Fuzzy multidimensional scaling

Pierre-Alexandre Hébert [a], Marie-Hélène Masson [a,b,*],

Thierry Denœux [a]

[a] *Université de Technologie de Compiègne*

U.M.R CNRS 6599 Heudiasyc

BP 20529 - F-60205 Compiègne Cedex - France

[b] *Université de Picardie Jules Verne*

Abstract

Multidimensional scaling (MDS) is a data analysis technique for representing measurements of (dis)similarity among pairs of objects as distances between points in a low-dimensional space. MDS methods differ mainly according to the distance model used to scale the proximities. The most usual model is the Euclidean one, although a spherical model is often preferred to represent correlation measurements. These two distance models are extended to the case where dissimilarities are expressed as intervals or fuzzy numbers. Each object is then no longer represented by a point but by a crisp or a fuzzy region in the chosen space. To determine these regions, two algorithms are proposed and illustrated using typical datasets. Experiments demonstrate the ability of the methods to represent both the structure and the vagueness of dissimilarity measurements.

Key words:

Fuzzy data analysis, Multidimensional scaling, Fuzzy dissimilarity, Fuzzy correlation

1 Introduction

Multidimensional Scaling (MDS) (Schiffman et al., 1981; Cox and Cox, 1994; Borg and Groenen, 1997) is a classical tool in data analysis. It aims at building a map of objects only described by a proximity matrix (similarities or dissimilarities). The idea is to represent the objects in a low dimensional space (usually Euclidean) in such a way that the interpoint distances reflect, in some sense, the dissimilarities: similar objects are mapped close to each other, whereas dissimilar objects are represented as points distant from one another. This method has been used for many years in various fields (e.g., marketing, psychometric, psychology, chemistry) to produce product maps, stimulus-response maps, concept maps, molecular maps, etc. Its fundamental purpose is to uncover any hidden structure that might be residing in a complex data set, through easily interpretable graphical displays. Fuzzy multidimensional scaling, as described in this paper, is an attempt to generalize classical MDS to the analysis of proximities expressed as fuzzy numbers. Such data can arise in various situations (Denœux and Masson, 2004):

- The initial data may consist in fuzzy feature vectors whose components are fuzzy numbers. Such a data type may be useful to describe a set of entities or the range of a variable observed during a certain period. It may be also a good way to summarize a large amount of data in data mining applications (see, for instance, Diday and Bock (2000)).
- The dissimilarities may be directly elicited from human evaluators who have some difficulty in precisely quantifying the proximity of two objects. A fuzzy number rather than a real value may be more suitable to account for the vagueness of the evaluation.

* Corresponding author. Heudiasyc, BP 20529 - F-60205 Compiègne Cedex - France. Tel.: +33 3 44 23 49 28. Fax: + 33 3 44 23 44 77;
Email address: mmasson@hds.utc.fr (Marie-Hélène Masson).

- The dissimilarities may be measured independently by several sensors, so that the available information concerning the dissimilarity between any two objects takes the form of an empirical distribution. One way to analyze such data is to describe the distribution by a fuzzy number computed from some fractiles of the distribution.

In the past few years, a great deal of attention has been paid to the statistical analysis of imprecise or fuzzy data (see, e.g., Kruse and Meyer (1987); Viertl (1996); Gebhardt et al. (1998); Diamond and Tanaka (1998); Bertoluzza et al. (2002); Denœux and Masson (2004)). In Denoeux and Masson (2000) and Masson and Denœux (2002), generalizations of Euclidean MDS to interval-valued and fuzzy data were presented. This work is reviewed in this paper, and the approach is extended to spherical MDS, a technique useful to analyze fuzzy correlation data such as introduced in Hébert et al. (2003) and Denœux et al. (2005).

The paper is organized as follows. Classical multidimensional scaling methods are first briefly recalled in Section 2. Two conventional distance models are presented: the Euclidean model and the spherical one. Section 3 reviews previous work concerning the generalization of the Euclidean model to interval-valued data (Section 3.1) and to fuzzy data (Section 3.2). In each case, two approaches are presented: least-square fitting that attempts to represent dissimilarities as faithfully as possible on average, and "possibilistic" fitting which guarantees a certain relationship between distances and dissimilarities. New material concerning spherical MDS is then introduced in Section 4. Again, the cases of interval-valued and fuzzy data are addressed, successively, in Sections 4.1 and 4.2, and both least-squares and possibilistic procedures are described in each case. All models and algorithms are illustrated using various real data sets. Section 5 concludes the paper.

2 Classical multidimensional scaling

2.1 Generalities

Given an arbitrary set of points in a p-dimensional space, it is easy to calculate a symmetric matrix containing distances between all points. The basic objective of MDS methods is to solve the inverse problem: given a square matrix $\Delta = (\delta_{ij})$ expressing the dissimilarities between n objects, the problem is to represent each object i by a point $[x_{i1}, ..., x_{ip}]$ in a p-dimensional space such that the interpoint distances reflect, according to some criterion, the input dissimilarities. The representation space is generally chosen to be Euclidean so that the distance between two objects i and j can be computed as:

$$d_{ij} = \sqrt{\sum_{l=1}^{p}(x_{il} - x_{jl})^2} \ . \tag{1}$$

The aim is then to find a $n \times p$ matrix $X = (x_{il})$ such that the distance matrix $D(X)$ computed from X approximates Δ. The quality of approximation is measured via a loss function, usually referred to as a *stress* function, one of the simplest form of which is:

$$\sigma(X) = \sum_{i<j}(d_{ij} - \delta_{ij})^2 \ . \tag{2}$$

Starting from an initial random position, X can be obtained through iterative minimization of (2).

Example 1 *Color data set.* An experiment by Helm (1964) reported in Borg and Groenen (1997, page 360) about the perception of colors by human subjects is considered. Ten colored objects (with different hues but constant brightness and saturation) were presented to different subjects who were asked to rate the perceived dissimilarities. In this example, only the response of the first subject is considered. Classical scaling on these data leads to the two-dimensional representation in Figure 1. It can be seen that the proximities of

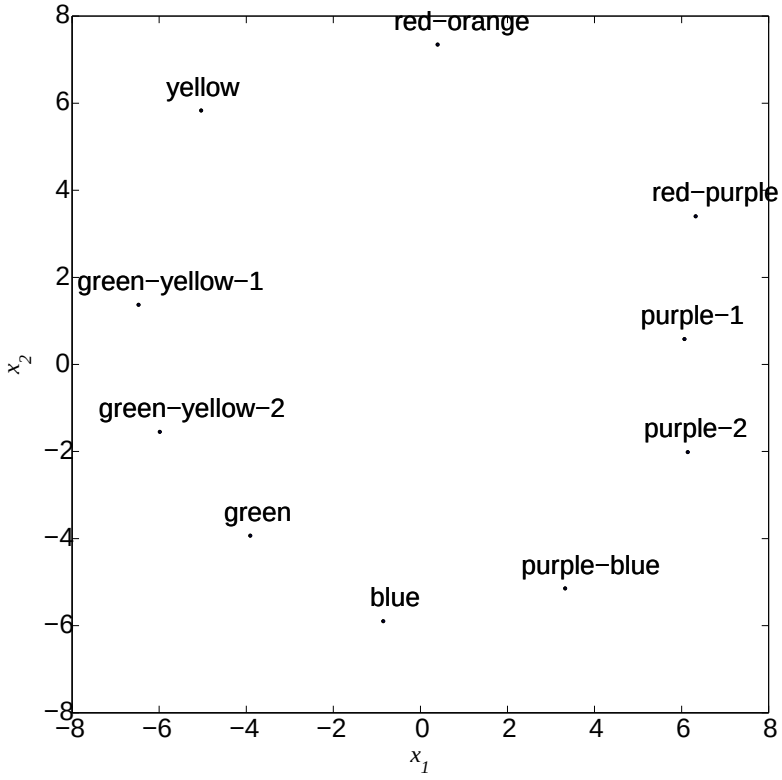

Fig. 1. Two-dimensional configuration for the color data.

the colors in the plane agree with common sense, and, as already reported by Eckman (1954), that they are positioned around a circle. The approximation quality can also be checked using the Shepard diagram shown in Figure 2, which represents the reconstructed distances as a function of the dissimilarities.

MDS is a generic term that includes many different specific models and algorithms. The models differ mainly according to:

- the way (quantitative or qualitative) the dissimilarities are considered (metric and nonmetric approaches);
- the distance model.

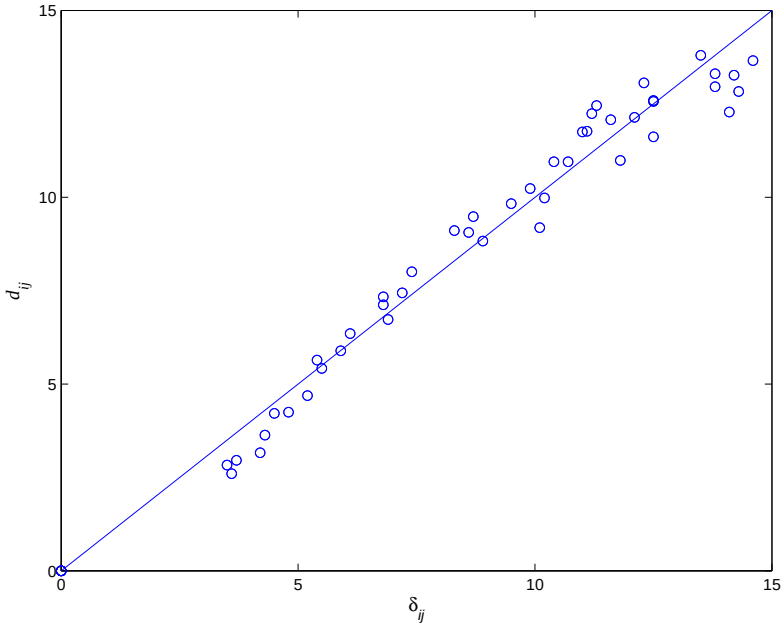

Fig. 2. Shepard diagram for the color data.

2.2 Metric vs. nonmetric approaches

It is often considered sufficient to represent dissimilarities *up to a certain transformation*. For that purpose, the raw dissimilarities δ_{ij} in (2) may often be replaced by pseudo-dissimilarities or *disparities* using an appropriate transformation f such that $\hat{d_{ij}} = f(\delta_{ij})$. Common choices for the transformation f are the affine transformation ($\hat{d_{ij}} = a\delta_{ij} + b$), the logarithmic transformation ($\hat{d_{ij}} = a\log(\delta_{ij}) + b$) and the exponential transformation ($\hat{d_{ij}} = a\exp(\delta_{ij}) + b$). Coefficients a and b are new parameters to be determined during the optimization process. Anyway, in all these cases, the method is referred to as *metric* since the dissimilarities are considered in a quantitative way. Often, especially in social sciences, only the rank order of the dissimilarities is considered meaningful. In this case, the disparities are computed using a transformation referred to as isotonic regression (Kruskal, 1964) which insures that $d_{ij} \leq d_{kl}$ whenever $\delta_{ij} \leq \delta_{kl}$. This *nonmetric* or *ordinal* approach will not be considered further in this paper, although a nonmetric MDS procedure for interval-valued data was proposed in Denoeux and Masson (2000).

2.3 Spherical scaling

Spherical scaling was introduced by Cox and Cox (1991) as an alternative to Euclidean scaling. It is used to find configurations of objects that do not need any "edge points" and is thus especially suitable to visualize correlations between statistical variables. We begin the explanation of the method with the representation on a circle (scaling in $\mathbb{R}^2$), after which spherical scaling will be presented.

Let us assume that the available data consists in a $n \times n$ matrix $T = (\tau_{ij})$ where τ_{ij} denotes the correlation between two variables i and j. The idea is to represent each variable by a vector of unit length in such a way that the cosine of the angle between two vectors is as close as possible to the correlation between the corresponding variables. The problem may be formalized as follows: let the polar coordinates of a variable i be given by $(1, \theta_i)$. The cosine of the angle ϕ_{ij} between two variables i and j is the scalar product of the Cartesian coordinates. It is given by:

$$\cos\phi_{ij} = cos\theta_i \cos\theta_j + \sin\theta_i \sin\theta_j \ . \tag{3}$$

The set $\Theta = (\theta_1, \theta_2, ..., \theta_n)$ of angles, starting from an initial random guess, can be determined by iterative minimization of the following stress function:

$$\sigma(\Theta) = \sum_{i<j} (\cos\phi_{ij} - \tau_{ij})^2 \ . \tag{4}$$

The generalization to multidimensional scaling on the two-dimensional surface of a sphere is straightforward: each variable is represented by spherical coordinates $(1, \theta_{i1}, \theta_{i2})$ which are equivalent in Cartesian coordinates to $\mathbf{x_i} = (\cos\theta_{i1} \sin\theta_{i2}, \sin\theta_{i1} \sin\theta_{i2}, \cos\theta_{i1})$. The cosine of the angle ϕ_{ij} between two variables i and j can be computed by the scalar product $< \mathbf{x_i}, \mathbf{x_j} >$ and the criterion (4) is minimized with respect to $\Theta = (\theta_{11}, \theta_{12}, ..., \theta_{n1}, \theta_{n2})$.

Note that spherical scaling can be seen, alternatively, as a constrained form of principal component analysis where the component loadings are constrained to have length one. This problem can thus be attacked by minimizing a least-squares criterion under equality constraints, as proposed by Pietersz and Groenen (2004). However, our extension of spherical scaling to interval and fuzzy data, described in Section 4, is based on spherical coordinates, following the approach introduced in Cox and Cox (1991).

3 Fuzzy Euclidean scaling

3.1 Interval-valued data

3.1.1 Model

In this section, we assume the dissimilarities to be given in the form of intervals $[\delta_{ij}] = [\delta_{ij}^-, \delta_{ij}^+]$. Each interval is interpreted as the set of possible values for the true unknown dissimilarity δ_{ij}. Since the objects are imprecisely located with respect to each other, it is natural to generalize the Euclidean multidimensional scaling so as to represent an object, no longer as a point, but as a region R_i in $\mathbb{R}^p$. The minimum and maximum distances between two regions R_i and R_j are then defined by:

$$d_{ij}^- = \min_{\mathbf{x}_i \in R_i, \mathbf{x}_j \in R_j} \|\mathbf{x}_i - \mathbf{x}_j\| \tag{5}$$

$$d_{ij}^+ = \max_{\mathbf{x}_i \in R_i, \mathbf{x}_j \in R_j} \|\mathbf{x}_i - \mathbf{x}_j\|. \tag{6}$$

In practice, a parameterized shape has to be chosen for regions R_i. Two models are presented in Denoeux and Masson (2000), corresponding to the representation of objects as hyperspheres and as hyperboxes. Only the hypersphere model will be described here (the hypersphere model is further investigated in

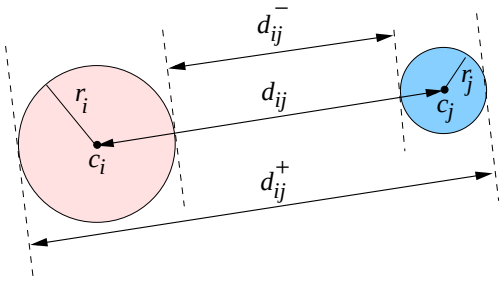

Fig. 3. Minimum and maximum distance between two regions.

the paper by Groenen et al. (2005) in this issue.). In this model, each region is parameterized by a center $\mathbf{c_i} \in \mathbb{R}^\mathbf{p}$ and a radius $r_i \in \mathbb{R}_+$. The representation of d_{ij}^- and d_{ij}^+ in this case is given in Figure 3. As it can be easily seen from this figure, d_{ij}^- and d_{ij}^+ are defined by the following equations:

$$d_{ij}^- = \max(0, d_{ij} - r_i - r_j) \tag{7}$$
$$d_{ij}^+ = d_{ij} + r_i + r_j, \tag{8}$$

where $d_{ij} = \|\mathbf{c}_i - \mathbf{c}_j\|$ denotes the Euclidean distance between the centers $\mathbf{c}_i$ and $\mathbf{c}_j$. The problem is then to determine the centers and the radii such that the interval-valued distances represent, in some sense, the dissimilarities.

Remark 1 Note that a MDS model with hypersphere representation was already proposed by Okada and Imaizumi (1987) (we thank the anonymous referee for drawing our attention on this work). However, the problem addressed in Okada and Imaizumi (1987) was not to scale imprecise dissimilarities, but asymmetric dissimilarity data. The radii of the hyperspheres allow to compute asymmetric distances, the expressions of which differ from those given by Equations (7)-(8).

3.1.2 Least-squares fitting

A straightforward generalization of the conventional Euclidean multidimensional scaling presented in Section 2.1 consists in minimizing the following

stress function:

$$\sigma'(\mathcal{R}) = \sum_{i<j}(d_{ij}^- - \delta_{ij}^-)^2 + \sum_{i<j}(d_{ij}^+ - \delta_{ij}^+)^2, \tag{9}$$

where $\mathcal{R}$ denotes the set of n regions $\{R_1, \ldots, R_n\}$. The $n(p+1)$ model parameters (n centers defined by p coordinates and n radii) can then be determined by minimizing $\sigma'(\mathcal{R})$ with respect to $\mathcal{R}$, using an iterative gradient descent algorithm. Note that this is a constrained optimization problem, since the radii have to be positive. To avoid using constrained optimization procedures, a trick consists in replacing each radius r_i by ρ_i^2, where ρ_i is a new parameter to be determined. Studying the optimality conditions of the problem, as explained in Masson and Denœux (2002), gives interesting insight into the model. In particular, two points are remarkable:

- if all the dissimilarities are precise (i.e. $\delta_{ij}^- = \delta_{ij}^+$), the model leads to null radii, thereby generalizing the classical model;
- otherwise, each radius r_k is linearly related to the quantity

$$s_k = \sum_{i \neq k}(\delta_{ik}^+ - \delta_{ik}^-), \tag{10}$$

which is a measure of the global imprecision of the assessed dissimilarities between object k and all other objects. This observation is fundamental, as it allows to relate the size of the region R_i describing object i, to the imprecision of the data regarding that object.

3.1.3 Possibilistic fitting

The previous algorithm was intended to approximate the input dissimilarities in the least squares sense. Inspired from fuzzy regression models initiated by Tanaka et al. (1982), another way to fit the model was proposed in Masson and Denœux (2002). In the basic Tanaka's interval regression model, a linear model with interval coefficients is postulated, leading to interval-valued predictions.

The regression coefficients are determined through a linear programming (LP) formulation so as to minimize the sum of the widths of the predicted intervals, under the constraint that they include the real or interval-valued output data.

Our possibilistic MDS fitting algorithm is based on a similar idea. Let us suppose that the centers $\mathbf{c_i}$ of the hyperspheres have already been determined, for example by minimizing (9). We may attempt to find the smallest radii r_i such that the following condition is satisfied:

$$[\delta_{ij}^-, \delta_{ij}^+] \subseteq [d_{ij}^-, d_{ij}^+] \quad \forall i, j. \tag{11}$$

This is a conservative approach, since the distance interval $[d_{ij}^-, d_{ij}^+]$ between regions i and j can then be interpreted as a "pessimistic" representation of the dissimilarity interval $[\delta_{ij}^-, \delta_{ij}^+]$. This leads to the following optimization problem:

$$\min_{\mathbf{r}} \sum_{i=1}^{n} r_i \tag{12}$$

subject to:

$$d_{ij}^- \leq \delta_{ij}^- \quad \forall i, j \tag{13}$$

$$d_{ij}^+ \geq \delta_{ij}^+ \quad \forall i, j \tag{14}$$

$$r_i \geq 0 \quad \forall i = 1, n \ . \tag{15}$$

In (12), $\mathbf{r}$ denotes the vector of radii $(r_1, r_2, ..., r_n)^t$. Using the expressions of d_{ij}^- and d_{ij}^- given by (7) and (8), constraints (13) and (14) may be written as

$$\max(0, d_{ij} - r_i - r_j) \leq \delta_{ij}^- \tag{16}$$

$$r_i + r_j \geq \delta_{ij}^+ - d_{ij}, \tag{17}$$

which may be expressed in a more compact form as

$$r_i + r_j \geq \max(d_{ij} - \delta_{ij}^-, \delta_{ij}^+ - d_{ij}) \quad \forall i, j. \tag{18}$$

The minimization of (12) over $\mathbf{r}$ under the constraints (15) and (18) is a linear programming problem once the d_{ij} are fixed. It is trivial to observe that this

problem always has a feasible solution, since $d_{ij}^- \to 0$ and $d_{ij}^+ \to \infty$ when r_i and $r_j \to \infty$. Thus, the parameters of the model can be obtained for any input dissimilarities.

REMARK 2 In contrast to least-squares fitting, possibilistic fitting does not lead to null radii in case of precise but erroneous input dissimilarities: in fact, the obtained representation reflects both the *imprecision* in the data (the widths of the input dissimilarities) and the *goodness-of-fit* of the model (i.e., the choice of the Euclidean distance, the dimensionality of the configuration, and the estimation errors).

EXAMPLE 2 *Towns data set.* We carried the following experiment: a member of our research group was asked to evaluate the distances between 8 European cities. Because of the difficulty of precisely assessing the distances, the subject was allowed to express his evaluations using intervals. These intervals are given in Table 1. The results obtained from both methods are given in Figures 4 and 5. Note that the orientation of axes is arbitrary and is the result of a subjective choice of classical north/south and east/west orientation (any rotation or reflection would keep inter-point distances unchanged). The possibilistic fitting was initialized using the centers obtained from the least-squares procedure. Figure 4 suggests that the respondent had more difficulty to assess large distances, which is reflected by the representation of the peripheral cities (Dublin, Berlin, Madrid and Rome) by larger circles. The least-squares model is thus able to render the overall vagueness (imprecision) in the input data. In the representation obtained by possibilistic fitting, the circles representing each of the cities are larger than those obtained in the least-squares model. This result was expected, since the possibilistic model reflects both imprecision and uncertainty that might reside in the input data. Figure 6 shows a modified Shepard diagram for this example, in which the lower and upper distances are plotted against the upper and lower dissimilarities. It may be checked that the inclusion constraints are satisfied.

	Paris	Dublin	London	Frankfort	Berlin	Marseille	Rome
Paris	0						
Dublin	[850;1050]	0					
London	[250;450]	[450;650]	0				
Frankfort	[500;700]	[1300;1700]	[600;800]	0			
Berlin	[900;1100]	[1700;2300]	[1000;1400]	[450;650]	0		
Marseille	[800;1000]	[1800-2400]	[1100;1400]	[1000;1200]	[1600;2000]	0	
Rome	[1400;1800]	[2200;2800]	[1800;2100]	[1000;1200]	[1700;2300]	[700;900]	0
Madrid	[1500;1900]	[1700;2300]	[1700;2000]	[1500;2500]	[2100;2800]	[900;1100]	[1200;1800]

Table 1
Interval-valued distances estimated by the human subject.

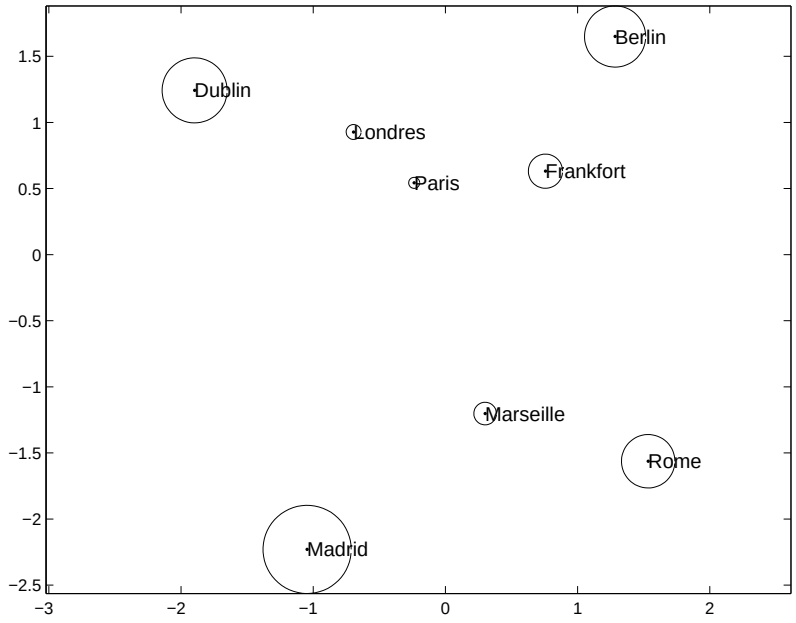

Fig. 4. Town data set: configuration obtained from a least-squares fitting.

REMARK 3 Note that, following a similar line of reasoning, one could attempt to solve the dual problem of *maximizing* the volume of hyperspheres, under the constraints:

$$[d_{ij}^-, d_{ij}^+] \subseteq [\delta_{ij}^-, \delta_{ij}^+] \quad \forall i, j. \tag{19}$$

This is again a LP problem, which, however, does not always have a solution. In fact, the significance of this approach appears to be mostly theoretical: experiments have shown that the existence of a solution is seldom satisfied in practice and that it leads to hardly interpretable representations. For that reason, it will not be detailed further in this paper.

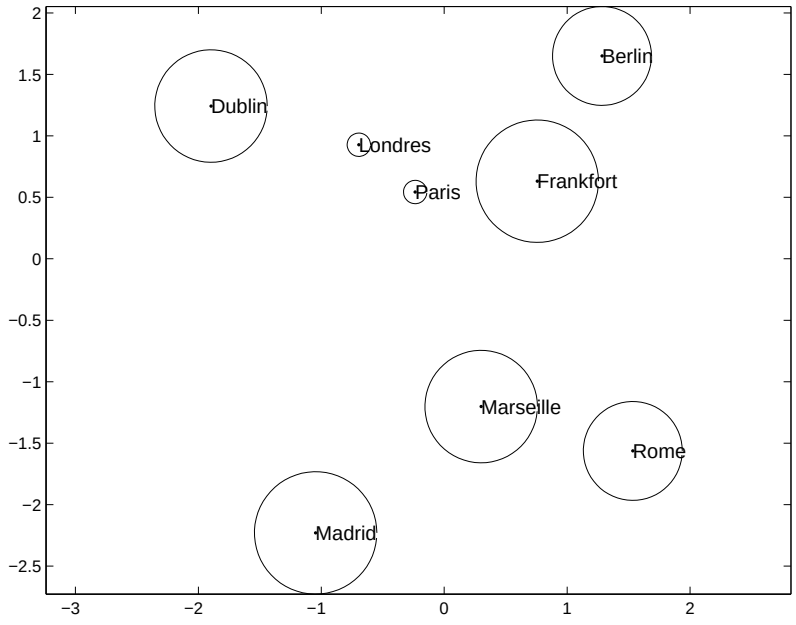

Fig. 5. Town data set: configuration obtained from a possibilistic fitting.

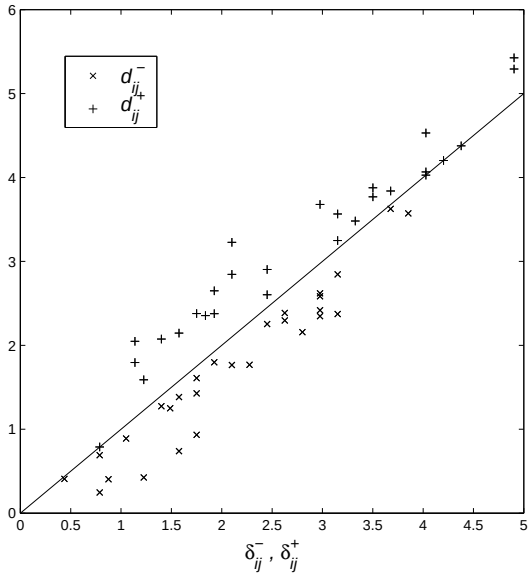

Fig. 6. Modified Sheppard diagram: minimum and maximum distances d_{ij}^- and d_{ij}^+ (y-axis) as a function of lower and upper dissimilarities δ_{ij}^- and δ_{ij}^+ (x-axis).

3.2 Fuzzy data

3.2.1 Model

We assume in this part that each dissimilarity is now expressed as a fuzzy number. As explained in the introduction, such data may come from a lin-

guistic evaluation by a single human subject (such as "very close", "quite different", etc.), or from the synthesis of responses from a panel of assessors. It is now natural to represent each object by a fuzzy region $\widetilde{R}_i$ in $\mathbb{R}^p$ defined by a fuzzy membership function $\mu_{\widetilde{R}_i}$. Applying the extension principle (Zadeh, 1975), the fuzzy distance between two fuzzy regions $\widetilde{R}_i$ et $\widetilde{R}_j$ can be defined as:

$$\mu_{\widetilde{d}_{ij}}(w) = \sup_{\mathbf{x},\mathbf{y}\in\mathbb{R}^p} \min(\mu_{\widetilde{R}_i}(\mathbf{x}), \mu_{\widetilde{R}_j}(\mathbf{y})), \tag{20}$$

where the supremum is computed under the constraint $\|\mathbf{x}-\mathbf{y}\| = w$. If $\widetilde{R}_i$ and $\widetilde{R}_j$ are multidimensional fuzzy numbers (Kaufmann and Gupta, 1991, page 146), each α-cut of $\widetilde{d}_{ij}$ is a closed interval ${}^{\alpha}\widetilde{d}_{ij} = [{}^{\alpha}\widetilde{d}_{ij}^{-}, {}^{\alpha}\widetilde{d}_{ij}^{+}]$, whose bounds are respectively the minimum and maximum distances between the α-cuts of $\widetilde{R}_i$ and $\widetilde{R}_j$. In the line of the previous section, we choose to represent each object by a fuzzy region whose α-cuts are concentric hyperspheres of radii ${}^{\alpha}r_i$ and center $\mathbf{c}_i$, so that

$${}^{\alpha}\widetilde{d}_{ij}^{-} = \max(0, d_{ij} - {}^{\alpha}r_i - {}^{\alpha}r_j) \tag{21}$$

$${}^{\alpha}\widetilde{d}_{ij}^{+} = d_{ij} + {}^{\alpha}r_i + {}^{\alpha}r_j, \tag{22}$$

where d_{ij} denotes, as before, the Euclidean distance between centers $\mathbf{c}_i$ and $\mathbf{c}_j$.

3.2.2 Least squares fitting

To fit the model, a set $\{\alpha_i\}_{i=1,c}$ of predetermined levels of $\alpha - cuts$ has to be chosen with the convention:

$$1 = \alpha_1 > \ldots > \alpha_c = 0 \tag{23}$$

Then, the stress function (9) is extended in the following way:

$$\sigma''(\widetilde{\mathcal{R}}) = \sum_{k=1}^{c} \sum_{i<j} ({}^{\alpha_k}\widetilde{d}_{ij}^{-} - {}^{\alpha_k}\widetilde{\delta}_{ij}^{-})^2 + \sum_{k=1}^{c} \sum_{i<j} ({}^{\alpha_k}\widetilde{d}_{ij}^{+} - {}^{\alpha_k}\widetilde{\delta}_{ij}^{+})^2, \tag{24}$$

where $\widetilde{\mathcal{R}}$ denotes the set of the fuzzy regions $\widetilde{R}_i$, and ${}^0\widetilde{x}$ represents, by convention, the support of fuzzy number $\widetilde{x}$. Note that this stress function is equivalent to the fuzzy least-squares criterion proposed by Diamond (1988) and extended by Ming et al. (1997). The number of parameters of the model is $n(p+c)$: n centers defined by p coordinates c_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, p$ and nc radii ${}^{\alpha_k}r_i$, $i = 1, \ldots, n$, $k = 1, \ldots, c$. The new parameterization

$$ {}^{\alpha_k}r_i = \sum_{h=1}^{k} {}^{\alpha_h}\rho_i^2 \tag{25} $$

allows to account for the positivity constraints ${}^{\alpha_k}r_i \geq 0$, $\forall i, k$ and the monotonicity constraints ${}^{\alpha_{k+1}}r_i \geq {}^{\alpha_k}r_i$, thus transforming a constrained minimization problem into an unconstrained one.

3.2.3 Possibilistic fitting

Following the idea developed in Section 3.1.3, we generalize condition (11) as

$$ \widetilde{\delta}_{ij} \subseteq \widetilde{d}_{ij}, \quad \forall i, j \tag{26} $$

where $\subseteq$ now denotes the standard fuzzy set inclusion, i.e.

$$ \mu_{\widetilde{\delta}_{ij}} \leq \mu_{\widetilde{d}_{ij}}, \quad \forall i, j. \tag{27} $$

Since $\widetilde{\delta}_{ij}$ and $\widetilde{d}_{ij}$ are fuzzy numbers, this condition may be expressed as

$$ [{}^{\alpha_k}\widetilde{\delta}_{ij}^{-}, {}^{\alpha_k}\widetilde{\delta}_{ij}^{+}] \subseteq [{}^{\alpha_k}\widetilde{d}_{ij}^{-}, {}^{\alpha_k}\widetilde{d}_{ij}^{+}] \quad \forall i, j, k. \tag{28} $$

As in Section 3.1.3, we assume that the centers $\mathbf{c}_i$, $i = 1, \ldots, n$ have been determined using, e.g., the least-squares procedure described in the previous section. The problem is then to find the "smallest" fuzzy regions $\widetilde{R}_i$ satisfying condition (28). Following the same line of reasoning as in Section 3.1.3, this can done by solving the following LP problem:

$$ \min_{\mathbf{r}} \sum_{k=1}^{c} \sum_{i=1}^{n} {}^{\alpha_k}r_i \tag{29} $$

subject to:

$$ {}^{\alpha_k}r_i + {}^{\alpha_k}r_j \geq \max(d_{ij} - {}^{\alpha_k}\delta_{ij}^-, {}^{\alpha_k}\delta_{ij}^+ - d_{ij}) \quad \forall i,j,k \tag{30} $$

$$ {}^{\alpha_0}r_i \geq 0, \quad \forall i \tag{31} $$

$$ {}^{\alpha_k}r_i \leq {}^{\alpha_{k+1}}r_i, \quad \forall i, \forall k < c, \tag{32} $$

where $\mathbf{r}$ in (29) denotes the vector of all parameters ${}^{\alpha_k}r_i$.

Example 3 We turn back to Example 1 about the perception of colors. We now consider the answers of several subjects. They were classified into two groups: some of them had a normal color vision, whereas the other had a color-deficient vision. Two separate analyses were conducted on these two groups. The perception of each group was summarized using a triangular fuzzy number computed from the minimum, maximum and mean responses of the subjects defining, respectively, the lower and upper bounds of the support and the core of the fuzzy number. Figure 7 presents the results obtained using $c = 2$ α-cut levels (support and core) using possibilistic fitting. It can be seen that the color annular structure is well-recovered for the first group of subject. Moreover, the small size of the darker circles indicates that the Euclidean model is well-adapted to the input data, and that the mean responses of the subjects were very precise. Some colors (like green-yellow-1, green-yellow-2 and green or red-purple, purple-1 and purple-2) are logically less discriminated than others. With the color-deficient subjects, a greater confusion among colors can be noticed. The annular structure is slightly distorted, confirming similar results reported by Helm. As compared to the configuration resulting from the color-normal subjects, it is evident that the answers of the second group of subjects are confused and erroneous, which is indicated by larger cores and supports of the fuzzy regions. This result is confirmed by the left-hand side and right-hand side of Figure 8 showing the membership of some dissimilarities and corresponding reconstructed distances for the two groups of subjects. The agreement between dissimilarities and distances is obviously much worse for

the blind-color group, which is compensated in the possibility model by greater imprecision.

The least-squares method was also applied to the same data as above. As before, only $c = 2$ levels were used. The configurations obtained for the normal and color-blind group are shown in Figure 9. Once again, the annular structure pointed out by Helm is well recovered, with some distortion in the color-blind group. The regions representing each color are more precise than those obtained using the possibility model, and their core is reduced to a point, which results from the choice of triangular fuzzy numbers to represent the input dissimilarities. Interestingly, the configurations for the two groups are more similar than those obtained with the possibility model. This may be explained by the fact that, in the least-squares model, the imprecision of the representation (i.e., the size and fuzziness of the regions) does not reflect estimations errors (stemming from the inadequacy of the Euclidean model), but the imprecision of the input dissimilarities, which is here about the same for the two groups. The left-hand side and right-hand side of Figure 10 representing the membership functions of some dissimilarities and reconstructed distances for the two groups, clearly show that reconstruction errors are larger in the color-blind group.

4 Fuzzy spherical scaling

4.1 Interval-valued data

4.1.1 Model

Let us now assume the data to consist in interval-valued correlations between n variables. Each interval $[\tau_{ij}] = [\tau_{ij}^{-}; \tau_{ij}^{+}]$ is interpreted as a set of possible values for the true unknown correlation value τ_{ij} between variables i and j. Such data

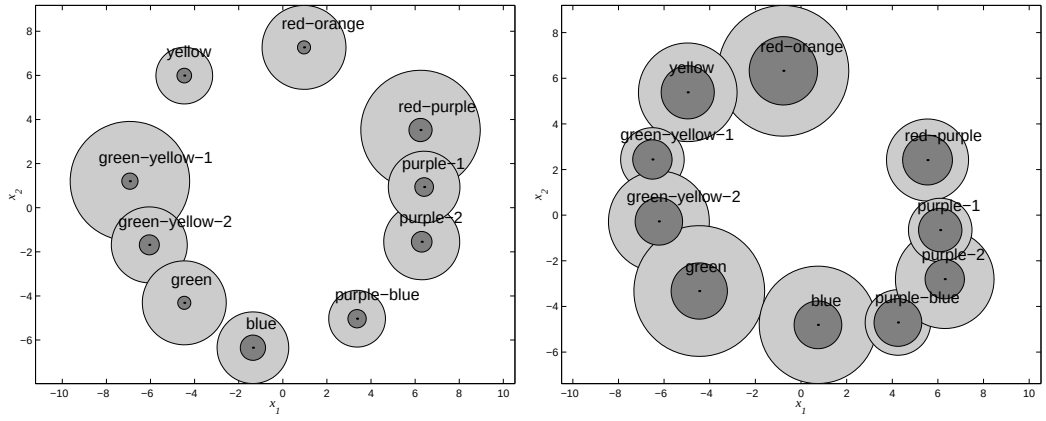

Fig. 7. Color data: possibilistic model. (left) normal subjects; (right) color-deficient subjects. The supports are colored in light grey whereas the dark grey represent the cores of the fuzzy regions.

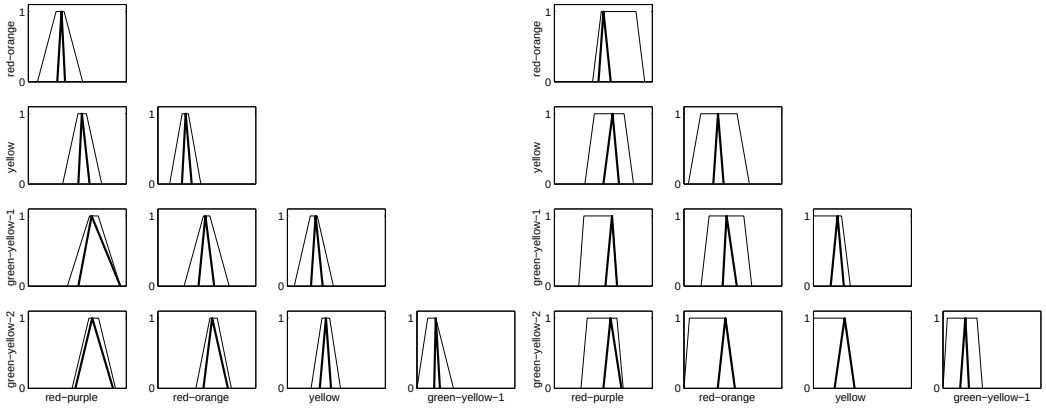

Fig. 8. Color data: possibilistic model. Membership functions of $\tilde{\delta_{ij}}$ (bold lines) and $\tilde{d_{ij}}$ for 10 pairs of colors seen by normal subjects (left) and color-deficient subjects (right).

typically arise when objects are described by interval-valued attributes. The Kendal's rank correlation coefficient was extended to such data in Hébert et al. (2003); Denœux et al. (2005), but other correlation coefficients could be considered as well (see, e.g., Liu and Kao (2002)). As in classical spherical scaling, the objects will be represented on a hypersphere $\mathbf{S}^p$ of radius 1 in the space of dimension p: $\mathbf{S}^2$ is a circle and $\mathbf{S}^3$ is a sphere.

In the case of precise numerical data, the classical spherical scaling method

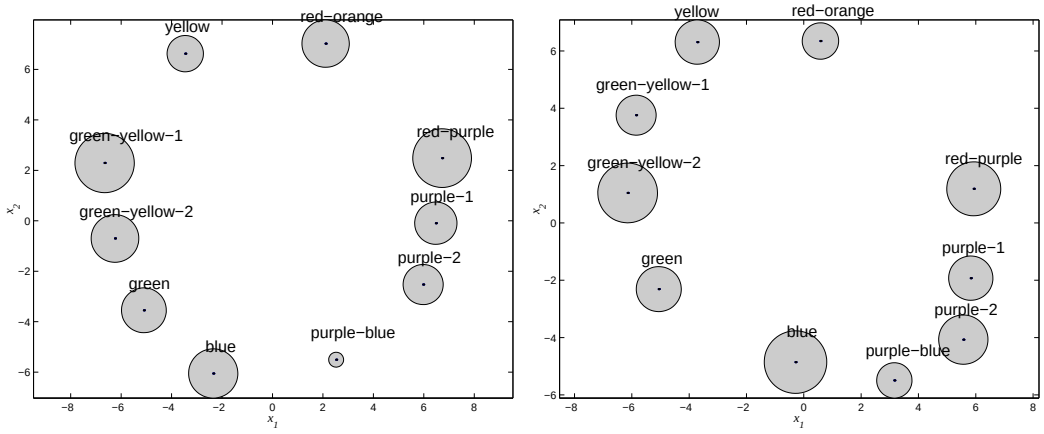

Fig. 9. Color data: least-squares model. (left) normal subjects; (right) color-deficient subjects. Supports of the fuzzy regions.

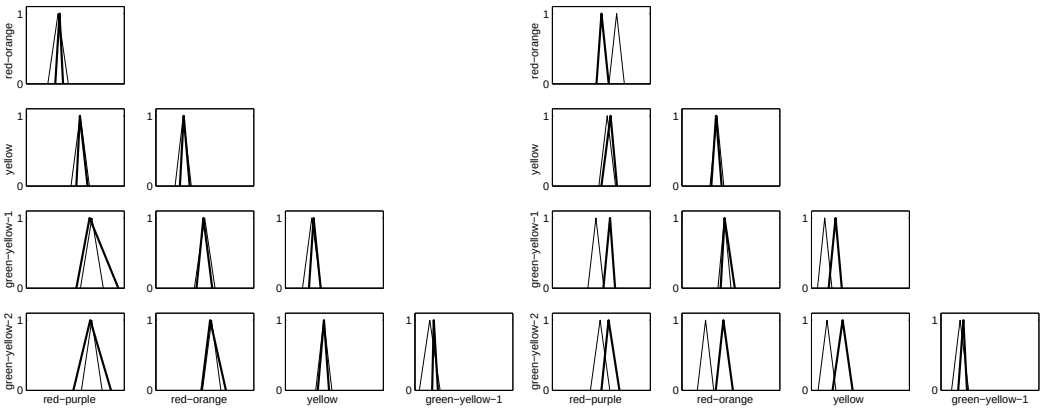

Fig. 10. Color data: least-squares model. Membership functions of $\tilde{\delta_{ij}}$ (bold lines) and $\tilde{d_{ij}}$ for 10 pairs of colors seen by normal subjects (left) and color-deficient subjects (right).

allows to approximate the correlation value τ_{ij} by the cosine of the angle $\widehat{(\mathbf{x}_i, \mathbf{x}_j)}$ (see Section 2.3). This angle, defined in $[0, \pi]$, is the angle between the points $\mathbf{x}_i$ and $\mathbf{x}_j$ in the plane specified by the three points $\mathbf{x}_i$, $\mathbf{x}_j$ and the center O of $\mathbf{S}^p$. As in the Euclidean case, it is natural to represent the imprecise location of each object i by a region S_i of $\mathbf{S}^p$. Therefore, a pair of regions S_i and S_j located on $\mathbf{S}^p$ induces a set of cosines, that may conveniently be characterized by its minimum and maximum:

$$\min_{\mathbf{x}_i \in S_i, \mathbf{x}_j \in S_j} \cos\left(\widehat{\mathbf{x}_i, \mathbf{x}_j}\right) = \cos\left(\max_{\mathbf{x}_i \in S_i, \mathbf{x}_j \in S_j} \left(\widehat{\mathbf{x}_i, \mathbf{x}_j}\right)\right) \tag{33}$$

$$\max_{\mathbf{x}_i \in S_i, \mathbf{x}_j \in S_j} \cos\left(\widehat{\mathbf{x}_i, \mathbf{x}_j}\right) = \cos\left(\min_{\mathbf{x}_i \in S_i, \mathbf{x}_j \in S_j} \left(\widehat{\mathbf{x}_i, \mathbf{x}_j}\right)\right). \tag{34}$$

The maximal and minimal angles, which correspond, respectively, to the minimal and maximal possible cosines, are denoted as follows:

$$\phi_{ij}^{-} = \max_{\mathbf{x}_i \in S_i, \mathbf{x}_j \in S_j} \left(\widehat{\mathbf{x}_i, \mathbf{x}_j}\right) \tag{35}$$

$$\phi_{ij}^{+} = \min_{\mathbf{x}_i \in S_i, \mathbf{x}_j \in S_j} \left(\widehat{\mathbf{x}_i, \mathbf{x}_j}\right). \tag{36}$$

In practice, we decide to parameterize the shape of each region S_i by a center $\mathbf{c}_i$ in $\mathbf{S}^p$ and an imprecision angle $\beta_i \in [0, \pi]$ as follows:

$$S_i = \left\{ \mathbf{x} \in \mathbf{S}^p \ / \ \widehat{(\mathbf{x}, \mathbf{c}_i)} \leq \beta_i \right\}. \tag{37}$$

Each region S_i is thus a circular arc when $p = 2$, and a spherical cap when $p = 3$. The interval $[\phi_{ij}^{+}, \phi_{ij}^{-}]$ obviously defines the complete set of angles $\left(\widehat{\mathbf{x}_i, \mathbf{x}_j}\right)$ such that $\mathbf{x}_i \in S_i$ and $\mathbf{x}_j \in S_j$. Since the cosine function is decreasing, the interval $[\cos(\phi_{ij}^{-}), \cos(\phi_{ij}^{+})]$ defines the whole set of possible cosines obtained with pairs $(\mathbf{x}_i, \mathbf{x}_j) \in (S_i, S_j)$.

Let ϕ_{ij} denote the angle $\left(\widehat{\mathbf{c}_i, \mathbf{c}_j}\right)$. As already explained in the crisp case, it can be computed as the arccosine of the scalar product of the centers $\mathbf{c}_i$ and $\mathbf{c}_j$:

$$\phi_{ij} = \arccos \langle \mathbf{c}_i, \mathbf{c}_j \rangle . \tag{38}$$

For any dimension p, the pair of extreme angles $(\phi_{ij}^{+}, \phi_{ij}^{-})$ may clearly be measured in the plane $(O, \mathbf{c}_i, \mathbf{c}_j)$. Figure 11 gives a representation of this plane, where it can be seen that ϕ_{ij}^{-} and ϕ_{ij}^{+} are defined by the following equations, regardless of the dimension p:

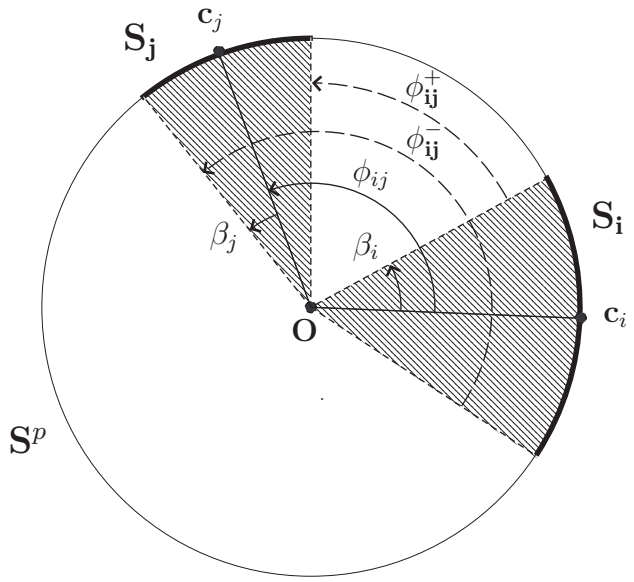

Fig. 11. Maximal and minimal angles between two regions.

$$\phi_{ij}^- = \min\left(\pi, \phi_{ij} + \beta_i + \beta_j\right) \tag{39}$$

$$\phi_{ij}^+ = \max\left(0, \phi_{ij} - \beta_i - \beta_j\right). \tag{40}$$

It is necessary to introduce the min and max functions, in order to deal with two particular cases:

- $\exists(\mathbf{x}_i, \mathbf{x}_j) \in (S_i, S_j) \;/\; \widehat{(\mathbf{x}_i, \mathbf{x}_j)} = \pi$: the maximal angle π corresponding to opposite locations is reached, so ϕ_{ij}^- is equal to π;
- $\exists(\mathbf{x}_i, \mathbf{x}_j) \in (S_i, S_j) \;/\; \widehat{(\mathbf{x}_i, \mathbf{x}_j)} = 0$: the minimal angle 0 corresponding to identical locations is reached, so ϕ_{ij}^+ is equal to 0.

Consequently, the problem is to determine the centers $\mathbf{c}_i$ and the imprecision angle β_i of each object i, such that the cosine of the interval-valued angular distances represent the interval-valued correlations.

The center $\mathbf{c}_i$ located on the sphere $\mathbf{S}^p$ will be characterized by the spherical coordinates $(1, \theta_{i1}, \ldots, \theta_{i(p-1)})$, where $\theta_{i(p-1)} \in [0, 2\pi]$ and $\theta_{iq} \in [0, \pi]$ for all $q < p - 1$.

Figure 12 shows a pair of objects in the case $p = 3$: $\mathbf{S}^3$ is a sphere, and

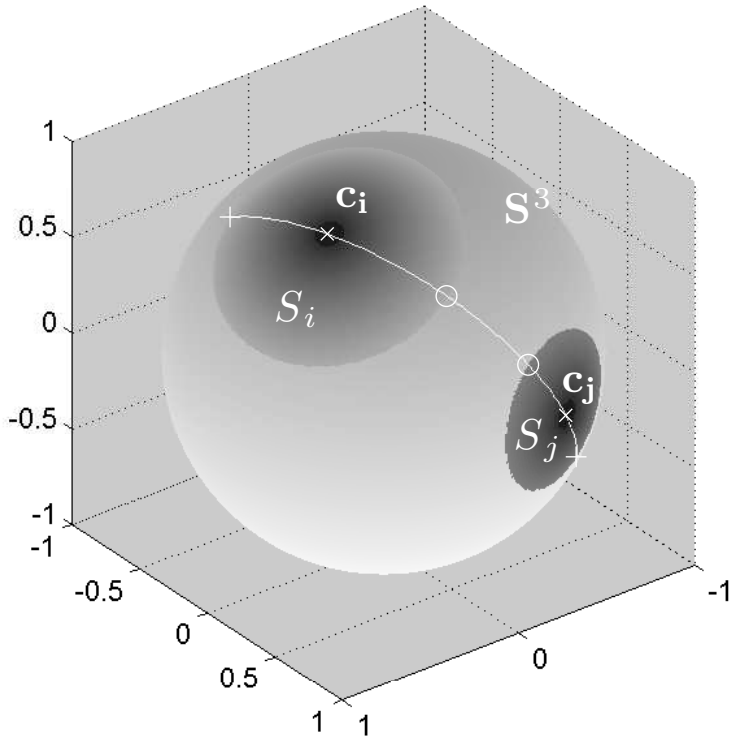

Fig. 12. Representation of two regions on the sphere S^3.

the shapes of regions S_i and S_j are spherical caps. Both centers $\mathbf{c}_i$ and $\mathbf{c}_j$ are marked with the $\times$ sign. Included in the plane $(O, \mathbf{c}_i, \mathbf{c}_j)$, the white arc delimited by the $+$ signs connects the furthest points of S_i and S_j, whereas its subset delimited by the o signs connects the nearest points: the corresponding angles are equal, respectively, to ϕ_{ij}^- and ϕ_{ij}^+.

4.1.2 Least-squares fitting

A straightforward generalization of the conventional spherical multidimensional scaling presented in Section 2.3 consists in minimizing the following criterion:

$$\sigma'(\mathcal{S}) = \sum_{i<j} \left(\cos(\phi_{ij}^-) - \tau_{ij}^-\right)^2 + \sum_{i<j} \left(\cos(\phi_{ij}^+) - \tau_{ij}^+\right)^2, \qquad (41)$$

where $\mathcal{S}$ denotes the set of n regions $\{S_1, \ldots, S_n\}$. The np model parameters (n centers defined by $p-1$ coordinates θ_{iq} and n angles β_i) can be determined by minimizing $\sigma'(\mathcal{S})$ with respect to $\mathcal{S}$, using an iterative gradient descent algorithm. The partial derivatives w.r.t. each parameter are given in Appendix A. It is a constrained optimization problem since the imprecision angles must belong to $[0, \pi]$. To account for this constraint, each parameter β_i is replaced by a real $b_i \in \mathbb{R}$, which is transformed by a derivable and increasing function $e(x) : \mathbb{R} \to [0, \pi]$. We choose in practice the function

$$e(x) = \frac{\pi}{1 + \exp(-x)} \ .$$

4.1.3 Possibility fitting

The possibilistic procedure described in Section 3.1 can be adapted to fit the interval spherical model. First, we assume that the centers of the regions S_i have already been computed, for example by minimizing (41). Next, we look for the smallest imprecision angles such that the following condition is satisfied:

$$[\tau_{ij}^-, \tau_{ij}^+] \subseteq \left[\cos(\phi_{ij}^-), \cos(\phi_{ij}^+)\right], \ \forall i, j. \tag{42}$$

The angular distances $[\phi_{ij}^+, \phi_{ij}^-]$ may then be interpreted as a "pessimistic" representation of the correlation interval $[\tau_{ij}^-, \tau_{ij}^+]$. The optimization problem is then specified as follows:

$$\text{Minimizing } \sum_{i=1}^{n} \beta_i, \tag{43}$$

subject to:

$$\cos(\phi_{ij}^-) \leq \tau_{ij}^-, \ \forall i, j \tag{44}$$

$$\cos(\phi_{ij}^+) \geq \tau_{ij}^+, \ \forall i, j \tag{45}$$

$$\beta_i \geq 0, \ \forall i \tag{46}$$

$$\beta_i \leq \pi, \ \forall i. \tag{47}$$

Using (39) and (40), constraints (44) and (45) may be simplified as:

$$\cos(\phi_{ij}^-) \leq \tau_{ij}^- \Leftrightarrow \cos(\min(\pi, \phi_{ij} + \beta_i + \beta_j)) \leq \tau_{ij}^- \quad (48)$$
$$\Leftrightarrow \min(\pi, \phi_{ij} + \beta_i + \beta_j) \geq \arccos(\tau_{ij}^-) \quad (49)$$
$$\Leftrightarrow \phi_{ij} + \beta_i + \beta_j \geq \arccos(\tau_{ij}^-) \quad (50)$$
$$\Leftrightarrow \beta_i + \beta_j \geq \arccos(\tau_{ij}^-) - \phi_{ij} \ , \quad (51)$$

and

$$\cos(\phi_{ij}^+) \geq \tau_{ij}^+ \Leftrightarrow \cos(\max(0, \phi_{ij} - \beta_i - \beta_j)) \geq \tau_{ij}^+ \quad (52)$$
$$\Leftrightarrow \max(0, \phi_{ij} - \beta_i - \beta_j) \leq \arccos(\tau_{ij}^+) \quad (53)$$
$$\Leftrightarrow \phi_{ij} - \beta_i - \beta_j \leq \arccos(\tau_{ij}^+) \quad (54)$$
$$\Leftrightarrow \beta_i + \beta_j \geq \phi_{ij} - \arccos(\tau_{ij}^+). \quad (55)$$

Consequently, both constraints (44) and (45) can be expressed in the following compact form:

$$\beta_i + \beta_j \geq \max\left(\arccos(\tau_{ij}^-) - \phi_{ij}, \phi_{ij} - \arccos(\tau_{ij}^+)\right), \ \forall i, j. \quad (56)$$

The minimization of criterion (43) under the constraints (46), (47) and (56) is a linear programming problem. This problem always has a feasible solution: if $\beta_i \to \pi$ and $\beta_j \to \pi$ then $\phi_{ij}^- \to \pi$ and $\phi_{ij}^+ \to 0$; consequently $\cos(\phi_{ij}^-) \to -1$ and $\cos(\phi_{ij}^+) \to 1$.

4.2 Fuzzy data

4.2.1 Model

We assume in this part some correlations to be expressed as fuzzy intervals $\widetilde{\tau}_{ij}$. Concepts of fuzzy correlation were introduced in Liu and Kao (2002) and Hébert et al. (2003); Denœux et al. (2005) to measure the degree of association between fuzzy-valued attributes, by applying the extension principle (Zadeh, 1975). The data being fuzzy, each object i will now be represented by a fuzzy region $\widetilde{S}_i$ in $\mathbf{S}^p$, defined by a fuzzy membership function $\mu_{\widetilde{S}_i}$. The extension

principle allows to define the fuzzy angular distance $\widetilde{\phi}_{ij}$ between two regions $\widetilde{S}_i$ and $\widetilde{S}_j$:

$$\mu_{\widetilde{\phi}_{ij}}(w) = \sup_{\mathbf{x},\mathbf{y}\in\mathbf{S}^p \,/\, \widehat{(\mathbf{x},\mathbf{y})}=w} \min(\mu_{\widetilde{S}_i}(\mathbf{x}), \mu_{\widetilde{S}_j}(\mathbf{y})). \tag{57}$$

In the line of the fuzzy Euclidean case, we choose to represent each object i by a fuzzy spherical cap (respectively, a fuzzy circular arc in the case $p = 2$), whose α-cuts are nested caps (respectively, nested circular arcs) with center $\mathbf{c}_i$ and imprecision angle ${}^{\alpha}\beta_i$. Then, ${}^{\alpha}\widetilde{\phi}_{ij}$ is the closed interval $[{}^{\alpha}\widetilde{\phi}^{+}_{ij}, {}^{\alpha}\widetilde{\phi}^{-}_{ij}]$ such that:

$${}^{\alpha}\widetilde{\phi}^{-}_{ij} = \min\left(\pi, \phi_{ij} + {}^{\alpha}\beta_i + {}^{\alpha}\beta_j\right) \tag{58}$$

$${}^{\alpha}\widetilde{\phi}^{+}_{ij} = \max\left(0, \phi_{ij} - {}^{\alpha}\beta_i - {}^{\alpha}\beta_j\right), \tag{59}$$

where ϕ_{ij} denotes, as before, the angle $\widehat{(\mathbf{c}_i, \mathbf{c}_j)}$.

4.2.2 Least-squares fitting

We apply the same convention of using a set $\{\alpha_k\}$ of predetermined levels of α-cuts, such that:

$$1 \geq \alpha_1 > \ldots > \alpha_c > 0. \tag{60}$$

The stress function (41) is extended in the following way:

$$\sigma''(\widetilde{\mathcal{S}}) = \sum_{k=1}^{c} \sum_{i<j} \left(\cos({}^{\alpha_k}\widetilde{\phi}^{-}_{ij}) - {}^{\alpha_k}\widetilde{\tau}^{-}_{ij}\right)^2 + \sum_{k=1}^{c} \sum_{i<j} \left(\cos({}^{\alpha_k}\widetilde{\phi}^{+}_{ij}) - {}^{\alpha_k}\widetilde{\tau}^{+}_{ij}\right)^2, \tag{61}$$

where $\widetilde{\mathcal{S}}$ denotes the set of the fuzzy spherical caps $\widetilde{S}_i$. The number of parameters of the model is $n(p-1+c)$: n centers defined by $(p-1)$ spherical coordinates θ_{ij} and nc imprecision angles ${}^{\alpha_k}\beta_i$. The following parameterization is used:

$$
\begin{aligned}
{}^{\alpha_k}\beta_i &= e\left({}^{\alpha_1}b_i + \sum_{h=2}^{k} {}^{\alpha_h}{b_i}^2\right) && (62)\\
&= \frac{\pi}{1+\exp\left(-\left({}^{\alpha_1}b_i + \sum_{h=2}^{k} {}^{\alpha_h}{b_i}^2\right)\right)}, && (63)
\end{aligned}
$$

which automatically takes care of the constraints ${}^{\alpha_k}\beta_i \in [0,\pi]$, $\forall i,k$, and ${}^{\alpha_k}\beta_i \leq {}^{\alpha_{k+1}}\beta_i$, $\forall i$, $\forall k < c$.

The obtained minimization problem is unconstrained, and may easily be solved by an iterative descent algorithm. Partial derivatives formulas are detailed in Appendix B.

EXAMPLE 4 *Sensory data set*[1] . We consider in this part a real fuzzy data set coming from the evaluation of a set of 8 products by a subject, according to 5 visual descriptors: *OPAQue*, *BRILliant*, *GRANUlose*, *BRIGht*, and *NACRe*. For each descriptor, the assessor was asked to assess the intensity level of its perception for each product, on a continuous scale $[0,100]$. The classical method consists in acquiring precise numerical scores. However, such scores typically exhibit high variability, and the evaluation needs to be repeated several times. Reducing the number of repetitions would allow to reduce the cost of the experiment. In this experiment, the subject was asked to provide not only a precise numerical value for each product and each attribute (corresponding to the "most possible" value), but also an interval that surely contained the "true" value. Together with the point estimate, this interval defines a triangular fuzzy number, which may be seen as a fuzzy score.

An expected result of such an analysis is deeper understanding of the relationship between descriptors. In the case of crisp data, the similarity or dissimilarity of each pair of descriptors is expressed by a correlation measure. In the case of fuzzy data, fuzzy correlations measures may be obtained by applying

[1] The authors would like to thank PSA Peugeot Citröen company for providing this data set.

Products	Opaque	Brilliant	Granulose	Bright	Nacre
1	(96,99.9,99.9)	(97.7,100,100)	(96,99.9,99.9)	(96.8,100,100)	(26.7,30,53.4)
2	(19.4,28,31.4)	(60.5,69.4,75.1)	(26.1,37.4,43.7)	(64.4,74.1,80.8)	(59.5,68.1,78)
3	(36,48.3,68.8)	(3.2,10.3,14.6)	(17.7,25.7,31.9)	(42.7,54.4,67.5)	(66.9,78,86)
4	(94.8,96.2,99.1)	(43.9,56.5,63.6)	(84.9,92.1,96.4)	(92.3,94,100)	(58.9,71,77.1)
5	(48.1,58.1,65.5)	(26.7,42.3,48.9)	(18.5,35.3,46.8)	(50.9,59.7,71)	(42.3,58.5,64.9)
6	(46.8,56,65.7)	(0,0.3,4.2)	(48.5,60.3,68.8)	(31.9,44.8,62.8)	(51.7,68.1,72.8)
7	(0,0,0)	(44,50.9,62.6)	(0.3,0.3,0.3)	(0.1,0.1,0.1)	(54.2,71.2,83.5)
8	(6.8,12.4,22.4)	(81.5,85.4,96.6)	(79.2,85.8,95.2)	(16.1,20.6,40.1)	(82.1,91.9,99.9)

Table 2
Fuzzy scores assigned by the expert to the 8 products according to the 5 descriptors.

Zadeh's extension principle. Kendall's correlation coefficient was chosen for this data set, because the ordinal information appears more relevant than the quantitative one. However, it would also be possible to use fuzzy Pearson's correlations (Liu and Kao, 2002). The application of Zadeh's extension principle to extend Kendall's rank correlation to fuzzy numbers is described in Denœux et al. (2005) and Hébert et al. (2003).

The triangular fuzzy numbers obtained for each descriptor are given in Table 2, and they are represented graphically in Figure 13. The computed fuzzy Kendall's tau values $\widetilde{\tau}_{ij}$ are shown in Figure 14. We note that each interval ${}^{\alpha}\widetilde{\tau}_{ij}$ is the convex hull of the set of possible Kendall's coefficients that may be obtained by values taken in the α-cuts of the fuzzy scores according to the descriptors i and j.

The 0^+ and 0.9 level cuts of the fuzzy regions obtained with spherical MDS for $p = 3$ are represented in Figure 15. Shepard's diagram (Figure 16) shows a quite good approximation of the Kendall's tau values. Figure 15 immediately reveals a rather high correlation between descriptors *Opaque*, *Granulose* and *Bright* which appear as a cluster on the sphere. The other salient feature is the high degree of imprecision of the representation for descriptor *Nacre*, due to the large support of its Kendall's coefficients as shown in Figure 14. This imprecision may be interpreted as a lack of discriminant power. In Table 2 and Figure 13, we may indeed verify that this descriptor does not allow to distinguish between the products: a lot of pairs (i, j) of objects have overlap-

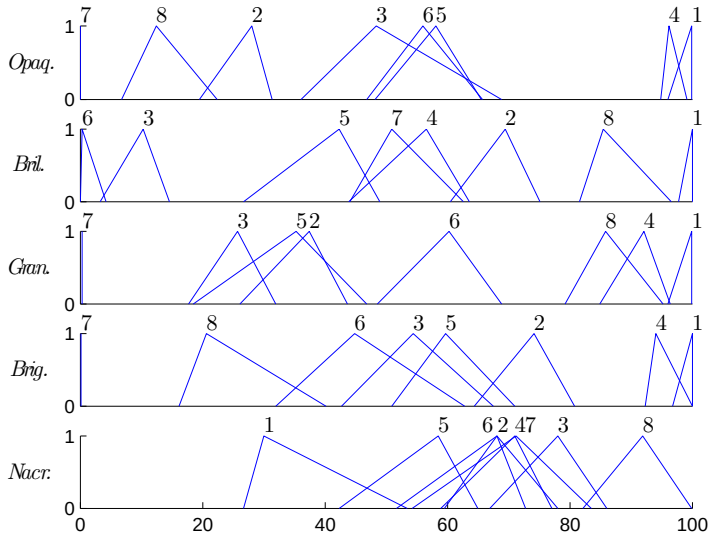

Fig. 13. Sensory data set.

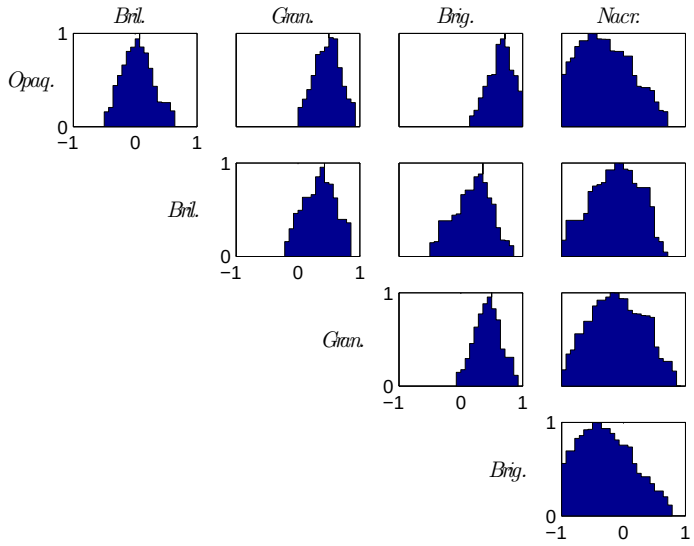

Fig. 14. Membership functions of the fuzzy Kendall's coefficients for the sensory data set.

ping fuzzy score supports; for each (i, j), this means that both orders $i \succ j$ and $i \prec j$ are possible. Note that the *Nacre* descriptor is also clearly opposed to the four others: on the sphere, it is possible to draw diameters joining the cap representing this descriptor and the four others. This is justified by the

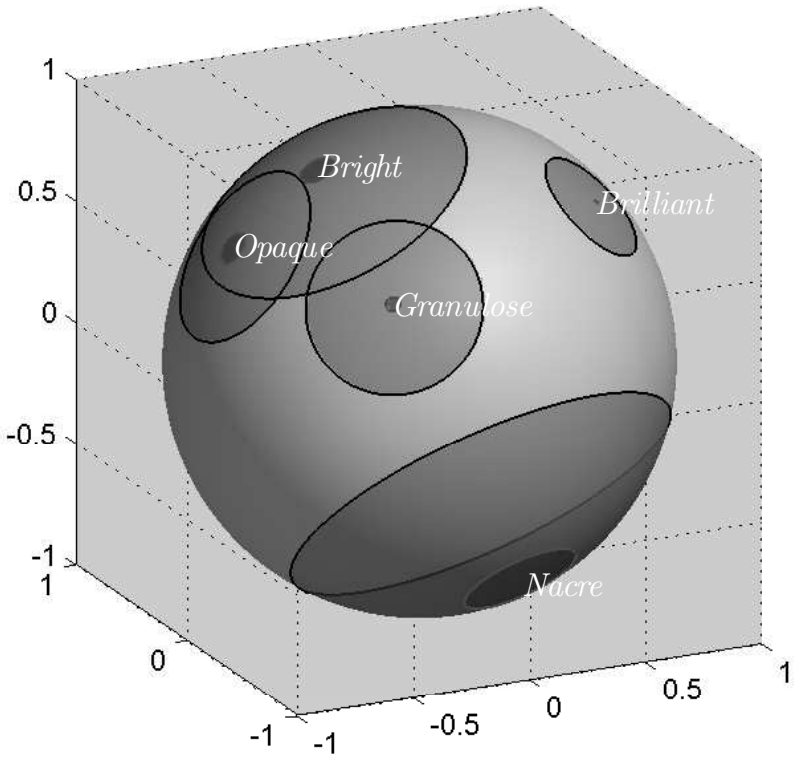

Fig. 15. Sensory data: spherical MDS, least-squares model, α-cuts $\{0^+, 0.9\}$.

possibility of strong negative correlations as shown in Figure 14.

4.2.3 Possibilistic fitting

As in the Euclidean case, constraint (42) is generalized by:

$$\widetilde{\tau}_{ij} \subseteq \cos(\widetilde{\phi}_{ij}),\ \forall i, j. \tag{64}$$

where $\subseteq$ denotes the standard fuzzy set inclusion, i.e.:

$$\mu_{\widetilde{\tau}_{ij}} \leq \mu_{\cos(\widetilde{\phi}_{ij})}, \quad \forall i, j. \tag{65}$$

Since $\widetilde{\tau}_{ij}$ and $\widetilde{\phi}_{ij}$ are fuzzy numbers, and the cosine function is decreasing on $[0, \pi]$, this condition may be expressed as:

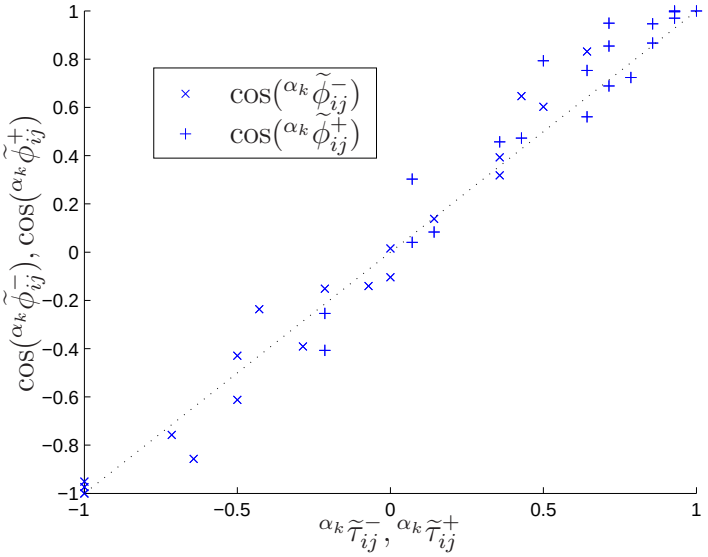

Fig. 16. Sensory data: Shepard's diagram of the spherical MDS, least-squares model.

$$[^{\alpha_k}\widetilde{\tau}_{ij}^{-}, ^{\alpha_k}\widetilde{\tau}_{ij}^{+}] \subseteq \left[\cos\left(\max(^{\alpha_k}\widetilde{\phi}_{ij})\right), \cos\left(\min(^{\alpha_k}\widetilde{\phi}_{ij})\right)\right] \tag{66}$$

$$\subseteq \left[\cos(^{\alpha_k}\widetilde{\phi}_{ij}^{-}), \cos(^{\alpha_k}\widetilde{\phi}_{ij}^{+})\right] \quad \forall i, j, k. \tag{67}$$

As in Section 4.1.3, we assume that the centers $\mathbf{c}_i, i = 1, \ldots, n$ have already been computed using, e.g., the least-squares algorithm of the previous section. The problem is to determine the "smallest" fuzzy regions S_i satisfying condition (66). This may be obtained by solving the following LP problem:

$$\text{Minimizing } \sum_{k=1}^{c}\sum_{i=1}^{n} {}^{\alpha_k}\beta_i, \tag{68}$$

subject to:

$$^{\alpha_k}\beta_i \leq {}^{\alpha_{k+1}}\beta_i, \ \forall k < c, \ \forall i, \tag{69}$$

$$^{\alpha_1}\beta_i \geq 0, \ \forall i, \tag{70}$$

$$^{\alpha_c}\beta_i \leq \pi, \ \forall i, \tag{71}$$

$$^{\alpha_k}\beta_i + {}^{\alpha_k}\beta_j \geq \max\left(\arccos(^{\alpha_k}\widetilde{\tau}_{ij}^{-}) - \phi_{ij}, \phi_{ij} - \arccos(^{\alpha_k}\widetilde{\tau}_{ij}^{+})\right), \ \forall i, j, k. \tag{72}$$

Example 5 *Sensory data set.* Figure 17 shows the representation of the sensory data set of Example 4 obtained using the possibilistic fitting procedure.

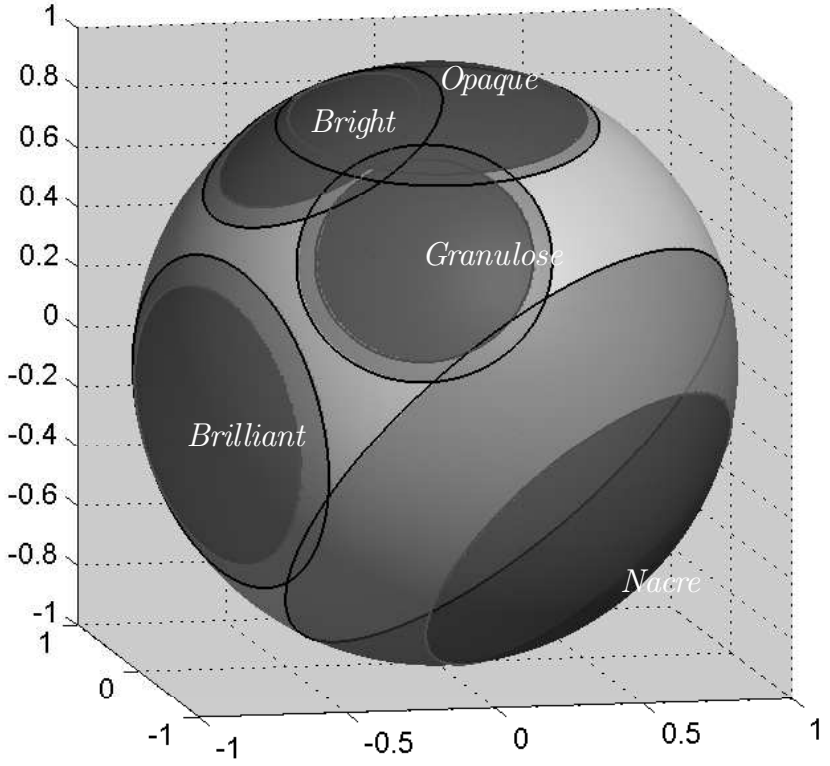

Fig. 17. Sensory data: spherical MDS, possibility model, α-cuts $\{0^+, 0.9\}$.

As expected, the imprecision of this representation is higher than using the least-squares method, which is also confirmed by the Shepard's diagram shown in Figure 18. This is particularly true for the spherical caps corresponding to 0.9-level cuts, which are much larger than in Figure 15. Otherwise, both representations basically lend themselves to the same interpretation.

5 Conclusion

This article focused on the problem of visualizing interval and fuzzy dissimilarity data. The basic principle underlying our approach is that imprecise dissimilarities should be represented by regions instead of points in the chosen space, which led us to propose extensions of standard MDS procedures. Two

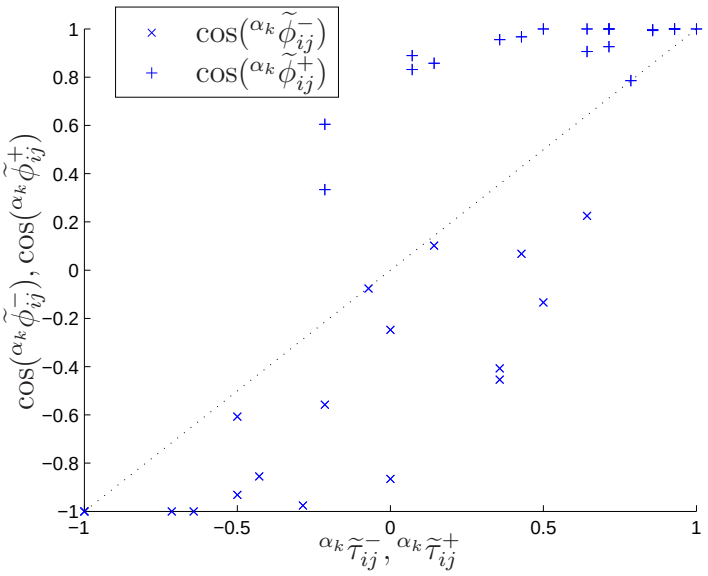

Fig. 18. Sensory data: Shepard's diagram of the spherical MDS, possibility model.

kinds of representation spaces were investigated, depending on the interpretation of the dissimilarity data: those corresponding to distances are better represented in a Euclidean space, whereas a spherical space is more relevant in the case of correlation measurements. The obtained representations have been presented in low dimensions, but they may be easily extended to higher dimensions, using planar projections.

Two approaches were proposed to compute optimal configurations. The least-squares procedure aims at minimizing the approximation errors, thus providing readable "compromise" solutions. In the configurations obtained using this method, the size of the regions reflects the imprecision of the input data: objects whose proximities to others have been specified in a less precise way tend to be represented by larger regions.

The second approach (refered to as the possibilistic procedure) starts from fixed centers determined, e.g., by the previous method, and computes the smallest regions such that reconstructed distance intervals include observed dissimilarity intervals. We thus obtain an exact, "worst-case", representation,

in contrast to the compromise solution provided by the least-squares method alone. The induced excess of reconstructed imprecision reflects the estimation errors, i.e., the inadequacy of the model. However, this interesting property is obtained at the cost of very high imprecision of the representation, which may make it less readable, and higher sensitivity to input data. In practice, these two methods can be seen as complementary, and we recommend to apply them jointly, as they provide different views of the data. They are currently being applied to the analysis of subjective evaluations collected during sensory testings for the car industry.

References

Bertoluzza, C., Gil, M. A., Eds, D. A. R., 2002. Statistical analysis and management of fuzzy data. Physica-Verlag, Heidelberg.

Borg, I., Groenen, P.J.F., 1997. Modern multidimensional scaling. Springer, New-York.

Cox, T. F., Cox, M. A. A., 1991. Multidimensional scaling on a sphere. Communications in Statistics 20, 2943–2953.

Cox, T. F., Cox, M. A. A., 1994. Multidimensional Scaling. Chapman and Hall, London.

Denoeux, T., Masson, M.-H., 2000. Multidimensional scaling of interval-valued dissimilarity data. Pattern Recognition Letters 21, 83–92.

Denœux, T., Masson, M.-H., 2004. Principal component analysis of fuzzy data using autoassociative neural networks. IEEE Transactions on Fuzzy Systems 12 (3), 336–349.

Denœux, T., Masson, M.-H., Hébert, P.-A., 2005. Nonparametric rank-based statistics and significance tests for fuzzy data. Fuzzy Set and Systems (To appear).

Diamond, P., 1988. Fuzzy least squares. Information Sciences 46, 141–157.

Diamond, P., Tanaka, H., 1998. Fuzzy regression analysis. In: Slowinski, R. (Ed.), Fuzzy sets in decision analysis, operations research and statistics. Kluwer Academic Publishers, Boston, pp. 349–387.

Diday, E., Bock, H., 2000. Analysis of symbolic data. Exploratory methods for extracting statistical information from complex data. Springer Verlag, Heidelberg.

Eckman, G., 1954. Dimension of color vision. Journal of Psychology 38, 367–474.

Gebhardt, J., Gil, M. A., Kruse, R., 1998. Fuzzy set-theoretic methods in statistics. In: Slowinski, R. (Ed.), Fuzzy sets in decision analysis, operations research and statistics. Kluwer Academic Publishers, Boston, pp. 311–347.

Groenen, P.J.F, Winsberg, S., Rodriguez, O., Diday, E., 2005. SymScal: symbolic multidimensional scaling of interval dissimilarities. Tech. Rep. EI 2005-15, Econometric Institute.

Hébert, P.-A., Masson, M.-H., Denœux, T., 2003. Fuzzy rank correlation between fuzzy numbers. In: Proceedings of the 10th IFSA World congress. Istanbul, Turkey, pp. 224–227.

Helm, C. E., 1964. A multidimensional ration scaling analysis of perceived color relations. Journal of the Optical Society of America 54, 256–262.

Kaufmann, A., Gupta, M. M., 1991. Introduction to fuzzy arithmetic. Theory and applications. International Thomson Computer Press, London.

Kruse, R., Meyer, K. D., 1987. Statistics with vague data. Reidel, Dordrecht.

Kruskal, J. B., 1964. Nonmetric multidimensional scaling: a numerical method. Psychometrika 29, 115–129.

Liu, S.-T., Kao, C., 2002. Fuzzy measures for correlation coefficient of fuzzy numbers. Fuzzy Sets and Systems 128 (2), 267–275.

Masson, M.-H., Denœux, T., 2002. Multidimensional scaling of fuzzy dissimilarity data. Fuzzy sets and systems 128 (33), 55–68.

Ming, M., Friedman, M., Kandel, A., 1997. General fuzzy least squares. Fuzzy sets and systems 88, 107–118.

Okada, A., Imaizumi, T., 1987. Geometric models for asymmetric similarity data. Behaviormetrika 21, 81–96.

Pietersz, R., Groenen, P., 2004. Rank reduction of correlation matrices by majorization. Quantitative Finance 4, 649–662.

Schiffman, S., Reynolds, M., Young, F., 1981. Introduction to Multidimensional Scaling: Theory, Methods and Applications. Academic Press, New-York.

Tanaka, H., Uejima, S., Asa, K., 1982. Fuzzy linear regression models. IEEE Trans. Systems, Man and Cybernetics 12, 903–907.

Viertl, R., 1996. Statistical methods for non-precise data. CRC Press, Boca

Raton.

Zadeh, L. A., 1975. The concept of a linguistic variable and its application to approximate reasoning (part 1). Information Sciences 8, 199–249.

A Spherical interval model: stress function derivatives for least square fitting

In the interval model, the following stress function has to be minimized with respect to the centers and the imprecision angles of the regions:

$$\sigma'(\mathcal{S}) = \sum_{i<j} \left(\cos(\phi_{ij}^-) - \tau_{ij}^-\right)^2 + \sum_{i<j} \left(\cos(\phi_{ij}^+) - \tau_{ij}^+\right)^2, \tag{A.1}$$

A.1 Partial derivatives w.r.t. the spherical coordinates of center $\mathbf{c}_i$

Denoting as θ_{ik} ($k < p-1$) the k-th spherical coordinate of vector $\mathbf{c}_i$, we have

$$\frac{\partial \sigma'(\mathcal{S})}{\partial \theta_{ik}} = \sum_{i<j} \frac{\partial \sigma'(\mathcal{S})}{\partial \phi_{ij}^-} \frac{\partial \phi_{ij}^-}{\partial \theta_{ik}} + \sum_{i<j} \frac{\partial \sigma'(\mathcal{S})}{\partial \phi_{ij}^+} \frac{\partial \phi_{ij}^+}{\partial \theta_{ik}}, \tag{A.2}$$

with

$$\frac{\partial \sigma'(\mathcal{S})}{\partial \phi_{ij}^-} = -2\left(\cos(\phi_{ij}^-) - \tau_{ij}^-\right)\sin(\phi_{ij}^-) \tag{A.3}$$

$$\frac{\partial \sigma'(\mathcal{S})}{\partial \phi_{ij}^+} = -2\left(\cos(\phi_{ij}^+) - \tau_{ij}^+\right)\sin(\phi_{ij}^+), \tag{A.4}$$

and, from (39) and (40)

$$\frac{\partial \phi_{ij}^-}{\partial \theta_{ik}} = \frac{\partial \phi_{ij}}{\partial \theta_{ik}} 1_{[0,\pi]}(\phi_{ij} + \beta_i + \beta_j) \tag{A.5}$$

$$\frac{\partial \phi_{ij}^+}{\partial \theta_{ik}} = \frac{\partial \phi_{ij}}{\partial \theta_{ik}} 1_{[0,\pi]}(\phi_{ij} - \beta_i - \beta_j). \tag{A.6}$$

Finally,

$$\frac{\partial \phi_{ij}}{\partial \theta_{ik}} = \frac{\partial \arccos < \mathbf{c}_i, \mathbf{c}_j >}{\partial \theta_{ik}} \tag{A.7}$$

$$= -\frac{1}{\sqrt{1 - < \mathbf{c}_i, \mathbf{c}_j >^2}} \cdot \frac{\partial < \mathbf{c}_i, \mathbf{c}_j >}{\partial \theta_{ik}} \tag{A.8}$$

$$= -\frac{1}{\sin \phi_{ij}} \cdot \frac{\partial \sum_{l=1}^{p} c_{il} c_{jl}}{\partial \theta_{ik}} \tag{A.9}$$

$$= -\frac{1}{\sin \phi_{ij}} \cdot \sum_{l=1}^{p} c_{jl} \frac{\partial c_{il}}{\partial \theta_{ik}} \tag{A.10}$$

In the case $\mathbf{S}^p$ is a circle ($p = 2$):

	c_{il}	$\frac{\partial c_{il}}{\partial \theta_{i1}}$
$l = 1$	$\cos \theta_{i1}$	$-\sin \theta_{i1}$
$l = 2$	$\sin \theta_{i1}$	$\cos \theta_{i1}$

In the case $\mathbf{S}^p$ is a sphere ($p = 3$):

	c_{il}	$\frac{\partial c_{il}}{\partial \theta_{i1}}$	$\frac{\partial c_{il}}{\partial \theta_{i2}}$
$l = 1$	$\cos \theta_{i1} \sin \theta_{i2}$	$-\sin \theta_{i1} \sin \theta_{i2}$	$\cos \theta_{i1} \cos \theta_{i2}$
$l = 2$	$\sin \theta_{i1} \sin \theta_{i2}$	$\cos \theta_{i1} \sin \theta_{i2}$	$\sin \theta_{i1} \cos \theta_{i2}$
$l = 3$	$\cos \theta_{i2}$	0	$-\sin \theta_{i2}$

A.2 Partial derivatives w.r.t. parameters b_i

We have

$$\frac{\partial \sigma'(\mathcal{S})}{\partial b_i} = \sum_{i<j} \frac{\partial \sigma'(\mathcal{S})}{\partial \phi_{ij}^-} \frac{\partial \phi_{ij}^-}{\partial b_i} + \frac{\partial \sigma'(\mathcal{S})}{\partial \phi_{ij}^+} \frac{\partial \phi_{ij}^+}{\partial b_i}, \tag{A.11}$$

with

$$\frac{\partial \phi_{ij}^{-}}{\partial b_i} = \frac{\partial \phi_{ij}^{-}}{\partial \beta_i}\frac{\partial \beta_i}{\partial b_i} \tag{A.12}$$

$$= 1_{[0,\pi]}(\phi_{ij} + \beta_i + \beta_j) \cdot \frac{\partial \beta_i}{\partial b_i} \tag{A.13}$$

$$\frac{\partial \phi_{ij}^{+}}{\partial b_i} = \frac{\partial \phi_{ij}^{+}}{\partial \beta_i}\frac{\partial \beta_i}{\partial b_i} \tag{A.14}$$

$$= -1_{[0,\pi]}(\phi_{ij} - \beta_i - \beta_j) \cdot \frac{\partial \beta_i}{\partial b_i}\ . \tag{A.15}$$

Finally,

$$\frac{\partial \beta_i}{\partial b_i} = -\pi \frac{\exp(-b_i)}{(1+\exp(-b_i))^2} \cdot (-1) \tag{A.16}$$

$$= \frac{\pi \exp(-b_i)}{(1+\exp(-b_i))^2}. \tag{A.17}$$

B Spherical fuzzy model: stress function derivatives for least square fitting

In the fuzzy model, the following stress function has to be minimized with respect to the centers and the nested imprecision angles of the regions:

$$\sigma''(\widetilde{\mathcal{S}}) = \sum_{k=1}^{c}\sum_{i<j}\left(\cos({}^{\alpha_k}\widetilde{\phi}_{ij}^{-}) - {}^{\alpha_k}\widetilde{\tau}_{ij}^{-}\right)^2 + \sum_{k=1}^{c}\sum_{i<j}\left(\cos({}^{\alpha_k}\widetilde{\phi}_{ij}^{+}) - {}^{\alpha_k}\widetilde{\tau}_{ij}^{+}\right)^2, \tag{B.1}$$

with

$${}^{\alpha}\widetilde{\phi}_{ij}^{-} = \min(\pi, \phi_{ij} + {}^{\alpha}\beta_i + {}^{\alpha}\beta_j) \tag{B.2}$$

$${}^{\alpha}\widetilde{\phi}_{ij}^{+} = \max(0, \phi_{ij} - {}^{\alpha}\beta_i - {}^{\alpha}\beta_j), \tag{B.3}$$

and

$${}^{\alpha_k}\beta_i = \frac{\pi}{1+\exp\left(-\left({}^{\alpha_1}b_i + \sum_{h=2}^{k} {}^{\alpha_h}{b_i}^2\right)\right)}. \tag{B.4}$$

B.1 Partial derivatives w.r.t. the spherical coordinates of center $\mathbf{c}_i$

Denoting as θ_{ik} $(k < p-1)$ the k-th spherical coordinate of vector $\mathbf{c}_i$, we have

$$\frac{\partial \sigma''(\mathcal{S})}{\partial \theta_{ik}} = \sum_{l=1}^{c} \sum_{i<j} \frac{\partial \sigma''(\mathcal{S})}{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{-}} \frac{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{-}}{\partial \theta_{ik}} + \sum_{l=1}^{c} \sum_{i<j} \frac{\partial \sigma''(\mathcal{S})}{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{+}} \frac{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{+}}{\partial \theta_{ik}}, \tag{B.5}$$

with

$$\frac{\partial \sigma''(\mathcal{S})}{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{-}} = -2 \left(\cos({}^{\alpha_l}\widetilde{\phi}_{ij}^{-}) - {}^{\alpha_k}\widetilde{\tau}_{ij}^{-} \right) \sin({}^{\alpha_l}\widetilde{\phi}_{ij}^{-}) \tag{B.6}$$

$$\frac{\partial \sigma''(\mathcal{S})}{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{+}} = -2 \left(\cos({}^{\alpha_l}\widetilde{\phi}_{ij}^{+}) - {}^{\alpha_k}\widetilde{\tau}_{ij}^{+} \right) \sin({}^{\alpha_l}\widetilde{\phi}_{ij}^{+}), \tag{B.7}$$

and

$$\frac{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{-}}{\partial \theta_{ik}} = \frac{\partial \phi_{ij}}{\partial \theta_{ik}} 1_{[0,\pi]}(\phi_{ij} + {}^{\alpha_l}\beta_i + {}^{\alpha_l}\beta_j) \tag{B.8}$$

$$\frac{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{+}}{\partial \theta_{ik}} = \frac{\partial \phi_{ij}}{\partial \theta_{ik}} 1_{[0,\pi]}(\phi_{ij} - {}^{\alpha_l}\beta_i - {}^{\alpha_l}\beta_j). \tag{B.9}$$

Finally, $\dfrac{\partial \phi_{ij}}{\partial \theta_{ik}}$ is given by (A.7)-(A.10).

B.2 Partial derivatives w.r.t. parameters ${}^{\alpha_h}b_i$

We have

$$\frac{\partial \sigma''(\mathcal{S})}{\partial {}^{\alpha_h}b_i} = \sum_{l=1}^{c} \sum_{i<j} \frac{\partial \sigma''(\mathcal{S})}{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{-}} \frac{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{-}}{\partial {}^{\alpha_h}b_i} + \sum_{l=1}^{c} \sum_{i<j} \frac{\partial \sigma''(\mathcal{S})}{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{+}} \frac{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{+}}{\partial {}^{\alpha_h}b_i}, \tag{B.10}$$

with

$$\frac{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{-}}{\partial {}^{\alpha_h}b_i} = \frac{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{-}}{\partial {}^{\alpha_l}\beta_i} \frac{\partial {}^{\alpha_l}\beta_i}{\partial {}^{\alpha_h}b_i} \tag{B.11}$$

$$= 1_{[0,\pi]}(\phi_{ij} + {}^{\alpha_l}\beta_i + {}^{\alpha_l}\beta_j) \cdot \frac{\partial {}^{\alpha_l}\beta_i}{\partial {}^{\alpha_h}b_i} \tag{B.12}$$

$$\frac{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{+}}{\partial b_i} = \frac{\partial {}^{\alpha_l}\widetilde{\phi}_{ij}^{+}}{\partial {}^{\alpha_l}\beta_i} \frac{\partial {}^{\alpha_l}\beta_i}{\partial {}^{\alpha_h}b_i} \tag{B.13}$$

$$= -1_{[0,\pi]}(\phi_{ij} - {}^{\alpha_l}\beta_i - {}^{\alpha_l}\beta_j) \cdot \frac{\partial {}^{\alpha_l}\beta_i}{\partial {}^{\alpha_h}b_i} \ . \tag{B.14}$$

Then,

$$\frac{\partial^{\alpha_l}\beta_i}{\partial^{\alpha_h}b_i} = \frac{-\pi \exp\left(-\left({}^{\alpha_1}b_i + \sum_{k=2}^{l} {}^{\alpha_k}{b_i}^2\right)\right)}{\left(1+\exp\left(-\left({}^{\alpha_1}b_i + \sum_{k=2}^{l} {}^{\alpha_k}{b_i}^2\right)\right)\right)^2} \cdot \frac{\partial\left({}^{\alpha_1}b_i + \sum_{k=2}^{l} {}^{\alpha_k}{b_i}^2\right)}{\partial^{\alpha_h}b_i}\ , \quad \text{(B.15)}$$

with

$$\frac{\partial\left({}^{\alpha_1}b_i + \sum_{k=2}^{l} {}^{\alpha_k}{b_i}^2\right)}{\partial^{\alpha_h}b_i} = \begin{cases} 1, & \text{if } h = 1; \\ 2({}^{\alpha_h}b_i), & \text{if } h \in \{2,\ldots,l\}; \\ 0, & \text{otherwise.} \end{cases} \quad \text{(B.16)}$$