Two- and three-way component models for LR fuzzy data in a possibilistic framework

Paolo Giordani

Department of Statistics, Probability and Applied Statistics
University of Rome "La Sapienza"
P.le Aldo Moro, 5, 00185 Rome, Italy
paolo.giordani@uniroma1.it

Abstract

In this work we address the data reduction problem for fuzzy data. In particular, following a possibilistic approach, several component models for handling two- and three-way fuzzy data sets are introduced. The two-way models are based on classical Principal Component Analysis (PCA), whereas the three-way ones on three-way generalizations of PCA, as Tucker3 and CANDECOMP/PARAFAC. The here-proposed models exploit the potentiality of the possibilistic regression. In fact, the component models for fuzzy data can be seen as particular regression analyses between a set of observed fuzzy variables (response variables) and a set of unobservable crisp variables (explanatory variables). In order to show how the models work, the results of an application to a three-way fuzzy data set are illustrated.

Key words: Two- and three-way fuzzy data sets, Principal Component Analysis, Tucker3, CANDECOMP/PARAFAC, Possibilistic approach

1 Introduction

During the last years, several scientific studies have dealt with topics from both Statistics and Fuzzy Set Theory. Among the statistical methods for fuzzy experimental data, Principal Component Analysis (PCA) has received a lot of attention. In general, the basic aim of PCA is to synthesize huge amounts of data (a collection of I observation units on which J quantitative variables are registered) by finding a low number of unobserved variables, called *components*. These components are constructed from the observed variables in such a way that they maintain most of the information contained in the observed variables. In the literature, several generalizations of PCA for fuzzy experimental data are available. The first work is proposed in [13]. Later, a few papers devoted to PCA for fuzzy data following a least-squares approach can be found in [3,4,8]. Finally, just one paper [5] attempts to handle the data reduction problem for fuzzy data in a possibilistic setting.

In Statistics, PCA has been generalized in a three-way perspective. Three-way data usually refer to measurements related to three modes [1]. For instance, one may think about I observation units on which J (quantitative) variables are measured at K occasions. The occasions can be different points in time or, in general different measurable conditions. It should be noted that methods for two-way data (I observation units $\times$ J variables) may be used to manage three-way data. This can be done either by analyzing all the two-way data sets contained in the three-way data set separately (different PCA's for each occasion) and by aggregating two of the three modes. In the latter case, for instance, one may perform a PCA on the two-way data set with I objects and JK 'variables'. Here, the JK new variables refer to all the possible combinations of the J variables at the K occasions. Unfortunately, these approaches do not offer a deep and complete analysis of the three-way data set because they do not capture relevant information connected with the three-way interactions in the data. To this purpose, diverse methods for managing multi-way data have been proposed, among which the Tucker3 [12] and CANDECOMP/PARAFAC (independently proposed in [1] and in [6]) models. Their peculiarity is the capability of finding different components for each mode and analyzing the interrelations among them.

As far as the author knows, the extension of three-way models for handling fuzzy experimental data still remain to be done. In this paper, we aim at generalizing the Tucker3 and CANDECOMP/PARAFAC models for dealing with fuzzy data following the possibilistic approach. In this respect, we will extend the work in [5]. The three-way models will be set-up by formulating non-linear programming problems in which the fuzziness of the models is minimized. However, we first develop suitable generalizations of the two-way problem for all the possible families of LR fuzzy numbers.

The paper is organized as follows. In the next section classical PCA and the Tucker3 and CANDECOMP/PARAFAC models for crisp (non-fuzzy) data are illustrated. In Section 3, first we recall the possibilistic approach to PCA in [5]. Since this paper deals with (two-way) symmetric fuzzy data, we then generalize it to asymmetric LR_1 and LR_2 fuzzy data. Section 4 is devoted to the three-way extensions of the models described in Section 3. Finally, in Section 5, the results of an application of the proposed models to a three-way fuzzy data set are discussed.

[1] In three-way analysis, it is standard to use the term 'mode' to refer to a set of entities. Unfortunately, in fuzzy analysis, the same term refers to the point(s) of the fuzzy number with maximal membership function. Even if this could be confusing, we decided to ignore it. In fact, the meaning of the term can be easily depicted from the context.

2 Two- and three-way component models

PCA is a well-known statistical tool for analyzing a two-way crisp data set, say $\mathbf{X}$, of order $I \times J$ where I and J denote the numbers of observation units and variables, respectively. PCA aims at summarizing $\mathbf{X}$ by extracting S ($<J$) components, which are linear combinations of the observed variables. From a mathematical point of view, this consists of approximating the data matrix by means of two matrices $\mathbf{A}$ ($I \times S$) and $\mathbf{F}$ ($J \times S$) of rank S. The matrices $\mathbf{A}$ and $\mathbf{F}$ are usually referred to as the *component score matrix* and *component loading matrix*, respectively. The component score matrix gives the scores of the observation units on the components, whereas the component loading matrix is fruitful in order to interpret the components: high component loadings express that the observed variables are strictly related to the components. In scalar notation, PCA can be formulated as

$$x_{ij} \cong \sum_{s=1}^{S} a_{is} f_{js}, \tag{1}$$

$i = 1, \ldots, I; j = 1, \ldots, J$; where x_{ij} is the generic element of $\mathbf{X}$ and a_{is} and f_{js} are the generic elements of $\mathbf{A}$ and $\mathbf{F}$, respectively. In matrix notation, we have

$$\mathbf{X} \cong \mathbf{A}\mathbf{F}'. \tag{2}$$

The optimal component matrices are obtained by minimizing

$$\|\mathbf{X} - \mathbf{A}\mathbf{F}\|_F \,, \tag{3}$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

In the three-way framework, suppose to deal with a crisp data set concerning I observation units on which J variables at K occasions have been collected. We thus have a three-way data matrix $\mathbf{X}$ of order $I \times J \times K$. The three-way data matrix $\mathbf{X}$ can be seen as the collection of all the K two-way matrices (frontal slices) of order $I \times J$ corresponding to entities $k = 1, \ldots, K$, of the occasion mode. In PCA, $\mathbf{X}$ is summarized by means of the matrices $\mathbf{A}$ and $\mathbf{F}$. The matrix $\mathbf{A}$ summarizes the observation units and $\mathbf{F}$ the variables. In the three-way models, the matrix $\mathbf{A}$, now of order $I \times P$, still aims at summarizing the individuals, whereas the matrix $\mathbf{F}$ is suitably replaced by two component matrices for summarizing the variables ($\mathbf{B}$ of order $J \times Q$) and the occasions ($\mathbf{C}$ of order $K \times R$), where P, Q and R are the numbers of components for the observation unit, variable and occasion mode, respectively. Differently from PCA in which each component summarizing the observation units is

uniquely related to a component summarizing the variables, in the three-way context the components summarizing each mode are related to all the components summarizing the remaining two modes. Such interactions among all the components of the modes are stored in the so-called 'core' matrix $\mathbf{G}$ of order $P \times Q \times R$.
The scalar formulation of the Tucker3 model is as follows:

$$x_{ijk} \cong \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} a_{ip} b_{jq} c_{kr} g_{pqr}, \tag{4}$$

$i = 1, \ldots, I; j = 1, \ldots, J; k = 1, \ldots, K$; where x_{ijk}, a_{ip}, b_{jq}, c_{kr} and g_{pqr} are the generic elements of $\mathbf{X}$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{G}$, respectively. In matrix notation, we get:

$$\mathbf{X}_a \cong \mathbf{A}\mathbf{G}_a \left(\mathbf{C}' \otimes \mathbf{B}'\right), \tag{5}$$

where the symbol $\otimes$ denotes the Kronecker product and $\mathbf{X}_a$ of order $I \times JK$ and $\mathbf{G}_a$ of order $P \times QR$ are two-way matrices obtained by 'matricizing' $\mathbf{X}$ and $\mathbf{G}$, respectively. This consists of rearranging a three-way matrix into a two-way one. Specifically, we set up $\mathbf{X}_a$ by arranging the frontal slices $\mathbf{X}_{..k}, k = 1, \ldots, K$, next to each other:

$$\mathbf{X}_a = \left[\mathbf{X}_{..1} \cdots \mathbf{X}_{..k} \cdots \mathbf{X}_{..K} \right]. \tag{6}$$

Analogously, using the R frontal slices of $\mathbf{G}$, for the core we have:

$$\mathbf{G}_a = \left[\mathbf{G}_{..1} \cdots \mathbf{G}_{..r} \cdots \mathbf{G}_{..R} \right]. \tag{7}$$

Further details can be found in [9].
The optimal component matrices of the Tucker3 model are obtained by minimizing

$$\left\| \mathbf{X}_a - \mathbf{A}\mathbf{G}_a \left(\mathbf{C}' \otimes \mathbf{B}'\right) \right\|_F. \tag{8}$$

See, for more details, [10].
The CANDECOMP/PARAFAC model represents a generalization of classical PCA by adding the component matrix for the occasions:

$$x_{ijk} \cong \sum_{s=1}^{S} a_{is} b_{js} c_{ks}, . \tag{9}$$

It is important to observe that the CANDECOMP/PARAFAC model can be seen as a constrained version of the Tucker3 model imposing the core to be superdiagonal ($g_{pqr} = 1$ if $p = q = r$, $g_{pqr} = 0$ otherwise). As a consequence, a matrix formulation of CANDECOMP/PARAFAC is

$$\mathbf{X}_a \cong \mathbf{A}\mathbf{I}_a \left(\mathbf{C}' \otimes \mathbf{B}'\right), \tag{10}$$

where $\mathbf{I}_a$ is the matrix of order $S \times S^2$ obtained by juxtaposing the S frontal slices of the unit superdiagonal array. For a deeper insight on three-way methods, see [11]
In choosing between CANDECOMP/PARAFAC and Tucker3, one should observe that the former is more parsimonious but less general than the latter. Therefore, a researcher may act as follows. After choosing the number of components, one can perform CANDECOMP/PARAFAC. If the solution does not capture the relevant information, which relies in the data, one should try to further increase the number of components. However, it may happen that the CANDECOMP/PARAFAC model is too restrictive and, thus, it should be replaced by the Tucker3 model.

Remark 1: Three-way methods as constrained PCA's
The three-way methods, such as Tucker3 and CANDECOMP/PARAFAC, can be seen as a classical PCA on the data matrix $\mathbf{X}_a$ with constrained loadings. Specifically, after imposing that

$$\mathbf{F} = \left(\mathbf{C} \otimes \mathbf{B}\right)\mathbf{G}_a', \tag{11}$$

and substituting (11) in (5), we get

$$\mathbf{X}_a \cong \mathbf{A}\mathbf{F}', \tag{12}$$

in which $\mathbf{X}_a$ is decomposed into the matrices of the component scores and loadings as $\mathbf{X}$ in (2). Therefore, the Tucker3 model can be seen as a PCA with component loadings constrained to be equal to $\left(\mathbf{C} \otimes \mathbf{B}\right)\mathbf{G}_a'$. Such a constraint is fruitful in order to capture the three-way interactions in the data. Also CANDECOMP/PARAFAC is a constrained PCA on $\mathbf{X}_a$. Taking into account that the core matrix reduces to an identity matrix, the constraint on the component loadings given in (11) is replaced by

$$\mathbf{F} = \left(\mathbf{C} \otimes \mathbf{B}\right)\mathbf{I}_a' = \left(\mathbf{C} \odot \mathbf{B}\right), \tag{13}$$

where the symbol $\odot$ denotes the so-called Khatri-Rao product (columnwise Kronecker product).

In this paper, we shall extend Tucker3 (and, indeed, CANDECOMP/PARAFAC) to fuzzy data following a possibilistic point of view by using the PCA extension for fuzzy data proposed in [5], which will be briefly described in the next section.

3 Two-way Possibilistic PCA for fuzzy data

3.1 Fuzzy data

The two-way possibilistic PCA for fuzzy data as proposed in [5] limits its attention to the class of symmetric LR_1 fuzzy numbers. In general, an LR_1 fuzzy number is recognized by the triple $\tilde{X} = (m, l, r)_{LR}$ where m denotes the mode and l and r the left and right spreads, respectively, with the following membership function:

$$\mu_{\tilde{X}}(x) = \begin{cases} L\left(\frac{m-x}{l}\right) & x \leq m \ (l > 0) \\ R\left(\frac{x-m}{r}\right) & x \geq m \ (r > 0), \end{cases} \tag{14}$$

where L and R are continuous strictly decreasing functions on $[0, 1]$ called shape functions, which must fulfil additional requirements. For instance, with respect to $L(z)$, $z = \frac{m-x}{l}$: $L(0) = 1$, $0 < L(z) < 1$ for $0 < z < 1$, $L(1) = 0$. A particular case of LR_1 fuzzy numbers is the triangular one (with triangular membership function). In fact, if L and R are of the form

$$L(z) = R(z) = \begin{cases} 1 - z & 0 \leq z \leq 1 \\ 0 & \text{otherwise}, \end{cases} \tag{15}$$

then $\tilde{X} = (m, l, r)_{LR}$ is a triangular fuzzy number, with membership function:

$$\mu_{\tilde{X}}(x) = \begin{cases} 1 - \frac{m-x}{l} & x \leq m \ (l > 0) \\ 1 - \frac{x-m}{r} & x \geq m \ (\text{r} > 0). \end{cases} \tag{16}$$

A symmetric fuzzy number (L_1) is a particular LR_1 fuzzy number for which $L = R$ and $l = r$ and is denoted as $\tilde{X} = (m, l)_L$. Therefore, (14) is replaced by

$$\mu_{\tilde{X}}(x) = L\left(\frac{m-x}{l}\right) \quad m - l \leq x \leq m + l \ (\text{l} > 0). \tag{17}$$

with $L(z) = L(-z)$. Moreover, a symmetric triangular fuzzy number can be defined by replacing (16) with

$$\mu_{\tilde{X}}(x) = 1 - \frac{m-x}{l} \quad (l > 0). \tag{18}$$

The so-called LR_2 fuzzy number can also be introduced. It is recognized by the quadruple $\tilde{X} = (m_1, m_2, l, r)_{LR}$ where m_1 and m_2 denote the left and right modes, respectively, and l and r the left and right spreads, respectively, with the following membership function:

$$\mu_{\tilde{X}}(x) = \begin{cases} L\left(\frac{m_1-x}{l}\right) & x \leq m_1 \ (l > 0) \\ 1 & m_1 \leq x \leq m_2 \\ R\left(\frac{x-m_2}{r}\right) & x \geq m_2 \ (r > 0), \end{cases} \tag{19}$$

If L and R are of the form given in (15) then, by substituting them in (19), we get the so-called *trapezoidal* fuzzy number.

3.2 The PCA model for L_1 *data*

In [5], the available fuzzy data are stored into the fuzzy data matrix $\tilde{\mathbf{X}} = (\mathbf{M}, \mathbf{L})_L$ of order $I \times J$ where $\mathbf{M}$ and $\mathbf{L}$ are the matrices of the modes and the spreads, respectively. Each row of $\tilde{\mathbf{X}}$ ($\tilde{\mathbf{x}}_i = (\tilde{x}_{i1}, \ldots, \tilde{x}_{iJ}), i = 1, \ldots, I$) corresponds to an observation unit. Note that, as each observation unit is characterized by J non-interactive ([7]) fuzzy variables, it can be represented by a hyperrectangle in $\Re^J$. The membership function of $\tilde{\mathbf{x}}_i$ is given by $\mu_{\tilde{\mathbf{X}}_i}(\mathbf{u}_i) = \min_{j=1,\ldots,J} \mu_{\tilde{X}_{ij}}(u_{ij})$.

The possibilistic PCA can be seen as a regression analysis with fuzzy coefficients between J observed fuzzy variables (the dependent variables) and S unobservable crisp variables (the explanatory variables), which are the components resulting from PCA:

$$\tilde{\mathbf{X}} = \tilde{\mathbf{A}}\mathbf{F}', \tag{20}$$

where $\tilde{\mathbf{A}}$ is the fuzzy component score matrix of order $I \times S$ whose generic element is a symmetric fuzzy number denoted as $\left(a_{is}^M, a_{is}^L\right)_L$, $i = 1, ..., I$; $s = 1, ..., S$; (let $\mathbf{A}^M$ and $\mathbf{A}^L$ be, respectively, the matrices of the modes and of the spreads for the component scores) and $\mathbf{F}$ is the crisp component loading matrix of order $J \times S$. Therefore, as a natural extension of classical PCA, it is assumed that the (fuzzy) component score matrix takes into

account the fuzziness of the observed data involved. Note that in PCA the extracted components are uncorrelated. The same holds in the possibilistic PCA after imposing the columnwise orthonormality of $\mathbf{F}$. As a consequence, the obtained possibility distributions of the component scores (the coefficients of the regression model) are non-interactive.
The membership function of the generic estimated datum $\tilde{x}^*_{ij} = \left(\mathbf{a}_i^M \mathbf{f}'_j, \mathbf{a}_i^L \left|\mathbf{f}'_j\right|\right)_L$ is

$$\mu_{\tilde{X}^*_{ij}}(u_{ij}) = L\left(\frac{u_{ij} - \mathbf{a}_i^M \mathbf{f}'_j}{\mathbf{a}_i^L \left|\mathbf{f}'_j\right|}\right), \tag{21}$$

for $\mathbf{f}_j \neq \mathbf{0}$ where $\mathbf{a}_i^M$, $\mathbf{a}_i^L$, $i = 1, ..., I$, and $\mathbf{f}_j$, $j = 1, ..., J$, are the generic rows of $\mathbf{A}^M$, $\mathbf{A}^L$ and $\mathbf{F}$, respectively. If $\mathbf{f}_j = \mathbf{0}$ and $u_{ij} = 0$, then $\mu_{\tilde{X}^*_{ij}}(u_{ij}) = 1$ and, if $\mathbf{f}_j = \mathbf{0}$ and $u_{ij} \neq 0$, then $\mu_{\tilde{X}^*_{ij}}(u_{ij}) = 0$. The h-level set can then be obtained as

$$\left[\tilde{x}^*_{ij}\right]_h = \left[\mathbf{a}_i^M \mathbf{f}'_j - \left|L_{ij}^{-1}(h)\right| \mathbf{a}_i^L \left|\mathbf{f}'_j\right|, \mathbf{a}_i^M \mathbf{f}'_j + \left|L_{ij}^{-1}(h)\right| \mathbf{a}_i^L \left|\mathbf{f}'_j\right|\right], \tag{22}$$

$i = 1, ..., I$, $j = 1, ..., J$. Taking into account the non-interactivity assumption, the estimated vectors pertaining to the observation units have membership functions given by $\mu_{\tilde{\mathbf{X}}^*_i}(\mathbf{u}_i) = \min_{j=1,\ldots,J} \mu_{\tilde{X}^*_{ij}}(u_{ij})$, $i = 1, ..., I$.
In order to formulate the principal component model, it is also assumed that:
- the observed data $\tilde{\mathbf{x}}_i = (\tilde{x}_{i1}, \ldots, \tilde{x}_{iJ})$, $i = 1, \ldots, I$, can be represented by model (20);
- given a threshold value h, the observed data $\tilde{\mathbf{x}}_i = (\tilde{x}_{i1}, \ldots, \tilde{x}_{iJ})$, are included in the h-level set of the estimated data, $\tilde{\mathbf{x}}^*_i = (\tilde{x}^*_{i1}, \ldots, \tilde{x}^*_{iJ})$, $i = 1, \ldots, I$;
- the sum of spreads of the estimated data is considered as an objective function:

$$J_{2L} = \sum_{i=1}^{I}\sum_{j=1}^{J} \mathbf{a}_i^L \left|\mathbf{f}'_j\right|. \tag{23}$$

The possibilistic PCA for symmetric fuzzy data is then formulated as

$$\begin{aligned}
&\min\ J_{2L} \\
&\text{s.t.}\ \ \mathbf{M} + \mathbf{L} * \mathbf{H}_L \leq \mathbf{A}^M \mathbf{F}' + \left(\mathbf{A}^L \left|\mathbf{F}'\right|\right) * \mathbf{H}_L, \\
&\qquad \mathbf{M} - \mathbf{L} * \mathbf{H}_L \geq \mathbf{A}^M \mathbf{F}' - \left(\mathbf{A}^L \left|\mathbf{F}'\right|\right) * \mathbf{H}_L, \\
&\qquad \mathbf{F}'\mathbf{F} = \mathbf{I}, \\
&\qquad \mathbf{A}^L \geq \mathbf{\Psi},
\end{aligned} \tag{24}$$

where $\mathbf{H}_L$ is the $(I \times J)$ matrix with generic element $\left|L_{ij}^{-1}(h)\right|$, $i = 1, ..., I$, $j = 1, ..., J$ and $*$ denotes the Hadamard product (elementwise matrix product). In (24) we impose that the component loading matrix is columnwise orthonormal and that $\mathbf{A}^L$ is greater than a pre-specified matrix $\boldsymbol{\Psi}$ with non-negative elements. The elements of $\tilde{\mathbf{A}}$ are the coordinates of the low-dimensional (S-dimensional) hyperrectangles (with edges parallel to the new axes) associated to the observation units in the subspace spanned by the component loadings. If one does not want low dimensional hyperrectangles having S-dimensional volume equal to 0 because of at least one component score spread being equal to 0, it is fruitful to set some elements of $\boldsymbol{\Psi}$ strictly greater than zero.

3.3 The PCA model for LR$_1$ *fuzzy data*

Based on the previous results, we now extend the possibilistic PCA to the wider class of LR$_1$ fuzzy data. Thus, let us suppose that the observed fuzzy data matrix of order $I \times J$ is now $\tilde{\mathbf{X}} = (\mathbf{M}, \mathbf{L}, \mathbf{R})_{LR}$, where $\mathbf{R}$ is the additional matrix of the *right* spreads ($\mathbf{L}$ is the *left* spread matrix). The I observation units $(\tilde{\mathbf{x}}_i, i = 1, \ldots, I)$ are now characterized by J non-interactive *asymmetric* fuzzy numbers $(\tilde{x}_{i1}, \ldots, \tilde{x}_{iJ}), i = 1, \ldots, I$.
In this case, the PCA model can be formalized as

$$\tilde{\mathbf{X}} = \tilde{\mathbf{A}}\mathbf{F}' \tag{25}$$

where $\tilde{\mathbf{A}}$ is the component score matrix the generic element of which is $\left(a_{is}^M, a_{is}^L, a_{is}^R\right)$, where a_{is}^M, a_{is}^L and a_{is}^R are the generic elements of the component score matrices for, respectively, the modes, the left spreads and the right spreads, $i = 1, ..., I$; $s = 1, ..., S$. It follows that now the elements of $\tilde{\mathbf{A}}$ are asymmetric fuzzy numbers. As for the previous model, $\mathbf{F}$ is the crisp matrix of the component loadings. The membership function of the generic estimated datum is

$$\mu_{\tilde{X}_{ij}^*}(x) = \begin{cases} L\left(\left(\mathbf{a}_i^M \mathbf{f}_j' - x\right) \big/ \mathbf{a}_i^L \left|\mathbf{f}_j'\right|\right), & x \le \mathbf{a}_i^M \mathbf{f}_j', \\ R\left(\left(x - \mathbf{a}_i^M \mathbf{f}_j'\right) \big/ \mathbf{a}_i^R \left|\mathbf{f}_j'\right|\right), & x \ge \mathbf{a}_i^M \mathbf{f}_j', \end{cases} \tag{26}$$

for $\mathbf{f}_j \neq \mathbf{0}$ where $\mathbf{a}_i^M$, $\mathbf{a}_i^L$ and $\mathbf{a}_i^R$ are the i-th rows of $\mathbf{A}^M$, $\mathbf{A}^L$ and $\mathbf{A}^R$, respectively. If $\mathbf{f}_j = \mathbf{0}$ and $x = 0$, we set $\mu_{\tilde{x}_{ij}^*}(x) = 1$ and, if $\mathbf{f}_j = \mathbf{0}$ and $x \neq 0$, $\mu_{\tilde{x}_{ij}^*}(x) = 0$.
Note that (26) represents the asymmetric generalization of (21). The h-level set can then be easily obtained as

$$\left[\tilde{x}_{ij}^*\right]_h = \left[\mathbf{a}_i^M \mathbf{f}_j' - \left|L_{ij}^{-1}(h)\right| \mathbf{a}_i^L \left|\mathbf{f}_j'\right|, \mathbf{a}_i^M \mathbf{f}_j' + \left|R_{ij}^{-1}(h)\right| \mathbf{a}_i^R \left|\mathbf{f}_j'\right|\right], \tag{27}$$

$i = 1, ..., I$; $j = 1, ..., J$. As for the symmetric case, we get that the estimated vectors pertaining to the observation units have membership functions given by $\mu_{\tilde{\mathbf{X}}_i^*}(\mathbf{u}_i) = \min_{j=1,...,J} \mu_{\tilde{X}_{ij}^*}(u_{ij})$, $i = 1, ..., I$. Once more, in order to formulate the principal component model, it is also assumed that:
- the observed data $\tilde{\mathbf{x}}_i = (\tilde{x}_{i1}, \ldots, \tilde{x}_{iJ}), i = 1, \ldots, I$, can be represented by model (25);
- given a threshold value h, the observed data $\tilde{\mathbf{x}}_i = (\tilde{x}_{i1}, \ldots, \tilde{x}_{iJ})$, are included in the h-level set of the estimated data, $\tilde{\mathbf{x}}_i^* = (\tilde{x}_{i1}^*, \ldots, \tilde{x}_{iJ}^*), i = 1, \ldots, I$;
- the sum of spreads of the estimated data is considered as an objective function:

$$J_{2LR} = \sum_{i=1}^{I} \sum_{j=1}^{J} \left(\mathbf{a}_i^L + \mathbf{a}_i^R\right) \left|\mathbf{f}'_j\right|. \tag{28}$$

Therefore, the optimal matrices $\tilde{\mathbf{A}}$ and $\mathbf{F}$ should be constructed in such a way that

$$[\tilde{\mathbf{x}}_i]_h \subseteq [\tilde{\mathbf{x}}_i^*]_h \,, \tag{29}$$

$i = 1, \ldots, I$. We already pointed out that each observation unit can be represented as a hyperrectangle with $Z = 2^J$ vertices. The inclusion constraints in (29) hold whenever the inclusion constraints at the h-level hold for all the vertices associated to the observed and estimated hyperrectangles. It follow that

$$\begin{aligned} m_{ij} + r_{ij} q_{zj}^R \left|R_{ij}^{-1}(h)\right| &\leq \mathbf{a}_i^M \mathbf{f'}_j + \mathbf{a}_i^R \left|\mathbf{f'}_j\right| q_{zj} \left|R_{ij}^{-1}(h)\right| \text{ if } q_{zj}^R = 1, \\ m_{ij} - l_{ij} q_{zj}^L \left|L_{ij}^{-1}(h)\right| &\geq \mathbf{a}_i^M \mathbf{f'}_j - \mathbf{a}_i^L \left|\mathbf{f'}_j\right| q_{zj} \left|L_{ij}^{-1}(h)\right| \text{ if } q_{zj}^L = 1, \end{aligned} \tag{30}$$

$i = 1, \ldots, I; j = 1, \ldots, J; z = 1, \ldots, Z$. Note that q_{zj}^L and q_{zj}^R are the generic elements of the matrices $\mathbf{Q}^L$ and $\mathbf{Q}^R$. In the general case of J variables, the rows of $\mathbf{Q}^L$ contain all the possible J-dimensional vectors of 0 and 1. The number of rows of $\mathbf{Q}^L$ is equal to that of the vertices of a J-dimensional hyperrectangle, which is Z. $\mathbf{Q}^R$ has the same structure of $\mathbf{Q}^L$ but the elements equal to 0 and 1 have switched places. Every pair of rows of $\mathbf{Q}^L$ and $\mathbf{Q}^R$ allows us to describe a given vertex. For instance, if $J = 3$, we have

$$\mathbf{Q}^L = \begin{bmatrix} 1\ 1\ 1 \\ 1\ 1\ 0 \\ 1\ 0\ 1 \\ 0\ 1\ 1 \\ 0\ 0\ 0 \\ 0\ 0\ 1 \\ 0\ 1\ 0 \\ 1\ 0\ 0 \end{bmatrix} \tag{31}$$

and the first row refers to the vertex of all the lower bounds.
Now, we can simplify the inclusion constraints in (30). To do so, we observe that, among the 2^J inclusions (one for each vertex), the attention has to be paid only to the vertices of the lower and upper bounds. In fact,

$$\begin{aligned} &\mathbf{M} - \left(\mathbf{L}\mathbf{Q}_1^L\right) * \mathbf{H}_L = \mathbf{M} - \mathbf{L} * \mathbf{H}_L \leq \mathbf{M} - \left(\mathbf{L}\mathbf{Q}_z^L\right) * \mathbf{H}_L, \\ &\mathbf{M} + \left(\mathbf{R}\mathbf{Q}_{2^{J-1}+1}^R\right) * \mathbf{H}_R = \mathbf{M} + \mathbf{R} * \mathbf{H}_R \geq \mathbf{M} + \left(\mathbf{R}\mathbf{Q}_z^R\right) * \mathbf{H}_R, \end{aligned} \tag{32}$$

where $\mathbf{Q}_z^L$ and $\mathbf{Q}_z^R$ are diagonal matrices whose non-zero elements are those of the z-th rows of $\mathbf{Q}^L$ and $\mathbf{Q}^R$, respectively, $z = 1, \ldots, Z$. It follows that the inclusion constraints in (32) hold if $\tilde{\mathbf{A}}$ and $\mathbf{F}$ are such that the observed lower bounds are higher than the estimated ones and the observed upper bounds are lower than the estimated ones because

$$\begin{aligned} &\mathbf{M}^* - \mathbf{L}^* * \mathbf{H}_L \leq \mathbf{M} - \mathbf{L} * \mathbf{H}_L \leq \mathbf{M} - \left(\mathbf{L}\mathbf{Q}_z^L\right) * \mathbf{H}_L, \\ &\mathbf{M}^* + \mathbf{R}^* * \mathbf{H}_R \geq \mathbf{M} + \mathbf{R} * \mathbf{H}_R \geq \mathbf{M} + \left(\mathbf{R}\mathbf{Q}_z^R\right) * \mathbf{H}_R, \end{aligned} \tag{33}$$

$z = 1, \ldots, Z$. Taking into account that $\mathbf{M}^* = \mathbf{A}^M\mathbf{F}'$, $\mathbf{L}^* = \mathbf{A}^L\mathbf{F}'$ and $\mathbf{R}^* = \mathbf{A}^R\mathbf{F}'$, it follows that (33) can be rewritten as

$$\begin{aligned} &\mathbf{A}^M\mathbf{F}' - \left(\mathbf{A}^L\mathbf{F}'\right) * \mathbf{H}_L \leq \mathbf{M} - \mathbf{L} * \mathbf{H}_L, \\ &\mathbf{A}^M\mathbf{F}' + \left(\mathbf{A}^R\mathbf{F}'\right) * \mathbf{H}_R \leq \mathbf{M} + \mathbf{R} * \mathbf{H}_R. \end{aligned} \tag{34}$$

So far, we have shown that the inclusion constraints in (29) can be simplified by means of (34). In addition, we impose that the component loadings

are columnwise orthonormal ($\mathbf{F}'\mathbf{F} = \mathbf{I}$) and that the spreads of the component scores are greater than pre-specified matrices with non-negative elements ($\mathbf{A}^L \geq \mathbf{\Psi}^L$ and $\mathbf{A}^R \geq \mathbf{\Psi}^R$).
By considering all the above constraints, we then get the following minimization problem:

$$
\begin{aligned}
\min\ & J_{2LR} \\
\text{s.t.}\ & \mathbf{M} + \mathbf{R} * \mathbf{H}_R \leq \mathbf{A}^M\mathbf{F}' + \left(\mathbf{A}^R\,|\mathbf{F}'|\right) * \mathbf{H}_R, \\
& \mathbf{M} - \mathbf{L} * \mathbf{H}_L \geq \mathbf{A}^M\mathbf{F}' - \left(\mathbf{A}^L\,|\mathbf{F}'|\right) * \mathbf{H}_L, \\
& \mathbf{F}'\mathbf{F} = \mathbf{I}, \\
& \mathbf{A}^L \geq \mathbf{\Psi}^L, \\
& \mathbf{A}^R \geq \mathbf{\Psi}^R.
\end{aligned}
\tag{35}
$$

It is fruitful to observe that, as for the symmetric case, the low dimensional plot of the observation units (as low dimensional hyperrectangles) provide information about the position of the observation units and the associated uncertainty. More specifically, the modes can be plotted onto the subspace spanned by the (columnwise orthonormal) matrix $\mathbf{F}$ using $\mathbf{A}^M$ as coordinates, whereas the bounds of the low dimensional hyperrectangles can be attained using the spreads of the component scores, leading to the interval $a_{is}^M - a_{is}^L$ and $a_{is}^M + a_{is}^R$, $i = 1, \ldots, I$; $s = 1, \ldots, S$. For each axis (component) and each observation unit, the width of the S-dimensional hyperrectangle, can be interpreted as a measure of uncertainty of the position of the observation unit for the axis (component) involved. If one is interested in observing the configuration of the observation units onto the obtained subspace, one should observe the position of the modes. In fact, in the asymmetric case, by observing the configuration of the hyperrectangles, one may find misleading results. In fact, it may happen that, for some observation units, one spread of the component scores is high and the other one is very near to zero. With respect to the first component, this may cause that, for instance, some (low dimensional) hyperrectangles move to the left side when the left and right spreads of the involved component scores are, respectively, high and very near to zero. It follows that, as a_{i1}^M can be found near to the upper bound $a_{i1}^M + a_{i1}^R$, $i = 1, ..., I$, the position of the (low dimensional) hyperrectangles gives misleading information about the low dimensional configuration of observations. Therefore, the size of the low dimensional hyperrectangles is useful in order to analyze the uncertainty associated to the observation units, whereas the position of the modes is relevant in order to analyze the configuration of the observation units.
Note that this ambiguity does not occur in the symmetric case because the bounds for the s-th component, $s = 1, \ldots, S$, are $a_{is}^M - a_{is}^L$ and $a_{is}^M + a_{is}^L$, $i = 1, \ldots, I$. Thus, with respect to the i-th observation unit, $i = 1, \ldots, I$, and

the s-th component, $s = 1, \ldots, S$, the mode is located in the middle point of the interval $\left(a_{is}^M - a_{is}^L, a_{is}^M + a_{is}^L\right)$. Therefore, in the symmetric case, the configuration of the observation units can be analyzed by means of the position of both the modes and the hyperrectangles.

Proposition 1: There always exists a solution of the NLP problem in (35).

Proof: Given a feasible solution for $\mathbf{F}$ such that the loadings are columnwise orthonormal, we can find a feasible solution for $\tilde{\mathbf{A}}$ according to the constraints in (35) by taking sufficiently large positive matrices for $\mathbf{A}_L$ and $\mathbf{A}_R$.

q.e.d.

Remark 2: Triangular membership function
When a triangular fuzzy number is used, $|L^{-1}(h)| = 1 - h$ and $|R^{-1}(h)| = 1 - h$.

Remark 3: Number of extracted components
In order to detect the optimal number of extracted components, we suggest to choose S such that it can be considered optimal in performing classical PCA on the (crisp) mode matrix.

Remark 4: Preprocessing
If necessary, it is advisable to preprocess the modes by subtracting the mean and dividing by the standard deviation of the variable at hand, whereas the spreads can be preprocessed by dividing them by the standard deviation of the related mode.

Remark 5: Lower problem
The problem (35) is usually referred to as the *Upper* Problem. As for the symmetric case, the *Lower* PCA Model can also be formulated as

$$\begin{aligned}
&\max\ J_{2LR} \\
&\text{s.t.}\quad \mathbf{M} + \mathbf{R} * \mathbf{H}_R \geq \mathbf{A}^M\mathbf{F}' + \left(\mathbf{A}^R\,|\mathbf{F}'|\right) * \mathbf{H}_R, \\
&\qquad\ \ \mathbf{M} - \mathbf{L} * \mathbf{H}_L \leq \mathbf{A}^M\mathbf{F}' - \left(\mathbf{A}^L\,|\mathbf{F}'|\right) * \mathbf{H}_L, \\
&\qquad\ \ \mathbf{F}'\mathbf{F} = \mathbf{I}, \\
&\qquad\ \ \mathbf{A}^L \geq \mathbf{\Psi}^L, \\
&\qquad\ \ \mathbf{A}^R \geq \mathbf{\Psi}^R,
\end{aligned} \tag{36}$$

in which one aims at finding the largest fuzzy set which satisfies

$$\left[x_{ij}\right]_h \supseteq \left[x_{ij}^*\right]_h, \tag{37}$$

$i = 1, ..., I$; $j = 1, ..., J$. Note that the NLP problem in (36) may not have any optimal solution.

3.4 The PCA model for LR_2 fuzzy data

For the sake of completeness, we briefly propose the PCA model for LR_2 fuzzy data. Now, the available data are stored in the fuzzy data matrix $\tilde{\mathbf{X}} = (\mathbf{M}_1, \mathbf{M}_2, \mathbf{L}, \mathbf{R})_{LR}$ of order $I \times J$, where $\mathbf{M}_1$ and $\mathbf{M}_2$ are the matrices of the left and right modes, respectively. In this case, the I observation units ($\tilde{\mathbf{x}}_i, i = 1, \ldots, I$) are characterized by J non-interactive LR_2 fuzzy numbers.
It follows that the PCA model can be formalized as

$$\tilde{\mathbf{X}} = \tilde{\mathbf{A}}\mathbf{F}' \tag{38}$$

where $\tilde{\mathbf{A}}$ is the component score matrix the generic element of which is $\left(a_{is}^{M_1}, a_{is}^{M_2}, a_{is}^{L}, a_{is}^{R}\right)$, where $a_{is}^{M_1}$, $a_{is}^{M_2}$, a_{is}^{L} and a_{is}^{R} are the generic elements of the component score matrices for, respectively, the left modes, the right modes, the left spreads and the right spreads, $i = 1, ..., I$; $s = 1, ..., S$. Therefore, as a natural extension of the previous cases, the elements of $\tilde{\mathbf{A}}$ are LR_2 fuzzy numbers, while $\mathbf{F}$ is still the crisp matrix of the component loadings.
Similarly to the previous section, by introducing the same assumptions and constraints, it is easy to show that the PCA problem consists of extending (35) as

$$\begin{aligned}
&\min\ J_{2LR} \\
&\text{s.t.}\ \ \mathbf{M}_2 + \mathbf{R} * \mathbf{H}_R \leq \mathbf{A}^{M_2}\mathbf{F}' + \left(\mathbf{A}^R\,|\mathbf{F}'|\right) * \mathbf{H}_R, \\
&\qquad \mathbf{M}_1 - \mathbf{L} * \mathbf{H}_L \geq \mathbf{A}^{M_1}\mathbf{F}' - \left(\mathbf{A}^L\,|\mathbf{F}'|\right) * \mathbf{H}_L, \\
&\qquad \mathbf{F}'\mathbf{F} = \mathbf{I}, \\
&\qquad \mathbf{A}^L \geq \mathbf{\Psi}^L, \\
&\qquad \mathbf{A}^R \geq \mathbf{\Psi}^R.
\end{aligned} \tag{39}$$

4 Three-way Possibilistic PCA for fuzzy data

This section is devoted to suitable extensions of Tucker3 and CANDECOMP/PARAFAC for handling fuzzy data following a possibilistic approach. In particular, three different analyses will be proposed according to the type of fuzzy data involved (L_1, LR_1 and LR_2 fuzzy numbers).

4.1 Fuzzy data

Let us suppose that the available data pertain to I observation units on which J non-interactive fuzzy variables are collected at K occasions. Each observation unit is thus recognized by JK fuzzy scores (the scores of all the variables at all the occasions): $\tilde{\mathbf{x}}_i = (\tilde{x}_{i11}, \ldots, \tilde{x}_{iJ1}, \ldots, \tilde{x}_{i1K}, \ldots, \tilde{x}_{iJK}), i = 1, \ldots, I$. From a geometrical point of view, each observation can be represented as a hyperrectangle in $\Re^{JK}$ identified by 2^{JK} vertices. Therefore, all the assumptions which were used in the two-way cases still hold since the main difference only relies in the number of vertices of the hyperrectangles.

4.2 The Three-way PCA model for L_1 *fuzzy data*

The first three-way extension of PCA in a possibilistic setting concerns the case in which the fuzzy variables are of L_1 type. Thus, we assume that the data are stored in the three-way fuzzy data matrix $\tilde{\mathbf{X}} = (\mathbf{M}, \mathbf{L})_L$ of order $I \times J \times K$, where $\mathbf{M}$ and $\mathbf{L}$ are the three-way matrices of the modes and the spreads, respectively. Following (6), the two-way supermatrices of the data $\tilde{\mathbf{X}}_a$, and, indeed, of the modes, $\mathbf{M}_a$, and the spreads, $\mathbf{L}_a$, can be defined. Using the Tucker3 model, the three-way PCA model can be expressed as

$$\tilde{\mathbf{X}}_a = \tilde{\mathbf{A}}\mathbf{G}_a(\mathbf{C}' \otimes \mathbf{B}') = \tilde{\mathbf{A}}\mathbf{F}', \tag{40}$$

where $\mathbf{B}$, $\mathbf{C}$ are the crisp component matrices for the variables and occasions, respectively, and $\mathbf{G}_a$ is the observation unit mode matricized version of the crisp core matrix $\mathbf{G}$. Finally, $\tilde{\mathbf{A}}$ is the fuzzy component matrix for the observation units. In particular, each element of $\tilde{\mathbf{A}}$ is an L_1 fuzzy number. Thus, the component matrix $\tilde{\mathbf{A}}$ captures the uncertainty associated with the observed data. In particular the generic element of $\tilde{\mathbf{A}}$ is the symmetric fuzzy number $\left(a_{ip}^M, a_{ip}^L\right)_L$, $i = 1, ..., I$; $p = 1, ..., P$; where a_{ip}^M and a_{ip}^L are the generic elements of the matrices of the modes ($\mathbf{A}^M$) and of the spreads ($\mathbf{A}^L$), respectively. If the CANDECOMP/PARAFAC model is performed, (40) is replaced by

$$\tilde{\mathbf{X}}_a = \tilde{\mathbf{A}}\mathbf{I}_a(\mathbf{C}' \otimes \mathbf{B}') = \tilde{\mathbf{A}}(\mathbf{C}' \odot \mathbf{B}') = \tilde{\mathbf{A}}\mathbf{F}'. \tag{41}$$

The matrix $\mathbf{F}$ is the constrained component loading matrix defined in (11) and, implicitly, in (40) for the Tucker3 model; and in (13) and, implicitly, in (41) for the CANDECOMP/PARAFAC model. As for the two-way case, we impose that the constrained component loadings $\mathbf{F}$ are columnwise orthonormal ($\mathbf{F}'\mathbf{F}=\mathbf{I}$). It follows that, the obtained possibility distributions of the component scores are non-interactive.

Let $\tilde{x}^*_{ijk} = \left(\mathbf{a}_i^M \mathbf{f'}_{jk}, \mathbf{a}_i^L \left|\mathbf{f'}_{jk}\right|\right)_L$ be the generic estimated datum, where $\mathbf{f}_{jk}$ is the jk-row of $\mathbf{F}$. Its membership function is

$$\mu_{\tilde{X}^*_{ijk}}(u_{ijk}) = L\left(\frac{u_{ijk} - \mathbf{a}_i^M \mathbf{f}'_{jk}}{\mathbf{a}_i^L \left|\mathbf{f}'_{jk}\right|}\right), \tag{42}$$

for $\mathbf{f}_{jk} \neq \mathbf{0}$ where $\mathbf{a}_i^M$, $\mathbf{a}_i^L$, $i = 1, ..., I$, and $\mathbf{f}_{jk}$, $j = 1, ..., J$, $k = 1, ..., K$, are the generic rows of $\mathbf{A}^M$, $\mathbf{A}^L$ and $\mathbf{F}$, respectively. If $\mathbf{f}_{jk} = \mathbf{0}$ and $u_{ijk} = 0$, then $\mu_{\tilde{X}^*_{ijk}}(u_{ijk}) = 1$ and, if $\mathbf{f}_{jk} = \mathbf{0}$ and $u_{ijk} \neq 0$, then $\mu_{\tilde{X}^*_{ijk}}(u_{ijk}) = 0$. Starting from the membership function in (42), the associated h-level set is then computed as

$$\left[\tilde{x}^*_{ijk}\right]_h = \left[\mathbf{a}_i^M \mathbf{f}'_{jk} - \left|L^{-1}_{ijk}(h)\right| \mathbf{a}_i^L \left|\mathbf{f}'_{jk}\right|, \mathbf{a}_i^M \mathbf{f}'_{jk} + \left|L^{-1}_{ijk}(h)\right| \mathbf{a}_i^L \left|\mathbf{f}'_{jk}\right|\right]. \tag{43}$$

In order to formulate the possibilistic Tucker3 model (and similarly the CANDECOMP/PARAFAC model) we also assume that:
- the observed data $\tilde{\mathbf{x}}_i = (\tilde{x}_{i11}, \ldots, \tilde{x}_{iJ1}, \ldots, \tilde{x}_{i1K}, \ldots, \tilde{x}_{iJK}), i = 1, \ldots, I$, can be represented by model (40) (or (41));
- given a threshold value h, the observed data $\tilde{\mathbf{x}}_i = (\tilde{x}_{i11}, \ldots, \tilde{x}_{iJ1}, \ldots, \tilde{x}_{i1K}, \ldots, \tilde{x}_{iJK})$, $i = 1, \ldots, I$, are included in the h-level set of the estimated data, $\tilde{\mathbf{x}}^*_i = (\tilde{x}^*_{i11}, \ldots, \tilde{x}^*_{iJ1}, \ldots, \tilde{x}^*_{i1K}, \ldots, \tilde{x}^*_{iJK})$, $i = 1, \ldots, I$;
- the sum of spreads of the estimated data is considered as an objective function:

$$J_{3L} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \mathbf{a}_i^L \left|\mathbf{f}'_{jk}\right|. \tag{44}$$

Once more, the optimal component matrices are determined in such a way that (44) is minimized and the h-level of the estimated data contains the corresponding h-level of the observed data:

$$[\tilde{\mathbf{x}}_i]_h \subseteq [\tilde{\mathbf{x}}^*_i]_h, i = 1, \ldots, I, \tag{45}$$

with

$$\mu_{\tilde{\mathbf{X}}^*_i}(\mathbf{u}_i) = \min_{j=1,\ldots,J} \mu_{\tilde{X}^*_{ij}}(u_{ij}), i = 1, \ldots, I. \tag{46}$$

It should be clear that, in the three-way framework, we deal with a minimization problem similar to the two-way one. In fact, we only need to add that the component loadings are constrained according to (11) or (13). More specifically, we can observe that the function to be minimized in (44) coincides with that in (23). As for the two-way case, the inclusion constraints between observed and estimated fuzzy data can be exploited by considering all the 2^{JK} vertices of the hyperrectangles associated to the observed and estimated data. Using the same manipulations as adopted for the two-way case, the inclusion constraints can be reduced to considering the constraints between observed and estimated upper and lower bounds. We then get

$$\begin{aligned} &\mathbf{M}_a + \mathbf{L}_a * \mathbf{H}_L \leq \mathbf{A}^M \mathbf{F}' + \left(\mathbf{A}^L \left|\mathbf{F}'\right|\right) * \mathbf{H}_L, \\ &\mathbf{M}_a - \mathbf{L}_a * \mathbf{H}_L \geq \mathbf{A}^M \mathbf{F}' - \left(\mathbf{A}^L \left|\mathbf{F}'\right|\right) * \mathbf{H}_L. \end{aligned} \tag{47}$$

After imposing the non-negativity of the spreads of the component scores for the observation units ($\mathbf{A}^L \geq \mathbf{\Psi}$), we then obtain the following problem for the Tucker3 model:

$$\begin{aligned} &\min\ J_{3L} \\ &\text{s.t.}\ \ \mathbf{M}_a + \mathbf{L}_a * \mathbf{H}_L \leq \mathbf{A}^M \mathbf{F}' + \left(\mathbf{A}^L \left|\mathbf{F}'\right|\right) * \mathbf{H}_L, \\ &\qquad \mathbf{M}_a - \mathbf{L}_a * \mathbf{H}_L \geq \mathbf{A}^M \mathbf{F}' - \left(\mathbf{A}^L \left|\mathbf{F}'\right|\right) * \mathbf{H}_L, \\ &\qquad \mathbf{F} = \left(\mathbf{C} \otimes \mathbf{B}\right) \mathbf{G}'_a, \\ &\qquad \mathbf{F}'\mathbf{F} = \mathbf{I}, \\ &\qquad \mathbf{A}^L \geq \mathbf{\Psi}. \end{aligned} \tag{48}$$

Note that the requirement that $\mathbf{F}$ is columnwise orthonormal can be replaced by the constraints that the component matrices $\mathbf{B}$ and $\mathbf{C}$ are columnwise orthonormal and the core matrix is rowwise orthonormal. This increases the function to be minimized, but it helps in order to interpret the extracted components. If the CANDECOMP/PARAFAC model is performed, the third constraint should be replaced by (13).

Proposition 2: There always exists a solution of the NLP problem in (48).

Proof: A feasible solution for $\mathbf{F}$ such that the constraints in (48) are satisfied can be found, for instance in the Tucker3 case, by choosing columnwise orthonormal matrices $\mathbf{B}$ and $\mathbf{C}$ and a rowwise orthonormal matrix $\mathbf{G}_a$. In fact, we have that $\mathbf{F}'\mathbf{F} = \mathbf{G}_a \left(\mathbf{C}' \otimes \mathbf{B}'\right) \left(\mathbf{C} \otimes \mathbf{B}\right) \mathbf{G}'_a = \mathbf{G}_a \left(\mathbf{C}'\mathbf{C} \otimes \mathbf{B}'\mathbf{B}\right) \mathbf{G}'_a = \mathbf{G}_a\mathbf{G}'_a = \mathbf{I}$. Now, given a feasible solution for $\mathbf{F}$, we can find a feasible solution for $\tilde{\mathbf{A}}$ according to the constraints in (48) by taking a sufficiently large positive matrix for $\mathbf{A}^L$.

q.e.d.

Remark 6: Numbers of extracted components
The choice of the number of extracted components is a complex issue in the three-way framework. See, for more details, [10]. However, a plausible suggestion is to perform the classical three-way model on the (crisp) mode matrix and to consider optimal, in the possibilistic case, the same numbers of extracted components chosen in the crisp case.

Remark 7: Preprocessing
Also the preprocessing step is rather complex in the three-way case. In fact, for three-way data, it is not obvious how each of the modes (the sets of entities) should be dealt with in centering and/or scaling the data. See, for further details, [9]. If necessary, we suggest to preprocess the modes by centering and/or scaling them. Then, if the modes are scaled, the spreads can be scaled by using the same scaling factor adopted for the modes.

Remark 8: Lower problem
The problem (48) is usually referred to as the *Upper* Problem. The *Lower* Tucker3 Model can also be formulated as the solution of

$$\begin{aligned} &\max\ J_{3L} \\ &\text{s.t.}\quad \mathbf{M}_a + \mathbf{L}_a * \mathbf{H}_L \geq \mathbf{A}^M\mathbf{F}' + \left(\mathbf{A}^L\left|\mathbf{F}'\right|\right) * \mathbf{H}_L, \\ &\qquad\ \ \mathbf{M}_a - \mathbf{L}_a * \mathbf{H}_L \leq \mathbf{A}^M\mathbf{F}' - \left(\mathbf{A}^L\left|\mathbf{F}'\right|\right) * \mathbf{H}_L, \\ &\qquad\ \ \mathbf{F} = (\mathbf{C}\otimes\mathbf{B})\,\mathbf{G}'_a, \\ &\qquad\ \ \mathbf{F}'\mathbf{F} = \mathbf{I}, \\ &\qquad\ \ \mathbf{A}^L \geq \mathbf{\Psi}, \end{aligned} \tag{49}$$

in which one aims at finding the largest fuzzy set which satisfies

$$\left[x_{ijk}\right]_h \supseteq \left[x^*_{ijk}\right]_h, \tag{50}$$

$i = 1, ..., I$; $j = 1, ..., J$; $k = 1, ..., K$. Note that the NLP problem in (49) may not have any optimal solution.

4.3 The Three-way PCA model for LR_1 *fuzzy data*

The generalization of the three-way models for LR_1 fuzzy data following a possibilistic approach is easily obtained taking into account (11) or (13) and the results introduced in Section 3.3. The observed data are stored in the three-way matrix $\tilde{\mathbf{X}} = (\mathbf{M}, \mathbf{L}, \mathbf{R})_{LR}$ of order $I \times J \times K$ with LR_1 fuzzy numbers.

The possibilistic Tucker3 model can be expressed as

$$\tilde{\mathbf{X}}_a = \tilde{\mathbf{A}}\mathbf{G}_a\left(\mathbf{C}' \otimes \mathbf{B}'\right) = \tilde{\mathbf{A}}\mathbf{F}^{'}, \tag{51}$$

and the possibilistic CANDECOMP/PARAFAC model as

$$\tilde{\mathbf{X}}_a = \tilde{\mathbf{A}}\mathbf{I}_a\left(\mathbf{C}' \otimes \mathbf{B}'\right) = \tilde{\mathbf{A}}\left(\mathbf{C}' \odot \mathbf{B}'\right) = \tilde{\mathbf{A}}\mathbf{F}^{'}, \tag{52}$$

where, of course, $\tilde{\mathbf{A}}$ is the LR_1 fuzzy component matrix for the observation units.
The aim is to determine component matrices such that the observed data are included in the estimated ones. In particular, we assume that:
- the observed data $\tilde{\mathbf{x}}_i, i = 1, \ldots, I$, can be represented by model (51) or (52);
- given a threshold value h, the observed data $\tilde{\mathbf{x}}_i$ are included in the h-level set of the estimated data, $\tilde{\mathbf{x}}_i^*, i = 1, \ldots, I$;
- the sum of spreads of the estimated data is considered as an objective function:

$$J_{3LR} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\mathbf{a}_i^L + \mathbf{a}_i^R\right)\left|\mathbf{f}'_{jk}\right|. \tag{53}$$

The estimated datum has the following membership function:

$$\mu_{\tilde{X}^*_{ijk}}(x) = \begin{cases} L\left(\left(\mathbf{a}_i^M\mathbf{f'}_{jk} - x\right)\Big/\mathbf{a}_i^L\left|\mathbf{f'}_{jk}\right|\right), x \leq \mathbf{a}_i^M\mathbf{f'}_{jk}, \\ R\left(\left(x - \mathbf{a}_i^M\mathbf{f'}_{jk}\right)\Big/\mathbf{a}_i^R\left|\mathbf{f'}_{jk}\right|\right), x \geq \mathbf{a}_i^M\mathbf{f'}_{jk}, \end{cases} \tag{54}$$

for $\mathbf{f}_{jk} \neq \mathbf{0}$. If $\mathbf{f}_{jk} = \mathbf{0}$ and $x = 0$, we set $\mu_{\tilde{x}^*_{ijk}}(x) = 1$ and, if $\mathbf{f}_{jk} = \mathbf{0}$ and $x \neq 0$, $\mu_{\tilde{x}^*_{ijk}}(x) = 0$.
From (54), the h-level can be computed as

$$\left[\tilde{x}^*_{ijk}\right]_h = \left[\mathbf{a}_i^M\mathbf{f}'_{jk} - \left|L^{-1}_{ijk}(h)\right|\mathbf{a}_i^L\left|\mathbf{f}'_{jk}\right|, \mathbf{a}_i^M\mathbf{f}'_{jk} + \left|R^{-1}_{ijk}(h)\right|\mathbf{a}_i^R\left|\mathbf{f}'_{jk}\right|\right], \tag{55}$$

$i = 1, ..., I$; $j = 1, ..., J$; $k = 1, ..., K$. In the Tucker3 case, by imposing that the component loadings are columnwise orthonormal ($\mathbf{F}'\mathbf{F} = \mathbf{I}$) and satisfy (11), that the spreads of the component scores are greater than pre-specified matrices with non-negative elements ($\mathbf{A}^L \geq \mathbf{\Psi}^L$ and $\mathbf{A}^R \geq \mathbf{\Psi}^R$) and taking into account the inclusion constraints between observed and estimated data, with obvious notation, we then get the following minimization problem:

$$
\begin{aligned}
\min\ & J_{3LR} \\
\text{s.t.}\ & \mathbf{M}_a + \mathbf{R}_a * \mathbf{H}_R \leq \mathbf{A}^M \mathbf{F}' + \left(\mathbf{A}^R \left|\mathbf{F}'\right|\right) * \mathbf{H}_R, \\
& \mathbf{M}_a - \mathbf{L}_a * \mathbf{H}_L \geq \mathbf{A}^M \mathbf{F}' - \left(\mathbf{A}^L \left|\mathbf{F}'\right|\right) * \mathbf{H}_L, \\
& \mathbf{F} = \left(\mathbf{C} \otimes \mathbf{B}\right) \mathbf{G}_a', \\
& \mathbf{F}'\mathbf{F} = \mathbf{I}, \\
& \mathbf{A}^L \geq \mathbf{\Psi}^L, \\
& \mathbf{A}^R \geq \mathbf{\Psi}^R.
\end{aligned}
\tag{56}
$$

Again, the orthonormality of $\mathbf{F}$ can be replaced by imposing the orthonormality of $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{G}$. Obviously, in the CANDECOMP/PARAFAC case, one should replace (11) by (13). It is fruitful to observe that the three-way PCA for LR_1 fuzzy data may lead to ambiguous low dimensional plots of the observation units, as it happened for the two-way case. Thus, in order to study the configuration of the observation units in the obtained low dimensional space, one should study the position of the modes (nor of the hyperrectangles).

Proposition 3: There always exists a solution of the NLP problem in (56).

Proof: Given a feasible solution for $\mathbf{F}$ such that the constraints in (56) are satisfied (see Proposition 2), we can find a feasible solution for $\tilde{\mathbf{A}}$ according to the constraints in (56) by taking sufficiently large positive matrices for $\mathbf{A}^L$ and $\mathbf{A}^R$.

q.e.d.

Remark 9: Numbers of extracted components
As for the symmetric three-way case, the optimal numbers of components can be found by performing the classical Tucker3 or CANDECOMP/PARAFAC model on the modes.

Remark 10: Preprocessing
The preprocessing procedure described in Remark 7 can be adopted (obviously scaling both the left and right spreads).

Remark 11: Lower problem
Together with the *Upper* Problem in (56), the *Lower* problem for Tucker3 can

be formulated as

$$
\begin{aligned}
&\max\ J_{3L} \\
&\text{s.t.}\quad \mathbf{M}_a + \mathbf{R}_a * \mathbf{H}_R \geq \mathbf{A}^M \mathbf{F}' + \left(\mathbf{A}^R \left|\mathbf{F}'\right|\right) * \mathbf{H}_R, \\
&\qquad\ \ \mathbf{M}_a - \mathbf{L}_a * \mathbf{H}_L \leq \mathbf{A}^M \mathbf{F}' - \left(\mathbf{A}^L \left|\mathbf{F}'\right|\right) * \mathbf{H}_L, \\
&\qquad\ \ \mathbf{F} = \left(\mathbf{C} \otimes \mathbf{B}\right) \mathbf{G}_a', \\
&\qquad\ \ \mathbf{F}'\mathbf{F} = \mathbf{I}, \\
&\qquad\ \ \mathbf{A}^L \geq \mathbf{\Psi}_L, \\
&\qquad\ \ \mathbf{A}^R \geq \mathbf{\Psi}_R.
\end{aligned}
\tag{57}
$$

4.4 The Three-way PCA model for LR_2 *fuzzy data*

To conclude, we also propose the three-way possibilistic component model for LR_2 fuzzy data, which is able to synthesize the three-way fuzzy data matrix $\tilde{\mathbf{X}} = (\mathbf{M}_1, \mathbf{M}_2, \mathbf{L}, \mathbf{R})_{LR}$. Thus, each observation unit is now characterized by JK non-interactive LR_2 fuzzy numbers (J fuzzy variables collected at K occasions).
As for the previous models, we have (performing the Tucker3 model)

$$
\tilde{\mathbf{X}}_a = \tilde{\mathbf{A}} \mathbf{G}_a \left(\mathbf{C}' \otimes \mathbf{B}'\right) = \tilde{\mathbf{A}} \mathbf{F}', \tag{58}
$$

but, of course, $\tilde{\mathbf{A}}$ is the component score matrix with LR_2 fuzzy numbers. Under the same assumptions and constraints utilized in the previous sections, we get the following problem:

$$
\begin{aligned}
&\min\ J_{3LR} \\
&\text{s.t.}\quad \mathbf{M}_{2a} + \mathbf{R}_a * \mathbf{H}_R \leq \mathbf{A}^{M_2} \mathbf{F}' + \left(\mathbf{A}^R \left|\mathbf{F}'\right|\right) * \mathbf{H}_R, \\
&\qquad\ \ \mathbf{M}_{1a} - \mathbf{L}_a * \mathbf{H}_L \geq \mathbf{A}^{M_1} \mathbf{F}' - \left(\mathbf{A}^L \left|\mathbf{F}'\right|\right) * \mathbf{H}_L, \\
&\qquad\ \ \mathbf{F} = \left(\mathbf{C} \otimes \mathbf{B}\right) \mathbf{G}_a', \\
&\qquad\ \ \mathbf{F}'\mathbf{F} = \mathbf{I}, \\
&\qquad\ \ \mathbf{A}^L \geq \mathbf{\Psi}^L, \\
&\qquad\ \ \mathbf{A}^R \geq \mathbf{\Psi}^R.
\end{aligned}
\tag{59}
$$

Similarly, the CANDECOMP/PARAFAC problem can be easily formulated.

5 Application

The data examined in this application refer to advertising on Internet [2]. This is usually done by means of banners. In particular, we can distinguish three types of banners: 'static' (a single image), 'dynamic' (a dynamic gif image) and 'interactive' (inducing the surfers to participate in polls, games, etc.). The collected data contain a subset of the "Web Advertising" data [2] in which twenty surfers express their judgments on the (static, dynamic and interactive) banners of a set of Web sites during six fortnights. For each combination of Web site, type of banner and fortnight, the median judgement among the twenty surfers is considered. Note that the surfers express their opinion according to five linguistic labels (Worst, Poor, Fair, Good, Best) which are fuzzified as in [2]. The adopted process of fuzzification is described in Table 1.

Table 1
Fuzzification of the linguistic labels (triangular fuzzy numbers)

Linguistic label	Fuzzy number
Worst (W)	$(3, 3, 1)_{LR}$
Poor (P)	$(4, 1.5, 1.5)_{LR}$
Fair (F)	$(6, 1, 0.5)_{LR}$
Good (G)	$(8, 1.75, 0.25)_{LR}$
Best (B)	$(10, 2, 0)_{LR}$

We then get a three-way LR_1 fuzzy data set where the observation units are $I = 9$ Web sites, the variables are $J = 3$ types of banners and the occasions are $K = 6$ fortnights. The data set is given in Table 2.

After running several Tucker3 and CANDECOMP/PARAFAC analyses, we decided to choose the CANDECOMP/PARAFAC model using $S = 2$ components because of its parsimony and, in the current data set, its capability to capture (with $S = 2$ components) the essential information underlying the observed data. We thus performed the three-way possibilistic PCA for LR_1 in (56) considering the appropriate modification for CANDECOMP/PARAFAC. Note that we imposed the columnwinse orthonormality of the component matrices $\mathbf{B}$ and $\mathbf{C}$ from which $\mathbf{F}'\mathbf{F} = \mathbf{I}$ and we set $h = 0.5$.

The optimal component matrices are given in Tables 3, 4 and 5. The extracted components can be interpreted as follows. By observing the component matrix for the variables, we can see that the first component is strictly related to the static and dynamic banners (scores equal to 0.73 and -0.67, respectively), whereas the role of the interactive banner is negligible (-0.12). Instead, the second component mainly reflects the interactive banners (0.99), while negli-

Table 2
Web Advertising Data (subset of $I = 9$ Web sites)

	Fortnight 1			Fortnight 2			Fortnight 3		
Web site	Sta	Dyn	Int	Sta	Dyn	Int	Sta	Dyn	Int
Iol.it	W	G	B	G	P	G	F	G	G
Kataweb.it	F	W	F	P	B	F	G	F	W
Yahoo.it	B	B	B	P	B	G	F	W	B
Altavista.com	B	P	F	F	F	G	G	G	F
Inwind.it	G	P	F	G	F	B	F	B	P
iBazar.it	G	W	G	P	B	P	G	P	P
Repubblica.it	G	P	P	W	F	W	F	F	G
Mediasetonline.it	P	F	G	B	W	F	B	G	F
Yahoo.com	F	G	B	F	F	G	P	B	F
	Fortnight 4			Fortnight 5			Fortnight 6		
Web site	Sta	Dyn	Int	Sta	Dyn	Int	Sta	Dyn	Int
Iol.it	P	W	P	F	B	F	F	F	F
Kataweb.it	F	P	F	P	G	W	G	P	B
Yahoo.it	W	B	G	B	P	G	G	F	F
Altavista.com	B	F	F	B	P	W	F	G	B
Inwind.it	F	B	P	P	F	P	P	F	G
iBazar.it	P	B	B	G	F	B	P	F	F
Repubblica.it	P	B	F	B	W	G	F	F	G
Mediasetonline.it	F	W	G	P	F	P	G	P	P
Yahoo.com	G	F	F	G	G	G	G	F	F

gible scores pertain to the static and dynamic banners. The component scores for the occasions (in Table 5) show that all the six fortnights affect both components. In fact, all the scores are noticeably different from zero (with positive or negative sign). It follows that the Web sites having high first component scores are those with high ratings for the static banners and low ratings for the dynamic banners at fortnights n.2, n.3 and n.4 (component scores for the occasion mode with positive sign) and with low ratings for the static banners and high ratings for the dynamic banners at fortnights n.1, n.5 and n.6 (component scores for the occasion mode with negative sign). In this respect, Iol.it, Mediasetonline.it and Yahoo.it have high first component scores. Conversely, Repubblica.it is characterized by a very low first component score. These re-

Table 3
Component matrix for the observation unit mode

Web site	Component 1	Component 2
Iol.it	$(5.44, 25.88, 15.14)_{LR}$	$(0.88, 8.80, 12.42)_{LR}$
Kataweb.it	$(-0.63, 27.70, 25.09)_{LR}$	$(3.92, 10.87, 5.48)_{LR}$
Yahoo.it	$(2.87, 31.62, 25.79)_{LR}$	$(-2.96, 0.00, 10.14)_{LR}$
Altavista.com	$(-2.98, 14.96, 32.87)_{LR}$	$(4.12, 9.95, 4.37)_{LR}$
Inwind.it	$(-0.93, 28.32, 15.71)_{LR}$	$(4.33, 5.50, 5.38)_{LR}$
iBazar.it	$(-0.01, 26.46, 18.01)_{LR}$	$(-0.99, 12.44, 14.23)_{LR}$
Repubblica.it	$(-4.58, 18.00, 10.63)_{LR}$	$(-1.06, 23.48, 7.35)_{LR}$
Mediasetonline.it	$(2.88, 27.57, 19.68)_{LR}$	$(-2.03, 2.14, 12.42)_{LR}$
Yahoo.com	$(-0.53, 22.49, 18.18)_{LR}$	$(-1.08, 4.01, 9.74)_{LR}$

Table 4
Component matrix for the variable mode

Banner	Component 1	Component 2
Static	0.73	-0.00
Dynamic	-0.67	-0.17
Interactive	-0.12	0.99

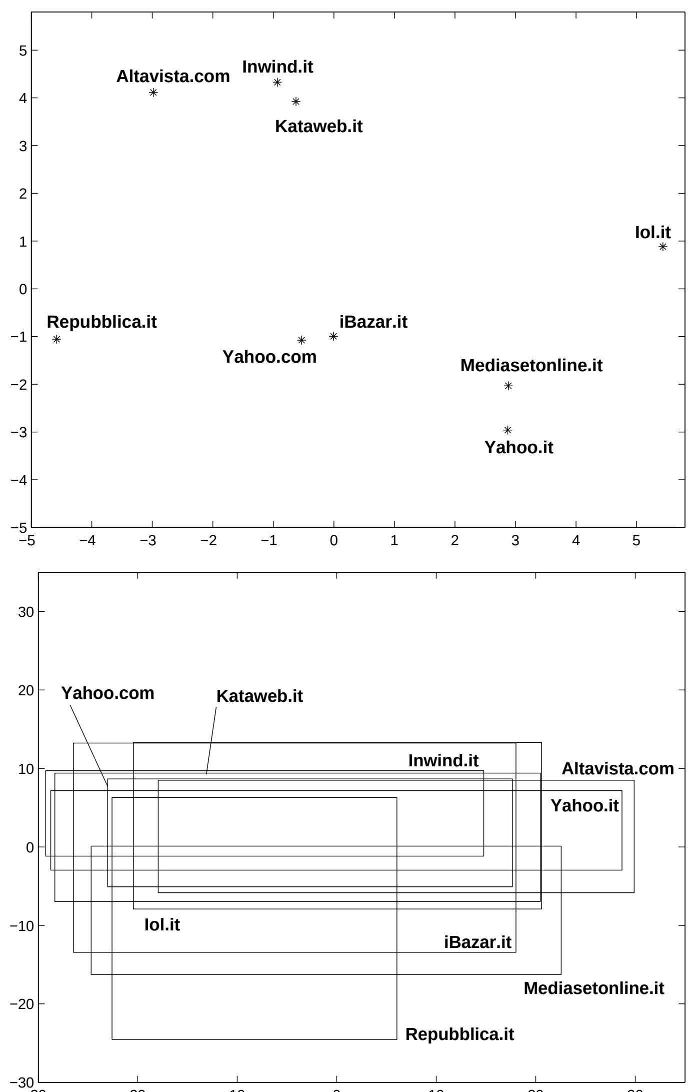
Table 5
Component matrix for the occasion mode

Fortnight	Component 1	Component 2
Fortnight 1	-0.41	-0.35
Fortnight 2	0.44	0.40
Fortnight 3	0.42	-0.43
Fortnight 4	0.44	-0.43
Fortnight 5	-0.40	-0.45
Fortnight 6	-0.33	0.38

sults are consistent to the original data set given in 2. For instance, the score of Iol.it is mainly influenced by the ratings for both the static and dynamic banners at fortnight n.1, for the static banners at fortnight n.2 and for the dynamic banners at fortnights n.4 and n.6. Instead, the score pertaining to Repubblica.it can be explained observing the ratings for the static banners at fortnight n.2, for the dynamic banners at fortnight n.4, and for both the static and dynamic banners at fortnight n.5.

In a similar way, the second component discriminates the Web sites with respect to the ratings of the interactive banners. In particular, high second component scores for the Web sites can be found in case of high ratings for the interactive banners at fortnights n.2 and n.6 and low ratings at fortnights n.1, n.3, n.4 and n.5. In this case, the Web sites with high second component scores are Altavista.com, Inwind.it and Kataweb.it, whereas the lowest scores pertain to Mediasetonline.it and Yahoo.it.
Starting from the the component score matrix for the Web sites given in Table 3 and taking into account that **F** is columnwise orthonormal, we provide the low dimensional configuration of the Web sites given in Figure 1. In particular,

Fig. 1. Low dimensional representation of the Web sites



on the upper side, we plotted the Web sites as a cloud of points, using the modes of the component scores for the Web sites as coordinates. This gives useful information in order to distinguish the Web sites according to their distance in the obtained low dimensional space. The farther two points are, the more the ratings of the two Web sites involved are different. Moreover, on the lower side of Figure 1, we plotted the Web sites as a cloud of low dimensional hyperrectangles (rectangles since $S = 2$) using $\mathbf{A}^M - \mathbf{A}^L$ and $\mathbf{A}^M + \mathbf{A}^R$ as coordinates. Such a plot is helpful in order to evaluate the uncertainty associated to every Web site. We can observe that, except for Repubblica.it, the first component scores are more uncertain than the second ones. This can be explained by taking into account the scores in matrix $\mathbf{B}$. Specifically, the first component mainly depends on the ratings (and the associated fuzziness) for two types of banners, while the second component on the ratings (and the associated fuzziness) for only one type. The sizes of the rectangles show which are the Web sites with (dis)-similar ratings. In particular, the bigger (smaller) a rectangle is, the more (less) the ratings vary among the types of banners and the fortnights. In this respect, Yahoo.com is represented by one of the smallest rectangles (in particular, by the rectangle with the smallest perimeter). By inspecting the original data pertaining to Yahoo.com, it is interesting to observe that 15 times (out of 18), the ratings are 'Good' or 'Fair'.

6 Conclusion

In this work, following a possibilistic approach, diverse component models for handling two- and three-way LR fuzzy data have been proposed. The two-way models are based on suitable extensions of classical PCA, whereas the three-way ones exploit the potentiality of the widely used Tucker3 and CANDECONP/PARAFAC models. The need for three-way methods arises in the attempt of comprehending the triple interactions among the modes that characterize the (three-way) data at hand.
In a fuzzy framework, the generalizations of the previous tools are attained by developing regression problems between a set of observed fuzzy variables (the data at hand) and a set of unobserved crisp variables (the components). We then get multivariate possibilistic regression analyses in which the observed data are the dependent variables and the components are the independent ones. The optimal component matrices can be obtained by solving classical nonlinear programming problems.
A nice property of two- and three-way component models is that they yield a low-dimensional representation of the observed data. In case of fuzzy data, this consists of plotting the observed hyperrectangles (associated to the observation units) as low-dimensional hyperrectangles in the given subspace. The hyperrectangles provide a double source of information: the location of the modes gives a configuration of the observation units onto the obtained subspace, whereas the spreads (the widths of the hyperrectangles) are a measure of the uncertainty associated to each observation.

The application to three-way fuzzy data has shown that the model works well in recovering the underlying structure of the three-way data involved. In the future, it will be interesting to develop three-way methods for (three-way) fuzzy data according to a least-squares point of view. This will be the main research issue in the near future.

References

[1] J.D. Carroll, J.J., Chang, Analysis of individual differences in multidimensional scaling via an n-way generalization of 'Eckart-Young' decomposition, Psychometrika 35 (1970) 283–319. P.

[2] R., Coppi, P., D'Urso, Three-way fuzzy clustering models for LR fuzzy time trajectories, Computat. Statist. Data Anal. 43 (2003) 149–177.

[3] R., Coppi, P. D'Urso, P. Giordani, Component models for fuzzy data, Technical report n.1, Department of Statistics, Probability and Applied Statistics, University of Rome "La Sapienza" (2004) submitted for publication.

[4] T. Denœux, M. Masson, Principal component analysis of fuzzy data using autoassociative neural networks, IEEE Trans. Fuzzy Systems 12 (2004) 336–349.

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

[6] R.A., Harshman Foundations of the PARAFAC procedure: models and conditions for an 'exploratory' multimode factor analysis, UCLA Working Papers in Phonetics 16 (1970) 1–84.

[7] E. Hisdal, Conditional possibilities, independence and noninteraction, Fuzzy Sets and Systems 1 (1978) 283–297.

[8] P. Giordani, H.A.L. Kiers, Principal component analysis of symmetric fuzzy data, Computat. Statist. Data Anal. 45 (2004) 519–548.

[9] H.A.L., Kiers, Towards a standardized notation and terminology in multiway analysis, J. Chem. 14 (2000) 105–122.

[10] H.A.L., Kiers, I., Van Mechelen, Three-way component analysis: principles and illustrative application, Psychological Methods, 6 (2001) 84–110.

[11] P.M., Kroonenberg, Three-mode principal component analysis: theory and applications, DSWO Press, Leiden, 1983.

[12] L.R., Tucker, Some mathematical notes on three-mode factor analysis, Psychometrika 31 (1966) 279–311.

[13] J. Watada, Y. Yabuuchi, Fuzzy principal component analysis and its application, Biomed. Fuzzy Hum. Sci. 3 (1997) 83-92.