Fuzzy Sets and Systems 1 (1978) 111–127.

FUZZY PARTITIONS AND RELATIONS; AN AXIOMATIC BASIS FOR CLUSTERING*

James C. BEZDEK

Mathematics Department, Utah State University, Logan, UT 84322, U.S.A.

J. Douglas HARRIS

Mathematics Department, Marquette University, Milwaukee, WI 53233, U.S.A.

Received 25 May 1977
Revised September 1977

In this paper some connections between fuzzy partitions and similarity relations are explored. A new definition of transitivity for fuzzy relations yields a relation-theoretic characterization of the class of all psuedo-metrics on a fixed (finite) data set into the closed unit interval. This notion of transitivity also links the triangle inequality to convex decompositions of fuzzy similarity relations in a manner which may generate new techniques for fuzzy clustering. Finally, we show that every fuzzy c-partition of a finite data set induces a psuedo-metric of the type described above on the data.

Key Words: Convex decompositions, Cluster analysis, Fuzzy relations, Pseudo-metrics, Transitivity.

1. Introduction

The extant theory and applications of fuzzy relations are contained in the papers of Zadeh [1]; Tamura *et al.* [2]; Kandel and Yelowitz [3]; and Dunn [4]. In particular, the transitive closure of [0, 1], reflexive, symmetric fuzzy relations is discussed as a basis for constructing hierarchical clusters in finite data sets. Dunn showed in [4] that this methodology yielded essentially the same results as the well-known graph-theoretic technique called the single linkage method (c.f. Duda and Hart [5]). The basis for this observation was that the notion of transitivity used in [1–3] for fuzzy relations is equivalent to the ultra-metric inequality. One of our main goals in the present work is to enlarge this theory by redefining fuzzy transitivity so that it becomes equivalent to the triangle inequality. The class of fuzzy similarity relations characterized in this way appears to be an important space for applications in clustering.

In Section 2 we review *hard* (i.e., non-fuzzy) partitions and equivalence relations for finite sets. Section 3 extends these ideas to fuzzy partitions and similarity relations, and introduces Max-$\triangle$ transitivity. In Section 4 we discuss convex decompositions of fuzzy similarity relations. Section 5 presents an application of the preceding ideas, defining a new method for clustering via convex decomposition of similarity relation matrices. In Section 6 the work of Bezdek and Harris [6] is continued by connecting fuzzy

*This research supported by NSF Grant Number DCR75–05014.

partitions with fuzzy similarity relations. We show that *every* fuzzy c-partition of a finite data set induces a pseudo-metric on the data. In this way individual fuzzy relationships can be obtained from fuzzy class memberships.

2. Hard partitions and relations

Let $X=\{x_1, x_2,\dots x_n\}$ be a finite data set. If, for a positive integer c, $2 \leqq c < n$, it is known (or assumed) that X contains representatives from c subclasses, then cluster analysis with respect to X is the problem of identifying the subclass labels, i.e., of partitioning X into c subsets (clusters).

A hard c-partition of X has three equivalent characterizations: sets, functions; and matrices. In what follows the description most convenient for us is in terms of matrices. Towards this end let V_{cn} be the usual vector space of real $c \times n$ matrices over V. Let u_{ik} be the ik^{th} element of $U \in V_{cn}$, and define

$$\mathscr{P}_c=\left\{U \in V_{cn} \middle| u_{ik} \in \{0,1\} \forall i,k;\ \sum_{i=1}^{c} u_{ik}=1\ \forall k;\ \sum_{k=1}^{n} u_{ik}>0\ \forall i\right\}. \tag{1}$$

Here u_{ik} is the membership of x_k in class i; $\mathscr{P}_c$ is exactly *non-degenerate hard c-partition space* for X, and the superset $\mathscr{P}_{co} \supset \mathscr{P}_c$ of matrices obtained by relaxing the last condition in (1) to $\sum_{k=1}^{n} u_{ik} \geqq 0\ \forall i$ is the corresponding degenerate space.

To each $U \in \mathscr{P}_c$ there corresponds a unique hard equivalence relation in the Cartesian product $X \times X$. Loosely speaking, we have, given $U \in \mathscr{P}_c$, the relation matrix $R=[r_{ij}]$ in V_{nn} defined by:

$$r_{ij}=\begin{cases}1; & u_{ki}=u_{kj}=1\ \exists k,\\ 0; & \text{otherwise.}\end{cases}$$

Since R is an equivalence relation, it satisfies three requirements:

$$r_{ii}=1 \quad \forall i, \qquad \text{(reflexivity)} \tag{2a}$$

$$r_{ij}=r_{ji} \quad \forall i \neq j, \qquad \text{(symmetry)} \tag{2b}$$

$$\left.\begin{array}{l} r_{ik}=1 \\ r_{kj}=1 \end{array}\right\} \Rightarrow r_{ij}=1 \quad \forall i,j. \qquad \text{(transitivity)} \tag{2c}$$

Let S and R be two equivalence relations. The composition of R followed by S will be denoted as $S \circ R$. In particular, $\circ$ denotes generalized matrix multiplication when S and R are relation matrices. Thus if $P=S \circ R \in V_{nn}$, then

$$p_{ij}=\mathop{\square}_{l=1}^{n} (s_{il} * r_{lj}), \tag{3}$$

where $(\Box, *)$ is the pair of operations defining $\circ$. In the sequel our interest lies with four matrix products:

$$\circ = (\textstyle\sum . \bullet) \sim \text{Sum-Product}, \tag{4a}$$

$$\circ = (\textstyle\sum . \wedge) \sim \text{Sum-Min}, \tag{4b}$$

$$\circ = (\vee . \bullet) \sim \text{Max-Product}, \tag{4c}$$

$$\circ = (\vee . \wedge) \sim \text{Max-Min}. \tag{4d}$$

Let I_n denote the $n \times n$ identity matrix, and define the ordering $A \leqq B \Leftrightarrow a_{ij} \leqq b_{ij} \forall i, j$; $A, B \in V_{nn}$. With these conventions and (3), (4), conditions (2) may be restated compactly as

$$I_n \leqq R, \qquad \text{(reflexivity)} \tag{2a)$'$($$

$$R = R^{\mathrm{T}}, \qquad \text{(symmetry)} \tag{2b)$'$($$

$$R = R(\vee . \bullet)R = R(\vee . \wedge)R = R^2. \qquad \text{(transitivity)} \tag{2c)$'$($$

The set of all *hard equivalence relations* on n data points we denote by

$$\mathscr{R}_n = \{R \in V_{nn} | r_{ij} \in \{0, 1\} \forall i, j; I_n \leqq R; R = R^{\mathrm{T}} = R(V . \wedge)R\}. \tag{5}$$

Another characterization (not used below) is in terms of graphs. We mention that the work of Rosenfeld [7] and Yeh and Bang [8] on fuzzy graphs might be substantially enriched by adopting for its basis the fuzzy relation structure described below.

3. Fuzzy relations and partitions

Physical and mathematical objections to $\mathscr{P}_c$(or $\mathscr{R}_n$) as a basis for pattern recognition models are discussed at length in [9]. For the present work our interest centers on the fact that each $U \in \mathscr{P}_c$ requires every $x_k \in X$ to belong unequivocally to precisely one partitioning subset of U; and the relation $R \in \mathscr{R}_n$ induced on $X \times X$ by U renders all x_k's in each equivalence class *indistinguishable* (from one another) to data in other classes, that is, totally related to one another and completely unrelated to members of other classes. Transitivity (2c) is particularly difficult to justify in many applications (c.f. Crowson [10]) and turns out to be the most subtle of properties (2) to generalize.

Zadeh in [11] originated the idea of allowing sets to have "fuzzy" boundaries. Following [11], we call a *membership function* $u_i : X \rightarrow [0, 1]$ a *fuzzy subset* of $X : u_i(x_k) \doteq u_{ik}$ is the *grade of membership* of x_k in u_i. For example, $u_A(8.98) = 0.96$ would imply for $r = 8.98$ a strong agreement with the properties characterizing the fuzzy subset $A = \{r \in R \mid r$ is "slightly less" than $9\}$. Fuzzy imbeddings for $\mathscr{P}_c$ and $\mathscr{R}_n$ can be constructed in a variety of ways. The most widely used imbedding for $\mathscr{P}_c$ is *non-degenerate fuzzy c-partition space:*

$$\mathscr{P}_{fc} = \left\{ U \in V_{cn} | u_{ik} \in [0, 1] \forall i, k; \sum_{i=1}^{c} u_{ik} = 1 \forall k; \sum_{k=1}^{n} u_{ik} > 0 \forall i \right\}. \tag{6}$$

C

At least three algorithms for generating fuzzy c-partitions of X are now known: those of Ruspini [12]; Woodbury [13]; and the fuzzy ISODATA algorithm discussed in [9]. More recently, the nature of the embeddings $\mathscr{P}_c \subseteq \mathscr{P}_{co} \subseteq \mathscr{P}_{fc}$ have been explored in [6] to exploit the fact that $\mathscr{P}_{fc}$ is the convex hull of $\mathscr{P}_{co}: \mathscr{P}_{fc} = \operatorname{conv}(\mathscr{P}_{co})$.

A matrix $U \in \mathscr{P}_{fc}$ of membership functions for points in X conveys the relationship each x_k bears to the c fuzzy subclasses partitioning X, but says nothing about relationships between individuals. To realize information of this kind we consider fuzzy relations in $X \times X$. A fuzzy relation in $X \times X$ is a membership function $\rho: X \times X \to [0, 1]$ whose values $\rho(x_i, x_j)$ denote the *strength of relationship* $\forall i$ and j between x_i and x_j. In a manner entirely analogous to the hard case above, ρ can be represented by an $n \times n$ relation matrix $R = [r_{ij}] = [\rho(x_i, x_j)]$. To generalize (2a) and (2b) we call a fuzzy relation R

$$\text{reflexive} \Leftrightarrow r_{ii} = 1\ \forall i; \quad \text{and} \quad \text{symmetric} \Leftrightarrow r_{ij} = r_{ji}\ \forall i \neq j. \tag{7a)-(7b}$$

Again compactly expressed as $I_n \leqq R$ and $R = R^{\mathrm{T}}$ respectively.

The extensions given in (7a) and (7b) are quite natural. Extending transitivity to fuzzy relations requires more thought. In [1] Zadeh proposed the following definition: a fuzzy relation R is

$$\max - * \mathit{transitive} \Leftrightarrow r_{ij} \geqq \bigvee_{l=1}^{n} (r_{il} * r_{lj})\ \forall i, j \tag{7c}$$

or equivalently, if and only if $R \geqq R(\vee\, . *)R \doteq R^2$, where $*$ could be either min ($\vee$) or ordinary product ($\bullet$). Following Zadeh we call

$$\mathscr{R}* = \{R \in V_{nn} | r_{ij} \in [0, 1]\ \forall i, j;\ I_n \leqq R : R = R^{\mathrm{T}} \geqq R(\vee\, . *)R\} \tag{8}$$

sets of *fuzzy similarity relations* in $X \times X$. Since $ab \leqq a \wedge b$ for $a, b \in [0, 1]$, max–min transitivity implies max–prod transitivity, i.e., $\mathscr{R}_{\wedge} \subseteq \mathscr{R}_{\bullet}$.

There is a more general operation $*$ with natural physical meaning which extends $\mathscr{R}_n$ to a maximal set of fuzzy similarity relations. Let us define a fuzzy similarity relation R to be

$$\text{max-}\triangle\ \mathit{transitive} \Leftrightarrow r_{ij} \geqq \bigvee_{l=1}^{n} ((r_{il} + r_{lj} - 1) \vee 0)\ \forall i, j. \tag{9}$$

To see that max-$\triangle$ transitivity is implied by max–prod transitivity, note that $(a + b - 1) \vee 0 \leqq ab$ for $a, b \in [0, 1]$, so

$$r_{ij} \geqq \bigvee_{l=1}^{n} (r_{il} \bullet r_{lj}) \geqq \bigvee_{l=1}^{n} (r_{il} + r_{lj} - 1) \vee 0.$$

In fact, if we define on $[0, 1) \times [0, 1]$ the operations

$$a \triangle b = (a + b - 1) \vee 0, \tag{10a}$$

$$a \square b = \frac{a+b}{2}, \tag{10b}$$

$$a \oplus b = a + b - ab, \tag{10c}$$

then the inequalities

$$(a+b-1) \vee 0 \leqq ab \leqq a \wedge b \leqq \frac{a+b}{2} \leqq a \vee b \leqq a+b-ab \leqq 1 \tag{11}$$

result in the following hierarchy of similarity relation spaces upon substitution of the appropriate operation for $*$ in (8).

$$\mathscr{R}_n \subseteq \mathscr{R}_{\oplus} \subseteq \mathscr{R}_{\vee} \subseteq \mathscr{R}_{\square} \subseteq \mathscr{R}_{\wedge} \subseteq \mathscr{R}_{\bullet} \subseteq \mathscr{R}_{\triangle}. \tag{12}$$

The type of transitivity employed will presumably be dictated by the application at hand. We contend that max-$\triangle$ transitivity is the most interesting type on both physical and mathematical grounds. To see how restrictive max–min transitivity is, for example, note that any m numbers $a_1 \leqq a_2 \leqq \ldots \leqq a_m$ can satisfy the requirement $a_i \geqq a_j \wedge a_k$, $\forall i, j, k$ distinct, if and only if $a_1 = a_2 = \ldots = a_{m-1} \leqq a_m$, because $a_1 \geqq (a_{m-1} \wedge a_m)$. Indeed, for the matrix

$$R(\lambda) = \begin{bmatrix} 1 & 0.8 & 0.7 \\ 0.8 & 1 & \lambda \\ 0.7 & \lambda & 1 \end{bmatrix} \text{ with } \lambda \in [0, 1], \tag{13}$$

it is easy to check that $R(\lambda) \in \mathscr{R}_{\wedge} \Leftrightarrow \lambda = 0.7$; whereas $R(\lambda) \in \mathscr{R}_{\triangle}\ \forall \lambda \in [0.5, 0.9]$. This illustrates how sparse the relations $\mathscr{R}_{\wedge}$ are among the relations $\mathscr{R}_{\triangle}$.

A physical interpretation of max-$\triangle$ transitivity can be made using (13) and a graphical representation of the relationships involved. Consider first $R(0.7)$ with $r_{12} = r_{21} = 0.8$; $r_{13} = r_{31} = 0.7$; and $r_{23} = r_{32} = 0.7$. Since $R(\lambda)$ is max–min transitive iff $\lambda = 0.7$, only one possibility is allowed for the mutual bonding provided by x_1 in linking x_2 to x_3: namely, that *all* of the relatives of x_3 responsible for the relationships $r_{31} = 0.7$ are shared through with x_2 (Fig. 1).

In other words, max–min transitivity calls for the optimal "alignment" of mutual relatives. An alternative way to think of this follows by answering the question:

(If $x\,R\,y = 0.7$, what should $y\,R\,z = (?)$ to imply that $x\,R\,z \geqq 0.7$?)

It is our contention that (?) should be 1: x should be related to z by at least $x\,R\,y \Leftrightarrow y$ and z are indistinguishable to outside observers $\Leftrightarrow y\,R\,z = 1$. For $y\,R\,z < 1$, y and z are *not* fully equivalent to each other, and the mathematical restrictiveness manifested by max–min transitivity amounts to assuming the optimal alignment displayed in Fig. 1.

On the other hand, max-$\triangle$ transitivity for the matrix $R(\lambda)$ allows for the "worst" alignment, as illustrated in Fig. 2.

The max-$\triangle$ alignment in Fig. 2 assumes that x_2 and x_3 must be coupled by only 50 of

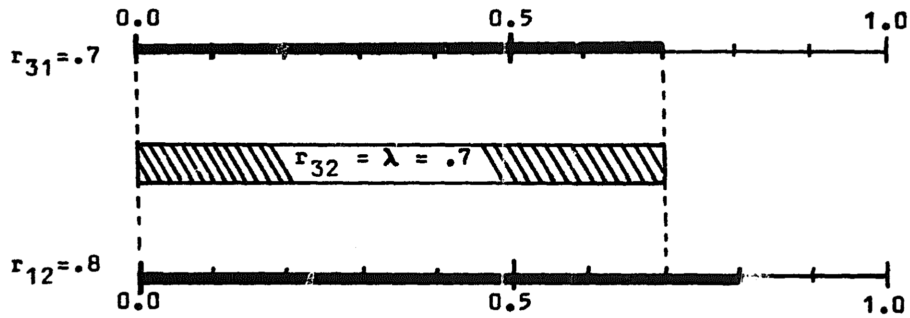

Fig. 1. Graphical illustration of max-$\wedge$ transitivity

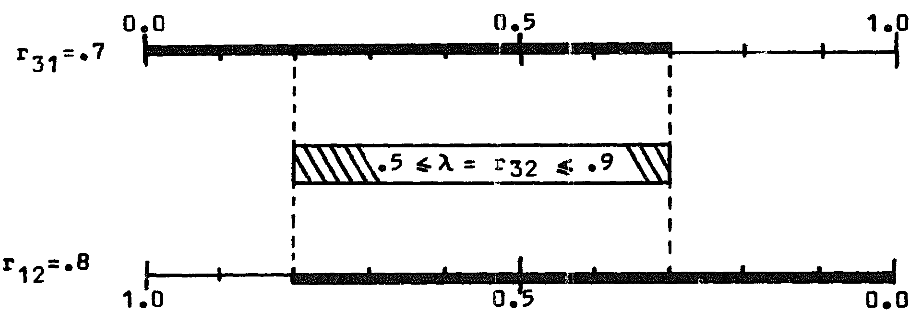

Fig. 2. Graphical illustration of max-$\triangle$ transitivity.

each 100 relatives shared with x_1, the most *unoptimistic* alignment of mutual bonds is used, instead of all 70 as in Fig. 1.

There is an interesting mathematical property of $R_\triangle$ which also argues for its use. Zadeh demonstrated in [1] that max–min transitivity was equivalent to the ultra-metric inequality for the function $d(x_i, x_j)=1-r_{ij}$, where $R=[r_{ij}]\in\mathscr{R}_\wedge$. Since $\mathscr{R}_\wedge\subseteq\mathscr{R}_\triangle$ and the ultra-metric inequality implies the triangle inequality, one might suspect that max-$\triangle$ transitivity is equivalent to the triangle inequality. This is precisely the case:

Theorem 3.1.

$$R\in\mathscr{R}_\triangle \Leftrightarrow \left\{\begin{matrix}\textit{The function } d: X\times X\to[0,1] \textit{ defined}\\ \textit{by } d(x_i,x_j)=1-r_{ij} \textit{ is a pseudo-metric.}\end{matrix}\right\} \tag{14}$$

Proof. Let $R=[r_{ij}]=[1-d_{ij}]$, where we put $d_{ij}\doteq d(x_i,x_j)\,\forall i,j$

(i) $r_{ij}\in[0,1]\,\forall i,j\Leftrightarrow d_{ij}\in[0,1]\,\forall i,j$.

(ii) Suppose $i=j: r_{ii}=1\Leftrightarrow d_{ii}=0$; thus $I_n\leqq R\Rightarrow d_{ii}=0\,\forall i$, and conversely, $d_{ii}=0\,\forall i\Rightarrow I_n\leqq R$.

Note, however, that $r_{ij}=1$ with $i\neq j\Rightarrow d_{ij}=0$ with $x_i\neq x_j$, so d is a pseudo-metric at best.

(iii) Symmetry for d follows from $R=R^{\mathrm{T}}$; symmetry for R from $d_{ij}=d_{ji}\,\forall_{i,j}$.

(iv) Finally, we show that max-$\triangle$ transitivity is equivalent to the triangle inequality:

$$\begin{aligned}
&0 \vee (r_{ij}+r_{jk}-1) \leqq r_{ik}\ \forall \text{ distinct } i, j, k && \Leftrightarrow \\
&r_{ij}+r_{jk} \leqq 1 \leqq 1+r_{ik} && \Leftrightarrow \\
&(1-d_{ij})+(1-d_{jk}) \leqq 1+(1-d_{ik}) && \Leftrightarrow \\
&-(d_{ij}+d_{jk}) \leqq -d_{ik} && \Leftrightarrow \\
&d_{ik} \leqq d_{ij}+d_{jk}. \ \square
\end{aligned}$$

Theorem 3.1 shows that $\mathscr{R}_\triangle$ is the set of fuzzy similarity relations which induce pseudo-metrics on $X \times X$ in exactly the same way that $\mathscr{R}_\wedge$ induces pseudo ultra-metrics on $X \times X$, and substantiates our supposition that it is the maximal set in the hierarchy (12).

4. Convex decompositions of similarity relations

Let conv$(\mathscr{R}_n)$ denote the convex hull of hard equivalence relations. Our goal in this section is to place conv$(\mathscr{R}_n)$ in hierarchy (12), and exhibit a relationship between convex decomposition and max-$\triangle$ transitivity. We begin with:

Theorem 4.1. *For $n=3$ we have*

$$\operatorname{conv}(\mathscr{R}_3)=\mathscr{R}_\triangle. \tag{15}$$

Proof. $R \in \operatorname{conv}(\mathscr{R}_3) \Leftrightarrow \exists$ scalars $\{c_1, c_2, c_3, c_4, c_5\} \subset [0,1]$ so that

$$R=\begin{bmatrix} 1 & y_1 & y_2 \\ & 1 & y_3 \\ & & 1 \end{bmatrix} = c_1 \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 0 \\ & & 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 & 0 & 1 \\ & 1 & 0 \\ & & 1 \end{bmatrix} + \\ + c_4 \begin{bmatrix} 1 & 0 & 0 \\ & 1 & 1 \\ & & 1 \end{bmatrix} + c_5 \begin{bmatrix} 1 & 0 & 0 \\ & 1 & 0 \\ & & 1 \end{bmatrix} \tag{16}$$

with $\sum_{k=1}^{5} c_k = 1$. Here and in the sequel we omit the lower triangular portion of symmetric relation matrices. We use y_1, y_2, y_3 for r_{12}, r_{13}, and r_{23} in R to shorten the notation.

From (2a) and (2b) it follows that $I \leqq R$ and $R=R^{\mathrm{T}}$; further, every entry of R is a partial sum of c_k's, so $0 \leqq r_{ij} \leqq 1\ \forall i, j$. It remains to be seen that R is max-$\triangle$ transitive. We are to show that $1+y_i \geqq y_j+y_k\ \forall$ distinct i, j, k. Towards this end we have from (16) the equations

$$y_1=c_1+c_2, \tag{17a}$$

$$y_2=c_1+c_3, \tag{17b}$$

$$y_3=c_1+c_4. \tag{17c}$$

Adding these and using $\sum_{k=1}^{5} c_k = 1$ yields

$$y_1 + y_2 + y_3 = 2c_1 + 1 - c_5. \tag{18}$$

Using (17) and (18), the convex coefficients c_1, c_2, c_3, and c_4 can be written in terms of c_5 and the entries of R as

$$c_1 = 1/2(y_1 + y_2 + y_3 + c_5 - 1), \tag{19a}$$

$$c_2 = 1/2(1 + y_1 - (y_2 + y_3 + c_5)), \tag{19b}$$

$$c_3 = 1/2(1 + y_2 - (y_1 + y_3 + c_5)), \tag{19c}$$

$$c_4 = 1/2(1 + y_3 - (y_1 + y_2 + c_5)). \tag{19d}$$

Now consider, for example, (19b): since $0 \leqq c_2 \leqq 1$ it is necessary that $1 + y_1 - (y_2 + y_3 + c_5) \geqq 0$, or equivalently that $1 + y_1 - (y_2 + y_3) \geqq c_5$. But $0 \leqq c_5 \leqq 1$, so $1 + y_1 - (y_2 + y_3) \geqq c_5 \geqq 0 \Rightarrow 1 + y_1 \geqq (y_2 + y_3)$. Similar arguments with (19c) and (19d) establish th required inequalities, so R is max-$\triangle$ transitive, and hence $\mathrm{conv}(\mathscr{R}_3) \subseteq \mathscr{R}_\triangle$.

Conversely, suppose $R \in \mathscr{R}_\triangle$. We must prove that R can be decomposed as in (16). From $r_{ij} \in [0,1]\ \forall\, i, k$ and $I_n \leqq R$ it follows that $\sum_{k=1}^{5} c_k = 1$ with $c_k \in [0,1]\ \forall\, k$. If we can choose the c_k's to satisfy these constraints and equations (19), then $R = R^{\mathrm{T}}$ guarantees that $R = \sum_{k=1}^{5} c_k R_k$ with $R_k \in \mathrm{conv}(\mathscr{R}_3)$ as exhibited at (16).

For c_2, c_3, and c_4 to be non-negative it is necessary from (19b), (19c), and (19d) that $c_5 \leqq 1 + y_i - (y_j + y_k)$: for c_1 to be non-negative, that $c_5 \geqq 1 - (y_1 + y_2 + y_3)$. Thus

$$1 - (y_1 + y_2 + y_3) \leqq c_5 \leqq (1 + y_i) - (y_j + y_k)\ \forall i, j, k \text{ distinct.} \tag{20}$$

But $R \in \mathscr{R}_\triangle \Rightarrow (1 + y_i) - (y_j + y_k) \geqq 0\ \forall i, j, k$ distinct, so we can *always* choose convex coefficients as follows: assuming without loss that $y_3 = y_1 \wedge y_2 \wedge y_3$, we choose

$c_1 = y_3$	(largest possible c_1)	(21a)
$c_2 = y_1 - y_3$	(least possible c_2)	(21b)
$c_3 = y_2 - y_3$	(least possible c_3)	(21c)
$c_4 = 0$	(least possible c_4)	(21d)
$c_5 = (1 + y_3) - (y_1 + y_2)$	(largest possible c_5)	(21e)

For this choice of c_k's we have a convex decomposition of R, so

$\mathscr{R}_\triangle \subseteq \mathrm{conv}(\mathscr{R}_3)$. □

Before proceeding to the general case, we elaborate the possibilities for all convex decompositions with $n=3$. First, the coefficients (21) can always be chosen, yielding the largest c_1 and c_5 with least c_2, c_3, c_4. There are several cases for coefficients, depending on the relationship of $(y_1+y_2+y_3)$ to 1: again assuming $y_3=y_1 \wedge y_2 \wedge y_3$, we find using the constraints $0 \leqq c_i \leqq 1\ \forall i$, that

Case 1. If $y_1+y_2+y_3 \geqq 1$, then $c_5=0$ and $c_1=1/2((y_1+y_2+y_3)-1)$ are minimum, while c_2, c_3, c_4 from equations (21) become largest.

Case 2. If $y_1+y_2+y_3 \leqq 1$, then $c_5=1-(y_1+y_2+y_3)$ and $c_1=0$ are minimum, whereas $c_2=y_1$, $c_3=y_2$, $c_4=y_3$ are again largest.

Case 3. If $1+y_3=y_1+y_2$, $c_5=0$, and the decomposition (16) is unique.

Case 4. If $y_3=0$, $y_1+y_2 \leqq 1$, then $c_5=1-(y_1+y_2)$, and the decomposition (16) is unique.

For $n>3$ we find that max-$\triangle$ transitivity is necessary but not sufficient for R to be in $\mathrm{conv}(\mathscr{R}_n)$. If $\subset$ indicates *proper* subset, we have

Theorem 4.2. *For* $n>3$,

$$\mathrm{conv}(\mathscr{R}_n) \subset \mathscr{R}_\triangle.$$

Proof. Let $R \in \mathrm{conv}(\mathscr{R}_n)$, say $R=\sum_{k=1}^{p} c_k R_k$ with $\sum_{k=1}^{p} c_k=1$; $0 \leqq c_k \leqq 1\ \forall k$, and each $R_k \in \mathscr{R}_n$. Let R_{ijk} be any 3×3 principal submatrix of R, i.e.,

$$R=\begin{bmatrix} 1 \begin{bmatrix} 1 & y_i & y_j \\ & 1 & y_k \\ & & 1 \end{bmatrix} = R_{ijk} & \\ & \ddots \\ & & 1 \end{bmatrix}.$$

The decomposition of R also effects a convex decomposition of $R_{ijk} \in \mathrm{conv}(\mathscr{R}_3)=\mathscr{R}_\triangle$ by Theorem 4.1. Thus $1+y_i \geqq y_j+y_k\ \forall i, j, k$ distinct and for every principal (3×3) submatrix in R. Therefore $R \in \mathscr{R}_\triangle\ \forall n$.

To see that (22) is *proper* for $n>3$, we note that the matrix

$$R(\beta)=\begin{bmatrix} 1 & 0.3 & 0.6 & 0 \\ & 1 & 0.7 & 0 \\ & & 1 & \beta \\ & & & 1 \end{bmatrix}; \qquad \beta \in [0,1],$$

is in $\mathrm{conv}(\mathscr{R}_4) \Leftrightarrow \beta=0$, but lies in $\mathscr{R}_\triangle$ for all $\beta \in [0, 0.30]$. $\square$

To place $\mathrm{conv}(\mathscr{R}_n)$ in the hierarchy (12), we note that Zadeh exhibits in [1] a nested, non-convex decomposition of every $R \in \mathscr{R}_\wedge$ by hard relations in $\mathscr{R}_n$, viz., $R = \bigcup_k d_k R_k$, with $0 < d_k \leqq 1$, $R_1 \leqq R_2 \leqq \ldots \forall k$. Defining $c_{k+1} = d_{k+1} - d_k$, $k = 0, 1, 2, \ldots$ and replacing $\bigcup_k$ by $\sum_k$ yields convex decomposition $R = \sum_k c_k R_k$, hence we have

Theorem 4.3. *For* $n > 2$

$$\mathscr{R}_\wedge \subset \operatorname{conv}(\mathscr{R}_n). \tag{23}$$

To see that containment in (23) is proper for $n > 2$, recall that $R(\lambda)$ at (13) lies in $\mathscr{R}_\wedge \Leftrightarrow \lambda = 0.7$, whereas $R(\lambda) \in \mathscr{R}_\triangle$ for $\lambda \in [0.5, 0.9]$, and by Theorem 4.1 $\mathscr{R}_\triangle = \operatorname{conv}(\mathscr{R}_3)$. Accordingly, the placement of $\operatorname{conv}(\mathscr{R}_n)$ in (12) is

$$\mathscr{R}_\wedge \subset \operatorname{conv}(\mathscr{R}_n) \subset \mathscr{R}_\triangle \; \forall n > 3. \tag{24}$$

The position of $\operatorname{conv}(\mathscr{R}_n)$ with respect to $\mathscr{R}_\bullet$—not yet known—seems relatively unimportant, since the sets exhibited in (24) are the ones most useful in the applications.

Theorems 3.1, 4.1, and 4.3 combine to exhibit a rather interesting inter-relationship between the triangle inequality, max-$\triangle$ transitivity, and convexity. Before proceeding we summarize the relationships between $\mathscr{R}_n$, $\mathscr{R}_\wedge$, ($\operatorname{conv}(\mathscr{R}_n)$, and $\mathscr{R}_\triangle$):

If $n = 1: \mathscr{R}_n = \mathscr{R}_\wedge = \operatorname{conv}(\mathscr{R}_n) = \mathscr{R}_\triangle$,
If $n = 2: \mathscr{R}_n \subset \mathscr{R}_\wedge = \operatorname{conv}(\mathscr{R}_n) = \mathscr{R}_\triangle$,
If $n = 3: \mathscr{R}_n \subset \mathscr{R}_\wedge \subset \operatorname{conv}(\mathscr{R}_n) = \mathscr{R}_\triangle$,
If $n > 3: \mathscr{R}_n \subset \mathscr{R}_\wedge \subset \operatorname{conv}(\mathscr{R}_n) \subset \mathscr{R}_\triangle$.

In each case, the proper imbeddings are *nowhere dense* in the indicated superset, illustrating how sparse each set of similarity relations is in the next larger set for $n > 3$.

5. Clustering by convex decomposition

At present the only clustering procedure based on fuzzy relations appears to be the one described in [1–4]. A concise summary of this method follows: beginning with a reflexive, symmetric fuzzy relation matrix R, its *transitive closure* $\bar{R}$ is obtained by any of three algorithms: ($\vee . \wedge$) composition iteration, $R \leqq R^2 \leqq R^3 \ldots \leqq R^q = \bar{R} = R^k \in \mathscr{R}_\wedge \; \forall k \geqq q \leqq n-1$, Zadeh [1], or Tamura *et al.* [2]; a column-row scanning algorithm, Kandel and Yelowitz [3]; or Prim's minimal spanning tree algorithm, Dunn [4].

Once $\bar{R}$ is obtained, its entries are used to define a nested sequence of hard relations $R_{ij} \in \mathscr{R}_n$ by thresholding at levels in-between successive values of $\bar{r}_{ij}$. Thus one might have from the entry $\bar{r}_{ij}$ the hard relation $x_p R_{ij} x_q \Leftrightarrow \bar{r}_{pq} \geqq \bar{r}_{ij} \; \forall p, q$. In this fashion one may construct a nested sequence of hard equivalence relations (therefore hard c-partitions of X) in $X \times X$, which ultimately yield a partition tree or dendogram. While this method appeared at first to be quite novel, it was shown by Dunn in [4] that because max–min transitivity is equivalent to the ultra-metric inequality, the resultant hierarchies of hard clusters were in fact a subset of *single-linkage* hierarchies, from a well known graph-theoretic method for hard clustering [5]. Thus, no apparent advantage was gained, and further, the density of $\mathscr{R}_\wedge$ in $\mathscr{R}_\triangle$ is so slight that hierarchies generated this way are severely limited.

As an alternative, if $R \in \operatorname{conv}(\mathscr{R}_n)$, we can use its convex decomposition for clustering as follows: suppose $R = \sum_{k=1}^{p} c_k R_k$. Each $R_k \in \mathscr{R}_n$ is isomorphic to a hard c-partition of X, say $U_k \in \mathscr{P}_c$. Note that c, the number of clusters in X, is in general a function of k, so there will be no hope that $\sum_{k=1}^{p} c_k U_k$ is well-defined, although when it is, the resultant U lies in $\mathscr{P}_{fc}$. Thus from $R = \sum_{k=1}^{p} c_k R_k$ there follows the sequence

$$\left\{(c_k, U_k) \middle| U_k \in \mathscr{P}_c \; \forall 1 \leqq k \leqq p; \; \sum_{k=1}^{p} c_k = 1\right\}.$$

Since the decomposition $\sum c_k R_k$ exhibits the "percentage" of each R_k needed to build up fuzzy relation R, we interpret c_k as an indicator of the relative merit of the associated U_k as a c-partitioning of X. Note that this also provides a method for choosing c, the number of clusters most likely to exhibit substructure in X; and finally, observe that the partitions $\{U_k\}$ generated this way are *not* nested hierarchically. We exemplify both methods using the matrix $R(\beta)$ appearing above with $\beta = 0$.

Example.

$$R = \begin{bmatrix} 1 & 0.3 & 0.6 & 0 \\ & 1 & 0.7 & 0 \\ & & 1 & 0 \\ & & & 1 \end{bmatrix} \tag{25}$$

Composing R with itself using $\circ = (\vee \,.\, \wedge)$, we find

$$R^2 = \begin{bmatrix} 1 & 0.6 & 0.6 & 0 \\ & 1 & 0.7 & 0 \\ & & 1 & 0 \\ & & & 1 \end{bmatrix}; \qquad R^3 = \begin{bmatrix} 1 & 0.6 & 0.6 & 0 \\ & 1 & 0.7 & 0 \\ & & 1 & 0 \\ & & & 1 \end{bmatrix} = R^2 = \bar{R}$$

Hence the transitive closure of R is $\bar{R} = R^2 = R^3 \ldots$. A typical hierarchy of hard clusters derived from $\bar{R}$ using the $\bar{r}_{ij}$'s as thresholds is

$$\begin{pmatrix} x_i R_{0.59} x_j \Leftrightarrow \bar{r}_{ij} \geqq 0.59 \Rightarrow \{1,2,3\} \cup \{4\} \\ x_i R_{0.65} x_j \Leftrightarrow \bar{r}_{ij} \geqq 0.65 \Rightarrow \{1\} \cup \{2,3\} \cup \{4\} \\ x_2 R_{0.71} x_j \Leftrightarrow \bar{r}_{ij} \geqq 0.71 \Rightarrow \{1\} \cup \{2\} \cup \{3\} \cup \{4\} \end{pmatrix}. \tag{26}$$

In (26) we represent hard partition

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

more compactly, for example, as $\{1,2,3\} \cup \{4\}$, etc. Another way to represent (26) is by a dendogram, as in Fig. 3 below.

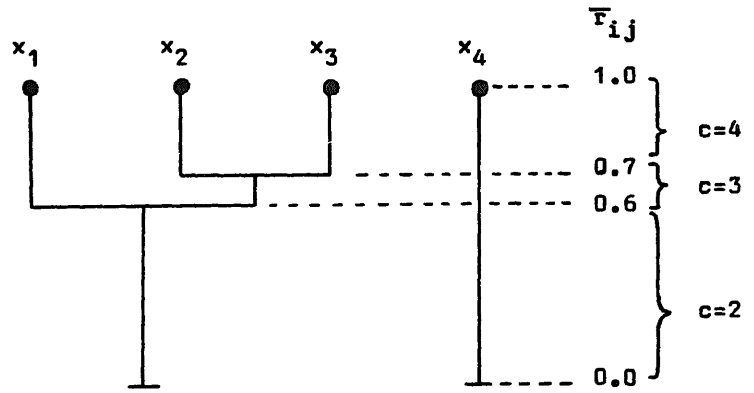

Fig. 3. Dendogram for clustering by transitive closure.

If we define $v(c)$, $c=2,3,4$, as the *jump in relation strength* between levels of clusters, we find that $v(2)=0.6$; $v(3)=0.10$; and $v(4)=0.30$. One commonly assumes that the order of the values $\{v(c)\}$ indicates the relative attractiveness of choices for c: in the present instance we infer from (26) that single-linkage hierarchies will manifest a strong preference for $c=2$, $\{1,2,3\} \cup \{4\}$; and for a second choice, $c=4$, $\{1\} \cup \{2\} \cup \{3\} \cup \{4\}$. Returning to (25), we agree that $\{x_4\}$ should always be isolated from $\{x_1, x_2, x_3\}$, but whether total segregation ($c=4$) is preferable to 3-partitions of X on the basis of the relationships shown in R is questionable.

Because column 4 of R is special, R in (25) has the unique convex decomposition

$$R = 0.3\underbrace{\begin{bmatrix} 1 & 1 & 1 & 0 \\ & 1 & 1 & 0 \\ & & 1 & 0 \\ & & & 1 \end{bmatrix}}_{R_1} + 0.4\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ & 1 & 1 & 0 \\ & & 1 & 0 \\ & & & 1 \end{bmatrix}}_{R_2} + 0.3\underbrace{\begin{bmatrix} 1 & 0 & 1 & 0 \\ & 1 & 0 & 0 \\ & & 1 & 0 \\ & & & 1 \end{bmatrix}}_{R_3}. \tag{27}$$

From (27) we obtain the sequence of clusterings

$$\left(\begin{array}{l} c_1=0.3;\ c=2;\ \{1,2,3\} \cup 4 \\ c_2=0.4;\ c=3;\ \{1\} \cup \{2,3\} \cup \{4\} \\ c_3=0.3;\ c=3;\ \{1,3\} \cup \{2\} \cup \{4\} \end{array}\right). \tag{28}$$

System (28) conveys a strikingly different portrait of substructure in X suggested by R than (26). From (28) and $c_2 = 0.4$ we infer that $c = 3$; $\{1\} \cup \{2, 3\} \cup \{4\}$ is the optimal clustering of X; and that $c = 2$ or 3 with $c_2 = c_3 = 0.3$ are equally likely as second choices. Note that the clusters obtained in this way *never* admit $c = 4$, a fact we find more consistent with our intuitive understanding of relationships in R. Moreover, the closeness of c_1, c_2, and c_3 indicates that R is fuzzier than the tree in Fig. 3 might imply.

A precise characterization of $\text{conv}(\mathscr{R}_n)$ is needed to make this technique applicable in general; Theorem 4.2 shows that max-$\triangle$ transitivity is necessary but not sufficient. Furthermore, an efficient algorithm, such as those described for convex decomposition of $U \in \mathscr{P}_{fc}$ in [6] has yet to be found. Nonetheless, our example seems to justify further efforts in this direction.

6. A connection between $\mathscr{P}_{fc}$ and $\mathscr{R}_\triangle$

As noted above, it is not generally possible to recover a $U \in \mathscr{P}_{fc} = \text{conv}(\mathscr{P}_{co})$ from an $R \in \text{conv}(\mathscr{R}_n)$. The difficulty lies with degeneracy in $\mathscr{P}_{co}$. In other words, the diagram in Fig. 4 is impossible ($\text{conv}(\mathscr{P}_{c0})$ cannot be isomorphic to $\text{conv}(\mathscr{R}_n)$).

$$
\begin{tikzcd}
\mathscr{P}_c \arrow[r, hook] \arrow[d, leftrightarrow] & \mathscr{P}_{co} \arrow[r, hook] & \text{conv}(\mathscr{P}_{co}) \arrow[d, leftrightarrow, dashed] \\
\mathscr{R}_n \arrow[rr, hook] & & \text{conv}(\mathscr{R}_n)
\end{tikzcd}
$$

Fig. 4.

Nonetheless, there is a mapping from $\mathscr{P}_{fc} \to \mathscr{R}_\triangle$ which provides a very nice method for deducing individual relationships in $X \times X$ from class membership in $U \in \mathscr{P}_{fc}$. To begin, we define the mapping $T_\circ : \mathscr{P}_{fc} \to V_{nn}$ by

$$T(U) = U^{\mathrm{T}} \circ U: \qquad (\circ) = \text{matrix multiplication as in (4).} \tag{29}$$

First, we prove that $T_\circ[\mathscr{P}_c]$, the image of hard c-partition space in V_{cn}, lies in hard equivalence relation space in V_{nn}:

Theorem 6.1. *For $T_\circ(U)$ defined in (29), we have for every $c = 1, 2, 3, \ldots n$ with n fixed.*

$$T_\circ[\mathscr{P}_c] \subset \mathscr{R}_n. \tag{30}$$

Proof. Let $\sim$ denote the equivalence relation in $X \times X$ induced by $U \in \mathscr{P}_c$. Thus $x_i \sim x_j \Leftrightarrow x_i$ and x_j are in the same subset of U in $\mathscr{P}_c$. We verify (30) for $(\circ) = (\sum_\bullet \wedge)$: the other products are similar. If $r_{ij}=\sum_{l=1}^{c}(u_{li}\wedge u_{lj})$, we have $\forall i, j$

$$\begin{aligned} r_{ij}=1 \Leftrightarrow \quad & (u_{li}\wedge u_{lj})=1 \quad \exists l, \\ \Leftrightarrow \quad & u_{li}=u_{lj}=1 \quad \exists l, \\ \Leftrightarrow & x_i \sim x_j. \end{aligned}$$

Otherwise, $r_{ij}=0$. Thus $R=U^{\mathrm{T}}(\sum_{\bullet}\wedge)U$ is equivalence relation $\sim$.

Corollary. *For fixed $n\in\mathbf{N}$, we have*

$$\bigcup_{i=1}^{n} T_{\circ}[\mathscr{P}_i]=\mathscr{R}_n. \tag{31}$$

It is clear from (31) that $T_{\circ}$ is not usually invertible. In fact, one may throw any number of degenerate images $T_{\circ}[\mathscr{P}_{io}]$ into the union (31) without altering its size. All one can assert when passing from $\mathscr{R}_n$ back to $\mathscr{P}_{co}$ is that there is a largest c, $1\leqq c\leqq n$, so that some

$$U\in\mathscr{P}_c \xrightarrow{T} R\in\mathscr{R}_n.$$

It is interesting that among the class of maps specified in (29) *only one* carries fuzzy c-partitions into similarity relations:

Theorem 6.2. *For $\circ=(\sum_{\bullet}\wedge)$ we have for $c=1,2,\ldots,n$ and n fixed,*

$$T_{\sum_{\bullet}\wedge}[\mathscr{P}_{fc}]\subseteq\mathscr{R}_{\triangle}. \tag{32}$$

Proof. Let $U\in\mathscr{P}_{fc}$, first, we note from (3) that in general

$$r_{ij}=\mathop{\square}_{l=1}^{c}(u_{li}*u_{lj}) \qquad \forall i,j.$$

Considering any of the products (4), we find that $R=U^{T}(\square_{\bullet}*)U$ is not reflexive (that is, that $r_{ii}<1\ \forall i$) unless $\square_{\bullet}*=\sum_{\bullet}\wedge$. To verify that $U^{T}(\sum_{\bullet}\wedge)U\in\mathscr{R}_{\triangle}$, we rewrite r_{ij} using the identity $a\wedge b=1/2(a+b-|a-b|)$: thus

$$\begin{aligned} r_{ij}&=\sum_{l=1}^{c}(u_{li}\wedge u_{lj}) \\ &=\sum_{l=1}^{c}\{1/2(u_{li}+u_{lj}-|u_{li}-u_{lj}|)\} \end{aligned}$$

$$r_{ij}=1-1/2\sum_{l=1}^{c}|u_{li}-u_{lj}|. \tag{33}$$

Now $u_{li} \in [0,1]\ \forall l$ and i, and since $\sum_{l=1}^{c} u_{li} = 1\ \forall i$, $\sum_{l=1}^{c} |u_{li} - u_{lj}| \in [0,2]$. Thus

(i) $0 \leqq r_{ij} \leqq 1\ \forall i, j$.
(ii) $r_{ii} = 1\ \forall i$ follows directly from (33).
(iii) $r_{ij} = r_{ji}\ \forall i \neq j$ because $u_{li} \wedge u_{lj} = u_{lj} \wedge u_{li}\ \forall i, j$.
(iv) Finally we check max-$\triangle$ transitivity: $\forall i, j, k$ we have

$$r_{ij} + r_{jk} - 1 \leqq r_{ik} \quad \Leftrightarrow$$

$$\sum_{l=1}^{c} \{(u_{li} \wedge u_{lj}) + (u_{lj} \wedge u_{lk}) - (u_{lj} \wedge u_{lj})\} \leqq \sum_{l=1}^{c} (u_{li} \wedge u_{lk}). \tag{34}$$

To verify (34) it suffices to see that for each term

$$(u_{li} \wedge u_{lj}) + (u_{lj} \wedge u_{lk}) - (u_{lj} \wedge u_{lj}) \leqq (u_{li} \wedge u_{lk}). \tag{35}$$

Let $a = u_{li}$, $b = u_{lk}$, and $c = u_{lj}$. Then (35) becomes $(a \wedge c) + (b \wedge c) - (a \wedge b) \leqq c$. Equivalently,

$$1/2\{(a + c - |a - c|) + (b + c - |b - c|) - (a + b - |a - b|)\} \leqq c \quad \Leftrightarrow$$

$$c + 1/2\{|a - b| - |a - c| - |b - c|\} \leqq c \quad \Leftrightarrow$$

$$1/2\{|a - b| - |a - c| - |b - c|\} \leqq 0 \quad \Leftrightarrow$$

$$-|a - b| + |a - c| + |b - c| \geqq 0 \quad \Leftrightarrow$$

$$|a - c| + |c - b| \geqq |a - b|. \tag{37}$$

Since (37) is just the triangle inequality $\forall$ real a, b, c, it follows that (34) holds, i.e. R is max-$\triangle$-transitive. □

Combining Theorems 3.1 and 6.2, we have a way to induce from every fuzzy c-partition of X a pseudo-metric on the data:

$$U \xrightarrow{T_{\Sigma,\wedge}} R = U^T(\textstyle\sum \cdot \wedge)U \rightarrow d_{ij} = 1 - r_{ij}. \tag{38}$$

As an example, suppose we find from the clustering algorithm fuzzy ISODATA [9] that an optimal 3-partitioning of X is

$$U = \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \begin{bmatrix} 0.3 & 0.9 & 0.85 & 0.10 & 0.11 \\ 0.5 & 0.05 & 0 & 0.25 & 0.78 \\ 0.2 & 0.05 & 0.15 & 0.65 & 0.11 \end{bmatrix} \end{array} \in \mathscr{P}_{f3}. \tag{39}$$

Aside from the information conveyed by (39) about the membership of each x_i in the 3 fuzzy clusters of X, one may wonder about individual relationships suggested by this partitioning. Applying $T_{\Sigma.\wedge}$ to U, we have the fuzzy similarity relation

$$T_{\Sigma.\wedge}(U)=U^T(\Sigma.\wedge)U=R=\begin{bmatrix} 1 & 0.40 & 0.45 & 0.55 & 0.72 \\ & 1 & 0.90 & 0.20 & 0.21 \\ & & 1 & 0.25 & 0.22. \\ & & & 1 & 0.44 \\ & & & & 1 \end{bmatrix} \tag{40}$$

From (40) we infer that x_2 and x_3 are most strongly related, $r_{23}=0.90$; that the bond between x_2 and x_4 is weakest, $r_{24}=0.20$. These conclusions seem corroborated by an application of intuition to the memberships exhibited in (39). Moreover, we have by Theorem 3.1 that

$$D=1-R=\begin{bmatrix} 0 & 0.60 & 0.55 & 0.45 & 0.18 \\ & 0 & 0.10 & 0.80 & 0.79 \\ & & 0 & 0.75 & 0.78 \\ & & & 0 & 0.56 \\ & & & & 0 \end{bmatrix} \tag{41}$$

is a pseudo-metric in $X\times X$ which ostensibly provides a natural measure of distance or dissimilarity in $X\times X$ induced by fuzzy partition (39).

Finally, note that we find in the proof of Theorem 6.2 a natural interpretation for the relation induced on $X\times X$ by $T_{\Sigma.\wedge}$. To discuss it, let $\|\cdot\|_1$ denote the l_1 norm of $\mathbf{y}\in\mathbf{R}^c$, viz., $\|\mathbf{y}\|_1=\sum_{i=1}^c |y_i|$. Then if $\mathbf{U}^{(i)}$ denotes the i^{th} column of $U\in\mathscr{P}_{fc}$, we have for r_{ij} the expression

$$r_{ij}=1-1/2\left(\sum_{l=1}^{c}|u_{li}-u_{lj}|\right)=1-1/2(\|\mathbf{U}^{(i)}-\mathbf{U}^{(j)}\|_1). \tag{42}$$

Thus the pseudo-metric induced on $X\times X$ by $U^T(\Sigma.\wedge)U$, $d_{ij}=1-r_{ij}$, is just half of the l_1 distance between membership columns of U:

$$d(x_i,x_j)=1/2(\|\mathbf{U}^{(i)}-\mathbf{U}^{(j)}\|_1). \tag{43}$$

In (43) it is evident why d is necessarily a pseudo-metric: $d_{ij}=0\Leftrightarrow\mathbf{U}^{(i)}=\mathbf{U}^{(j)}$, but clearly this can happen when x_i and x_j may be distinct. Since $d_{ij}=0\Leftrightarrow r_{ij}=1$, we have for $i\neq j$ that $x_iRx_j=1\Leftrightarrow x_i$ and x_j have identical *membership vectors* in the c fuzzy clusters of U. This is a very appealing generalization of the physical and mathematical meanings of $x_iRx_j=1$ for $R\in\mathscr{R}_n$. Conversely, from (42) it follows that $r_{ij}=0$ for $i\neq j\Leftrightarrow\|\mathbf{U}^{(i)}-\mathbf{U}^{(j)}\|_1=2\Leftrightarrow\mathbf{U}^{(i)}$ and $\mathbf{U}^{(j)}$ are *disjoint*: wherever x_i has membership in a fuzzy cluster in U, x_j cannot, and vice-versa. This also extends the meaning of $x_iRx_j=0$ for $R\in\mathscr{R}_n$ quite naturally.

7. Conclusions

Our imbedding of hard equivalence relations yields a new type of transitivity with very interesting properties. Max-$\triangle$ transitivity is equivalent to the triangle inequality, and is necessary for fuzzy similarity relations admitting a convex decomposition by equivalence relations. Furthermore, each similarity relation of this type induces a pseudo-metric in $X \times X$. For relations constructed from fuzzy c-partitions of X via $U^T(\sum . \wedge)U$, this pseudo-metric is half of the l_1 distance between membership vectors (for points in the data) for the fuzzy partition used. Examples of applications to cluster analysis and cluster validity seem to support our contention that the space of similarity relations characterized by max-$\triangle$ transitivity is an important one for applications in pattern recognition. A complete characterization of the convex hull of hard equivalence relation space together with an implementable decomposition algorithm will provide a new clustering method based on fuzzy similarity relations which seems to hold great promise: we hope to make this the subject of a future investigation.

References

[1] L.A. Zadeh, Similarity relations and fuzzy orderings, Inf. Sci. 3 (1971) 177–200.
[2] S. Tamura, S. Higuchi and K. Tanaka, Pattern classification based on fuzzy relations, IEEE Trans. SMC (1971) 61–66.
[3] A. Kandel and L. Yelowitz, Fuzzy chains, IEEE Trans. SMC (1971) 61–66.
[4] J. Dunn, A graph theoretic analysis of pattern classification via Tamura's fuzzy relation, IEEE Trans. SMC (1974) 310–313.
[5] R. Duda and P. Hart, Pattern Classification and Scene Analysis (Wiley-Interscience, New York, 1973).
[6] J. Bezdek and D. Harris, Convex decompositions of fuzzy partitions (in review) Inf. Sci. (1977).
[7] A. Rosenfeld, Fuzzy graphs, in: L.A. Zadeh et al., eds., Fuzzy Sets and their Applications to Cognitive and Decision Processes (Academic Press, New York, 1975) 77–96.
[8] R. Yeh and S. Bang, Fuzzy relations, fuzzy graphs, and their applications to clustering analysis, in: L.A. Zadeh et al., eds., Fuzzy Sets and Their Applications to Cognitive and Decision Processes (Academic Press, New York, 1975) 125–170.
[9] J. Bezdek, Mathematical models for systematics and taxonomy, in: G. Estabrock, ed., Proc. Eighth Int. Conf. Numerical Taxonomy (Freeman, SF, 1975) 143–166.
[10] R. Crowson, Classification and Biology (Atherton Press, New York, 1970).
[11] L.A. Zadeh, Fuzzy sets, Inf. Cont. (1965) 338–353.
[12] E. Ruspini, A new approach to clustering, Inf. Cont. (1969) 22–32.
[13] M. Woodbury and J. Clive, Clinical pure types as a fuzzy partition, J. Cybernetics (1974) 111–121.