Multidimensional scaling of interval-valued dissimilarity data

T. Denœux and M. Masson

Université de Technologie de Compiègne

U.M.R CNRS 6599 Heudiasyc

BP 20529 - F-60205 Compiègne cedex - France

email: Thierry.Denoeux@hds.utc.fr

August 31, 1999

Abstract

Multidimensional scaling is a well-known technique for representing measurements of dissimilarity among objects as points in a p-dimensional space. In this paper, this method is extended to the case where dissimilarities are only known to lie within certain intervals. Each object is then no longer represented as *point*, but as a *region* of $\mathbb{R}^p$, in such a way that the minimum and maximum distances between two regions approximate the lower and upper bounds of the dissimilarity interval between the two objects. Experiments with real data demonstrate the ability of this method to represent both the structure and the precision of dissimilarity measurements.

Keywords: Multidimensional scaling, Interval-valued data, Exploratory data analysis, Data visualization.

1 Introduction

Multidimensional scaling (MDS) is a classical technique for representing measurements of dissimilarity among objects as distances between points of a low-dimensional space [3, 1]. Typically, a $n \times n$ symmetric matrix $\Delta = (\delta_{ij})$ of dissimilarities is given, and one seeks a configuration of n points $\{\mathbf{x}_i, i = 1, \ldots, n\}$ in $\mathbb{R}^p$, such that the $n \times n$ matrix $D = (d_{ij})$ of Euclidean distances between the n points approximates Δ in some specified way. This is usually achieved by minimizing a loss function measuring the discrepancies between the observed and reconstructed dissimilarities. Multidimensional scaling techniques are generally classified as *metric* and *non metric*. In metric MDS, one seeks to match dissimilarities by distances through iterative minimization of a stress function such as:

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

where X is the $n \times p$ data matrix, whose rows contain the coordinates of the data points. In contrast, non metric MDS only aims at preserving the order of dissimilarities, i.e., one seeks to have $d_{ij} \leq d_{i'j'}$ whenever $\delta_{ij} \leq \delta_{i'j'}$. This is achieved by minimizing the discrepancies between distances d_{ij} and transformed dissimilarities $f(\delta_{ij})$, where f is an arbitrary increasing function [7]. Typically, the sum of squares of the differences is used, and f is found by isotonic regression. The following normalized stress function is then minimized over both the configuration X and f:

$$\sigma_1(X, f) = \frac{\sum_{i<j} (d_{ij} - f(\delta_{ij}))^2}{\sum_{i<j} d_{ij}^2}. \tag{2}$$

In pattern recognition, MDS is generally used as a data exploration technique for, e.g., detecting clusters in a data set [10]. The matrix Δ is then often obtained by computing distances among a set of q-dimensional feature vectors $\mathbf{y}_i, i = 1, \ldots, n$, and MDS is then essentially used as a dimensionality reduction and data visualization technique. Another application of MDS is found in psychological experiments or sensory evaluation studies, in which the goal is to discover underlying dimensions (forming a "subjective space") that would explain dissimilarity judgments provided by humans.

In this paper, it is proposed to extend MDS to the case where dissimilarities δ_{ij} are only known to lie within certain intervals $[\delta_{ij}] = [\delta_{ij}^-, \delta_{ij}^+]$. Such data may arise in a variety of situations, including the following ones:

1. The initial data may consist in interval-valued feature vectors $[\mathbf{y}_i]$, whose components are defined as intervals $[y_{i,k}] = [y_{i,k}^-, y_{i,k}^+]$. Such a data type may be used to describe sets of entities (e.g., the weights of pupils in a classroom), the range of a variable observed during a certain period (e.g., minimum and maximum daily temperatures), or to account for measurement uncertainty. Each dissimilarity interval $[\delta_{ij}]$ may then be defined, e.g., as the range of $\|\mathbf{y}_i - \mathbf{y}_j\|$, for all $\mathbf{y}_i \in [\mathbf{y}_i]$ and $\mathbf{y}_j \in [\mathbf{y}_j]$. The extension of data analysis techniques such as clustering and principal component analysis to such data has been studied by several authors [5, 6, 9, 2].

2. In some applications, subjective assessments of dissimilarities are directly elicited from human subjects, who may have difficulties in quantifying the proximity of certain pairs of objects by a precise value, even measured on an ordinal scale. It may then be preferable to account for the indeterminacy of dissimilarity judgments by explicitly allowing some imprecision in the respondents' answers.

3. Finally, a third situation is that in which dissimilarity judgments are collected from N individuals or devices. In that case, the disparity between two objects i and j is characterized by an empirical distribution of N values $\delta_{ij,1}, \ldots, \delta_{ij,N}$, where $\delta_{ij,k}$ denotes the dissimilarity given by judge k. Classical methods for analyzing such data are based either on the introduction of a "subject space", such as in the INDSCAL model or variants thereof, or on a comparison between individual configurations using Procrustes analysis [3]. Although these approaches may on some occasions provide interesting insights into the structure of the data, they usually lead to complex results requiring careful interpretation and a good deal of expertise. An alternative approach could be to globally describe the distribution of N dissimilarity values for each object pair by an interval, such as the interquartile range or the mean plus or minus one standard deviation, and proceed with the analysis of the resulting $n \times n$ matrix of interval-valued dissimilarities.

In the sequel, a general approach to the analysis of such data is presented. The main idea behind this method is to represent each object no longer as a *point* but as a *region* of a p-dimensional space, in such a way that the minimum and maximum distances between two regions approximate the lower and upper bounds of the dissimilarity interval between the two objects. Several variants of this method will be presented, differing by the shape of the regions (hyperspheres or hyperboxes), and by the type of scaling (metric or non metric). Finally, the efficiency of the proposed scheme will be demonstrated using typical data sets.

2 The method

2.1 Principle approach

Let us 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 the two regions:

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

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

We want the regions R_i, $i = 1, \ldots, n$ to be defined in such a way that the discrepancies between dissimilarity intervals $[\delta_{ij}^-, \delta_{ij}^+]$ and distance intervals $[d_{ij}^-, d_{ij}^+]$ are minimized. In a metric approach, this may be achieved, for instance, by generalizing the stress function of Eq. (1) as:

$$\sigma'(\mathcal{R}) = \sum_{i<j} (d_{ij}^- - \delta_{ij}^-)^2 + \sum_{i<j} (d_{ij}^+ - \delta_{ij}^+)^2, \quad (5)$$

where $\mathcal{R}$ denotes the set of n regions $\{R_1, \ldots, R_n\}$, and minimizing $\sigma'(\mathcal{R})$ with respect to $\mathcal{R}$ by gradient descent.

Alternatively, one may attempt to preserve only the rank order of lower and upper dissimilarities by generalizing the normalized stress of Eq. (2) as:

$$\sigma_1'(\mathcal{R}, f) = \frac{\sum_{i<j} (d_{ij}^- - f(\delta_{ij}^-))^2 + \sum_{i<j} (d_{ij}^+ - f(\delta_{ij}^+))^2}{\sum_{i<j} (d_{ij}^-)^2 + \sum_{i<j} (d_{ij}^+)^2}, \quad (6)$$

and minimizing $\sigma'_1(\mathcal{R}, f)$ with respect to both $\mathcal{R}$ and f. As in standard non metric MDS, this may be achieved by alternating gradient descent minimization over the configuration, and isotonic regression for the determination of optimal f.

In practice, a parameterized shape has to be chosen for regions R_i. In the following section, two models are successively presented, corresponding to the representation of objects as hyperspheres and as hyperboxes, respectively. Expressions for the derivatives of $\sigma'(\mathcal{R})$ and $\sigma'_1(\mathcal{R}, f)$ with respect to the parameters of each model are also given.

2.2 Hypersphere model

A very simple representation is obtained by defining each region R_i as a hypersphere (a circle when $p = 2$) with center $\mathbf{c}_i$ and radius r_i (Fig. 1). We then obtain a model with $n(p+1)$ parameters (n centers defined by p coordinates each, and n radii).

It is trivial to remark that, in that case, d^-_{ij} and d^+_{ij} defined by Eqs (3) and (4) are equal, respectively, to:

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

$$d^+_{ij} = d_{ij} + r_i + r_j, \tag{8}$$

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

The iterative minimization of stress functions $\sigma'(\mathcal{R})$ and $\sigma'_1(\mathcal{R}, f)$ defined, respectively, by Eqs (5) and (6), requires the computation of their derivatives with respect to each parameter, which may be conducted as follows. Introducing new variables ρ_i, $i = 1, \ldots, n$ such that $r_i = \rho_i^2$, to ensure the condition $r_i \geq 0$, and denoting as c_{ik} the

k-th coordinate of vector $\mathbf{c}_i$, we have

$$\frac{\partial \sigma'(\mathcal{R})}{\partial c_{ik}} = \sum_{j \neq i} \frac{\partial \sigma'(\mathcal{R})}{\partial d_{ij}^-} \frac{\partial d_{ij}^-}{\partial c_{ik}} + \sum_{j \neq i} \frac{\partial \sigma'(\mathcal{R})}{\partial d_{ij}^+} \frac{\partial d_{ij}^+}{\partial c_{ik}} \quad k = 1, p \tag{9}$$

$$\frac{\partial \sigma'(\mathcal{R})}{\partial \rho_i} = \sum_{j \neq i} \frac{\partial \sigma'(\mathcal{R})}{\partial d_{ij}^-} \frac{\partial d_{ij}^-}{\partial \rho_i} + \sum_{j \neq i} \frac{\partial \sigma'(\mathcal{R})}{\partial d_{ij}^+} \frac{\partial d_{ij}^+}{\partial \rho_i}, \tag{10}$$

with

$$\frac{\partial \sigma'(\mathcal{R})}{\partial d_{ij}^-} = 2(d_{ij}^- - \delta_{ij}^-) \tag{11}$$

$$\frac{\partial \sigma'(\mathcal{R})}{\partial d_{ij}^+} = 2(d_{ij}^+ - \delta_{ij}^+) \tag{12}$$

and

$$\frac{\partial d_{ij}^-}{\partial c_{ik}} = \frac{c_{ik} - c_{jk}}{d_{ij}} 1_{[0,+\infty)}(d_{ij} - \rho_i^2 - \rho_j^2) \tag{13}$$

$$\frac{\partial d_{ij}^+}{\partial c_{ik}} = \frac{c_{ik} - c_{jk}}{d_{ij}} \tag{14}$$

$$\frac{\partial d_{ij}^-}{\partial \rho_i} = -2\rho_i \, 1_{[0,+\infty)}(d_{ij} - \rho_i^2 - \rho_j^2) \tag{15}$$

$$\frac{\partial d_{ij}^+}{\partial \rho_i} = 2\rho_i. \tag{16}$$

The gradient of $\sigma_1'(\mathcal{R}, f)$ may be computed in a similar fashion. Noting

$$\sigma_1'(\mathcal{R}, f) = \frac{A}{B},$$

Eqs (11) and (12) are now replaced by:

$$\frac{\partial \sigma_1'(\mathcal{R}, f)}{\partial d_{ij}^-} = \frac{2(d_{ij}^- - f(\delta_{ij}^-))B + 2d_{ij}^- A}{B^2} \tag{17}$$

$$\frac{\partial \sigma_1'(\mathcal{R}, f)}{\partial d_{ij}^+} = \frac{2(d_{ij}^+ - f(\delta_{ij}^+))B + 2d_{ij}^+ A}{B^2}. \tag{18}$$

Remark: The function $x \mapsto \max(0, x)$ is differentiable everywhere except at $x = 0$. This fact need not be considered as far as the minimization of $\sigma'(\mathcal{R})$ or $\sigma_1'(\mathcal{R}, f)$ is concerned.

2.3 Hyperbox model

As an alternative to the above model, it is possible to define each region R_i representing object i in the p-dimensional space as a hyperbox (Fig. 2) with center $\mathbf{c}_i$ and edge lengths ℓ_{ik}, $k = 1, \ldots, p$:

$$R_i = \{\mathbf{x} \in \mathbb{R}^p \mid c_{ik} - \frac{\ell_{ik}}{2} \leq x_k \leq c_{ik} + \frac{\ell_{ik}}{2}, k = 1, p\}.$$

This defines a more flexible model than the previous one, with $2np$ parameters. Additionally, such regions correspond to *symbolic objects* according to the terminology introduced by Diday et al. [5, 2].

The expressions for the minimum and maximum distances between points located in each of two hyperboxes are given by the following proposition.

PROPOSITION 1

Let R_i and R_j be two hyperboxes with parameters $(\mathbf{c}_i, \ell_{i1} \ldots, \ell_{ip})$ and $(\mathbf{c}_j, \ell_{j1} \ldots, \ell_{jp})$. Let d_{ij}^- and d_{ij}^+ be the minimum and maximum distances between any $\mathbf{x}_i \in R_i$ and any $\mathbf{x}_j \in R_j$. Then:

$$d_{ij}^- = \frac{1}{4}\sqrt{\sum_{k=1}^{p} (\ell_{ik} + \ell_{jk} - 2|c_{ik} - c_{jk}| - |\ell_{ik} + \ell_{jk} - 2|c_{ik} - c_{jk}|\,|)^2} \qquad (19)$$

$$d_{ij}^+ = \frac{1}{2}\sqrt{\sum_{k=1}^{p} (\ell_{ik} + \ell_{jk} + 2|c_{ik} - c_{jk}|)^2}, \qquad (20)$$

where $|\cdot|$ denotes the absolute value.

Proof: We seek the minimum and maximum of

$$d(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{\sum_{k=1}^{p} |x_{ik} - x_{jk}|^2}$$

under the constraints

$$c_{ik} - \frac{\ell_i}{2} \leq x_{ik} \leq c_{ik} + \frac{\ell_i}{2} \qquad k = 1, p$$

$$c_{jk} - \frac{\ell_j}{2} \le x_{jk} \le c_{jk} + \frac{\ell_j}{2} \qquad k = 1, p.$$

Let us denote, for $h = i, j$ and $k \in \{1, \dots, p\}$:

$$x_{hk}^- = c_{hk} - \frac{\ell_h}{2} \qquad x_{hk}^+ = c_{hk} + \frac{\ell_h}{2}.$$

The range of $x_{ik} - x_{jk}$ under the above constraints is the interval

$$[x_{ik}^- - x_{jk}^+, x_{ik}^+ - x_{jk}^-].$$

Using the following property:

$$\begin{aligned}
\min_{a \le x \le b} |x| &= \max(0, a, -b) \\
\max_{a \le x \le b} |x| &= \max(b, -a),
\end{aligned}$$

the range of $|x_{ik} - x_{jk}|$ is found to be:

$$[\max(0, x_{ik}^- - x_{jk}^+, x_{jk}^- - x_{ik}^+), \max(x_{ik}^+ - x_{jk}^-, x_{jk}^+ - x_{ik}^-)].$$

Applying the formula:

$$\max(a, b) = \frac{1}{2}(a + b + |a - b|) \quad \forall a, b \in \mathbb{R}$$

to each bound, squaring, summing over k and taking the square root yields the desired result. $\square$

Expressions of the gradients of $\sigma'(\mathcal{R})$ and $\sigma_1'(\mathcal{R}, f)$ with respect to the model parameters are similar to those obtained in the previous model, except for the derivatives of d_{ij}^- and d_{ij}^+. To ensure the condition $\ell_{ik} \ge 0$ for all i and k, let us introduce new variables λ_{ik} such that $\ell_{ik} = \lambda_{ik}^2$. Let

$$\begin{aligned}
a_{ijk}^- &= \lambda_{ik}^2 + \lambda_{jk}^2 - 2|c_{ik} - c_{jk}| \\
a_{ijk}^+ &= \lambda_{ik}^2 + \lambda_{jk}^2 + 2|c_{ik} - c_{jk}|.
\end{aligned}$$

We obtain, from Eqs (19) and (20):

$$
\begin{aligned}
\frac{\partial d_{ij}^-}{\partial \lambda_{ik}} &= \frac{\lambda_{ik}(a_{ijk}^- - |a_{ijk}^-|)(1-\text{sgn}(a_{ijk}^-))}{2d_{ij}^-} && (21)\\
\frac{\partial d_{ij}^-}{\partial c_{ik}} &= \frac{\text{sgn}(c_{jk}-c_{ik})(a_{ijk}^- - |a_{ijk}^-|)(1-\text{sgn}(a_{ijk}^-))}{2d_{ij}^-} && (22)\\
\frac{\partial d_{ij}^+}{\partial \lambda_{ik}} &= \frac{a_{ijk}^+ \lambda_{ik}}{d_{ij}^+} && (23)\\
\frac{\partial d_{ij}^+}{\partial c_{ik}} &= \frac{a_{ijk}^+ \,\text{sgn}(c_{jk}-c_{ik})}{d_{ij}^+} && (24)
\end{aligned}
$$

Remarks:

1. In the case of a unidimensional feature space ($p = 1$), the hypersphere and hyperbox models reduce to a single one, which can be checked by verifying the equivalence between Eqs (7-8) and (19-20), with $\ell_i = 2r_i$.

2. If we start with an initial configuration $\{[\mathbf{y}_i], i = 1, \ldots, n\}$ of n two-dimensional interval-valued feature vectors, compute an interval-valued distance matrix Δ, and apply our interval metric MDS method with the hyperbox model and $p = 2$, it may be wondered whether we obtain the initial configuration as a solution, up to a translation. It is clear that the initial true configuration has zero stress, so it is actually a (not necessarily unique) global minimum of the stress function. However, this global minimum is not guaranteed to be attained by any gradient-based algorithm. The same remark obviously applies in the case of standard MDS of precise dissimilarity data (even if a perfect configuration exists, there is no guarantee to find it using a local optimization procedure).

3 Experiments

3.1 Oil data

Ichino's *Fats and oil* data set presented in Table 1 has been used by several authors [6, 2] as a typical example of a data set involving interval-valued features. It is composed of 8 objects described by 4 features taking interval values (a fifth nominal feature in the original data has been discarded from the present analysis). The objective of the analysis is to find a meaningful feature space with reduced dimension, allowing to reveal the underlying structure of the data.

In order to apply our approach, an interval-valued distance matrix Δ was first computed using Eqs (19) and (20) (Table 2). Results obtained with the metric hypersphere and hyperbox models are shown in Figs (3-4) and (5-6), respectively. In each case, an initial configuration was obtained by applying the classical scaling method [1] to the matrix of mean distances. Figs (4) and (6) show the reasonably good quality of lower and upper distance approximations in a two-dimensional feature space ($p = 2$). The stress value obtained with the hyperbox model ($\sigma' = 1.86$) is less than that obtained with the hypersphere model ($\sigma' = 6.43$), as expected given the respective numbers of free parameters in both models. The clustering structure revealed by Figs (3) and (5) is coherent with that found by Cazes *et al.* [2] using their extension of principal component analysis to interval-valued data.

3.2 Vowel data

The application of our technique to the analysis of dissimilarities between classes, defined as sets of objects, may be illustrated using Deterding's speech data [4, 8].

This data set has often been used as a benchmark for classification techniques. It consists of 528 vectors of 10 features extracted from speech signals. There are 11 classes corresponding to eleven vowels as indicated Table 3. Eight speakers were used in this experiment. The objective of our analysis is to map the 11 classes onto a 2-dimensional space so as to represent both the proximities between classes, and the spread of these classes.

For that purpose, interval-valued dissimilarities between each pair of classes were computed as follows. For each pair (i, j) of classes, the distance between each vector of class i and each vector of class j was computed. If n_i and n_j denote the number of examples in class i and j, respectively, we then obtained a distribution of $n_i n_j$ distances. This lower and upper distances between classes i and j were then defined as, respectively, the first and ninth deciles of this distribution. The result of this computation is shown in Table 4.

Two-dimensional non-metric configurations with the corresponding Shepard diagrams are shown in Figs 7 and 8 for the hypersphere model, and in Figs 9 and 10 for the hyperbox model. As expected, both configurations respect the phonetic similarities between vowels, while revealing the variability of speech signals in each class. As in the previous example, the hyperbox model yielded a smaller value of normalized stress ($\sigma_1' = 0.004$ against 0.011 for the hypersphere model). However, the hypersphere model may be prefered, at least in this example, for the better readability of the corresponding configuration (although it does not lead to symbolic objects, as opposed to the hypersphere model).

4 Conclusion

The problem of visualizing interval-valued dissimilarity data has been addressed. In the proposed method, each object is represented as a region in a p-dimensional space, in such a way that the minimum and maximum distances between two regions approximate the lower and upper dissimilarities between the two corresponding objects. Different models are obtained by varying the shape of the regions (e.g., hyperspheres or hyperboxes), and the stress function (leading, e.g., to metric or non metric scaling). Experiments with real data have demonstrated the ability of this method to produce configurations which attempt to preserve both the structure of dissimilarities and the imprecision of the available data. Potential applications of this method include the visualization of sets of objects described by multiple interval-valued attributes, and the analysis of subjective similarity assessments, such as encountered in sensory evaluation studies.

References

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

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

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

[4] D. H. Deterding. *Speaker Normalisation for Automatic Speech Recognition*. PhD thesis, University of Cambridge, 1989.

[5] K. Chidananda Gowda and E. Diday. Symbolic clustering using a new similarity measure. *IEEE Transactions on Systems, Man and Cybernetics*, 22(2):368–378, 1992.

[6] M. Ichino and H. Yaguchi. Generalized Minkowski metrics for mixed type data analysis. *IEEE Transactions on Systems, Man and Cybernetics*, 24(4):698–708, 1994.

[7] J. B. Kruskal. Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, 29(2):115–129, 1964.

[8] P. M. Murphy and D. W. Aha. *UCI Repository of machine learning databases [Machine-readable data repository]*. University of California, Department of Information and Computer Science., Irvine, CA, 1994.

[9] P. Nagabhushan, K. Chidananda Gowda, and E. Diday. Dimensionality reduction of symbolic data. *Pattern Recognition Letters*, 16:219–223, 1995.

[10] B. D. Ripley. *Pattern Recognition and Neural networks*. Cambridge University Press, Cambridge, 1996.

Table 1: Fats and oils data.

Object	Specific gravity	Freezing point	Iodine value	Saponification value
1. linseed oil	[0.930,0.935]	[-27,-18]	[170,204]	[118,196]
2. perilla oil	[0.930,0.937]	[-5,-4]	[192,208]	[188,197]
3. cottonseed oil	[0.916,0.918]	[-6,-1]	[99,113]	[189,198]
4. sesame oil	[0.920,0.926]	[-6,-4]	[104,116]	[187,193]
5. camellia oil	[0.916,0.917]	[-21,-15]	[80,82]	[189,193]
6. olive oil	[0.914,0.919]	[0,6]	[79,90]	[187,196]
7. beef tallow	[0.860,0.870]	[30,38]	[40,48]	[190,199]
8. hog fat	[0.858,0.864]	[22,32]	[53,77]	[190,202]

Table 2: Fats and oils data: interval-valued distances from Table 1.

	1	2	3	4	5	6	7
2	[0.7,6.3]						
3	[1.3,6.7]	[1.5,2.2]					
4	[1.1,6.3]	[1.4,2.1]	[0.1,1.0]				
5	[1.6,6.3]	[2.1,2.6]	[0.5,1.4]	[0.6,1.2]			
6	[1.7,6.7]	[1.9,2.6]	[0.2,1.2]	[0.3,1.2]	[0.8,1.5]		
7	[3.9,8.1]	[3.7,4.7]	[2.4,3.4]	[2.6,3.6]	[2.8,3.7]	[2.0,3.1]	
8	[3.5,8.2]	[3.4,4.5]	[2.2,3.2]	[2.4,3.5]	[2.6,3.6]	[1.9,3.0]	[0.1,1.5]

Table 3: Vowel data: words used in recording the vowels

	vowel	word		vowel	word
1	i	heed	7	O	hod
2	I	hid	8	C:	hoard
3	E	head	9	U	hood
4	A	had	10	u:	who'd
5	a:	hard	11	3:	heard
6	Y	hud			

Table 4: Vowel data: interval-valued distance matrix.

	1	2	3	4	5	6	7	8	9	10
2	[1.8,5.1]									
3	[2.2,5.0]	[1.4,4.4]								
4	[3.1,4.6]	[2.4,4.0]	[1.7,3.3]							
5	[3.6,4.8]	[3.1,4.2]	[2.6,3.7]	[1.8,2.9]						
6	[3.5,4.5]	[2.7,3.9]	[2.3,3.3]	[1.5,2.7]	[1.4,2.8]					
7	[3.4,5.1]	[3.3,4.4]	[2.8,3.9]	[2.4,3.4]	[1.6,3.7]	[2.0,3.4]				
8	[3.1,5.9]	[3.6,5.1]	[3.3,4.7]	[3.1,4.1]	[2.3,4.1]	[2.7,3.8]	[1.9,4.0]			
9	[3.3,4.8]	[2.9,4.4]	[2.6,4.1]	[2.3,3.6]	[2.1,3.2]	[2.1,3.2]	[1.8,3.4]	[1.9,4.2]		
10	[3.0,5.0]	[2.9,4.6]	[2.7,4.4]	[2.6,3.9]	[2.6,3.6]	[2.4,3.6]	[2.5,3.7]	[2.2,4.1]	[1.7,3.8]	
11	[3.0,4.2]	[2.4,3.8]	[2.1,3.3]	[1.6,2.9]	[1.8,2.9]	[1.4,2.9]	[2.2,3.4]	[3.0,4.4]	[1.7,4.2]	[2.4,4.9]

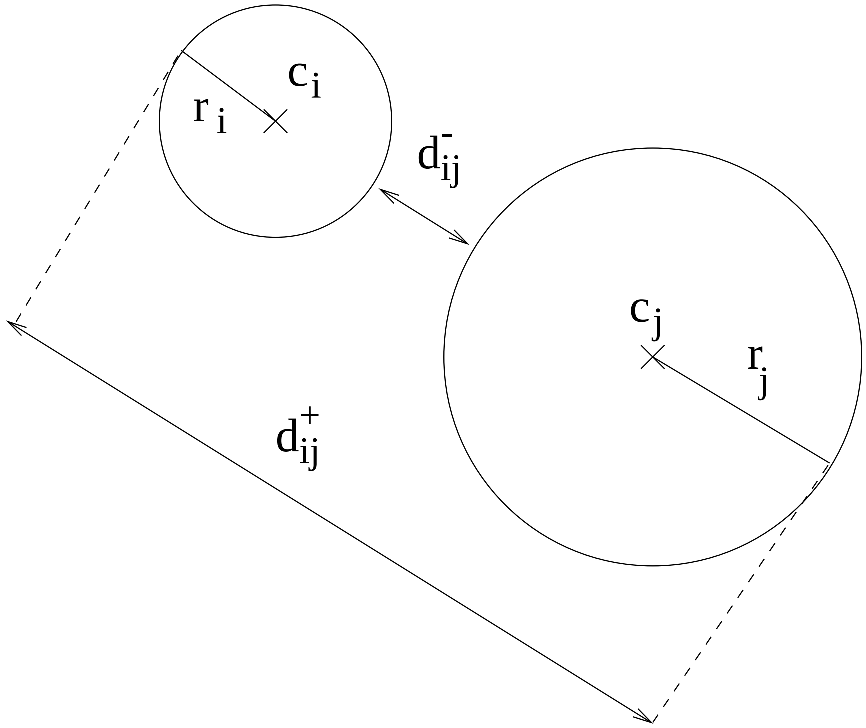

Figure 1: Minimum and maximum distances between two spheric regions.

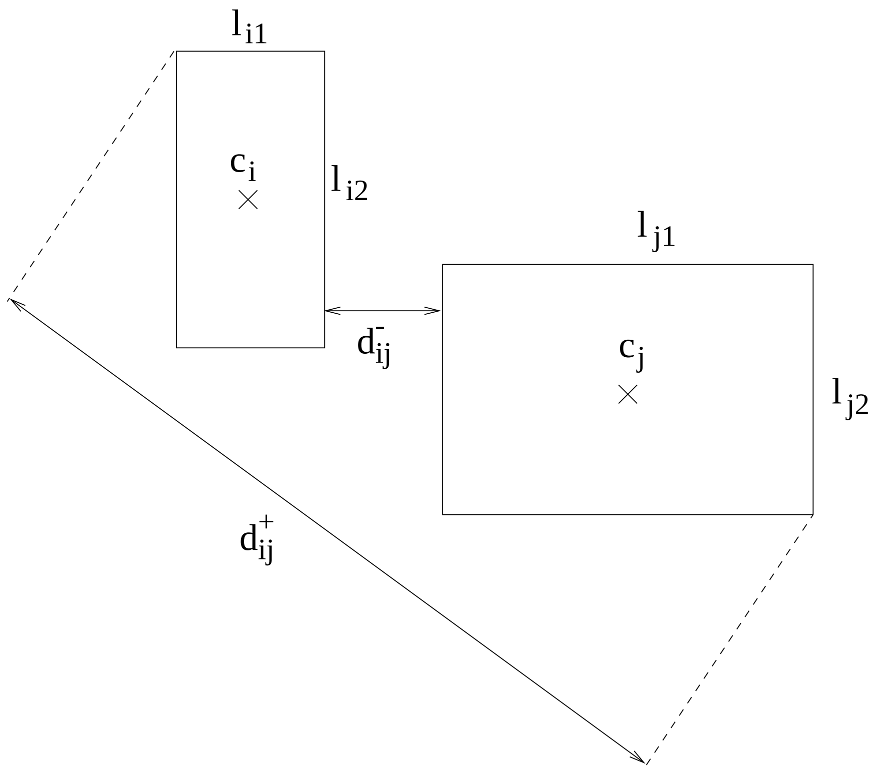

Figure 2: Minimum and maximum distances between two boxes.

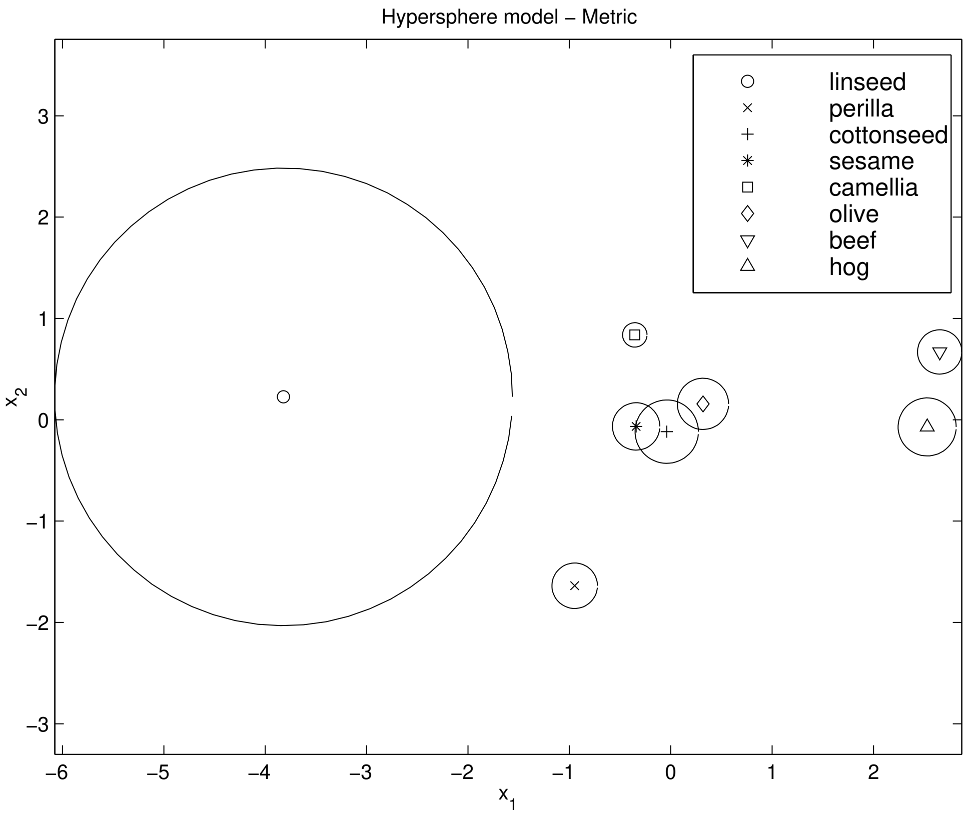

Figure 3: Two-dimensional metric configuration for the Fats and Oil data (hypersphere model). Each circle represents one of the 8 objects. The lower and upper Euclidean distances between any two circles approximate the distances between the corresponding objects.

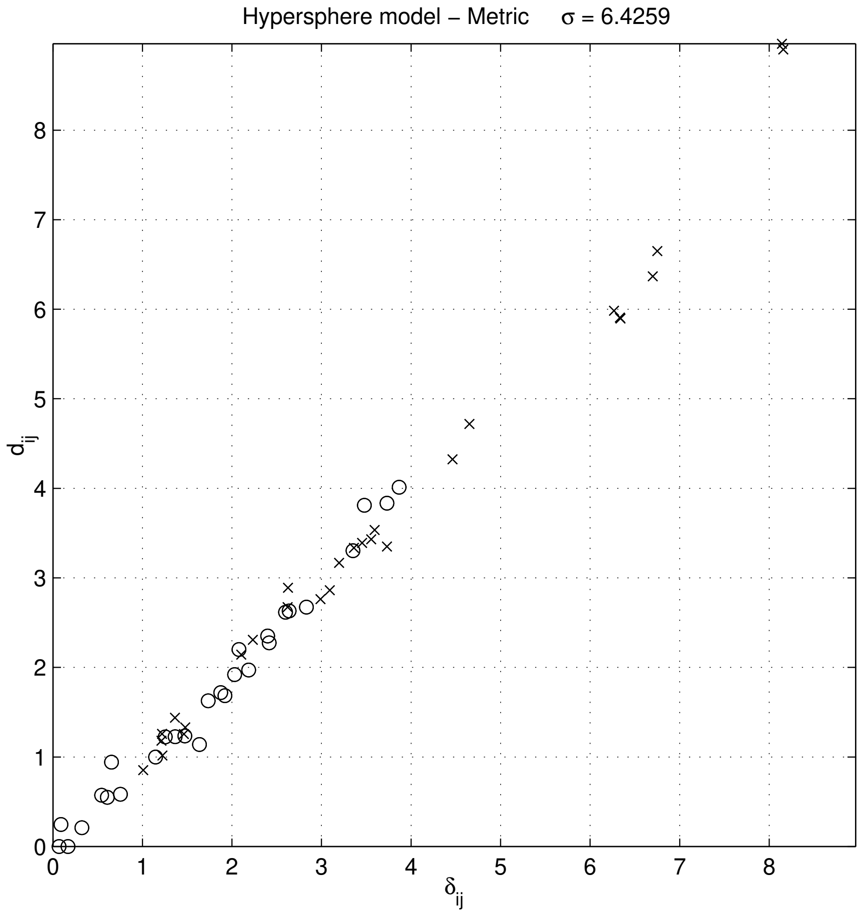

Figure 4: Real distances (x-axis) vs. reconstructed distances (y-axis) for the Fats and Oil data using the hypersphere model. (o: δ_{ij}^- vs. d_{ij}^-; x: δ_{ij}^+ vs. d_{ij}^+). The closeness of the points to the first diagonal reflects the good quality of the model.

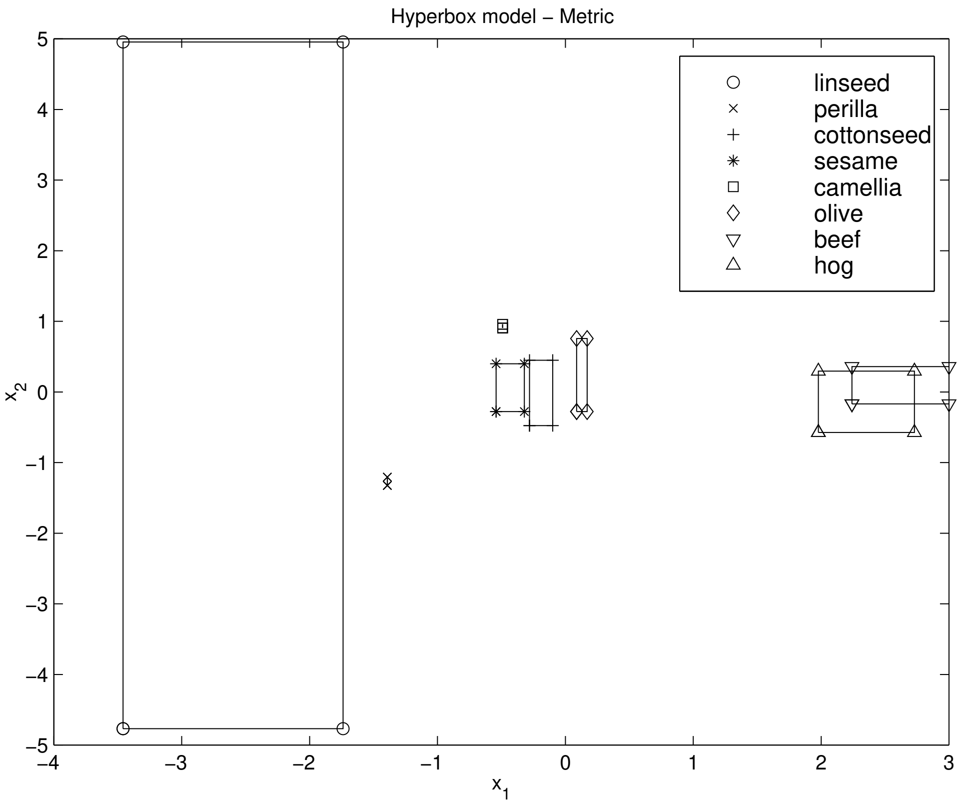

Figure 5: Two-dimensional metric configuration for the Fats and Oil data (hyperbox model). Each rectangle represents one of the 8 objects. The lower and upper Euclidean distances between any two rectangles approximate the distances between the corresponding objects.

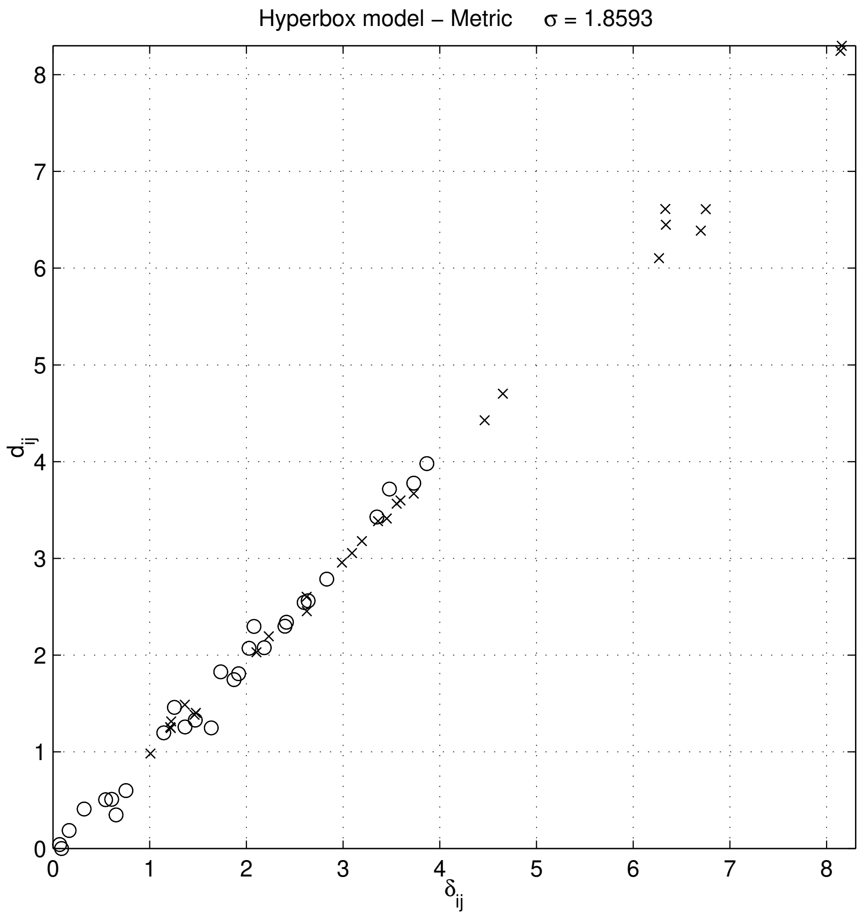

Figure 6: Real distances (x-axis) vs. reconstructed distances (y-axis) for the Fats and Oil data using the hyperbox model. (o: δ_{ij}^{-} vs. d_{ij}^{-}; x: δ_{ij}^{+} vs. d_{ij}^{+}).

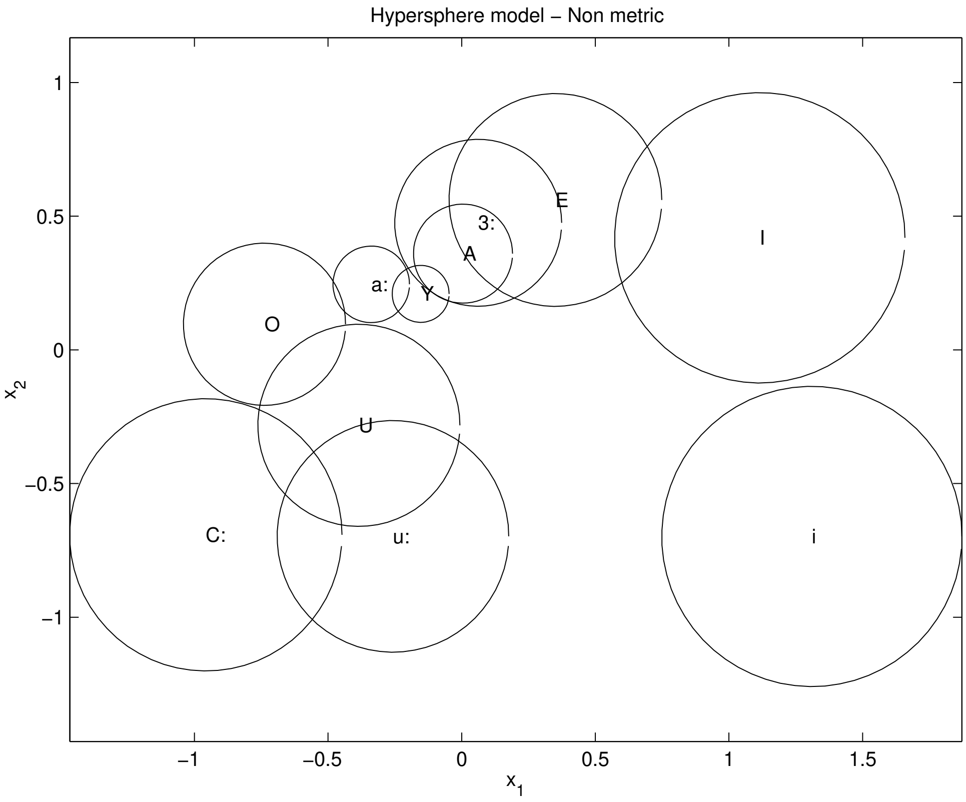

Figure 7: Two-dimensional non metric configuration for the Vowel data (hypersphere model).

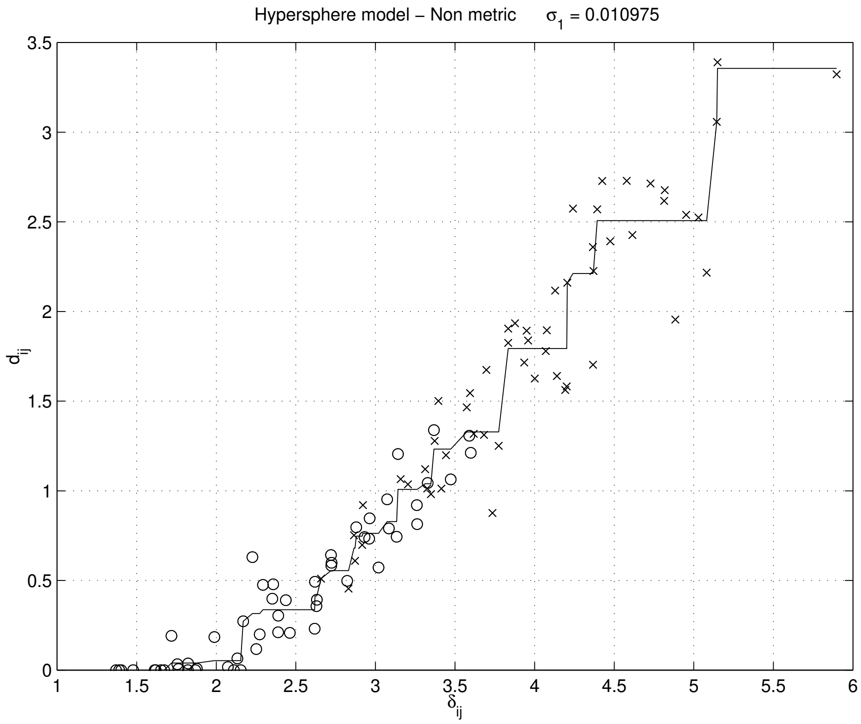

Figure 8: Shepard diagram for the Vowel data using the hypersphere model (o: lower distances; x: upper distances; -: isotonic regression function f.).

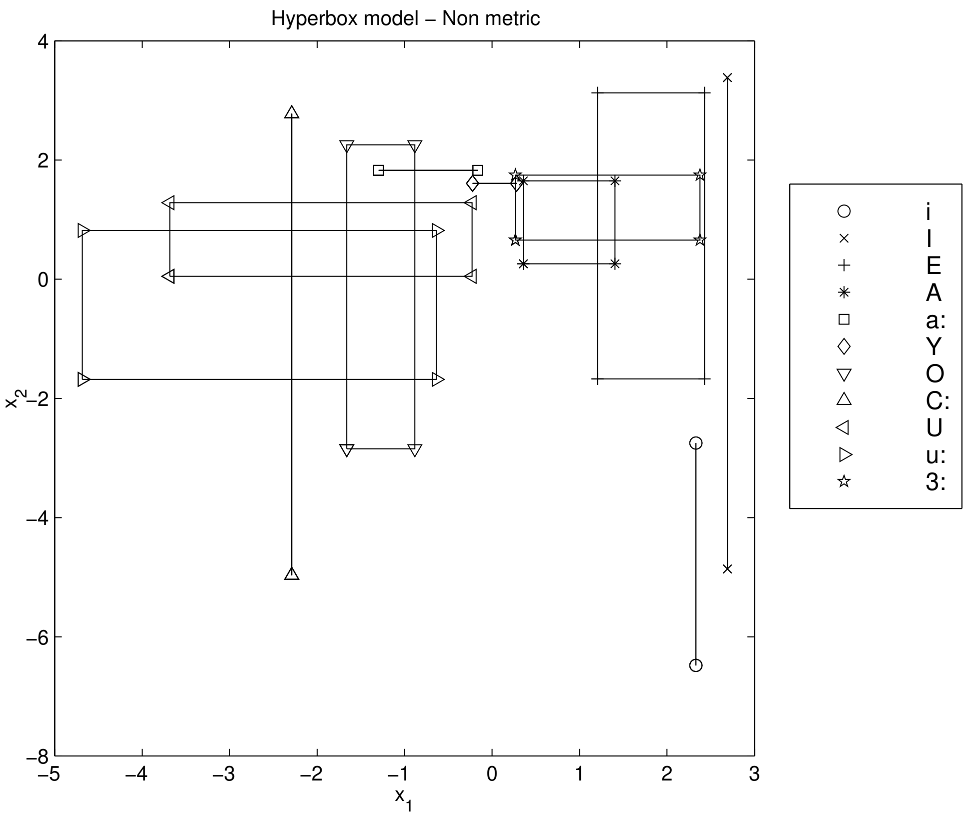

Figure 9: Two-dimensional non metric configuration for the Vowel data (hyperbox model).

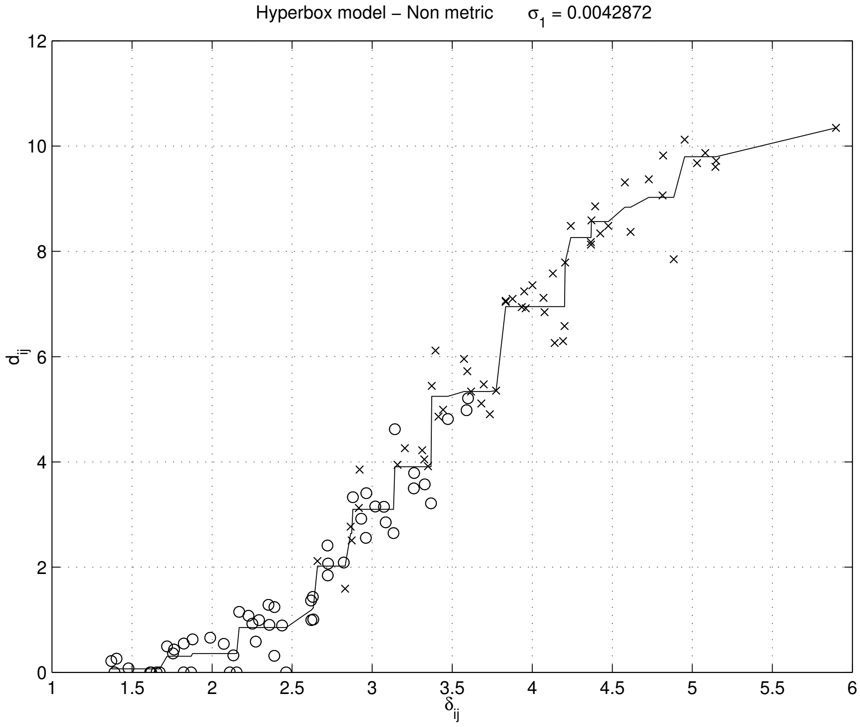

Figure 10: Shepard diagram for the Vowel data using the hyperbox model (o: lower distances; x: upper distances; -: isotonic regression function f.).