Multidimensional scaling of fuzzy dissimilarity data

M. Masson and T. Denœux

Université de Technologie de Compiègne
U.M.R CNRS 6599 Heudiasyc
BP 20529 - F-60205 Compiègne cedex - France
email: `Mylene.Masson@@hds.utc.fr`

Abstract

Multidimensional scaling is a well-known technique for representing measurements of dissimilarity among objects as distances between points in a p-dimensional space. In this paper, this method is 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. 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.
Keywords: Fuzzy data analysis, Multidimensional scaling, Fuzzy dissimilarity.

1 Introduction

Multidimensional scaling (MDS) [1] [3] is classical technique for analyzing similarity or dissimilarity measurements among a set of n objects. MDS attempts to represent dissimilarities as *distances* between points in a low-dimensional Euclidean space, each point corresponding to one object. It provides easily interpretable graphical displays (*i.e.* maps of objects), in which similar objects are mapped close to each other, and dissimilar objects are represented as points distant from one another. An early use of MDS has been dimension reduction [13]. In this approach, pairwise distances are computed between objects, originally described as feature vectors in a high-dimensional feature space. We then seek a configuration of points in a lower dimensional space such that the inter-point distances approximate as well as possible the original distances. However, much of the theory related to MDS has been developed for analyzing proximity data that arise in social sciences, or in disciplines like product development and marketing. In this case, the dissimilarities between the objects, referred to as stimuli (e.g. linguistic concepts, food products, acoustical sounds, colors, smells, etc), are often provided by a human subject who is asked to rate the perceived dissimilarity between all possible pairs of objects on a specified scale (e.g. score between 0 for "no difference" and 10 for "maximal difference"). The goal is to visualize the perceived differences and to discover underlying dimensions (forming a "subjective space") that would explain dissimilarity judgments. In such contexts, the input data consists of a

$n \times n$ squared matrix of real numbers, $\Delta = (\delta_{ij})$, δ_{ij} being the dissimilarity between object i and the object j.

In this paper, we consider the case where dissimilarities are no longer real values, as in the classical theory of MDS, but *intervals*, or, more generally, *fuzzy numbers.* A first justification for considering imprecise dissimilarities lies in the difficulty for a judge to quantify the proximity of certain pairs of objects. It may then be preferable to account for the indeterminacy of dissimilarity judgment by explicitly allowing some imprecision or vagueness in the answers. A judge can be asked to provide an interval, or to express his answers using linguistic labels, which can in turn be converted into fuzzy numbers.

A second justification for considering interval-valued or fuzzy data is to provide an alternative tool for the analysis of three-way dissimilarity data. This kind of data arises when N evaluators (e.g. a panel of consumers) are used. The dissimilarity between two objects i and j is then characterized by an empirical distribution of N values $\delta_{ij,1}, \ldots, \delta_{ij,N}$, where $\delta_{ij,k}$ denotes the dissimilarity between objects i and j, assessed by judge k. The input data then consists of a collection of N dissimilarity matrices $\Delta^k = (\delta^k_{ij})$, $k = 1, \ldots, N$. A simple way to analyze such data is to average the N matrices. However, it is clear that a lot of information is lost during the averaging process. Another popular method is the INDSCAL model (INdividual DIfferences SCALing) [2]. This model comprises two spaces: the object space and the subject space (of the same dimension). The evaluators are assumed to share a common object space explaining the observed dissimilarities, but they can weight differently the underlying dimensions. The coordinates of each judge in the subject space are given by the weights. The INDSCAL model provides interesting insights into the structure of the data, but it usually produces complex results requiring careful interpretation and a good deal of expertise. A potentially simpler and more intuitive approach could be to globally describe the distribution of N dissimilarity values for each object pair by an interval or a fuzzy number, and to proceed with the analysis of the resulting $n \times n$ interval-valued or fuzzy dissimilarity matrix. Such an approach is made possible by the techniques introduced in this paper.

To account for imprecision and vagueness in the dissimilarity data, we propose to represent each object no longer as a point in $\mathbb{R}^p$, but as a region (crisp or fuzzy) accounting for both imprecision and uncertainty in the input data.

Two models are presented. They differ by the nature of the link imposed between the original dissimilarities and the approximating distances in the Euclidean configuration space. The first model is derived from a simple least-squares approach, and is a natural generalization of classical MDS. The other model, referred to as the *possibility model*, is inspired from a fuzzy regression method introduced by Tanaka [14] [9].

In the next section, we start by a brief description of the standard multidimensional scaling method. We then explain in Section 3 how to derive the two models in the case of interval-valued dissimilarities. Several examples are used to illustrate the properties of both methods. In Section 4, the algorithms are extended to the case where the dissimilarities are fuzzy numbers. As an illustration, these techniques are finally applied in Section 5 to a classical dataset from a perception experiment.

2 Multidimensional Scaling

In this section, we briefly outline the basic principles of the standard MDS approach, using a hypothetical data set as an illustration. For a more complete description of the many theoretical and practical aspects of MDS, the reader is invited to refer to recent monographs on this subject such as, e.g., Refs. [1], [3] and [15].

Let us assume that the available data consists in a $n \times n$ dissimilarity matrix $\Delta = (\delta_{ij})$, where δ_{ij} denotes the dissimilarity between objects i and j. The problem is to find the coordinates of n objects in a p-dimensional *configuration space*, in the form of a $n \times p$ matrix $X = (x_{il})$, such that the inter-point distances matrix $D(X) = (d_{ij}(X))$ approximates as well as possible Δ. The Euclidean distance is often chosen to measure proximities in the configuration space. We then have:

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

However, other distance measures can be considered as well. To assess the quality of the approximation of Δ by $D(X)$, the following loss function, known in the literature as the *Stress* function [11], is used:

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

This function is invariant under any isometric transformations (rotations, translations and reflections). The optimal configuration of points X, starting from an initial random guess, is obtained through iterative minimization of (2).

Example 1 As an example, assume that a human subject was asked to mentally assess pairwise distances between ten European cities. The estimated distances are taken as the centers of the intervals shown in Table 1 (the processing of interval-valued distances is deferred until the next section). Given these estimates, we seek to infer the relative positions of the cities, i.e., to reconstruct the map of Figure 1. The dissimilarity matrix was normalized so that its standard deviation equals one. The resulting map is shown in Figure 2. 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 quality of the representation can be judged from a scatterplot, usually referred to as the *Shepard diagram*, given in Figure 3. This plot shows the reconstructed distances on the vertical axis versus the original dissimilarities on the horizontal axis. In this case, the points in the Shepard diagram are close to the diagonal, which indicates that Figure 2 is a relatively good representation of the data. Large deviations from a perfect linear fit would indicate the inability of the Euclidean model to correctly represent the data[1] or a bad choice of the configuration dimensionality.

[1] This may occur, in particular, when the input dissimilarities are interpreted as distances in a psychological space, which does not necessarily possess a Euclidean structure [1].

3 Scaling of interval-valued dissimilarities

Let us now assume the available data to consist in a symmetric matrix $\Delta = ([\delta_{ij}])$ of interval-valued dissimilarities. Each interval $[\delta_{ij}] = [\delta_{ij}^-, \delta_{ij}^+]$ represents a set of plausible values for the dissimilarity between objects i and j.

To account for the imprecision in the determination of dissimilarities, we seek a representation of each object i as a *region* R_i in a p-dimensional feature space. Let us denote as d_{ij}^- and d_{ij}^+, respectively, the minimum and maximum Euclidean distances between any two regions R_i and R_j:

$$d_{ij}^- = \min_{\mathbf{x}_i \in R_i, \mathbf{x}_j \in R_j} \|\mathbf{x}_i - \mathbf{x}_j\| \tag{3}$$

$$d_{ij}^+ = \max_{\mathbf{x}_i \in R_i, \mathbf{x}_j \in R_j} \|\mathbf{x}_i - \mathbf{x}_j\|. \tag{4}$$

In practice, a parameterized shape has to be chosen for regions R_i. A very simple representation is obtained by defining each region R_i as a hypersphere (a circle when $p = 2$) with center $\mathbf{c}_i \in \mathbb{R}^p$ and radius r_i (Fig. 4). We then obtain a model with $n(p+1)$ parameters (n centers defined by p coordinates each, and n radii).

It is easy to see that, in that case, d_{ij}^- and d_{ij}^+ defined by Eqs (3) and (4) are, respectively, equal to:

$$d_{ij}^- = \max(0, d_{ij} - r_i - r_j) \tag{5}$$

$$d_{ij}^+ = d_{ij} + r_i + r_j, \tag{6}$$

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. Two algorithms are presented in the sequel, differing as to the nature of the relationship imposed between distances in the configuration space, and the input dissimilarities.

3.1 Least squares model

In a first approach, one may attempt to minimize the discrepancies between dissimilarity intervals $[\delta_{ij}^-, \delta_{ij}^+]$ and distance intervals $[d_{ij}^-, d_{ij}^+]$. For that purpose, the stress function (2) may be generalized as:

$$\sigma'(\mathcal{R}) = \sum_{i<j} (d_{ij}^- - \delta_{ij}^-)^2 + \sum_{i<j} (d_{ij}^+ - \delta_{ij}^+)^2, \tag{7}$$

where $\mathcal{R}$ denotes the set of n regions $\{R_1, \ldots, R_n\}$. The model parameters can then be determined by minimizing $\sigma'(\mathcal{R})$ with respect to $\mathcal{R}$, using an iterative gradient descent algorithm.

Details concerning this approach may be found in Ref. [4], where the analytical expression of the gradient of (7) with respect to the model parameters is given. Note that the minimization of $\sigma'(\mathcal{R})$ is a constrained optimization problem, since the radii r_i, $i = 1, \ldots, n$ must be kept positive. However, the conditions $r_i \geq 0$ may be enforced by introducing new variables ρ_i such that $r_i = \rho_i^2$; an unconstrained optimization procedure can then be applied.

Some insight into this model may be gained by studying the optimality conditions for the minimization of (7). Assuming the radii to be small, the derivatives of (7) with respect to r_k $(k = 1, n)$ are given by:

$$\frac{\partial \sigma'}{\partial r_k} = 2 \sum_{i \neq k} (d_{ik} + r_i + r_k - \delta^+_{ik}) - 2 \sum_{i \neq k} (d_{ik} - r_i - r_k - \delta^-_{ik}). \tag{8}$$

The optimality conditions can then be written as:

$$\sum_{i \neq k} (2r_i + 2r_k - (\delta^+_{ik} - \delta^-_{ik})) = 0 \qquad k = 1, \ldots, n, \tag{9}$$

or equivalently:

$$\sum_{i \neq k} (\delta^+_{ik} - \delta^-_{ik}) = 2(n-2)r_k + 2 \sum_{i=1}^{n} r_i \qquad k = 1, \ldots, n. \tag{10}$$

Summing (10) over k leads to the following result:

$$\sum_{k=1}^{n} r_k = \frac{1}{4n-4} \sum_{k=1}^{n} \sum_{i \neq k} (\delta^+_{ik} - \delta^-_{ik}). \tag{11}$$

Eqs (10) and (11) have interesting implications. First, (11) shows that, when applied to standard dissimilarity data ($\delta^-_{ij} = \delta^+_{ij} = \delta_{ij}$), the least-squares model leads to null radii ($r_i = 0, \forall i$). The regions R_i are then reduced to their centers $\mathbf{c}_i$, and the method is then strictly equivalent to the standard MDS method minimizing (2).

Secondly, it follows from (10) that the radii r_k are linearly related to the quantities

$$s_k = \sum_{i \neq k} (\delta^+_{ik} - \delta^-_{ik}), \tag{12}$$

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.

Example 2 To illustrate this point, let us come back to the interval-valued inter-city distances given in Table 1. The representation of the data obtained using the least-squares model is shown in Figure 5 (a). This figure 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 linear relationship between the radii r_k and the imprecision measures s_k, resulting from (10), is confirmed experimentally in this case (Figure 6). The least-squares model is thus able to render the overall vagueness (imprecision) in the input data.

3.2 Possibility model

The previous model gives a representation in which the distances approximate "as well as possible" (in the least squares sense) the dissimilarities. We seek in this section to give, in some sense, an *exact* representation of the input data. For that purpose, let us assume that the region centers $\mathbf{c}_i$ have already been determined by minimization of (2). Hence, the distances d_{ij} between centers are given. By analogy with Tanaka's possibilistic fuzzy regression model [14], one may attempt to find the smallest radii such that the following condition is satisfied:

$$[\delta_{ij}^-, \delta_{ij}^+] \subseteq [d_{ij}^-, d_{ij}^+] \quad \forall i, j. \tag{13}$$

This is a more 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 $[\delta_{ij}^-, \delta_{ij}^+]$. It leads to the following optimization problem:

$$\min_{\mathbf{r}} \sum_{i=1}^{n} r_i \tag{14}$$

subject to:

$$d_{ij}^- \leq \delta_{ij}^- \quad \forall i, j \tag{15}$$

$$d_{ij}^+ \geq \delta_{ij}^+ \quad \forall i, j \tag{16}$$

$$r_i \geq 0 \quad \forall i = 1, n \ , \tag{17}$$

In (14), $\mathbf{r}$ denotes the vector of radii $(r_1, r_2, ..., r_n)^t$. Using the expressions of d_{ij}^- and d_{ij}^- given by (5) and (6), constraints (15) and (16) may be written as

$$\max(0, d_{ij} - r_i - r_j) \leq \delta_{ij}^- \tag{18}$$

$$r_i + r_j \geq \delta_{ij}^+ - d_{ij}, \tag{19}$$

which may be expressed in a more compact form as

$$r_i + r_j \geq \max(d_{ij} - \delta_{ij}^-, \delta_{ij}^+ - d_{ij}) \quad \forall i, j. \tag{20}$$

The minimization of (14) under the constraints (17) and (20) is a linear programming (LP) problem. 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 1 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:

$$[d_{ij}^-, d_{ij}^+] \subseteq [\delta_{ij}^-, \delta_{ij}^+] \quad \forall i, j. \tag{21}$$

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.

EXAMPLE 3 Let us once again turn back to the data in Table 1. To illustrate the behavior of the possibility model in the presence of precise dissimilarity data, we shall first use the centers of the intervals shown in Table 1 as point estimates of inter-city distances (as in Example 1). The configuration obtained using the possibility model is shown in Figure 7. Note that, contrary to the least squares model, the possibility model does not produce pointwise representations of the objects in that case. Since we assume that $\delta_{ij}^- = \delta_{ij}^+ = \delta_{ij}$, Eq. (20) can now be rewritten as:

$$r_i + r_j \geq |d_{ij} - \delta_{ij}|, \tag{22}$$

which shows that the sizes of the regions are related to the estimation errors in the input dissimilarity matrix, or reveal the inadequacy of the specified model (i.e., the choice of the Euclidean distance and the dimensionality of the configuration) to the input data. Figure 8 shows a modified Shepard diagram for this example, in which the lower and upper distances d_{ij}^- and d_{ij}^+ are plotted against the dissimilarities δ_{ij}. It may be checked in this diagram that we have $d_{ij}^- \leq \delta_{ij} \leq d_{ij}^+$ for all $i \neq j$.

Let us now consider the application of possibility model to the *interval-valued dissimilarities* reported in Table 1. The configuration obtained is shown in Figure 5 (b). As expected, the circles representing each of the cities are larger than those obtained in the least-squares model. This is due to the fact that the possibility model represents both the imprecision in the data (as the least-squares model), *and the goodness-of-fit of the model.* To better show this, let us introduce the following quantities:

$$e_i = \sum_{j \neq i} |\delta_{ij} - d_{ij}|, \tag{23}$$

which can be used to measure the ability of the chosen model to reconstruct the distances involving object i. Figure 9 shows the radii r_i produced by the possibility model, plotted against the e_i and the s_i defined in (12). It is clear from this graph that the representation of an object i by a large circle is due either to a great imprecision in the input dissimilarity data regarding object i (measured by s_i), or to large differences between input and reconstructed distances between object i and the other objects (measured by e_i). In this example, the possibility model is thus able to reflect the two sources of uncertainty in input data:

- *imprecision* or *vagueness*, stemming from the difficulty for a human subject to precisely estimate the distance between two cities, and

- *uncertainty*, which is linked to estimation errors: the knowledge of the subject in geography may be limited, he/she may be completely wrong in his/her estimation.

4 Extension to fuzzy dissimilarities

In this section, the least-squares and possibilistic methods described in the previous section are extended to *fuzzy* dissimilarity data. More precisely, we shall assume each judgment of dissimilarity between two objects i and j, to be represented by a fuzzy number $\widetilde{\delta}_{ij}$. As already mentioned in Section 1, such data may arise either from the elicitation of dissimilarities as linguistic assessments (such as "very close", "quite different", etc.), or as a means to models a distribution of answers provided, e.g., by a panel of subjects.

To deal with this new kind of input data, we propose to represent each object by a fuzzy region $\widetilde{R}_i$ of $\mathbb{R}^p$. Using the extension principle [16] [6], the fuzzy distance between two fuzzy regions $\widetilde{R}_i$ et $\widetilde{R}_j$ can be defined as:

$$\mu_{\widetilde{d}_{ij}}(w) = \sup_{\mathbf{x},\mathbf{y}\in\mathbb{R}^p} \min(\mu_{\widetilde{R}_i}(\mathbf{x}), \mu_{\widetilde{R}_j}(\mathbf{y})), \tag{24}$$

where the supremum is computed under the constraint $\|\mathbf{x}-\mathbf{y}\| = w$.

More specifically, if $\widetilde{R}_i$ and $\widetilde{R}_j$ are multidimensional fuzzy numbers [10, p.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 $\widetilde{R}_j$.

Following the approach adopted in Section 3 for the representation of interval-valued data, we may consider regions $\widetilde{R}_i$ whose α-cuts are concentric hypersheres of radii ${}^\alpha r_i$ and center $\mathbf{c}_i$. We thus have

$$\alpha > \alpha' \Rightarrow {}^\alpha r_i \leq {}^{\alpha'} r_i \quad \forall \alpha, \alpha' \in [0,1]. \tag{25}$$

The bounds of the interval ${}^\alpha\widetilde{d}_{ij}$ are in that case expressed as follows:

$$\begin{aligned} {}^\alpha\widetilde{d}^-_{ij} &= \max(0, d_{ij} - {}^\alpha r_i - {}^\alpha r_j) && (26)\\ {}^\alpha\widetilde{d}^+_{ij} &= d_{ij} + {}^\alpha r_i + {}^\alpha r_j, && (27) \end{aligned}$$

where d_{ij} denotes, as before, the Euclidean distance between the centers $\mathbf{c}_i$ and $\mathbf{c}_j$. The next sections describe how to extend the previous algorithms to determine the parameters of the different regions such that the fuzzy distances represent, in some sense, the input fuzzy dissimilarities.

4.1 Least squares model

Let $\{\alpha_i\}_{i=1,c}$ denote a set of c predetermined levels of α-cut such that:

$$1 = \alpha_1 > \ldots > \alpha_c = 0 \tag{28}$$

The stress function (7) can be generalized as follows:

$$\sigma''(\widetilde{\mathcal{R}}) = \sum_{k=1}^{c}\sum_{i<j}({}^{\alpha_k}\widetilde{d}^-_{ij} - {}^{\alpha_k}\widetilde{\delta}^-_{ij})^2 + \sum_{k=1}^{c}\sum_{i<j}({}^{\alpha_k}\widetilde{d}^+_{ij} - {}^{\alpha_k}\widetilde{\delta}^+_{ij})^2, \tag{29}$$

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 [5] and extended by Ming and al. [12]. 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 $n \times c$ radii ${}^{\alpha_k}r_i$, $i = 1,\ldots,n$, $k = 1,\ldots,c$.

The minimization of $\sigma''(\widetilde{\mathcal{R}})$ is a constrained optimization problem, since the ${}^{\alpha_k}r_i$, $k = 1,\ldots,c$ are linked by $c-1$ inequality constraints of the form (25). However, these constraints may be easily taken into account by introducing new variables ${}^{\alpha_k}\rho_i \in \mathbb{R}$, $k = 1, \ldots, c$, such that:

$$ {}^{\alpha_k}r_i = \sum_{h=1}^{k} {}^{\alpha_h}\rho_i^2. \tag{30} $$

This reformulation transforms the minimization of (29) into an unconstrained non linear optimization problem.

4.2 Possibility model

We assume again c values α_k satisfying (28) to be chosen. In order to extend the LP formulation from interval-valued to fuzzy dissimilarities, we generalize condition (13) as

$$ \widetilde{\delta}_{ij} \subseteq \widetilde{d}_{ij}, \quad \forall i, j \tag{31} $$

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

$$ \mu_{\widetilde{\delta}_{ij}} \leq \mu_{\widetilde{d}_{ij}}, \quad \forall i, j. \tag{32} $$

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

$$ [{}^{\alpha_k}\widetilde{\delta}_{ij}^-, {}^{\alpha_k}\widetilde{\delta}_{ij}^+] \subseteq [{}^{\alpha_k}\widetilde{d}_{ij}^-, {}^{\alpha_k}\widetilde{d}_{ij}^+] \quad \forall i, j, k. \tag{33} $$

As in the Section 3.2, 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 (33). Following the same line of reasoning as in Section 3.2, this can done by solving, successively, the following LP problems, starting with $k = 1$ up to $k = c$:

$$ \min_{{}^{\alpha_k}\mathbf{r}} \sum_{i=1}^{n} {}^{\alpha_k}r_i \tag{34} $$

subject to:

$$ {}^{\alpha_k}r_i + {}^{\alpha_k}r_j \geq \max(d_{ij} - {}^{\alpha_k}\delta_{ij}^-, {}^{\alpha_k}\delta_{ij}^+ - d_{ij}) \quad \forall i, j \tag{35} $$

$$ {}^{\alpha_k}r_i \geq \begin{cases} 0 & \text{if } k = 1 \\ {}^{\alpha_{k-1}}r_i & \text{if } k > 1 \end{cases} \quad \forall i = 1, n. \tag{36} $$

with ${}^{\alpha_k}\mathbf{r} = ({}^{\alpha_k}r_1, \ldots, {}^{\alpha_k}r_n)^t$.

The application of the least-squares and possibility model will now be demonstrated using a real data set.

5 Application

5.1 Problem description

We consider in this section an experiment by Helm [8] reported in [1]. Different subjects were asked to judge the similarity of ten colors ranging over the entire spectrum from red to purple. The subjects were classified into two groups: some of them had a normal color vision, whereas the others had a color-deficient vision. Several experiments, including the work of Eckman [7], have shown that the vision of colors can be represented as a circular configuration in a two-dimensional space. Helm applied a classical MDS algorithm on the average dissimilarity matrix over the color-normal subjects and recovered this annular structure. Then, the same study was conducted based on the answers of color-deficient subjects; it clearly showed how the color vision of deficient-color subjects was modified compared to normal subjects.

The idea in this section is to take benefit from a fuzzy description of dissimilarities to study the effect of the variability in the subject answers. A triangular fuzzy number was chosen to represent each dissimilarity. The lower bound, upper bound and central value of each fuzzy number were, respectively, chosen to be the minimum, the maximum and the mean response over all subjects. Two separated analyses were conducted, one with five color-normal subjects, one with the same number of color-deficient subjects.

5.2 Results with the possibility model

We first present the results obtained using the possibility model described in Section 4.2. The input dissimilarities being defined as triangular fuzzy numbers, we chose to consider only $c = 2$ levels (the support and the core) for the optimization of the fuzzy regions.

Figure 10 presents the results obtained with the first group of subjects. It can be seen that the color annular structure is well-recovered. 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.

Fig. 11 reports the results obtained with the color-deficient subjects. The possibility model clearly indicates a greater confusion among colors. 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 is confirmed by Figures 12 and 13 showing the membership functions of some input dissimilarities $\widetilde{\delta}_{ij}$ and the corresponding reconstructed distances $\widetilde{d}_{ij}$, for the two groups of subjects. The agreement between dissimilarities and distances is obviously much worse for the color-blind group, which is compensated in the possibility model by greater imprecision.

5.3 Results with the least-squares model

The least-squares method was also applied to the same data as above. As before, only $c = 2$ levels were used for calculation of the generalized stress function (29). The configurations obtained for the normal and color-blind group are shown in Figures 14 and 15, respectively. Once again, the annular structure pointed out by Helm is well recovered, with some distorsion 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, and rather surprisingly at first glance, 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. Figures 16 and 17, 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.

6 Concluding remarks

Although the theory and practice of MDS has been well developed during the past decades (as reflected by recent monographs such as [1], [3] and [15]), the problem of dealing explicitly with vagueness, imprecision and ambiguity in dissimilarity judgments does not seem to have received all the attention it deserves. Indeed, the literature on MDS, particularly in application domains such as sensory analysis, abounds with examples showing the difficulty for human subjects to provide precise dissimilarity assessments (see, e.g., [1]). The most frequent strategy to cope with this problem has been to reduce the data to some kind of "average" or "compromise" matrix, and to proceed with standard MDS algorithm. It is clear, however, that a great deal of information is usually lost in this process.

In this paper, two fuzzy versions of the standard MDS method have been proposed. The techniques developed allow the analysis of imprecise dissimilarities, expressed or modeled by intervals or fuzzy numbers. As a result of dealing explicitly with the imperfectness of the data, graphical displays are produced, in which each object is no longer represented as a point, but as a crisp or a fuzzy region. The size and fuzziness of these regions provide meaningful information regarding the imprecision and ambiguity of the input data.

How do our two models compare, and which one is better for which application ? Obviously, the possibilistic model is a pessimistic one, since it imposes all the reconstructed distances to enclose the input dissimilarities. On some occasions, this may be asking too much, and may result in overly large regions. However, the possibilistic model does seem to provide interesting insights into the structure of the data, thanks to the representation of both imprecision, and estimation errors (reflecting the partial inadequacy of the model). In contrast, the least-squares model provides "approximate" or "compromise" solutions, which may sometimes be more readable. The configurations produced are closer to those that would be obtained using standard MDS, except that they carry additional information (the size of the regions) reflecting the imprecision of the data. In short, the two models appear to be complementary and equally useful. They are currently being applied to the analysis of subjective evaluations collected during sensory testings for the car industry.

Acknowledgments

The authors thank the anonymous referees for their valuable comments and suggestions.

References

[1] I. Borg and P. Groenen. Modern Multidimensional scaling, Springer, New-York, 1997.

[2] J. D. Carrol and J. J. Chang. Analysis of individual differences in multidimensional scaling via n-way generalization of Eckart-Young decomposition, Psychometrika 35 (1970) 283-319.

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

[4] T. Denœux and M. Masson. Multidimensional scaling of interval-valued dissimilarity data, Pattern Recognition Letters 21 (2000) 83-92.

[5] P. Diamond. Fuzzy least squares, Information Sciences 46 (1988) 141-157.

[6] D. Dubois and H. Prade. Possibility Theory: An approach to computerized processing of uncertainty, Plenum Press, New-York, 1988.

[7] G. Ekman. Dimension of color vision, Journal of Psychology 38 (1954) 367-474.

[8] C. E. Helm. A multidimensional ration scaling analysis of perceived color relations, Journal of the Optical Society of America 54 (1964) 256-262.

[9] J. Kacprzik and M. Fedrizzi. Fuzzy Regression Analysis, Physica-Verlag, Heidelberg, 1992.

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

[11] J. B. Kruskal. Multidimensional scaling by optimizing goodness-of-fit to a nonmetric hypothesis, Psychometrika 29 (1964) 1-27.

[12] M. Ming, M. Friedman and A. Kandel. General fuzzy least squares, Fuzzy sets and systems 88 (1997) 107-118.

[13] J. W. Sammon Jr. A non-linear mapping for data structure analysis. IEEE Transactions on Computers 18 (1969) 401-409.

[14] H. Tanaka, S. Uejima and K. Asai. Fuzzy linear regression models, IEEE Trans. Systems, Man and Cybernetics 12 (1982) 903-907.

[15] F. W. Young and R. M. Hamer. Theory and Applications of Multidimensional Scaling, Eribaum Associates, Hillsdale, NJ, 1994.

[16] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning (Part 1), Information Sciences 8 (1975) 199-249.

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

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

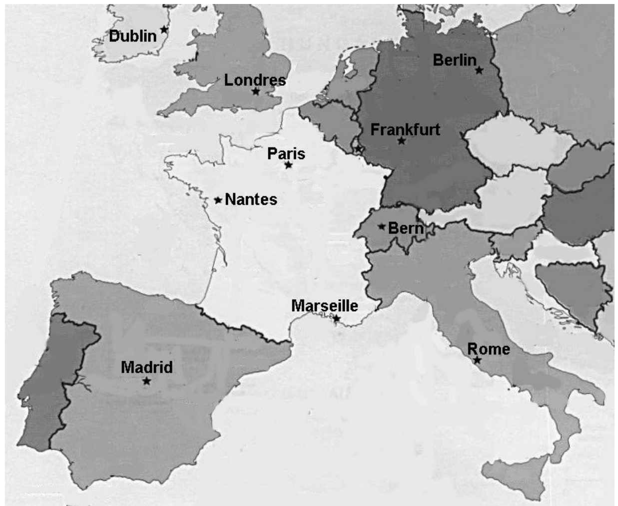

Figure 1: Map of Europe

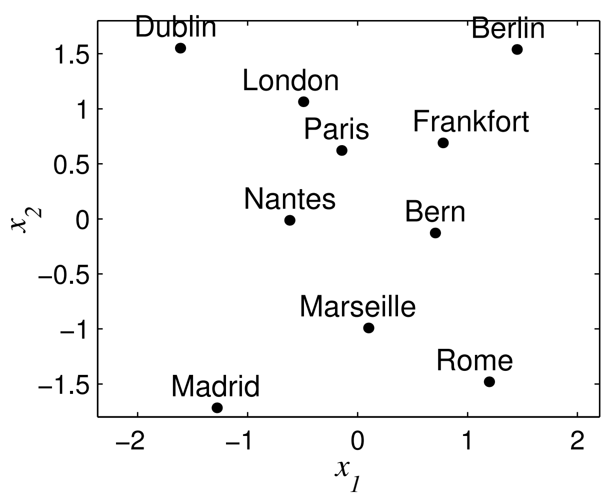

Figure 2: "Classical" MDS of the cities dataset: reconstructed map.

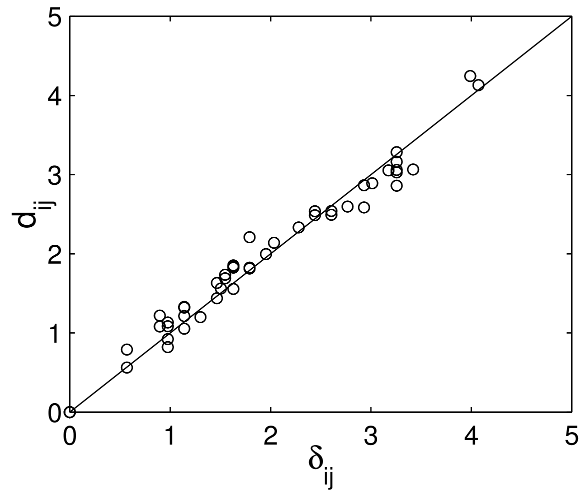

Figure 3: "Classical" MDS of the cities dataset: Shepard diagram.

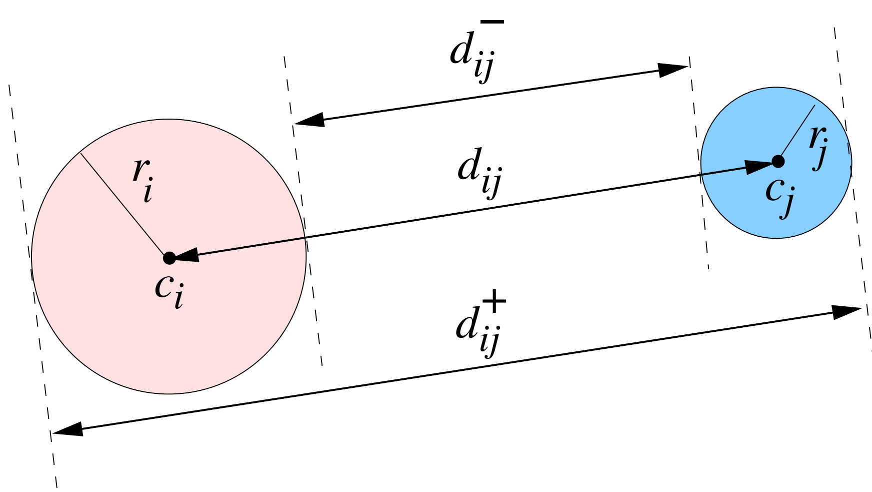

Figure 4: Maximum and minimum distances between two regions.

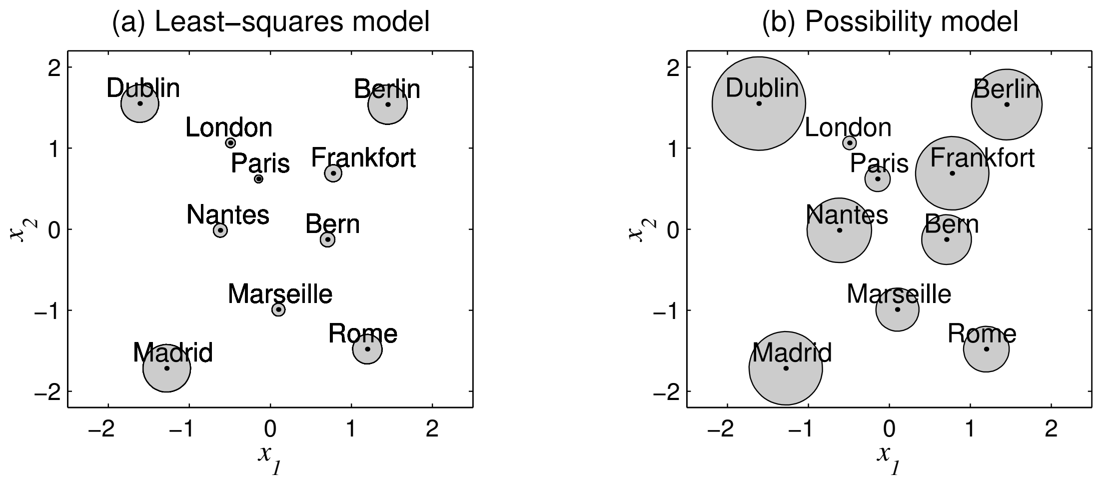

Figure 5: Multidimensional scaling of the cities dataset (interval-valued dissimilarities) using the least-squares model (a) and the possibility model (b).

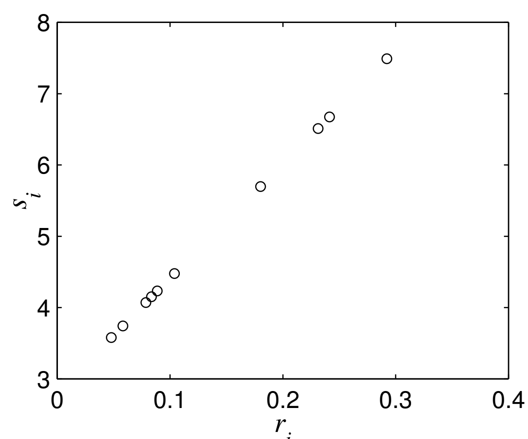

Figure 6: Radii of the configurations versus imprecision (s_i) for the least-squares model applied to the interval-valued city data.

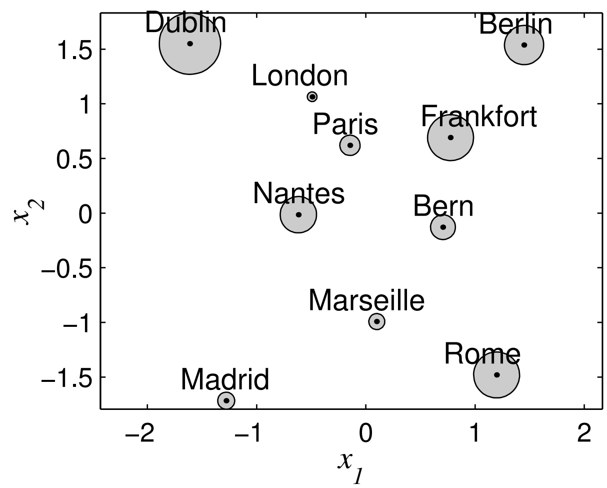

Figure 7: Multidimensional scaling of the cities dataset (real-valued dissimilarities) using the possibility model: reconstructed map.

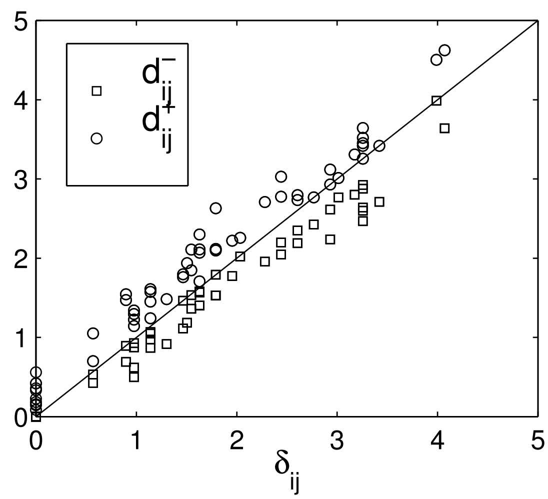

Figure 8: Multidimensional scaling of the cities dataset (real-valued dissimilarities) using the possibility model: Shepard diagram.

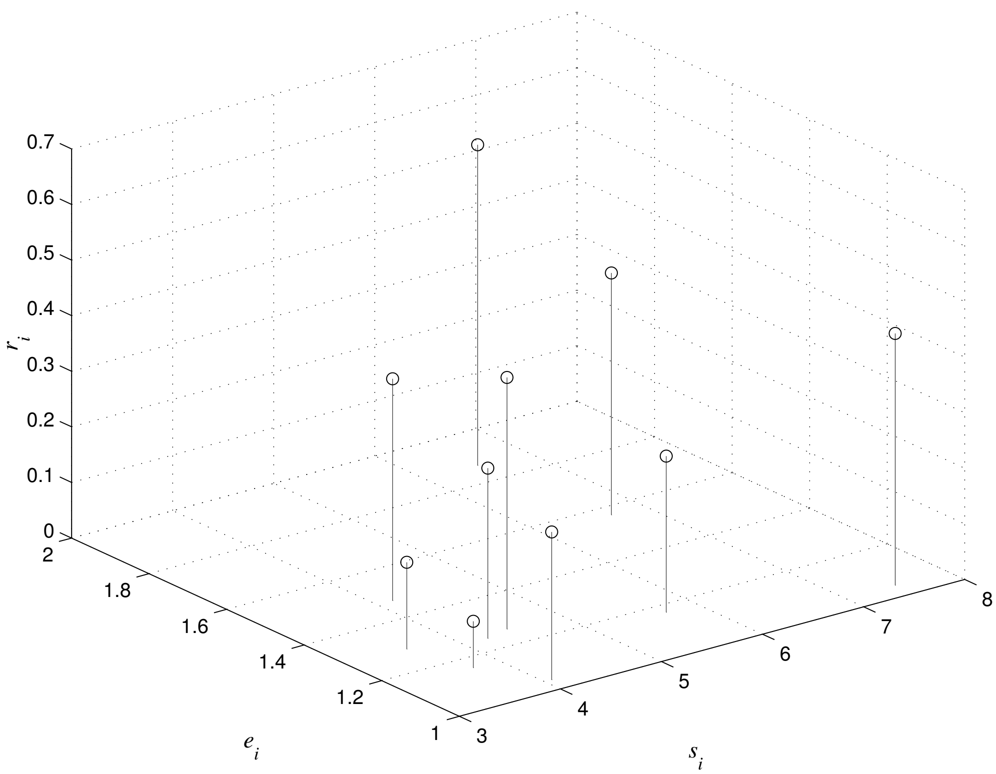

Figure 9: Radii of the configurations versus imprecision (s_i) and uncertainty e_i for the possibility model applied to the interval-valued city data.

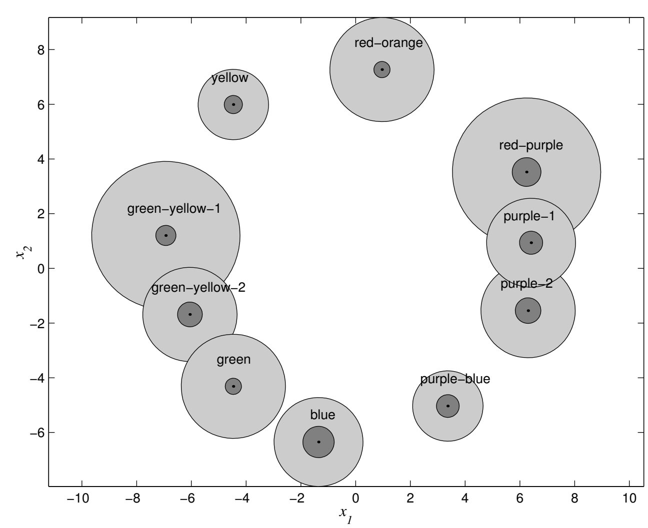

Figure 10: Reconstructed colors configuration using the possibility model in the case of normal subjects (light grey: supports; dark grey: level 1 α-cut).

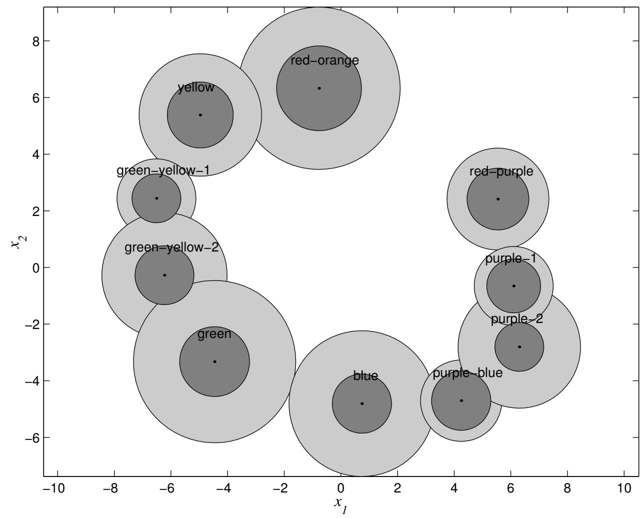

Figure 11: Reconstructed colors configuration using the possibility model in the case of color-deficient subjects (light: supports ; dark grey: level 1 α-cut).

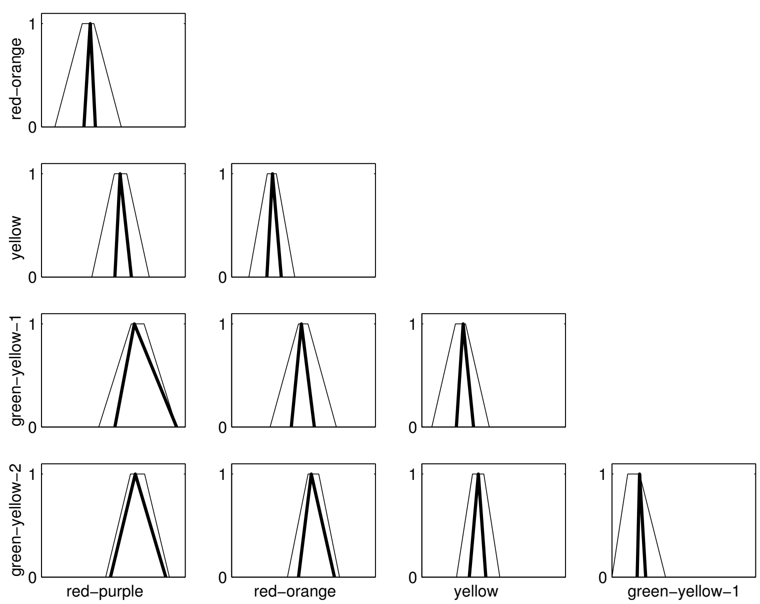

Figure 12: Membership functions of $\widetilde{\delta}_{ij}$ (bold lines) and $\widetilde{d}_{ij}$ for 10 pairs of colors seen by color-deficient subjects (possibility model).

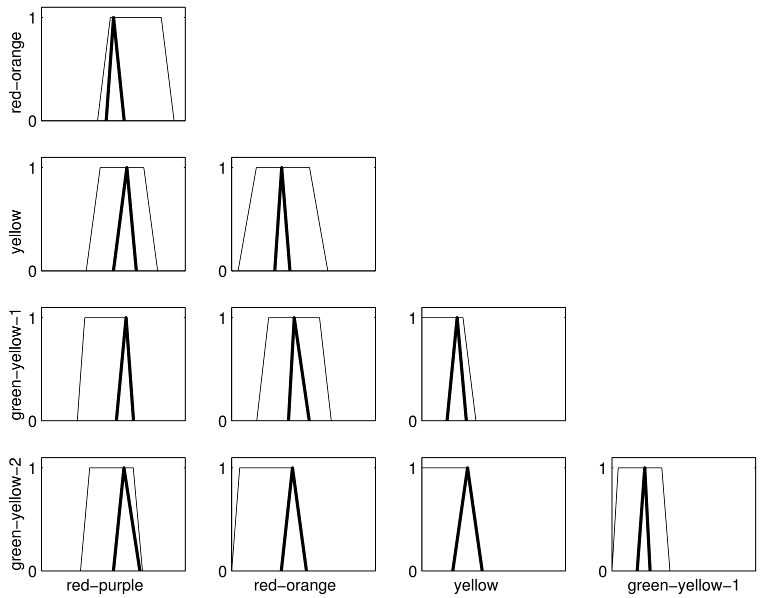

Figure 13: Membership functions of $\widetilde{\delta}_{ij}$ (bold lines) and $\widetilde{d}_{ij}$ for 10 pairs of colors seen by color-deficient subjects (possibility model).

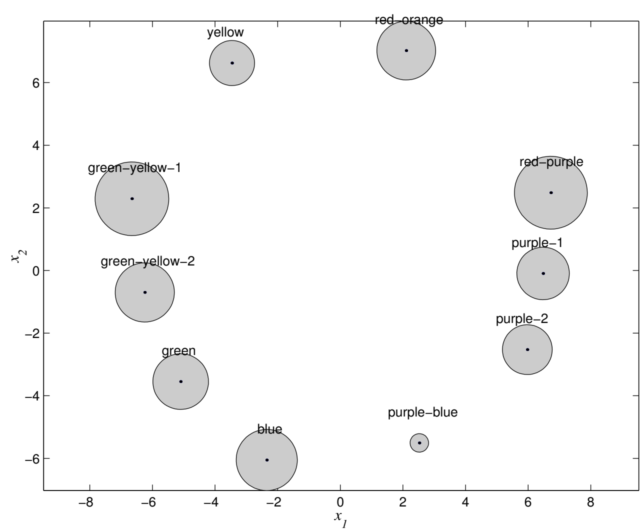

Figure 14: Reconstructed colors configuration using the least-squares model in the case of normal subjects (light grey: supports; dark grey: level 1 α-cut).

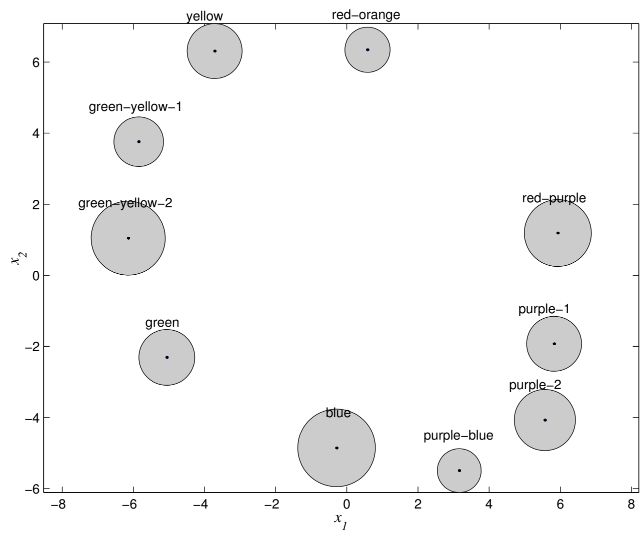

Figure 15: Reconstructed colors configuration using the least-squares model in the case of color-deficient subjects (light: supports ; dark grey: level 1 α-cut).

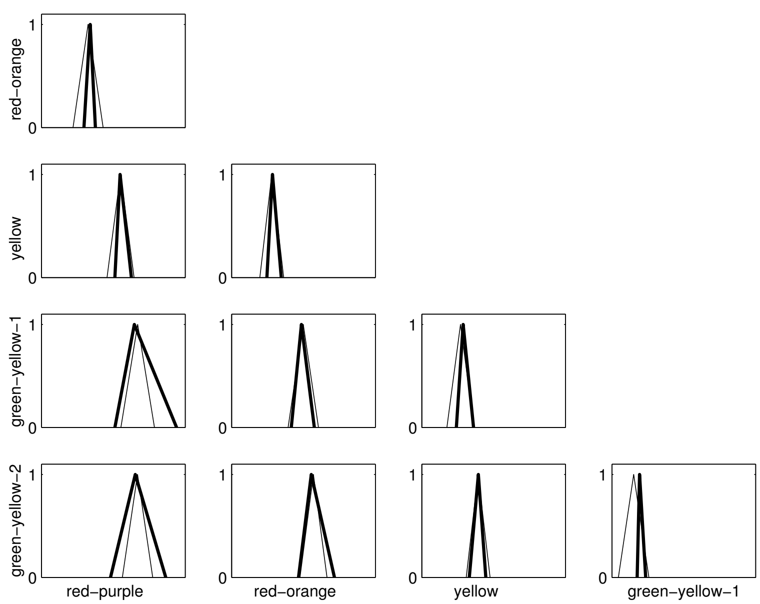

Figure 16: Membership functions of $\widetilde{\delta}_{ij}$ (bold lines) and $\widetilde{d}_{ij}$ for 10 pairs of colors seen by normal subjects (least-squares model).

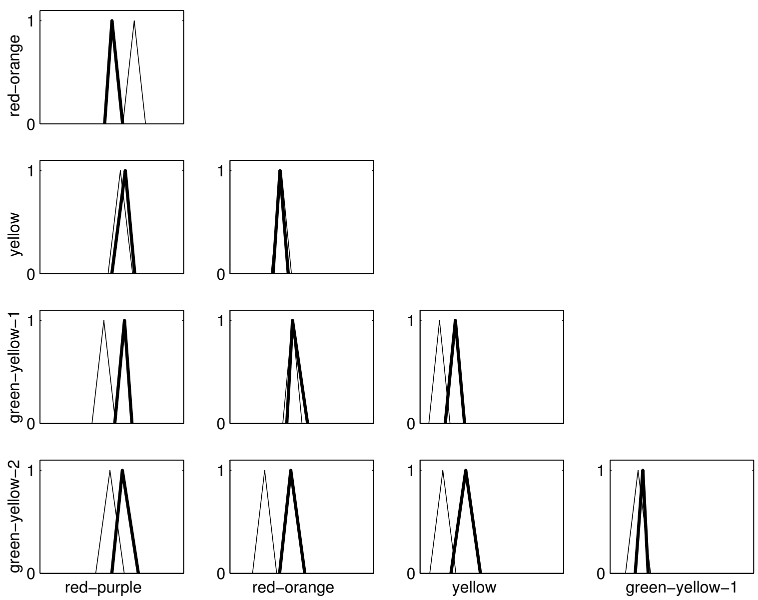

Figure 17: Membership functions of $\widetilde{\delta}_{ij}$ (bold lines) and $\widetilde{d}_{ij}$ for 10 pairs of colors seen by color-deficient subjects (least-squares model).