FUZZY PARTITION INDUCED BY POSITIVE SEMI-DEFINITE SIMILARITY MATRIX

Slávka Bodjanová, Institute of Economy, Bratislava, Czechoslovakia

The aim of this contribution is to present a simple attempt to fuzzy partitioning of a finite set of units (objects or variables) where the similarities between units are expressed by positive semi-definite (p.s.d.) similarity matirx.

1. Introduction

Let $X = \{x_1,\dots,x_n\}$ be a finite set of units. Let us denote by $S = (s_{ij})$ similarity matirx of X. Many of similarity matrices are p.s.d.matrices (see e.g. [3]). In this paper we shall consider only p.s.d. similarity matrices of normalized similarity coefficients, hence $s_{jj} = s(x_j,x_j) = 1$.

If $S = \text{diag}\{1,\dots,1\}$ then there is no structure of X and the units of X are fully dissimilar. If $S \neq \text{diag}\{1,\dots,1\}$ we may suppose that the similarities between units are explained by a small number of unknown underlying prototyps to which the units have different belongings (from zero - there is no similarity between unit and prototyp till one - both the unit and the prototyp are identical). The prototyps are fuzzy subsets of X (e.g. good, average, bad students...).

Let us denote by u_i the ith fuzzy prototyp of X and by $u_i(x_j) = u_{ij}$ the membership degree of $x_j \in X$ to the fuzzy set u_i, $i=1,\dots,k$, $j=1,\dots,n$, $k \leq n$. It is naturally to require that different prototyps are dissimilar as much as is possible. In terms of fuzzy sets

$$(\bigcup_{i=1}^{t} u_i) \cap u_{i+1} = \emptyset \quad \text{for } t=2,\dots,k-1 \qquad (1.1)$$

hence the fuzzy sets $u_1, u_2, \dots, u_k$ are disjoint.

The intersection and the union of the fuzzy sets u_i and u_m are defined in this paper as

$$(u_i \cap u_m)(x_j) = \max\{u_{ij} + u_{mj} - 1, 0\} \quad \text{for all } x_j \in X \qquad (1.2)$$

$$(u_i \cup u_m)(x_j) = \min\{u_{ij} + u_{mj}, 1\} \quad \text{for all } x_j \in X \qquad (1.3)$$

The observed data set X is crips set, hence $X(x_j) = 1$ for all $x_j \in X$.

The degree of representation of unit $x_j \in X$ by fuzzy prototyps $u_1,\dots,u_k$ can be expressed by

$$\sum_{i=1}^{k} u_i(x_j) = \sum_{i=1}^{k} u_{ij} \qquad (1.4)$$

and the degree of "unicity" of unit $x_j \in X$ can be expressed by

$$X(x_j) - \sum_{i=1}^{k} u_i(x_j) = 1 - \sum_{i=1}^{k} u_{ij} \qquad (1.5)$$

We introduce so called residual fuzzy subset of X (denoted by u_{res}) which membership function is defined by (1.5). The following properties hold:

$$\left(\bigcup_{i=1}^{k} u_i\right) \cap u_{res} = \emptyset \qquad (1.6)$$

$$\left(\bigcup_{i=1}^{k} u_i\right) \cup u_{res} = X \qquad (1.7)$$

It can be easily shown that the family of fuzzy sets $u_1,\dots,u_k,u_{res}$ satisfying the conditions (1.1), (1.6), (1.7) is fuzzy (k+1)-partition of X defined by Bezdek [1] as follows:

Definition 1.1

Let us denote the usual vector space of real k×n matrices by V_{kn}. Then the fuzzy k-partition space associated with a set of n units is defined by

$$P_{fk} = \left\{ U \in V_{kn};\ u_{ij} \in \langle 0,1\rangle \text{ for all } i,j;\ \sum_i u_{ij} = 1 \text{ for all } j;\ 0 \le \sum_j u_{ij} \text{ for all } i \right\} \qquad (1.8)$$

If two units $x_p, x_q \in X$ are very similar (they have hight coefficient of similarity in S) they would have similar belonginges to fuzzy prototyps of X. We can define similarity measure between units of X induced by fuzzy partition U of X as follows:

Definition 1.2

Let $U \in P_{fk}$ be a fuzzy partition of X. Then a function $s^*: X \times X \to R^+$ defined by

$$s^*(x_p,x_q) = A\left(\sum_{i=1}^{k} f(u_{ip},u_{iq})\right)^r + B \quad \text{for all } x_p,x_q \in X \qquad (1.9)$$

where A,B,r are real constants and f is real function defined such that s^* is a similarity measure, is called similarity measure induced by fuzzy partition U.

Examples of normalized similarity measures s^*:

a/ $$s^*(x_p,x_q) = 1 - \frac{1}{2}\sum_i |u_{ip} - u_{iq}| \qquad (1.10)$$

b/ $$s^*(x_p,x_q) = 1 - \frac{1}{2}\left(\sum_i (u_{ip} - u_{iq})^2\right)^{1/2} \qquad (1.11)$$

c/ $$s^*(x_p,x_q) = \sum_i (u_{ip} \cdot u_{iq})^{1/2} \qquad (1.12)$$

Let us denote by S^* the matrix of normalized similarity coefficients induced by fuzzy partition U. We are interested to obtain the fuzzy partition U of fuzzy prototyps of X which minimizes criterion function

$$\phi(S,S^*) = \sum_{p=1}^{n-1} \sum_{q=p+1}^{n} \left| s(x_p,x_q) - s^*(x_p,x_q) \right| \qquad (1.13)$$

where S is a similarity matrix of X.

2. Fuzzy partition of k principal clusters

Let S be a p.s.d. matrix of normalized similarity coefficients of $X = \{x_1,\dots,x_n\}$. Due to spectral decomposition theorem

$$S = \sum_{i=1}^{n} \lambda_i \cdot \gamma_i \cdot \gamma_i' \qquad (2.1)$$

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$ are eigenvalues of S with orthonormal eigenvectors $\gamma_i = (\gamma_{i1}, \ldots, \gamma_{in})'$, i=1,...,n. Suppose that the first k eigenvalues are positive. Then we can define matrix $U \in V_{k+1,n}$ as follows:

for j=1,...,n: $\quad u_{ij} = \lambda_i \cdot \gamma_{ij}^2 \quad$ for i=1,...,k $\qquad (2.2)$

$$u_{k+1,j} = 1 - \sum_{i=1}^{k} u_{ij} \qquad (2.3)$$

Theorem 2.1

The matrix $U \in V_{k+1,n}$ defined by(2.2) and (2.3)is fuzzy (k+1) -partition of $X = \{x_1, \ldots, x_n\}$.

The proof is evident.

Definition 2.1

The fuzzy partition defined by (2.2) and (2.3) is called the fuzzy partition of k principal clusters. The fuzzy sets (fuzzy clusters) u_i, i=1,...,k are called principal clusters or prototyps and the fuzzy cluster u_{k+1} is called residual cluster.

Every principal cluster has its own structure determined by the sign of γ_{ij}, i=1,...,k,j=1,...,n. Hence

$$u_i = u_i^+ \cup u_i^- \qquad (2.4)$$

where $\quad u_{ij}^+ = u_{ij}$ if $\gamma_{ij} > 0 \quad$ else $\quad u_{ij}^+ = 0 \qquad (2.5)$

$\quad u_{ij}^- = u_{ij}$ if $\gamma_{ij} < 0 \quad$ else $\quad u_{ij}^- = 0 \qquad (2.6)$

Let us define the similarity coefficient induced by fuzzy partition of k principal clusters as follows: for $x_p, x_q \in X$:

$$s^*(x_p, x_q) = \sum_{i \in I_{pq}} \sqrt{u_{ip}} \cdot \sqrt{u_{iq}} - \sum_{i \notin I_{pq}} \sqrt{u_{ip}} \cdot \sqrt{u_{iq}} \qquad (2.7)$$

where $I_{pq} = \{ i \in \{1, \ldots, k\} : \gamma_{ip} \cdot \gamma_{iq} > 0 \}$.

Theorem 2.2

Let S be a p.s.d. matrix of normalized similarity coefficients of X, let U be the fuzzy partition of k principal clusters of X and let S^* be the matrix of induced similarity coefficients defined by (2.7). Then the criterion function (1.13) equals zero.

The proof is evident.

In practice the problem is to decide how many principal clusters it is worth fitting to the data. We can show that the process of successive identification of q principal clusters (q=1,...,k) produces a hierarchical sequence of fuzzy partitions. The hierarchical clustering is defined by Dumitrescu in [2].

Let a sequence $U_{(0)}, U_{(1)}, \ldots, U_{(t)}$ of fuzzy partitions of X be obtained by means of the following algorithm:

Algorithm 2.1

Step 1. Let S be a p.s.d. similarity matrix of range k, let $\varepsilon \in (0,1)$ be a given small number and put t=0.
Let $U_{(0)} = \{u_{res(0)}\}$, where for all $x_j \in X$: $u_{res(0)}(x_j) = 1$

Step 2. Increment t by 1 and put
$U_{(t)} = \{u_1, \dots, u_t, u_{res(t)}\}$
where for all $x_j \in X$ $u_i(x_j) = \lambda_i \, \gamma_{ij}^2$, i=1,...,t
(λ_i and γ_{ij} are defined by (2.1))
$u_{res(t)}(x_j) = u_{res(t-1)}(x_j) - u_t(x_j)$

Step 3. Let us compare S by S^* (defined by (2.7)) .
If $\max_{x_p, x_q \in X} | s(x_p,x_q) - s^*(x_p,x_q) | < \varepsilon$ or t=k
then stop. Else go to Step2.

Theorem 2.3

The sequence $U_{(0)}, \dots, U_{(t)}$ of fuzzy partitions obtained by means of Algorithm 2.1 is hierarchical.
The proof is evident.

Theorem 2.4

The cardinality of the ith principal cluster equals ith eigenvalue of similarity matrix S.
Proof:

$$\text{card } u_i = \sum_j u_{ij} = \sum_j \lambda_i \cdot \gamma_{ij}^2 = \lambda_i \cdot \sum_j \gamma_{ij}^2 = \lambda_i \cdot \gamma_i' \cdot \gamma_i = \lambda_i$$

Corollary 2.1

Let U be a fuzzy partition of k principal clusters of X. Then

i/ $\text{card } u_1 \geq \dots \geq \text{card } u_k$ (2.8)

ii/ $\sum_{i=1}^{k} \text{card } u_i + \text{card } u_{res} = \text{card } X$ (2.9)

Definition 2.2

Let U be a fuzzy partition of k principal clusters of $X = \{x_1, \dots, x_n\}$.
Let $\varepsilon \in (0,1)$. If $t < k$ is such integer that

$$\frac{\sum_{i=1}^{t} \text{card } u_i}{\text{card } X} \geq 1 - \varepsilon \quad \text{and} \quad \frac{\sum_{i=1}^{t+1} \text{card } u_i}{\text{card } X} < \varepsilon \qquad (2.10)$$

we say that the first t principal clusters are $(1-\varepsilon)$- sufficient to describe classification of n units.

The result of fuzzy hierarchical clustering can be the fuzzy partition which is $(1-\varepsilon)$- sufficient to describe classification of units of X if we modify the stop criterion in Algorithm 2.1 by (2.10) .

3. Discussion

a/ It is important to note that the structure of k principal clusters is indeterminate with respect to rotation. The choice of rotation can be realized due to additional information about data or due to minimization of fuzziness of fuzzy partition in order to facilitate the interpretation of the solution.

b/ Let $X = \{x_1, \ldots, x_n\}$ be a set of n variables each measured on a set of p objects. Let k-principal factor model hold for X. It can be easily shown that the k-principal factor model can be considered as a fuzzy partition of k principal clusters of X.

c/ There can be find some connection between fuzzy partition of k principal clusters and the method of principal coordinate anaysis in multidimensional scaling.

References:

[1] Bezdek,J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York 1981

[2] Dumitrescu,D.: On Fuzzy Partitions in Cluster Analysis, Cluj-Napoca 1983

[3] Mardia,K.V. - Kent,J.T. - Bibby,J.M.: Multivariate Analysis, Academic Press, New York 1979