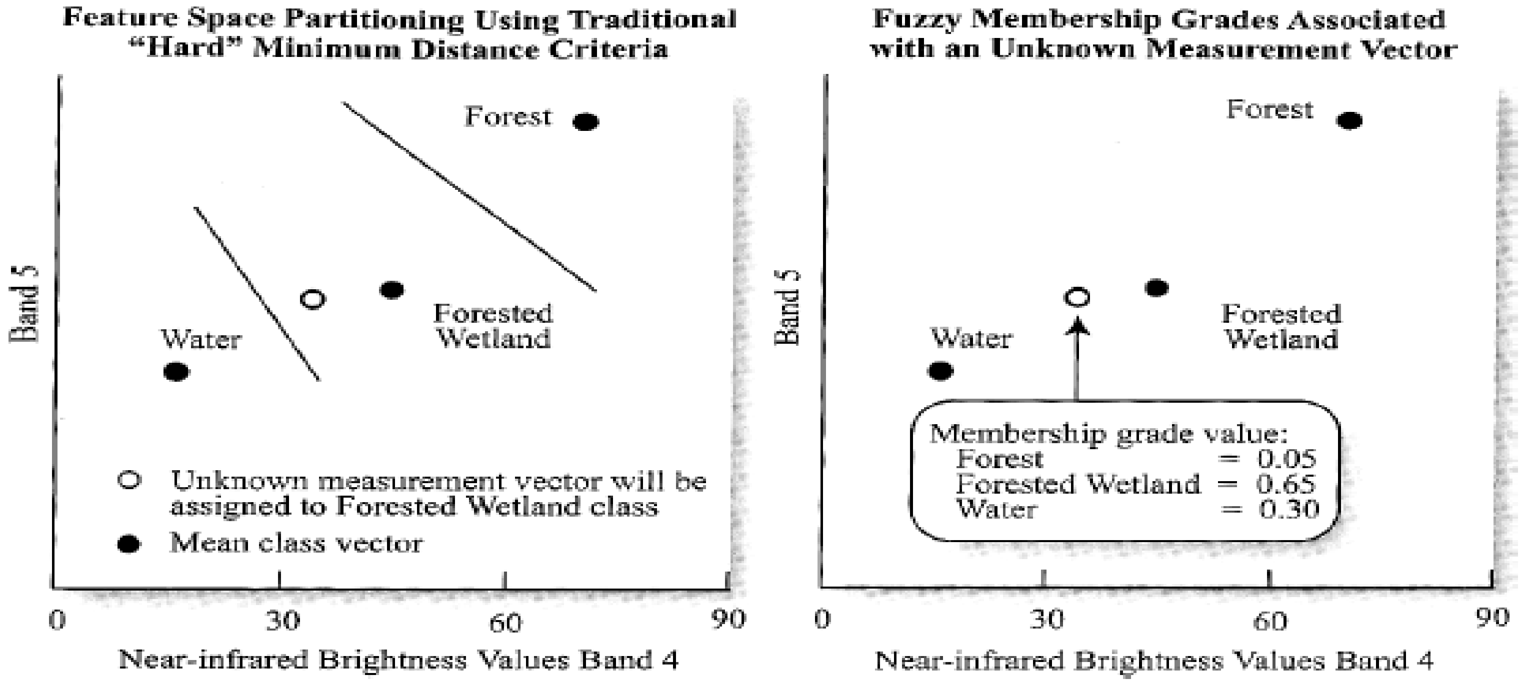
Fuzzy Clustering

Hard vs Soft Classification

Hard- versus soft-classifiers



Why use soft-classifiers?

- Sub-pixel classification
- Uncertainty of classification/scheme
- Incorporating ancillary data (hardeners)

Fuzzy Clustering

Let $X = \{x_1, x_2, \ldots, x_n\}$ be a set of given data. A fuzzy pseudopartition or fuzzy c-partition of X is a family of fuzzy subsets of X, denoted by $\mathcal{P} = \{A_1, A_2, \ldots, A_c\}$, which satisfies

$$\sum_{i=1}^{c} A_i(x_k) = 1$$

for all $k \in \mathbb{N}_n$ and

$$0 < \sum_{k=1}^{n} A_i(x_k) < n$$

for all $i \in \mathbb{N}_c$, where c is a positive integer.

For instance, given $X = \{x_1, x_2, x_3\}$ and

$$A_1 = .6/x_1 + 1/x_2 + .1/x_3,$$

$$A_2 = .4/x_1 + 0/x_2 + .9/x_3,$$

$$\mathbf{v}_i = \frac{\sum_{k=1}^{n} [A_i(\mathbf{x}_k)]^m \mathbf{x}_k}{\sum_{k=1}^{n} [A_i(\mathbf{x}_k)]^m}$$

$$J_m(\mathcal{P}) = \sum_{k=1}^{n} \sum_{i=1}^{c} [A_i(\mathbf{x}_k)]^m \|\mathbf{x}_k - \mathbf{v}_i\|^2,$$

Fuzzy c-means

The algorithm is based on the assumption that the desired number of clusters c is given and, in addition, a particular distance, a real number $m \in (1, \infty)$, and a small positive number ε, serving as a stopping criterion, are chosen.

Step 1. Let $t = 0$. Select an initial fuzzy pseudopartition $\mathcal{P}^{(0)}$.

Step 2. Calculate the c cluster centers $\mathbf{v}_1^{(t)}, \ldots, \mathbf{v}_c^{(t)}$ by (13.3) for $\mathcal{P}^{(t)}$ and the chosen value of m.

Step 3. Update $\mathcal{P}^{(t+1)}$ by the following procedure: For each $\mathbf{x}_k \in X$, if $\|\mathbf{x}_k - \mathbf{v}_i^{(t)}\|^2 > 0$ for all $i \in \mathbb{N}_c$, then define

$$A_i^{(t+1)}(\mathbf{x}_k) = \left[\sum_{j=1}^{c}\left(\frac{\|\mathbf{x}_k - \mathbf{v}_i^{(t)}\|^2}{\|\mathbf{x}_k - \mathbf{v}_j^{(t)}\|^2}\right)^{\frac{1}{m-1}}\right]^{-1};$$

if $\|\mathbf{x}_k - \mathbf{v}_i^{(t)}\|^2 = 0$ for some $i \in I \subseteq \mathbb{N}_c$, then define $A_i^{(t+1)}(\mathbf{x}_k)$ for $i \in I$ by any nonnegative real numbers satisfying

$$\sum_{i \in I} A_i^{(t+1)}(\mathbf{x}_k) = 1,$$

and define $A_i^{(t+1)}(\mathbf{x}_k) = 0$ for $i \in \mathbb{N}_c - I$.

Step 4. Compare $\mathcal{P}^{(t)}$ and $\mathcal{P}^{(t+1)}$. If $|\mathcal{P}^{(t+1)} - \mathcal{P}^{(t)}| \le \varepsilon$, then stop; otherwise, increase t by one and return to **Step 2**.

In Step 4, $|\mathcal{P}^{(t+1)} - \mathcal{P}^{(t)}|$ denotes a distance between $\mathcal{P}^{(t+1)}$ and $\mathcal{P}^{(t)}$ in the space $\mathbb{R}^{n \times c}$. An example of this distance is

$$|\mathcal{P}^{(t+1)} - \mathcal{P}^{(t)}| = \max_{i \in \mathbb{N}_c, k \in \mathbb{N}_n} |A_i^{(t+1)}(\mathbf{x}_k) - A_i^{(t)}(\mathbf{x}_k)|.$$

P=

	A1	A2	A3	
X1	0.2	0.4	0.4	=1
X2	0.3	0.5	0.2	
X3	0.1	0.9	0.0	
X4	0.8	0.1	0.1	
X5	0.9	0.1	0.0	

A1(X1) =0.2

U_{A1}(X1)=0.2

$$v_i = \frac{\sum_{k=1}^{n}[A_i(x_k)]^m x_k}{\sum_{k=1}^{n}[A_i(x_k)]^m}$$

Sample calculation

- Ai (Xk)   A1(X1)=0.2
- X1=(1,1)=(x11,x12)
- X2=(2,3)
- X3=(-4,-3)
- X4=(0,0)
- X5=(4,-4)

V1=(V11,V12), V2=(V21,V22)

V11= [A1(X1)*A1(X1)*X11 + A1(X2)*A1(X2)*X21 + A1(X3)*A1(X3)*X31 + A1(X4)*A1(X4)*X41 + A1(X5)*A1(X5)*X51]/[A1(X1)*A1(X1) + A1(X2)*A1(X2) + A1(X3)*A1(X3) + A1(X4)*A1(X4) + A1(X5)*A1(X5)]

V11= [0.2*0.2*1 + 0.3*0.3*2 + 0.1*0.1* -4 + 0.8*0.8 *0 + 0.9*0.9 * 4]/ [0.2*0.2+ 0.3*0.3+ 0.1*0.1+ 0.8*0.8+ 0.9*0.9] = 3.42/1.59 = 2.15

V12= -1.86;   V1 = {2.15, -0.02}; V2= ?

P=	A1	A2	A3	
X1	0.2	0.4	0.4	=1
X2	0.3	0.5	0.2	
X3	0.1	0.9	0.0	
X4	0.8	0.1	0.1	
X5	0.9	0.1	0.0	

$$v_i = \frac{\sum_{k=1}^{n} [A_i(x_k)]^m x_k}{\sum_{k=1}^{n} [A_i(x_k)]^m}$$

It will be Continued....