Annals of Mathematics and Artificial Intelligence
https://doi.org/10.1007/s10472-024-09924-y

Weighted and Choquet L^p distance representation of comparative dissimilarity relations on fuzzy description profiles

Giulianella Coletti[1] · **Davide Petturiti**[2] · **Bernadette Bouchon-Meunier**[3]

Accepted: 8 January 2024

Abstract
We consider comparative dissimilarity relations on pairs on fuzzy description profiles, the latter providing a fuzzy set-based representation of pairs of objects. Such a relation expresses the idea of "no more dissimilar than" and is used by a decision maker when performing a case-based decision task under vague information. We first limit ourselves to those relations admitting a weighted L^p distance representation, for which we provide an axiomatic characterization in case the relation is complete, transitive and defined on the entire space of pairs of fuzzy description profiles. Next, we switch to the more general class of comparative dissimilarity relations representable by a Choquet L^p distance, parameterized by a completely alternating normalized capacity.

Keywords Dissimilarity relation · Fuzzy description profiles · Weighted L^p distance · Choquet L^p distance

Mathematics Subject Classification (2010) 91B08 · 68T37 · 03E72

1 Introduction

The large availability of data and the increasing pervasiveness of Artificial Intelligence through machine learning and deep learning techniques, made case-based decision paradigms more and more common in recent years. In this context, similarity and dissimilarity mea-

✉ Davide Petturiti
davide.petturiti@unipg.it

Giulianella Coletti
giulianella.coletti@unipg.it

Bernadette Bouchon-Meunier
bernadette.bouchon-meunier@lip6.fr

1 Dip. Matematica e Informatica, University of Perugia, Perugia, Italy

2 Dip. Economia, University of Perugia, Perugia, Italy

3 LIP6, Sorbonne Université-CNRS, Paris, France

Published online: 24 January 2024

sures are important to evaluate a degree of resemblance between two or more objects. A plethora of similarity/dissimilarity measures are available in the literature and the choice of one of them is done each time two images, cases, objects, situations, texts or data must be compared. Indeed, similarity/dissimilarity measures enter into play each time case-based decision-making needs to be performed.

We are interested in measuring the similarity/dissimilarity of general objects characterized by a profile formed by a finite number of attributes or features. Then any object is identified by a vector which is binary (if the features can only belong or not to the object) or, as it has been more recently preferred, with elements in [0, 1] (if a partial degree of membership is accepted). In this setting, one can simply compare the resulting *fuzzy description profiles*, rather than the objects themselves. For that, many papers on this subject present in the literature (see for instance [15, 19, 25]) discuss the opportunity of considering as dissimilarity measure a pseudo-distance function [4] or a measure of comparison, as studied in [6] generalizing Tversky's contrast model [26] (for a general parameterized form, see [15]). Usually the comparison is made for a particular environment and "a posteriori" (on the basis of the obtained results), focusing on one or more specific properties.

With the purpose of providing conscious reasons to use a particular similarity measure on the basis of the semantics behind this choice, in [7–9] two classes of similarity measures have been studied by using the paradigm of measurement theory. These classes are very large and contain as particular cases almost all the known measures in the sense of Tversky's contrast model and its generalizations. The study starts from the concept of comparative similarity which is a binary relation between pairs of objects expressing the primitive idea of "no more similar than" and provides the conditions that this relation needs to respect, when choosing a measure of this class.

The aim of this paper is to make an equivalent study for some classes of dissimilarity measures that cannot be derived from similarity measures analyzed in [7–9]. Here, we focus on those dissimilarity measures that do not depend on the specific values taken by the degrees of membership of each feature of a pair of objects but only on the feature-wise distances between them, by analyzing the *comparative dissimilarity* relation they represent. This class contains many distances like the *weighted L^p distances* and the *Choquet L^p distances* with respect to a completely alternating normalized capacity. The reason for studying the comparative dissimilarity relation is to make explicit what are the qualitative conditions an agent tacitly accepts when he/she chooses a dissimilarity measure of these types.

It turns out that, though the quoted class of dissimilarity measures is characterized by a set of axioms easy to justify from a behavioral point of view, the class is too wide. This is why we first restrict ourselves to the distinguished subclass of weighted L^p distances for $p \in [1, +\infty)$, depending on a vector of parameters. These parameters express the importance, with respect to dissimilarity, assigned to the various features that describe the objects and are indirectly assessed through the comparative dissimilarity.

We provide a complete characterization for every $p \in [1, +\infty)$, by relying on a *rationality* condition that acts on finite subsets of comparisons. Such condition turns out to be necessary and sufficient to obtain a weighted L^p distance representation, provided the set of comparisons is finite. Nevertheless, rationality, alone, is not sufficient to get a weighted L^p distance representation in case the set of comparisons is infinite: rationality actually implies a weaker condition equivalent to almost representability under non-triviality. We show that, if we add a suitable axiom of equivalence involving constant fuzzy description profiles, then rationality becomes necessary and sufficient to get a unique weighted L^p distance representation. The characterization we give generalizes some preliminary results involving the case $p = 1$, that appeared in [10].

Next, still for $p \in [1, +\infty)$, we provide an analogous characterization of dissimilarity relations on fuzzy description profiles that are representable by a Choquet L^p distance, with respect to a completely alternating normalized capacity. This set of distances contains the weighted L^p distances and permits to model interactions among features. As a byproduct, we also get a characterization of dissimilarity relations representable by the unweighted L^∞ distance.

The paper is structured as follows. Section 2 introduces the necessary preliminaries on comparative dissimilarity relations and their numerical representations. Section 3 presents the axioms a comparative dissimilarity relation is asked to satisfy in order to have a numerical representation that only depends on feature-wise distances between fuzzy description profiles. Section 4 characterizes the comparative dissimilarity relations that are representable by a weighted L^p distance, while Section 5 considers those that are representable by a Choquet L^p distance with respect to a completely alternating normalized capacity. Next, Section 6 provides a discussion about introduced axioms, their logical grouping and their purpose in an elicitation task. Finally, Section 7 gathers our conclusions.

2 Preliminaries

Let $\mathcal{H} = \{h_1, \dots, h_m\} = \{h_k\}_{k \in I}$ be a set of $m \geq 2$ attributes (also referred to as features), indexed by the set $I = \{1, \dots, m\}$, each of which is present in an object with a degree of membership $\mu_k(\cdot) \in [0, 1]$.

Let $\mathcal{Y} = [0, 1]^m$ be the set of all *fuzzy description profiles*: objects are identified through vectors $X = (x_1, \dots, x_m) \in \mathcal{Y}$, where $x_k \in [0, 1]$ expresses the degree of membership of attribute k in the considered object. In other words, fuzzy description profiles in $\mathcal{Y}$ can be regarded as membership functions of fuzzy subsets of the set $\mathcal{H}$ of m attributes.

Since the attributes can be expressed by a vague characterization, we can regard each of them as a fuzzy subset of a corresponding hidden variable. So each $X \in \mathcal{Y}$ is a projection of the Cartesian product of m possibly fuzzy subsets of m variables. For instance if the attributes h_1 and h_2 represent a person as "old" and "fat", every $X = (x_1, x_2)$ is a projection of the Cartesian product of the fuzzy sets "old" and "fat" of variables "age" and "weight", both taking values in $\mathbb{R}$. Therefore, in our setting we have two types of fuzzy sets: *(i)* single attributes are seen as fuzzy subsets of the reference set related to the corresponding variable; *(ii)* object description profiles are fuzzy subsets of $\mathcal{H}$, formed by the evaluations of the attribute memberships on the object values for each related variable.

Given two fuzzy description profiles $X, Y \in \mathcal{Y}$, seen as fuzzy subsets of $\mathcal{H}$, we adopt as fuzzy inclusion the classic concept introduced by Zadeh [27]:

$$X \subseteq Y \text{ if and only if } x_k \leq y_k, \text{ for all } k \in I. \tag{1}$$

In other words, we have that $X \subseteq Y$ if and only if every attribute is no more present in X than it is in Y. So, in what follows the relation $\subseteq$ on $\mathcal{Y}$ is identified with the partial order relation $\leq$ on $\mathcal{Y}$, where the inequality is component-wise. In the rest of the paper, we also write $X < Y$ if and only if $x_k < y_k$, for all $k \in I$.

We denote by $\mathcal{X} \subset \mathcal{Y}$ the set of *crisp description profiles*, i.e., $\mathcal{X} = \{0, 1\}^m$, and for any $X \in \mathcal{Y}$, we consider the *support* $s_X = \{k \in I : x_k > 0\}$, so, in particular, $\underline{0}$ is the fuzzy description profile with $s_X = \emptyset$. More generally, if $\varepsilon \in [0, 1]$, then $\underline{\varepsilon}$ denotes the element of $\mathcal{Y}$ whose components are all equal to ε.

For every $0 \leq \delta \leq x_k$ and $0 \leq \eta \leq 1 - x_k$ we denote by $x_k^{-\delta}$ the value $x_k - \delta$, and by x_k^{η} the value $x_k + \eta$, and consider the elements of $\mathcal{Y}$: $X_k^{-\delta} = (x_1, \ldots, x_k^{-\delta}, \ldots, x_m)$, and $X_k^{\eta} = (x_1, \ldots, x_k^{\eta}, \ldots, x_m)$.

Given $X, Y \in \mathcal{Y}$, we denote by $|X - Y|$ the element of $\mathcal{Y}$ whose k-th component is $|x_k - y_k|$, by X^c the element of $\mathcal{Y}$, whose k-th component is $1 - x_k$, which is referred to as the *complement* of X. For $p \in [1, +\infty)$, we denote by X^p the element of $\mathcal{Y}$ whose k-th component is x_k^p and by $|X - Y|^p$ the element of $\mathcal{Y}$ whose k-th component is $|x_k - y_k|^p$.

Let us now consider a *comparative dissimilarity* that is a binary relation $\precsim$ on $\mathcal{Y}^2$, with the following meaning: for all $(X, Y), (X', Y') \in \mathcal{Y}^2$,

$$(X, Y) \precsim (X', Y') \Longleftrightarrow \text{“}X \text{ is no more dissimilar to } Y \text{ than } X' \text{ is dissimilar to} Y'\text{”}. \tag{2}$$

The relations $\sim$ and $\prec$ are then induced by $\precsim$ in the usual way: $(X, Y) \sim (X', Y')$ stands for $(X, Y) \precsim (X', Y')$ and $(X', Y') \precsim (X, Y)$, while $(X, Y) \prec (X', Y')$ stands for $(X, Y) \precsim (X', Y')$ and not $(X', Y') \precsim (X, Y)$.

In a case-based decision task, the decision maker has a knowledge base $\mathcal{K} = \{X_1, \ldots, X_n\}$ of prototypical fuzzy description profiles. Thus, given a new fuzzy description profile X, the relation $\precsim$ is normally used to find the less dissimilar prototype X_{i^*}, that satisfies

$$(X, X_{i^*}) \precsim (X, X_i), \quad \text{for all } i \neq i^*.$$

The relation $\precsim$ on $\mathcal{Y}^2$ is said to be:

complete: if $(X, Y) \precsim (X', Y')$ or $(X', Y') \precsim (X, Y)$, for all $(X, Y), (X', Y') \in \mathcal{Y}^2$;
transitive: if $(X, Y) \precsim (X', Y')$ and $(X', Y') \precsim (X'', Y'')$ implies $(X, Y) \precsim (X'', Y'')$, for all $(X, Y), (X', Y'), (X'', Y'') \in \mathcal{Y}^2$;
weak order: if it is complete and transitive;
nontrivial: if $(X, Y) \prec (X', Y')$ for some $(X, Y), (X', Y') \in \mathcal{Y}^2$.

If $\precsim$ is assumed to be complete, then $\sim$ and $\prec$ are the symmetric and the asymmetric parts of $\precsim$, respectively.

We recall that in the literature there is not a commonly accepted definition (and nomenclature) for dissimilarity measures between fuzzy sets. In [12–14] a thorough analysis of axioms that a dissimilarity measure should satisfy is carried out together with a comparison between different definitions. Here, we adopt a very broad definition where a *dissimilarity measure* is a function $D : \mathcal{Y}^2 \to \mathbb{R}$ satisfying: for all $X, Y \in \mathcal{Y}$,

(i) $D(X, Y) \geq 0$;
(ii) $D(X, Y) = D(Y, X)$;
(iii) $X = Y \Longrightarrow D(X, Y) = 0$.

We notice that *(i)–(iii)* are necessary properties that all dissimilarity measures considered in this paper satisfy, though they are not sufficient to completely capture the notion of dissimilarity. Further properties according to [12–14] for the numerical function D will be singled out in the next section.

Definition 1 Let $\precsim$ be a comparative dissimilarity and $D : \mathcal{Y}^2 \to \mathbb{R}$ a dissimilarity measure. We say that D **represents** $\precsim$ if and only if, for all $(X, Y), (X', Y') \in \mathcal{Y}^2$, it holds

$$\begin{cases} (X, Y) \precsim (X', Y') \Longrightarrow D(X, Y) \leq D(X', Y'), \\ (X, Y) \prec (X', Y') \Longrightarrow D(X, Y) < D(X', Y'). \end{cases}$$

If only the first implication is satisfied, we say that D **almost represents** $\precsim$.

As is well-known, if $\precsim$ is complete, the above conditions of representability can be summarized as follows:

$$(X, Y) \precsim (X', Y') \iff D(X, Y) \leq D(X', Y').$$

Proposition 1 *If $\precsim$ is a trivial complete dissimilarity relation on $\mathcal{Y}^2$, i.e., the asymmetric part $\prec$ is empty, then a dissimilarity measure D satisfying (i)–(iii) represents $\precsim$ if and only if $D(X, Y) = 0$, for all $X, Y \in \mathcal{Y}$.*

Proof Since D represents $\precsim$ and $\prec$ is empty, for all $X, Y, X', Y' \in \mathcal{Y}$, it holds

$$D(X, Y) = D(X', Y').$$

Finally, by property *(iii)* we derive that, for all $X, Y \in \mathcal{Y}$, it holds

$$D(X, Y) = D(X, X) = D(Y, Y) = 0,$$

thus the claim follows. □

The previous proposition shows that the case of a trivial complete dissimilarity relation $\precsim$ is not interesting, as it can be uniquely represented by a trivial dissimilarity measure which is constantly equal to 0. Therefore, in what follows we will always assume that $\precsim$ is nontrivial.

3 Basic axioms

Given a comparative dissimilarity relation $\precsim$ on $\mathcal{Y}^2$, in the following we propose a set of axioms that reveal to be necessary and sufficient for $\precsim$ to have a dissimilarity measure representation inside a suitable class of dissimilarity measures.

The next axiom is a necessary condition for the existence of any real-valued function representing a binary relation.

(FD0) $\precsim$ is a weak order on $\mathcal{Y}^2$.

The completeness of relation $\precsim$ can be removed and required only in some specific cases: we assume it for simplicity. We note that, under axiom **(FD0)**, $\sim$ is an equivalence relation and $\prec$ a strict order on the quotient set.

The next axiom requires the comparative degree of dissimilarity to be independent of the common increase or decrease of the presence/absence of the features in the objects of a pair. In fact, what is discriminant in assessing the comparative degrees of dissimilarity is the distance between the two membership degrees assigned to each feature.

(FD1) For all $X, Y \in \mathcal{Y}$, for all $k \in I$, for all $\varepsilon \leq \min(x_k, y_k)$, it holds:

$$(X, Y) \sim (X_k^{-\varepsilon}, Y_k^{-\varepsilon}).$$

The next example, inspired by an example given in [5], shows a situation of three pairs that the axioms **(FD0)** and **(FD1)** require to be equivalent.

Example 1 Let us consider a comparative dissimilarity among banks in the Euro zone. As prescribed by Basel II and Basel III accords [2, 3], banks should be rated in a way to point out their ability to pay debts they have contracted with other financial institutions. Ratings take into account qualitative, quantitative and performance information and the way they are calculated may vary, depending on the credit agency. Here, we assume that banks are described by the following attributes extracted from Basel accords:

- h_1: high quality of the enterprise;
- h_2: low cost of interest paid;
- h_3: compliance with the terms of repayment of received credit;
- h_4: absence of inactive accounts or with a negative balance;
- h_5: good use of credit lines.

Consider the following fuzzy description profiles related to six different banks of the Euro zone.

$\mathcal{H}$	h_1	h_2	h_3	h_4	h_5
X	0.5	0.4	0.9	0.6	0.1
Y	0.4	0.8	0.3	0.8	0.2
X'	0.3	0.2	0.8	0.2	0.05
Y'	0.2	0.6	0.2	0.4	0.15
X''	0.1	0	0.6	0	0
Y''	0	0.4	0	0.2	0.1

Axioms **(FD0)** and **(FD1)** require that one must retain $(X, Y) \sim (X', Y') \sim (X'', Y'')$, that is all the pairs of banks (X, Y), (X', Y'), (X'', Y'') should be judged as equally dissimilar. ♦

The next axiom is a local strong form of symmetry.

(FD2) For all $X, Y \in \mathcal{Y}$, for all $k \in I$, denoting $X'_k = (x_1, \ldots, y_k, \ldots, x_m)$ and $Y'_k = (y_1, \ldots, x_k, \ldots, y_m)$, it holds:

$$(X, Y) \sim (X'_k, Y'_k).$$

Example 2 Refer to the features in Example 1 and consider the fuzzy description profiles below, related to 4 different banks of the Euro zone.

$\mathcal{H}$	h_1	h_2	h_3	h_4	h_5
X''	0.1	0	0.6	0	0
Y''	0	0.4	0	0.2	0.1
X'''	0.1	0.4	0.6	0.2	0.1
Y'''	0	0	0	0	0

Accepting axioms **(FD0)** and **(FD2)** implies to set $(X'', Y'') \sim (X''', Y''')$. ♦

As the following proposition shows, for a weak order, local symmetry implies symmetry. We note that the transitivity is necessary and that the converse does not hold.

Proposition 2 *Let $\precsim$ be a comparative dissimilarity on $\mathcal{Y}^2$. If $\precsim$ satisfies axioms* **(FD0)** *and* **(FD2)***, then, for every $X, Y \in \mathcal{Y}$ one has: $(X, Y) \sim (Y, X)$.*

Proof The proof trivially follows by applying at most m times **(FD2)** and **(FD0)**. □

The next proposition shows that under axioms **(FD0)** and **(FD1)**, all pairs of identical fuzzy description profiles belong to the same equivalence class.

Proposition 3 *Let $\precsim$ be a comparative dissimilarity on $\mathcal{Y}^2$. If $\precsim$ satisfies axioms* **(FD0)** *and* **(FD1)***, then for every $X \in \mathcal{Y}$ one has: $(\underline{1}, \underline{1}) \sim (X, X) \sim (\underline{0}, \underline{0})$.*

Proof For every $X \in \mathcal{Y}$, in particular $X = \underline{1}$, apply m times axiom **(FD1)** taking $\varepsilon = x_k$ and then use **(FD0)**. □

The following axiom is a boundary condition. It provides a natural left limitation: "the elements of each pair (X, Y) are at least dissimilar to each other as an element of the pair is from itself". On the other hand, for the right limitation it is not enough to refer to any pair (X, X^c) formed by a profile and its complement, but it is required that the profile X is crisp, or equivalently that the supports of X and X^c are disjoint.
(FD3) The following conditions hold:

a) for every $X, Y \in \mathcal{Y}$, $(X, X) \precsim (X, Y)$ and $(Y, Y) \precsim (X, Y)$, and if $x_k \neq y_k$, for all $k \in I$, the comparisons must be strict;
b) for every $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$, $(X, Y) \precsim (X, X^c)$, and if $s_X \subseteq s_Y$ and $y_k < 1$, for all $k \in s_Y \setminus s_X$, the comparison must be strict.

The following is a monotonicity axiom.
(FD4) For all $X, Y \in \mathcal{Y}$, for all $k \in I$, such that $x_k \leq y_k$, for all $0 < \varepsilon \leq x_k$ and $0 < \eta \leq 1 - y_k$, it holds:

$$(X, Y) \precsim (X_k^{-\varepsilon}, Y) \quad \text{and} \quad (X, Y) \precsim (X, Y_k^{\eta}).$$

The following Theorem 4 shows that the introduced axioms are necessarily satisfied by any comparative dissimilarity agreeing with a dissimilarity measure, taking into account the distances of the degree of membership of each feature in the compared fuzzy description profiles. The same axioms become necessary and sufficient together with the following structural axiom **(Q)**, known as Debreu's condition [16], which assures the representability of a weak order $\precsim$ by a real function.
(Q) There is a countable $\prec$-dense set $\mathcal{Z} \subseteq \mathcal{Y}^2$ (i.e., for all $(X, Y), (X', Y') \in \mathcal{Y}^2$, with $(X, Y) \prec (X', Y')$, there exists $(X'', Y'') \in \mathcal{Z}$, such that $(X, Y) \prec (X'', Y'') \prec (X', Y')$).

Theorem 4 *Let $\precsim$ be a nontrivial comparative dissimilarity relation on $\mathcal{Y}^2$. Then, the following statements are equivalent:*

(i) $\precsim$ *satisfies* **(FD0)–(FD4)** *and* **(Q)**;
(ii) there exists a function (unique under strictly increasing transformations of $[0, 1]$*)* $\Phi : \mathcal{Y}^2 \to [0, 1]$ *representing* $\precsim$ *in the sense of Definition 1 and a function* $\varphi : \mathcal{Y} \to [0, 1]$ *such that:*

a) for all $X, Y \in \mathcal{Y}$*, it holds*

$$\Phi(X, Y) = \Phi(|X - Y|, \underline{0}) = \varphi(|X - Y|);$$

b) $Z \leq Z' \Longrightarrow \varphi(Z) \leq \varphi(Z')$*, for every* $Z, Z' \in \mathcal{Y}$*;*
c) $\varphi(\underline{0}) = 0$ *and* $\varphi(\underline{1}) = 1$*;*
d) $\varphi(Z) = \varepsilon \in \{0, 1\}$ *implies* $z_k = \varepsilon$ *for at least one* $k \in I$*, for every* $Z \in \mathcal{Y}$*.*

Proof We first prove that *(i)* $\Longrightarrow$ *(ii)*. Axioms **(FD0)** and **(Q)** are sufficient conditions for the existence of a function $\Phi : \mathcal{Y}^2 \to \mathbb{R}$ representing $\precsim$ [22]. Now, applying at most m times **(FD1)** with $\varepsilon = \min(x_k, y_k)$ and at most $\frac{m}{2}$ times **(FD2)** we get, by **(FD0)**, that $(X, Y) \sim (|X - Y|, \underline{0})$. So, $\precsim$ induces in $\mathcal{Y}$ a strict order among the equivalence classes represented by $(|X - Y|, \underline{0})$. Then, since Φ represents $\precsim$ we have $\Phi(X, Y) = \Phi(|X - Y|, \underline{0})$. Thus it is sufficient to define, for all $Z \in \mathcal{Y}$, $\varphi(Z) = \Phi(Z, \underline{0})$ and note that φ satisfies condition *a)*.

We now prove the validity of statement *b)*. Let $Z = |X - Y| \leq Z' = |X' - Y'|$. Taking into account Proposition 3, if we start from $(Z', \underline{0})$ and apply **(FD4)** for $k = 1, \ldots, m$ with $\varepsilon = z'_k - z_k$, then we derive $(Z, \underline{0}) \precsim (Z', \underline{0})$, and since Φ represents $\precsim$ we get *b)*.

Conditions *c)* and *d)* follow by axiom **(FD3)**, by noting that $\underline{0} = |X - X| = |Y - Y| \leq |X - Y|$, for all $X, Y \in \mathcal{Y}$, while $\underline{1} = |X - X^c| \geq |X - Y|$, for all $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$. Thus, by the previous point we derive that $(\underline{0}, \underline{0}) \precsim (|X - Y|, \underline{0}) \precsim (\underline{1}, \underline{0})$, for all $X, Y \in \mathcal{Y}$, and $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$ by non-triviality. Since Φ is unique under strictly increasing transformations, we can consider particular Φ and φ taking values in $[0, 1]$ and such that $\varphi(\underline{0}) = 0$ and $\varphi(\underline{1}) = 1$. Next, for all $X, Y \in \mathcal{Y}$, we can have $(X, X) \sim (X, Y)$ or $(Y, Y) \sim (X, Y)$ only when there is at least one $k \in I$ with $x_k = y_k$, thus $\varphi(Z) = 0$ implies that there is at least one $k \in I$ with $z_k = |x_k - y_k| = 0$. Analogously, for all $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$, we can have $(X, Y) \sim (X, X^c)$ only when there is at least one $k \in I$ with $k \in s_X \setminus s_Y$, or $k \in s_Y \setminus s_X$ and $y_k = 1$, thus $\varphi(Z) = 1$ implies that there is at least on $k \in I$ with $z_k = |x_k - y_k| = 1$.

Let us consider now the implication *(ii)* $\Longrightarrow$ *(i)*. Every binary relation $\precsim$ representable by a real function satisfies axiom **(FD0)** and **(Q)** [22]. We must prove that $\precsim$ satisfies axioms **(FD1)–(FD4)**. Taking into account the representability of $\precsim$ by Φ we deduce that condition *a)* in *(ii)* implies **(FD1)** and **(FD2)** whereas condition *b)* in *(ii)* implies **(FD4)**. To prove condition *a)* of axiom **(FD3)** it is sufficient to consider that, for all $X, Y \in \mathcal{Y}$, Φ assigns 0 to all the elements of the equivalence class of $(\underline{0}, \underline{0})$, that contains the pairs (X, X) and (Y, Y), while it cannot contain any pair (X, Y) with $x_k \neq y_k$, for all $k \in I$. Analogously, condition *b)* of axiom **(FD3)** follows since, for all $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$, Φ assigns 1 to all the elements of the equivalence class of $(\underline{1}, \underline{0})$, that contains the pair (X, X^c), while it cannot contain any pair (X, Y) with $s_X \subseteq s_Y$ and $y_k < 1$, for all $k \in s_Y \setminus s_X$. □

The previous theorem singles out a class of functions $D : \mathcal{Y}^2 \to [0, 1]$, obtained from Φ through strictly increasing transformation of $[0, 1]$, all representing the same $\precsim$. The axioms **(FD0)–(FD4)** assure that every such D satisfies the properties *(i)–(iii)* that we reported in Section 2 for a broad-sense dissimilarity measure, together with the following additional properties analyzed in [12–14]:

(iv) $D(X, X) = 0$, for all $X \in \mathcal{Y}$;
(v) $D(X, X^c) = 1$, for all $X \in \mathcal{X}$;
(vi) $D(X, Z) \geq D(X, Y)$ and $D(X, Z) \geq D(Y, Z)$, for all $X, Y, Z \in \mathcal{Y}$ with $X \leq Y \leq Z$.

We point out that a function $D : \mathcal{Y} \to [0, 1]$ satisfying properties *(i)–(vi)* is called a *distance measure* in [12], not referring to the usual metric sense. Thus, the dissimilarity measures studied in this paper are particular distance measures between fuzzy sets.

4 Representation by a weighted L^p distance

Condition *(ii)* of Theorem 4 identifies a too wide and therefore too general class of functions. In the following we will study those relations representable by the elements of a particular subclass of functions Φ that is the class of the weighted L^p distances, for $p \in [1, +\infty)$.

Every element of this class takes into account the different weights that the decision maker or the field expert assigns to each feature through $\precsim$. Nevertheless, interactions among features are not admitted by this class of dissimilarity measures.

Definition 2 Let $p \in [1, +\infty)$. A **weighted** L^p **distance** is a function $D^p_{\boldsymbol{\alpha}} : \mathcal{Y}^2 \to [0, 1]$ parameterized by $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)$ with $\alpha_k \geq 0$, $k = 1, \ldots, m$, and $\sum_{k=1}^m \alpha_k = 1$, defined, for every $X, Y \in \mathcal{Y}$, as

$$D^p_{\boldsymbol{\alpha}}(X, Y) = \left(\sum_{k=1}^m \alpha_k |x_k - y_k|^p \right)^{\frac{1}{p}}.$$

The corresponding **p-th power weighted L^p distance** is a function $W^p_{\boldsymbol{\alpha}} : \mathcal{Y}^2 \to [0,1]$ defined, for every $X, Y \in \mathcal{Y}$, as

$$W^p_{\boldsymbol{\alpha}}(X,Y) = (D^p_{\boldsymbol{\alpha}}(X,Y))^p = \sum_{k=1}^{m} \alpha_k |x_k - y_k|^p.$$

In particular, for $p = 1$ we get a *weighted Manhattan distance*, and for $p = 2$ we get a *weighted Euclidean distance*.

For every $p \in [1, +\infty)$, we notice that the function $f(x) = x^p$, for all $x \in [0,1]$, is continuous, strictly increasing, and such that $f(0) = 0$ and $f(1) = 1$. Therefore, f is invertible and f^{-1} is defined, for all $x \in [0,1]$, as $f^{-1}(x) = x^{\frac{1}{p}}$, which is continuous, strictly increasing, and such that $f^{-1}(0) = 0$ and $f^{-1}(1) = 1$. This implies that, for all $(X,Y), (X',Y') \in \mathcal{Y}^2$, it holds

$$D^p_{\boldsymbol{\alpha}}(X,Y) \leq D^p_{\boldsymbol{\alpha}}(X',Y') \Longleftrightarrow W^p_{\boldsymbol{\alpha}}(X,Y) \leq W^p_{\boldsymbol{\alpha}}(X',Y'). \tag{3}$$

As is well-known, for a fixed $\boldsymbol{\alpha}$, defining the relation $X \equiv_0 Y \Longleftrightarrow D^p_{\boldsymbol{\alpha}}(X,Y) = 0$, then $D^p_{\boldsymbol{\alpha}}$ turns out to be a metric on the quotient space $\mathcal{Y}_{/\equiv_0}$. In other terms, for all equivalence classes $[X], [Y], [Z] \in \mathcal{Y}_{/\equiv_0}$ we have that

(i) $D^p_{\boldsymbol{\alpha}}(X,Y) = 0$ if and only if $[X] = [Y]$;
(ii) $D^p_{\boldsymbol{\alpha}}(X,Y) = D^p_{\boldsymbol{\alpha}}(Y,X)$;
(iii) $D^p_{\boldsymbol{\alpha}}(X,Y) \leq D^p_{\boldsymbol{\alpha}}(X,Z) + D^p_{\boldsymbol{\alpha}}(Z,Y)$.

In particular, $D^p_{\boldsymbol{\alpha}}$ is a metric on the whole $\mathcal{Y}$ if (and only if) $\boldsymbol{\alpha}$ is strictly positive.

4.1 Rationality principle

The next axiom highlights "the constraint accepted" to obtain that the function representing our comparative dissimilarity $\precsim$ belongs to the particular subclass of weighted L^p distances.

(R-p) For all $n \in \mathbb{N}$, for all $(X_1,Y_1), \ldots, (X_n,Y_n), (X'_1,Y'_1), \ldots, (X'_n,Y'_n) \in \mathcal{Y}^2$ with $(X_i,Y_i) \precsim (X'_i,Y'_i), i = 1, \ldots, n-1$, and $(X_n,Y_n) \prec (X'_n,Y'_n)$, there are no $\lambda_1, \ldots, \lambda_n > 0$ with $\sum_{i=1}^n \lambda_i = 1$ such that:

$$\sum_{i=1}^{n} \lambda_i |X'_i - Y'_i|^p \leq \sum_{i=1}^{n} \lambda_i |X_i - Y_i|^p.$$

Let us notice that if $\precsim$ is trivial, then **(R-p)** is vacuously satisfied as **(R-p)** requires the presence of at least one strict comparison in order to be applied. In other terms, non-triviality of $\precsim$ is not implied by **(R-p)** and must be explicitly required. The following proposition lists some immediate properties implied by **(R-p)**.

Proposition 5 *Let $p \in [1, +\infty)$. Let $\precsim$ be a nontrivial complete comparative dissimilarity relation on $\mathcal{Y}^2$ satisfying* **(R-p)**. *Then, for all $X, Y, X', Y' \in \mathcal{Y}$, the following statements hold:*

(i) it must be $(X,Y) \sim (|X-Y|, \underline{0})$;
(ii) if $|X-Y| \leq |X'-Y'|$, then it must be $(X,Y) \precsim (X',Y')$;
(iii) it must be $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$.

Proof Statement *(i)*. Since $|X - Y|^p = ||X - Y| - \underline{0}|^p$, by **(R-**$p$**)** it cannot be neither $(X, Y) \prec (|X - Y|, \underline{0})$, nor $(|X - Y|, \underline{0}) \prec (X, Y)$, and since $\precsim$ is complete, it must be $(X, Y) \sim (|X - Y|, \underline{0})$.

Statement *(ii)*. If $|X - Y| \leq |X' - Y'|$ and $(X', Y') \prec (X, Y)$, then $|X - Y|^p \leq |X' - Y'|^p$ and **(R-**p**)** is violated, thus it cannot be $(X', Y') \prec (X, Y)$. Then, the statement follows by the completeness of $\precsim$.

Statement *(iii)*. By non-triviality, there are $(X, Y), (X', Y') \in \mathcal{Y}^2$ such that $(X, Y) \prec (X', Y')$. Notice that it cannot be $(\underline{1}, \underline{0}) \prec (\underline{0}, \underline{0})$ by statement *(ii)*. Suppose $(\underline{1}, \underline{0}) \precsim (\underline{0}, \underline{0})$. Denote by $q = \max_{k=1,\dots,m} \{|x'_k - y'_k|^p - |x_k - y_k|^p\}$ and take $\lambda_1, \lambda_2 \in (0, 1)$ with $\lambda_1 \geq \frac{q}{1+q}$ and $\lambda_2 = 1 - \lambda_1$. Then we have that

$$\lambda_1 |\underline{0} - \underline{0}|^p + \lambda_2 |X' - Y'|^p \leq \lambda_1 |\underline{1} - \underline{0}|^p + \lambda_2 |X - Y|^p,$$

violating condition **(R-**p**)**. Thus, since $\precsim$ is complete, **(R-**p**)** implies $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$. □

The above axiom has an easy interpretation. It asserts that if you have n pairs (X_i, Y_i) of fuzzy profiles and you judge the elements of each of them no more dissimilar than those of other n pairs (X'_i, Y'_i), with at least one strict comparison, combining in a positive convex combination the $|X_i - Y_i|^p$'s and the $|X'_i - Y'_i|^p$'s you cannot obtain two fuzzy profiles Z and Z' such that they satisfy $Z' \leq Z$.

In the next example we provide a comparative dissimilarity assessment which violates the above rationality principle.

Example 3 Let $p \in [1, +\infty)$. Referring to the features in Example 1, let us consider the following profiles:

$\mathcal{H}$	h_1	h_2	h_3	h_4	h_5
X_1	1/2	1	1/4	3/4	1/10
Y_1	1/2	2/3	1/4	1/2	1/10
X_2	1	1	1/6	1/2	1/2
Y_2	1/2	1	1/6	1/2	1/2
X_3	2/3	2/3	1/3	0	1/4
Y_3	1/6	1	1/3	0	1/4
X_4	1/8	1	1/2	1/2	0
Y_4	1/8	1	1/3	1/4	0
X_5	1	1/8	0	0	1
Y_5	1/2	1/8	0	1/4	1
X_6	0	1/2	1/3	0	0
Y_6	0	1/6	1/3	0	2/3
X_7	1/4	1/6	1/6	1	1/3
Y_7	1/4	1/6	0	1	1
X_8	0	1/3	1/4	3/4	0
Y_8	1/2	0	1/4	1/2	0

Suppose now we assign the following reasonable relation: $(X_1, Y_1) \prec (X_2, Y_2)$, $(X_3, Y_3) \prec (X_4, Y_4)$, $(X_5, Y_5) \prec (X_6, Y_6)$, $(X_7, Y_7) \prec (X_8, Y_8)$.

It is easy to prove that the relation violates axiom **(R-p)**. By trivial computations, taking all λ_i's equal to $1/4$, one obtains

$$\sum_{i\in\{1,3,5,7\}} \frac{1}{4}|X_i - Y_i|^p = \left(\frac{2}{4\cdot 2^p}, \frac{2}{4\cdot 3^p}, \frac{1}{4\cdot 6^p}, \frac{2}{4\cdot 4^p}, \frac{2^p}{4\cdot 3^p}\right) = \sum_{i\in\{2,4,6,8\}} \frac{1}{4}|X_i - Y_i|^p.$$

♦

The next theorem shows that, for a nontrivial complete $\precsim$, condition **(R-p)** implies all the axioms from **(FD0)** to **(FD4)**. Nevertheless, since condition **(R-p)** deals with finite sets of pairs, condition **(Q)** is not guaranteed to hold. This is why **(R-p)** does not assure the representability of $\precsim$ on the whole $\mathcal{Y}^2$, but only its almost representability, as will be shown in Subsection 4.3.

Theorem 6 *Let $p \in [1, +\infty)$. Let $\precsim$ be a nontrivial complete comparative dissimilarity relation on $\mathcal{Y}^2$ satisfying* **(R-p)**. *Then $\precsim$ is transitive and axioms* **(FD1)–(FD4)** *hold.*

Proof To prove transitivity note that, for all $(X, Y), (X', Y'), (X'', Y'') \in \mathcal{Y}^2$, it holds

$$\tfrac{1}{3}\left(|X - Y|^p + |X' - Y'|^p + |X'' - Y''|^p\right) = \tfrac{1}{3}\left(|X' - Y'|^p + |X'' - Y''|^p + |X - Y|^p\right).$$

Hence, if $(X, Y) \precsim (X', Y')$, $(X', Y') \precsim (X'', Y'')$ and $(X'', Y'') \precsim (X, Y)$, then **(R-$p$)** implies that none of the comparisons can be strict, thus we get transitivity.

To prove **(FD1)** and **(FD2)** it is sufficient to note that $|X-Y| = |X_k^{-\varepsilon} - Y_k^{-\varepsilon}| = |X'_k - Y'_k|$, for all $X, Y \in \mathcal{Y}$. Thus, statement *(i)* of Proposition 5 and transitivity of $\precsim$ imply that $(X, Y) \sim (X_k^{-\varepsilon}, Y_k^{-\varepsilon})$ and $(X, Y) \sim (X'_k, Y'_k)$.

To prove **(FD3)** notice that $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$ by statement *(iii)* of Proposition 5. Condition *a)* follows since $\underline{0} = |X - X| = |Y - Y| \le |X - Y|$, for all $X, Y \in \mathcal{Y}$, thus, by statement *(ii)* of Proposition 5 it must be $(X, X) \precsim (X, Y)$ and $(Y, Y) \precsim (X, Y)$. Moreover, it cannot neither be $(X, Y) \precsim (X, X)$ nor $(X, Y) \precsim (Y, Y)$ if $x_k \neq y_k$, for all $k \in I$, as we have $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$. Indeed, denoting by $q = \min_{k\in I} |x_k - y_k|^p$, we get

$$\frac{1}{1+q}|X - X|^p + \frac{q}{1+q}|\underline{1} - \underline{0}|^p \le \frac{1}{1+q}|X - Y|^p + \frac{q}{1+q}|\underline{0} - \underline{0}|^p,$$
$$\frac{1}{1+q}|Y - Y|^p + \frac{q}{1+q}|\underline{1} - \underline{0}|^p \le \frac{1}{1+q}|X - Y|^p + \frac{q}{1+q}|\underline{0} - \underline{0}|^p,$$

that contradict **(R-p)**. Analogously, condition *b)* follows since $\underline{1} = |X - X^c| \ge |X - Y|$, for all $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$, that implies $(X, Y) \precsim (X, X^c)$ by statement *(ii)* of Proposition 5. Moreover, it cannot be $(X, X^c) \precsim (X, Y)$ if $s_X \subseteq s_Y$ and $y_k < 1$, for all $k \in s_Y \setminus s_X$, as we have $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$. Indeed, denoting by $q = 1 - \max_{k\in I} |x_k - y_k|^p$, we get

$$\frac{1}{1+q}|X - Y|^p + \frac{q}{1+q}|\underline{1} - \underline{0}|^p \le \frac{1}{1+q}|X - X^c|^p + \frac{q}{1+q}|\underline{0} - \underline{0}|^p,$$

that contradicts **(R-p)**.

Finally, axiom **(FD4)** holds since $|X - Y| \le |X_k^{-\varepsilon} - Y|$ and $|X - Y| \le |X - Y_k^{\eta}|$, for all $X, Y \in \mathcal{Y}$, for all $k \in I$, such that $x_k \le y_k$, for all $0 < \varepsilon \le x_k$ and $0 < \eta \le 1 - y_k$. Thus, by statement *(ii)* of Proposition 5 it must be $(X, Y) \precsim (X_k^{-\varepsilon}, Y)$ and $(X, Y) \precsim (X, Y_k^{\eta})$. □

4.2 Representability theorems

In the following we characterize the weighted L^p distance representability of a nontrivial complete comparative dissimilarity relation on a finite set of fuzzy description profiles.

Theorem 7 *Let $p \in [1, +\infty)$. Let $\precsim$ be a nontrivial complete comparative dissimilarity relation on a finite $\mathcal{F} \subset \mathcal{Y}^2$. Then, the following statements are equivalent:*

(i) $\precsim$ *satisfies* **(R-**p**)**;
(ii) there exists a weight vector $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m)$ with $\alpha_k \geq 0$, $k = 1, \dots, m$, and $\sum_{k=1}^m \alpha_k = 1$ such that, for all $(X, Y), (X', Y') \in \mathcal{F}$, it holds that

$$(X, Y) \precsim (X', Y') \Longleftrightarrow D^p_{\boldsymbol{\alpha}}(X, Y) \leq D^p_{\boldsymbol{\alpha}}(X', Y').$$

Proof Since $\mathcal{F}$ is finite, the binary relation $\precsim$ amounts to a finite number of comparisons. Consider the sets

$$\mathcal{S} = \{((X, Y), (X', Y')) \in \mathcal{F}^2 : (X, Y) \prec (X', Y')\},$$
$$\mathcal{W} = \{((X, Y), (X', Y')) \in \mathcal{F}^2 \setminus \mathcal{S} : (X, Y) \precsim (X', Y')\},$$

with $s = \text{card}\,\mathcal{S}$ and $w = \text{card}\,\mathcal{W}$, and fix two enumerations $\mathcal{S} = \{((X_j, Y_j), (X'_j, Y'_j))\}_{j\in J}$ and $\mathcal{W} = \{((X_h, Y_h), (X'_h, Y'_h))\}_{h\in H}$ with $J = \{1, \dots, s\}$ and $H = \{1, \dots, w\}$.

Condition *(ii)* is equivalent to the solvability of the following linear system

$$\begin{cases} \mathbf{A}\boldsymbol{\beta} > \mathbf{0}, \\ \mathbf{B}\boldsymbol{\beta} \geq \mathbf{0}, \\ \boldsymbol{\beta} \geq \mathbf{0}, \end{cases}$$

with unknown $\boldsymbol{\beta} \in \mathbb{R}^{m\times 1}$, and $\mathbf{A} \in \mathbb{R}^{s\times m}$ and $\mathbf{B} \in \mathbb{R}^{w\times m}$, where the s rows of $\mathbf{A}$ are the vectors $|X'_j - Y'_j|^p - |X_j - Y_j|^p$, for all $j \in J$, while the w rows of $\mathbf{B}$ are the vectors $|X'_h - Y'_h|^p - |X_h - Y_h|^p$, for all $h \in H$. Indeed, if we have a weight vector $\boldsymbol{\alpha}$ satisfying *(ii)*, then setting $\boldsymbol{\beta} = \boldsymbol{\alpha}^T$ we get a solution of the above system. On the converse, if $\boldsymbol{\beta}$ is a solution of the above system, then defining $\alpha_k = \frac{\beta_k}{\sum_{i=1}^m \beta_i}$, we get a weight vector $\boldsymbol{\alpha}$ satisfying *(ii)*.

By the Motzkin's theorem of the alternative [24], the solvability of the above system is equivalent to the non-solvability of the following system

$$\begin{cases} \boldsymbol{\mu}\mathbf{A} + \boldsymbol{\nu}\mathbf{B} \leq \mathbf{0}, \\ \boldsymbol{\mu}, \boldsymbol{\nu} \geq \mathbf{0}, \\ \boldsymbol{\mu} \neq \mathbf{0}, \end{cases}$$

with unknowns $\boldsymbol{\mu} \in \mathbb{R}^{1\times s}$ and $\boldsymbol{\nu} \in \mathbb{R}^{1\times w}$. In particular, the first inequality reduces to

$$\sum_{j\in J} \mu_j(|X'_j - Y'_j|^p - |X_j - Y_j|^p) + \sum_{h\in H} \nu_h(|X'_h - Y'_h|^p - |X_h - Y_h|^p) \leq \mathbf{0}.$$

Therefore, it easy to see that, dividing both sides of the inequality by the sum of the μ_j's and the ν_h's, the non-solvability of the above system is equivalent to condition **(R-**p**)**. □

Remark 1 We note that, in the hypotheses of previous theorem, if $(X, Y), (X', Y') \in \mathcal{F}$ and it holds $|X - Y|^p < |X' - Y'|^p$ then it must be $(X, Y) \prec (X', Y')$. In particular, if $(\underline{0}, \underline{0}), (\underline{1}, \underline{0}) \in \mathcal{F}$, then it must be $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$, as already noticed.

Consider now the case where $\precsim$ is a nontrivial complete relation on $\mathcal{Y}^2$. In this case, axiom **(R-**p**)** is not sufficient to assure representability of $\precsim$ by a weighted L^p distance $D^p_{\boldsymbol{\alpha}}$ on the whole $\mathcal{Y}^2$, as the following Example 4 shows.

Example 4 Let $p \in [1, +\infty)$. We consider objects described by only two features, i.e., $m = 2$ and $I = \{1, 2\}$, and we introduce the following dissimilarity relation: for every $(X, Y), (X', Y') \in \mathcal{Y}^2$,

$$(X, Y) \precsim (X', Y') \iff \begin{cases} |x_1 - y_1| < |x'_1 - y'_1| \\ \text{or} \\ |x_1 - y_1| = |x'_1 - y'_1| \text{ and } |x_2 - y_2| \leq |x'_2 - y'_2|. \end{cases} \tag{4}$$

We first notice that, for every $X, Y \in \mathcal{Y}$, it holds that $(X, Y) \sim (|X - Y|, \underline{0})$ as $|x_1 - y_1| = ||x_1 - y_1| - 0|$ and $|x_2 - y_2| = ||x_2 - y_2| - 0|$. Moreover, we have that $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$.

The relation $\precsim$ is a nontrivial weak order on $\mathcal{Y}^2$ and its quotient relation on $\mathcal{Y}^2_{/\sim}$ corresponds to the lexicographic order on $\mathcal{Y}$. As is well-known (see, e.g., [22]), we cannot find a real function $\varphi : \mathcal{Y} \to \mathbb{R}$ representing the lexicographic order on $\mathcal{Y}$. Hence, by Theorems 4 and 6 we cannot find a weight vector $\boldsymbol{\alpha}$ such that $D^p_{\boldsymbol{\alpha}}$ represents $\precsim$ on the whole $\mathcal{Y}^2$.

Now, we show that $\precsim$ satisfies **(R-**p**)**. Let $\mathcal{G} \subset \mathcal{Y}^2$ be arbitrary and finite, and take $\mathcal{F} = \mathcal{G} \cup \{(\underline{0}, \underline{0}), (\underline{1}, \underline{0})\}$. Therefore, the restriction of $\precsim$ to $\mathcal{F}$ is a nontrivial complete weak order, that we denote by the same symbol.

Consider a weight vector $\boldsymbol{\alpha} = (\alpha, 1 - \alpha)$ with $\alpha \in (0, 1)$. The strictly compared pairs can be partitioned as

$$\begin{aligned} \mathcal{S}_1 = \{((X, Y), (X', Y')) \in \mathcal{F}^2 \,:\, & (X, Y) \prec (X', Y'), \\ & |x_1 - y_1| < |x'_1 - y'_1|,\ |x_2 - y_2| \leq |x'_2 - y'_2|\}, \\ \mathcal{S}_2 = \{((X, Y), (X', Y')) \in \mathcal{F}^2 \,:\, & (X, Y) \prec (X', Y'), \\ & |x_1 - y_1| < |x'_1 - y'_1|,\ |x_2 - y_2| > |x'_2 - y'_2|\}, \\ \mathcal{S}_3 = \{((X, Y), (X', Y')) \in \mathcal{F}^2 \,:\, & (X, Y) \prec (X', Y'), \\ & |x_1 - y_1| = |x'_1 - y'_1|,\ |x_2 - y_2| < |x'_2 - y'_2|\}. \end{aligned}$$

Choosing $\alpha \in (0, 1)$ such that

$$\alpha < \min_{((X,Y),(X',Y')) \in \mathcal{S}_2} \frac{|x_2 - y_2|^p - |x'_2 - y'_2|^p}{|x'_1 - y'_1|^p - |x_1 - y_1|^p + |x_2 - y_2|^p - |x'_2 - y'_2|^p},$$

we have that, for all $(X, Y), (X', Y') \in \mathcal{F}$,

$$(X, Y) \precsim (X', Y') \iff D^p_{\boldsymbol{\alpha}}(X, Y) \leq D^p_{\boldsymbol{\alpha}}(X', Y').$$

Hence, by Theorem 7 the relation $\precsim$ satisfies **(R-**p**)** on $\mathcal{F}$ and by the arbitrariness of the choice of $\mathcal{G}$, $\precsim$ satisfies **(R-**p**)** on the whole $\mathcal{Y}^2$. ♦

Nevertheless, axiom **(R-**p**)** guarantees representability only on every finite subset $\mathcal{Y}^2$, by virtue of Theorem 7. This implies that the parameter $\boldsymbol{\alpha}$ characterizing $D^p_{\boldsymbol{\alpha}}$ depends on the particular finite subset $\mathcal{F}$. To remedy this problem we can follow two paths: *(1)* reinforce the hypotheses or *(2)* weaken the requirements on the result, contenting ourselves with obtaining an almost-representation. For the first strategy we introduce a further axiom which requires that in each equivalence class containing $(|X - Y|, \underline{0})$ there must be one pair $(\underline{\varepsilon}, \underline{0})$. Indeed, having fixed $\boldsymbol{\alpha}$, if

$$D^p_{\boldsymbol{\alpha}}(X, Y) = D^p_{\boldsymbol{\alpha}}(|X - Y|, \underline{0}) = \left(\sum_{k=1}^{m} \alpha_k |x_k - y_k|^p \right)^{\frac{1}{p}} = \varepsilon,$$

where $\varepsilon \in [0,1]$, then

$$D^p_{\boldsymbol{\alpha}}(\underline{\varepsilon}, \underline{0}) = \left(\sum_{k=1}^{m} \alpha_k \varepsilon^p\right)^{\frac{1}{p}} = \varepsilon,$$

thus the pairs (X, Y), $(|X - Y|, \underline{0})$ and $(\underline{\varepsilon}, \underline{0})$ should be judged as equally dissimilar.
(FD5) For all $(X, Y) \in \mathcal{Y}^2$ there exists $\varepsilon \in [0,1]$, such that $(X, Y) \sim (\underline{\varepsilon}, \underline{0})$.

Given a nontrivial complete dissimilarity relation $\precsim$ on the whole $\mathcal{Y}^2$, we notice that **(FD5)** is a "richness axiom" used to guarantee both that all strict comparisons are preserved by the numerical representation of $\precsim$ and that the numerical representation of $\precsim$ is unique. Indeed, asking only **(R-**p**)** to hold, by applying a compacteness argument one can only prove the existence of a (not necessarily unique) vector $\boldsymbol{\alpha}$ whose corresponding $D^p_{\boldsymbol{\alpha}}$ almost represents $\precsim$.

Theorem 8 *Let $p \in [1, +\infty)$. Let $\precsim$ be a nontrivial complete comparative dissimilarity relation on $\mathcal{Y}^2$. Then, the following statements are equivalent:*

(i) *$\precsim$ satisfies* **(R-**p**)** *and* **(FD5)***;*
(ii) *there exists a weight vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)$ with $\alpha_k \geq 0$, $k = 1, \ldots, m$, and $\sum_{k=1}^m \alpha_k = 1$ such that, for all $(X, Y), (X', Y') \in \mathcal{Y}^2$, it holds*

$$(X, Y) \precsim (X', Y') \Longleftrightarrow D^p_{\boldsymbol{\alpha}}(X, Y) \leq D^p_{\boldsymbol{\alpha}}(X', Y').$$

Moreover, the weight vector $\boldsymbol{\alpha}$ is unique.

Proof The implication *(ii)* $\Longrightarrow$ *(i)* is easily proven, therefore we only prove *(i)* $\Longrightarrow$ *(ii)*.

For every finite $\mathcal{F} \subset \mathcal{Y}^2$ such that the restriction of $\precsim$ to $\mathcal{F}$ is nontrivial, Theorem 7 implies the existence of a weight vector $\boldsymbol{\alpha}^{\mathcal{F}} = (\alpha^{\mathcal{F}}_1, \ldots, \alpha^{\mathcal{F}}_m)$ with $\alpha^{\mathcal{F}}_k \geq 0$ and $\sum_{k=1}^m \alpha^{\mathcal{F}}_k = 1$, such that the corresponding $W^p_{\boldsymbol{\alpha}^{\mathcal{F}}}$ represents the restriction of $\precsim$ to $\mathcal{F}$. Notice that, by Proposition 5 every finite subset of $\mathcal{Y}^2$ containing $(\underline{0}, \underline{0})$ and $(\underline{1}, \underline{0})$ meets non-triviality, as it must be $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$.

Next, axiom **(FD5)** implies that, for all $(X, Y) \in \mathcal{Y}^2$, there exists $\varepsilon_{(X,Y)} \in [0,1]$, such that $(X, Y) \sim (\underline{\varepsilon_{(X,Y)}}, \underline{0})$. In particular, denoting by E_k the element of $\mathcal{Y}$ whose k-th component is 1 and the others are 0, we have that there exists $\beta_k \in [0,1]$ such that $(E_k, \underline{0}) \sim (\underline{\beta_k}, \underline{0})$. Now, for every $(X, Y), (X', Y') \in \mathcal{Y}^2$ we consider the finite subset of $\mathcal{Y}^2$

$$\begin{aligned}\mathcal{F} = \{&(X, Y), (X', Y'), (\underline{\varepsilon_{(X,Y)}}, \underline{0}), (\underline{\varepsilon_{(X',Y')}}, \underline{0}),\\ &(E_1, \underline{0}), \ldots, (E_m, \underline{0}), (\underline{\beta_1}, \underline{0}), \ldots, (\underline{\beta_m}, \underline{0}), (\underline{0}, \underline{0}), (\underline{1}, \underline{0})\}.\end{aligned}$$

By the previous point we have that there is a weight vector $\boldsymbol{\alpha}^{\mathcal{F}} = (\alpha^{\mathcal{F}}_1, \ldots, \alpha^{\mathcal{F}}_m)$ with $\alpha^{\mathcal{F}}_k \geq 0$ and $\sum_{k=1}^m \alpha^{\mathcal{F}}_k = 1$ such that

$$\begin{aligned}(X, Y) \precsim (X', Y') &\Longleftrightarrow W^p_{\boldsymbol{\alpha}^{\mathcal{F}}}(X, Y) \leq W^p_{\boldsymbol{\alpha}^{\mathcal{F}}}(X', Y'),\\ (X, Y) \sim (\underline{\varepsilon_{(X,Y)}}, \underline{0}) &\Longleftrightarrow W^p_{\boldsymbol{\alpha}^{\mathcal{F}}}(X, Y) = W^p_{\boldsymbol{\alpha}^{\mathcal{F}}}(\underline{\varepsilon_{(X,Y)}}, \underline{0}) = \varepsilon^p_{(X,Y)},\\ (X', Y') \sim (\underline{\varepsilon_{(X',Y')}}, \underline{0}) &\Longleftrightarrow W^p_{\boldsymbol{\alpha}^{\mathcal{F}}}(X', Y') = W^p_{\boldsymbol{\alpha}^{\mathcal{F}}}(\underline{\varepsilon_{(X',Y')}}, \underline{0}) = \varepsilon^p_{(X',Y')}.\end{aligned}$$

Moreover, for all $k = 1, \ldots, m$, we have

$$(E_k, \underline{0}) \sim (\underline{\beta_k}, \underline{0}) \Longleftrightarrow W^p_{\boldsymbol{\alpha}^{\mathcal{F}}}(E_k, \underline{0}) = \alpha^{\mathcal{F}}_k = \beta^p_k = W^p_{\boldsymbol{\alpha}^{\mathcal{F}}}(\underline{\beta_k}, \underline{0}),$$

thus the numbers $\beta_k \in [0, 1]$ are such that $\beta_k^p \geq 0$ and $\sum_{k=1}^m \beta_k^p = 1$. So, defining $\alpha_k = \beta_k^p$ we get

$$\varepsilon^p_{(X,Y)} = \sum_{k=1}^m \alpha_k |x_k - y_k|^p \quad \text{and} \quad \varepsilon^p_{(X',Y')} = \sum_{k=1}^m \alpha_k |x'_k - y'_k|^p,$$

and also

$$\varepsilon_{(X,Y)} \leq \varepsilon_{(X',Y')} \Longleftrightarrow \varepsilon^p_{(X,Y)} \leq \varepsilon^p_{(X',Y')}.$$

Hence, there exists a weight vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)$ with $\alpha_k \geq 0$ and $\sum_{k=1}^m \alpha_k = 1$ such that, for all $(X, Y), (X', Y') \in \mathcal{Y}^2$, it holds that

$$(X, Y) \precsim (X', Y') \Longleftrightarrow D^p_{\boldsymbol{\alpha}}(X, Y) \leq D^p_{\boldsymbol{\alpha}}(X', Y'),$$

and such weight vector is unique. Indeed, suppose there exists $\boldsymbol{\alpha}' = (\alpha'_1, \ldots, \alpha'_m)$ with $\alpha'_k \geq 0$ and $\sum_{k=1}^m \alpha'_k = 1$, such that $D^p_{\boldsymbol{\alpha}'}$ represents $\precsim$ on the whole $\mathcal{Y}^2$ and $\boldsymbol{\alpha}' \neq \boldsymbol{\alpha}$. For $k = 1, \ldots, m$, it holds that

$$(E_k, \underline{0}) \sim (\underline{\beta_k}, \underline{0}) \Longleftrightarrow W^p_{\boldsymbol{\alpha}'}(E_k, \underline{0}) = \alpha'_k = \alpha_k = W^p_{\boldsymbol{\alpha}'}(\underline{\beta_k}, \underline{0}),$$

reaching in this way a contradiction. □

In the particular case of a uniform distribution of weights $\boldsymbol{\alpha}_u = \left(\frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m}\right)$, the weighted L^p distance $D^p_{\boldsymbol{\alpha}_u}$ turns out to be a strictly increasing transformation of the *unweighted* L^p *distance* (also known as *Minkowski distance* of order p) defined, for all $X, Y \in \mathcal{Y}$, as

$$D^p(X, Y) = \left(\sum_{k=1}^m |x_k - y_k|^p\right)^{\frac{1}{p}}. \tag{5}$$

Therefore, for every $(X, Y), (X', Y') \in \mathcal{Y}^2$, it holds that

$$D^p_{\boldsymbol{\alpha}_u}(X, Y) \leq D^p_{\boldsymbol{\alpha}_u}(X', Y') \Longleftrightarrow D^p(X, Y) \leq D^p(X', Y'). \tag{6}$$

In particular, for $p = 1$ we get the *Manhattan distance*, and for $p = 2$ we get the *Euclidean distance*.

It turns out that, to get an unweighted L^p distance representation, in the presence of **(R-**p**)** and **(FD5)**, it is necessary and sufficient to add one of the following two axioms:

(U1) Denoting by E_k the element of $\mathcal{Y}$ whose k-th component is 1 and the others are 0, it holds that

$$(E_1, \underline{0}) \sim (E_2, \underline{0}) \sim \cdots \sim (E_m, \underline{0});$$

(U2) For every $X, Y \in \mathcal{Y}$ and every $i, j \in I$ with $i < j$, denoting by $X^{ij} = (x_1, \ldots, x_j, \ldots, x_i, \ldots, x_m)$ and $Y^{ij} = (y_1, \ldots, y_j, \ldots, y_i, \ldots, y_m)$, it holds that

$$(X, Y) \sim (X^{ij}, Y^{ij}).$$

Theorem 9 *Let* $p \in [1, +\infty)$. *Let* $\precsim$ *be a nontrivial complete comparative dissimilarity relation on* $\mathcal{Y}^2$. *Then, the following statements are equivalent:*

(i) $\precsim$ *satisfies* **(R-**p**)**, **(FD5)** *and* **(U1)**;
(ii) $\precsim$ *satisfies* **(R-**p**)**, **(FD5)** *and* **(U2)**;
(iii) for all $(X, Y), (X', Y') \in \mathcal{Y}^2$, *it holds*

$$(X, Y) \precsim (X', Y') \Longleftrightarrow D^p(X, Y) \leq D^p(X', Y').$$

Proof By (6), condition *(iii)* is equivalent to the representability of $\precsim$ by a weighted L^p distance $D^p_{\boldsymbol{\alpha}_u}$ with uniform weight vector $\boldsymbol{\alpha}_u$. In turn, *(iii)* $\Longleftrightarrow$ *(i)* directly comes from the proof of Theorem 8 since, for any weight vector $\boldsymbol{\alpha}$, $W^p_{\boldsymbol{\alpha}}(E_k, \underline{0}) = \alpha_k$. The implication *(iii)* $\Longrightarrow$ *(ii)* is easily proved, thus we only prove *(ii)* $\Longrightarrow$ *(iii)*. If $\precsim$ is representable by a weighted L^p distance $D^p_{\boldsymbol{\alpha}}$ with a weight vector $\boldsymbol{\alpha}$, then

$$\begin{aligned}(X, Y) \sim (X^{ij}, Y^{ij}) &\Longleftrightarrow W^p_{\boldsymbol{\alpha}}(X, Y) = W^p_{\boldsymbol{\alpha}}(X^{ij}, Y^{ij}) \\ &\Longleftrightarrow \alpha_i(|x_i - y_i|^p - |x_j - y_j|^p) \\ &\qquad + \alpha_j(|x_j - y_j|^p - |x_i - y_i|^p) = 0,\end{aligned}$$

and since this holds for every $X, Y \in \mathcal{Y}$ and every $i, j \in I$ with $i < j$, we get that $\alpha_1 = \alpha_2 = \ldots = \alpha_m$, that is $\boldsymbol{\alpha} = \boldsymbol{\alpha}_u$. □

4.3 Almost representability theorems

Here, we weaken the requirement of representability by a weighted L^p distance, only requiring almost representability. For this purpose, we consider the following weaker condition.

(AR-p) For all $n \in \mathbb{N}$, for all $(X_1, Y_1), \ldots, (X_n, Y_n), (X'_1, Y'_1), \ldots, (X'_n, Y'_n) \in \mathcal{Y}^2$ with $(X_i, Y_i) \precsim (X'_i, Y'_i)$, $i = 1, \ldots, n$, there are no $\lambda_1, \ldots, \lambda_n > 0$ with $\sum_{i=1}^n \lambda_i = 1$ such that:

$$\sum_{i=1}^n \lambda_i |X'_i - Y'_i|^p < \sum_{i=1}^n \lambda_i |X_i - Y_i|^p.$$

Let us stress that, for a complete relation $\precsim$ on a finite $\mathcal{F} \subset \mathcal{Y}^2$ satisfying **(AR-p)**, if $|X - Y|^p < |X' - Y'|^p$, then it must be $(X, Y) \prec (X', Y')$. In other terms, condition **(AR-p)** implies non-triviality, contrary to **(R-p)** for which non-triviality must be explicitly required. Thus, for instance, if $(\underline{0}, \underline{0}), (\underline{1}, \underline{0}) \in \mathcal{F}$, then condition **(AR-$p$)** forces us to set $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$ and so $\precsim$ must be nontrivial.

Theorem 10 *Let $p \in [1, +\infty)$. Let $\precsim$ be a nontrivial complete comparative dissimilarity relation on a finite $\mathcal{F} \subset \mathcal{Y}^2$. Then, the following statements are equivalent:*

(i) *$\precsim$ satisfies* **(AR-p)***;*
(ii) *there exists a weight vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)$ with $\alpha_k \geq 0$, $k = 1, \ldots, m$, and $\sum_{k=1}^m \alpha_k = 1$ such that, for all $(X, Y), (X', Y') \in \mathcal{F}$, it holds that*

$$(X, Y) \precsim (X', Y') \Longrightarrow D^p_{\boldsymbol{\alpha}}(X, Y) \leq D^p_{\boldsymbol{\alpha}}(X', Y').$$

Proof Since $\mathcal{F}$ is finite, the binary relation $\precsim$ amounts to a finite number of comparisons. Consider the set

$$\mathcal{W} = \{((X, Y), (X', Y')) \in \mathcal{F}^2 : (X, Y) \precsim (X', Y')\},$$

with $w = \text{card}\,\mathcal{W}$, and fix an enumeration $\mathcal{W} = \{((X_h, Y_h), (X'_h, Y'_h))\}_{h \in H}$ with $H = \{1, \ldots, w\}$.

Condition *(ii)* is equivalent to the solvability of the following linear system

$$\begin{cases} \mathbf{A}\boldsymbol{\beta} \geq \mathbf{0}, \\ \boldsymbol{\beta} \geq \mathbf{0}, \\ \boldsymbol{\beta} \neq \mathbf{0}, \end{cases}$$

with unknown $\boldsymbol{\beta} \in \mathbb{R}^{m \times 1}$, and $\mathbf{A} \in \mathbb{R}^{w \times m}$, where the w rows of $\mathbf{A}$ are the vectors $|X'_h - Y'_h|^p - |X_h - Y_h|^p$, for all $h \in H$. Indeed, if we have a weight vector $\boldsymbol{\alpha}$ satisfying *(ii)*, then setting $\boldsymbol{\beta} = \boldsymbol{\alpha}^T$ we get a solution of the above system. On the converse, if $\boldsymbol{\beta}$ is a solution of the above system, then defining $\alpha_k = \frac{\beta_k}{\sum_{i=1}^m \beta_i}$, we get a weight vector $\boldsymbol{\alpha}$ satisfying *(ii)*.

By the Gale's theorem of the alternative [18, 24], the solvability of the above system is equivalent to the non-solvability of the following system

$$\begin{cases} \boldsymbol{\mu}\mathbf{A} < \mathbf{0}, \\ \boldsymbol{\mu} \geq \mathbf{0}, \end{cases}$$

with unknown $\boldsymbol{\mu} \in \mathbb{R}^{1\times w}$. In particular, the first inequality reduces to

$$\sum_{h\in H} \mu_h(|X'_h - Y'_h|^p - |X_h - Y_h|^p) < \mathbf{0}.$$

Therefore, dividing both sides of the inequality by the sum of the μ_h's, it easy to see that the non-solvability of the above system is equivalent to condition **(AR-**p**)**. □

It is immediate to verify that condition **(R-**p**)** implies **(AR-**p**)** under non-triviality, but the converse does not hold. To see this, we provide the following toy example, inspired by a well-known example given in [21].

Example 5 Let $p \in [1, +\infty)$. Take $m = 5$, that is $I = \{1, \ldots, 5\}$, and consider the finite set $\mathcal{F} = \{(X_i, \underline{0}) \,:\, i = 1, \ldots, 6\}$ where

	h_1	h_2	h_3	h_4	h_5
X_1	1	0	1	0	0
X_2	0	1	0	0	1
X_3	1	1	0	0	0
X_4	0	0	1	1	0
X_5	0	1	1	0	0
X_6	1	0	0	0	0

Consider the comparisons $(X_1, \underline{0}) \prec (X_2, \underline{0})$, $(X_3, \underline{0}) \prec (X_4, \underline{0})$, $(X_5, \underline{0}) \prec (X_6, \underline{0})$ that can be extended to the nontrivial weak order $\precsim$ on $\mathcal{F}$ such that

$$(X_5, \underline{0}) \prec (X_6, \underline{0}) \prec (X_1, \underline{0}) \sim (X_3, \underline{0}) \prec (X_2, \underline{0}) \sim (X_4, \underline{0}).$$

Taking $\boldsymbol{\alpha} = (0.2, 0.1, 0.1, 0.3, 0.3)$ we get

$$D^p_{\boldsymbol{\alpha}}(X_1, \underline{0}) = 0.3^{\frac{1}{p}},$$
$$D^p_{\boldsymbol{\alpha}}(X_2, \underline{0}) = 0.4^{\frac{1}{p}},$$
$$D^p_{\boldsymbol{\alpha}}(X_3, \underline{0}) = 0.3^{\frac{1}{p}},$$
$$D^p_{\boldsymbol{\alpha}}(X_4, \underline{0}) = 0.4^{\frac{1}{p}},$$
$$D^p_{\boldsymbol{\alpha}}(X_5, \underline{0}) = 0.2^{\frac{1}{p}},$$
$$D^p_{\boldsymbol{\alpha}}(X_6, \underline{0}) = 0.2^{\frac{1}{p}},$$

thus $D^p_{\boldsymbol{\alpha}}$ is easily seen to almost represent $\precsim$, therefore $\precsim$ satisfies **(AR-**p**)** by virtue of Theorem 10. Notice $D^p_{\boldsymbol{\alpha}}$ does not represent $\precsim$ since $(X_5, \underline{0}) \prec (X_6, \underline{0})$ but $D^p_{\boldsymbol{\alpha}}(X_5, \underline{0}) = D^p_{\boldsymbol{\alpha}}(X_6, \underline{0})$.

It actually holds that for the relation $\precsim$ there is no weight vector $\boldsymbol{\alpha}$ such that $D^p_{\boldsymbol{\alpha}}$ represents it. Indeed, since the following system

$$\begin{cases} \alpha_1 + \alpha_3 < \alpha_2 + \alpha_5, \\ \alpha_1 + \alpha_2 < \alpha_3 + \alpha_4, \\ \alpha_2 + \alpha_3 < \alpha_1, \\ \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 = 1, \\ \alpha_k \geq 0,\ k = 1, \dots, 5, \end{cases}$$

has no solution, the three comparisons $(X_1, \underline{0}) \prec (X_2, \underline{0})$, $(X_3, \underline{0}) \prec (X_4, \underline{0})$, $(X_5, \underline{0}) \prec (X_6, \underline{0})$ cannot be represented simultaneously by any weighted L^p distance. In turn, this implies that $\precsim$ does not satisfy **(R-p)** by virtue of Theorem 7. ♦

Contrarily to the representability case, if we have a nontrivial complete relation $\precsim$ on the whole $\mathcal{Y}^2$, condition **(AR-p)** alone turns out to be necessary and sufficient to the almost representability of $\precsim$ with a weighted L^p distance.

Theorem 11 *Let $p \in [1, +\infty)$. Let $\precsim$ be a nontrivial complete comparative dissimilarity relation on $\mathcal{Y}^2$. Then, the following statements are equivalent:*

(i) *$\precsim$ satisfies* **(AR-p)***;*
(ii) *there exists a weight vector $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m)$ with $\alpha_k \geq 0$, $k = 1, \dots, m$, and $\sum_{k=1}^m \alpha_k = 1$ such that, for all $(X, Y), (X', Y') \in \mathcal{Y}^2$, it holds that*

$$(X, Y) \precsim (X', Y') \Longrightarrow D^p_{\boldsymbol{\alpha}}(X, Y) \leq D^p_{\boldsymbol{\alpha}}(X', Y').$$

Proof The implication *(ii)* $\Longrightarrow$ *(i)* follows by Theorem 10. Thus, suppose that *(i)* holds. Theorem 10 implies, for every finite $\mathcal{G} \subseteq \mathcal{Y}^2$ and $\mathcal{F} = \mathcal{G} \cup \{(\underline{0}, \underline{0}), (\underline{1}, \underline{0})\}$, the compatibility of the following system

$$\begin{cases} \mathbf{A}\boldsymbol{\alpha} \geq \mathbf{0}, \\ \boldsymbol{\alpha} \geq \mathbf{0}, \\ \sum_{k=1}^m \alpha_k = 1, \end{cases}$$

where $\boldsymbol{\alpha} \in \mathbb{R}^{m \times 1}$ is the unknown and $\mathbf{A} \in \mathbb{R}^{w \times m}$ is defined as in the proof of Theorem 10. Let $\mathcal{A}_{\mathcal{F}}$ be the set of solutions of the above system, which is easily seen to be a non-empty closed subset of the compact space $[0, 1]^m$ endowed with the product topology. The family

$$\mathfrak{A} = \left\{\mathcal{A}_{\mathcal{F}} \,:\, \mathcal{F} = \mathcal{G} \cup \{(\underline{0}, \underline{0}), (\underline{1}, \underline{0})\},\ \text{finite } \mathcal{G} \subseteq \mathcal{Y}^2\right\},$$

is easily shown to possess the finite intersection property, therefore $\bigcap \mathfrak{A} \neq \emptyset$, and this implies the existence of a weight vector $\boldsymbol{\alpha} \in \bigcap \mathfrak{A}$ whose corresponding $D^p_{\boldsymbol{\alpha}}$ satisfies *(ii)*. □

Let us stress that Theorem 11 does not assure the uniqueness of the weight vector $\boldsymbol{\alpha}$. The uniqueness is achieved if we further require condition **(FD5)**.

Theorem 12 *Let $p \in [1, +\infty)$. Let $\precsim$ be a nontrivial complete comparative dissimilarity relation on $\mathcal{Y}^2$. Then, the following statements are equivalent:*

(i) *$\precsim$ satisfies* **(AR-p)** *and* **(FD5)***;*
(ii) *there exists a weight vector $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m)$ with $\alpha_k \geq 0$, $k = 1, \dots, m$, and $\sum_{k=1}^m \alpha_k = 1$ such that, for all $(X, Y), (X', Y') \in \mathcal{Y}^2$, it holds that*

$$(X, Y) \precsim (X', Y') \Longrightarrow D^p_{\boldsymbol{\alpha}}(X, Y) \leq D^p_{\boldsymbol{\alpha}}(X', Y').$$

Moreover, the weight vector $\boldsymbol{\alpha}$ is unique.

Proof The existence of $\boldsymbol{\alpha}$ follows by Theorem 11 while uniqueness is proved as in the proof of Theorem 8, since almost representability preserves all equivalences in comparative dissimilarity. □

5 Representation by a Choquet L^p distance

One of the main drawbacks of weighted L^p distances is the fact that they cannot consider interactions among attributes. Inspired by considerations on similarity measures [1, 11], we can generalize weighted L^p distances by referring to the Choquet integral computed with respect to a completely alternating normalized capacity. We recall that a set function $\nu : 2^I \to [0, 1]$ is a *completely alternating normalized capacity* if it satisfies:

(i) $\nu(\emptyset) = 0$ and $\nu(I) = 1$;
(ii) for all $n \geq 2$ and $A_1, \ldots, A_n \in 2^I$ it holds that

$$\nu\left(\bigcap_{j=1}^{n} A_j\right) \leq \sum_{\emptyset \neq J \subseteq \{1,\ldots,n\}} (-1)^{|J|+1} \nu\left(\bigcup_{j \in J} A_j\right).$$

We notice that a completely alternating normalized capacity is *additive* if it satisfies condition *(ii)* as an equality. In this case, ν is completely characterized by its values on the atoms of the algebra 2^I that can be identified with the vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)$.

As is well-known (see [20]), every completely alternating normalized capacity ν is characterized by the *Möbius inverse* of its dual capacity, which is a set function $\pi : 2^I \to [0, 1]$ such that,

$$\pi(\emptyset) = 0, \quad \sum_{B \in 2^I} \pi(B) = 1 \quad \text{and} \quad \nu(A) = \sum_{B \cap A \neq \emptyset} \pi(B), \quad \text{for all } A \in 2^I. \tag{7}$$

Notice that the function π allows us to attach a weight to groups of features, thus it can be used to model interactions among some of them.

We recall (see [20]) that, for all $X \in \mathcal{Y}$, the *Choquet integral* of X with respect to ν is defined as

$$\oint X \, \mathrm{d}\nu = \sum_{k=1}^{m} (x_{\sigma(k)} - x_{\sigma(k+1)}) \nu(A_k^\sigma), \tag{8}$$

where σ is a permutation of I such that $x_{\sigma(1)} \geq \ldots \geq x_{\sigma(m)}$, $A_k^\sigma = \{\sigma(1), \ldots, \sigma(k)\}$ for $k = 1, \ldots, m$, and $x_{\sigma(m+1)} = 0$.

Definition 3 Let $p \in [1, +\infty)$. A **Choquet L^p distance** is a function $D_\nu^p : \mathcal{Y}^2 \to [0, 1]$ parameterized by a completely alternating normalized capacity ν, defined, for every $X, Y \in \mathcal{Y}$, as

$$D_\nu^p(X, Y) = \left(\oint |X - Y|^p \, \mathrm{d}\nu\right)^{\frac{1}{p}}.$$

The corresponding **p-th power Choquet L^p distance** is a function $W_\nu^p : \mathcal{Y}^2 \to [0, 1]$ defined, for every $X, Y \in \mathcal{Y}$, as

$$W_\nu^p(X, Y) = (D_\nu^p(X, Y))^p = \oint |X - Y|^p \, \mathrm{d}\nu.$$

Also in this case, for all $X, Y \in \mathcal{Y}$, it holds that

$$D_\nu^p(X,Y) \leq D_\nu^p(X',Y') \Longleftrightarrow W_\nu^p(X,Y) \leq W_\nu^p(X',Y'). \tag{9}$$

Moreover, as shown in [17], for a fixed ν, defining the relation $X \equiv_0 Y \Longleftrightarrow D_\nu^p(X,Y) = 0$, then D_ν^p turns out to be a metric on the quotient space $\mathcal{Y}_{/\equiv_0}$.

Remark 2 As follows from [17], to get a Choquet L^p distance D_ν^p, it is sufficient that ν is 2-alternating, i.e., it satisfies the above condition *(ii)* only for $n = 2$. Nevertheless, we stick to completely alternating capacities since they are sufficiently general to accommodate (as shown below) also the unweighted L^∞ distance and, at the same time, they can be characterized by an axiom that is simpler to justify.

Let us notice that in case of an additive ν, a Choquet L^p distance reduces to a weighted L^p distance according to Definition 2, parameterized by the vector $\boldsymbol{\alpha}$ of the values of ν on the atoms of 2^I. More generally, introducing

$$\mathcal{C}_\nu = \left\{\boldsymbol{\alpha} \in [0,1]^m \ : \ \sum_{k \in A} \alpha_k \leq \nu(A), \text{ for all } A \in 2^I, \ \sum_{k=0}^{m} \alpha_k = 1\right\}, \tag{10}$$

which is the *anti-core* induced by ν (see [20]), for all $X, Y \in \mathcal{Y}$, we have that

$$D_\nu^p(X,Y) = \left(\max_{\boldsymbol{\alpha} \in \mathcal{C}_\nu} W_{\boldsymbol{\alpha}}^p(X,Y)\right)^{\frac{1}{p}} = \max_{\boldsymbol{\alpha} \in \mathcal{C}_\nu} D_{\boldsymbol{\alpha}}^p(X,Y). \tag{11}$$

We point out that for $p = 1$ and ν_∞ defined, for all $A \in 2^I$, as

$$\nu_\infty(A) = \begin{cases} 0 \text{ if } A = \emptyset, \\ 1 \text{ otherwise}, \end{cases} \tag{12}$$

then $D_{\nu_\infty}^1$ reduces to the *unweighted* L^∞ *distance* (also known as *Čebyšëv distance*) defined, for all $X, Y \in \mathcal{Y}$, as

$$D_{\nu_\infty}^1(X,Y) = D^\infty(X,Y) = \max_{k=1,\ldots,m} |x_k - y_k|. \tag{13}$$

We notice that

$$\mathcal{C}_{\nu_\infty} = \mathcal{A} := \left\{\boldsymbol{\alpha} \in [0,1]^m \ : \ \sum_{k=0}^{m} \alpha_k = 1\right\}, \tag{14}$$

where $\mathcal{A}$ is the set of all possible weight vectors, therefore, by the properties of the Choquet integral (see, e.g., [17]), we have that the unweighted L^∞ distance can be expressed, for all $X, Y \in \mathcal{Y}$, as

$$D^\infty(X,Y) = \lim_{p \to +\infty} D^p(X,Y) = \max_{\boldsymbol{\alpha} \in \mathcal{A}} D_{\boldsymbol{\alpha}}^1(X,Y). \tag{15}$$

5.1 Rationality principle

Below we report the normative condition that must be accepted if we want that the function representing our comparative dissimilarity $\precsim$ belongs to the particular subclass of Choquet L^p distances with respect to a completely alternating normalized capacity.

(CR-p) For all $n \in \mathbb{N}$, for all $(X_1, Y_1), \ldots, (X_n, Y_n), (X'_1, Y'_1), \ldots, (X'_n, Y'_n) \in \mathcal{Y}^2$ with $(X_i, Y_i) \precsim (X'_i, Y'_i)$, $i = 1, \ldots, n-1$, and $(X_n, Y_n) \prec (X'_n, Y'_n)$, setting $V_i = |X_i - Y_i|^p$ and $V'_i = |X'_i - Y'_i|^p$, there are no $\lambda_1, \ldots, \lambda_n > 0$ with $\sum_{i=1}^n \lambda_i = 1$, such that:

$$\sum_{i=1}^{n} \lambda_i \max_{k \in B} v_k^{i\,\prime} \leq \sum_{i=1}^{n} \lambda_i \max_{k \in B} v_k^i, \quad \text{for all } B \in 2^I \setminus \{\emptyset\}.$$

In analogy with **(R-p)**, axiom **(CR-p)** does not imply non-triviality as it is vacuously satisfied if there are no strict comparisons. Thus, we need to explicitly assume that $\precsim$ is nontrivial. In this case, an analog of Proposition 5 holds.

Proposition 13 *Let $p \in [1, +\infty)$. Let $\precsim$ be a nontrivial complete comparative dissimilarity relation on $\mathcal{Y}^2$ satisfying* **(CR-p)**. *Then, for all $X, X', Y, Y' \in \mathcal{Y}$, the following statements hold:*

(i) it must be $(X, Y) \sim (|X - Y|, \underline{0})$;
(ii) if $|X - Y| \leq |X' - Y'|$, then it must be $(X, Y) \precsim (X', Y')$;
(iii) it must be $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$.

Proof Statement *(i)*. Since $V = |X - Y|^p = ||X - Y| - \underline{0}|^p = V'$, we get that

$$\max_{k \in B} v_k{}' = \max_{k \in B} v_k, \quad \text{for all } B \in 2^I \setminus \{\emptyset\}.$$

Therefore, by **(CR-p)** it cannot be neither $(X, Y) \prec (|X - Y|, \underline{0})$, nor $(|X - Y|, \underline{0}) \prec (X, Y)$, and since $\precsim$ is complete, it must be $(X, Y) \sim (|X - Y|, \underline{0})$.

Statement *(ii)*. If $|X - Y| \leq |X' - Y'|$ and $(X', Y') \prec (X, Y)$, then setting $V = |X - Y|^p \leq |X' - Y'|^p = V'$ we get that

$$\max_{k \in B} v_k \leq \max_{k \in B} v_k{}', \quad \text{for all } B \in 2^I \setminus \{\emptyset\},$$

and **(CR-p)** is violated, thus it cannot be $(X', Y') \prec (X, Y)$. Then, the statement follows by the completeness of $\precsim$.

Statement *(ii)*. By non-triviality, there are $(X, Y), (X', Y') \in \mathcal{Y}^2$ such that $(X, Y) \prec (X', Y')$. Notice that it cannot be $(\underline{1}, \underline{0}) \prec (\underline{0}, \underline{0})$ by statement *(ii)*. Suppose $(\underline{1}, \underline{0}) \precsim (\underline{0}, \underline{0})$. Let $V_1 = |\underline{1} - \underline{0}|^p$, $V_2 = |X - Y|^p$, $V'_1 = |\underline{0} - \underline{0}|^p$, $V'_2 = |X' - Y'|^p$. Denote by $q = \max_{B \in 2^I \setminus \{\emptyset\}} \left\{ \max_{k \in B} v_k^{2\prime} - \max_{k \in B} v_k^2 \right\}$ and take $\lambda_1, \lambda_2 \in (0, 1)$ with $\lambda_1 \geq \frac{q}{1+q}$ and $\lambda_2 = 1 - \lambda_1$. Then we have

$$\sum_{i=1}^{2} \lambda_i \max_{k \in B} v_k^{i\,\prime} \leq \sum_{i=1}^{2} \lambda_i \max_{k \in B} v_k^i, \quad \text{for all } B \in 2^I \setminus \{\emptyset\},$$

violating condition **(CR-p)**. Thus, **(CR-p)** requires to set $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$. □

Axiom **(CR-p)** asserts that if you have n pairs (X_i, Y_i) of fuzzy profiles and you judge the elements of each of them no more dissimilar than those of other n pairs (X'_i, Y'_i), with at least one strict comparison, combining in a positive convex combination the maxima of $|X_i - Y_i|^p$'s and the maxima of $|X'_i - Y'_i|^p$'s on each non-empty subset of features, you cannot obtain a situation of weak dominance, uniformly over $2^I \setminus \{\emptyset\}$.

Also in this case we have that condition **(CR-p)** is stronger than **(FD1)-(FD4)** under non-triviality and completeness.

Theorem 14 *Let $p \in [1, +\infty)$. Let $\precsim$ be a nontrivial complete comparative dissimilarity relation on $\mathcal{Y}^2$ satisfying* **(CR-p)**. *Then $\precsim$ is transitive and axioms* **(FD1)–(FD4)** *hold.*

Proof To prove transitivity, for all $(X, Y), (X', Y'), (X'', Y'') \in \mathcal{Y}^2$, let $U = |X - Y|^p$, $V = |X' - Y'|^p$ and $Z = |X'' - Y''|^p$. For every $B \in 2^I \setminus \{\emptyset\}$, then we have

$$\frac{1}{3}\left(\max_{k\in B} u_k + \max_{k\in B} v_k + \max_{k\in B} z_k\right) = \frac{1}{3}\left(\max_{k\in B} v_k + \max_{k\in B} z_k + \max_{k\in B} u_k\right).$$

Hence, if $(X, Y) \precsim (X', Y')$, $(X', Y') \precsim (X'', Y'')$ and $(X'', Y'') \precsim (X, Y)$, then **(CR-**$p$**)** implies that none of the comparisons can be strict, thus we get transitivity.

To prove **(FD1)** and **(FD2)** it is sufficient to note that $|X - Y| = |X_k^{-\varepsilon} - Y_k^{-\varepsilon}| = |X'_k - Y'_k|$, for all $X, Y \in \mathcal{Y}$. Thus, statement *(i)* of Proposition 13 and transitivity of $\precsim$ imply that $(X, Y) \sim (X_k^{-\varepsilon}, Y_k^{-\varepsilon})$ and $(X, Y) \sim (X'_k, Y'_k)$.

To prove **(FD3)** notice that $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$ by statement *(iii)* of Proposition 13. Condition *a)* follows since $\underline{0} = |X - X| = |Y - Y| \leq |X - Y|$, for all $X, Y \in \mathcal{Y}$, thus, by statement *(ii)* of Proposition 13 it must be $(X, X) \precsim (X, Y)$ and $(Y, Y) \precsim (X, Y)$. Moreover, it cannot be neither $(X, Y) \precsim (X, X)$ nor $(X, Y) \precsim (Y, Y)$ if $x_k \neq y_k$, for all $k \in I$, as we have $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$. Indeed, setting $U = |X - Y|^p$, $V = |X - X|^p$, $Z = |Y - Y|^p$, $W = |\underline{0} - \underline{0}|^p$, $R = |\underline{1} - \underline{0}|^p$, and denoting by $q = \min\limits_{B\in 2^I\setminus\{\emptyset\}} \max\limits_{k\in B} u_k$, for all $B \in 2^I \setminus \{\emptyset\}$, we get

$$\begin{aligned}
\frac{1}{1+q}\max_{k\in B} v_k + \frac{q}{1+q}\max_{k\in B} r_k &\leq \frac{1}{1+q}\max_{k\in B} u_k + \frac{q}{1+q}\max_{k\in B} w_k,\\
\frac{1}{1+q}\max_{k\in B} z_k + \frac{q}{1+q}\max_{k\in B} r_k &\leq \frac{1}{1+q}\max_{k\in B} u_k + \frac{q}{1+q}\max_{k\in B} w_k,
\end{aligned}$$

that contradict **(CR-**p**)**. Analogously, condition *b)* follows since $\underline{1} = |X - X^c| \geq |X - Y|$, for all $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$, that implies $(X, Y) \precsim (X, X^c)$ by statement *(ii)* of Proposition 13. Moreover, it cannot be $(X, X^c) \precsim (X, Y)$ if $s_X \subseteq s_Y$ and $y_k < 1$, for all $k \in s_Y \setminus s_X$, as we have $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$. Indeed, setting $V_1 = |X - X^c|^p$, $V'_1 = |X - Y|^p$, $V_2 = |\underline{0} - \underline{0}|^p$, $V'_2 = |\underline{1} - \underline{0}|^p$, and denoting by $q = 1 - \max\limits_{B\in 2^I\setminus\{\emptyset\}} \max\limits_{k\in B} v_k^{1'}$, we get

$$\frac{1}{1+q}\max_{k\in B} v_k^{1'} + \frac{q}{1+q}\max_{k\in B} v_k^{2'} \leq \frac{1}{1+q}\max_{k\in B} v_k^{1} + \frac{q}{1+q}\max_{k\in B} v_k^{2},$$

that contradicts **(CR-**p**)**.

Finally, axiom **(FD4)** holds since $|X - Y| \leq |X_k^{-\varepsilon} - Y|$ and $|X - Y| \leq |X - Y_k^{\eta}|$, for all $X, Y \in \mathcal{Y}$, for all $k \in I$, such that $x_k \leq y_k$, for all $0 < \varepsilon \leq x_k$ and $0 < \eta \leq 1 - y_k$. Thus, by statement *(ii)* of Proposition 13 it must be $(X, Y) \precsim (X_k^{-\varepsilon}, Y)$ and $(X, Y) \precsim (X, Y_k^{\eta})$. □

As it happens with axiom **(R-**p**)**, axiom **(CR-**p**)**, alone, guarantees representability by a Choquet L^p distance only if $\precsim$ is defined on a finite set of fuzzy description profiles.

Theorem 15 *Let* $p \in [1, +\infty)$. *Let* $\precsim$ *be a nontrivial complete comparative dissimilarity relation on a finite* $\mathcal{F} \subset \mathcal{Y}^2$. *Then, the following statements are equivalent:*

(i) $\precsim$ *satisfies* **(CR-**p**)***;*

(ii) there exists a completely alternating normalized capacity ν *such that, for all* $(X, Y), (X', Y') \in \mathcal{F}$*, it holds that*

$$(X, Y) \precsim (X', Y') \Longleftrightarrow D_\nu^p(X, Y) \leq D_\nu^p(X', Y').$$

Proof Since $\mathcal{F}$ is finite, the binary relation $\precsim$ amounts to a finite number of comparisons. Consider the sets

$$\begin{aligned}
\mathcal{S} &= \{((X, Y), (X', Y')) \in \mathcal{Y}^2 : (X, Y) \prec (X', Y')\},\\
\mathcal{W} &= \{((X, Y), (X', Y')) \in \mathcal{Y}^2 \setminus \mathcal{S} : (X, Y) \precsim (X', Y')\},
\end{aligned}$$

with $s = \text{card}\,\mathcal{S}$ and $w = \text{card}\,\mathcal{W}$, and fix two enumerations $\mathcal{S} = \{((X_j, Y_j), (X'_j, Y'_j))\}_{j\in J}$ and $\mathcal{W} = \{((X_h, Y_h), (X'_h, Y'_h))\}_{h\in H}$ with $J = \{1, \ldots, s\}$ and $H = \{1, \ldots, w\}$.

Let π be the Möbius inverse associated to ν through (7). Fixing the enumeration $2^I \setminus \{\emptyset\} = \{B_1, \ldots, B_{2^m-1}\}$, for every $X \in \mathcal{Y}$, we have (see [20])

$$\oint X\,\mathrm{d}\nu = \sum_{i=1}^{2^m-1} \left(\max_{k\in B_i} x_k\right) \pi(B_i).$$

Define $V'_j = |X'_j - Y'_j|^p$ and $V_j = |X_j - Y_j|^p$, for all $j \in J$, and $U'_h = |X'_h - Y'_h|^p$ and $U_h = |X_h - Y_h|^p$, for all $h \in H$.

Condition *(ii)* is equivalent to the solvability of the following linear system

$$\begin{cases} \mathbf{A}\boldsymbol{\beta} > \mathbf{0}, \\ \mathbf{B}\boldsymbol{\beta} \geq \mathbf{0}, \\ \boldsymbol{\beta} \geq \mathbf{0}, \end{cases}$$

with unknown $\boldsymbol{\beta} \in \mathbb{R}^{(2^m-1)\times 1}$, and $\mathbf{A} \in \mathbb{R}^{s\times(2^m-1)}$ and $\mathbf{B} \in \mathbb{R}^{w\times(2^m-1)}$, where $\mathbf{A} = (a_{ji})$ and $\mathbf{B} = (b_{hi})$ are such that

$$\begin{aligned} a_{ji} &= \max_{k\in B_i} v_k^{j'} - \max_{k\in B_i} v_k^{j}, \\ b_{hi} &= \max_{k\in B_i} u_k^{h'} - \max_{k\in B_i} u_k^{h}. \end{aligned}$$

Indeed, if we have a completely alternating normalized capacity ν satisfying *(ii)*, the corresponding Möbius inverse π is such that setting $\beta_i = \pi(B_i)$ we get a solution of the above system. On the converse, if $\boldsymbol{\beta}$ is a solution of the above system, then defining $\pi(B_i) = \frac{\beta_i}{\sum_{l=1}^{2^m-1} \beta_l}$, we get a Möbius inverse π inducing a completely alternating normalized capacity ν that satisfies *(ii)*.

By the Motzkin's theorem of the alternative [24], the solvability of the above system is equivalent to the non-solvability of the following system

$$\begin{cases} \boldsymbol{\mu}\mathbf{A} + \boldsymbol{\nu}\mathbf{B} \leq \mathbf{0}, \\ \boldsymbol{\mu}, \boldsymbol{\nu} \geq \mathbf{0}, \\ \boldsymbol{\mu} \neq \mathbf{0}, \end{cases}$$

with unknowns $\boldsymbol{\mu} \in \mathbb{R}^{1\times s}$ and $\boldsymbol{\nu} \in \mathbb{R}^{1\times w}$. In particular, the first inequality reduces, for $i = 1, \ldots, 2^m - 1$, to

$$\sum_{j\in J} \mu_j \left(\max_{k\in B_i} v_k^{j'} - \max_{k\in B_i} v_k^{j}\right) + \sum_{h\in H} \nu_h \left(\max_{k\in B_i} u_k^{h'} - \max_{k\in B_i} u_k^{h}\right) \leq 0,$$

thus, dividing both sides of the inequality by the sum of the μ_j's and the ν_h's, the non-solvability of the above system is equivalent to condition **(CR-p)**. □

Remark 3 In the hypotheses of previous theorem, if $(X, Y), (X', Y') \in \mathcal{F}$ and it holds $|X - Y|^p < |X' - Y'|^p$ then it must be $(X, Y) \prec (X', Y')$. So, also in this case, if $(\underline{0}, \underline{0}), (\underline{1}, \underline{0}) \in \mathcal{F}$, then it must be $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$, as noted above.

Theorems 7 and 15 imply that every nontrivial complete comparative dissimilarity relation that satisfies **(R-p)** also satisfies **(CR-p)**, but the converse does not hold, as the following example shows.

Example 6 Let $p \in [1, +\infty)$. Take the finite set $\mathcal{F} = \{(X_i, \underline{0}) : i = 1, \ldots, 6\}$ and the nontrivial weak order $\precsim$ on $\mathcal{F}$ of Example 5. It has been shown that such relation does not satisfy **(R-**p**)** as there is no weight vector $\boldsymbol{\alpha}$ such that the corresponding weighted L^p distance represents $\precsim$.

Consider the function $\pi : 2^I \to [0, 1]$ such that $\pi(\{4\}) = \pi(\{5\}) = 0.2$, $\pi(\{4, 5\}) = \pi(\{2, 3, 4, 5\}) = 0.1$, $\pi(\{1, 4, 5\}) = 0.4$ and 0 otherwise. By setting $\nu(A) = \sum_{B \cap A \neq \emptyset} \pi(B)$, for all $A \in 2^I$, we get a completely alternating normalized capacity (see, e.g., [20]). The Choquet L^p distance with respect to ν is such that

$$\begin{aligned}
D^p_\nu(X_1, \underline{0}) &= 0.5^{\frac{1}{p}},\\
D^p_\nu(X_2, \underline{0}) &= 0.8^{\frac{1}{p}},\\
D^p_\nu(X_3, \underline{0}) &= 0.5^{\frac{1}{p}},\\
D^p_\nu(X_4, \underline{0}) &= 0.8^{\frac{1}{p}},\\
D^p_\nu(X_5, \underline{0}) &= 0.1^{\frac{1}{p}},\\
D^p_\nu(X_6, \underline{0}) &= 0.4^{\frac{1}{p}},
\end{aligned}$$

thus D^p_ν represents $\precsim$, which satisfies **(CP-**p**)** by virtue of Theorem 15. ♦

Also in this case, axiom **(FD5)** must be added to **(CR-**p**)** to have Choquet L^p distance representability, when $\precsim$ is defined on the whole $\mathcal{Y}^2$.

Theorem 16 *Let* $p \in [1, +\infty)$. *Let* $\precsim$ *be a nontrivial complete comparative dissimilarity relation on* $\mathcal{Y}^2$. *Then, the following statements are equivalent:*

(i) $\precsim$ *satisfies* **(CR-**p**)** *and* **(FD5)***;*
(ii) there exists a completely alternating normalized capacity ν *such that, for all* $(X, Y), (X', Y') \in \mathcal{Y}^2$, *it holds*

$$(X, Y) \precsim (X', Y') \Longleftrightarrow D^p_\nu(X, Y) \leq D^p_\nu(X', Y').$$

Moreover, the capacity ν *is unique.*

Proof The implication *(ii)* $\Longrightarrow$ *(i)* is easily proven, therefore we only prove *(i)* $\Longrightarrow$ *(ii)*.

For every finite $\mathcal{F} \subset \mathcal{Y}^2$ such that the restriction of $\precsim$ to $\mathcal{F}$ is nontrivial, Theorem 15 implies the existence of a completely alternating normalized capacity $\nu^{\mathcal{F}}$, such that the corresponding $W^p_{\nu^{\mathcal{F}}}$ represents the restriction of $\precsim$ to $\mathcal{F}$. Notice that, by Proposition 13 every finite subset of $\mathcal{Y}^2$ containing $(\underline{0}, \underline{0})$ and $(\underline{1}, \underline{0})$ meets non-triviality, as it must be $(\underline{0}, \underline{0}) \prec (\underline{1}, \underline{0})$.

Fix the enumeration $2^I \setminus \{\emptyset\} = \{B_1, \ldots, B_{2^m-1}\}$. Next, axiom **(FD5)** implies that, for all $(X, Y) \in \mathcal{Y}^2$, there exists $\varepsilon_{(X,Y)} \in [0, 1]$, such that $(X, Y) \sim (\underline{\varepsilon_{(X,Y)}}, \underline{0})$. In particular, for $i = 1, \ldots, 2^m - 1$, denoting by E_{B_i} the element of $\mathcal{Y}$ which is 1 for every $k \in B_i$ and 0 otherwise, we have that there exists $\beta_{B_i} \in [0, 1]$ such that $(E_{B_i}, \underline{0}) \sim (\underline{\beta_{B_i}}, \underline{0})$. Now, for every $(X, Y), (X', Y') \in \mathcal{Y}^2$ we consider the finite subset of $\mathcal{Y}^2$

$$\begin{aligned}
\mathcal{F} = \{&(X, Y), (X', Y'), (\underline{\varepsilon_{(X,Y)}}, \underline{0}), (\underline{\varepsilon_{(X',Y')}}, \underline{0}),\\
&(E_{B_1}, \underline{0}), \ldots, (E_{B_{2^m-1}}, \underline{0}), (\underline{\beta_{B_1}}, \underline{0}), \ldots, (\underline{\beta_{B_{2^m-1}}}, \underline{0}), ,\\
&(\underline{0}, \underline{0}), (\underline{1}, \underline{0})\}.
\end{aligned}$$

By the previous point, there is a completely alternating normalized capacity $\nu^{\mathcal{F}}$ such that

$$(X,Y) \precsim (X',Y') \iff W^p_{\nu^{\mathcal{F}}}(X,Y) \le W^p_{\nu^{\mathcal{F}}}(X',Y'),$$
$$(X,Y) \sim (\underline{\varepsilon_{(X,Y)}},\underline{0}) \iff W^p_{\nu^{\mathcal{F}}}(X,Y) = W^p_{\nu^{\mathcal{F}}}(\underline{\varepsilon_{(X,Y)}},\underline{0}) = \varepsilon^p_{(X,Y)},$$
$$(X',Y') \sim (\underline{\varepsilon_{(X',Y')}},\underline{0}) \iff W^p_{\nu^{\mathcal{F}}}(X',Y') = W^p_{\nu^{\mathcal{F}}}(\underline{\varepsilon_{(X',Y')}},\underline{0}) = \varepsilon^p_{(X',Y')}.$$

Moreover, for all $i = 1,\ldots,2^m - 1$, we have

$$(E_{B_i},\underline{0}) \sim (\underline{\beta_{B_i}},\underline{0}) \iff W^p_{\nu^{\mathcal{F}}}(E_{B_i},\underline{0}) = \nu^{\mathcal{F}}(B_i) = \beta^p_{B_i} = W^p_{\nu^{\mathcal{F}}}(\underline{\beta_{B_i}},\underline{0}),$$

thus the set function $\nu : 2^I \to [0,1]$ defined as $\nu(\emptyset) = 0$ and $\nu(B_i) = \beta^p_{B_i}$, is a completely alternating normalized capacity. So, we get

$$\varepsilon^p_{(X,Y)} = \oint |X-Y|^p \,\mathrm{d}\nu \quad \text{and} \quad \varepsilon^p_{(X',Y')} = \oint |X'-Y'|^p \,\mathrm{d}\nu.$$

and also

$$\varepsilon_{(X,Y)} \le \varepsilon_{(X',Y')} \iff \varepsilon^p_{(X,Y)} \le \varepsilon^p_{(X',Y')}.$$

Hence, there exists a completely alternating normalized capacity ν such that, for all $(X,Y),(X',Y') \in \mathcal{Y}^2$, it holds that

$$(X,Y) \precsim (X',Y') \iff D^p_\nu(X,Y) \le D^p_\nu(X',Y'),$$

and such capacity is unique. Indeed, suppose there exists a completely alternating normalized capacity $\nu' : 2^I \to [0,1]$, such that $D^p_{\nu'}$ represents $\precsim$ on the whole $\mathcal{Y}^2$ and $\nu' \neq \nu$. For $i = 1,\ldots,2^m - 1$, it holds that

$$(E_{B_i},\underline{0}) \sim (\underline{\beta_{B_i}},\underline{0}) \iff W^p_{\nu'}(E_{B_i},\underline{0}) = \nu'(B_i) = \nu(B_i) = W^p_{\nu'}(\underline{\beta_{B_i}},\underline{0}),$$

reaching in this way a contradiction. □

As pointed out before, if the capacity ν coincides with ν_∞ defined as in (12), then $D^1_{\nu_\infty}$ reduces to the unweighted L^∞ distance, according to (13).

It turns out that, to get an unweighted L^∞ distance representation, in the presence of **(CR-1)** and **(FD5)**, it is necessary and sufficient to add one of the following two axioms:

(I1) Denoting by E_B the element of $\mathcal{Y}$ which is 1 for every $k \in B$ and 0 otherwise, it holds that

$$(E_B,\underline{0}) \sim (\underline{1},\underline{0}), \quad \text{for all } B \in 2^I \setminus \{\emptyset\};$$

(I2) For every $X, Y \in \mathcal{Y}$ it holds that

$$(\underline{|x_k - y_k|},\underline{0}) \precsim (X,Y), \quad \text{for } k = 1,\ldots,m,$$

with at least a symmetric comparison.

Theorem 17 *Let $\precsim$ be a nontrivial complete comparative dissimilarity relation on $\mathcal{Y}^2$. Then, the following statements are equivalent:*

(i) $\precsim$ *satisfies* **(CR-1)**, **(FD5)** *and* **(I1)**;
(ii) $\precsim$ *satisfies* **(CR-1)**, **(FD5)** *and* **(I2)**;
(iii) for all $(X,Y),(X',Y') \in \mathcal{Y}^2$, it holds

$$(X,Y) \precsim (X',Y') \iff D^\infty(X,Y) \le D^\infty(X',Y').$$

Proof By (13), condition *(iii)* is equivalent to the representability of $\precsim$ by a Choquet L^1 distance $D^1_{\nu_\infty}$ with capacity ν_∞ defined as in (12). In turn, *(iii)* $\Longleftrightarrow$ *(i)* directly comes from the proof of Theorem 16 since, for any completely alternating normalized capacity ν, $W^1_\nu(E_B, \underline{0}) = \nu(B)$. The implication *(iii)* $\Longrightarrow$ *(ii)* is easily proved, thus we only prove *(ii)* $\Longrightarrow$ *(iii)*. If $\precsim$ is representable by a Choquet L^1 distance D^1_ν with completely alternating normalized capacity ν, then

$$(\underline{|x_k - y_k|}, \underline{0}) \precsim (X, Y) \Longleftrightarrow D^1_\nu(\underline{|x_k - y_k|}, \underline{0}) \leq D^1_\nu(X, Y)$$
$$\Longleftrightarrow |x_k - y_k| \leq D^1_\nu(X, Y),$$

and since this holds for $k = 1, \ldots, m$ with at least an equality, we get $D^1_\nu(X, Y) = \max_{k=1,\ldots,m} |x_k - y_k|$. $\square$

Remark 4 Also in the more general context of Choquet L^p distances with respect to a completely alternating normalized capacity, the demand of representability can be weakened by requiring only almost representability. This is done by referring to the following condition.

(ACR-p) For all $n \in \mathbb{N}$, for all $(X_1, Y_1), \ldots, (X_n, Y_n), (X'_1, Y'_1), \ldots, (X'_n, Y'_n) \in \mathcal{Y}^2$ with $(X_i, Y_i) \precsim (X'_i, Y'_i)$, $i = 1, \ldots, n$, for all $\lambda_1, \ldots, \lambda_n > 0$ with $\sum_{i=1}^n \lambda_i = 1$, setting $V_i = |X_i - Y_i|^p$ and $V'_i = |X'_i - Y'_i|^p$, it does not hold:

$$\sum_{i=1}^{n} \lambda_i \max_{k \in B} v_k^{i\,\prime} < \sum_{i=1}^{n} \lambda_i \max_{k \in B} v_k^i, \quad \text{for all } B \in 2^I \setminus \{\emptyset\}.$$

Proceeding in a similar way to the proof of Theorems 15 and 16, it is easy to prove that analogous versions of Theorems 10, 11 and 12 hold, where **(AR-p)** and $D^p_{\boldsymbol{\alpha}}$ are replaced by **(ACR-p)** and D^p_ν, respectively.

6 Discussion

The axioms introduced so far can be logically divided into five different groups: **(FD0)–(FD4)**; **(Q)** and **(FD5)**; **(R-p)** and **(CR-p)**; **(U1)** and **(U2)**; **(I1)** and **(I2)**.

Axioms **(FD0)–(FD4)** are the most qualitative in nature and are necessary in order to have a representing function $D : \mathcal{Y}^2 \to [0, 1]$ which satisfies at least properties *(i)–(vi)*, recalled in Sections 2 and 3. We point out that *(i)–(vi)* are a minimal set of properties, that we expect from a notion of dissimilarity measure.

Axiom **(Q)** is a purely technical axiom assuring that the order structure of $\precsim$ is not "too fine" with respect to the canonical order structure of the real numbers. Its justification encounters the same issues we have with an analogous axiom appearing in utility theory (see, e.g., [23]).

The ensemble **(FD0)–(FD4)** and **(Q)** turns out to be necessary and sufficient to the representability of $\precsim$ by means of an element of a large class of functions, all satisfying properties *(i)–(vi)*, that actually depend only on the component-wise distance of two fuzzy description profiles, that is $D(X, Y) = \varphi(|X - Y|)$, for all $X, Y \in \mathcal{Y}$.

For a fixed $p \in [1, +\infty)$, axiom **(R-p)** allows us to arrive to a representation by means of a weighted L^p distance $D^p_{\boldsymbol{\alpha}}$. Such axiom is not purely qualitative as it considers strict convex combinations of p-th power component-wise distances between any nontrivial finite set of comparisons among fuzzy description profiles. The strict convex combination operation is the responsible for the metric properties of $D^p_{\boldsymbol{\alpha}}$. Given a nontrivial complete dissimilarity relation $\precsim$ on $\mathcal{Y}^2$, axiom **(R-p)** implies all axioms **(FD0)–(FD4)**, while **(Q)** is not implied due to the finite scope of **(R-p)**.

Axiom **(FD5)** is another technical axiom, with a role analogous to **(Q)**. Since **(FD5)** is still based on comparisons, we reserve for it a name in compliance with **(FD0)–(FD4)**. It turns out that, in presence of **(R-p)**, for a nontrivial complete dissimilarity relation $\precsim$, **(FD5)** is actually stronger than **(Q)**, as it assures the uniqueness of the representation $D^p_{\boldsymbol{\alpha}}$, besides existence. We point out that the pair $(\underline{\varepsilon}, \underline{0})$ appearing in **(FD5)** plays a similar role of the *certainty equivalent* in the classical expected utility theory [23]. Referring to the element E_k of $\mathcal{Y}$ whose k-th component is 1 and the others are 0, axiom **(FD5)** can be used to single out an elicitation procedure.

For $k = 1$, we ask the agent to single out a number $\alpha_1 = \beta_1^p \in [0, 1]$, such that $(E_1, \underline{0}) \sim (\underline{\beta}_1, \underline{0})$. Then, for $k = 2, \ldots, m-1$, we ask the agent to provide a number $\alpha_k = \beta_k^p \in \left[0, 1 - \sum_{h=1}^{k-1} \alpha_h\right]$, such that $(E_k, \underline{0}) \sim (\underline{\beta}_k, \underline{0})$. Finally, we set $\alpha_m = 1 - \sum_{h=1}^{m-1} \alpha_h$. Hence, in the end we get a vector $\boldsymbol{\alpha}$ of non-negative numbers summing up to 1 that can be used to define a $D^p_{\boldsymbol{\alpha}}$.

Axioms **(U1)** and **(U2)** are equivalent for a nontrivial complete dissimilarity relation $\precsim$ satisfying **(R-p)** and **(FD5)**. They have a qualitative nature and assure that $\boldsymbol{\alpha}$ reduces to a uniform weight distribution $\boldsymbol{\alpha}_u$. In this case we get an unweighted L^p distance representation of $\precsim$. In particular, the p-th power unweighted L^p distance W^p representing $\precsim$ turns out to be an *additive dissimilarity measure*, according to [12–14], also referred to as *local divergence measure*.

Axiom **(CR-p)** is actually a weakening of axiom **(R-p)**, still implying axioms **(FD0)–(FD4)** for a nontrivial complete dissimilarity relation $\precsim$. In all representation results, substituting **(CR-p)** to **(R-p)** we get the representability of $\precsim$ by means of a Choquet L^p distance D^p_ν with respect to a completely alternating normalized capacity ν. Also in this case, the ensemble **(CR-p)** and **(FD5)** is equivalent to the Choquet L^p distance representability of $\precsim$ and axiom **(FD5)** can be used to elicit a completely alternating normalized capacity ν or, equivalently, the Möbius inverse π of its dual capacity. For every $B \in 2^I \setminus \{\emptyset\}$, denote by E_B the element of $\mathcal{Y}$ which is 1 for every $k \in B$ and 0 otherwise. Moreover, we consider the partition $2^I \setminus \{\emptyset\} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_m$, where $\mathcal{B}_j$ is the collection of subsets of I with cardinality j. For every $\mathcal{B}_j$, we fix an enumeration $\mathcal{B}_j = \left\{B_{j,1}, \ldots, B_{j,\binom{m}{j}}\right\}$.

For $j = 1$ and $k = 1$, we ask the agent to provide a number $\pi(B_{1,1}) = 1 - \beta_{1,1}^p \in [0, 1]$, such that $(E_{B^c_{1,1}}, \underline{0}) \sim (\underline{\beta_{1,1}}, \underline{0})$. For $k = 2, \ldots, m$, we ask the agent to provide a number $\pi(B_{1,k}) = 1 - \beta_{1,k}^p \in \left[0, 1 - \sum_{h=1}^{k-1} \pi(B_{1,h})\right]$, such that $(E_{B^c_{1,k}}, \underline{0}) \sim (\underline{\beta_{1,k}}, \underline{0})$. Next, for $j = 2, \ldots, m-1$ and $k = 1, \ldots, \binom{m}{j}$, we ask the agent to single out a number $\pi(B_{j,k}) \in \left[0, 1 - \sum_{i=1}^{j} \sum_{h=1}^{k-1} \pi(B_{i,h})\right]$, such that $1 - \beta_{j,k}^p = \sum_{A \subseteq B_{j,k}} \pi(A)$ and $(E_{B^c_{j,k}}, \underline{0}) \sim (\underline{\beta_{j,k}}, \underline{0})$. Finally, we set $\pi(I) = 1 - \sum_{i=1}^{m-1} \sum_{h=1}^{\binom{m}{i}} \pi(B_{i,h})$. Hence, setting $\pi(\emptyset) = 0$, in the end we get a function $\pi : 2^I \to [0, 1]$ summing up to 1, which is the Möbius inverse (see, e.g., [20]) of the dual of a completely alternating normalized capacity ν that can be used to define a D^p_ν.

In the context of Choquet L^p distance representations, axioms **(I1)** and **(I2)** are equivalent for a nontrivial complete dissimilarity relation $\precsim$ satisfying **(CR-p)** and **(FD5)**. They have a qualitative nature and assure that ν reduces to a vacuous completely alternating normalized capacity ν_∞. In this case we get an unweighted L^∞ distance representation of $\precsim$.

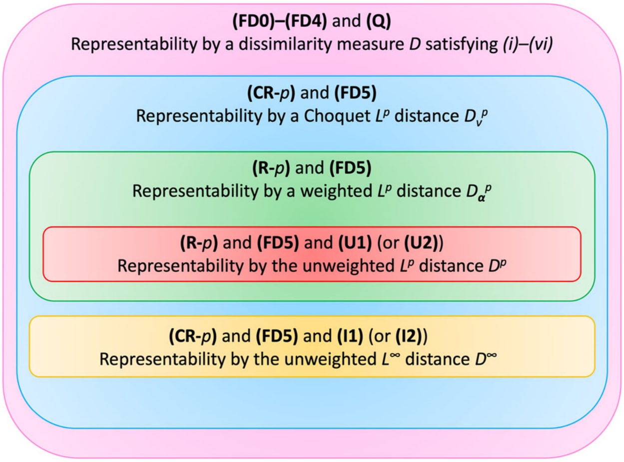

Fig. 1 Axioms and representations of a nontrivial complete dissimilarity relation $\precsim$ on $\mathcal{Y}^2$

Figure 1 summarizes the role of axioms in the family of dissimilarity measures dealt with in this paper. We conclude recalling that, both axioms **(AR-**p**)** and **(ACR-**p**)** are a weakening of **(R-**p**)** and **(CR-**p**)**, respectively, that imply non-triviality but only assure almost representability of $\precsim$.

7 Conclusions

In this paper we characterize comparative dissimilarities on fuzzy description profiles, representable by elements of a class of dissimilarity measures only depending on the attribute-wise distance. This very large class contains all weighted L^p distances, for $p \in [1, +\infty)$, and, in particular, the weighted Manhattan distances and the weighted Euclidean distance. Next, we show that this class also contains the wider class of Choquet L^p distances, for $p \in [1, +\infty)$, with respect to a completely alternating normalized capacity. As a byproduct, this allows us to characterize those comparative dissimilarities representable by the unweighted L^∞ distance.

Acknowledgements The first two authors are members of the INdAM-GNAMPA research group. The second author has been partially supported by the MUR PRIN 2022 project "Models for dynamic reasoning under partial knowledge to make interpretable decisions" (grant number 2022AP3B3B) funded by the European Union - Next Generation EU.

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

Data Availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Conflict of Interests The authors declare that they have no conflict of interest.

References

1. Baioletti, M., Coletti, G., Petturiti, D.: Weighted Attribute Combinations Based Similarity Measures. In: Greco, S., Bouchon-Meunier, B., Coletti, G., Fedrizzi, M., Matarazzo, B., Yager, R.R. (eds.) Advances in Computational Intelligence, pp. 211–220. Springer, Berlin, Heidelberg (2012)
2. Basel Committee on Banking Supervision: International Convergence of Capital Measurement and Capital Standards: A Revised Framework. Technical report, Bank for International Settlements (2004)
3. Basel Committee on Banking Supervision: Basel III: A global regulatory framework for more resilient banks and banking systems. Technical report, Bank for International Settlements (2011)
4. Beg, I., Ashraf, S.: Fuzzy dissimilarity and distance functions. Fuzzy Inf. Eng. **1**(2), 205–217 (2009)
5. Bouchon-Meunier, B., Coletti, G.: How to choose a fuzzy similarity measure in decision-making? Asian J. Econ. Banking **4**, 37–48 (2020)
6. Bouchon-Meunier, B., Rifqi, M., Bothorel, S.: Towards general measures of comparison of objects. Fuzzy Sets Syst. **84**, 143–153 (1996)
7. Coletti, G., Bouchon-Meunier, B.: Fuzzy similarity measures and measurement theory. In: IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), New Orleans, LA, USA, pp. 1–7 (2019)
8. Coletti, G., Bouchon-Meunier, B.: A study of similarity measures through the paradigm of measurement theory: the classic case. Soft Comput. **23**(16), 6827–6845 (2019)
9. Coletti, G., Bouchon-Meunier, B.: A study of similarity measures through the paradigm of measurement theory: the fuzzy case. Soft Comput. **24**(15), 11223–11250 (2020)
10. Coletti, G., Petturiti, D., Bouchon-Meunier, B.: A Measurement Theory Characterization of a Class of Dissimilarity Measures for Fuzzy Description Profiles. In: Lesot, M.-J., Vieira, S., Reformat, M.Z., Carvalho, J.P., Wilbik, A., Bouchon-Meunier, B., Yager, R.R. (eds.) Information Processing and Management of Uncertainty in Knowledge-Based Systems, pp. 258–268. Springer, Cham (2020)
11. Coletti, G., Petturiti, D., Vantaggi, B.: Fuzzy Weighted Attribute Combinations Based Similarity Measures. In: Antonucci, A., Cholvy, L., Papini, O. (eds.) Symbolic and Quantitative Approaches to Reasoning with Uncertainty, pp. 364–374. Springer, Cham (2017)
12. Couso, I., Sánchez, L.: Additive similarity and dissimilarity measures. Fuzzy Sets Syst. **322**, 35–53 (2017)
13. Couso, I., Sánchez, L.: A note on "similarity and dissimilarity measures between fuzzy sets: A formal relational study" and "additive similarity and dissimilarity measures". Fuzzy Sets Syst. **390**, 183–187 (2020)
14. Couso, I., Garrido, L., Sánchez, L.: Similarity and dissimilarity measures between fuzzy sets: a formal relational study. Inf. Sci. **229**, 122–141 (2013)
15. De Baets, B., Janssens, S., Meyer, H.D.: On the transitivity of a parametric family of cardinality-based similarity measures. Int. J. Approximate Reason. **50**(1), 104–116 (2009)
16. Debreu, P.: Representation of Preference Ordering by a Numerical Function. Wiley, New York (1954)
17. Denneberg, D.: Non-Additive Measure and Integral. Kluwer Academic Publshers, Dordrecht (1994)
18. Gale, D.: The Theorey of Linear Economics Models. The University of Chicago Press, Chicago (1989)
19. Goshtasby, A.A.: Similarity and Dissimilarity Measures, pp. 7–66. Springer, London (2012)
20. Grabisch, M.: Set Functions. Games and Capacities in Decision Making. Springer, Cham (2016)
21. Kraft, C.H., Pratt, J.W., Seidenberg, A.: Intuitive probability on finite sets. Ann. Math. Stat. **30**(2), 408–419 (1959)
22. Krantz, D., Luce, R., Suppes, P., Tversky, A.: Foundations of Measurement. Academic Press, San Diego and London (1971)
23. Kreps, D.: Notes On The Theory Of Choice. Westview Press, Boulder, CO (1988)
24. Mangasarian, O.L.: Nonlinear Programming. Classics in Applied Mathematics, vol. 10. SIAM, Philadelphia (1994)

25. Prasetyo, H., Purwarianti A.: Comparison of distance and dissimilarity measures for clustering data with mix attribute types. In: 2014 The 1st International Conference on Information Technology, Computer, and Electrical Engineering, pp. 276–280 (2014)
26. Tversky, A.: Features of similarity. Psychological Rev. **84**, 327–352 (1977)
27. Zadeh, L.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)