



A family of distances for preference–approvals

Alessandro Albano¹  · José Luis García-Lapresta² · Antonella Plaia¹ · Mariangela Sciandra¹

Accepted: 23 September 2022
© The Author(s) 2022

Abstract

A preference–approval on a set of alternatives consists of a weak order on that set and, additionally, a cut-off line that separates acceptable and unacceptable alternatives. In this paper, we propose a new method for defining the distance between preference–approvals taking into account jointly the disagreements in preferences and approvals for each pair of alternatives. The proposed distance is compared to the existing distance functions to deal with clustering problems. Specifically, we prove that our metric improves the estimated clusters in terms of both stability and accuracy.

Keywords Preferences · Approval voting · Preference–approvals · Distances · Clustering

1 Introduction

In social choice theory, preference rankings and approvals are two popular ways to collect the preferences of a group of agents on a set of alternatives. Preference rankings order the alternatives from best to worst without distinguishing between acceptable and unacceptable alternatives. That is, if a is ranked above b , we can only infer that a is preferred to b , but we cannot infer anything about their absolute acceptability. In contrast, the approval voting system (Brams & Fishburn, 1978) consists of separating the set of acceptable alternatives from the set of unacceptable alternatives without considering preferences neither over acceptable nor over unacceptable alternatives.

✉ Alessandro Albano
alessandro.albano@unipa.it

José Luis García-Lapresta
lapresta@eco.uva.es

Antonella Plaia
antonella.plaia@unipa.it

Mariangela Sciandra
mariangela.sciandra@unipa.it

¹ Department of Economics, Business and Statistics, University of Palermo, Viale delle Scienze, Edificio 13, 90129 Palermo, Sicily, Italy

² IMUVA, PRESAD Research Group, Departamento de Economía Aplicada, Universidad de Valladolid, Valladolid, Spain

Preference rankings and approval voting are related, but they are basically different types of information and cannot be inferred from each other.

In this paper, we focus on preference–approval structures. They combine preferences over the alternatives, through a weak order, and establish which alternatives are acceptable (Brams, 2008, Chapter 3; Brams & Sanver, 2009; Sanver, 2010). In preference–approval structures, voters can pay attention to which alternatives are acceptable and simultaneously rank-order them. Voters may either rank-order unacceptable alternatives or avoid declaring their preferences about them¹ by (implicitly) showing indifference between these alternatives².

The distance and correlation between two preference rankings are of particular interest within this framework. Kendall's correlation coefficient (Kendall, 1948) is likely the most well-known ranking metric. Kendall's correlation coefficient is a measure of similarity that can be linearly transformed into a measure of dissimilarity (i.e., the Kendall's distance), which counts the number of pairwise disagreements between two rankings. Emond and Mason (2002, p. 20) demonstrated that when indifference between alternatives is allowed (weak orders), Kendall's distance violates the triangle inequality. Moreover, the Kendall's correlation of the all ties ranking with any other weak order is undefined, resulting in a 0/0.

The Spearman's distance (Spearman, 1987), another famous ranking metric, between two rank vectors is calculated by taking the square root of the sum of the squared rank differences. Spearman's distance uses rank values as if they were mathematical variates, which leads to anomalous behaviour. Indeed, Spearman distance suffers from sensitivity to irrelevant alternatives (see Emond, 1997, p. 4; Emond & Mason, 2000, p. 16). In short, including additional irrelevant objects in the ranking exercise may alter the maximum agreement solution.

Kemeny and Snell (1962) took a different approach to this problem. They defined a set of additional axioms that should be applicable to any distance measure between two weak orders, and introduced a distance that satisfies all these constraints. Besides the classical properties (positivity, symmetry, identity of indiscernibles and triangle inequality), the Kemeny distance is not affected by a relabeling of alternatives (neutrality), and it is consistent in measurement as the number of objects varies (i.e. it is not sensible to irrelevant alternatives). The Kemeny distance is a city block and a geodesic distance in the permutation polytope (see Heiser, 2004). It takes the shortest path between two rankings.

Generally, extensions to ranking measures have mainly focused on the definition of weighted distances (see García-Lapresta & Pérez-Román, 2010, Albano & Plaia 2021, Plaia et al., 2021). In the last years, there has been a dramatic increase in recent publications about preference–approval structures and the introduction of consensus and distance measures in that setting.

Erdamar et al. (2014) introduced a family of distances in the preference–approval setting, and they applied them to measure the consensus in that framework. Kamwa (2019) studied the propensity of the preference–approval voting of electing the Condorcet winner/loser when they exist.

Dong et al. (2021) established some axioms implying the existence of a distinct distance function of preference–approval systems. They investigated a preferences aggregation model in the context of group decision-making based on the proposed axiomatic distance function.

Kruger and Sanver (2021) investigated the compatibility between ordinal and evaluative approaches to social choice theory under two weak assumptions: respect for unanimity and

¹ This is the case of fallback voting in Brams and Sanver (2009).

² If the number of alternatives is large, voters may have difficulties to rank-order all the alternatives (see Dummett 1984, p. 243).

independence of evaluation of each alternative. They claimed that there is an incompatibility between the two, and described some options whenever the second assumption is relaxed.

Long et al. (2021) developed a two-stage consensus reaching method for multi-attribute group decision making problems with preference–approval structures, promoting the efficiency of consensus reaching.

Barokas and Sprumont (2021) extended the classing Borda count to rank alternative in preference–approval setting, constructing an axiomatization of a new aggregation procedure called *broken Borda rule*.

In this paper, we propose a new distance for preference–approvals, following the axiomatic approach of the Kemeny distance. However, while the Kemeny distance can only consider the *preference–discordance*, our approach takes into account the *approval–discordance* as well, and use an aggregation function to combine the two types of information for each pair of alternatives.

We show that using, as an aggregation function, the family of *weighed power means* (a class of weighted quasiarithmetic means) brings the benefit of many interesting properties. The final aggregated distance will thus be the sum of the pairwise preference–approval discordances. Furthermore, we show that our distance respects the fundamental properties to be defined as a metric and that, under certain assumptions, it has a precise geometric interpretation.

Our proposal can be regarded as the generalization of the Erdamar et al. (2014) distance measure, with the two coinciding for a specific parameter setting. However, we show that the proposed distance family has some advantages over the existing one as it is more versatile and performs better in cluster analysis.

Finally, the proposed metric is used to cluster a set of preference–approvals into homogeneous groups, considering the whole 2-dimensional universe of preference–approvals and a real case study.

The paper is organized as follows. Section 2 is devoted to introduce basic notation, preference–approvals and the codifications used throughout the paper. Section 3 includes our proposal for measuring distances between preference–approvals and some results. Section 4 offers some applications to the clustering task. Finally, Sect. 5 concludes the paper with some remarks.

2 Preliminaries

Let $X = \{x_1, \dots, x_n\}$ a finite set of alternatives, with $n \geq 2$. A *weak order* (or *complete preorder*) on X is a complete and transitive binary relation on X . A *linear order* on X is an antisymmetric weak order on X . With $W(X)$ and $L(X)$ we denote the set of weak and linear orders on X , respectively. Given $R \in W(X)$, with \succ and \sim we denote the asymmetric and the symmetric parts of R , respectively: $x_i \succ x_j$ if not $x_j R x_i$, and $x_i \sim x_j$ if $x_i R x_j$ and $x_j R x_i$.

Given a set Y , with $\mathcal{P}(Y)$ we denote its power set, i.e., $I \in \mathcal{P}(Y) \Leftrightarrow I \subseteq Y$. In turn, with $\#Y$ we denote the cardinality of Y .

2.1 Preference–approval

Consider that a set of voters $V = \{v_1, \dots, v_m\}$, with $m \geq 2$, have to express their opinions over X . We assume that each voter ranks the alternatives in X by means of a weak order and,

additionally, assesses each alternative as either acceptable or unacceptable by partitioning X into A , the set of *acceptable* alternatives, and $U = X \setminus A$, the set of *unacceptable* alternatives, where A and U can be empty sets.

We also assume the following consistency condition: given two alternatives x_i and x_j , if x_j is acceptable and x_i is ranked above x_j , then x_i should be acceptable as well.

Definition 1 A *preference–approval* on X is a pair $(R, A) \in W(X) \times \mathcal{P}(X)$ satisfying the following condition:

$$\forall x_i, x_j \in X \left((x_i R x_j \text{ and } x_j \in A) \Rightarrow x_i \in A \right).$$

With $\mathcal{R}(X)$ we denote the set of preference–approvals on X .

Remark 1 If $(R, A) \in \mathcal{R}(X)$, then the following conditions are satisfied:

1. $\forall x_i, x_j \in X \left((x_i \in A \text{ and } x_j \in U) \Rightarrow x_i \succ x_j \right)$.
2. $\forall x_i, x_j \in X \left((x_i R x_j \text{ and } x_i \in U) \Rightarrow x_j \in U \right)$.

We now illustrate preference–approval structures through the following example.

Example 1 Let us consider $(R, A) \in \mathcal{R}(\{x_1, \dots, x_8\})$ represented by

$$\begin{array}{c} x_4 \\ x_1 \ x_6 \\ \hline x_2 \\ x_3 \\ x_5 \ x_7 \ x_8. \end{array}$$

It means that alternatives in the same row are indifferent, alternatives in upper rows are preferred to those located in lower rows, alternatives above the line are acceptable, i.e., $A = \{x_1, x_2, x_4, x_6\}$, and those below the line are unacceptable, i.e., $U = \{x_3, x_5, x_7, x_8\}$.

Table 1 includes the number of possible approvals, linear orders, weak orders and preference–approvals when the number of alternatives is $n = 2, 3, \dots, 10$.

It is well-known that the total number of approvals (subsets of X) and linear orders are 2^n and $n!$, respectively. The number of weak orders is $n!(\log_2 e)^{n+1}/2$ (see Good, 1980). The formula for calculating the number of preference–approvals has never been defined in the

Table 1 Number of approvals, linear orders, weak orders and preference–approvals

n	Approvals	Linear orders	Weak orders	Preference–approvals
2	4	2	3	8
3	8	6	13	44
4	16	24	75	308
5	32	120	541	2612
6	64	720	4683	25,988
7	128	5040	47,293	296,564
8	256	40,320	545,835	3,816,548
9	512	362,880	7,087,261	54,667,412
10	1024	3,628,800	102,247,563	862,440,068

literature. For the first time, the exact number of preference–approvals for $n = 2, 3, \dots, 10$ alternatives is reported herein in Table 1. The formula to compute the exact number of preference–approvals on a set of n alternatives is

$$\omega(n) = \sum_{r=0}^n (r + 1)! S_n^{(r)}, \tag{1}$$

where $S_n^{(r)}$ is a Stirling integer (number) of the second kind defined by David and Barton (1962, p. 294), Abramowitz and Stegun (1964, p. 824) and more thoroughly in Fisher and Yates (1953, p. 78), while r denotes the number of distinct positions in a weak order on n alternatives, also known as *buckets*. For example, considering four alternatives, if two are tied for first place and the other two are tied for third place, we can say that the number of distinct positions, or buckets, is two.

Table 2 shows the quotients between preference–approvals and approvals. In turn, Table 3 shows the quotients between preference–approvals and weak orders.

It is clear that the expressivity of voters explodes with preference–approvals.

Table 2 Quotients between preference–approvals and approvals

n	Quotients
2	2
3	5.5
4	19.25
5	81.62
6	406.06
7	2316.91
8	14,908.39
9	106,772.29
10	842,226.63

Table 3 Quotients between preference–approvals and weak orders

n	Quotients
2	2.67
3	3.38
4	4.11
5	4.83
6	5.55
7	6.27
8	6.99
9	7.71
10	8.43

2.2 Codifications

Assigning positions to alternatives in linear orders is trivial because indifferences among distinct alternatives are not allowed. Given $R \in L(X)$, the *position* of each alternative $x_i \in X$ in R is defined through the mapping $P_R : X \rightarrow \{1, \dots, n\}$ that assigns 1 to the alternative ranked first, 2 to the alternative ranked second, and so on.

There are different ways of assigning positions to the alternatives in weak orders. One of them is used by García-Lapresta and Pérez-Román (2011) and it is based on Smith (1973), Black (1976) and Cook and Seiford (1982).

Given $R \in W(X)$, the position of $x_i \in X$ in R is assigned through the mapping $P_R : X \rightarrow [1, n]$ defined as

$$P_R(x_i) = n - \#\{x_k \in X \mid x_i \succ x_k\} - \frac{1}{2} \cdot \#\{x_k \in X \setminus \{x_i\} \mid x_i \sim x_k\}. \quad (2)$$

Given $A \subseteq X$, the *indicator function* (or *characteristic function*) of A , $I_A : X \rightarrow \{0, 1\}$, is defined as

$$I_A(x_i) = \begin{cases} 1, & \text{if } x_i \in A, \\ 0, & \text{if } x_i \in X \setminus A. \end{cases} \quad (3)$$

Remark 2 Every preference–approval $(R, A) \in \mathcal{R}(\{x_1, \dots, x_n\})$ can be codified in terms of $P_R(x_i)$ [Eq. (2)] and $I_A(x_i)$ [Eq. (3)] as follows:

$$(P_R(x_1), P_R(x_2), \dots, P_R(x_n)) \& (I_A(x_1), I_A(x_2), \dots, I_A(x_n))). \quad (4)$$

Example 2 Consider the preference–approval $(R, A) \in \mathcal{R}(\{x_1, x_2, x_3, x_4\})$ represented by

$$\begin{array}{c} x_4 \\ x_1 \\ \underline{x_2} \\ x_3 \end{array}$$

Following Eq. (4), (R, A) is codified as $(2, 3, 4, 1) \& (1, 1, 0, 1)$.

The *sign function*, $\text{sgn} : \mathbb{R} \rightarrow \{-1, 0, 1\}$, is defined as

$$\text{sgn}(a) = \begin{cases} 1, & \text{if } a > 0, \\ 0, & \text{if } a = 0, \\ -1, & \text{if } a < 0. \end{cases}$$

Taking into account May (1952) and Fishburn (2015), we now introduce an index that codifies the order between two alternatives in a weak order $R \in W(X)$:

$$O_R(x_i, x_j) = \begin{cases} 1, & \text{if } x_i \succ x_j, \\ 0, & \text{if } x_i \sim x_j, \\ -1, & \text{if } x_j \succ x_i. \end{cases} \quad (5)$$

It is worth noting that the index $O_R(x_i, x_j)$ is also known in the literature as a_{ij} (see Kemeny & Snell, 1962, p. 11) or *score matrix* (see Emond & Mason, 2002). In this paper, to avoid confusion with the approval-discordance notation of Eq. (8), we chose to use the notation $O_R(x_i, x_j)$.

3 The proposal

Given two preference–approvals $((R_1, A_1), (R_2, A_2)) \in \mathcal{R}(X)$ and two generic alternatives $x_i, x_j \in X$, we now introduce two indices that measure the discordances between these alternatives with respect to preference and approvals, respectively.

The *preference–discordance* between x_i and x_j is defined as

$$p_{ij} = \frac{1}{2} \cdot |\operatorname{sgn}(P_{R_1}(x_j) - P_{R_1}(x_i)) - \operatorname{sgn}(P_{R_2}(x_j) - P_{R_2}(x_i))|. \quad (6)$$

Taking into account Eqs. (5), (6) can be defined in an equivalent and simpler way:

$$p_{ij} = \frac{1}{2} \cdot |O_{R_1}(x_i, x_j) - O_{R_2}(x_i, x_j)|, \quad (7)$$

and therefore, $p_{ij} \in \{0, 0.5, 1\}$.

The *approval–discordance* between x_i and x_j is defined as

$$a_{ij} = \frac{1}{2} \cdot (|I_{A_1}(x_i) - I_{A_2}(x_i)| + |I_{A_1}(x_j) - I_{A_2}(x_j)|), \quad (8)$$

and again $a_{ij} \in \{0, 0.5, 1\}$.

In both cases, the values of 0, 0.5 and 1 indicate a null, moderate and high discordance, respectively. In order to generate a global measure of discordance between two alternatives, we consider an aggregation function (see Beliakov et al., 2007; Grabisch et al., 2009; Ramík & Vlach, 2012, Sect. 2, among others).

Definition 2 Given an aggregation function $h : [0, 1] \times [0, 1] \rightarrow [0, 1]$, the distance associated with h , $D : \mathcal{R}(X) \times \mathcal{R}(X) \rightarrow [0, 1]$, is defined as

$$D((R_1, A_1), (R_2, A_2)) = \frac{2}{n \cdot (n - 1)} \cdot \sum_{\substack{i, j=1 \\ i < j}}^n h(p_{ij}, a_{ij}). \quad (9)$$

Among the huge variety of aggregation functions, in this proposal we consider a class of weighted quasiarithmetic means³: the family of *weighted power means*, $h : [0, 1] \times [0, 1] \rightarrow [0, 1]$, defined as

$$h(x, y) = (\lambda \cdot x^r + (1 - \lambda) \cdot y^r)^{\frac{1}{r}}, \quad (10)$$

where $\lambda \in [0, 1]$ and $r > 0$.

Remark 3 Weighted power means, defined in Eq. (10), have interesting properties [see, for instance, Beliakov et al. (2007, pp. 45–47)]:

1. *Continuity*: h is continuous.
2. *Monotonicity*: $(x \leq x' \text{ and } y \leq y') \Rightarrow h(x, y) \leq h(x', y')$, for all $x, y, x', y' \in [0, 1]$.
3. *Idempotency*: $h(x, x) = x$ for every $x \in [0, 1]$.
4. *Compensativeness*: $\min\{x, y\} \leq h(x, y) \leq \max\{x, y\}$ for all $x, y \in [0, 1]$.
5. *Comparability*: h is increasing in r .

³ They are defined as $h(x, y) = g^{-1}(\lambda \cdot g(x) + (1 - \lambda) \cdot g(y))$, where g is a generating function (see, for instance, Ostasiewicz & Ostasiewicz, 2000; Beliakov et al., 2007, Section 2.3).

Table 4 Values of h for $\lambda = 0.5$

(x, y)	$h(x, y)$					
	$r = 0.5$	$r = 1$	$r = 1.5$	$r = 2$	$r = 5$	$r = 10$
(0, 0)	0	0	0	0	0	0
(0, 0.5)	0.12	0.25	0.31	0.35	0.44	0.47
(0, 1)	0.25	0.50	0.63	0.71	0.87	0.93
(0.5, 0)	0.12	0.25	0.31	0.35	0.44	0.47
(0.5, 0.5)	0.50	0.50	0.50	0.50	0.50	0.50
(0.5, 1)	0.73	0.75	0.77	0.79	0.88	0.93
(1, 0)	0.25	0.50	0.63	0.71	0.87	0.93
(1, 0.5)	0.73	0.75	0.77	0.79	0.88	0.93
(1, 1)	1	1	1	1	1	1

Table 5 Values of h for $\lambda = 0.75$

(x, y)	$h(x, y)$					
	$r = 0.5$	$r = 1$	$r = 1.5$	$r = 2$	$r = 5$	$r = 10$
(0, 0)	0	0	0	0	0	0
(0, 0.5)	0.03	0.12	0.20	0.25	0.38	0.44
(0, 1)	0.06	0.25	0.40	0.50	0.76	0.87
(0.5, 0)	0.28	0.38	0.41	0.43	0.47	0.49
(0.5, 0.5)	0.50	0.50	0.50	0.50	0.50	0.50
(0.5, 1)	0.61	0.62	0.64	0.66	0.77	0.87
(1, 0)	0.56	0.75	0.83	0.87	0.94	0.97
(1, 0.5)	0.86	0.88	0.89	0.90	0.95	0.97
(1, 1)	1	1	1	1	1	1

6. *Symmetry*: $h(x, y) = h(y, x)$ for all $x, y \in [0, 1] \Leftrightarrow \lambda = 0.5$.

7. $\lim_{r \rightarrow \infty} h(x, y) = \max\{x, y\}$.

8. $\lim_{r \rightarrow 0} h(x, y) = x^\lambda \cdot y^{1-\lambda}$ (weighted geometric mean).

Notice that the inputs of h in Eq. (9) are the pairs of 0, 0.5, 1. In Tables 4 and 5 we show the values of h for these pairs and different values of the parameter r for $\lambda = 0.5, 0.75$, respectively.

According to Tables 4 and 5, the parameter r governs the penalty for each pair of values. Indeed, as r increases, so does the value of $h(p_{ij}, a_{ij})$. As a result, taking an excessively large r value results in very similar penalties and reduces the weight of high discordance compared to moderate discordance.

Taking into account Eq. (9) with the aggregation function h in Eq. (10), we now introduce the family of distances on preference–approvals that we analyze in the present paper.

Definition 3 Given $\lambda \in [0, 1]$ and $r > 0$, the distance associated with λ and r is the mapping $D_\lambda^r : \mathcal{R}(X) \times \mathcal{R}(X) \rightarrow [0, 1]$ defined as

$$D_\lambda^r((R_1, A_1), (R_2, A_2)) = \frac{2}{n \cdot (n-1)} \cdot \sum_{\substack{i, j=1 \\ i < j}}^n \left(\lambda \cdot p_{ij}^r + (1-\lambda) \cdot a_{ij}^r \right)^{\frac{1}{r}}. \quad (11)$$

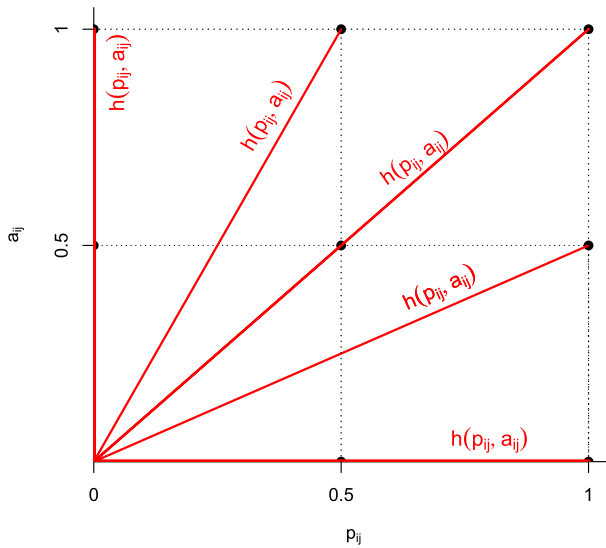


Fig. 1 Preference–approval plane

Remark 4 When $r = 2$ and $\lambda = 0.5$, the geometric interpretation of $h(p_{ij}, a_{ij})$ is related to the Euclidean distance.

Figure 1 reports the *preference–approval plane*, that is a Euclidean plane having on the x -axis the preference–discordance, p_{ij} , and on the y -axis the approval–discordance, a_{ij} .

If $r = 2$ and $\lambda = 0.5$, then $h(p_{ij}, a_{ij})$ is proportional to the Euclidean distance between (p_{ij}, a_{ij}) and the origin, $(0, 0)$, $d((p_{ij}, a_{ij}), (0, 0))$:

$$h(p_{ij}, a_{ij}) = \sqrt{0.5 \cdot (p_{ij}^2 + a_{ij}^2)} = \sqrt{0.5} \cdot d((p_{ij}, a_{ij}), (0, 0)), \text{ i.e.,}$$

$$h(p_{ij}, a_{ij}) \propto d((p_{ij}, a_{ij}), (0, 0)).$$

This means that the aggregation function h can be interpreted as a proper distance in the preference–approval plane. As a result, the point of greatest discordance, $(1, 1)$, will be the farthest from the origin of the axes. Conversely, $(0, 0)$ represents the point of greatest agreement. The red segments in Fig. 1 are proportional to the values $h(p_{ij}, a_{ij})$ for each $p_{ij}, a_{ij} \in \{0, 0.5, 1\}$.

Thus, the aggregated distance $D_{0.5}^2((R_1, A_1), (R_2, A_2))$ [see Eq. (11)] can be interpreted as the sum of $\frac{n \cdot (n-1)}{2}$ Euclidean distances in the preference–approval plane. That is,

$$D_{0.5}^2((R_1, A_1), (R_2, A_2)) = \sqrt{0.5} \cdot \sum_{\substack{i,j=1 \\ i < j}}^n d((p_{ij}, a_{ij}), (0, 0)).$$

Proposition 1 D_λ^r is a metric on $\mathcal{R}(X)$ for all $\lambda \in (0, 1)$ and $r \geq 1$. That is, for all $(R_1, A_1), (R_2, A_2) \in \mathcal{R}(X)$ the following conditions are satisfied⁴:

1. Positivity: $D_\lambda^r((R_1, A_1), (R_2, A_2)) \geq 0$.

⁴ If $0 < r < 1$, then D_λ^r reduces to a distance since the triangle inequality does not hold.

2. Symmetry: $D_\lambda^r((R_1, A_1), (R_2, A_2)) = D_\lambda^r((R_2, A_2), (R_1, A_1))$.
3. Identity of indiscernibles: $D_\lambda^r((R_1, A_1), (R_2, A_2)) = 0 \Leftrightarrow (R_1, A_1) = (R_2, A_2)$.
4. Triangle inequality: $D_\lambda^r((R_1, A_1), (R_3, A_3)) \leq D_\lambda^r((R_1, A_1), (R_2, A_2)) + D_\lambda^r((R_2, A_2), (R_3, A_3))$, for every $(R_3, A_3) \in \mathcal{R}(X)$.

Remark 5 If $\lambda \in \{0, 1\}$, then D_λ^r is not a metric.

If $\lambda = 0$, let $(R_1, A_1), (R_2, A_1) \in \mathcal{R}(X)$ be such that $R_1 \neq R_2$. Then, we have $D_\lambda^r((R_1, A_1), (R_2, A_1)) = 0$.

If $\lambda = 1$, let $(R_1, A_1), (R_1, A_2) \in \mathcal{R}(X)$ be such that $A_1 \neq A_2$. Then, we have $D_\lambda^r((R_1, A_1), (R_1, A_2)) = 0$.

Consequently, if $\lambda \in \{0, 1\}$, then D_λ^r does not verify the identity of indiscernibles, hence it is not a metric.

Proposition 2 demonstrate that our proposal can be considered as a generalization of the preference–approval distance proposed by Erdamar et al. (2014).

Given two preference–approvals $((R_1, A_1), (R_2, A_2)) \in \mathcal{R}(X)$, its distance, $d_\lambda((R_1, A_1), (R_2, A_2))$, is generated from the preference distance and the approval distance marginally, and eventually aggregate them by a convex combination.

The authors measure the disagreement between preferences by using the Kemeny metric (Kemeny, 1959), d_K :

$$d_K(R_1, R_2) = \sum_{\substack{i,j=1 \\ i < j}}^n |\text{sgn}(P_{R_1}(x_j) - P_{R_1}(x_i)) - \text{sgn}(P_{R_2}(x_j) - P_{R_2}(x_i))|.$$

Or, equivalently, by considering Eq. (5):

$$d_K(R_1, R_2) = \sum_{\substack{i,j=1 \\ i < j}}^n |O_{R_1}(x_i, x_j) - O_{R_2}(x_i, x_j)|. \quad (12)$$

Notice that $d_K(R_1, R_2) \in [0, n \cdot (n - 1)]$.

In turn, the approval disagreement is measured through the Hamming metric (Hamming, 1950), d_H :

$$d_H(A_1, A_2) = \sum_{i=1}^n |I_{A_1}(x_i) - I_{A_2}(x_i)|. \quad (13)$$

Notice that $d_H(A_1, A_2) \in [0, n]$.

In order to aggregate d_K and d_H as a global distance, the two metrics are normalized to the same codomain $[0, 1]$ via dividing by their maximum distances.

The mappings $d_R : \mathcal{R}(X) \times \mathcal{R}(X) \rightarrow [0, 1]$ and $d_A : \mathcal{R}(X) \times \mathcal{R}(X) \rightarrow [0, 1]$ are defined as

$$d_R((R_1, A_1), (R_2, A_2)) = \frac{d_K(R_1, R_2)}{n \cdot (n - 1)},$$

$$d_A((R_1, A_1), (R_2, A_2)) = \frac{d_H(A_1, A_2)}{n}.$$

The two normalized distances are eventually aggregated in a final preference–approval distance, $d_\lambda : \mathcal{R}(X) \times \mathcal{R}(X) \rightarrow [0, 1]$, defined as

$$d_\lambda((R_1, A_1), (R_2, A_2)) = \lambda \cdot d_R((R_1, A_1), (R_2, A_2)) + (1 - \lambda) \cdot d_A((R_1, A_1), (R_2, A_2)), \tag{14}$$

where $\lambda \in [0, 1]$ is a parameter used to control the relative relevance of the two components.

Taking into account Eqs. (12) and (13), (14) can be re-written as

$$d_\lambda((R_1, A_1), (R_2, A_2)) = \frac{\lambda}{n \cdot (n - 1)} \cdot \sum_{\substack{i,j=1 \\ i < j}}^n |O_{R_1}(x_i, x_j) - O_{R_2}(x_i, x_j)| + \frac{1 - \lambda}{n} \cdot \sum_{i=1}^n |I_{A_1}(x_i) - I_{A_2}(x_i)|. \tag{15}$$

Proposition 2 For all $(R_1, A_1), (R_2, A_2) \in \mathcal{R}(X)$ and $\lambda \in [0, 1]$ it holds

$$D_\lambda^1((R_1, A_1), (R_2, A_2)) = d_\lambda((R_1, A_1), (R_2, A_2)).$$

Note that Proposition 2 is valid for weighted power means. They are the proper weighted quasiarithmetic means that allow us to generalize the distance between preference–approvals introduced by Erdamar et al. (2014).

In Proposition 2, we have shown that $D_\lambda^r = d_\lambda$ when $r = 1$. We now show that is not true if $r \neq 1$.

Proposition 3 If $r \neq 1$, $D_\lambda^r((R_1, A_1), (R_2, A_2)) = d_\lambda((R_1, A_1), (R_2, A_2))$ for all $(R_1, A_1), (R_2, A_2) \in \mathcal{R}(X)$ and $\lambda \in [0, 1]$ is not true.

Proof Let us consider the case of two alternatives. Notice that in Eq. (11), when $n = 2$, D_λ^r reduces to the h function computed in $i = 1$ and $j = 2$. That is, $D_\lambda^r((R_1, A_1), (R_2, A_2)) = h(p_{12}, a_{12}) = (\lambda \cdot p_{12}^r + (1 - \lambda) \cdot a_{12}^r)^{\frac{1}{r}}$. By Proposition 2, we have $D_\lambda^1((R_1, A_1), (R_2, A_2)) = d_\lambda((R_1, A_1), (R_2, A_2)) = \lambda \cdot p_{12} + (1 - \lambda) \cdot a_{12}$.

If we force the equality $D_\lambda^1((R_1, A_1), (R_2, A_2)) = D_\lambda^r((R_1, A_1), (R_2, A_2))$, we have $\lambda \cdot p_{12} + (1 - \lambda) \cdot a_{12} = (\lambda \cdot p_{12}^r + (1 - \lambda) \cdot a_{12}^r)^{\frac{1}{r}}$, i.e.,

$$(\lambda \cdot p_{12} + (1 - \lambda) \cdot a_{12})^r = \lambda \cdot p_{12}^r + (1 - \lambda) \cdot a_{12}^r. \tag{16}$$

We have to prove that there exist $p_{12}, a_{12} \in \{0, 0.5, 1\}$ and $\lambda \in [0, 1]$ such that Eq. (16) is not true for any $r \neq 1$.

If $p_{12} = 1$ and $a_{12} = 0$, then Eq. (16) becomes $\lambda^r = \lambda$, and it is true if and only if $\lambda \in \{0, 1\}$. In all the other cases, if $r \neq 1$, then Eq. (16) is false. \square

4 Clustering tasks

This section shows how the proposed distance can be used to study the universe of preference–approvals and to determine clusters.

Section 4.1 examines the universe of preference approvals in the case of two alternatives in order to observe how the values of r and λ affect the creation of homogeneous clusters. Afterwards, the influence of the two parameters r and λ when the number of alternatives n varies is investigated.

Section 4.2 provides an application on real data, to investigate how the countries of the European Union can be clustered into groups, according to their preference–approvals on nine alternatives concerning social values. The dataset used comes from the Eurobarometer website⁵.

4.1 Universe of preference–approvals

Let us consider the 2-dimensional preference–approval universe where the set of alternatives is $X = \{x_1, x_2\}$. Following Eq. (4), the preference–approvals (R_i, A_i) , $i = 1, 2, \dots, 8$, are represented by two 2-dimensional vectors:

$$(2, 1) \& (1, 1) \equiv \frac{x_2}{x_1} \quad (2, 1) \& (0, 1) \equiv \frac{x_2}{x_1}$$

$$(2, 1) \& (0, 0) \equiv \frac{x_2}{x_1} \quad (1, 2) \& (1, 1) \equiv \frac{x_1}{x_2}$$

$$(1, 2) \& (1, 0) \equiv \frac{x_1}{x_2} \quad (1, 2) \& (0, 0) \equiv \frac{x_1}{x_2}$$

$$(1.5, 1.5) \& (1, 1) \equiv \frac{x_1}{x_2} \quad (1.5, 1.5) \& (0, 0) \equiv \frac{x_1}{x_2}$$

The distances between preference–approvals on two alternatives for $r = 1$ and $\lambda = 0.5$ (Fig. 2) and $\lambda = 0.75$ (Fig. 3) are reported in the heatmaps.

Increasing the value of λ emphasizes the discordance in the preference part, and modifies the relationships between the corresponding preference–approvals. Indeed, when $\lambda = 0.75$, there is an increase in the intensity of the distances at the top-right hand side of the graph, which concerns the triples

$$(2, 1) \& (1, 1), (2, 1) \& (0, 1), (2, 1) \& (0, 0)$$

and

$$(1, 2) \& (0, 0), (1, 2) \& (1, 1), (1, 2) \& (1, 1).$$

The hierarchical relationship between objects is reported in Fig. 4; the dendrograms show how the hierarchical clustering of the eight preference–approvals changes based on D'_λ .

Figure 4 shows that the value of λ strongly influences the hierarchical aggregation of preference–approvals.

A similar analysis can be carried out by varying the value of r . In Fig. 5 the distances between the corresponding preference–approvals, for $r = 2$ and $\lambda = 0.5$ are shown.

⁵ <https://europa.eu/eurobarometer/screen/home>.

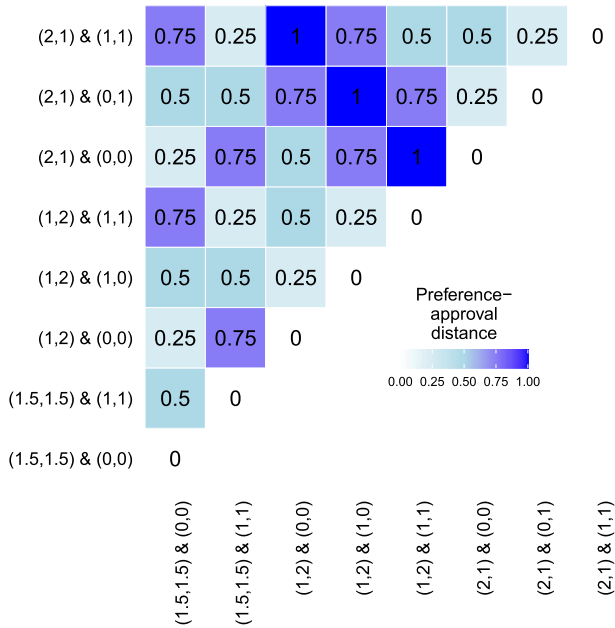


Fig. 2 Distances between preference–approvals for 2 alternatives, $r = 1$ and $\lambda = 0.5$

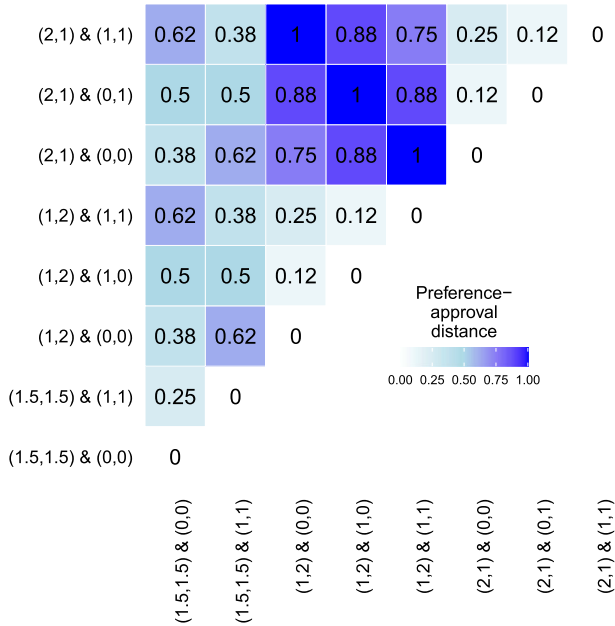


Fig. 3 Distances between preference–approvals for 2 alternatives, $r = 1$ and $\lambda = 0.75$

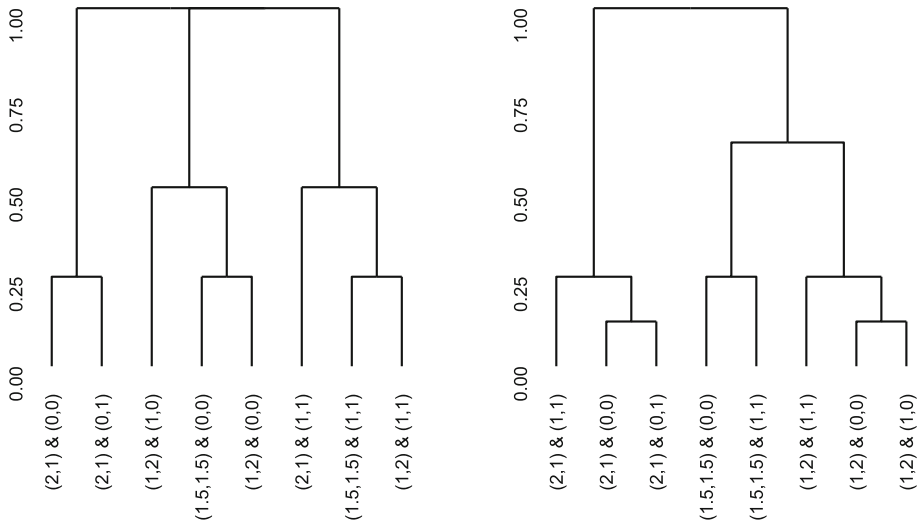


Fig. 4 Hierarchical clustering dendrogram for 2 alternatives, $r = 1$, $\lambda = 0.5$ (left) and $\lambda = 0.75$ (right)

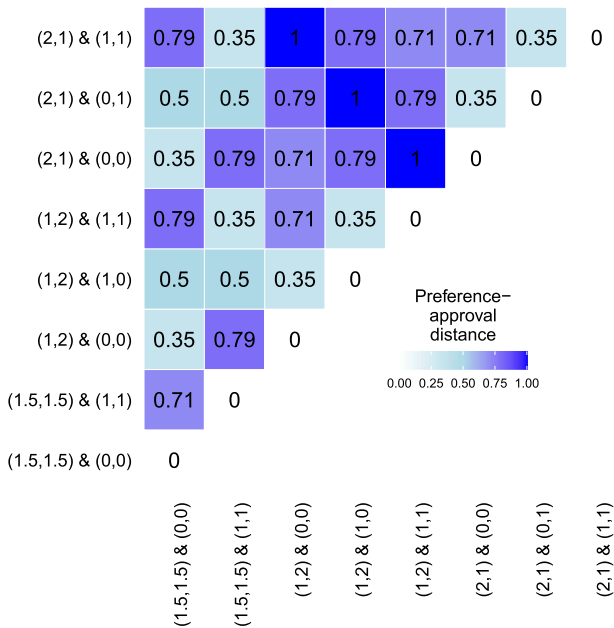


Fig. 5 Distance between preference–approvals for 2 alternatives, $r = 2$, $\lambda = 0.5$

Compared to Figs. 2 and 5 shows a general increase of distances determined by the increase of r . In particular,

$$D_{\lambda}^2((R_1, A_1), (R_2, A_2)) \geq D_{\lambda}^1((R_1, A_1), (R_2, A_2)),$$

for all $(R_1, A_1), (R_2, A_2) \in \mathcal{R}(X)$. This is due to h being increasing in r .

The dendrograms between preference–approvals objects are reported in Fig. 6.

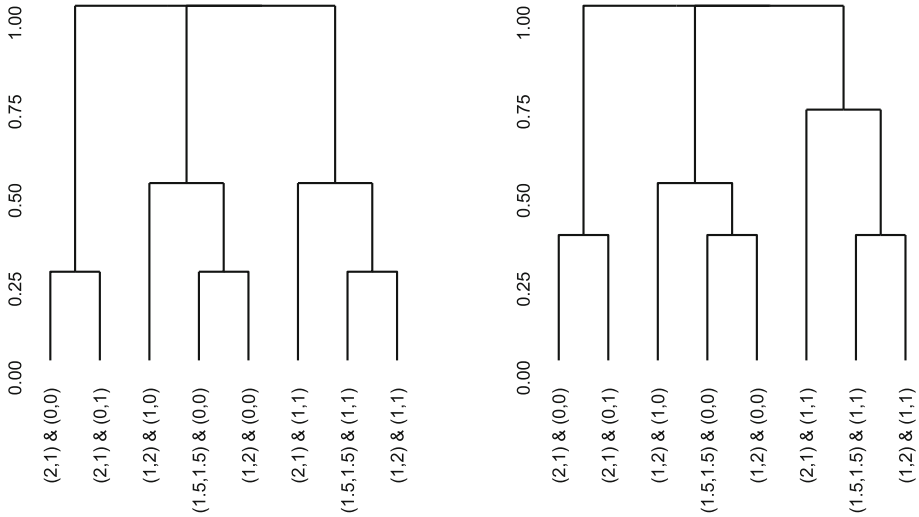


Fig. 6 Hierarchical clustering dendrogram for 2 alternatives, $r = 1$ (left), $r = 2$ (right) and $\lambda = 0.5$

Table 6 Cophenetic dendrogram correlations for $n = 2$, $r = 1, 1.5, 2, 5, 10$ and $\lambda = 0.5$

$n = 2$	1	1.5	2	5	10
1	1				
1.5	0.99	1			
2	0.98	1	1		
5	0.94	0.97	0.98	1	
10	0.90	0.94	0.97	1	1

Figure 6 shows that an increase in r contributes differently (with respect to an increase in λ) to the change of the hierarchical aggregation structure. In fact, the two dendrograms merge preference–approvals in the same way. What changes is the “height” at which there is the aggregation or, in other words, the distance to be tolerated to aggregate two preference–approvals. Note that this happens only for two alternatives.

Tables 6, 7, 8 and 9 show the cophenetic correlation coefficient⁶ (see Sokal & Rohlf, 1962; Schlee, 1973, pp. 278–284) between dendrograms, for $n = 2, 3, 4, 5$ and $\lambda = 0.5$. The cophenetic coefficient was computed in R using the dendextend package (Galili, 2015).

Tables 6, 7, 8 and 9 show that dendrogram correlations are strictly related to the values of r and n . Overall, the correlations between dendrograms tend to decrease as r increases. This is especially evident when we examine the first column of each table, which reports the correlation between dendrograms obtained with $r = 1$ and dendrograms obtained with $r = 1.5, 2, 5, 10$. In terms of the number of alternatives, it should be noted that as n increases, the dendrogram correlations generally decrease with an oscillatory trend.

In other words, Tables 6, 7, 8 and 9 highlight that the parameter r has a considerable influence, not only on the resulting values of the proposed distance D'_λ , but also on the

⁶ The cophenetic correlation coefficient is a measure of similarity between dendrograms. It is particularly used in biostatistics to investigate how faithfully a dendrogram preserves the pairwise distances between the original unmodeled data points, or also to study where raw data tends to occur in clumps or clusters. This coefficient has also been proposed as a nested cluster test (see Rohlf & Fisher, 1968; Saraçlı et al., 2013).

Table 7 Cophenetic dendrogram correlations for $n = 3$, $r = 1, 1.5, 2, 5, 10$ and $\lambda = 0.5$

$n = 3$	1	1.5	2	5	10
1	1				
1.5	0.76	1			
2	0.76	1	1		
5	0.63	0.68	0.70	1	
10	0.65	0.71	0.73	0.99	1

Table 8 Cophenetic dendrogram correlations for $n = 4$, $r = 1, 1.5, 2, 5, 10$ and $\lambda = 0.5$

$n = 4$	1	1.5	2	5	10
1	1				
1.5	0.85	1			
2	0.80	0.83	1		
5	0.71	0.76	0.88	1	
10	0.70	0.75	0.86	0.99	1

Table 9 Cophenetic dendrogram correlations for $n = 5$, $r = 1, 1.5, 2, 5, 10$ and $\lambda = 0.5$

$n = 5$	1	1.5	2	5	10
1	1				
1.5	0.80	1			
2	0.72	0.79	1		
5	0.61	0.73	0.81	1	
10	0.57	0.69	0.79	0.95	1

cluster structure discovered among the observations of the preference–approvals universe. Specifically, as n increases and the expressiveness of the voters explodes (Table 1), so does the discriminating power of r , allowing different clustering structures to be highlighted. Indeed, the proposed family of distances D_{λ}^r is more flexible than the existing one, and it ultimately comes down to a new parameter that can be exploited in various applications, such as maximizing the goodness of a clustering procedure.

To explore further this issue, let us consider a simulation study on the universe of 5 alternatives, which involves three steps:

- generate four groups of clustered preference–approvals;
- apply a hierarchical clustering algorithm for different values of r .
- compute an external validation index, the Adjusted Rand index (Hubert & Arabie, 1985), to investigate which value of r maximises the similarity between the estimated and the theoretical clusters.

Therefore, we aim to find the value of r that provides more reliable clusters, i.e. clusters that are more consistent with the data-generating process.

The number of preference–approvals (on five alternatives) generated within each cluster was determined by randomly drawing four values from a normal distribution $\mathcal{N}(50, 4)$ and converting them into integer numbers.

Orderings and approvals were generated individually and merged to produce the final set of preference–approvals. Specifically, orderings within each sub-partition were generated from a Mallows Model (Mallows, 1957), which was one of the earliest probability models

Table 10 Cluster central permutations

Cluster k	Central permutation σ_0
1	(4.5, 2, 4.5, 3, 1)
2	(1, 3, 4, 2, 5)
3	(4, 3, 1.5, 1.5, 5)
4	(3, 5, 2, 4, 1)

Table 11 Multinomial probability vectors

Cluster k	Approved alternatives i					
	0	1	2	3	4	5
1	0.10	0.40	0.25	0.15	0.05	0.05
2	0.10	0.10	0.30	0.30	0.15	0.05
3	0	0.05	0.10	0.25	0.35	0.25
4	0.35	0.30	0.15	0.10	0.05	0.05

suggested for rankings and it is still widely used in theoretical and applied research. It is an exponential model defined by a central permutation σ_0 and a dispersion parameter θ . When $\theta \neq 0$, σ_0 represents the mode of the distribution, i.e., the permutation with the highest probability of being generated. The probability of any other ranking decays exponentially with increasing distance to the central permutation. The dispersion parameter θ controls the steepness of this decline. The θ values for our simulation studies are $\{0, 0.5, 1, 1.5, 2\}$. Assuming that σ is a generic ranking, the probability for this ranking is function of θ , and it is given by:

$$\Pr(\theta) = \frac{\exp(-\theta d(\sigma, \sigma_0^{-1}))}{\psi(\theta)}, \tag{17}$$

where d is a ranking distance measure and $\psi(\theta)$ is a normalization constant.

We generated rankings assuming the Kemeny distance d_K . The cluster central permutations, σ_0 , used in the analysis are reported in Table 10.

Approvals, within each cluster are generated from four multinomial distributions, with probability vectors, p_{ik} , described in Table 11. Specifically, p_{ik} is the probability to draw i approved alternatives into the k -th cluster.

After deriving clusters, the adjusted Rand index (Hubert & Arabie, 1985) is used to assess their goodness. The adjusted Rand index is a measure of the similarity between two set of clusterings; it is the corrected-for-chance version of the Rand index (Rand, 1971). The correction uses the predicted similarity of all pair-wise comparisons between clusterings described by a random model to generate a baseline. Although the Rand Index can only provide values between 0 and +1 (0 when the two data clusterings do not agree on any pair of points, and 1 when data clusterings are exactly the same), the modified Rand Index can return negative values if the index is lower than the expected similarity of all pair-wise comparisons between clusterings specified by a random model.

The results (Table 12) are obtained by averaging the adjusted Rand index over ten randomly generated datasets for each value of θ .

Table 12 shows that, except for the case $\theta = 1.5$, our measure D_λ^r with $r \neq 1$ results in higher average adjusted Rand indices. Thus, $r \neq 1$ allows the true clustered structure of data to be found more accurately and provides more accurate clusters.

Table 12 Average adjusted Rand index over r and θ

		θ				
		0	0.5	1	1.5	2
r	1	0.093	0.267	0.591	0.847	0.801
	0.5	0.114	0.249	0.435	0.692	0.822
	2	0.080	0.320	0.611	0.665	0.602
	3	0.090	0.274	0.568	0.607	0.559
	4	0.082	0.313	0.569	0.588	0.530
	5	0.081	0.301	0.558	0.553	0.502
	7	0.089	0.273	0.534	0.542	0.502
	10	0.074	0.276	0.528	0.540	0.502

The bold highlights the highest values of the average adjusted Rand, column by column. That is, they indicate the best value of r conditionally to the dispersion parameter

4.2 A real data application

This subsection shows how the proposed metric can be used to perform cluster analysis on real data retrieved from the Eurobarometer website.

Since 1973, Eurobarometer has undertaken a series of public opinion polls on behalf of the European Commission and other European Union (EU) institutions. These polls cover a wide range of topics concerning the EU and its member countries. The data utilized in these analyses are specifically from question Q5 of the poll titled “Defending Democracy, Empowering citizens. Public Opinion at the legislature’s midpoint”⁷.

A group of voters, divided by countries, was asked to indicate which of the following values should the European Parliament defend as a matter of priority:

- x_1 : Equality between women and men.
- x_2 : The fight against discrimination and for the protection of minorities.
- x_3 : Tolerance and respect for diversity in society.
- x_4 : Solidarity between EU Member States and between its regions.
- x_5 : Solidarity between the EU and poor countries in the world.
- x_6 : The protection of human rights in the EU and worldwide.
- x_7 : Freedom of religion and belief.
- x_8 : Freedom of movement.
- x_9 : Freedom of speech and thought.

As a result, data are stored in a table (see Table 14) with 27 rows (one row for each EU member country) and 9 columns (each column representing an alternative of $X = \{x_1, \dots, x_9\}$). The total number of votes cast by the i -th country in favor of the j -th alternative is shown in the table’s generic cell ij .

In order to transform the original table into a set of preference–approvals, preferences and approvals need to be derived. For each country, the alternatives are ranked in order of popularity, beginning with the one that received the most votes and ending with the one that received the fewest. Furthermore, in order to generate a vector of approvals, those alternatives that received more votes than the national average were deemed acceptable.

For example, in Table 13 we show the votes expressed in France (the votes of all countries are included in Table 14).

⁷ <https://europa.eu/eurobarometer/surveys/detail/2612>.

Table 13 Votes in France

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
France	37	17	20	12	13	20	12	12	32

Since the votes' average is 19.44, the votes in France are transformed into a preference–approval codification [see Eq. (4)] as

$$(1, 5, 3.5, 8, 6, 3.5, 8, 8, 2) \& (1, 0, 1, 0, 0, 1, 0, 0, 1)$$

that can be visualized as follows:

$$\begin{array}{c} x_1 \\ x_9 \\ \hline x_3 \ x_6 \\ x_2 \\ x_5 \\ x_4 \ x_7 \ x_8. \end{array}$$

To run the cluster analysis, the distance matrix 27×27 was constructed using Eq. (11). All the alternatives seem important in this example, so a distinction between acceptable and unacceptable alternatives should not be interpreted as a distinction between valuable and not valuable, but instead as a distinction between more and less urgent. For this reason, $\lambda = 0.75$ was chosen in order to emphasize preference differences more than approvals.

A cluster-wise measure of cluster stability (Hennig, 2007) is used to jointly discover the optimal value of r and the optimal number of clusters k . Stability refers to the property of a meaningful and valid cluster that does not change easily when the data set is perturbed in a non-essential way. That is, when applied to many datasets collected from the same data distribution, a reliable clustering method should produce similar partitions. The cluster stability method (Hennig, 2007) employs three steps:

1. use various strategies to resample new data sets from the original and apply the hierarchical clustering method to each of them;
2. for every given original cluster, find the most similar cluster using the Jaccard coefficient (Jaccard, 1901) in the new data set and record the similarity value;
3. assess the cluster stability of every single cluster by the mean similarity taken over the resampled data sets.

The average cluster-wise stability is shown in Fig. 7 as a function of r (for $k = 2, 3, 4$ clusters). The procedure suggests that the most stable cluster configuration is $k = 2$ and $r = 2$. It is worth noting that, regardless of the value of k , $r > 1$ always leads to improved cluster stability. Indeed, with two clusters ($k = 2$) the value of r that maximizes stability is $r = 2$. Whereas with three or four clusters, the optimal solution is $r = 4$. In addition, as the number of clusters k increases, the average stability decreases.

For several reasons, stability is a particularly relevant cluster validation measure in this example for determining the best value of r . First, it is not possible to use external validation measures in this case as the true clustered structure of the EU countries is not known. At the same time, most internal validation measures employ the distance between observations (D'_λ) to assess the goodness of clusters. However, this may be an issue in our instance since the distance between observations (D'_λ) is influenced by r . Therefore, to determine which value of r yields more accurate clusters, a metric that is independent of r is desirable. Furthermore, cluster stability has been examined both theoretically and practically (Hennig, 2007; Von

Table 14 Votes in the EU

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
Belgium	31	15	16	23	12	24	10	14	34
Bulgaria	11	7	14	24	10	23	11	32	21
Czech Republic	12	6	10	22	8	28	4	23	25
Denmark	18	14	17	13	10	30	7	12	27
Germany	15	11	18	21	9	32	6	8	28
Estonia	11	12	14	14	5	20	8	29	24
Ireland	27	19	15	14	9	28	11	27	24
Greece	9	12	8	34	19	31	8	14	35
Spain	35	15	18	15	15	20	4	12	25
France	37	17	20	12	13	20	12	12	32
Croatia	14	14	14	20	16	25	12	28	28
Italy	25	17	14	21	11	20	9	21	29
Cyprus	25	14	5	24	21	37	10	11	24
Latvia	5	14	11	33	5	35	4	20	26
Lithuania	10	12	17	17	8	33	7	18	29
Luxembourg	23	19	17	18	13	19	7	16	22
Hungary	15	17	15	19	11	28	12	20	21
Malta	21	17	15	16	13	28	10	15	18
Netherlands	25	18	25	18	9	34	12	6	31
Austria	24	16	19	19	12	23	9	19	30
Poland	15	15	13	18	11	19	12	24	19
Portugal	32	22	16	30	20	27	7	5	17
Romania	14	12	12	20	16	24	14	28	22
Slovenia	15	8	23	19	9	32	5	24	31
Slovakia	19	10	10	20	9	21	15	35	28
Finland	15	14	15	14	6	30	7	17	26
Sweden	33	12	19	12	11	39	5	12	30

Luxburg, 2010; Ullmann et al., 2022), and it has been shown to be capable of distinguishing between meaningful stable and spurious clusters.

Figures 8 and 9 show the resulting dendrogram and clusters, respectively, obtained with $k = 2$ and $r = 2$.

The clustering procedure suggests that the EU countries can be separated into two large groups. Cluster 1 is mainly made up of Western European countries, whereas Cluster 2 of Eastern European countries.

To provide a more in-depth picture of how the EU countries express their views on the nine alternatives proposed, the two preference–approvals that represent the two clusters, that we call *representative preference–approvals*, are shown in Eq. (18).

To obtain the representative preference–approvals that summarize each cluster, preferences and approvals need to be aggregated. In each cluster, the set of preferences is combined into a unique weak order by deriving the average position for each alternative and ranking them according to it. Note that this aggregation method is equivalent to the the Borda count (Borda, 1781) extended to weak orders (see Smith, 1973; Black, 1976, Cook & Seiford, 1982).

Bootstrap Evaluation of Clusters

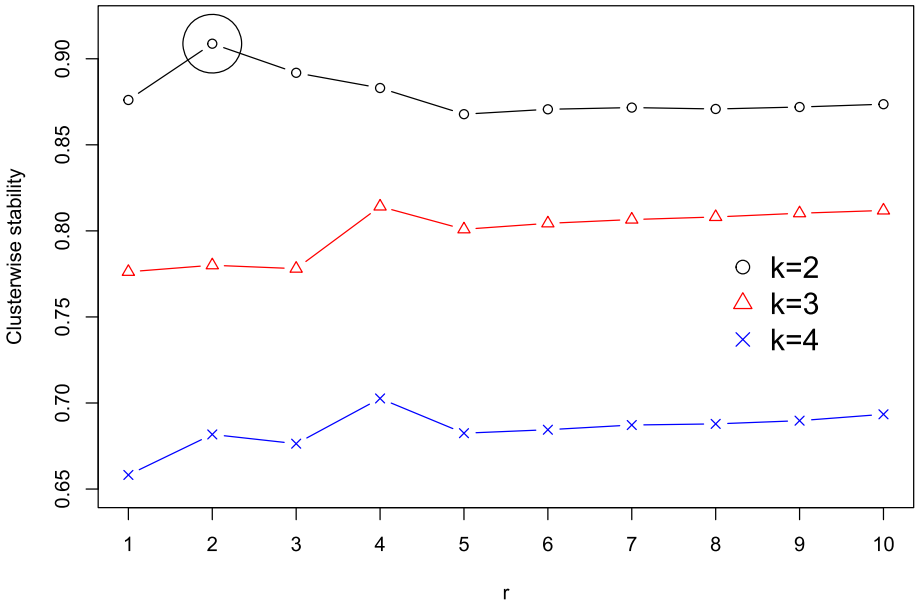
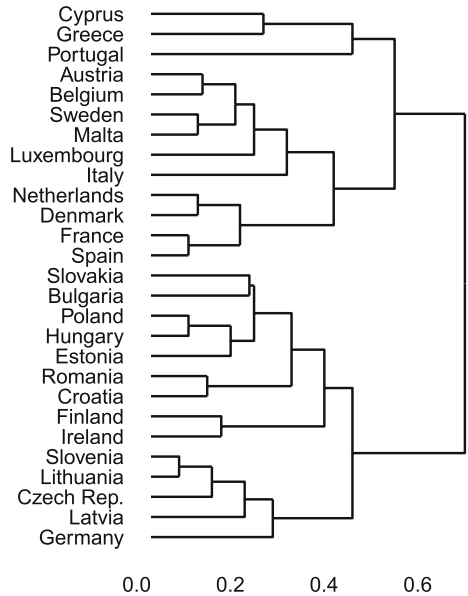


Fig. 7 Average cluster-wise stability over r

Fig. 8 EU cluster dendrogram



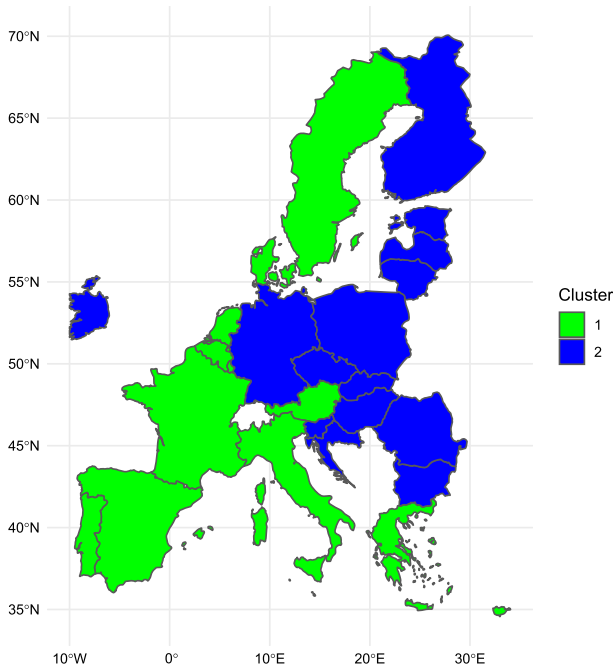


Fig. 9 Map of EU voters with clusters

In our example, the extended Borda count assigns a score to each alternative, for each country, the number of alternatives ranked below plus half of the number of alternatives that are indifferent to it:

$$B_R(x_i) = \#\{x_k \in X \mid x_i \succ x_k\} + \frac{1}{2} \cdot \#\{x_k \in X \setminus \{x_i\} \mid x_i \sim x_k\}.$$

Similarly, the set of approvals are combined into a unique approval vector by taking the average approval for each alternative, and then considering those alternatives whose average approval is greater than 0.5 as approved.

Cluster 1	Cluster 2	
x_1	x_6	
x_9	x_9	
<hr style="width: 50%; margin: 0;"/>	x_8	
x_4	x_4	
x_2	<hr style="width: 50%; margin: 0;"/>	
x_3	x_1 x_3	
x_8	x_2	
x_5	x_5	
x_7	x_7	(18)

It is worth noting that x_6 and x_9 , namely, “The protection of human rights in the EU and worldwide” and “Freedom of speech and thought”, respectively, are above the approval line in the two representative preference–approvals, indicating that they can be considered very urgent. Regarding x_1 , that is “Equality between women and men”, it is ranked at the top of the representative preference–approval of Cluster 1, while it is just below the approval line in

Table 15 Distance between countries and representative cluster preference–approvals

Country	Cluster 1	Cluster 2	Cluster assignment
Belgium	0.092	0.246	1
Bulgaria	0.487	0.192	2
Czech Rep.	0.342	0.036	2
Denmark	0.176	0.332	1
Germany	0.286	0.204	2
Estonia	0.355	0.195	2
Ireland	0.267	0.265	2
Greece	0.388	0.318	1
Spain	0.164	0.419	1
France	0.219	0.440	1
Croatia	0.387	0.156	2
Italy	0.200	0.225	1
Cyprus	0.281	0.370	1
Latvia	0.341	0.120	2
Lithuania	0.400	0.094	2
Luxembourg	0.144	0.359	1
Hungary	0.343	0.072	2
Malta	0.096	0.301	1
Netherlands	0.243	0.377	1
Austria	0.096	0.262	1
Poland	0.349	0.132	2
Portugal	0.367	0.537	1
Romania	0.463	0.204	2
Slovenia	0.415	0.105	2
Slovakia	0.354	0.219	2
Finland	0.320	0.123	2
Sweden	0.144	0.274	1

The bold values are the smallest distances between each country and the two clusters

the Cluster 2 representative preference–approval. Similarly, x_4 , that is “Solidarity between the EU Member States and between its regions”, is ranked fourth (above the approval line) in Cluster 2. Still, it is the first alternative below the approval line in Cluster 1. Furthermore, Cluster 2 prioritizes x_8 , that is “Freedom of movement”, which is at the end of the preference–approval of Cluster 1. Finally, in both the two representative preference–approvals, x_7 , that is “Freedom of religion and belief”, is ranked last.

Table 15 reports the $D_{0.75}^2$ distances of each country to the representative cluster preference–approvals.

It should be noted that, except for Greece, each country is closer to the preference–approval of its own cluster than the other. Despite being reasonable, this result is not trivial since the technique for obtaining the cluster preference–approval does not involve D_{λ}^x .

Some countries can be considered central in their clusters as they are very close to the representative preference–approval, e.g. Belgium (0.092), Austria (0.096), Malta (0.096) for Cluster 1, and the Czech Republic (0.036), Lithuania (0.094), Hungary (0.072) and Slovenia

(0.105) for Cluster 2. As a rule of thumb, the greater the distance from the own cluster preference–approval, the more the country disagrees with the other countries in its cluster. Finally, it is worth noting that some countries, such as Ireland, Italy and Greece, are located in the middle of the two clusters, as they have a similar distances to the two cluster preference–approvals.

5 Concluding remarks

In social choice theory, preference rankings and approvals are two popular ways to collect the preferences of a group of agents on a set of alternatives. In the preference–approval setting, each agent, in addition to ordering a set of alternatives from best to worst, submits a cut-off line to distinguish between acceptable and unacceptable. Within this framework, in this paper, we propose a new distance for preference–approvals, following the approach of the Kemeny distance.

Given two preference–approvals and two alternatives, we introduce two indices that measure the discordances between these alternatives with respect to preference and approvals, and an aggregation function belonging to the class of weighted power means to define a new distance. This new distance depends on two parameters. The effect of these parameters on the distance is analyzed and described through some heatmaps. The proposed distance can be used to study the universe of preference–approvals and to determine clusters of voters: how the two parameters characterizing the distances affect the clustering process is shown with some dendrograms and by the cophenetic correlations among them. We have shown that the new distance family offers some advantages compared to the existing distance function. Specifically, through a simulation study and the adjusted Rand index, we have proved that D_{λ}^r with $r \neq 1$ allows the true clustered structure of data to be found more accurately. Similarly, through a cluster-wise stability index, we have shown that D_{λ}^r with $r \neq 1$ produces more stable clusters on the real data example.

In future work, axiomatizing the new family of distance functions might prove important.

Moreover, future research should examine consensus measures based on distances between preference–approvals (see Erdamar et al., 2014), algorithms to determine representative preference–approvals efficiently (see D’Ambrosio, 2017), clustering on alternatives (see González del Pozo et al., 2017), and also reaching consensus processes (see Palomares et al., 2014; García-Lapresta & Pérez-Román, 2017; Chao et al., 2021, among others).

Acknowledgements The authors are grateful to two anonymous reviewers for their useful comments and suggestions, and also to the Spanish Agencia Estatal de Investigación (project PID2021-122506NB-I00) and the University of Palermo (projects: FFR_D16_PLAIA and FFR_D16_SCIANDRA) for their financial support.

Funding Open access funding provided by Università degli Studi di Palermo within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

A Appendix: Proofs

A.1 Proof of Proposition 1

1. Positivity holds since $h(p_{ij}, a_{ij}) \geq 0$ for all $i, j \in \{1, \dots, n\}$.
2. Symmetry holds since $p_{ij} = p_{ji}$ [see Eq. (7)] and $a_{ij} = a_{ji}$ [see Eq. (8)] for all $i, j \in \{1, \dots, n\}$.
3. Identity of indiscernibles: Obviously, $D'_\lambda((R_1, A_1), (R_1, A_1)) = 0$. If $D'_\lambda((R_1, A_1), (R_2, A_2)) = 0$, then $(\lambda \cdot p_{ij}^r + (1 - \lambda) \cdot a_{ij}^r)^{\frac{1}{r}} = 0$ for all $i, j \in \{1, \dots, n\}$. Since $p_{ij}, a_{ij} \geq 0$ and $\lambda \in (0, 1)$, we have $p_{ij} = a_{ij} = 0$ for all $i, j \in \{1, \dots, n\}$. Then, $O_{R_1}(x_i, x_j) = O_{R_2}(x_i, x_j)$, $I_{A_1}(x_i) = I_{A_1}(x_i)$ and $I_{A_1}(x_j) = I_{A_2}(x_j)$ for all $i, j \in \{1, \dots, n\}$. Consequently, $(R_1, A_1) = (R_2, A_2)$.
4. Triangle inequality: If we define

$$\begin{aligned}
 p'_{ij} &= \frac{1}{2} \cdot |O_{R_1}(x_i, x_j) - O_{R_2}(x_i, x_j)|, \\
 p''_{ij} &= \frac{1}{2} \cdot |O_{R_2}(x_i, x_j) - O_{R_3}(x_i, x_j)|, \\
 p'''_{ij} &= \frac{1}{2} \cdot |O_{R_1}(x_i, x_j) - O_{R_3}(x_i, x_j)|,
 \end{aligned}$$

then, we have

$$p'''_{ij} \leq p'_{ij} + p''_{ij}. \tag{19}$$

Similarly, if we define

$$\begin{aligned}
 a'_{ij} &= \frac{1}{2} \cdot (|I_{A_1}(x_i) - I_{A_2}(x_i)| + |I_{A_1}(x_j) - I_{A_2}(x_j)|), \\
 a''_{ij} &= \frac{1}{2} \cdot (|I_{A_2}(x_i) - I_{A_3}(x_i)| + |I_{A_2}(x_j) - I_{A_3}(x_j)|), \\
 a'''_{ij} &= \frac{1}{2} \cdot (|I_{A_1}(x_i) - I_{A_3}(x_i)| + |I_{A_1}(x_j) - I_{A_3}(x_j)|),
 \end{aligned}$$

then, we have

$$a'''_{ij} \leq a'_{ij} + a''_{ij}. \tag{20}$$

From Eqs. (19) and (20) it follows

$$p'''_{ij} + a'''_{ij} \leq p'_{ij} + p''_{ij} + a'_{ij} + a''_{ij}. \tag{21}$$

To prove the triangle inequality we need to show

$$h(p'''_{ij}, a'''_{ij}) \leq h(p'_{ij}, a'_{ij}) + h(p''_{ij}, a''_{ij}),$$

i.e.,

$$\begin{aligned}
 &(\lambda \cdot (p'''_{ij})^r + (1 - \lambda) \cdot (a'''_{ij})^r)^{\frac{1}{r}} \leq \\
 &(\lambda \cdot (p'_{ij})^r + (1 - \lambda) \cdot (a'_{ij})^r)^{\frac{1}{r}} + (\lambda \cdot (p''_{ij})^r + (1 - \lambda) \cdot (a''_{ij})^r)^{\frac{1}{r}}
 \end{aligned} \tag{22}$$

Raising the two members of the inequality by r , Eq. (22) is equivalent to

$$\lambda \cdot (p'''_{ij})^r + (1 - \lambda) \cdot (a'''_{ij})^r \leq \tag{23}$$

$$\left((\lambda \cdot (p'_{ij})^r + (1 - \lambda) \cdot (a'_{ij})^r)^{\frac{1}{r}} + (\lambda \cdot (p''_{ij})^r + (1 - \lambda) \cdot (a''_{ij})^r)^{\frac{1}{r}} \right)^r.$$

Taking into account that for all $a, b \geq 0$ and $r \geq 1$, (see Hardy et al., 1952, p. 32 for more details) it holds:

$$(a + b)^r \geq a^r + b^r,$$

we have

$$\begin{aligned} & \left((\lambda \cdot (p'_{ij})^r + (1 - \lambda) \cdot (a'_{ij})^r)^{\frac{1}{r}} + (\lambda \cdot (p''_{ij})^r + (1 - \lambda) \cdot (a''_{ij})^r)^{\frac{1}{r}} \right)^r \geq \\ & (\lambda \cdot (p'_{ij})^r + (1 - \lambda) \cdot (a'_{ij})^r) + (\lambda \cdot (p''_{ij})^r + (1 - \lambda) \cdot (a''_{ij})^r) = \\ & \lambda \cdot ((p'_{ij})^r + (p''_{ij})^r) + (1 - \lambda) \cdot ((a'_{ij})^r + (a''_{ij})^r). \end{aligned} \quad (24)$$

Because of Eqs. (19) and (20), we have

$$\begin{aligned} & \lambda \cdot ((p'_{ij})^r + (p''_{ij})^r) + (1 - \lambda) \cdot ((a'_{ij})^r + (a''_{ij})^r) \geq \\ & \lambda \cdot (p'''_{ij})^r + (1 - \lambda) \cdot (a'''_{ij})^r. \end{aligned} \quad (25)$$

Therefore, following Eqs. (24) and (25), we can write:

$$\begin{aligned} & \lambda \cdot (p'''_{ij})^r + (1 - \lambda) \cdot (a'''_{ij})^r \leq \\ & \lambda \cdot ((p'_{ij})^r + (p''_{ij})^r) + (1 - \lambda) \cdot ((a'_{ij})^r + (a''_{ij})^r) \leq \\ & \left((\lambda \cdot (p'_{ij})^r + (1 - \lambda) \cdot (a'_{ij})^r)^{\frac{1}{r}} + (\lambda \cdot (p''_{ij})^r + (1 - \lambda) \cdot (a''_{ij})^r)^{\frac{1}{r}} \right)^r. \end{aligned}$$

for all $i, j \in \{1, \dots, n\}$.

Hence,

$$D_{\lambda}^r((R_1, A_1), (R_3, A_3)) \leq D_{\lambda}^r((R_1, A_1), (R_2, A_2)) + D_{\lambda}^r((R_2, A_2), (R_3, A_3))$$

for all $(R_1, A_1), (R_2, A_2), (R_3, A_3) \in \mathcal{R}(X)$.

A.2 Proof of Proposition 2

The first distance can be expressed in the following way:

$$\begin{aligned} D_{\lambda}^1((R_1, A_1), (R_2, A_2)) &= \frac{2}{n \cdot (n - 1)} \cdot \sum_{\substack{i, j=1 \\ i < j}}^n h(p_{ij}, a_{ij}) \\ &= \frac{2}{n \cdot (n - 1)} \cdot \sum_{\substack{i, j=1 \\ i < j}}^n (\lambda \cdot p_{ij}^1 + (1 - \lambda) \cdot a_{ij}^1) \\ &= \frac{2 \cdot \lambda}{n \cdot (n - 1)} \cdot \sum_{\substack{i, j=1 \\ i < j}}^n \frac{|O_{R_1}(x_i, x_j) - O_{R_2}(x_i, x_j)|}{2} \\ &\quad + \frac{2 \cdot (1 - \lambda)}{n \cdot (n - 1)} \cdot \sum_{\substack{i, j=1 \\ i < j}}^n \frac{|I_{A_1}(x_i) - I_{A_2}(x_i)| + |I_{A_1}(x_j) - I_{A_2}(x_j)|}{2}. \end{aligned}$$

Taking into account Eq. (15), the equality between $D_\lambda^1((R_1, A_1), (R_2, A_2))$ and $d_\lambda((R_1, A_1), (R_2, A_2))$ holds if and only if

$$\begin{aligned} & \frac{1 - \lambda}{n \cdot (n - 1)} \cdot \sum_{\substack{i,j=1 \\ i < j}}^n |I_{A_1}(x_i) - I_{A_2}(x_i)| + |I_{A_1}(x_j) - I_{A_2}(x_j)| \\ &= \frac{1 - \lambda}{n} \cdot \sum_{i=1}^n |I_{A_1}(x_i) - I_{A_2}(x_i)|. \end{aligned} \tag{26}$$

Let us define $I_i = |I_{A_1}(x_i) - I_{A_2}(x_i)|$. Then,

- $\sum_{\substack{i,j=1 \\ i < j}}^n (|I_{A_1}(x_i) - I_{A_2}(x_i)| + |I_{A_1}(x_j) - I_{A_2}(x_j)|) = \sum_{\substack{i,j=1 \\ i < j}}^n (I_i + I_j),$
- $\sum_{i=1}^n |I_{A_1}(x_i) - I_{A_2}(x_i)| = \sum_{i=1}^n I_i.$

Therefore, the equality in Eq. (26) can be re-written as:

$$\begin{aligned} D_\lambda^1((R_1, A_1), (R_2, A_2)) &= d_\lambda((R_1, A_1), (R_2, A_2)) \Leftrightarrow \\ & \frac{1 - \lambda}{n \cdot (n - 1)} \cdot \sum_{\substack{i,j=1 \\ i < j}}^n (I_i + I_j) = \frac{1 - \lambda}{n} \cdot \sum_{i=1}^n I_i. \end{aligned} \tag{27}$$

To prove Eq. (27):

$$\begin{aligned} & \frac{1 - \lambda}{n \cdot (n - 1)} \cdot \sum_{\substack{i,j=1 \\ i < j}}^n (I_i + I_j) \\ &= \frac{1 - \lambda}{n \cdot (n - 1)} \cdot (I_1 + I_2 + I_1 + I_3 + \dots + I_2 + I_3 + \dots + I_{n-1} + I_n) \\ &= \frac{1 - \lambda}{n \cdot (n - 1)} (n - 1) \cdot (I_1 + I_2 + \dots + I_n) = \frac{1 - \lambda}{n} \cdot \sum_{i=1}^n I_i. \end{aligned}$$

The equality Eq. (27) is a necessary and sufficient condition to show that $D_\lambda^1 = d_\lambda$, for every $\lambda \in [0, 1]$.

References

Abramowitz, M., & Stegun, I. A. (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables* (Vol. 55). US Government Printing Office.

Albano, A., & Plaia, A. (2021). Element weighted Kemeny distance for ranking data. *Electronic Journal of Applied Statistical Analysis*, 14(1), 117–145.

Barokas, G., & Sprumont, Y. (2021). The broken Borda rule and other refinements of approval ranking. *Social Choice and Welfare*, 58(1), 187–199.

Beliakov, G., Pradera, A., & Calvo, T. (2007). *Aggregation functions: A guide for practitioners*. New York: Springer.

Black, D. (1976). Partial justification of the Borda count. *Public Choice*, 28, 1–15.

Borda, J.d. (1781). Mémoire sur les élections au scrutin: Histoire de l'académie royale des sciences. *Paris, France*, 12.

- Brams, S. J. (2008). Mathematics and democracy: Designing better voting and fair-division procedures. *Mathematical and Computer Modelling*, 48(9), 1666–1670.
- Brams, S. J., & Fishburn, P. C. (1978). Approval voting. *American Political Science Review*, 72(3), 831–847.
- Brams, S. J., & Sanver, M. R. (2009). Voting systems that combine approval and preference. In S. J. Brams, W. V. Gehrlein, & F. S. Roberts (Eds.), *The mathematics of preference, choice and order: Essays in honor of Peter C. Fishburn* (pp. 215–237). New York: Springer.
- Chao, X., Dong, Y., Kou, G., & Peng, Y. (2021). How to determine the consensus threshold in group decision making: a method based on efficiency benchmark using benefit and cost insight. *Annals of Operations Research*, 316, 1–35.
- Cook, W. D., & Seiford, L. M. (1982). On the Borda–Kendall consensus method for priority ranking problems. *Management Science*, 28(6), 621–637.
- D'Ambrosio, A., Mazzeo, G., Iorio, C., & Siciliano, R. (2017). A differential evolution algorithm for finding the median ranking under the Kemeny axiomatic approach. *Computers & Operations Research*, 82, 126–138.
- David, F. N., & Barton, D. E. (1962). *Combinatorial chance*. New York: Hafner.
- Dong, Y., Li, Y., He, Y., & Chen, X. (2021). Preference–approval structures in group decision making: Axiomatic distance and aggregation. *Decision Analysis*, 18(4), 273–295.
- Dummett, M. (1984). *Voting procedures*. Oxford: Oxford University Press.
- Emond, E. J. (1997). Maximum rank correlation as a solution concept in the m rankings problem with application to multi criteria decision analysis. In *DOR (CAM) Research Note RN 9705*.
- Emond, E. J., & Mason, D. W. (2000). *A new technique for high level decision support*. Operational Research Division: Department of National Defence Canada.
- Emond, E. J., & Mason, D. W. (2002). A new rank correlation coefficient with application to the consensus ranking problem. *Journal of Multi-Criteria Decision Analysis*, 11(1), 17–28.
- Erdamar, B., García-Lapresta, J. L., Pérez-Román, D., & Sanver, M. R. (2014). Measuring consensus in a preference-approval context. *Information Fusion*, 17, 14–21.
- Fishburn, P. C. (2015). *The theory of social choice*. Princeton: Princeton University Press.
- Fisher, R. A., & Yates, F. (1953). *Statistical tables for biological, agricultural and medical research*. Hafner Publishing Company.
- Galili, T. (2015). Dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics*, 31(22), 3718–3720.
- García-Lapresta, J. L., & Pérez-Román, D. (2010). Consensus measures generated by weighted Kemeny distances on weak orders. In *In 2010 10th international conference on intelligent systems design and applications* (pp. 463–468.) IEEE.
- García-Lapresta, J. L., & Pérez-Román, D. (2011). Measuring consensus in weak orders. In E. Herrera-Viedma, J. García-Lapresta, J. Kacprzyk, H. Nurmi, M. Fedrizzi, & S. Zadrozny (Eds.), *Consensual Processes* (pp. 213–234). New York: Springer.
- García-Lapresta, J. L., & Pérez-Román, D. (2017). A consensus reaching process in the context of non-uniform ordered qualitative scales. *Fuzzy Optimization and Decision Making*, 16(4), 449–461.
- González del Pozo, R., García-Lapresta, J. L., & Pérez-Román, D. (2017). *Clustering us 2016 presidential candidates through linguistic appraisals, Advances in Fuzzy Logic and Technology 2017* (pp. 143–153). Springer.
- Good, I. (1980). The number of orderings of n candidates when ties and omissions are both allowed. *Journal of Statistical Computation and Simulation*, 10(2), 159–159.
- Grabisch, M., Marichal, J. L., Mesiar, R., & Pap, E. (2009). *Aggregation functions* (Vol. 127). Cambridge: Cambridge University Press.
- Hamming, R. W. (1950). Error detecting and error correcting codes. *The Bell System Technical Journal*, 29(2), 147–160.
- Hardy, G. H., Littlewood, J. E., Pólya, G., & Pólya, G. (1952). *Inequalities*. Cambridge: Cambridge University Press.
- Heiser, W. J. (2004). Geometric representation of association between categories. *Psychometrika*, 69(4), 513–545.
- Hennig, C. (2007). Cluster-wise assessment of cluster stability. *Computational Statistics & Data Analysis*, 52(1), 258–271.
- Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of Classification*, 2(1), 193–218.
- Jaccard, P. (1901). Distribution de la flore alpine dans le bassin des dranges et dans quelques regions voisines. *Bulletin de la Société Vaudoise des Sciences Naturelles* 241–272.
- Kamwa, E. (2019). Condorcet efficiency of the preference approval voting and the probability of selecting the Condorcet loser. *Theory and Decision*, 87(3), 299–320.
- Kemeny, J. G. (1959). Mathematics without numbers. *Daedalus*, 88(4), 577–591.

- Kemeny, J. G., & Snell, J. (1962). *Mathematical models in the social sciences*. Blaisdall Publishing Company.
- Kendall, M. G. (1948). *Rank Correlation Methods*. Griffin.
- Kruger, J., & Sanver, M. R. (2021). An Arrovian impossibility in combining ranking and evaluation. *Social Choice and Welfare*, 57(3), 535–555.
- Long, J., Liang, H., Gao, L., Guo, Z., & Dong, Y. (2021). Consensus reaching with two-stage minimum adjustments in multi-attribute group decision making: A method based on preference-approval structure and prospect theory. *Computers & Industrial Engineering*, 158, 107349.
- Mallows, C. L. (1957). Non-null ranking models. I. *Biometrika*, 44(1/2), 114–130.
- May, K. O. (1952). A set of independent necessary and sufficient conditions for simple majority decision. *Econometrica*, 20, 680–684.
- Ostasiewicz, S., & Ostasiewicz, W. (2000). Means and their applications. *Annals of Operations Research*, 97(1), 337–355.
- Palomares, I., Estrella, F. J., Martínez, L., & Herrera, F. (2014). Consensus under a fuzzy context: Taxonomy, analysis framework AFRYCA and experimental case of study. *Information Fusion*, 20, 252–271.
- Plaia, A., Buscemi, S., & Sciandra, M. (2021). Consensus among preference rankings: A new weighted correlation coefficient for linear and weak orderings. *Advances in Data Analysis and Classification*, 15(4), 1015–1037.
- Ramík, J., & Vlach, M. (2012). Aggregation functions and generalized convexity in fuzzy optimization and decision making. *Annals of Operations Research*, 195(1), 261–276.
- Rand, W. M. (1971). Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, 66(336), 846–850.
- Rohlf, F. J., & Fisher, D. R. (1968). Tests for hierarchical structure in random data sets. *Systematic Biology*, 17(4), 407–412.
- Sanver, M. R. (2010). Approval as an intrinsic part of preference. In J. F. Laslier & M. R. Sanver (Eds.), *Handbook on Approval Voting, Studies in Choice and Welfare* (pp. 469–481). New York: Springer.
- Saraçlı, S., Doğan, N., & Doğan, İ. (2013). Comparison of hierarchical cluster analysis methods by cophenetic correlation. *Journal of Inequalities and Applications*, 2013(1), 1–8.
- Schlee, D. (1973). *Numerical Taxonomy. The Principles and Practice of Numerical Classification*. San Francisco: Freeman.
- Smith, J. H. (1973). Aggregation of preferences with variable electorate. *Econometrica*, 41(6), 1027–1041.
- Sokal, R. R., & Rohlf, F. J. (1962). The comparison of dendrograms by objective methods. *Taxon*, 11(2), 33–40.
- Spearman, C. (1987). The proof and measurement of association between two things. *The American Journal of Psychology*, 100(3/4), 441–471.
- Ullmann, T., Hennig, C., & Boulesteix, A.L. (2022). Validation of cluster analysis results on validation data: A systematic framework. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 12, e1444.
- Von Luxburg, U. (2010). Clustering stability: An overview. *Foundations and Trends in Machine Learning*, 2(3), 235–274.