



**HAL**  
open science

## Metrical properties of fuzzy relations

Bernadette Bouchon-Meunier, Gerard Cohen, P. Frankl

► **To cite this version:**

Bernadette Bouchon-Meunier, Gerard Cohen, P. Frankl. Metrical properties of fuzzy relations. Bulletin pour les sous-ensembles flous et leurs applications, 1981, 9, pp.39–47. hal-04561091

**HAL Id: hal-04561091**

**<https://hal.science/hal-04561091>**

Submitted on 26 Apr 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

METRICAL PROPERTIES OF FUZZY RELATIONS

B. BOUCHON\* - G. COHEN\*\* - P. FRANKL\*

ABSTRACT

We study distance relations  $d$  defined on a finite set  $E$ , associated with a fuzzy binary relation  $R$  verifying a generalized form of transitivity. We classify these distance relations with respect to  $k$ -ultrametricity and  $m$ -hypermetricity introducing the new criterium of  $l$ -supermetricity.

Cet article a été présenté au Colloque International sur la Théorie de l'Information qui s'est tenu à Budapest (Hongrie) du 24 au 28 Août 1981.

\*C.N.R.S., Université Paris VI, 4 Place Jussieu - 75230 Paris Cedex 05.  
FRANCE.

\*\* ENST, 46 Rue Barrault, 75013 Paris, FRANCE.

## 0. STATEMENT OF THE PROBLEM

Fuzzy binary relations generalize the classical deterministic binary relations between elements of a given population. They can be used in problems where two different elements are neither absolutely bound nor strictly unbound. They admit various applications in clustering, data analysis and classification problems ; they are softer tools than classical binary relations, since the association of two different elements can be modulated by numbers between 0 and 1.

In particular cases, these fuzzy relations yield distances between the members of the population. In this paper, we consider such relations which further verify a generalization of the transitivity property. We study metricity and hypermetricity in view of using relations weaker than the ultrametric one, most commonly considered.

We introduce supermetricity to simplify the study of hypermetric relations. A generalization of ultrametricity,  $k$ -ultrametricity, is presented; extremal  $k$ -ultrametric spaces are defined and characterized.

## 1. INTRODUCTION

Let  $E$  be a finite set of cardinality  $n$ . A fuzzy binary relation  $R$  on  $E$  is a mapping from  $E \times E$  to  $[0,1] \subset \mathbb{R}$ . More particularly, we consider a  $*$ -relation  $R$ , i.e.  $R$  is :

- definite : for  $x, y$  in  $E$ ,  $R(x, y) = 1$  if and only if  $x=y$ ,
- symmetric :  $R(x, y) = R(y, x) \forall x, y$ ,
- $*$ -transitive ([1], [9]) : for every  $x, y$  in  $E$   $R(x, y) \geq \sup_{z \in E} R(x, z) * R(y, z)$ , for an operation  $*$  from  $[0,1] \times [0,1]$  to  $[0,1]$ ,  $\vee$  denoting the supremum.

We further suppose that  $R$  verifies  $\forall x, y, z$  :  
 $R(x, z) * R(y, z) \geq ((R(x, z) + R(y, z) - 1) \vee 0)$  ; then  $d(x, y) = 1 - R(x, y)$  satisfies the triangular inequality ;  $d$  is the T-distance or distance associated with  $R$ .

## 2. PARTICULAR $*$ -RELATIONS AND ASSOCIATED DISTANCES

Definition 1 : The distance  $d$  is called :

- ultrametric (UM-distance or UM) if  $\forall x, y, z$   $d(x, y) \leq d(x, z) \vee d(z, y)$ ,
- $k$ -hypermtric (k-HM-distance or k-HM) [3], [6] if  $\forall x_1, \dots, x_k, y_1, \dots, y_{k+1}$  in  $E$   
 $\sum_{1 \leq i < j \leq k} d(x_i, x_j) + \sum_{1 \leq i < j \leq k+1} d(y_i, y_j) \leq \sum_{1 \leq i \leq k} \sum_{1 \leq j \leq k+1} d(x_i, y_j)$ ,
- hypermetric (HM-distance or HM) if  $d$  is  $k$ -hypermtric  $\forall k \in \mathbb{N}$ .
- four-point (FP-distance or FP) if  $\forall x, y, z, t$ ,  
 $d(x, y) + d(z, t) \leq (d(x, z) + d(y, t)) \vee (d(x, t) + d(y, z))$ ,
- Probabilistic (P-distance or P) if  $\forall x, y, z$   
 $d(x, y) \leq d(x, z) + d(z, y) - d(x, z) d(z, y)$ .

It is easy to check :

$$\forall a, b \in [0,1] \quad a+b-ab \geq a \vee b \geq \frac{a+b}{2} \geq a \wedge b \geq ab \geq (a+b-1) \vee 0. \quad (1)$$

As P and UM are associated with  $\ast$ -relations defined by :

$$a \ast_p b = ab$$

$$a \ast_{UM} b = a \wedge b,$$

we obtain :

$$UM \Rightarrow P \Rightarrow T. \quad (2)$$

We also remark :

$$1H \iff T \quad (3)$$

$$k\text{-HM} \implies (k-1)\text{-HM} \quad (4)$$

$$\text{and [6] } UM \Rightarrow FP \implies HM \implies T. \quad (5)$$

Comparing results (2) and (5), we try to insert links concerning FP, P and k-HM in the same chain.

Proposition 1 : a FP-distance is not necessarily a P-distance.

Proof : Consider  $E : \{P, Q, R, S\}$ .

$$d(P,Q) = d(S,Q) = 0.195, \quad d(P,S) = 0.01$$

$$d(P,R) = d(S,R) = d(R,Q) = 0.1.$$

Then  $d$  is FP, but

$$d(P,Q) > d(P,R) + d(R,Q) - d(P,R) d(R,Q).$$

Hence  $d$  is not a P-distance.

Proposition 2 : a P-distance is not necessarily a 2H-distance.

Proof : Consider  $E : \{P_1, P_2, Q_1, Q_2, Q_3\}$ .

$$\forall i, j \quad d(P_i, Q_j) = 0.51a, \quad d(Q_i, Q_j) = a$$

$$d(P_1, P_2) = 0.1a, \quad \text{for } a \in ]0, 1[.$$

Then  $d$  is P-distance for  $a$  small enough (for example  $a = 10^{-2}$ ), but is not 2H-distance, since

$$d(P_1, P_2) + d(Q_1, Q_2) + d(Q_1, Q_3) + d(Q_2, Q_3) = 3.1a$$

$$> \sum_{i=1, 2} d(P_i, Q_1) + d(P_i, Q_2) + d(P_i, Q_3) = 3.06a.$$

Looking at (1), we consider other distances, allowing to add an element to the chain (2).

.../...

Definition 2 : We associate IP-distance (or IP), JP-distance (or JP), and  $\gamma$ -distance (or  $\gamma$ -d) [4], with the following operations

$$a *_{IP} b = (a \wedge b) (a + b - ab)$$

$$a *_{JP} b = ab \frac{1 + a \wedge b}{a + b}$$

$$a *_{\gamma d} b = \frac{ab}{\gamma + (1-\gamma)(a+b-ab)} \quad 0 \leq \gamma \leq 1$$

Proposition 3 :

- i)  $UM \Rightarrow IP \Rightarrow P$
- ii)  $UM \Rightarrow JP \Rightarrow \gamma d \quad (\forall \gamma) \Rightarrow P$
- iii)  $IP \not\Rightarrow JP$  and  $JP \not\Rightarrow IP$
- iv)  $\exists \gamma$  such that  $IP \not\Rightarrow \gamma d$
- v)  $\forall \gamma \quad \gamma d \not\Rightarrow IP$
- vi)  $IP \not\Rightarrow FP$
- vii)  $JP \not\Rightarrow FP$ .

Proof : i) and ii) are simple.

For iii), iv) and v), take  $a = 10^{-1}$ ,  $b = 10^{-2}$  then  $a *_{\gamma d} b > a *_{IP} b$  for  $\gamma = 0$ , and  $a *_{IP} b < a *_{JP} b$ .

Take  $a = 0.9$  and  $b = 0.8$ , then  $a *_{IP} b > a *_{JP} b > a *_{\gamma} b \quad \forall \gamma$ .

For vi), take  $E : \{P, Q, R, S\}$  with  $d(P,S) = d(R,S) = d(Q,S) = d(P,R) = a$ . and  $d(P,Q) = a - \epsilon$ ,  $d(R,Q) = a + \epsilon$ ; then  $d$  is not FP, but is IP for e.g.  $a = 10^{-1}$ ,  $\epsilon = 10^{-4}$ .

For vii), consider the same example with  $a : 0.9$ ,  $\epsilon = 10^{-4}$ ; then  $d$  is JP.

### 3. A SUFFICIENT CONDITION FOR M-HYPERMETRICITY

Definition 3 : A distance  $d$  is called  $m$ -supermetric (or  $m$ -SM) if the following condition holds for any  $x, y, z$  in  $E$  :

$$d(x,y) \leq d(x,z) \vee d(y,z) + \frac{1}{m} (d(x,z) \wedge d(y,z)). \quad (6)$$

Proposition 4 :  $m$ -supermetricity is sufficient for  $m$ -hypermetricity.

Proof : Let  $d$  satisfy (6) and take any  $2m+1$  points  $A_1, A_2, \dots, A_m, B_0, B_1, \dots, B_m$  in  $E$ . Denote  $d(A_i, B_j)$  by  $A_i B_j$ .

.../...

By symmetry we can take  $\min_{1 \leq i \leq m, 0 \leq j \leq m} A_i B_j = A_m B_m$  and again by symmetry  $\min_{1 \leq i \leq t, 0 \leq j \leq m} A_i B_j = A_t B_t$  for all  $t \leq m$ . Then we get :

$$A_i A_t \leq A_i B_t + \frac{1}{m} A_t B_t \quad 1 \leq i < t \leq m$$

$$B_i B_t \leq A_t B_i + \frac{1}{m} A_t B_t \quad 0 \leq i < t \leq m$$

Summing on all  $i, j, t$  yields

$$1 \leq i < j \leq m \quad A_i A_j + 0 \leq i < j \leq m \quad B_i B_j \leq \sum_{1 \leq i < j \leq m} A_i B_j + \sum_{0 \leq i < j \leq m} A_j B_i + \sum_{1 \leq t \leq m} \frac{2t-1}{m} A_t B_t.$$

The right hand side term (RHS) is

$$\sum_{i \neq j, 0 \leq j \leq m, 1 \leq i \leq m} A_i B_j + \sum_{t \leq \frac{m}{2}} \frac{2t-1}{m} A_t B_t + \frac{2(m+1-t)-1}{m} A_{m+1-t} B_{m+1-t} + (m-2) \binom{m}{2} \frac{A_{m+1}}{2} \frac{B_{m+1}}{2}.$$

Now the second term of RHS is  $\sum_{t \leq \frac{m}{2}} \frac{2t-1}{m} A_t B_t + (1 - \frac{2t-1}{m}) A_{m+1-t} B_{m+1-t} + A_{m+1-t} B_{m+1-t}$  and is less than  $A_t B_t + A_{m+1-t} B_{m+1-t}$  since  $A_{m+1-t} B_{m+1-t} \leq A_t B_t$ . Hence  $\sum_{1 \leq i < j \leq m} A_i A_j + \sum_{0 \leq i < j \leq m} B_i B_j \leq \sum_{1 \leq i \leq m, 0 \leq j \leq m} A_i B_j$ .

Remark : The condition is best possible : take  $E =$

$\{A_1 \dots A_m, B_0 \dots B_m\}$ ,  $A_i B_j = 1$ ,  $A_i A_j = 1 + \alpha = B_i B_j$  for all  $i, j$ . Then  $d$  is  $m$ -HM iff  $\alpha \leq \frac{1}{m}$  because  $\sum_{1 \leq i < j \leq m} A_i A_j + \sum_{0 \leq i < j \leq m} B_i B_j = [\binom{m}{2} + \binom{m+1}{2}] (1 + \alpha) = m^2 (1 + \alpha)$  is not greater than  $\sum_{1 \leq i \leq m, 0 \leq j \leq m} A_i B_j = m(m+1)$  iff  $\alpha \leq \frac{1}{m}$ .

Corollary : A JP distance is 2-HM but not 3-HM. A IP-distance is 4-HM but not 5-HM.

Proof : After some computations, from Definition 2, we obtain for any triangle with sides  $x, y, z$  and  $y \geq z$  (say) :  
 for a JP distance :  $x \leq y + \frac{(1-y)^2 z}{2-z-y} \leq y + \frac{(1-y)^2 z}{2(1-y)} \leq y + \frac{z}{2}$ .  
 For a IP distance :  $x \leq y + yz - y^2 z = y + z(y-y^2)$   
 so  $x \leq y + \frac{1}{4} z$ .

.../...

Hence by Proposition 4 a JP distance is 2-HM and a IP-distance is 4-HM. To see that this is best possible take  $E = \{A_1, A_2, A_3, B_0, B_1, B_2, B_3\}$ ,  $A_i A_j = 0.014 = B_i B_j$ ,  $A_i B_j = 0.01$  :

$d$  is JP and not 3-HM. Take now :

$E = \{A_1, \dots, A_5, B_0, \dots, B_5\}$ ,  $A_i A_j = 5/8 = B_i B_j$ ,  $A_i B_j = 0.5$  ;  
 $d$  is IP and not 5-HM.

#### 4. EXTREMAL K-ULTRAMETRIC SPACES

The notion of  $k$ -ultrametricity is introduced in [2] with the following definition.

Definition 4 : A distance  $d$  is  $k$ -ultrametric ( $k$ -UM) if for any  $k+2$  points  $x_1, x_2, \dots, x_{k+2}$  in  $E$ , the  $\binom{k+2}{2}$  distances between them satisfy

$$d(x_i, x_j) \leq \vee d(x_s, x_t) \quad \text{for all } i, j \\ (s, t), \{s, t\} \neq \{i, j\}$$

i.e. the largest distance between any  $k+2$  points is realized at least twice.

The case  $k = 1$  corresponds to  $d$  being ultrametric ; a  $k$ -UM distance is also a  $(k+1)$  - UM distance. In [2], we show that  $k$ -UM is equivalent to weak  $k$ -ultrametricity [7].

Definition 5 : Let  $C(n, k)$  be the maximum number of different distances between the  $n$  points of a  $k$ -UM space  $E$ . Such an  $E$  is called extremal.

Proposition 5 :  $C(n, 1) = n-1$ .

This will be a consequence of the following.

Proposition 6 : If the complete graph  $K_n$  is coloured with at least  $n$  colours, there exists a tricolor triangle.

Proof : By induction. For  $n = 3$ , it is obvious. Suppose it is true for  $n-1$ . Let us consider a colouring of  $K_n$  with  $n$  colours. If for some vertex  $x$  of  $K_n$ , the deletion of  $x$  (and all edges incident to  $x$ ), suppresses only one colour, we are done by induction. So suppose that for all  $x$  this deletion suppresses at least two colours. Consider



the following  $n \times n$  incidence tableau  $T = (t_{ij})$  with  $t_{ij} = 1$  iff the deletion of the vertex  $x_i$  suppresses colour  $c_j$ . Every line contains at least two 1's, so the average number of 1's per column is two. But no colour can be incident to three vertices so there are exactly two 1's per column. This means every colour is used once, which is impossible for  $n \geq 4$ , since there are  $\binom{n}{2} > n$  edges. So  $C(n, k) \leq n-1$ . Now a  $(n-1)$ -coloration of  $K_n$  without tricolor triangle is easy to find.

Proposition 7 :  $C(n, t) = T(n, t+1) + 1$  for  $t \geq 2$ , where  $T(n, i)$  is the Turán number (maximum number of edges in a graph with  $n$  vertices not containing a  $K_i$ ).

The proof is deduced of the following result.

Proposition 8 [5]: If the edges of  $K_n$  are colored with  $T(n, t)+2$  colours, then there exists a  $\binom{t+1}{2}$ -colored  $K_{t+1}$  in  $K_n$ . Moreover the extremal graphs (containing  $T(n, t) + 1$  colors and no such  $K_{t+1}$ ) are all obtained by coloring the edges of a  $(t-1)$ -partite Turán graph with different colors and using an extra color for the remaining edges of  $K_n$ .

Remarks : To get the metric results, replace "color" by "distance", an "extra color" by "largest distance" in Prop. 8. Prop. 6 is slightly more general than Prop. 5, stating that  $n$  points in a ultrametric space realize at most  $n-1$  distances. In fact this holds as soon as all triangles are isoscele (but not necessarily with sides  $x, y, z$  such that  $x = y \geq z$ ). The triangular inequality is not even needed.  $T(n, t+1) \sim \frac{t-1}{t} \binom{n}{2}$  for  $n$  large.

For  $t \geq 2$  the structure of the extremal metric space is essentially unique from Proposition 8. For  $t = 1$  the number of maximal chains of partitions [8] gives an upper bound for the number of non-isomorph extremal ultrametric spaces on  $n$  points.

#### Relations between FP, k-UM, 1-SM, m-HM

Observe first that the strength of  $k$ -SM and  $k$ -HM increases with  $k$ , where as it is the opposite for  $k$ -UM. We have seen that  $k$ -SM  $\Rightarrow$   $m$ -HM  $\forall m \leq k$ .

The converse is false : take the Hamming distance ; it is hypermetric (m-HM,  $V_m$  [3] ) but not 1-SM for  $1 > 1$ .

Proposition 9 : FP, k-UM and 1-SM are not comparable.

Proof : FP  $\not\Rightarrow$  2-SM : take  $E : \{P, Q, R, S\}$ , with  $d(R,S) = 0.75$   
 $d(Q,R) = 0.5 + \epsilon$ ,  $d(P,R) = d(Q,S) = 0.5$ ,  $d(P,S) = 0.5 - \epsilon$ ,  $d(P,Q) = 0.25 - \epsilon$ .

FP  $\not\Rightarrow$  k-UM. Take  $E = \{P, Q, R, S, X_1, X_2, \dots, X_{n-4}\}$ ,  
 $d(X_i, R) = d(X_i, S) = 0.9$  ;  $d(X_i, P) = d(X_i, Q) = 0.7$  ;  $d(X_i, X_j) = 0.6$  ;  
 $d(R, Q) = d(Q, S) = d(P, S) = d(R, P) = 0.8$  ;  $d(R, S) = 1$  ;  $d(P, Q) = 0.6$ .

k-SM  $\not\Rightarrow$  k-UM. Take  $E = \{X_i\}$  and  $\{d(X_i, X_j)\} = \{1 - \frac{t}{2\binom{n}{2}k}, 1 \leq t \leq \binom{n}{2}\}$ .

k-SM  $\not\Rightarrow$  FP, k-UM  $\not\Rightarrow$  FP, k-UM  $\not\Rightarrow$  1-SM for  $k \geq 2$  are easy to check (for  $k \geq 2$  k-UM  $\not\Rightarrow$  triangular inequality).

#### REFERENCES

- [1] BEZDEK J.C., HARRIS J.D. : Fuzzy partitions and relations : an axiomatic basis for clustering ; Fuzzy sets and systems 1 (1978), pp 111-127.
- [2] BOUCHON B., COHEN G. : On Fuzzy Relations and Partitions, Policy Analysis and Information Systems. To appear.
- [3] DEZA M. (TYLKIN M.E.) : On Hamming Geometry of Unitary Cubes (Russian) Doklady Akad. Nauk SSSR 134 n° 5 (1960) 1037-1040.
- [4] DUBOIS D. : Quelques classes d'opérateurs remarquables pour combiner des ensembles flous, BUSEFAL n° 1 (1980), pp 29-35.
- [5] ERDŐS P., SOS V., SIMONOVITS : Anti-Ramsey Theorems, in Finite and Infinite Sets, Keszthely (1973) 633-644.
- [6] KELLY J.B. : Hypermetric Spaces, in The Geometry of Metric and Linear Spaces, Lect. Notes in Math 490, Springer Verlag (1975) pp 17-31.
- [7] KIM K.H., ROUSH F. W. : Ultrametric and Matrix Theory, J. of Math. Psy. 18 (1978) pp 195-203.
- [8] MONJARDET B. : Personal Communication.
- [9] ZADEH L.A. : Similarity Relations and Fuzzy Orderings, Inf. Sci. 3 (1971) , pp 177-200.