

EL7025/EL761/CI6308

Control Inteligente para Problemas Dinámicos de Transporte

Dra. Doris Sáez,

Departamento de Ingeniería Eléctrica

Universidad de Chile

Dr. Cristián Cortés

Departamento de Ingeniería Civil – Transporte

Universidad de Chile



Fuzzy Clustering



- Concepto de Cluster.
- Tipos de Clustering
 - Hard clustering.
 - Fuzzy Clustering.
- Algoritmo Fuzzy C-Means
- Ejemplo:

Concepto básico de Cluster

- ◆ Un cluster corresponde a un conjunto de datos que poseen alguna similitud entre ellos.
- ◆ En general, existen 2 tipos principales de Clustering:
 - Hard Clustering.
 - Clustering Difuso.

Concepto básico de Cluster

- ◆ Se introducen las funciones de pertenencia que tratan de cuantificar la pertenencia de un elemento a un cierto cluster. Este concepto cobra un sentido casi intuitivo durante el estudio del clustering difuso.

Hard Clustering

- ◆ Se tiene un set de datos Z que es particionado en “ c ” clusters o conjuntos, donde cada uno de estos conjuntos es disjunto de los $c-1$ restantes, ie:

$$\bigcup_{i=1}^c A_i = Z.$$

$$A_i \cap A_j = \emptyset, 1 \leq i \neq j \leq c.$$

$$\emptyset \subset A_i \subset Z, 1 \leq i \leq c.$$

Hard Clustering

- ◆ En términos de funciones de pertenencia:

$$\mu_{ik} \in \{0,1\}, \quad 1 \leq i \leq c, 1 \leq k \leq N.$$

$$\sum_{i=1}^c \mu_{ik} = 1, \quad 1 \leq k \leq N.$$

$$0 < \sum_{i=1}^N \mu_{ik} < N, \quad 1 \leq i \leq c$$

N : Número de datos.

c : Número de Clusters

Slide preparada por R. Zuñiga (2005)

Hard Clustering

- ♦ Una definición matemática precisa es:

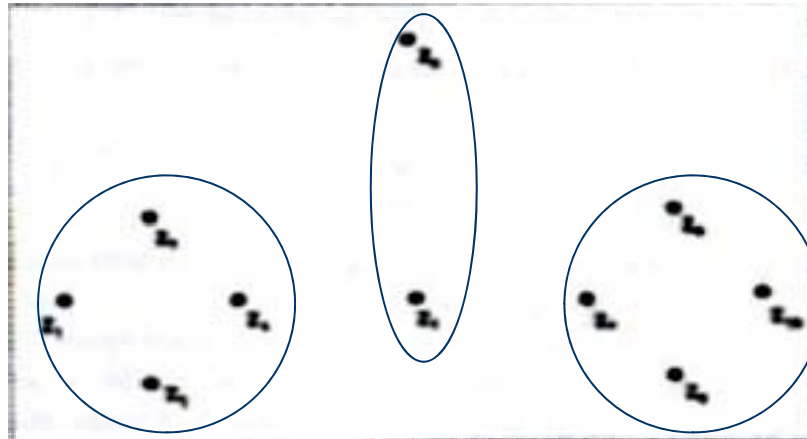
$Z = \{z_1, z_2, z_3, \dots, z_N\}$ set finito de datos.

$2 \leq c < N$, c es un entero; Entonces el hard Clustering de Z es :

$$M_{hc} = \left\{ U \in R^{c \times N} \mid \mu_{ik} \in \{0,1\}, \forall i, k; \sum_{i=1}^c \mu_{ik} = 1, \forall k; 0 < \sum_{i=1}^N \mu_{ik} < N, \forall i \right\}$$

Hard Clustering

- ◆ Ejemplo:



Fuzzy Clustering

- ◆ El clustering difuso corresponde a una generalización natural del Hard clustering, en donde se permite que los grados de pertenencia tomen valores continuos entre 0 y 1 para los distintos elementos que componen los clusters.

Fuzzy Clustering

- ◆ Las condiciones para establecer una partición difusa son:

$$\mu_{ik} \in [0,1], \quad 1 \leq i \leq c, \quad 1 \leq k \leq N$$

$$\sum_{i=1}^c \mu_{ik} = 1, \quad 1 \leq k \leq N$$

$$0 < \sum_{i=1}^N \mu_{ik} < N, \quad 1 \leq i \leq c$$

Fuzzy Clustering

- ♦ La definición matemática precisa es:

$Z = \{z_1, z_2, z_3, \dots, z_N\}$ set finito de datos.

$2 \leq c < N$, c es un entero; Entonces el Fuzzy Clustering de Z es :

$$M_{fc} = \left\{ U \in R^{c \times N} \mid \mu_{ik} \in [0,1], \forall i, k; \sum_{i=1}^c \mu_{ik} = 1, \forall k; 0 < \sum_{i=1}^N \mu_{ik} < N, \forall i \right\}$$

Fuzzy Clustering

- ◆ Otro tipo de partición posible es la llamada partición posibilística, que es muy similar a la vista anteriormente, pero la única restricción que existe sobre los grados de pertenencia es que cada uno sea mayor que cero, permitiendo alcanzar valores mayores a uno.

Algoritmos de Clustering Difuso

- ◆ Naturalmente, elegir las posibles particiones y los correspondientes grados de pertenencia no es un problema trivial de resolver.
- ◆ Existe la necesidad de crear algoritmos que generen las particiones y los grados de pertenencia de forma automática.

Fuzzy C-Means Clustering

- ◆ Este es el algoritmo más utilizado para realizar las particiones difusas y se basa en la optimización de un funcional que representa matemáticamente la existencia de similitud entre los datos.
- ◆ Los demás algoritmos son muy similares salvo pequeñas modificaciones a la función objetivo.

Fuzzy C-Means Clustering

- ♦ La funcional a minimizar es:

$$J(Z, U, V) = \sum_{i=1}^c \sum_{k=1}^N (\mu_{ik})^m (z_k - v_i)^T A(z_k - v_i)$$

con:

$U = [\mu_{ik}] \in M_{fc}$, es una matriz que describe la partición difusa de Z .

$V = [v_1, v_2, v_3, \dots, v_c]$, $v_i \in R^n$, es un vector que contiene los centros de los clusters llamados *prototipos*, los cuales deben ser determinados.

$m \in [1, \infty)$ corresponde a un coeficiente que determina cuan difusos son los clusters.

Fuzzy C-Means Clustering

- ◆ Existen varias formas de resolver el problema de optimización no-lineal, pero la forma más utilizada es el llamado FCM (Fuzzy C-means), que es una simple iteración a través de las condiciones de primer orden para los puntos estacionarios.
- ◆ Dichos puntos pueden ser encontrados agregando restricciones al funcional y planteando el lagrangeano correspondiente.

Fuzzy C-Means Clustering

$$\bar{J}(Z, U, V, \lambda) = \sum_{i=1}^c \sum_{k=1}^N (\mu_{ik})^m D_{ikA}^2 + \sum_{k=1}^N \lambda_k \left[\sum_{i=1}^c \mu_{ik} - 1 \right]$$

Se obtienen las condiciones de Kunh - Tucker.

Si $D_{ikA}^2 > 0$ y $m > 1$, entonces los $(U, V) \in M_{fc} \times R^{n \times c}$ optimos son :

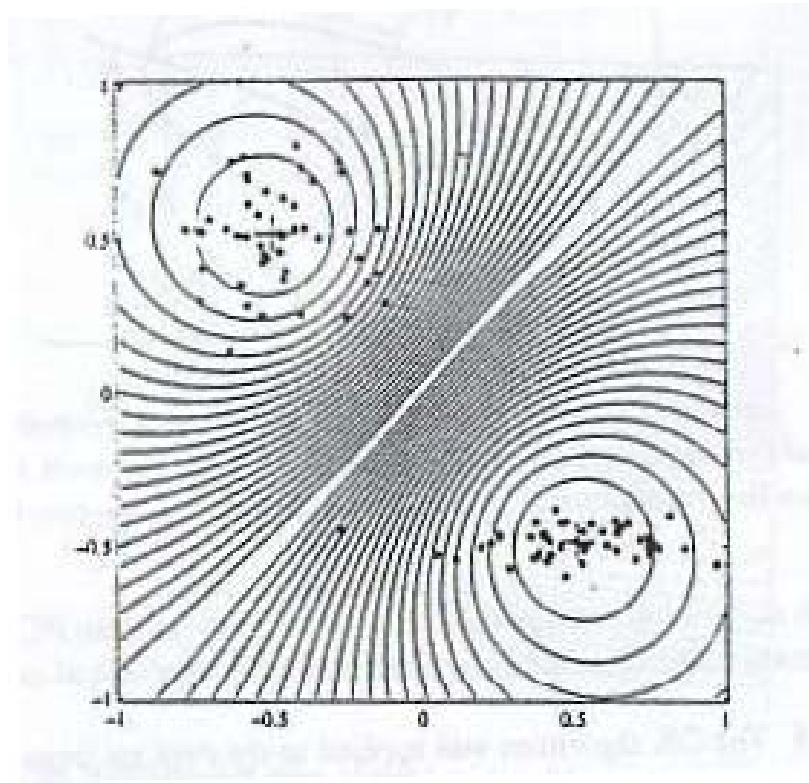
$$\mu_{ik} = \frac{1}{\sum_{j=1}^c \left(D_{ikA} / D_{jkA} \right)^{\frac{2}{m-1}}} \quad 1 \leq i \leq c, \quad 1 \leq k \leq N$$

$$v_i = \frac{\sum_{k=1}^N (\mu_{ik})^m z_k}{\sum_{k=1}^N (\mu_{ik})^m} \quad 1 \leq i \leq c$$

donde: $D_{ikA} = (z_k - v_i)^T A (z_k - v_i)$.

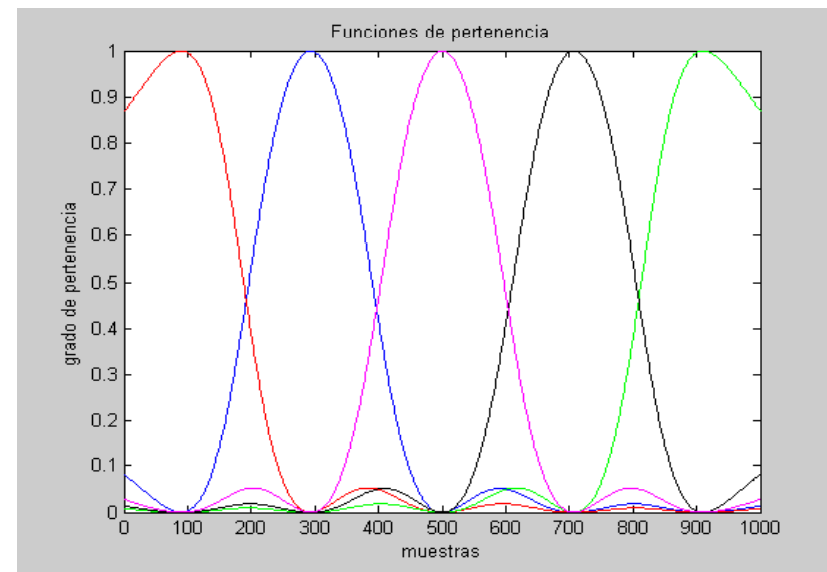
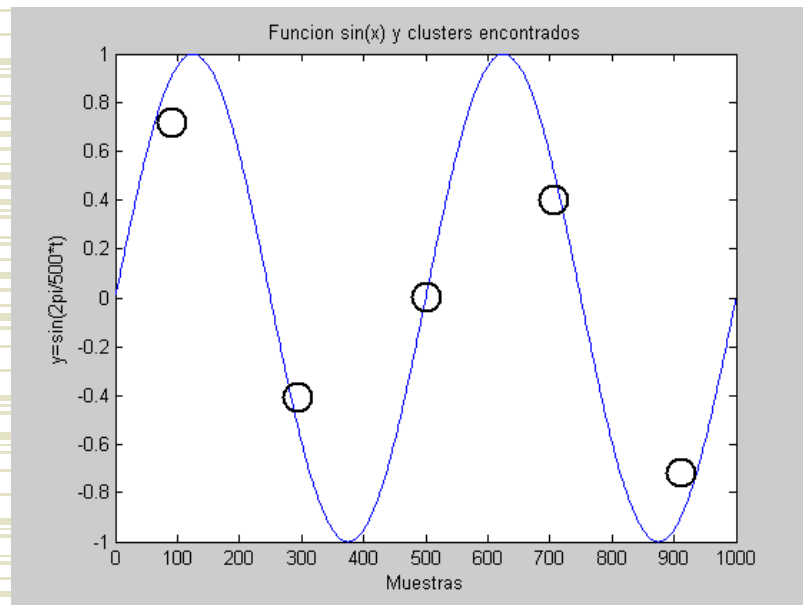
Fuzzy C-Means Clustering

- ◆ Ejemplo:

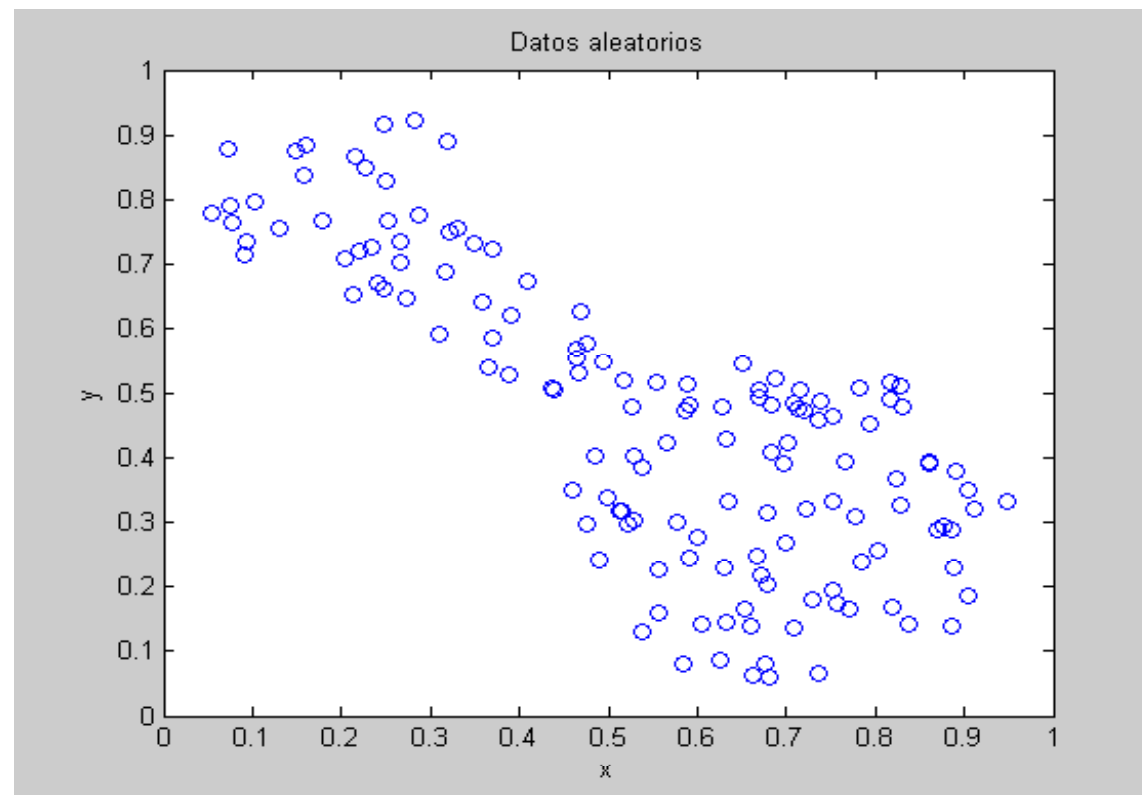


Slide preparada por R. Zuñiga (2005)

Fuzzy C-Means Clustering

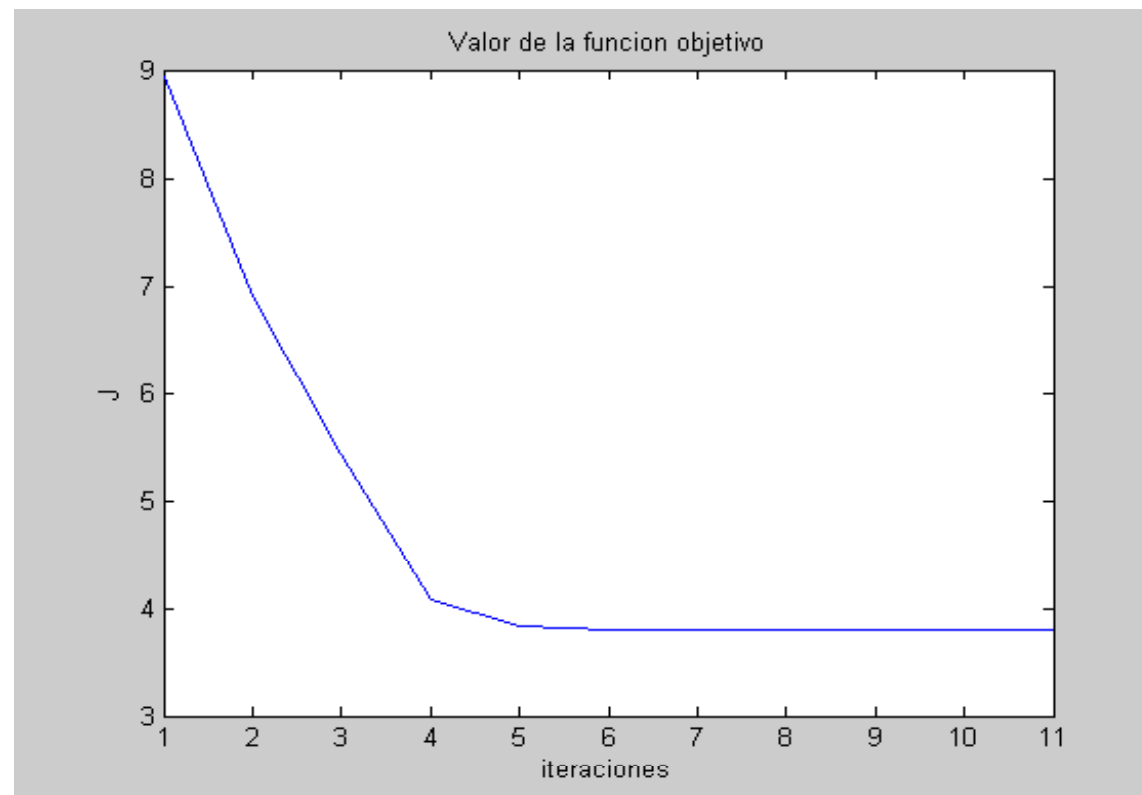


Fuzzy C-Means Clustering



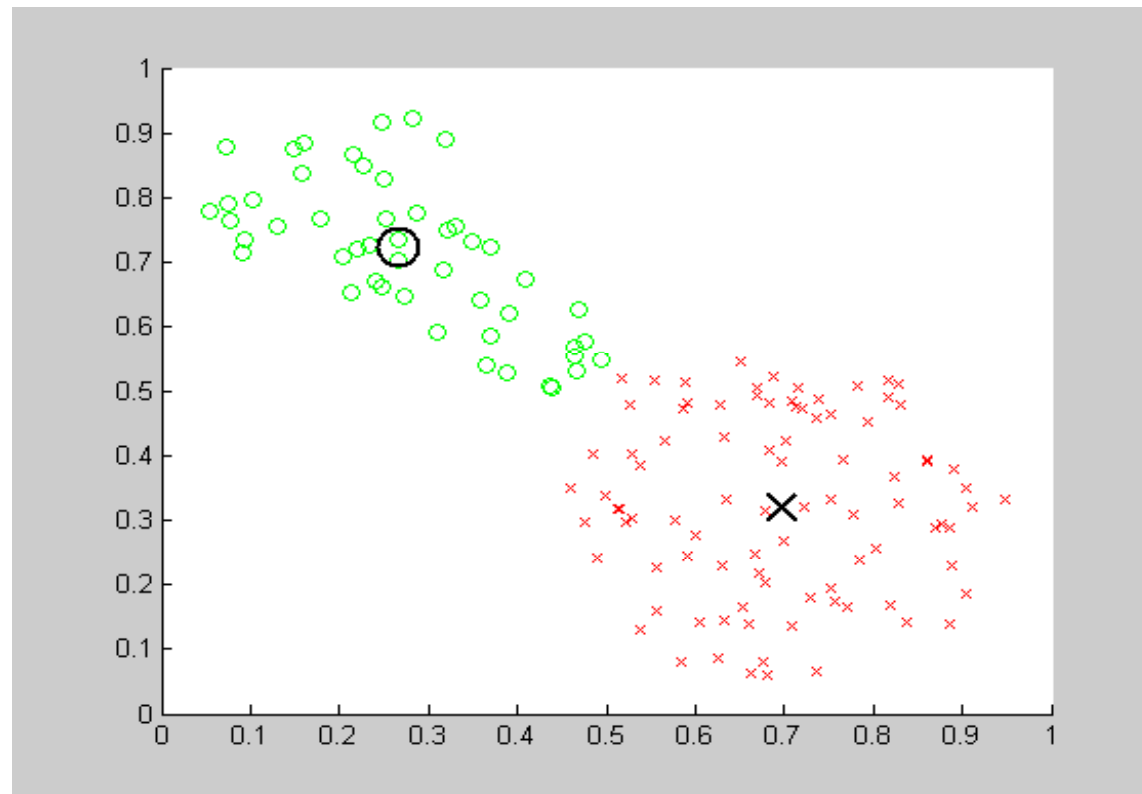
Slide preparada por R. Zuñiga (2005)

Fuzzy C-Means Clustering



Slide preparada por R. Zuñiga (2005)

Fuzzy C-Means Clustering

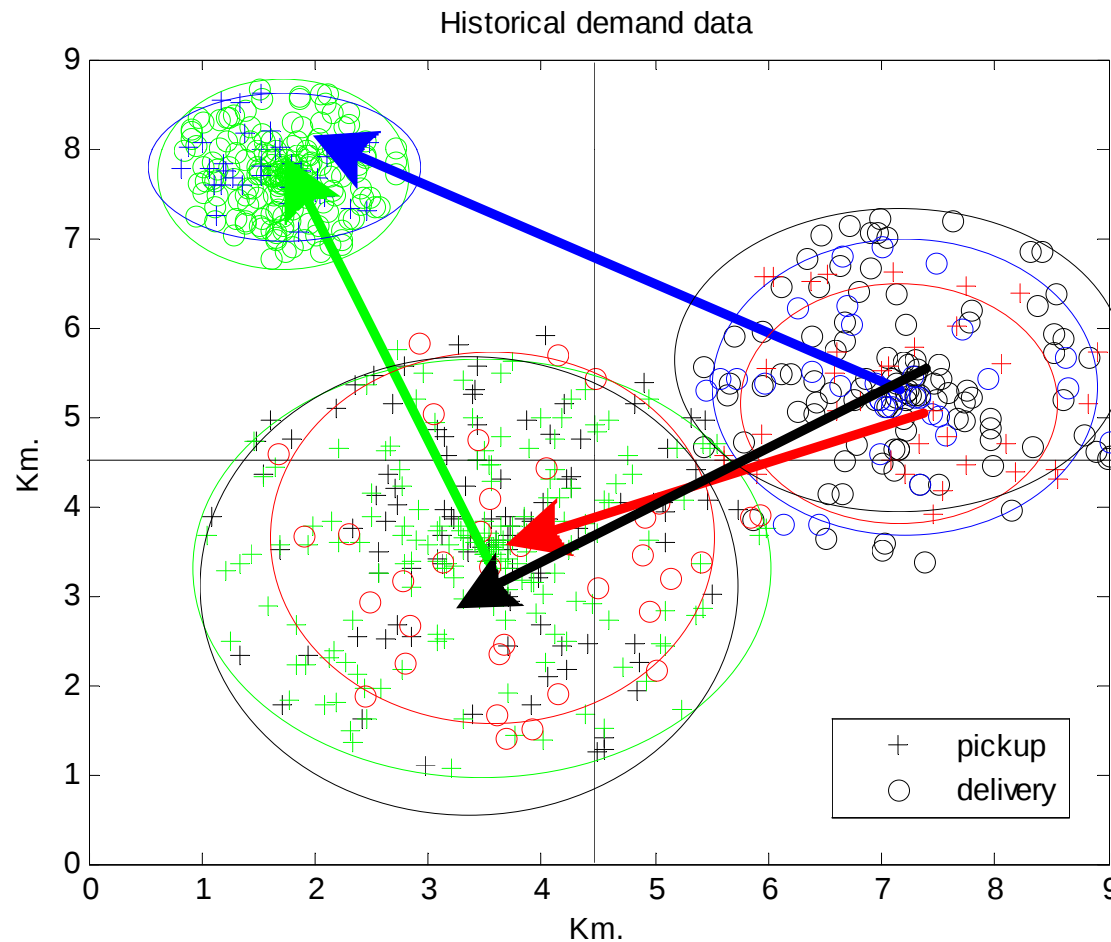


Diapositiva 22

U CHILE1 TOSHIBA, 28/08/2007

Fuzzy clustering: Predicción de demanda

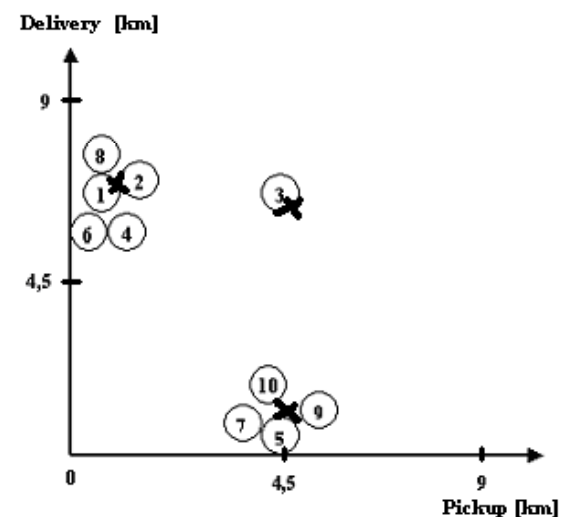
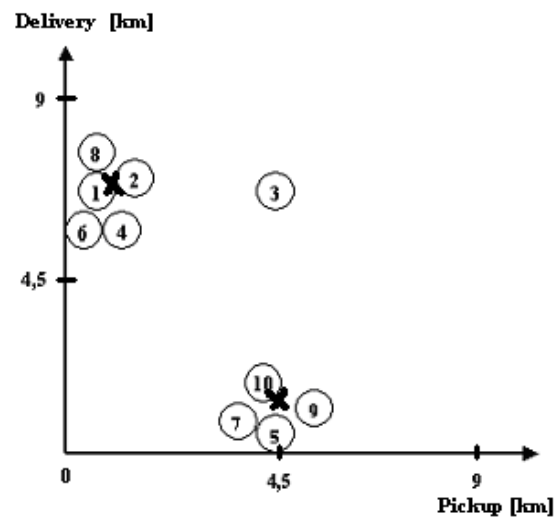
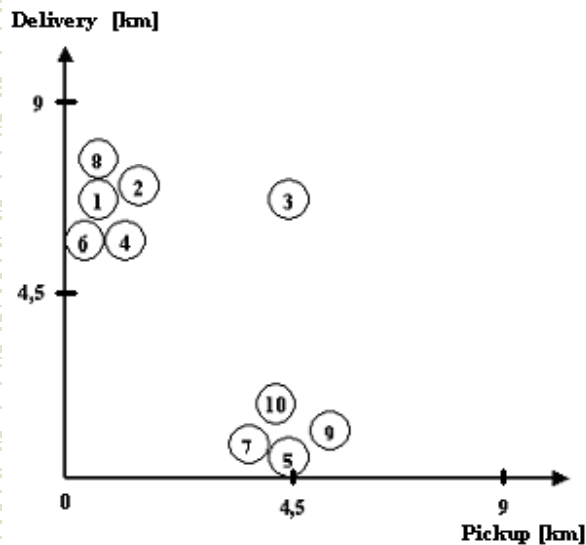
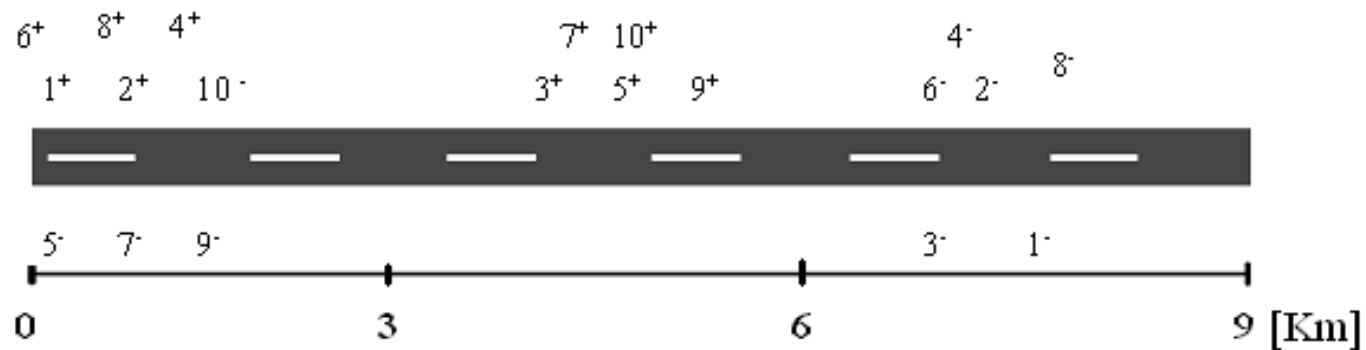
- ◆ Motivación: Diferentes configuraciones de demanda.



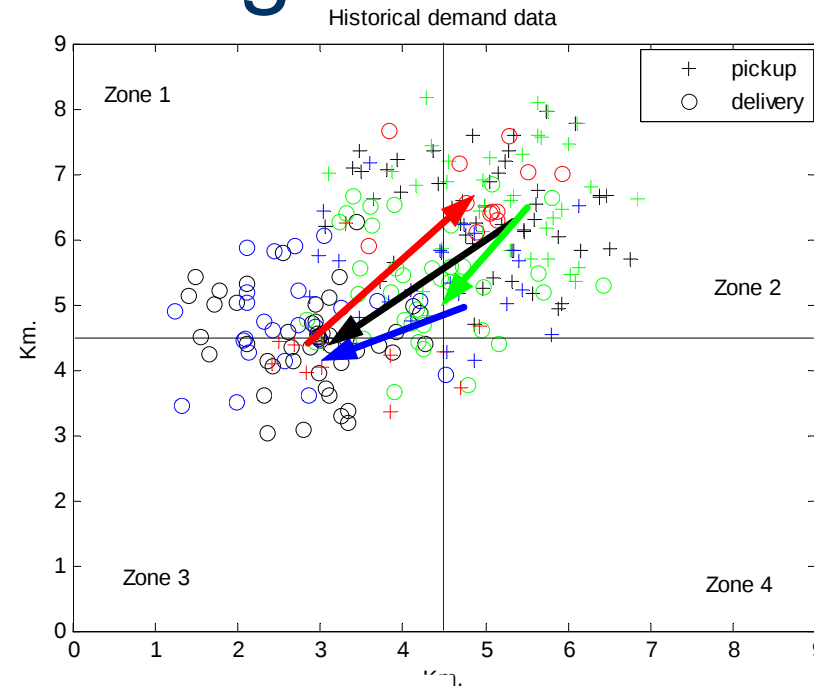
Fuzzy clustering: Predicción de demanda

- ◆ Metodología de zonificación sistemática
- ◆ División del espacio en regiones para una mejor representación de los patrones históricos de demanda, basándose en clustering difuso
- ◆ Permite clasificar las llamadas origen-destino en clusters flexibles y representativos.

Fuzzy clustering: Predicción de demanda



Fuzzy clustering: Predicción de demanda



X pickup	Y pickup	X delivery	Y delivery
4.5540	5.7155	2.9218	4.7514
3.7514	4.4812	5.2293	6.2232
4.7989	6.6121	3.0751	4.4972
5.2595	6.5057	4.3494	5.5161

Pickup and delivery coordinates Fuzzy zoning

X pickup	Y pickup	X delivery	Y delivery
6.75	6.75	6.75	6.75
2.25	6.75	2.25	6.75
6.75	6.75	2.25	2.25
6.75	6.75	2.25	6.75

Pickup and delivery coordinates. Classic zoning