Dimensionality reduction and visualization of interval and fuzzy data: a survey

Denœux, Thierry
Heudiasyc, UMR CNRS 6599
Université de Technologie de Compiègne
BP 20 529, Compiègne (60205),France
E-mail: Thierry.Denoeux@hds.utc.fr

Masson, Marie-Hélène
Heudiasyc, UMR CNRS 6599
Université de Picardie Jules Verne
BP 20 529, Compiègne (60205),France
E-mail: mmasson@hds.utc.fr

1. Introduction

One of the main tasks in exploratory data analysis is the search for a relevant low-dimensional feature space in which the original data can be mapped and displayed so as to uncover their underlying structure. Dimensionality reduction is also commonly used as a preprocessing step prior to other operations such as clustering, regression, or classifier design. Classical feature extraction and data visualization methods include principal component analysis (PCA), and multidimensional scaling (MDS). Given a $n \times p$ data matrix $\mathbf{X}$ whose rows are p-dimensional feature vectors observed for n objects, PCA finds $q < p$ new features, defined as linear combinations of the original features, with maximal dispersion. In contrast, MDS usually starts with a $n \times n$ matrix $\Delta = (\delta_{ij})$ of dissimilarities between pairs of objects, and finds a configuration of points in a q dimensional space such that the interpoint distances reflect the input dissimilarities.

Classically, the data are assumed to consist in precise measurements of variables (features) for observation units of a population of interest, or precisely defined similarities between objects. In recent years, however, the necessity in many applications to take into account the imprecision of the observation process has been increasingly recognized. This has given rise to new branches of data analysis such as symbolic data analysis (Bock and Diday , 2000) and fuzzy data analysis (see, e.g. Masson and Denœux (2002)). One direction of research in data analysis thus aims at extending classical data analysis techniques to imprecise data. The goal of this paper is to provide a survey of recent work in this area, with emphasis on PCA and MDS analysis of interval and fuzzy data.

An imprecise observation of a quantitative variable X may be represented in several ways. The simplest formalism is that of real intervals: in many situations, the value x_i of X for object i is not known precisely, but one can assert that x_i is contained in an interval $[x_i^-, x_i^+]$. If this formalism is adopted for all variables X_j, $j = 1, \ldots, p$, one obtains an interval data matrix $[\mathbf{X}] = ([x_{ij}^-, x_{ij}^+])$ or an interval dissimilarity matrix $[\Delta] = ([\delta_{ij}^-, \delta_{ij}^+])$. The extension of PCA and MDS to such data has been addressed in (Cazes et al. , 1997; Lauro and Palumbo, 2000; D'Urso and Giordani , 2004) and (Denoeux and Masson, 2000), respectively. These methods account for imprecision in the input data by representing each object as a region (a hyperbox or a hypersphere) of the extracted feature space.

A more refined way of representing partial or imprecise knowledge of a variable value is through the formalism of fuzzy sets. A fuzzy subset $\widetilde{F}$ of a set ω is defined by a membership function $\mu_{\widetilde{F}} : \Omega \rightarrow [0, 1]$. For all $\omega \in \Omega$, $\mu_{\widetilde{F}}(\omega)$ is interpreted as the degree of membership of element ω to the fuzzy set $\widetilde{F}$. The fuzzy set $\widetilde{F}$ is said to be normal when $\sup_{\omega \in \Omega} \mu_{\widetilde{F}}(\omega) = 1$. For all $\alpha \in (0, 1]$, the crisp set ${}^{\alpha}\widetilde{F} = \{\omega \in \Omega, \mu_{\widetilde{F}}(\omega) \geq \alpha\}$ is called the α-cut of $\widetilde{F}$. A fuzzy number is defined as a normal fuzzy

subset $\widetilde{x}$ of $\mathbb{R}$ with compact support, and whose α-cuts are closed intervals (Dubois and Prade , 1988). It generalizes both concepts of real numbers and closed intervals, and may be viewed as an elastic constraint acting on a certain variable which is only known to lie "around" a certain value. Fuzzy sets were initially introduced as a means to model ambiguity of natural language, but they can actually capture a wider range of partial information.

When observations of feature or dissimilarity values are best modeled by fuzzy numbers, one gets a fuzzy data matrix $\widetilde{\mathbf{X}} = (\widetilde{x}_{ij})$ or a fuzzy dissimilarity matrix $\widetilde{\Delta} = (\widetilde{\delta}_{ij})$. Methods for the analysis of interval data may often be easily generalized to fuzzy data. The extension of PCA and MDS to fuzzy data has been addressed by Denœux and Masson (2004), Giordani and Kiers (2004a), D'Urso and Giordani (2005) and Masson and Denœux (2002), respectively. In these extensions, each object is represented as a fuzzy region, i.e., a fuzzy subset of $\mathbb{R}^q$.

The rest of this paper is devoted to survey of dimensionality reduction and visualization methods for interval and fuzzy data. Because of space constraints, the emphasis will be put on interval data, and fuzzy extensions will only be briefly mentioned. PCA will first be addressed in Section 2. MDS will then be covered in a similar fashion in Section 3, and Section 4 will conclude the paper.

2. Principal component analysis

Let $\mathbf{X} = (x_{ij})$ be the numerical data matrix of order $(n \times p)$, where n denotes the number of observation units, and p the number of variables. Without loss of generality, we shall assume $\mathbf{X}$ to be centered, i.e., $\frac{1}{n}\sum_{i=1}^{n} x_{ij} = 0$ for all $j \in \{1, \ldots, p\}$.

With this assumption, we can think of the n data points as a cloud in $\mathbb{R}^p$, with center of gravity located at the origin. PCA attempts to find a q-dimensional subspace $\mathcal{L}$ of $\mathbb{R}^p$, with $q \leq p$, such that the orthogonal projections of the n points on $\mathcal{L}$ have maximal variance. The solution is known to be the subspace spanned by the q normalized eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_q$ of the sample covariance matrix $\mathbf{S} = \frac{1}{n}\mathbf{X}'\mathbf{X}$, associated with the first q largest eigenvalues. The matrix $\mathbf{U}_q = (\mathbf{u}_1, \ldots, \mathbf{u}_q)$ of order $(p \times q)$ is sometimes called the *component loading matrix*. The coordinates of the objects in the projected space are defined by matrix $\mathbf{Y} = \mathbf{X}\mathbf{U}_q$, often referred to as the *component score matrix*. Based on $\mathbf{Y}$, the best reconstruction of $\mathbf{X}$ in the least squares sense is $\widehat{\mathbf{X}} = \mathbf{Y}\mathbf{U}_q' = \mathbf{X}\mathbf{U}_q\mathbf{U}_q'$.

2.1 Simple Extensions

Let us now assume that we have an interval data matrix $[\mathbf{X}] = ([x_{ij}^-, x_{ij}^+])$ of order $n \times p$. How can PCA be extended to handle such data?

Answers to this question were first provided by Cazes et al. (1997), who introduced two simple extensions of PCA, referred to as vertices PCA (V-PCA) and centers PCA (C-PCA). Both methods consist in applying PCA to a single valued data matrix derived from $[\mathbf{X}]$. In V-PCA, each line i of $[\mathbf{X}]$ corresponding to a hyperrectangle R_i in $\mathbb{R}^p$ is replaced by the matrix $\mathbf{X}_i$ of size $2^p \times p$, whose rows are the vertices of R_i. We thus get a single-valued data matrix $\mathbf{X}_V$ of size $n2^p \times p$ by concatenating the $\mathbf{X}_i$s. The application of standard PCA to this matrix yields a component score matrix $\mathbf{Y}_V$ of size $n2^p \times q$, in which each line corresponds to the projection of a vertex of one of the hyperrectangles R_i. The representation of object i is then defined as the smallest hyperrectangle in $\mathbb{R}^q$ containing all component scores for vertices associated to observation i. We thus obtain an interval-valued component score matrix $[\mathbf{Y}_V] = ([y_{ik}^-, y_{ik}^+])$ of size $n \times q$, with

$$y_{ik}^- = \min_{\ell \in n_i} y_{\ell k},$$

$$y_{ik}^+ = \max_{\ell \in n_i} y_{\ell k},$$

where n_i denotes the set of all vertices for observation unit i. Note that the size of data matrix $\mathbf{X}_V$ grows exponentially with the dimension of the input feature space. However, Cazes et al. (1997) have shown that matrix $[\mathbf{Y}_V]$ can be computed without having to compute the component scores for all vertices, and the complexity of the calculations is not higher than that involved in the application of PCA to a single-valued data matrix of size $n \times p$.

Similarly to V-PCA, C-PCA transforms the original interval-valued data matrix $[\mathbf{X}]$ into a single-valued data matrix, this time replacing each interval $[x_{ij}^-, x_{ij}^+]$ by its center $x_{ij}^c = (x_{ij}^- + x_{ij}^+)/2$. Standard PCA is then applied to data matrix $\mathbf{X}_C = (x_{ij}^c)$ of size $n \times p$, yielding a component score matrix $\mathbf{Y}_C$ of size $n \times q$. As for V-PCA, the component scores for the vertices can then be computed, and the smallest covering rectangles can be determined, yielding an interval-valued component score matrix $[\mathbf{Y}_C]$. This method obviously has an intuitive appeal because of its simplicity. However, we may remark that it does not take into account the imprecision of the data in the feature extraction process and, consequently, it does not exploit all the available information in dectecting the underlying structure of the data.

V-PCA and C-PCA were extended to three-way interval data by Giordani and Kiers (2004b).

2.2 Least-squares approaches

Other extensions of PCA to interval data were later proposed by D'Urso and Giordani (2004) and Denœux and Masson (2004) using two different least-squares approaches. D'Urso and Giordani (2004) take as a starting point the fact that, in standard PCA, $\mathbf{YU}_q'$ is a solution to the following minimization problem:

$$\min_{\mathbf{AB}'} \|\mathbf{X} - \mathbf{AB}'\|^2$$

where A is a matrix of size $n \times q$, and B a matrix of size $p \times q$. Given an interval data matrix $[\mathbf{X}]$, they propose to generalize the above optimization problem by searching an interval score matrix $[\mathbf{A}]$ and a single-valued component load matrix $\mathbf{B}$ so as to minimize a squared Euclidean distance between the original data matrix $[\mathbf{X}]$ and the reconstructed data matrix $[\widehat{\mathbf{X}}] = [\mathbf{A}]\mathbf{B}'$. This happens to be a constrained non linear optimization problem, for which they propose an iterative algorithm. This method, called MR-PCA for Midpoint-Radius PCA, is extended to triangular fuzzy numbers in (Giordani and Kiers , 2004a).

Denœux and Masson (2004) take a different route, starting with the observation that, in standard PCA, $\mathbf{U}_q$ is a solution of the following minimization problem:

$$\min_{\mathbf{B}} \|\mathbf{X} - \mathbf{XBB}'\|^2,$$

where $\mathbf{B}$ is a matrix of order $p \times q$, which can be seen as a weight matrix in an autoassociative neural network with one hidden layer of q linear neurons, and one output layer of p linear neurons. Iterative minimization of a mean squared error function yields a weight matrix $\mathbf{B}$, which is identical to $\mathbf{U}_q$ up to an isometric transformation. By feeding such a network with interval inputs and applying interval arithmetics, one obtains fuzzy outputs. The weight matrix $\mathbf{B}$ can then be determined to optimize a mean squared error function as in the precise case. The method, called NN-PCA for Neural network PCA, was originally introduced in the more general setting of fuzzy intervals. Note that NN-PCA seems to be computationally more efficient than MR-PCA, as the optimization is performed in a much smaller parameter space.

An experimental comparison of V-PCA, C-PCA and NN-PCA was reported by Giordani and Kiers (2006a); no method was found to outperform the other two in all cases. The authors give some clues helping to select a method according to the data structure.

2.3 Possibilistic approach

Yet another approach was proposed by D'Urso and Giordani (2005) in the case of fuzzy data, based on an approach initially proposed by Tanaka et al. (1982) in the context of fuzzy regression analysis. To simplify the presentation, the principle of this method will be explained in the case of interval data. As before, let us assume that we have an $n \times p$ interval data marix $[\mathbf{X}] = ([x_{ij}]$.
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. The possibilistic approach to PCA described in D'Urso and Giordani (2005) is based on a similar idea. Given an interval component score matrix $[\mathbf{A}]$ and a single-valued component load matrix $\mathbf{B}$, the reconstructed data matrix is $[\widehat{\mathbf{X}}] = [\mathbf{A}]\mathbf{B}'$. D'Urso and Giordani (2005) proposed to determine $[A]$ and $\mathbf{B}$ so that the reconstructed interval data *include* the original data, i.e., $[x_{ij}] \subseteq [\widehat{x}_{ij}]$ for all i and i, and the imprecision is minimized. Imprecision may be measured by $\sum_{i=1}^{n}\sum_{j=1}^{p}(\widehat{x}_{ij}^{+} - \widehat{x}_{ij}^{-})$, which is minimized subject to inclusion constraints and orthogonality constraint $\mathbf{B}\mathbf{B}' = \mathbf{I}$, where $\mathbf{I}$ is the indentity matrix. Note that this is a constrained non linear optimization problem, and the number of variables can grow very large as n and p increase.
This method was originally introduced in (D'Urso and Giordani , 2005) for the case where observations are symmetric fuzzy numbers. It was generalized in (Giordani and Kiers , 2006b) to a more general class of fuzzy nmbers, and to three-way data.

3. Multidimensional scaling

Multidimensional scaling (MDS) (see, e.g., Borg and Groenen (1997)) is a classical tool for analyzing dissimilarity measurements among a set of objects. The idea is to produce a map of the objects in which each object is represented by a point in a low-dimensional space in such a way that the distances between points reflect in some sense the input dissimilarities between the objects. The representation space is often chosen to be Euclidean. The computation of the configuration of points, starting from an initial random guess, is generally based on the iterative minimization of a *stress function* expressing the discrepancy between the input dissimilarities and the distances. One of the simplest form of stress function is:

$$(1) \quad \sigma = \sum_{i<j}(d_{ij} - \delta_{ij})^2,$$

where d_{ij} and δ_{ij} denotes, respectively, the Euclidean distance and the dissimilarity between object i and object j. The latter approach is referred to as *Euclidean Scaling*. If there is no need of "edge points" (especially when visualizing correlations between statistical variables), an alternative approach, proposed by Cox and Cox (1991), consists in constraining each object to lie on the surface of a sphere. In this model, referred to as *spherical MDS*, the distance between two objects is measured as the cosine of the angle between the vectors of unit length defined by the objects coordinates. As before, a suitable stress function is minimized through an iterative scheme to derive the coordinates of the points on the sphere.

3.1. Geometrical models for imprecise dissimilarities

First, we assume that the input dissimilarities are given as interval values $[\delta_{ij}^{-}, \delta_{ij}^{+}]$. The extension to fuzzy dissimilarity data will be discussed later. The dissimilarity between two objects being imprecise, it seems natural to seek for an imprecise location of the objects, whatever the representation space

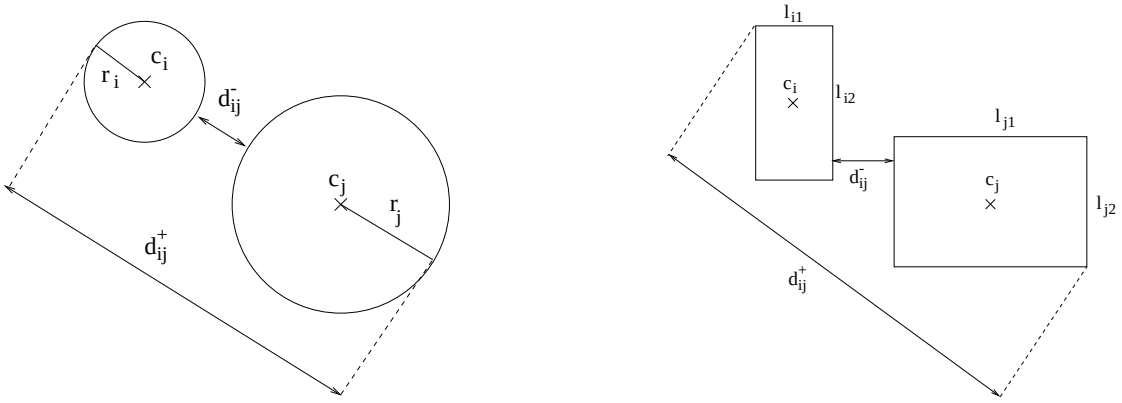

Figure 1. Minimum and maximum distances between two regions. (Left): hypersphere model; (Right): hyperbox model.

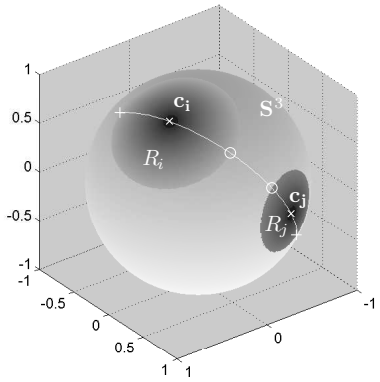

Figure 2. Spherical caps in the spherical space

(Euclidean or spherical) chosen. For that purpose, instead of associating precise coordinates to each object, one can choose to associate a region to it. Several region shapes with different parameterizations have been proposed. For instance, in a Euclidean space, Denoeux and Masson (2000) and Groenen et al. (2006) have proposed to represent each object as an hyperbox in a q-dimensional space. Each object i is thus represented by the center $\mathbf{c_i}$ of the box and q real values l_{ik} representing the size of the box along each each dimension $k = 1, q$ (see Figure 1). Alternatively, Denoeux and Masson (2000) have proposed to represent each object as a hypersphere, parameterized by a center $\mathbf{c_i}$ and a radius r_i (see Figure 1). Note that a similar model was proposed by Okada and Imaizumi (1987) to scale crisp asymmetric dissimilarities.

The distances in the hypersphere model are preserved under any isometric transformation of the configuration. This is not the case in the hyperbox model. One advantage of the hypersphere model is that the number of parameters is smaller than in the hyperbox model, so that the optimization is made simpler. The hyperbox model may be preferred, as underlined by Groenen et al. (2006), if the aim of the analysis is to discover interpretable relationships among objects in terms of underlying dimensions: as a result, the location of each object is expressed as a range of values on each dimension of the representation space.

For spherical MDS, Hébert et al. (2006) have proposed to represent the imprecise location of an object by a spherical cap defined by a center $\mathbf{c_i}$ representing the location of the object on the sphere, and an imprecision angle around this center (see Figure 2).

In each of these three representations, minimum and maximum distances between two regions R_i and

R_j are defined as:

$$(2) \quad d_{ij}^- = \min_{\mathbf{x}_i \in R_i, \mathbf{x}_j \in R_j} \|\mathbf{x}_i - \mathbf{x}_j\|$$

$$(3) \quad d_{ij}^+ = \max_{\mathbf{x}_i \in R_i, \mathbf{x}_j \in R_j} \|\mathbf{x}_i - \mathbf{x}_j\|.$$

Expressions of these distances as a function the parameters of the models are easily found. For instance, for the hypersphere model, as it can be seen from figure 1, d_{ij}^- and d_{ij}^+ are defined by the following equations:

$$(4) \quad d_{ij}^- = \max(0, d_{ij} - r_i - r_j)$$

$$(5) \quad d_{ij}^+ = d_{ij} + r_i + r_j,$$

where $d_{ij} = \|\mathbf{c}_i - \mathbf{c}_j\|$ denotes the Euclidean distance between the centers $\mathbf{c}_i$ and $\mathbf{c}_j$.

3.2. Fitting the models

Once the representation model is fixed, the problem is to determine the parameters (centers, radii, sizes of the box, or imprecision angles) in such a way that the minimum and maximum distances reflect, in some sense, the minimum and maximum dissimilarities. Two major approaches may be found in the literature. The first one is a straightforward generalization of conventional MDS. It consists in minimizing the following stress function:

$$(6) \quad \sigma' = \sum_{i<j} (d_{ij}^- - \delta_{ij}^-)^2 + \sum_{i<j} (d_{ij}^+ - \delta_{ij}^+)^2.$$

This function can be minimized using a simple gradient descent algorithm as initially proposed by Denœux and Masson (2000, 2002). However Groenen et al. (2006) have reported some convergence problems and developed for the hyperbox model a more efficient algorithm, based an iterative majorization. Simulations have shown that this algorithm, combined with multiple random starts, yields to good solutions. Note that they have also proposed a method (termed rational start) for initializing the parameters of the method, derived from the links between MDS and V-PCA (see Section 2.1). It is interesting to note that, if all the dissimilarities are precise (i.e. $\delta_{ij}^- = \delta_{ij}^+$), the model leads to null radii, thereby generalizing the classical model.

The minimization of (6) 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 by Masson and Denoeux (2002). It has been developed for the hypersphere representation in an Euclidean space and for spherical scaling. Its principle in the case of the hypersphere representation is detailed in the sequel. Let us suppose that the centers $\mathbf{c_i}$ of the hyperspheres have already been determined, for example by minimizing (6). One may attempt to find the smallest radii r_i such that the following condition is satisfied:

$$(7) \quad [\delta_{ij}^-, \delta_{ij}^+] \subseteq [d_{ij}^-, d_{ij}^+] \quad \forall i, j.$$

This principle can be formulated as the following optimization problem:

$$(8) \quad \min_{\mathbf{r}} \sum_{i=1}^{n} r_i$$

subject to:

$$(9) \quad d_{ij}^- \leq \delta_{ij}^- \quad \forall i, j$$

(10) $d_{ij}^{+} \geq \delta_{ij}^{+} \quad \forall i, j$

(11) $r_i \geq 0 \quad \forall i = 1, n$.

In (8), $\mathbf{r}$ denotes the vector of radii $(r_1, r_2, ..., r_n)^t$. The minimization of (8) over $\mathbf{r}$ under the constraints (9), (10) and (11) 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. Note that a similar approach has been reconsidered in the context of rough sets theory by Huang et al. (2006). 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). It is intended to provide exact, "worst-case", representations, in contrast to the compromise solutions derived from the least-squares method.

Least-square and possibilistic interval MDS were extended to fuzzy dissimilarities by Masson and Denœux (2002) for Euclidean scaling, and by Hébert et al. (2006) for spherical scaling. In that case, each object is represented by a fuzzy region $\widetilde{R}_i$ in $\mathbb{R}^q$ 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 $\widetilde{R}_j$ is defined as:

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

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 $\widetilde{R}_j$. This property can be exploited to optimize the fuzzy regions so that fuzzy distances between them approximate, or include, the input fuzzy dissimilarities.

4. Conclusion

Methods for dimensionality reduction and visualization of interval and fuzzy data have been reviewed. These methods extend PCA and MDS, two classical methods for analysing feature and dissimilarity data, in such a way that each object is no longer represented by a point, but by a crisp or a fuzzy region in a low-dimensional feature space. This makes it possible to account explicitly for the imprecision of input data, and may in some cases prevent the analyst from drawing wrong conclusions unsupported by the original data. Two fundamental mechanisms have been proposed for deriving such representation: one mechanism consists in finding optimal approximations, in the least squares sense, of the original data, whereas the other approach finds the most precise representations satisfying data consistency constraints. Extending more complex data analysis techniques to different types of imprecise data, using these or other principles, is believed to be a major challenge for the years to come.

References

Bock, H.-H., Diday E., editors., 2000. *Analysis of Symbolic Data. Exploratory methods for extracting statistical information from complex data.* Springer Verlag, Heidelberg.

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

Cazes, P., Chouakria, A., Diday, E., Schektman, Y., 1997. Extension de l'analyse en composantes principales à des données de type intervalle. *Revue de Statistique Appliquée*, 14(3):5–24, 1997.

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.

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

D'Urso, P., Giordani, P., 2004. A least squares approach to principal component analysis for interval valued data. *Chemometrics and Intelligent Laboratory Systems*, 70(2):179–192.

D'Urso, P., Giordani, P., 2005. A possibilistic approach to latent component analysis for symmetric fuzzy data. *Fuzzy Sets and Systems*, 150(2):285–305.

Giordani, P., Kiers, H.A.L., 2004. Principal component analysis of symmetric fuzzy data. *Computational Statistics and Data Analysis*, 45(3):519–548.

Giordani, P., Kiers, H.A.L., 2004. Three-way component analysis of interval-valued data. *Journal of Chemometrics*, 18(5):253–264.

Giordani, P., Kiers, H.A.L., 2006. A comparison of three methods for principal component analysis of fuzzy interval data. *Computational Statistics and Data Analysis*, 51(1):379–397.

Giordani, P., Kiers, H.A.L., 2006. Two- and three-way component models for LR fuzzy data in a possibilistic framework . *Fuzzy Sets and Systems*, 157(19):2648–2664.

Groenen, P.J.F., Winsberg, S., Rodrguez, O. Diday. E., 2006. I-Scal: Multidimensional scaling of interval dissimilarities. *Computational Statistics and Data Analysis*, 51(1), 2006, 360–378.

Hébert, P.-A. , Masson, M.-H., Denœux, T., 2006. Fuzzy multidimensional scaling. *Computational Statistics and Data Analysis*, 51(1), 2006, 335-359.

Huang, J.-J., Ong, C.-S., Tzeng, G-H., 2006. Interval multidimensional scaling for group decision using rough set concept. *Expert Systems with Applications*, 31(3), 525–530.

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

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

Lauro, C. N., Palumbo, F., 2000. Principal component analysis of interval data: a symbolic data analysis approach. *Computational Statistics*, 15(1), 73–87.

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

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

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.

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