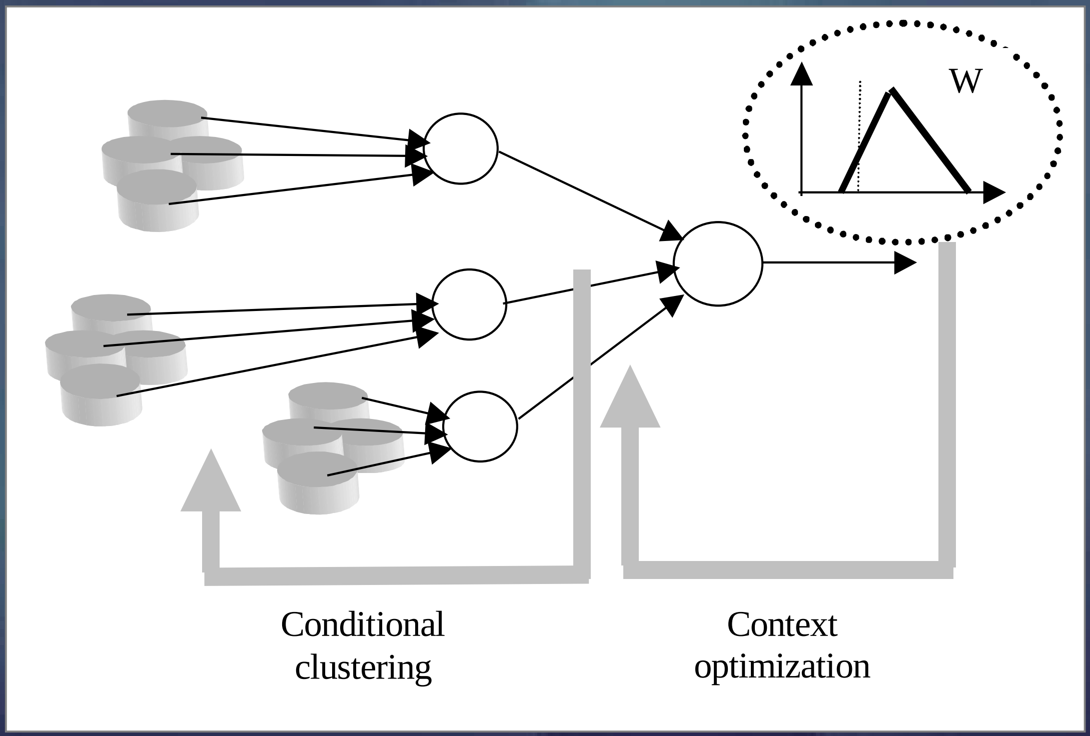
$$\sum_{t=1}^{p} z_t w_t + w_0$$

- modal value

$$\sum_{t=1}^{p} z_t w_{t+} + w_0$$

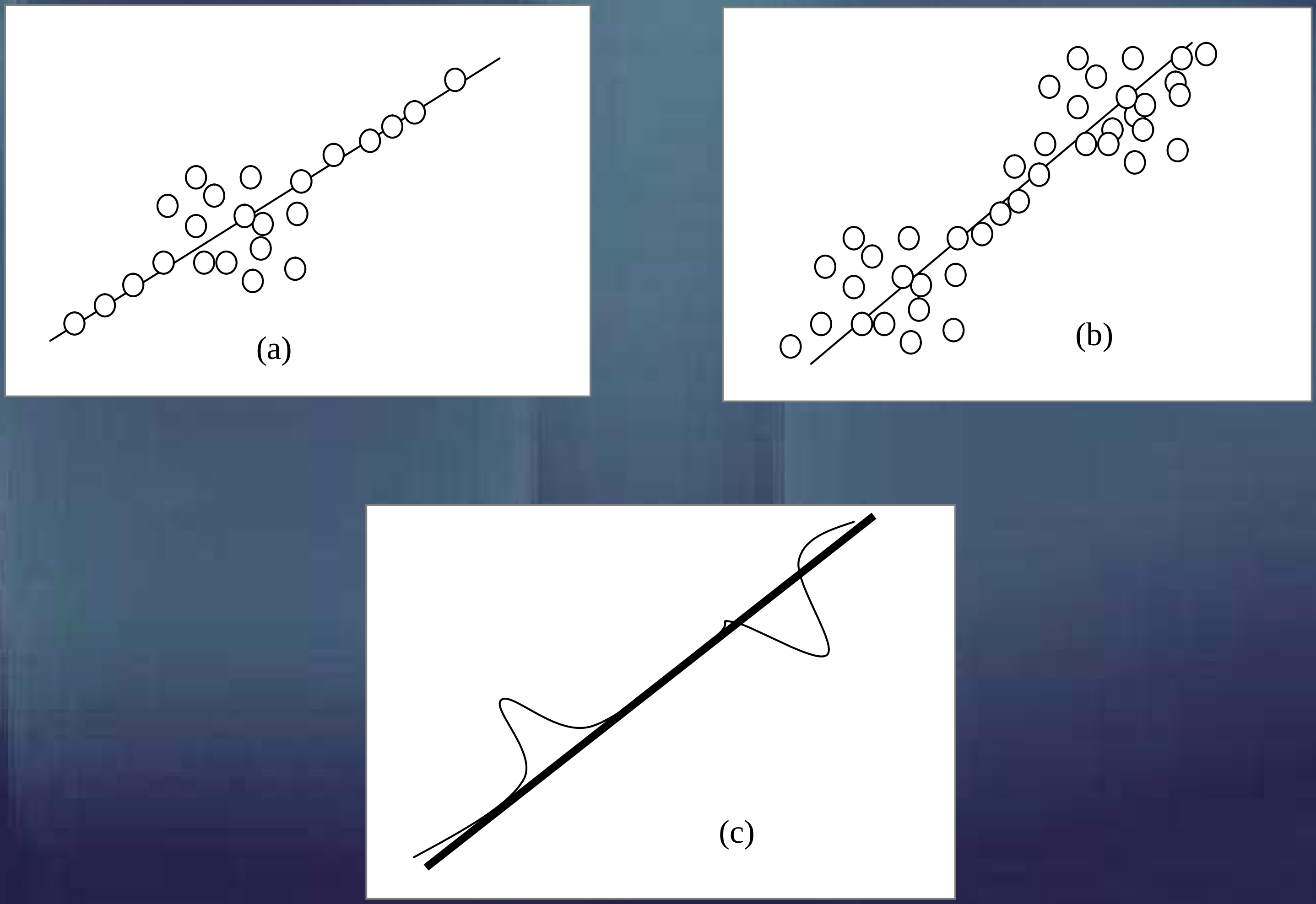
- upper bound

Refinement of contexts



$$\max_{\mathbf{P}} \frac{1}{N} \sum_{k=1}^{N} Y(\mathbf{x}_k)(\mathrm{target}_k) \quad \text{or} \quad \min_{\mathbf{P}} \frac{1}{N} \sum_{k=1}^{N} (b_k - a_k)$$

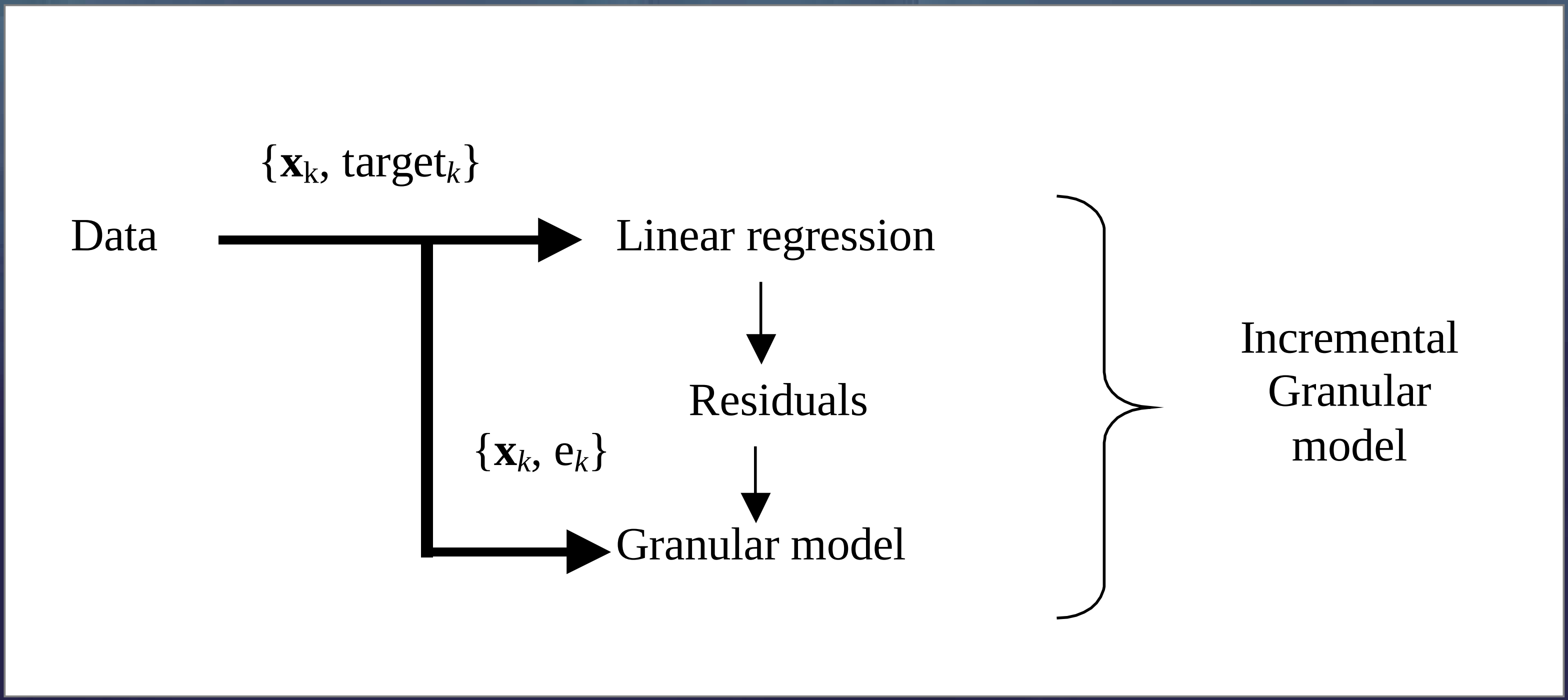
14.6 Incremental granular models



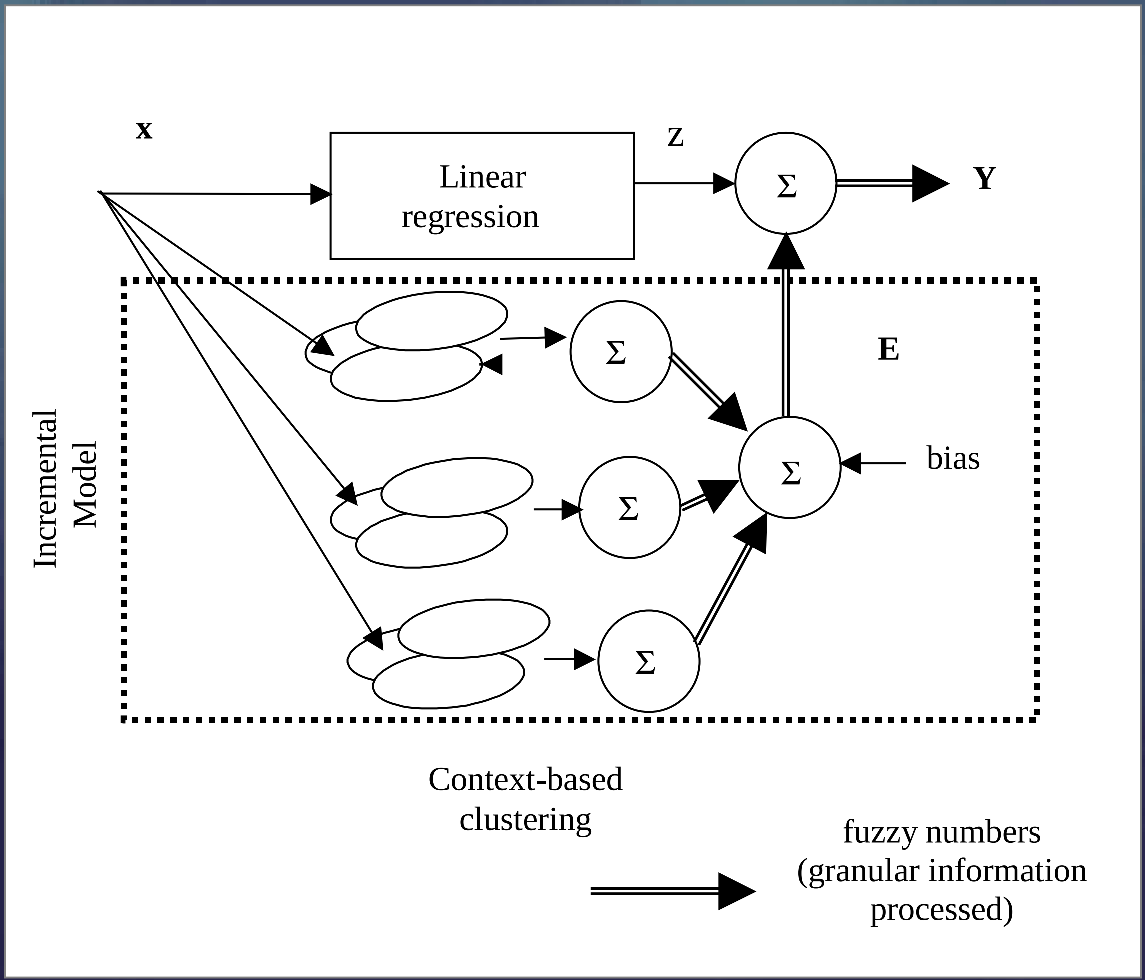
Fuzzy model = linear regression + local granular models

The principle of incremental fuzzy models and its design and architecture

General flow of development

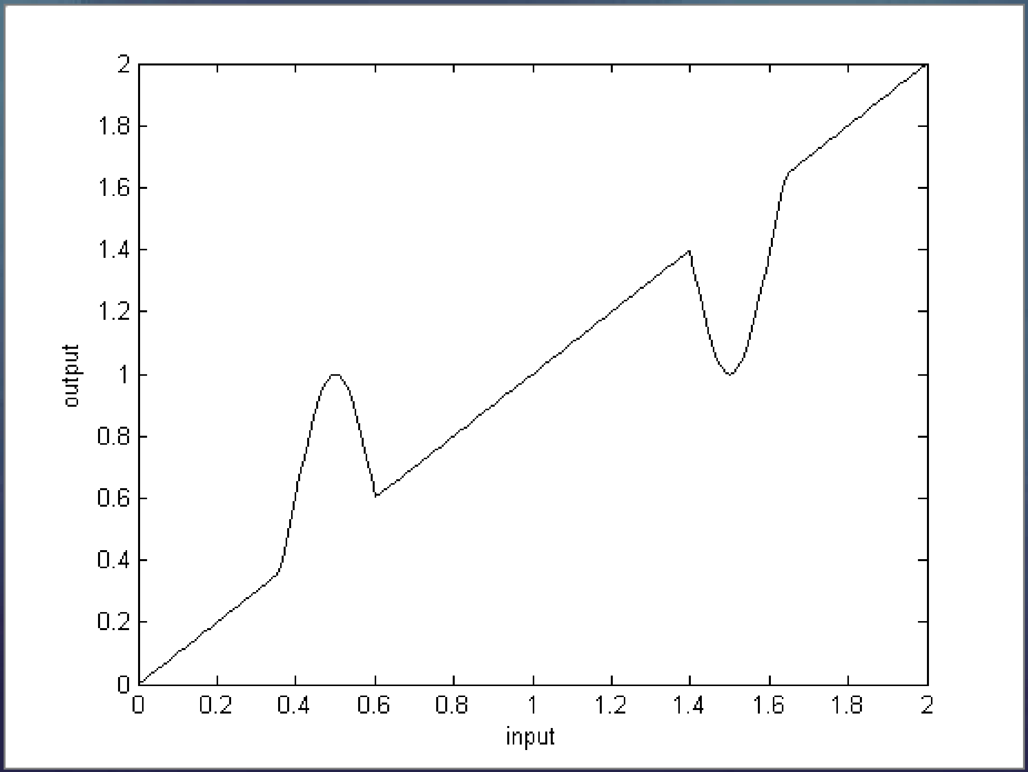


Overall flow processing of incremental granular models



Pedrycz and Gomide, FSE 2007

Example



$$G(x) = \exp\left(\frac{-(x-m)^2}{2\sigma^2}\right)$$

$$\mathrm{spiky}(x) = \begin{cases} \max(x, G(x)) & 0 \le x \le 1 \\ \min(x, -G(x)+2) & 1 < x \le 2 \end{cases}$$

$m = 0.5 \quad \sigma = 0.1$

RMSE values (means and standard deviation) – Training Data

		No. of contexts (p)			
		3	4	5	6
No. of clusters per context (c)	2	0.148±0.013	0.142 ± 0.018	0.136 ± 0.005	0.106 ± 0.006
	3	0.141 ± 0.012	0.131 ± 0.008	0.106 ± 0.008	0.087 ± 0.006
	4	0.143 ± 0.006	0.124 ± 0.007	0.095 ± 0.007	0.078 ± 0.005
	5	0.131 ± 0.012	0.111 ± 0.007	0.077 ± 0.008	0.073 ± 0.006
	6	0.126 ± 0.011	0.105 ± 0.005	0.072 ± 0.007	0.061 ± 0.007

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(y_k - \mathrm{target}_k)^2}$$

Pedrycz and Gomide, FSE 2007

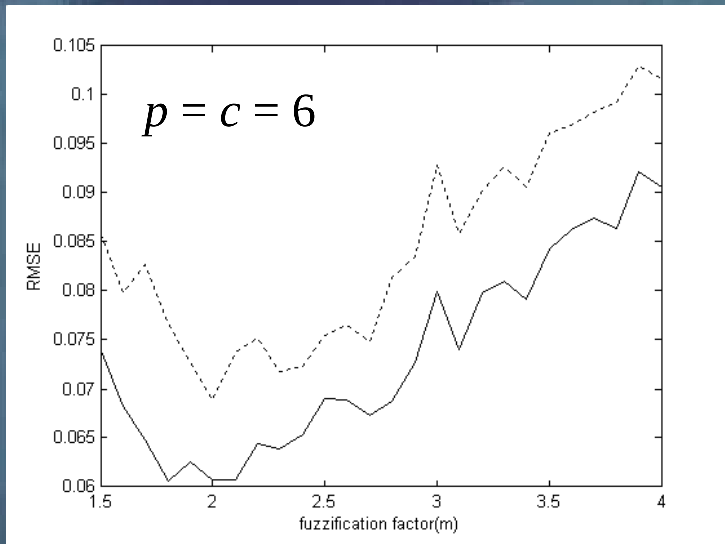
RMSE values (means and standard deviation) – Testing Data

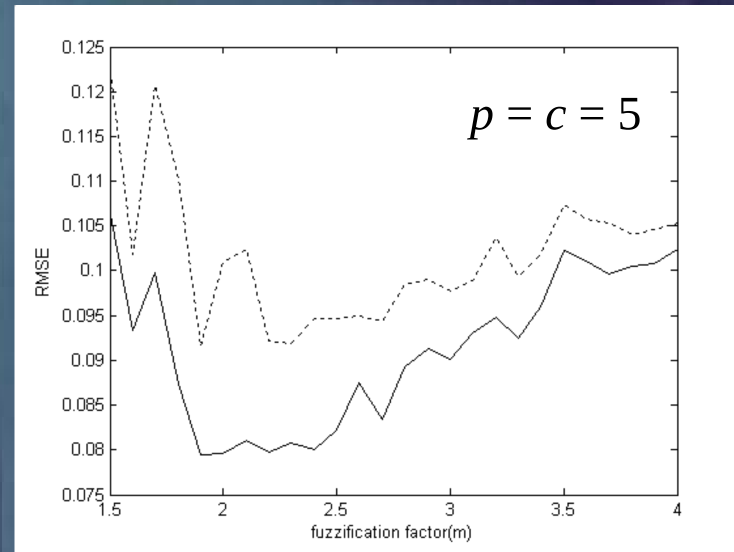
		No. of contexts (p)			
		3	4	5	6
No. of clusters per context (c)	2	0.142±0.016	0.139 ± 0.028	0.139 ± 0.012	0.114 ± 0.007
	3	0.131 ± 0.007	0.125 ± 0.017	0.115 ± 0.009	0.096 ± 0.009
	4	0.129 ± 0.014	0.126 ± 0.014	0.101 ± 0.009	0.085 ± 0.012
	5	0.123 ± 0.005	0.119 ± 0.016	0.097 ± 0.008	0.082 ± 0.010
	6	0.119 ± 0.016	0.114 ± 0.015	0.082 ± 0.011	0.069 ± 0.007

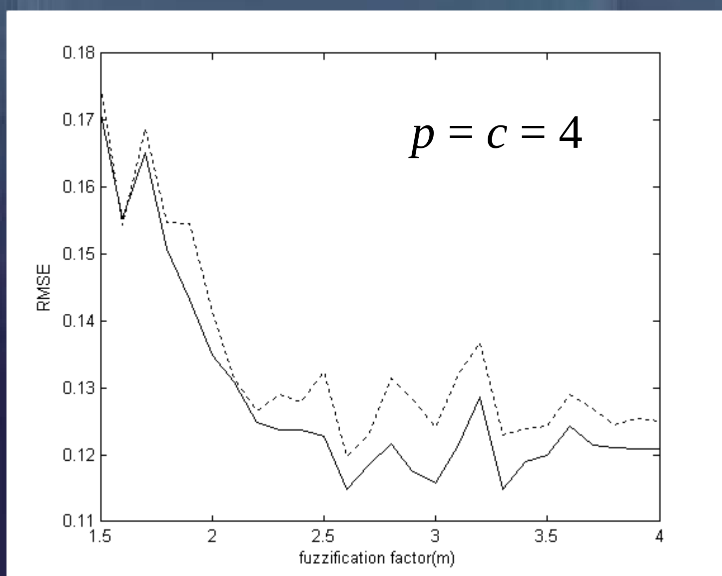
Pedrycz and Gomide, FSE 2007

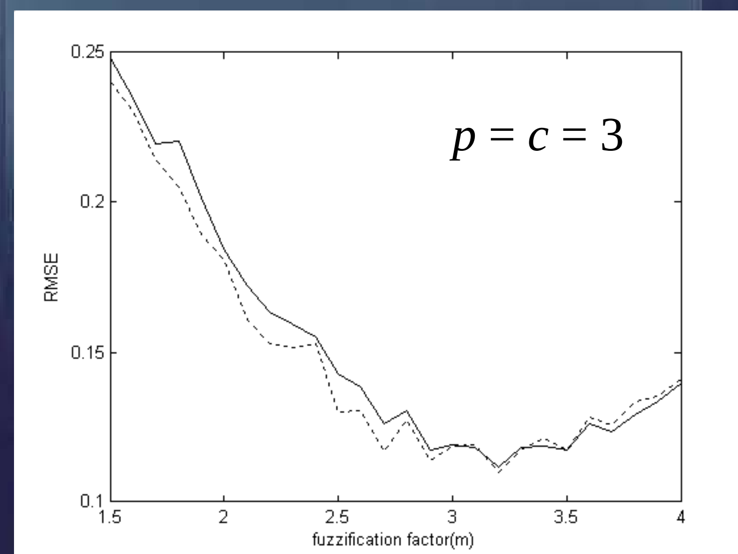
Optimal Values for fuzzification coefficient

		No. of contexts (*p*)			
		3	4	5	6
No. of clusters per context (*c*)	2	3.5	4.0	3.8	3.1
	3	3.2	3.9	3.5	3.1
	4	3.0	2.7	2.6	2.6
	5	3.1	2.8	2.2	2.4
	6	3.0	2.5	2.2	2.0









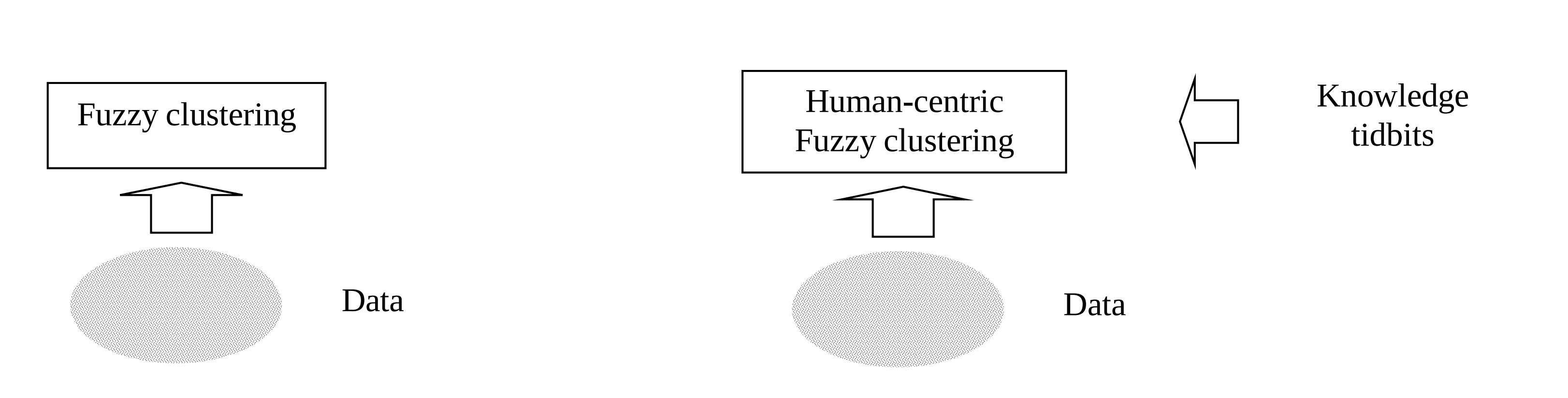
Training data

Testing data

Pedrycz and Gomide, FSE 2007

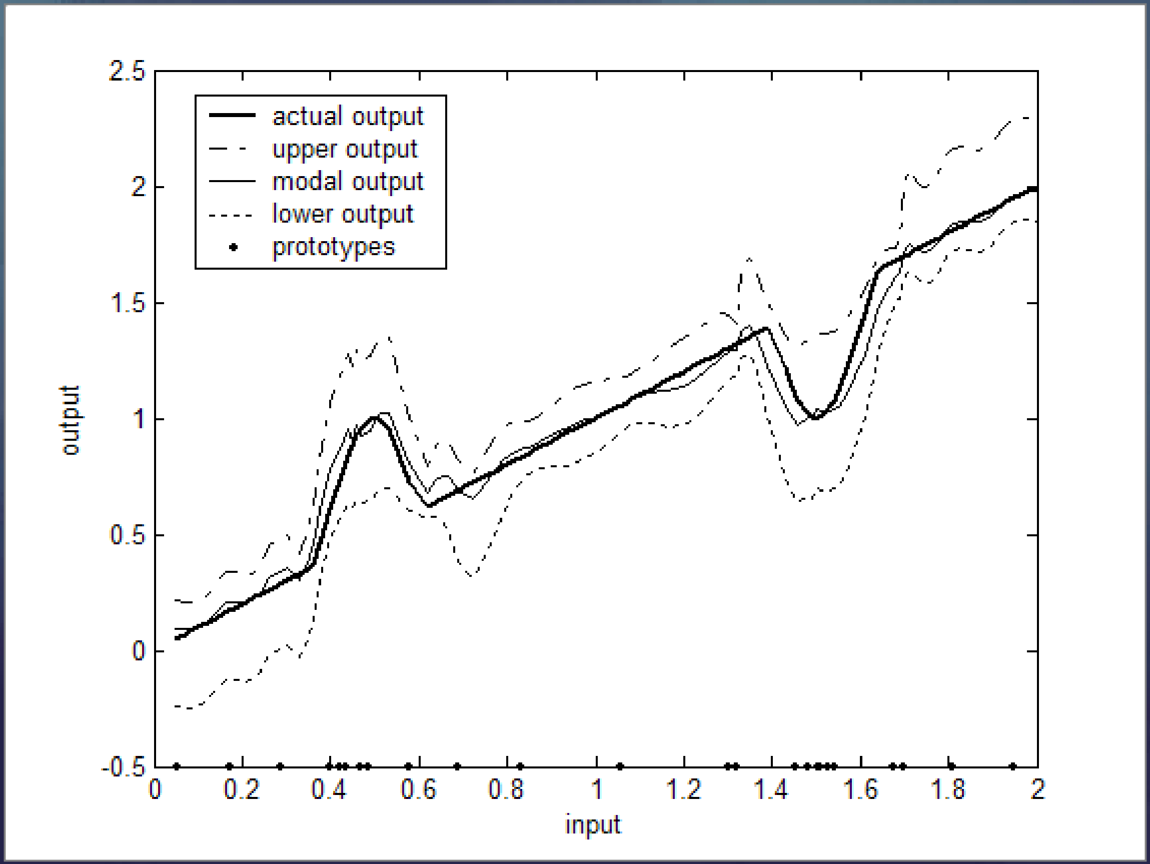
14.7 Human-Centric fuzzy clustering

Human-Centric clustering



- Human-Centric ≡ knowledge-Based clustering
- Clusters reflect human-driven customization
- Clustering algorithms consider knowledge about data

Example

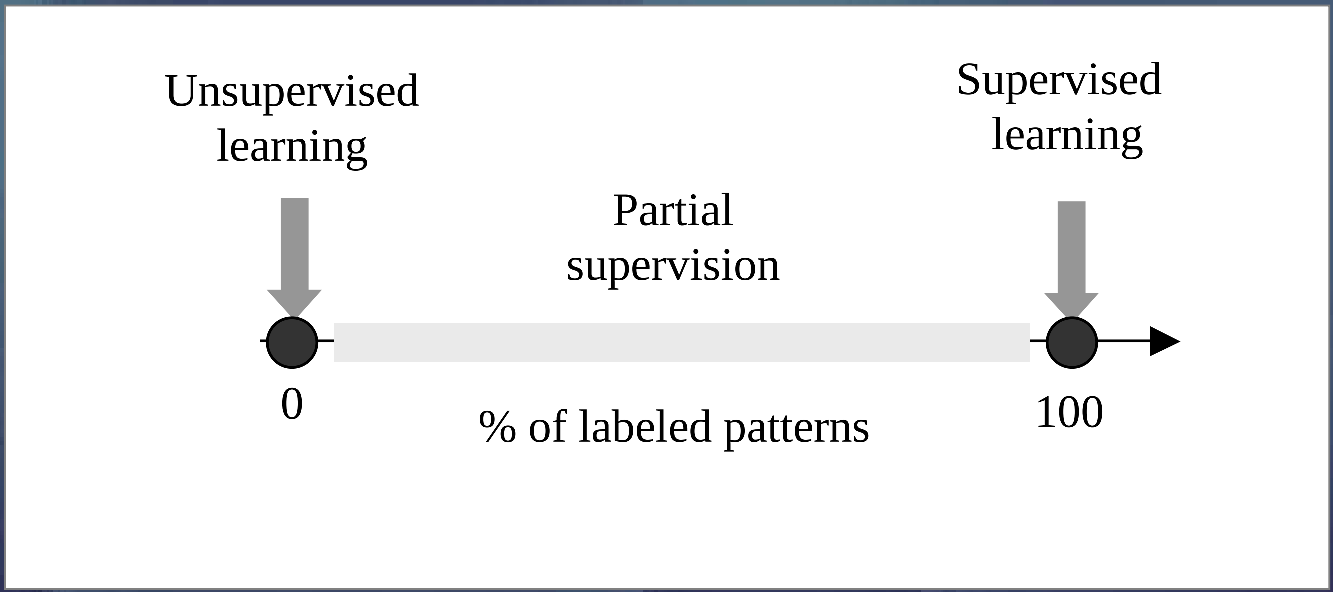


$c = 5.0$
$p = 5.0$
$m = 2.2$

Human-Centric clustering approaches

- fuzzy clustering with partial supervision
  - human-centric clusters
- proximity-based fuzzy clustering

Fuzzy clustering with partial supervision



- involves a subset of labeled patterns
- subset of labeled patterns comes with class membership

Clustering algorithm

$$Q = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^2 d_{ik}^2 + \alpha \sum_{i=1}^{c} \sum_{k=1}^{N} (u_{ik} - f_{ik} b_k)^2 d_{ik}^2$$

b = $(b_1, b_2,..,b_N)$

$b_k = 1$ if pattern $\mathbf{x}_k$ is labeled, $b_k = 0$ otherwise

$F = [f_{ik}]$ $i = 1, 2,....,c$; $k = 1, 2, ..., N$

F contains membership grades assigned to patterns
α = weight factor to capture effect of partial supervision

Development of human-centric clusters

$$V = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^2 d_{ik}^2 + \alpha \sum_{i=1}^{c} \sum_{k=1}^{N} (u_{ik} - f_{ik} b_k)^2 d_{ik}^2 - \lambda\left(\sum_{i=1}^{c} u_{ik} - 1\right)$$

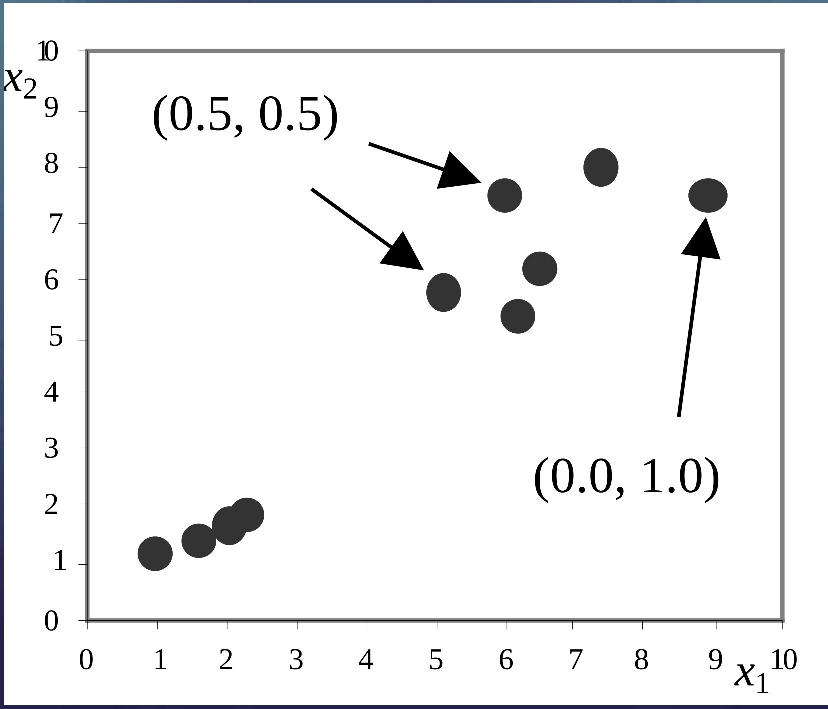
$$u_{ik} = \frac{1}{1+\alpha} \left[\frac{1 + \alpha\left(1 - b_k \sum_{i=1}^{c} f_{ik}\right)}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}}\right)^2} + \alpha f_{ik} b_k \right]$$

$$\mathbf{v}_s = \frac{\sum_{k=1}^{N} \psi_{sk} \mathbf{x}_k}{\sum_{k=1}^{N} \psi_{sk}}$$

$$\psi_{sk} = u_{ik}^2 + (u_{ik} - f_{ik} b_k)^2$$

Pedrycz and Gomide, FSE 2007

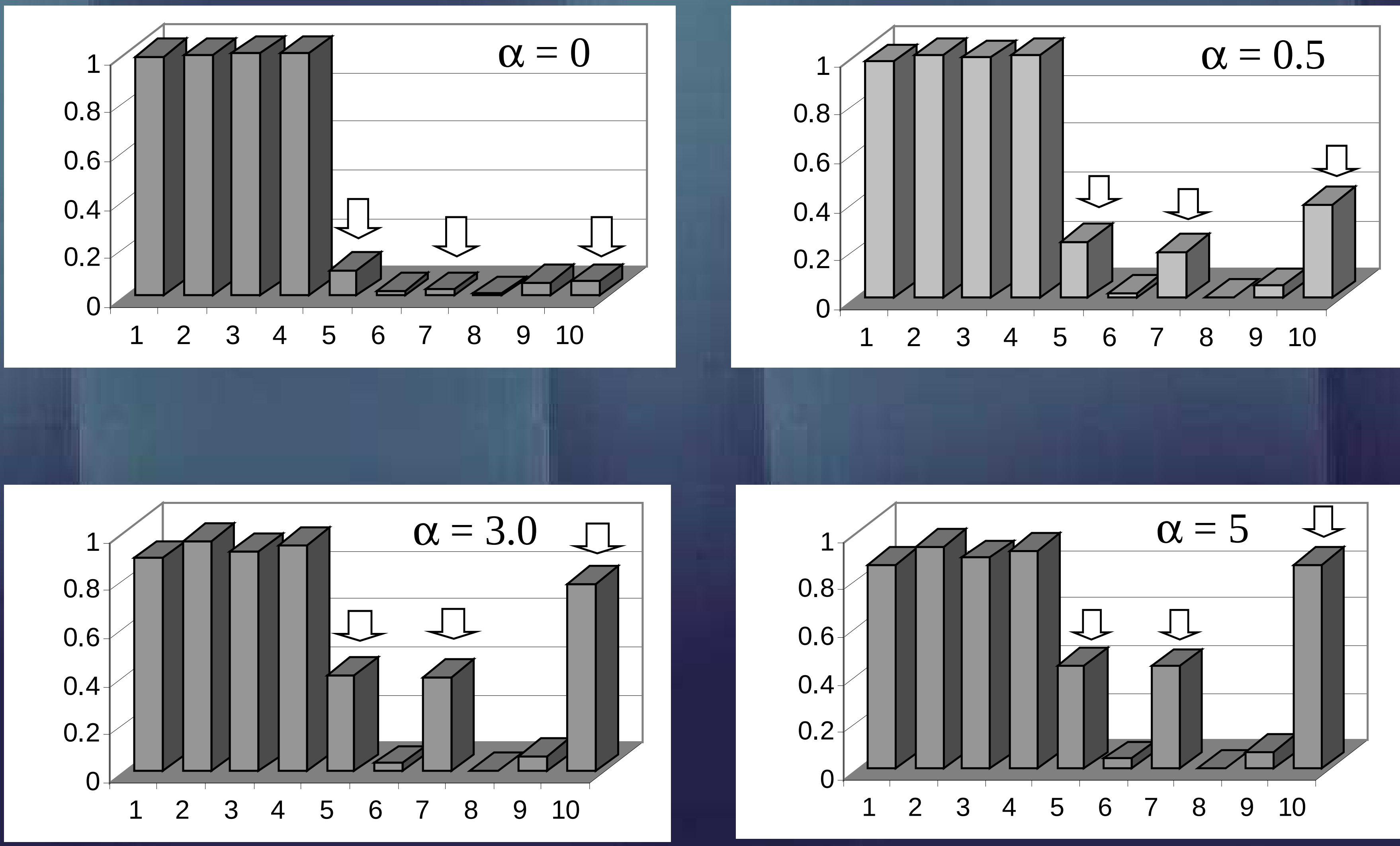
Example



Tidbits (hints)

Membership grades

Membership grades of patterns



Proximity-based fuzzy clustering

- Proximity between two objects (patterns)
  - $\text{prox}(a, b) = \text{prox}\,(b, a)$ symmetry
  - $\text{prox}(a, a) = 1$ reflexivity
- Collection of patterns: proximity relation (matrix form) P

Proximity- based fuzzy clustering

$$\hat{p}[k_1, k_2] = \sum_{i=1}^{c} \min(u_{ik_1}, u_{ik_2})$$

Patterns: $\mathbf{x}_{k1}$, $\mathbf{x}_{k2}$

$$\hat{P} = [\hat{p}[k_1, k_2]]$$

$k_1, k_2 = 1, \ldots., N$

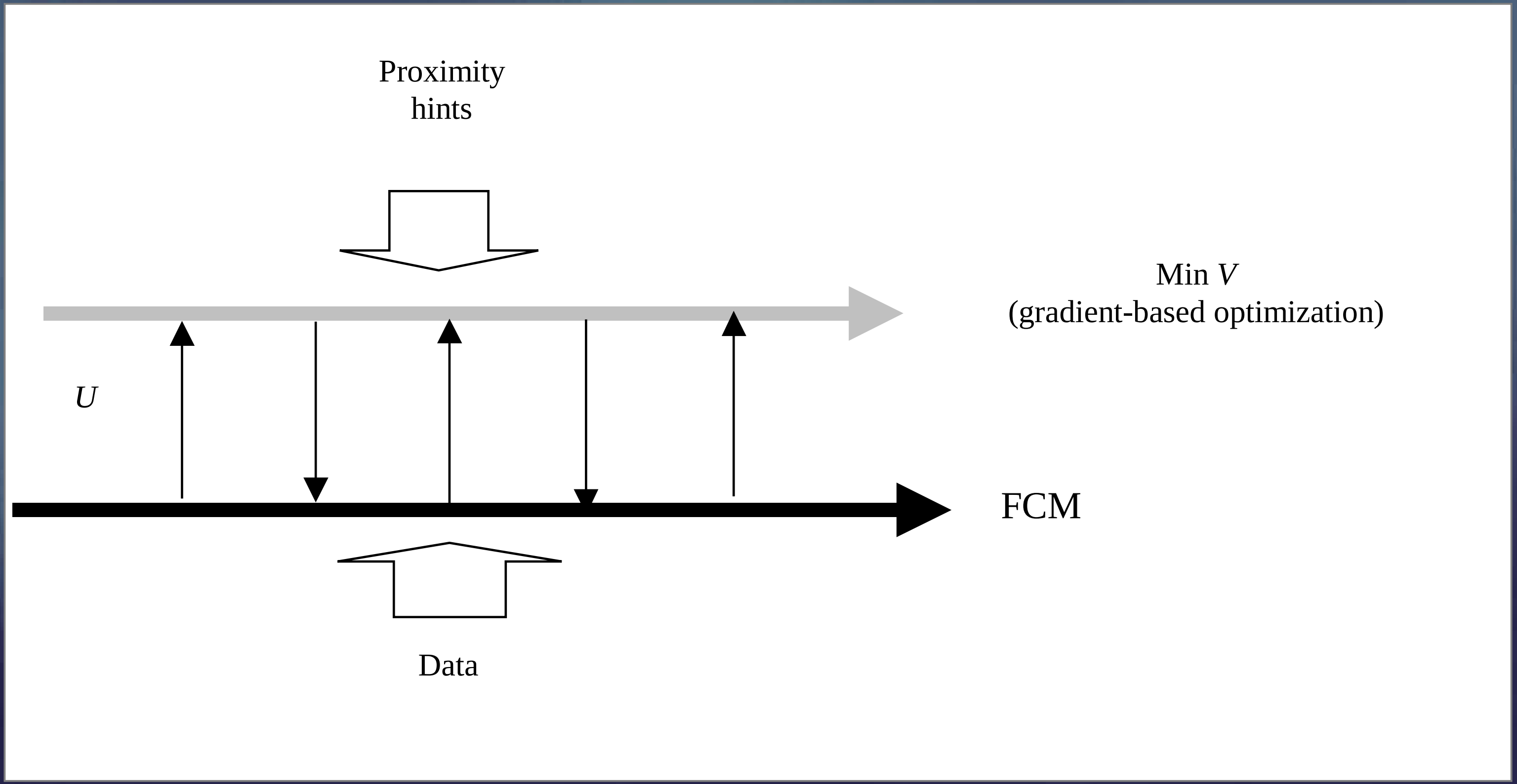
$$V = \sum_{k_1=1}^{N} \sum_{k_2=1}^{N} (\hat{p}[k_1, k_2] - p[k_1, k_2])^2 b[k_1, k_2] d[k_1, k_2]$$

P-FCM clustering algorithm

procedure P-FCM-CLUSTERING (**X**) **returns** cluster centers and partition matrix
input: data set $\mathbf{X}=\{\mathbf{x}_k, k=1,..,N\}$
local: fuzzification coefficient: *m*
thresholds: δ, ε

INITIALIZE-PARTITION-MATRIX
repeat until distance two successive partition matrices ≤ δ
run FCM
repeat until values of *V* over successive iterations ≤ ε
minimize *V*
compute u_{ik}
compute $\mathbf{v}_s$
return cluster centers and partition matrix

P-FCM optimization steps

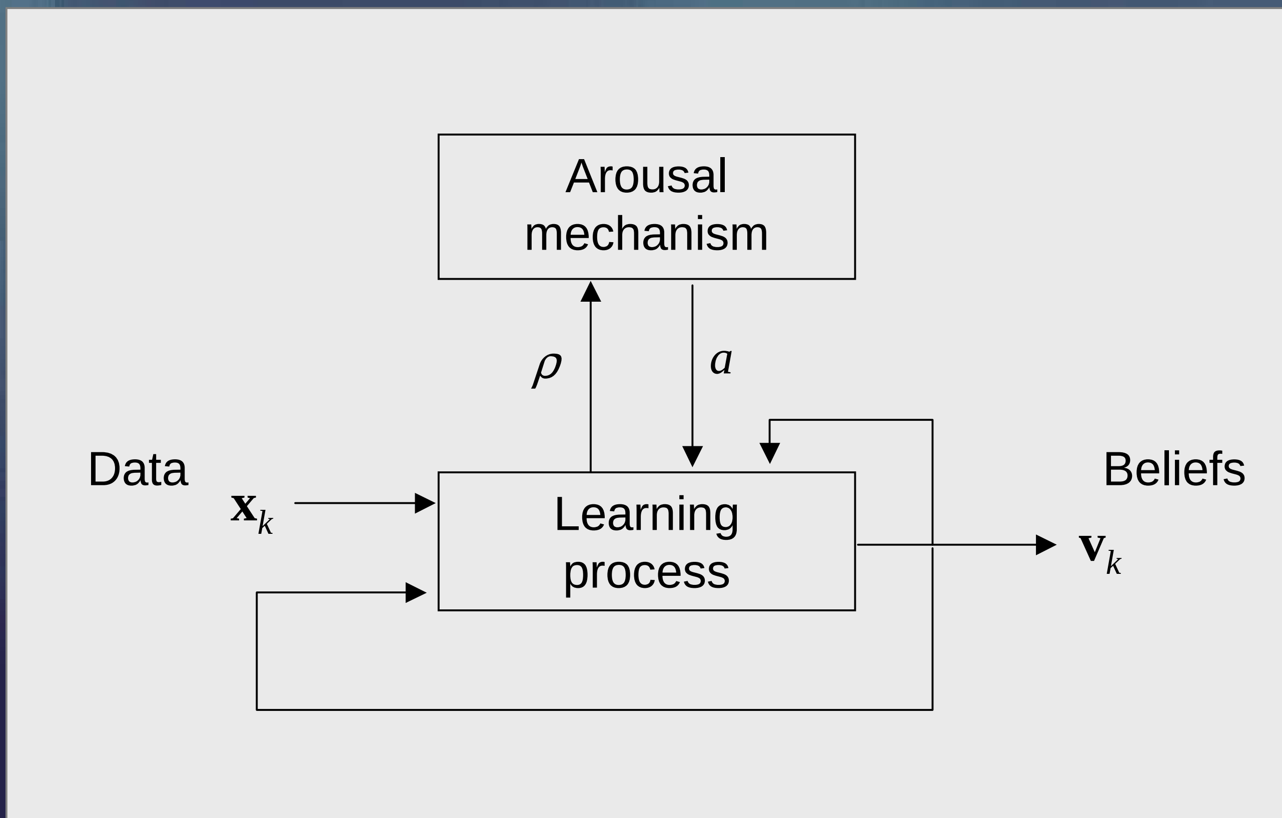


Interaction aspects of sources of information in the P-FCM

- P-FCM augments FCM adding extra optimization using patterns
- P-FCM reconcile structural and domain information
- Computationally, P-FCM does not affect size of original dataset
- P-FCM dwells on the core part of FCM optimization scheme

14.8 Participatory learning fuzzy clustering

Participatory learning



Participatory learning updates

$$\mathbf{v}_{k+1} = \mathbf{v}_k + \alpha(\rho)^{1-a_k}(\mathbf{x}_k - \mathbf{v}_k)$$

$$\rho_k = 1 - d_k \qquad d_k = \| \mathbf{x}_k - \mathbf{v}_k \|$$

$$a_{k+1} = a_k + \beta((1-\rho_{k+1}) - a_k)$$

$$\mathbf{x}_k \in [0,1]^n, \ \ a_k \in [0,1], \ \ \beta \in [0,1]$$

Distance measure and membership degree assignment

$$d_{ik} = (\mathbf{x}_k - \mathbf{v}_i)^T (\det(f_i)^{1/N} F_i^{-1})(\mathbf{x}_k - \mathbf{v}_i)$$

$$F_i = \frac{\sum_{k=1}^{N} u_{ik}^m (\mathbf{x}_k - \mathbf{v}_i)(\mathbf{x}_k - \mathbf{v}_i)^T}{\sum_{k=1}^{N} u_{ik}^m}$$

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} (d_{ik} / d_{jk})^{2/m-1}}$$

Mahalanobis distance

PL clustering algorithm (off-line)

procedure OFF-LINE-PARTICIPATORY (X) **returns** cluster centers and partition matrix
input: data set: $X = \{\mathbf{x}_k, k = 1,..,N\}$
local: cluster membership parameter: m
threshold: τ
learning rates: α, β
parameters: ε, l_{max}

V=INITIALIZE-CLUSTER-CENTERS(X)
$l = 1$
until stop = TRUE **do**
for $k = 1{:}N$
CLUSTER-LEARNING($\mathbf{x}_k$,V)
if $\|\Delta V\| \leq \varepsilon$ and $l \geq l_{max}$ **then** update U, set stop = TRUE
else $l = l + 1$
return V, U

PL clustering algorithm (on-line)

procedure OFF-LINE-PARTICIPATORY (**x**) **returns** cluster centers and partition matrix
input: data: **x**
local: cluster membership parameter: m
threshold: τ
learning rates: α, β
parameters: ε, l_{max}

V=INITIALIZE-CLUSTER-CENTERS(**x**)

do forever

CLUSTER-LEARNING(**x**, *V*)

return *V*, *U*

Clustering learning procedure

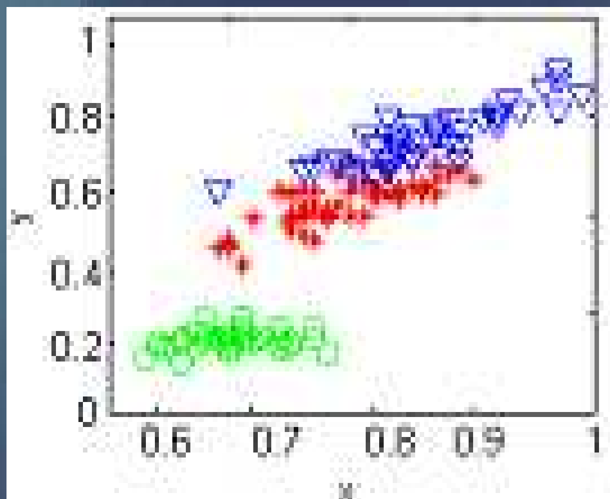
procedure CLUSTER-LEARNING(**x**) **returns** cluster centers and partition matrix

input: $\mathbf{x}_k = \mathbf{x}$

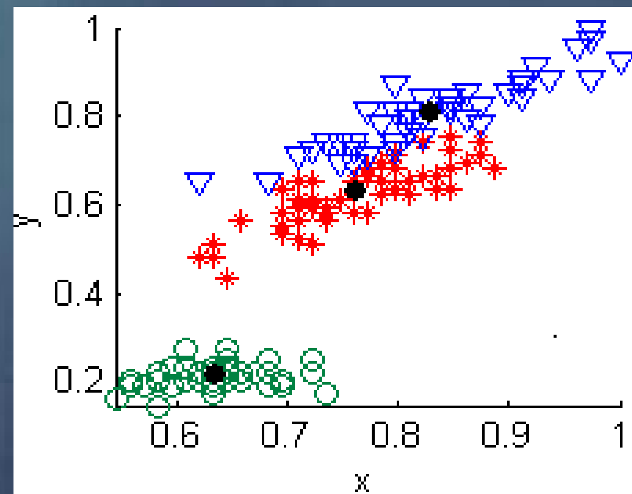
for $i = 1{:}c$
    compute d_{ik}
    compute ρ_{ik}
    compute a_{ik}
      if $a_{ik} \le \tau$ for all $i = 1,\ldots,c$
        then update $\mathbf{v}_s$, , compute U
        else create new cluster center
            for $i = 1{:}c$
                for $j = (i + 1){:}c$
                    compute ρ_{vi}
                    compute λ_{vi}
                    if $\lambda_{vi} \le 0.95\,\tau$
                      then eliminate $\mathbf{v}_i$ and update U
return V, U

$$s = \arg\max_{i}\{\rho_{ik}\}$$

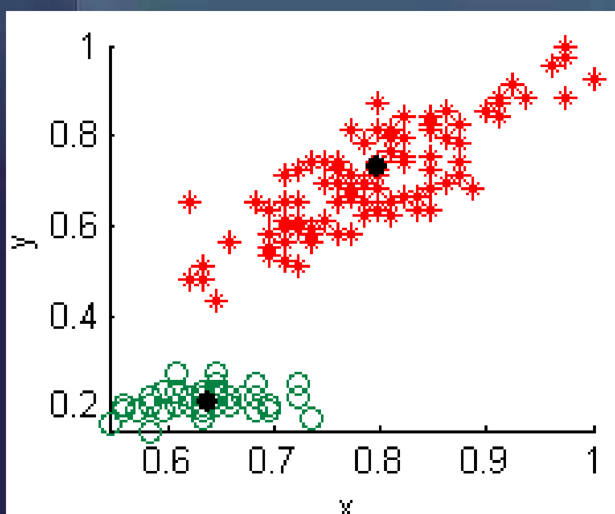
Example



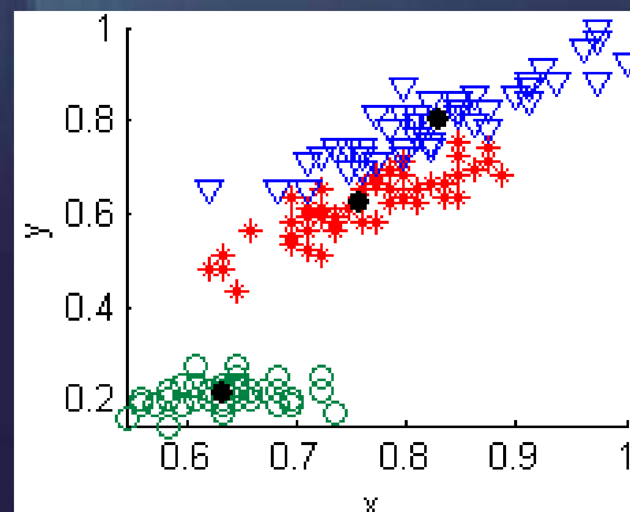
Data



GK



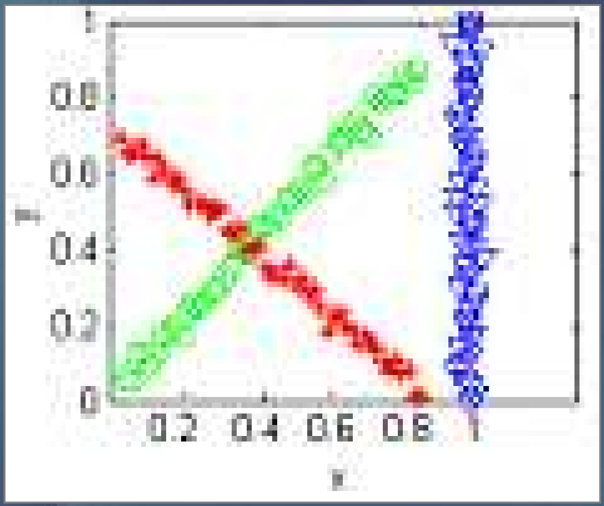
MKFM



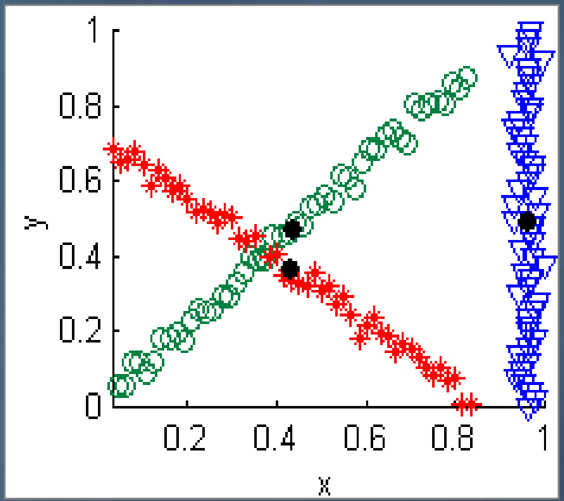
PL

Pedrycz and Gomide, FSE 2007

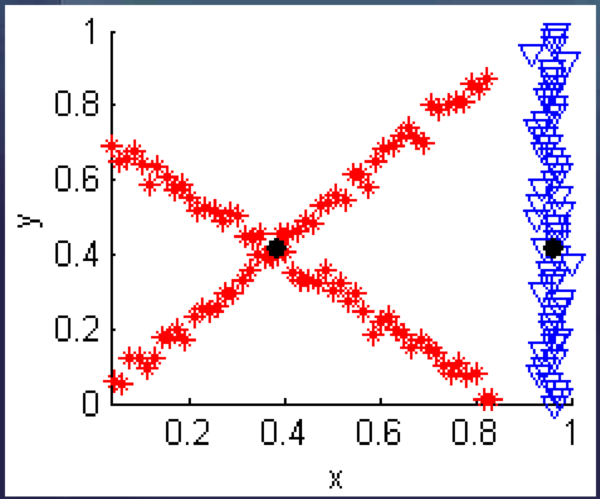
Example



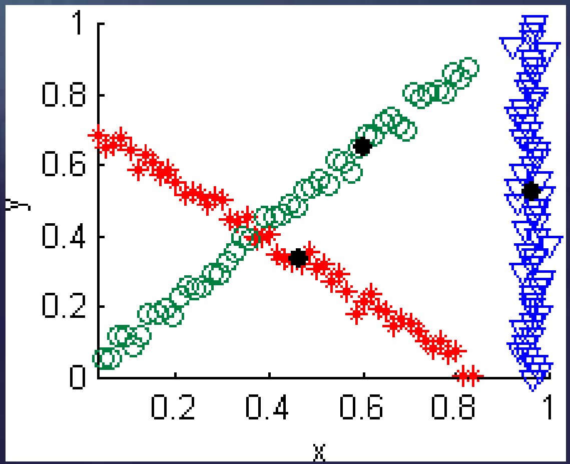
Data



GK



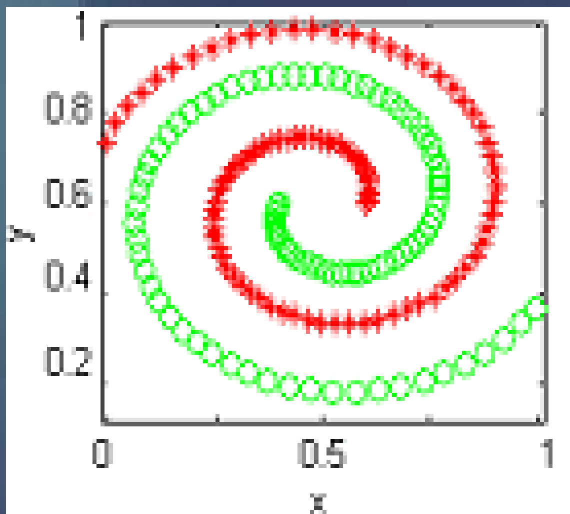
MKFM



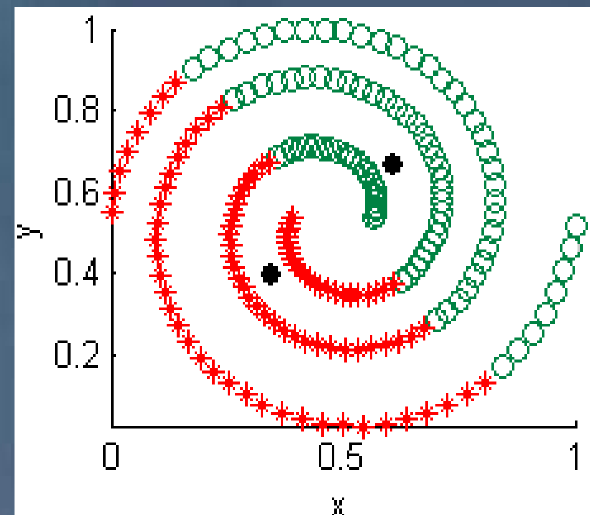
PL

Pedrycz and Gomide, FSE 2007

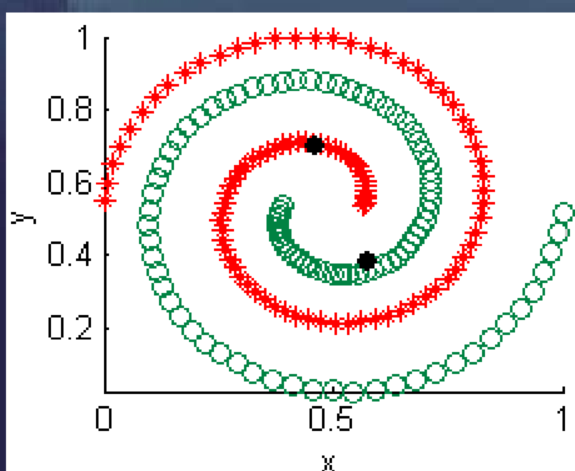
Example



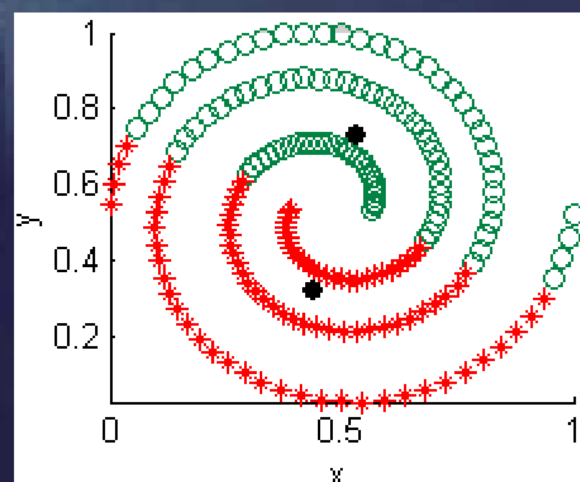
Data



GK



MKFM



PL

Pedrycz and Gomide, FSE 2007