

A New Index Based on Sparsity Measures for Comparing Fuzzy Partitions

Romain Quéré and Carl Frélicot

Mathématiques, Image et Applications – Université de La Rochelle
`{romain.queré,carl.frelicot}@univ-lr.fr`

Abstract. This article addresses the problem of assessing how close two strict and/or fuzzy partitions are. A new index based on a measurement of the sparsity of the contingency matrix crossing the partitions is proposed that satisfies the required properties formulated within the paper and presents a low complexity. It is compared to well-known existing indices of the literature, such as the Rand and the Jaccard indices, the transfert distance and some of their recent fuzzy counterparts.

Keywords: Cluster analysis, Rand index, Jaccard Index, Transfert distance, Sparsity measure, Fuzzy residual implications.

1 Introduction

A *partition* of a set $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ of n objects is a set of c non-empty subsets of X , called *clusters*, that group objects along common attributes they share. Partitions are usually characterized by a $(c \times n)$ *partition matrix* $U = (u_{ik})_{i=1,c;k=1,n}$, identified with a c -*partition* of X for the sake of simplicity. Each u_{ik} represents the degree with which the k^{th} datum is associated to the i^{th} cluster, each column \mathbf{u}_k gathers the degrees for the k^{th} object and each row U^i defines the i^{th} cluster. In this paper, we focus on *fuzzy/probabilistic* partitions such that $u_{ik} \in [0, 1]$ and $\sum_{i=1}^c u_{ik} = 1$, and on *strict* partitions such that u_{ik} are binary and sum up to unity, *e.g.*:

$$U_h = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 & \mathbf{u}_4 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{matrix} U_h^1 \\ U_h^2 \end{matrix} \text{ and } U_f = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 & \mathbf{u}_4 \\ 0.6 & 0.8 & 0.3 & 0.9 \\ 0.4 & 0.2 & 0.7 & 0.1 \end{pmatrix} \begin{matrix} U_f^1 \\ U_f^2 \end{matrix}.$$

Since clustering algorithms always produce a partition U even if there is no cluster structure in the data, assessing the quality of U is a problem of great interest. It can be tackled using the data itself, by mean of an *internal index* as in *cluster validity* [1, 2] or by assessing how close U is to a ground truth/ expert assessed (mostly strict) partition or a set of ordinary partitions, respectively by mean of an *external* and a *relative* index, both referred as *comparison indices* [3, 4]. This approach have been largely explored for both strict and fuzzy domains, see section 2. We propose to use sparsity measures [5] and fuzzy residual implications [6] to define a new fuzzy index in section 3. In section 4, numerical experiments show its good properties as compared to other indices.

2 Necessary Tools for the Index Construction

2.1 Comparing Partitions

Plenty of indices have been proposed in the literature for comparing partitions. Depending on the nature of the latter, *i.e.* strict or fuzzy, these indices rely on multiple and different techniques or theoretical frameworks. For the strict case, let us cite the well-known *Rand* and *Jaccard* indices [7] based on a set-theoretical approach and respectively denoted RI and JI hereafter, and the *transfert distance* TD based on graph theory [8]. Among fuzzy comparison indices, let us cite the recent Anderson et al. [9] and the Quéré and Frélicot's [10] extensions, both relying on fuzzy logics and denoted respectively RI_A , JI_A , RI_{QF} , JI_{QF} , the Huellermeier and Rifqi extension of the Rand index HR based on a geometrical approach [11], and the Campello's fuzzy extension of the transfert distance FTD [12]. The ideal index I^* , no matter the nature of the partitions U and V it is meant to compare, must satisfy the following properties (see Table 1 for mentioned indices): (I1) $I^*(U, V) = 1 \Leftrightarrow U \equiv V$ (*identity*), (I2) $I^*(U, V) \geq 0$ (*non-negativity*), (I3) $I^*(U, V) = I^*(V, U)$ (*symmetry*). Moreover, we consider that such an index should satisfy an additional informal property: (I4) $I^*(U, V) \gg I^*(U, W)$ if V is known to be much more closer to U than W (*dynamics*). If a practitioner decides whether or not two partitions are compatible by thresholding the index value, such an informal property ensures him that the index is known to present very different values while comparing close and distant partitions. Because of lack of space, we do not go further into the details of each index. It is not the purpose of this paper and we invite the interested reader to refer to surveys of quality [7, 9]. Yet, let us describe a well-known construction to go one step further in our proposition.

Table 1. Properties satisfied by some indices of the literature

Property	RI	JI	TD	RI_A	JI_A	RI_{QF}	JI_{QF}	HR	FTD
(I1) Identity	•	•	•					•	•
(I2) Non-negativity	•	•	•	•	•	•	•	•	•
(I3) Symmetry	•	•	•	•	•	•	•	•	•

Crossing a c -partition U and a r -partition V results in a $(c \times r)$ contingency matrix $N(U, V) = (n_{ij})_{i=1, c; j=1, r}$ whose general term n_{ij} represents the number of data being in the i^{th} cluster of U and in the j^{th} cluster of V . If both U and V are strict, the cardinal of the intersection between each pair (U^i, V^j) of clusters is given by [13]:

$$N(U, V) = U {}^tV. \quad (1)$$

where tV stands for the transpose of V . This is the basis of some set-theoretical indices such as the strict Rand and Jaccard indices, or the fuzzy Anderson et al. extension [9] where a new computation of N is proposed, whose elements n_{ij} are replaced by:

$$n_{ij}^\top(U, V) = \sum_{k=1}^n \top(u_{ik}, v_{jk}) \quad (2)$$

where \top is t-norm¹, see [14]. Basic t-norms are the minimum $\top_M(a, b) = \min(a, b)$ and the product $\top_P(a, b) = a b$. There also exist parametrized families of t-norms, *e.g.* the Hamacher's one $\top_{H_\gamma}(a, b) = \frac{ab}{\gamma + (1-\gamma)(a+b-ab)}$, $\gamma \in [0, +\infty)$. Note that n_{ij} induced by (1) is strictly equivalent to (2) computed with \top_P .

2.2 Sparsity Measures

A fundamental problem in many data analysis problems is to find a suitable representation of the data, say $\mathbf{y} = \{\mathbf{y}_1, \dots, \mathbf{y}_c\} \in \mathbb{R}^c$. A sparsity (or sparseness) measure aims at assessing to which extend most values in \mathbf{y} are close to zero while only few ones are non-zero so that they can be used to represent the data. Many sparsity measures are found in the literature, mainly coming from fields such as signal analysis, *e.g.* in [15]. For a comparison of fifteen well-known sparsity measures, the reader should refer to [5]. With no loss of generality, we restrict ourselves to sparsity measures $\mathcal{S} : \mathbb{R}^c \rightarrow [0, 1]$. Among the properties such measures may have, let us cite the two that are required for the comparison index we propose: (S1) adding a constant to each value decreases sparsity, (S2) as one value becomes infinite, as sparse as possible is the distribution. Two of the sparsity measures reviewed in [5] have these properties. The first one is the *Hoyer's sparsity measure* which is based on the relationship between the L_1 norm and the L_2 norms. It is defined as:

$$H(\mathbf{y}) = \frac{\sqrt{c} - \frac{\sum_{j=1}^c \mathbf{y}_j}{\sqrt{\sum_{j=1}^c \mathbf{y}_j^2}}}{\sqrt{c} - 1}. \quad (3)$$

It varies from 0, *i.e.* \mathbf{y} is not sparse, if all components are equal (up to signs) to unity if \mathbf{y} contains a single non-zero element. The second one is called *kurtosis sparsity measure* by analogy to the well-known measure of peakedness of a probability distribution. It is defined as:

$$\kappa_4(\mathbf{y}) = \frac{\sum_{j=1}^c \mathbf{y}_j^4}{\left(\sum_{j=1}^c \mathbf{y}_j^2\right)^2}. \quad (4)$$

In order to show how these two measures behave, we have driven a short experiment inspired by the work in [5]. Consider a vector \mathbf{y} of 500 values in $\{0, 1\}$ drawn from a Bernoulli distribution, so that 1 and 0 have a respective probability of occurrence of p and $q = 1 - p$. When q is barely null, \mathbf{y} is then composed of very few zeros while only a small number of values are 1 when q is close to 1. Thus as q increase so should the sparsity measure as exhibited in Fig. 1. One can observe that H presents more granularity than κ_4 which only gives a strong response for $q > 0.9$.

¹ A t-norm is binary operation on the unit interval $\top : [0, 1]^2 \rightarrow [0, 1]$ which is commutative, associative, non decreasing and has 1 for neutral element.

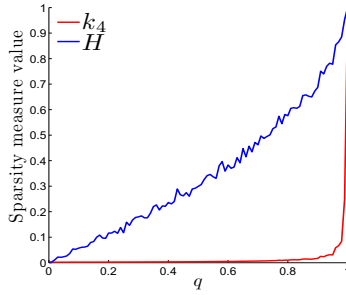


Fig. 1. Sparsity measures of 500 data drawn from a Bernoulli distribution

2.3 Fuzzy Implications

A fuzzy residual implication is an application $\mathcal{I} : [0, 1]^2 \rightarrow [0, 1]$, $(a, b) \mapsto \mathcal{I}(a, b)$, such that:

$$\mathcal{I}(a, b) = \sup_t \{t \in [0, 1] : \top(a, t) \leq b\} \quad (5)$$

where \top is t-norm. We speak about an implication function if \mathcal{I} is non-increasing in the first variable, non-decreasing in the second variable and $I(0, 0) = I(1, 1) = 1$, and $I(1, 0) = 0$, see [6] for a large survey on fuzzy implication functions. Within theses implications, the well-known *Gödel* is obtained with \top_M and given by:

$$\mathcal{I}_M(a, b) = \begin{cases} 1 & \text{if } b \geq a \\ b & \text{if } b < a \end{cases} \quad (6)$$

As well, parametrical fuzzy implications are defined, *e.g.* the Hamacher's ones, defined by [16]:

$$\mathcal{I}_{H_\gamma}(a, b) = \begin{cases} 1 & \text{if } b \geq a \\ \frac{b(\gamma + a - \gamma a)}{b(\gamma + a - \gamma a) + a - b} & \text{if } b < a \end{cases} \quad (7)$$

3 The New Index

Let us consider, for pedagogical purpose, the following two strict partitions :

$$U_h = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 & \mathbf{u}_4 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{matrix} U_h^1 \\ U_h^2 \end{matrix} \text{ and } V_h = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{matrix} V_h^1 \\ V_h^2 \\ V_h^3 \end{matrix}.$$

Both share in common the information that elements \mathbf{x}_2 and \mathbf{x}_4 belong to the same cluster, while \mathbf{x}_3 belongs to another one. Actually, the only difference between U_h and V_h is about \mathbf{x}_1 , which is grouped with \mathbf{x}_2 and \mathbf{x}_4 in U_h while it is put aside in its own cluster in V_h because it probably differs in some subtle ways of \mathbf{x}_2 and \mathbf{x}_4 . The idea here is that V_h can be seen as a refinement of U_h , so that V_h^1 is included in U_h^1 , as pointed out by the contingency matrix crossing the two partitions $N(U_h, V_h) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ by (1). Indeed, the 1^{st} row $\mathbf{n}_1 = (2 \ 1 \ 0)$ and the 1^{st} column ${}^t\mathbf{n}_1 = (2 \ 0)$ of $N(U_h, V_h)$ show that the only two elements of V_h^1

also belong to U_h^1 , leading thus to conclude that $V_h^1 \subset U_h^1$. Moreover, let us have a look at the matrices $N(U_h, U_h) = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$ and $N(V_h, V_h) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ crossing

each partition with itself. Both logically present only one non-zero element on each of their rows and columns, so that it can be directly connected with the concept of sparsity. This is the main idea of our proposition : the closest the partitions, the more sparse the rows and columns of their contingency matrix. This idea is also valid when crossing two fuzzy partitions U_f and V_f . The inner product between U_f^i and V_f^j induced by (1) will be high if and only if both clusters are similar and are not too fuzzy, *i.e.* they have a certain amount $n_s < n$ of same components close to 1, so that $U_f^i \cdot V_f^j$ tends to n_s . Follows that considering the whole set of clusters of U_f and V_f , the elements of N will be large for crossed fuzzy clusters having a lot of values in common, and will be small for frankly different ones, so that the idea of sparsity as defined in section 2.2 is preserved. The same reasoning holds while N is computed with (2), whatever the t-norm \top . The t-norm only slightly emphasizes or reduces the gaps between high and low values, in the same manner as exhibited in [10].

Concretely, the new index is constructed as follows. Given a sparsity measure \mathcal{S} , it is easy to compute $\mathcal{R}_{\mathcal{S}} = \{\mathcal{S}(\mathbf{n}_1), \dots, \mathcal{S}(\mathbf{n}_r)\}$ and $\mathcal{C}_{\mathcal{S}} = \{\mathcal{S}(\mathbf{t}_{\mathbf{n}_1}), \dots, \mathcal{S}(\mathbf{t}_{\mathbf{n}_c})\}$ from the contingency matrix N crossing two partitions U and V . For each set, we propose to combine the sparsities using a suitable aggregation function \mathcal{A} , *e.g.* the arithmetic means $\overline{\mathcal{R}}_{\mathcal{S}}$ and $\overline{\mathcal{C}}_{\mathcal{S}}$, to get two representative values. Many families of aggregation functions exist, see [17] for a recent monograph. In our proposition, we restrict to functions \mathcal{A} taking values in $[0, 1]$ while computed for $\mathcal{R}_{\mathcal{S}}$ and $\mathcal{C}_{\mathcal{S}}$, so that the resulting two representatives of the sparsities in row/column of N can be inputs of a fuzzy residual implication to assess whether partitions U and V are compatible or not. Therefore, we propose a new comparison index of strict/fuzzy partitions as follows:

$$QF_{(\mathcal{S}, \mathcal{A}, \mathcal{I})}(U, V) = \min \left(\mathcal{I}(\mathcal{A}(\mathcal{R}_{\mathcal{S}}), \mathcal{A}(\mathcal{C}_{\mathcal{S}})), \mathcal{I}(\mathcal{A}(\mathcal{C}_{\mathcal{S}}), \mathcal{A}(\mathcal{R}_{\mathcal{S}})) \right). \quad (8)$$

This index is in $[0, 1]$ by construction, and it is required that \mathcal{I} satisfies the so-called ordering property², so that $QF_{(\mathcal{S}, \mathcal{A}, \mathcal{I})}(U, V) = 1$ whenever $U \equiv V$. Finally, it is worthy on note that the asymptotic complexity of this new index is $\mathcal{O}(n)$, as Anderson et al. and Campello's ones, while the other considered indices are in $\mathcal{O}(n^2)$. To sum up, the proposed comparison index satisfies properties (I1), (I2), (I3) and (I4), and has a triple of user-defined parameters $(\mathcal{S}, \mathcal{A}, \mathcal{I})$:

- a sparsity measure $\mathcal{S} : \mathbb{R} \times \dots \times \mathbb{R} \mapsto [0, 1]$, *e.g.* the Hoyer H and the kurtosis κ_4 ones given by (3) and (4),
- an aggregation function $\mathcal{A} : [0, 1] \times \dots \times [0, 1] \mapsto [0, 1]$, *e.g.* the arithmetic mean $M(\mathcal{R}) = \frac{1}{\text{card}(\mathcal{R})} \sum_{s \in \mathcal{R}} s$,
- a fuzzy residual implication $\mathcal{I} : [0, 1] \mapsto [0, 1]$, *e.g.* the Gödel \mathcal{I}_M and the Hamacher $\mathcal{I}_{H_{\gamma}}$ ones respectively given by (6) and (7).

² $\forall a, b \in [0, 1], a \leq b$ iff $\mathcal{I}(a, b) = 1$.

4 Numerical Experiments

Some of the tested comparison indices require to choose a t-norm for their computation, some others use the product. For sake of simplicity and fairness, we choose \top_P whenever a t-norm is required.

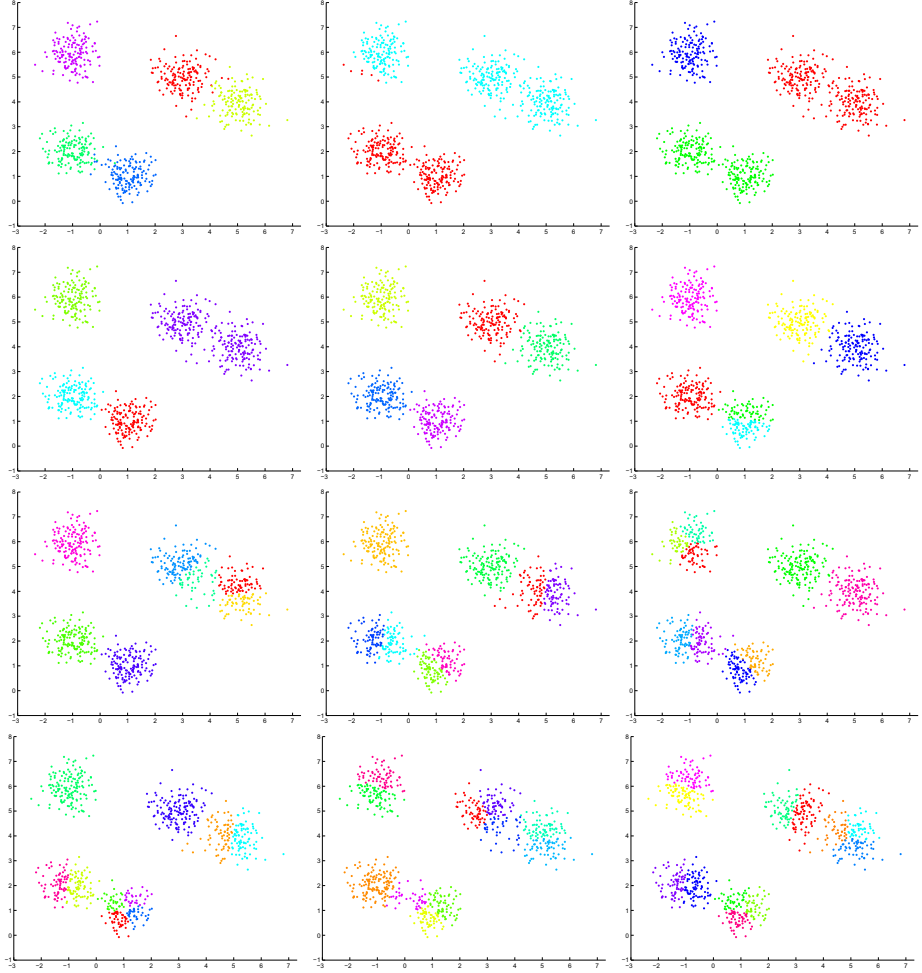


Fig. 2. From top to bottom and from left to right : ground truth clustering U_{k^*} and clusterings V_k ($k = 2, \dots, 12$) obtained with the standard k - means algorithm

4.1 Strict Partitions

First, we compare strict partitions of a 2-dimensional synthetic dataset composed of $k^* = 5$ Gaussian clusters centered at $(-1, 6)$, $(1, 1)$, $(-1, 2)$, $(5, 4)$ and $(3, 5)$ with the same standard deviation $\Sigma = \frac{1}{2}Id$, so that two pairs of them present a slight overlap, see Fig. 2 (*top-left*). Remaining subfigures present the clusterings V_k obtained with the standard k - means algorithm for $k = 2, \dots, 12$. Each

partition V_k is compared to the ground truth partition U_{k^*} using different strict indices, resulting in the curves plotted in Fig. 3 (left). The tested indices are the Rand Index RI , the Jaccard Index JI , the Transfert Distance TD , and the new sparsity based index $QF_{(S,A,I)}$ with different triples of parameters: $(\kappa_4, M, \mathcal{I}_M)$, (H, M, \mathcal{I}_M) , $(\kappa_4, M, \mathcal{I}_{H_5})$ and $(H, M, \mathcal{I}_{H_5})$. As expected, all the indices exhibit their maximum value when $k = k^* = 5$, but they do not have the same dynamics. In particular, the Rand index presents the smallest dynamics, so that we will stop using it and prefer the Jaccard indices in the remaining experiments. For $QF_{(S,A,I)}$, this property clearly depends on S . The difference of sparsity between the contingency matrices crossing U with each V_k is less marked with H because the Hoyer's sparsity measure presents more granularity than the kurtosis κ_4 in its measurement of the sparsity, as previously exhibited in Fig. 1. Another interesting point is that the proposed index better exhibit compatibility of partitions with respect to cluster refinements. Indeed, one can see that $QF_{(S,A,I)}$ considers that V_4 is closer to U_{k^*} more than the other indices do. We think it represents a slight improvement since V_4 differs with U_{k^*} about only one cluster, so that the clusters should be considered as being quite close, see cluster refinements in Fig. 2. Moreover, this behaviour is clearly reinforced with the Hamacher fuzzy implication \mathcal{I}_{H_5} , as shown for instance for V_6 and V_7 . This is because $\mathcal{I}_{H_5} > \mathcal{I}_M$ by construction. However, since $QF_{(S,M,\mathcal{I}_M)}$ and $QF_{(S,M,\mathcal{I}_{H_5})}$ mostly give analogous results for both sparsities measures H and κ_4 , the influence of the chosen fuzzy residual implication is no more studied.

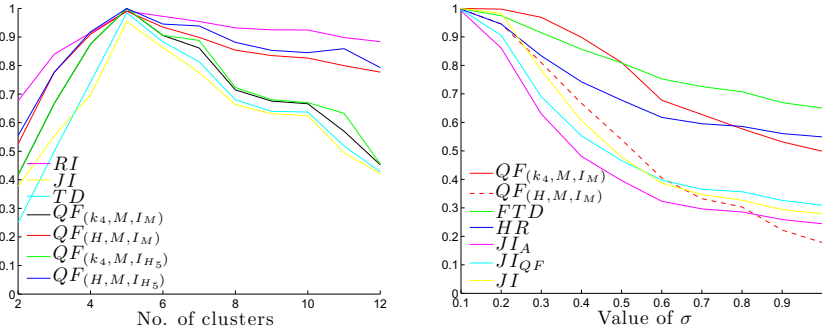


Fig. 3. Comparison of the ground truth strict partition U_{k^*} to the 11 strict partitions V_k shown in Fig. 2 (left). Comparison of the ground truth strict partition W_{k^*} to the 10 fuzzy 3-partitions V_{σ} with increasing overlap (right).

4.2 Strict vs. Fuzzy Partitions

Inspired by the work in [18], this second experiment aims at comparing a strict reference partition to a collection of fuzzy ones. Ten 3-dimensional datasets composed of $k^* = 3$ isotropic Gaussian clusters centered at $(1, 0, 0)$, $(0, 0, 0)$ and $(-1, 0, 0)$ are generated for increasing standard deviations $\sigma = \{\frac{1}{k}\}_{k=10,9,\dots,1}$, so that they evolve from no overlap to a strong one. The *Fuzzy C-Means (FCM)* algorithm is run for each dataset to produce 10 fuzzy 3-partitions V_{σ} , with a fuzzifier exponent and a termination parameter respectively set to 2 and 10^{-3} .

Fig. 3 (*right*) shows the resulting curves of comparison of those 10 partitions to the strict ground truth partition W_{k^*} for different indices. As expected, all indices achieve their maximum value for the smaller σ and decrease as σ increases. However, four of them present a higher dynamics so that their value for the most overlapping dataset is significantly lower: our fuzzy $QF_{(H,M,\mathcal{I}_M)}$, the Anderson et al. JIA , the strict³ JII , and the Quéré and Frélicot's $JIQF$. Hoyer's H sparsity measure gives a better result than κ_4 because increasing σ can be seen as increasing a kind of amount of noise within the data, and Hoyer's measure is known to overperform κ_4 in such cases, see [15].

4.3 Fuzzy Partitions : Real Datasets

The last experiment is driven for several datasets from the UCI Machine Learning Repository [19], presenting various characteristics in terms of: number n of observations, number p of attributes, true number c^* of classes/clusters and degree of overlap between clusters. Since it has lead to convergent outcomes, we only give the results obtained on the following three well-known ones:

- Fisher iris ($n = 150, p = 4, c^* = 3$, slight overlap between two classes),
- Pima diabetes ($n = 768, p = 8, c^* = 2$, strong overlap between both classes),
- Italian wine ($n = 178, p = 13, c^* = 3$, (slight overlap between 3 classes).

For each dataset, the FCM algorithm is run under the same parametrization as in the previous experiment to produce a reference fuzzy c^* -partition U_{c^*} and a collection of 14 fuzzy c -partitions U_c , c varying from 2 to 15. Each partition U_c is compared to U_{c^*} using the same indices than in the previous experiment. The resulting curves are plotted in Fig. 4 as a function of c . Unsurprisingly, each index reaches its maximum value at $c = c^*$ for each dataset and drops from its maximum toward an asymptotic value, with different dynamics. According to this criterion, our $QF_{(\kappa_4,M,\mathcal{I}_M)}$ outperforms the others (since even if HR and FTD exhibit higher values, they also present a poor dynamics). The performance of $QF_{(H,M,\mathcal{I}_M)}$ is not as good as in the previous experiment. The reason is that the number of clusters was not changing while it increases here, so that the partitions V_c frankly differ from one to another and lead $QF_{(\kappa_4,M,\mathcal{I}_M)}$ to be more discriminant thanks to the drastic behaviour of κ_4 , see Fig.1.

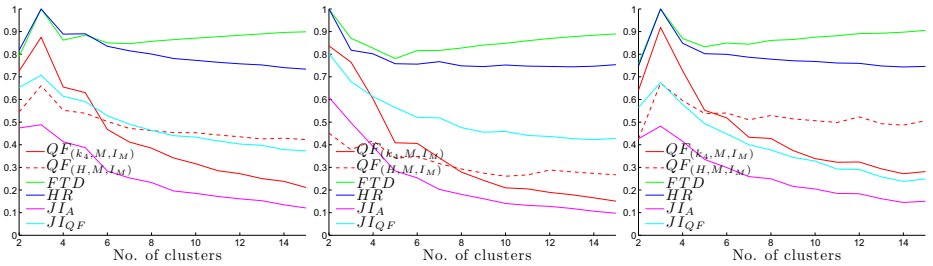


Fig. 4. Indices values obtained for the Iris, Pima and Wine datasets (from left to right)

³ For the strict Jaccard Index JII , hardened partitions obtained from fuzzy partitions V_σ are considered.

5 Conclusion

In this article, we propose a new index for comparing strict and fuzzy partitions, lying on the original idea of measuring the sparsity of the contingency matrix crossing two partitions. Its construction, involving sparsity measures and fuzzy residual implications, is simple but efficient, so that as shown by numerous experimental results, this index outperforms the existing ones, in particular with respect to the dynamics property. Moreover, by its low computational complexity, the proposed index could become a privileged tool for many practitioners.

References

- [1] Pal, N., Bezdek, J.: On cluster validity for the fuzzy c-means model. *IEEE Trans. on Fuzzy Systems* 3(3) (1995)
- [2] Wang, W., Zhang, Y.: On fuzzy cluster validity indices. *Fuzzy Sets and Systems* 158(19) (2007)
- [3] Borgelt, C.: Prototype-based classification and clustering Habilitation Thesis. Habilitation Thesis, Fakultät für Informatik der Otto von Guericke, Universität Magdeburg (2005)
- [4] Patrikainen, A.: Methods for comparing subspace clusterings. PhD thesis, Helsinki University of Technology (2005)
- [5] Hurley, N., Rickard, S.: Comparing measures of sparsity. *IEEE Trans. on Information Theory* 55(10) (2009)
- [6] Mas, M., Monserrat, M., Torrens, J., Trillas, E.: A survey on fuzzy implication functions. *IEEE Trans. on Fuzzy Systems* 15(6) (2007)
- [7] Albatineh, A., Niewiadomska-Bugaj, M., Mihalko, D.: On similarity indices and correction for chance agreement. *J. of Classification* 23 (2006)
- [8] Charon, I., Denoeud, L., Guenoche, A., Hudry, O.: Maximum transfer distance between partitions. *J. of Classification* 23(1) (2006)
- [9] Anderson, D., Bezdek, J., Popescu, M., Keller, J.: Comparing fuzzy, probabilistic, and possibilistic partitions. *IEEE Trans. on Fuzzy Systems* 18(5) (2010)
- [10] Quéré, R., Capitaine, H.L., Fraisseix, N., Frélicot, C.: On normalizing fuzzy coincidence matrices to compare fuzzy and/or possibilistic partitions with the rand index. In: 10th IEEE International Conference on Data Mining, pp. 977–982 (2010)
- [11] Hüllermeier, E., Rifqi, M.: A fuzzy variant of the rand index for comparing clustering structures. In: 13th IFSA World Congress (2009)
- [12] Campello, R.: Generalized external indexes for comparing data partitions with overlapping categories. *Pattern Recognition Letters* 31(9) (2010)
- [13] Brouwer, R.: Extending the rand, adjusted rand and jaccard indices to fuzzy partitions. *J. of Intelligent Information Systems* 32(3) (2009)
- [14] Klement, E., Mesiar, R.: Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms. Elsevier (2005)
- [15] Karvanen, J., Cichocki, A.: Measuring sparseness of noisy signals. In: 4th Int. Symp. on Independent Component Analysis and Blind Signal Separation (2003)
- [16] Le Capitaine, H., Frélicot, C.: Classification with reject options in a logical framework: a fuzzy residual implication approach. In: 13th IFSA World Congress (2009)

- [17] Grabisch, M., Marichal, J., Mesiar, R., Pap, E.: Aggregation Functions. Encyclopedia of Mathematics and its Applications. Cambridge University Press (2009)
- [18] Ceccarelli, M., Maratea, A.: A Fuzzy Extension of Some Classical Concordance Measures and an Efficient Algorithm for their Computation. In: Lovrek, I., Howlett, R.J., Jain, L.C. (eds.) KES 2008, Part III. LNCS (LNAI), vol. 5179, pp. 755–763. Springer, Heidelberg (2008)
- [19] Frank, A., Asuncion, A.: UCI machine learning repository (2010)