Soft Comput (2016) 20:749–762
DOI 10.1007/s00500-014-1538-8

METHODOLOGIES AND APPLICATION

A dimension reduction technique for two-mode non-convex fuzzy data

A. Calcagnì · L. Lombardi · E. Pascali

Published online: 16 December 2014

Abstract Fuzzy modeling and fuzzy statistics provide useful tools for handling empirical situations affected by vagueness and imprecision in the data. Several fuzzy statistical models and methods (e.g., fuzzy regression, fuzzy principal component analysis, fuzzy clustering) have been developed over the years. Generally the standard LR-fuzzy data representation has been used in these methods. However, several empirical contexts, such as human ratings and decision making, may show more complex fuzzy structures which cannot be successfully modeled by the LR representation. In all these cases another type of fuzzy data representation, the so-called LHIR representation, should be preferred instead. In particular, this novel representation allows to handle with fuzzy data which are characterized by non-convex membership functions. In this paper, we address the problem of summarizing large datasets characterized by two-mode non-convex fuzzy data. We introduce a novel dimension reduction technique (NCFCA) based on the framework of Component Analysis and Least squares programming. Finally, to better highlight some important characteristics of the proposed model, we apply NCFCA to three empirical datasets concerning behavioral and socio-economic issues.

Keywords Non-convex fuzzy data · Fuzzy component analysis · Fuzzy statistics · Data Mining · Fuzzy rating scales

Communicated by V. Loia.

Electronic supplementary material The online version of this article (doi:10.1007/s00500-014-1538-8) contains supplementary material, which is available to authorized users.

A. Calcagnì (✉) · L. Lombardi
Department of Psychology and Cognitive Science,
University of Trento, 38068 Rovereto, TN, Italy
e-mail: antonio.calcagni@unitn.it

L. Lombardi
e-mail: luigi.lombardi@unitn.it

E. Pascali
Department of Mathematics and Physics 'E. De Giorgi',
University of Salento, 73100 Lecce, Italy
e-mail: eduardo.pascali@unisalento.it

1 Introduction

In many research fields such as, for example, behavioral and social sciences, epidemiology, bioinformatics, engineering, and astronomy, researchers often have to deal with high-dimensional datasets which are usually represented by n (units) $\times$ m (variables) matrices. Country statistical profiles, socio-economic tables, chemical databases, survival tables, and self-report questionnaires, are all examples of this type of data structures (Eriksson 2006). In these contexts, it may be useful to reduce the dimensionality/complexity of large datasets. This may happen, for instance, when a researcher wants to enhance the efficiency and accuracy of a data analysis, or when s/he wants to extract the most relevant information from the available data. In all these cases, several *dimension reduction techniques* such as, for instance, Principal Component Analysis, Independent Component Analysis, Multidimensional Scaling, Cluster Analysis, and Latent Semantic Analysis, are available to perform statistical analysis on large data structures (Hastie et al. 2001). Among these options, Principal Component Analysis (PCA) is a well-known and widely used unsupervised variables transformation technique for linear dimensionality reduction. Its aim is to summarize a $n \times m$ data matrix into a new $n \times p$ reduced model matrix (with $p \ll m$) which reconstructs the information contained in the original data (Abdi and Williams 2010). Usually, PCA can be performed using different mathematical procedures like, eigenvalues decomposition, singular

values decomposition, low-rank approximations, and component analysis.

PCA has been mainly applied to standard crisp data. However, some researchers have extended the PCA framework also to more complex data (e.g., interval, symbolic, or fuzzy) to better model variables with vague and imprecise information (Lauro and Palumbo 2000; Douzal-Chouakria et al. 2011; Taheri 2003; Viertl 2011). A natural way to model imprecision and vagueness in empirical data is by means of the so-called fuzzy sets (Zimmermann 2001).

Conventionally, fuzzy sets have been described by *LR-type representations* (Dubois et al. 1988) which is primarily used for modeling convex-shaped fuzzy objects. However, in some empirical contexts like, human decision making and ratings, convex representations might not be capable to capture more complex structures in the data. A case of particular interest in such situations concerns some features of human judgments that are characterized by high levels of uncertainty in individuals' responses or evaluations. In solving decision-making problems, some individuals may largely hesitate in providing their responses. For example, in decision-making tasks it is not rare to observe response patterns in which uncertainty is related to the individual's choice between two possible alternatives (relevant options) among a larger set of alternatives (irrelevant options) (e.g., Greene and Haidt 2002; Magnuson 2005; Weber and Johnson 2009). Clearly, in this context convex representations are inappropriate to describe this type of information. Moreover, non-convexity seems to arise as a natural property in many applications based on fuzzy systems, such as fuzzy decision making and expert systems (Calcagnì et al. 2014; Garibaldi et al. 2004; Facchinetti and Pacchiarotti 2006; Reuter 2008). In this framework, the use of standard LR-type representations could be questionable. A possible way out consists in adopting ad-hoc data manipulation procedures to transform non-convex data into standard convex representations (e.g., using Graham Scan algorithm or Steiner symmetrization). However, one serious limitation of these data transformation procedures is that they can artificially mask relevant information carried out by the non-convexity property. Unfortunately, reduction dimension techniques for analyzing non-convex fuzzy data, as far as we know, have not been proposed yet in the literature. In this article we present a novel dimension reduction technique, called *non-convex fuzzy component analysis* (NCFCA), which is based on the frameworks of Component Analysis, CA (Meredith and Millsap 1985; Millsap and Meredith 1988) and standard least squares (LS) (Diamond 1988; Giordani and Kiers 2004). Unlike other fuzzy modeling procedure, NCFCA always guarantees a direct modeling of multidimensional fuzzy data with non-convex shapes based on 2-mode representations.

The reminder of the article is organized as follows. The second section is devoted to briefly recall the basic characteristics of convex as well as non-convex fuzzy data. The third section exposes the component analysis for non-convex fuzzy data together with its main features. Moreover, this section also describes some useful procedures for data fitting and model evaluation. The fourth section illustrates three applications of the proposed method to some behavioral and socio-economic data collected using different procedures (e.g., fuzzy scales of measurement and fuzzy measurement systems). Finally, the fifth section concludes this article providing some final remarks and suggestions for future extensions of our approach.

2 Non-convex fuzzy data

Before introducing the formal representation of non-convex fuzzy data, we briefly recall some basic features of the LR representation. In general, a fuzzy set $\tilde{A}$ can be described by its α-sets $\tilde{A}_\alpha = \{x \in U \mid \mu_{\tilde{A}}(x) > \alpha\}$ with $\alpha \in \ ]0, 1]$ and where U and $\mu_{\tilde{A}}$ indicate the universal set and the membership function of $\tilde{A}$, respectively. If the α-sets of $\tilde{A}$ are all convex sets, then $\tilde{A}$ is called a *convex fuzzy set*. The *support* of $\tilde{A}$ is denoted by $\tilde{A}_0 = \{x \in U \mid \mu_{\tilde{A}}(x) > 0\}$ whereas the collection $\tilde{A}_g = \{x \in U \mid \mu_{\tilde{A}}(x) = \max_{y \in U} \mu_{\tilde{A}}(y)\}$ of all its maximal points is called the *core* of $\tilde{A}$. The *height* of $\tilde{A}$ is defined as $hgt(\tilde{A}) = \max[\mu_{\tilde{A}}(x)]$, and if $hgt(\tilde{A}) = 1$, then $\tilde{A}$ is also called a *normal* fuzzy set. Now, if $\tilde{A}$ satisfies the conditions of normality, convexity, and unimodality (the core is a singleton), then $\tilde{A}$ is named *LR-fuzzy number* (Hanss 2005) and can be denoted by $\tilde{a}$. Moreover, the membership function of $\tilde{a}$ can be described by a couple of monotonic decreasing and left-continuous smooth functions L and R. Using these functions one can represent $\tilde{a}$ by a specific parametric representation. Note that different types of LR-fuzzy numbers can be defined, such as triangular fuzzy numbers, trapezoidal fuzzy numbers, and gaussian fuzzy numbers. In particular, for a trapezoidal fuzzy number the parametric representation is denoted as $\tilde{a} = (m_1, m_2, l, r)_{LR}$. The LR-tuple conveys the main information about the fuzzy set, namely its precision (by means of its core or modal values m_1 and m_2) and fuzziness (by means of l and r). More details about the formal properties of L and R together with other features about the LR representation can be found in Dubois et al. (1988). Figure 1a shows a graphical representation for a LR-fuzzy number with a trapezoidal shape. Unlike the convex case, the α-sets of a non-convex fuzzy set represent set-theoretical unions of at least two compact disjoint intervals. In general, we may observe different levels of non-convexity which can characterize the structure of a fuzzy set (e.g., fuzzy sets with two or more modes). However, in this contribution we will limit our attention to a simple but important type of non-convex fuzzy set called *2-mode fuzzy number* (Fig. 1b). More formally, let $\tilde{B}$ be a fuzzy set which obeys to (1) *normality* (at

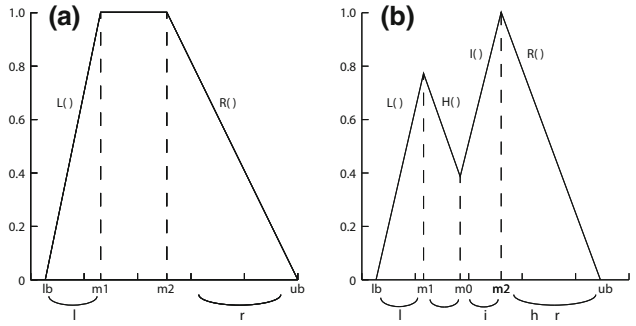

Fig. 1 **a** Convex trapezoidal LR fuzzy data and **b** non-convex 2-mode fuzzy data (LHIR representation)

least one of the points of its support has full membership value), (2) *non-convexity*, and (3) *bimodality* (it has no more than two modes), then the fuzzy number $\tilde{b}$ is called a 2-mode fuzzy number (Calcagnì et al. 2014). Note that, condition (3) restricts the non-convexity property to its simplest representation which corresponds to α-sets obtained by taking the union of at maximum two disjoint compact intervals. Like for the standard convex representation, also for the LHIR representation, non-convex fuzzy data can be described using a parametric format. To this purpose, let us consider four monotonic decreasing and left-continuous functions:

$$L:\mathbb{R}^+ \to [0,1] \quad H:\mathbb{R}^+ \to [0,1]$$
$$I:\mathbb{R}^+ \to [0,1] \quad R:\mathbb{R}^+ \to [0,1]$$

with the following properties:

$$L(\upsilon)\begin{cases} =0 & \text{if } \upsilon = 1 \\ =t_1 & \text{if } \upsilon = 0 \\ >0 & \text{if } \upsilon < 1 \\ <t_1 & \text{if } \upsilon > 0 \end{cases} \qquad H(\upsilon)\begin{cases} =t_1 & \text{if } \upsilon = 0 \\ =z & \text{if } \upsilon = 1 \\ >z & \text{if } \upsilon > 0 \\ <t_1 & \text{if } \upsilon < 1 \end{cases}$$

with: $\upsilon \in \mathbb{R}^+ \quad t_1 \in [0,1] \quad \text{and} \quad z < t_1$

$$I(\upsilon)\begin{cases} =z & \text{if } \upsilon = 1 \\ =t_2 & \text{if } \upsilon = 0 \\ >z & \text{if } \upsilon > 0 \\ <t_2 & \text{if } \upsilon < 1 \end{cases} \qquad R(\upsilon)\begin{cases} =0 & \text{if } \upsilon = 1 \\ =t_2 & \text{if } \upsilon = 0 \\ >0 & \text{if } \upsilon < 1 \\ <t_2 & \text{if } \upsilon > 0 \end{cases}$$

with: $\upsilon \in \mathbb{R}^+ \quad t_2 \in [0,1]$ and $\quad z < t_2$

Using L, H, I, and R, the membership function of $\tilde{b}$ can be described in a very general way as follows:

$$\mu_{\tilde{b}}(x) = \begin{cases} L\left(\frac{m_1 - x}{l}\right) & \text{if } x < m_1 \\ H\left(\frac{x - m_1}{h}\right) & \text{if } m_1 < x < m_0 \\ I\left(\frac{m_2 - x}{i}\right) & \text{if } m_0 < x < m_2 \\ R\left(\frac{x - m_2}{r}\right) & \text{if } x > m_2 \end{cases}$$

where m_1, m_2, and m_0 are the modal points and the middle point, respectively; l and r are the external spreads; h and i are the internal spreads with $h = (m_0 - m_1)$ and $i = (m_2 - m_0)$, respectively. Therefore, the parametric representation for the 2-mode fuzzy data can be expressed as:

$$\tilde{b} = \{(m_0, m_1, m_2, h, i, l, r); (\mu_{m_1}, \mu_{m_0}, \mu_{m_2})\}_{LHIR}$$

where $\mu_{m_1} = t_1$, $\mu_{m_0} = z$, and $\mu_{m_2} = t_2$ are the membership values for m_1, m_0, and m_2, respectively. In addition to conditions (1–3), the following structural properties guarantee the correct representation for a 2-mode fuzzy number: (4) $m_1 < m_0 < m_2$, (5) $l > 0$, (6) $r > 0$, (7) $t_1 < z < t_2$. Note that, the definitions of L, H, I, and R allow to consider different shapes for the 2-mode non-convex representation. For the simplest case, if L, H, I, and R are chosen to be linear functions:

$$L(\upsilon) = \max\{0, \quad (1-\upsilon)t_1\}$$
$$H(\upsilon) = \max\{0, \quad t_1 - \upsilon(t_1 - z)\}$$
$$I(\upsilon) = \max\{0, \quad t_2 - \upsilon(t_2 - z)\}$$
$$R(\upsilon) = \max\{0, \quad (1-\upsilon)t_2\}$$

we obtain the piecewise-linear 2-mode fuzzy number (Fig. 1b). The definition of 2-mode fuzzy number is flexible enough to capture also the convex cases. More precisely, when condition (7) is not met the 2-mode fuzzy number degenerates into a *trapezoidal fuzzy number* obeying to $m_1 < m_0 < m_2, l > 0, r > 0, t_1 = z = t_2 = 1$. By contrast, when conditions (4) and (7) are not met the 2-mode fuzzy number degenerates into a *triangular fuzzy number* obeying to $m_1 = m_0 = m_2, l > 0, r > 0$. Note that in this latter case the internal shape functions $H(\upsilon)$ and $I(\upsilon)$ do not take part in the model representation.

3 Non-convex fuzzy component analysis (NCFCA)

In this section we provide a detailed description of the NCFCA model. From a least squares viewpoint, the main idea is to reduce the dimensionality of the underlying structure of the non-convex fuzzy data by finding a set of *components* which minimize a specific distance between the empirical data and the fuzzy model data. For the sake of simplicity, in this article we describe a technique which is restricted to deal with piecewise-linear 2-mode representations and/or degenerated triangular and trapezoidal fuzzy data. Although some empirical contexts may require different representations for non-convex fuzzy data (e.g., quadratic 2-mode fuzzy numbers), in this contribution we introduce a dimension reduction technique for the most simple case first. However, the two-mode representation is of relevant interest for modeling data observed in many human decision-making applications (e.g., Greene and Haidt 2002; Weber and Johnson 2009). Moreover, the piecewise-linear two-mode representation still guarantees that the extension of the fuzzy principal compo-

nent analysis framework to the non-convex case still remains at a manageable level of technical complexity.

3.1 Model and data analysis

Let $\mathbf{X}$ be a n (units) $\times$ m (variables) data matrix representing the observed data. The generic element x_{ij} of $\mathbf{X}$ defines an array $x_{ij} = \{m_0, h, i, l, r, \mu_0, \mu_1, \mu_2\}_{ij}$ representing a parametrized fuzzy set. By adopting the parametric representation for 2-mode fuzzy data, the elements of $\mathbf{X}$ can be described by a collection of $n \times m$ matrices, $\mathbf{M_0}$, $\mathbf{H}$, $\mathbf{I}$, $\mathbf{L}$, $\mathbf{R}$, $\mathbf{MU_0}$, $\mathbf{MU_1}$, $\mathbf{MU_2}$, which contain the set of parameters involved in the LHIR representation. Therefore, the component model for 2-mode fuzzy data can be expressed as follows:

$$\begin{cases} \mathbf{M_0} &= \boldsymbol{\Psi}_{M_0}\boldsymbol{\Gamma}^T + \mathbf{E}_{M_0} \\ \mathbf{H} &= \boldsymbol{\Psi}_{H}\boldsymbol{\Gamma}^T + \mathbf{E}_{H} \\ \mathbf{I} &= \boldsymbol{\Psi}_{I}\boldsymbol{\Gamma}^T + \mathbf{E}_{I} \\ \mathbf{L} &= \boldsymbol{\Psi}_{L}\boldsymbol{\Gamma}^T + \mathbf{E}_{L} \\ \mathbf{R} &= \boldsymbol{\Psi}_{R}\boldsymbol{\Gamma}^T + \mathbf{E}_{R} \\ \mathbf{MU_0} &= \boldsymbol{\Psi}_{MU_0}\boldsymbol{\Gamma}^T + \mathbf{E}_{MU_0} \\ \mathbf{MU_1} &= \boldsymbol{\Psi}_{MU_1}\boldsymbol{\Gamma}^T + \mathbf{E}_{MU_1} \\ \mathbf{MU_2} &= \boldsymbol{\Psi}_{MU_2}\boldsymbol{\Gamma}^T + \mathbf{E}_{MU_2} \end{cases} \tag{1}$$

where $\boldsymbol{\Psi}_{M_0}$, $\boldsymbol{\Psi}_{H}$, $\boldsymbol{\Psi}_{I}$, $\boldsymbol{\Psi}_{L}$, $\boldsymbol{\Psi}_{R}$, $\boldsymbol{\Psi}_{MU_0}$, $\boldsymbol{\Psi}_{MU_1}$, and $\boldsymbol{\Psi}_{MU_2}$ denote $n \times p$ matrices of *score components*, $\boldsymbol{\Gamma}$ is an $m \times p$ matrix representing the *component loadings* whereas $\mathbf{E}_{M_0}$, $\mathbf{E}_{H}$, $\mathbf{E}_{I}$, $\mathbf{E}_{L}$, $\mathbf{E}_{R}$, $\mathbf{E}_{MU_0}$, $\mathbf{E}_{MU_1}$ and $\mathbf{E}_{MU_2}$ are $n \times m$ matrices of *residual terms*. In general, the decomposition $\boldsymbol{\Psi}_X\boldsymbol{\Gamma}^T$ yields the best p-rank approximation for the original matrix $\mathbf{X}$. The loading matrix $\boldsymbol{\Gamma}$ contains the coefficients which relate the original variables to the new components. From an algebraical point of view, $\boldsymbol{\Gamma}^T$ represents the basis of the subspaces $\mathbb{R}^p$ on which each fuzzy observation is projected. Moreover, $\boldsymbol{\Psi}_{M_0}, \boldsymbol{\Psi}_{H}, \boldsymbol{\Psi}_{I}, \boldsymbol{\Psi}_{L}, \boldsymbol{\Psi}_{R}, \boldsymbol{\Psi}_{MU_0}, \boldsymbol{\Psi}_{MU_1}$, and $\boldsymbol{\Psi}_{MU_2}$ are the matrices containing the coordinates of such projections. Note that in our model representation, the internal and external spreads together with the matrices for the membership values have all the same underlying component structure $\boldsymbol{\Gamma}$ (Millsap and Meredith 1988). In general, $\boldsymbol{\Gamma}$ can be understood as an intermediate representation among the midpoints in $\mathbf{M_0}$, the membership values in $\mathbf{MU_0}$, $\mathbf{MU_1}$, $\mathbf{MU_2}$, and the related left ($\mathbf{H}$, $\mathbf{L}$) and right ($\mathbf{I}$, $\mathbf{R}$) spreads, respectively. This should offer a good compromise between model flexibility and model simplicity for capturing the underlying structure of the data.

3.2 Parameters estimation

In NCFCA, each empirical observation can be considered as an object represented by an m-dimensional polytope in $\mathbb{R}^m$. By considering the main vertices and hedges of this object, its corresponding support is an interval in $\mathbb{R}$ (for $m = 1$), a rectangle in $\mathbb{R}^2$ (for $m = 2$), and a hyper-rectangle in $\mathbb{R}^m$ (for $m \geq 2$). Figure 2 shows an example of a generic fuzzy object in $\mathbb{R}^2$. It is interesting to note that the support of the fuzzy object is obtained by taking the union of two sub-rectangles representing the projection of the internal and external left and right spreads, respectively. Similarly, for both the fuzzy sets in the graphical representation, the projection of the modal values m_0 indicates the upper and lower bounds of these sub-rectangles. Finally, the external rectangle is obtained by joining the lower and upper bounds of the left and right sub-rectangles, respectively. Using this formal framework, the parameters estimation is obtained by minimizing a suitable loss function between the observed data and the model data. To this end, several measures for fuzzy data can be adopted (Bloch 1999). In our proposal, we resort to use a dissimilarity function based on the least squares criterion (Jahanshahloo et al. 2006; Yang and Ko 1996):

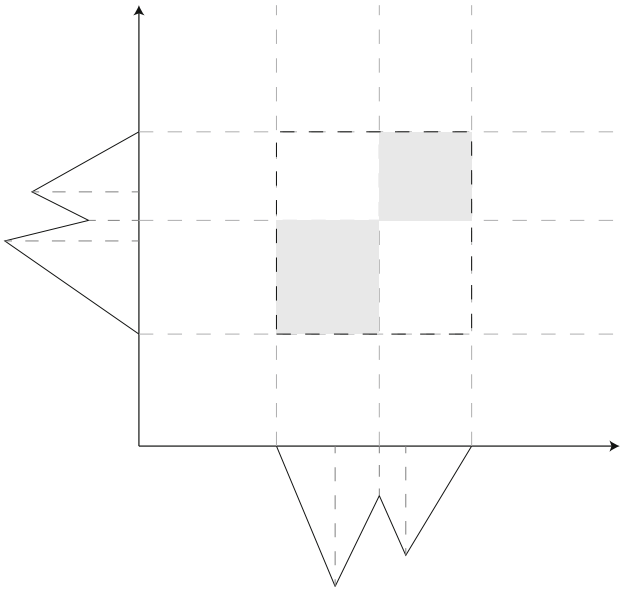

Fig. 2 Example of a fuzzy object in $\mathbb{R}^2$

$$\begin{aligned} \mathcal{D}^2 = &\sum_{k=1}^{2^m} \left\| (\mathbf{M_0} - \mathbf{M_0^*})\boldsymbol{\Phi}_k^L \right\|^2 \\ &+ \sum_{k=1}^{2^m} \left\| [(\mathbf{H} - \mathbf{H^*}) + (\mathbf{L} - \mathbf{L^*})]\boldsymbol{\Phi}_k^L + [(\mathbf{I} + \mathbf{I^*}) + (\mathbf{R} - \mathbf{R^*})]\boldsymbol{\Phi}_k^R \right\|^2 \\ &+ \sum_{k=1}^{2^m} \left\| (\mathbf{MU_1} - \mathbf{MU_1^*})\boldsymbol{\Phi}_k^L + (\mathbf{MU_2} - \mathbf{MU_2^*})\boldsymbol{\Phi}_k^R \right\|^2 \\ &+ \sum_{k=1}^{2^m} \left\| (\mathbf{MU_0} - \mathbf{MU_0^*})\boldsymbol{\Phi}_k^L \right\|^2. \end{aligned} \tag{2}$$

where $\mathbf{M_0^*} = \boldsymbol{\Psi}_{M_0}\boldsymbol{\Gamma}^T$, $\mathbf{H^*} = \boldsymbol{\Psi}_{H}\boldsymbol{\Gamma}^T$, $\mathbf{I^*} = \boldsymbol{\Psi}_{I}\boldsymbol{\Gamma}^T$, $\mathbf{L^*} = \boldsymbol{\Psi}_{L}\boldsymbol{\Gamma}^T$, $\mathbf{R^*} = \boldsymbol{\Psi}_{R}\boldsymbol{\Gamma}^T$, $\mathbf{MU_0^*} = \boldsymbol{\Psi}_{MU_0}\boldsymbol{\Gamma}^T$, $\mathbf{MU_1^*} = \boldsymbol{\Psi}_{MU_1}\boldsymbol{\Gamma}^T$, $\mathbf{MU_2^*} = \boldsymbol{\Psi}_{MU_2}\boldsymbol{\Gamma}^T$. Note that in the above function, $\boldsymbol{\Phi}_k^L$ and $\boldsymbol{\Phi}_k^R$ are $m \times m$ diagonal matrices which allow

to separately consider each distinct main vertex/hedge of the m-dimensional polytope. More precisely, the diagonals are equal to the rows of the Boolean structural matrices $\mathbf{\Phi}^L$ and $\mathbf{\Phi}^R$ of order $2^m \times m$ which are defined according to the following properties:

(a) $\mathbf{\Phi}_t^{L/R} + \mathbf{\Phi}_{2^{(m-1)}+t}^{L/R} = \mathbf{0}_m \;\; (t = 1, \ldots, 2^{m-1})$
(b) $\mathbf{\Phi}_k^L \cdot \mathbf{\Phi}_k^R = \mathbf{0}_{m\times m}$
(c) $\mathbf{\Phi}_k^{L/R} \cdot \mathbf{\Phi}_k^{R/L} = \mathbf{\Phi}_k^{L/R}$
(d) $\mathbf{\Phi}_k^L + \mathbf{\Phi}_k^R = \mathbf{I}_{m\times m}$
(e) $\sum_{k=1}^{2^m} \mathrm{Tr}(\mathbf{X}\mathbf{\Phi}_k^{L/R}) = 2^{m-1}\,\mathrm{Tr}(\mathbf{X})$.

For instance, when $m = 2$ these matrices take the following form:

$$\mathbf{\Phi}^L = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{\Phi}^R = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}$$

Therefore, to represent the lower bounds of the support of the m-dimensional object, we can set $\mathbf{LB} = \mathbf{M_0} - \mathbf{H}\mathbf{\Phi}_1^L - \mathbf{L}\mathbf{\Phi}_1^L + \mathbf{I}\mathbf{\Phi}_1^R + \mathbf{R}\mathbf{\Phi}_1^R$ which, in turn, is equivalent to write $\mathbf{LB} = \mathbf{M_0} - \mathbf{H}\mathbf{\Phi}_1^L - \mathbf{L}\mathbf{\Phi}_1^L$ after noticing that $\mathbf{I}\mathbf{\Phi}_1^R = \mathbf{0}$ and $\mathbf{R}\mathbf{\Phi}_1^R = \mathbf{0}$. The dissimilarity measure in (2) can also be simplified by expanding the kth term of the norms by the properties (b), (c) and (e):

$$\begin{aligned}
\mathcal{D}^2 = {} & 2^{m-1} \left\| (\mathbf{M_0} - \mathbf{\Psi}_{M_0}\mathbf{\Gamma}^T) \right\|^2 + 2^{m-1} \left\| (\mathbf{H} - \mathbf{\Psi}_H\mathbf{\Gamma}^T) \right\|^2 \\
& + 2^{m-1} \left\| (\mathbf{I} - \mathbf{\Psi}_I\mathbf{\Gamma}^T) \right\|^2 + 2^{m-1} \left\| (\mathbf{L} - \mathbf{\Psi}_L\mathbf{\Gamma}^T) \right\|^2 \\
& + 2^{m-1} \left\| (\mathbf{R} - \mathbf{\Psi}_R\mathbf{\Gamma}^T) \right\|^2 \\
& + 2^{m-1} \left\| (\mathbf{MU_0} - \mathbf{\Psi}_{MU_0}\mathbf{\Gamma}^T) \right\|^2 \\
& + 2^{m-1} \left\| (\mathbf{MU_1} - \mathbf{\Psi}_{MU_1}\mathbf{\Gamma}^T) \right\|^2 \\
& + 2^{m-1} \left\| (\mathbf{MU_2} - \mathbf{\Psi}_{MU_2}\mathbf{\Gamma}^T) \right\|^2 \\
& + 2^m \;\; \mathrm{Tr}[(\mathbf{H} - \mathbf{\Psi}_H\mathbf{\Gamma}^T)^T(\mathbf{L} - \mathbf{\Psi}_L\mathbf{\Gamma}^T)] + \\
& + 2^m \;\; \mathrm{Tr}[(\mathbf{I} - \mathbf{\Psi}_I\mathbf{\Gamma}^T)^T(\mathbf{R} - \mathbf{\Psi}_R\mathbf{\Gamma}^T)] \qquad (3)
\end{aligned}$$

where the structural matrices are simply replaced by appropriate weights. We use the Alternating Least Squares algorithm—ALS (Kiers and ten Berge 1989; Kiers 2002) to estimate the parameters contained in $\mathbf{\Psi}_{M_0}$, $\mathbf{\Psi}_H$, $\mathbf{\Psi}_I$, $\mathbf{\Psi}_L$, $\mathbf{\Psi}_R$, $\mathbf{\Psi}_{MU_0}$, $\mathbf{\Psi}_{MU_1}$, $\mathbf{\Psi}_{MU_2}$, $\mathbf{\Gamma}$. In particular, the final ALS solutions for the model (1) are

$$\begin{aligned}
\mathrm{vec}(\widehat{\mathbf{\Gamma}}) = {} & [\,(\mathbf{\Psi}_H^T\mathbf{\Psi}_H \otimes \mathbf{I}_{m\times m} + \mathbf{\Psi}_L^T\mathbf{\Psi}_L \otimes \mathbf{I}_{m\times m} \\
& + \mathbf{\Psi}_I^T\mathbf{\Psi}_I \otimes \mathbf{I}_{m\times m} + \mathbf{\Psi}_R^T\mathbf{\Psi}_R \otimes \mathbf{I}_{m\times m}) \\
& + (\mathbf{\Psi}_{M_0}^T\mathbf{\Psi}_{M_0} \otimes \mathbf{I}_{m\times m} + \mathbf{\Psi}_{MU_0}^T\mathbf{\Psi}_{MU_0} \otimes \mathbf{I}_{m\times m} \\
& + \mathbf{\Psi}_{MU_1}^T\mathbf{\Psi}_{MU_1} \otimes \mathbf{I}_{m\times m} + \mathbf{\Psi}_{MU_2}^T\mathbf{\Psi}_{MU_2} \otimes \mathbf{I}_{m\times m}) \\
& + 2(\mathbf{\Psi}_H^T\mathbf{\Psi}_L \otimes \mathbf{I}_{m\times m} + \mathbf{\Psi}_I^T\mathbf{\Psi}_R \otimes \mathbf{I}_{m\times m})\,]^{-1} \\
& \times \mathrm{vec}[\,(\mathbf{H}^T\mathbf{\Psi}_H + \mathbf{L}^T\mathbf{\Psi}_L + \mathbf{I}^T\mathbf{\Psi}_I + \mathbf{R}^T\mathbf{\Psi}_R) \\
& + (\mathbf{M_0}^T\mathbf{\Psi}_{M_0} + \mathbf{MU_0}^T\mathbf{\Psi}_{MU_0} + \mathbf{MU_1}^T\mathbf{\Psi}_{MU_1} \\
& + \mathbf{MU_2}^T\mathbf{\Psi}_{MU_2}) + (\mathbf{H}^T\mathbf{\Psi}_L + \mathbf{L}^T\mathbf{\Psi}_H + \mathbf{I}^T\mathbf{\Psi}_R + \mathbf{R}^T\mathbf{\Psi}_I)]; \qquad (4)
\end{aligned}$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_L) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T\,\mathrm{vec}(\mathbf{H} - \mathbf{\Psi}_H\mathbf{\Gamma}^T + \mathbf{L}); \quad (5)$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_R) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T\,\mathrm{vec}(\mathbf{I} - \mathbf{\Psi}_I\mathbf{\Gamma}^T + \mathbf{R}); \quad (6)$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_I) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T\,\mathrm{vec}(\mathbf{R} - \mathbf{\Psi}_R\mathbf{\Gamma}^T + \mathbf{I}); \quad (7)$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_H) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T\,\mathrm{vec}(\mathbf{L} - \mathbf{\Psi}_L\mathbf{\Gamma}^T + \mathbf{H}); \quad (8)$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_{M_0}) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T\,\mathrm{vec}(\mathbf{M_0}); \quad (9)$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_{MU_0}) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T\,\mathrm{vec}(\mathbf{MU_0}); \quad (10)$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_{MU_1}) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T\,\mathrm{vec}(\mathbf{MU_1}); \quad (11)$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_{MU_2}) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T\,\mathrm{vec}(\mathbf{MU_2}); \quad (12)$$

where vec(.) is the linear operator that converts a $n \times m$ matrix into an $mn \times 1$ vector, $\otimes$ denotes the Kronecker product, and $\mathbf{I}$ is an identity matrix of appropriate order. In fitting the unconstrained NFCA model, we adopted an iterative procedure based on standard stopping criteria and random initialization. However, one potential limitation of such algorithm is that, in some circumstances, it might not yield feasible solutions. In particular, if the model is fitted to empirical data which are largely corrupted by noise, the corresponding estimations might violate the natural constraints of the 2-mode fuzzy numbers (namely: $\mathbf{h}^*_j > \mathbf{0}_n$, $\mathbf{i}^*_j > \mathbf{0}_n$, $\mathbf{l}^*_j > \mathbf{0}_n$, $\mathbf{r}^*_j > \mathbf{0}_n$, $\boldsymbol{\mu}^*_{\mathbf{1}j} < \boldsymbol{\mu}^*_{\mathbf{0}j} < \boldsymbol{\mu}^*_{\mathbf{2}j}$). In these situations, a constrained version of the algorithm based on specific optimization techniques should instead be preferred (Giordani and Kiers 2007). However, a common and easy strategy to deal with eventual infeasible parameter estimates is to apply a post-hoc correction on the estimated parameters (Giordani and Kiers 2004). In particular, the estimates of eventual negative spreads could be set to zero whereas the estimated membership values of the fuzzy data could be row-wise normalized to satisfy their natural constraints.

3.3 Data interpretation and visualization

Once the estimated components are finally obtained, the results can be analyzed by inspecting the loading matrix $\widehat{\mathbf{\Gamma}}$ and/or by displaying the scores in a low-dimensional plot. In particular, the loadings can be understood as linear coefficients which express the magnitude of the relation between the observed variables and the estimated components. By contrast, the scores represent the projections of the fuzzy observation into the subspace spanned by $\widehat{\mathbf{\Gamma}}$. Like for standard PCA (or CA), also for NCFCA, the score plot represents an important visualization procedure that allows to assess the relationship among the projected units (e.g., by

analyzing the patterns of similarity or dissimilarity among the units). In what follows, we describe in more detail both the data pre-treatment technique and the data evaluation procedure adopted in NCFCA modeling.

3.3.1 Data pre-treatment

A common practice in multivariate analysis is to pre-process raw-data to obtain an improved and clean dataset. Two of the most important ways to pre-process raw-data are *centering* and *scaling*. Centering corresponds to a repositioning of the coordinate system such that the center of gravity of the cloud of data points becomes the origin. By contrast, scaling allows to re-distribute the data according to a specific factor (e.g., the standard deviation). In particular, scaling configures the original variables within a unique scale range without changing the original structure of the data. Centering and scaling can be performed for several reasons. For instance, centering may be applied to improve the fit of the model, remove noise from the data, avoid problems in the estimation procedures, etc. Similarly, scaling may be implemented to adjust for scale differences, reduce the inflation of small or big values in the data, improve the interpretation and visualization of the results, etc. (Bro and Smilde 2003; van den Berg et al. 2006). Several methods can be used for centering and scaling (e.g., auto-scaling, range-scaling, pareto-scaling, vast-scaling). In this contribution, we opted for mean-centering and pareto-scaling methods. In general, given a matrix $\mathbf{A}$, the mean-centered matrix $\mathbf{A}^*$ is obtained as $\mathbf{A}^* = \mathbf{A} - \mathbf{1}_{n\times m} \cdot \text{diag}(\bar{\mathbf{a}})$ whereas the pareto-scaling is performed by $\mathbf{A}^* = \mathbf{A} \cdot [\text{diag}(\sqrt{\text{std}(\mathbf{A})})]^{-1}$. Note that, $\bar{\mathbf{a}}$ is the vector containing the column means of $\mathbf{A}$, std(.) indicates the column-wise standard deviation whereas diag(.) is the operator which transforms a vector into a diagonal matrix. In particular, we pre-processed our data matrices according to the following steps: (1) $\mathbf{M_0}$ was simultaneously mean-centered and pareto-scaled by considering its means and standard deviations, (2) $\mathbf{H}$ and $\mathbf{I}$ were pareto-scaled by considering the standard deviations of $\mathbf{M_0}$, (3) $\mathbf{L}$ and $\mathbf{R}$ were pareto-scaled by considering the standard deviations of $\mathbf{M_1}$ and $\mathbf{M_2}$ respectively, (4) $\mathbf{MU_0}$, $\mathbf{MU_1}$, and $\mathbf{MU_2}$ were pareto-scaled by considering the standard deviations of $\mathbf{M_0}$, $\mathbf{M_1}$, and $\mathbf{M_2}$, respectively.

3.3.2 Rotation of $\widehat{\mathbf{\Gamma}}$

Unlike standard PCA (or CA), the NCFCA estimation procedure does not necessarily yield an orthonormal matrix $\widehat{\mathbf{\Gamma}}$. For this reason, a direct interpretation of $\widehat{\mathbf{\Gamma}}$ might be arduous for some datasets. However, by adopting an orthonormalization procedure such as, for example, the *modified Gram-Schmidt* algorithm, one can always define a rotation matrix $\mathbf{\Omega}$ such that $\widehat{\mathbf{\Gamma}}\mathbf{\Omega}$ is column-wise orthonormal (Trefethen and Bau 1997). In particular, the modified Gram–Schmidt algorithm requires to balance the estimated score matrices with the inverse of the transpose of $\mathbf{\Omega}$. For instance, by considering the case of $\mathbf{M_0}$ the balancing is performed as $\widehat{\mathbf{\Psi}}_{M_0}(\mathbf{\Omega}^T)^{-1}$. In addition, to facilitate the interpretation of the component structure, the analysis might also involve a rotation of $\widehat{\mathbf{\Gamma}}\mathbf{\Omega}$. Several techniques can be adopted to this purpose (Kiers 1997). In NCFCA modeling we adopted the well-known *Varimax rotation* which provides a very simple component structure where each original fuzzy variable is associated with a small set of components (Kaiser 1958). The rotation of $\widehat{\mathbf{\Gamma}}$ allows to simplify the interpretation of both numerical and graphical results of the model.

3.3.3 Model evaluation

In this subsection we illustrate some useful procedures to assess the performance and reliability of the NCFCA model.

Goodness of fit. To evaluate the performance of the NCFCA model, we considered the normalized index:

$$R = 1 - (A/B)$$

where:

$$\begin{aligned} A = {} & \left\|\mathbf{M_0} - \widehat{\mathbf{\Psi}}_{M_0}\widehat{\mathbf{\Gamma}}^T\right\|^2 + \left\|\mathbf{H} - \widehat{\mathbf{\Psi}}_{H}\widehat{\mathbf{\Gamma}}^T\right\|^2 + \left\|\mathbf{I} - \widehat{\mathbf{\Psi}}_{I}\widehat{\mathbf{\Gamma}}^T\right\|^2 \\ & + \left\|\mathbf{L} - \widehat{\mathbf{\Psi}}_{L}\widehat{\mathbf{\Gamma}}^T\right\|^2 + \left\|\mathbf{R} - \widehat{\mathbf{\Psi}}_{R}\widehat{\mathbf{\Gamma}}^T\right\|^2 \\ & + \left\|\mathbf{MU_0} - \widehat{\mathbf{\Psi}}_{MU_0}\widehat{\mathbf{\Gamma}}^T\right\|^2 + \left\|\mathbf{MU_1} - \widehat{\mathbf{\Psi}}_{MU_1}\widehat{\mathbf{\Gamma}}^T\right\|^2 \\ & + \left\|\mathbf{MU_2} - \widehat{\mathbf{\Psi}}_{MU_2}\widehat{\mathbf{\Gamma}}^T\right\|^2 + \text{Tr}[(\mathbf{H} - \widehat{\mathbf{\Psi}}_{H}\widehat{\mathbf{\Gamma}}^T)^T(\mathbf{L} - \widehat{\mathbf{\Psi}}_{L}\widehat{\mathbf{\Gamma}}^T)] \\ & + \text{Tr}[(\mathbf{I} - \widehat{\mathbf{\Psi}}_{I}\widehat{\mathbf{\Gamma}}^T)^T(\mathbf{R} - \widehat{\mathbf{\Psi}}_{R}\widehat{\mathbf{\Gamma}}^T)], \\ B = {} & \|\mathbf{M_0}\|^2 + \|\mathbf{H}\|^2 + \|\mathbf{I}\|^2 + \|\mathbf{L}\|^2 + \|\mathbf{R}\|^2 + \|\mathbf{MU_0}\|^2 \\ & + \|\mathbf{MU_1}\|^2 + \|\mathbf{MU_2}\|^2 + \text{Tr}[\mathbf{H}^T\mathbf{L}] + \text{Tr}[\mathbf{I}^T\mathbf{R}]. \end{aligned}$$

In the R index, A indicates the residual sum of squares and B the observed sum of squares. This index takes values in [0, 1] and is directly related to the number p of components extracted by the NCFCA model. In particular, high values for this index indicate that the NCFCA model almost exactly reconstructs the original data matrices whereas low values for R suggest that more components should be extracted to obtain a satisfactory reconstruction of the original data (Giordani 2010; Bro and Smilde 2003). Therefore, the proposed index gives us additional information about the quality of the NCFCA performance in reproducing the original information stored in the data.

Reliability. To asses the accuracy of the NCFCA solutions, we used a non-parametric bootstrap procedure for component analysis (Coppi et al. 2006; Kiers 2004). In particular, in the non-parametric bootstrap Q samples (with $Q \geq 1000$) of size n were row-wise randomly drawn (with replacement) from the original matrices $\mathbf{M_0}$, $\mathbf{H}$, $\mathbf{I}$, $\mathbf{L}$, $\mathbf{R}$, $\mathbf{MU_0}$, $\mathbf{MU_1}$,

and $\mathbf{MU_2}$. For each qth sample, the loading matrix $\widehat{\mathbf{\Gamma}}^q$ was derived by applying the NCFCA procedure on the sample matrices $\mathbf{M_0}^q$, $\mathbf{H}^q$, $\mathbf{I}^q$, $\mathbf{L}^q$, $\mathbf{R}^q$, $\mathbf{MU_0}^q$, $\mathbf{MU_1}^q$, and $\mathbf{MU_2}^q$. To make bootstrap solutions optimally comparable, $\widehat{\mathbf{\Gamma}}^q$ was rotated to match as close as possible the original $\widehat{\mathbf{\Gamma}}$. Such rotation was obtained by finding a rotation matrix $\mathbf{\Omega}^q$ that minimizes the risk $\|\widehat{\mathbf{\Gamma}}^q\mathbf{\Omega}^q - \widehat{\mathbf{\Gamma}}\|^2$ with the following optimal solution $\mathbf{\Omega}^q = (\widehat{\mathbf{\Gamma}}^{qT}\widehat{\mathbf{\Gamma}}^q)^{-1}\widehat{\mathbf{\Gamma}}^{qT}\widehat{\mathbf{\Gamma}}$. These steps were then repeated for Q times. Finally, the ensuing rotate sample matrices $\widehat{\mathbf{\Gamma}}^1, \widehat{\mathbf{\Gamma}}^2, \ldots, \widehat{\mathbf{\Gamma}}^Q$ were used for computing (by resampling) the standard errors or percentile intervals for each estimated parameter in the model. In general, the lower the standard errors, the greater the accuracy of the model.

3.3.4 Score plot

The score plot is the graphical representation of the original n units in the $\mathbb{R}^p$ subspace. Unlike standard PCA (or CA), in the NCFCA framework each statistical unit is represented by a hyper-rectangle in $\mathbb{R}^p$. There are several methods that can be considered for score plotting such as, for example, maximum covering area rectangle—MCAR (Cazes et al. 1997), parallel edge connected shapes—PECS (Irpino et al. 2003) and Polytope Representation (Le-Rademacher and Billard 2012). In this contribution we adopted the MCAR approach which is a simple and fast graphical technique to represent interval, symbolic, or fuzzy data. In particular, in the simple two-dimensional case, MCAR allows to illustrate the statistical units by means of rectangles in $\mathbb{R}^2$ whereas the information associated to the membership functions is usually not represented.[1] More formally, in our context MCAR was applied as follows. Once the loading matrix was column-wise orthonormalized and the score matrices balanced, each non-convex fuzzy datum was described as the union of two rectangles (or hyper-rectangles) referring to the left and right internal and external spreads (see Fig. 2). The vertices representing the lower and upper bounds of the external rectangles were obtained by the score matrices using the following formula:

$$\begin{aligned}\mathbf{\Lambda}_i = {} & \mathbf{\Phi}^L_{2^p\times p}[\mathrm{diag}(\widehat{\mathbf{\Psi}}_{M_0 i}) - \kappa \,|\, \mathrm{diag}(\widehat{\mathbf{\Psi}}_{Hi})\,| - \kappa \,|\, \mathrm{diag}(\widehat{\mathbf{\Psi}}_{Li})|] \\ & + \mathbf{\Phi}^R_{2^p\times p}[\,\mathrm{diag}(\widehat{\mathbf{\Psi}}_{M_0 i}) + \kappa \,|\, \mathrm{diag}(\widehat{\mathbf{\Psi}}_{Ii})\,| \\ & + \kappa |\,\mathrm{diag}(\widehat{\mathbf{\Psi}}_{Ri})|] \end{aligned} \quad (13)$$

whereas the inner left and right rectangles were depicted by considering the midpoints in $\widehat{\mathbf{\Psi}}_{M_0}$ as upper and lower vertices, respectively. Note that $\mathbf{\Phi}^L_{2^p\times p}$ and $\mathbf{\Phi}^R_{2^p\times p}$ are Boolean structural matrices of order $2^p \times p$ having the same structures and properties of those described in Sect. 3.2, whereas $|.|$ indicates the absolute value. Note that in (13) the scalar $\kappa \in\ ]0, 1]$ acts as a resizing factor which allows to reduce the eventual oversize effect of the plotted rectangles.

[1] Note, however, that the membership functions always contribute to the orientation of the axes in $\mathbb{R}^p$ even if they are not directly illustrated in the graphical representation.

Table 1 Example 1: loading matrix $\widehat{\mathbf{\Gamma}}$ with standard errors in parenthesis ($Q = 5{,}000$)

Clinical items/variables	Comp. 1	Comp. 2
(x_1) I am good at controlling negative and positive emotions	**−0.38** (0.03)	0.00 (0.05)
(x_2) I am worried that I will never realize my ambitions	**−0.50** (0.04)	0.10 (0.08)
(x_3) It is important that human relations are based upon trust	−0.01 (0.04)	**−0.98** (0.16)
(x_4) It is important that I am competent in everything I do	−0.34 (0.04)	−0.10 (0.08)
(x_5) I am worried to be bad	**−0.43** (0.04)	−0.01 (0.05)
(x_6) I am worried to lose my close friends	**−0.39** (0.05)	0.09 (0.10)
(x_7) I like to be alone when I am working with a problem	**−0.39** (0.03)	−0.10 (0.04)

Loadings higher than 0.35 (in absolute value) are in boldface type

4 Applications

In this section we describe three applications to illustrate the main features of the NCFCA analysis. All the algorithms developed for these applications are available upon request to the authors.

4.1 Example 1: psychological assessment of worry

In this first example we analyzed a real dataset about the psychology of worry (Stöber and Joormann 2001). In clinical psychology, the assessment of worry is usually characterized by high levels of imprecision and vagueness in the data. The clinical inventory was composed by seven items (see Table 1) and administered to a group of 10 undergraduate students from the University of Trento (Italy). The scores were collected using a computerized interface based on the mouse tracking methodology (Calcagnì and Lombardi 2014; Johnson et al. 2012). In particular, for each of the seven items, participants were told that a pseudo-circular scale with five response levels (strongly disagree, disagree, neither, agree, strongly agree) would be presented on the screen, and that they were asked to choose which of these responses was the most appropriate for the presented item. After participants clicked a start button, a window with the text of the item appeared at the top of the screen. Next the scale with the five levels appeared while the cursor was allocated to the center of the screen. Participants gave their responses by mouse-clicking the chosen level of the scale. Meanwhile,

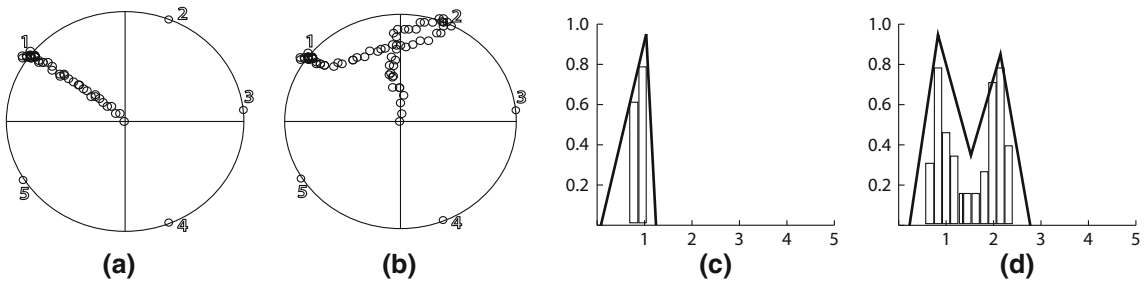

Fig. 3 *Mouse tracker*: empirical patterns of mouse movements (**a**, **b**) and their associated histograms and fuzzy sets (**c**, **d**). *Note* the two patterns **a** and **b** are different but share a same finale response (*1* strongly disagree), the numbers encode the five levels of the scale (*1* strongly disagree, *2* disagree, *3* neither, *4* agree, *5* strongly agree), whereas the histograms were rescaled to provide a better comparison with the fuzzy sets

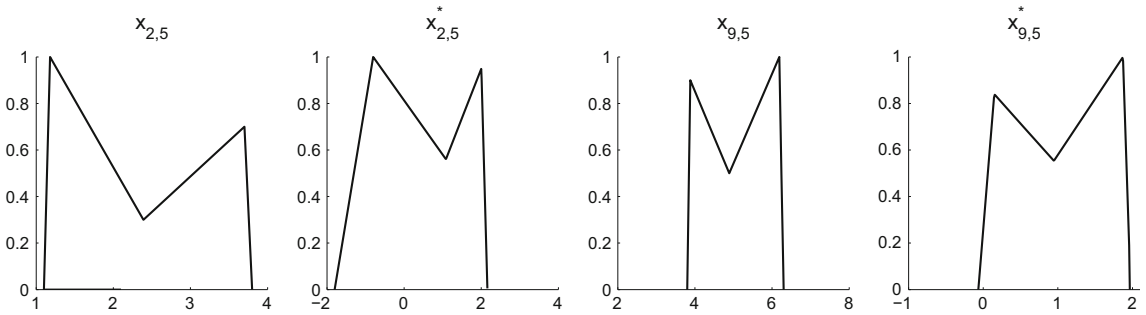

Fig. 4 *Example 1*: observed (x) and model reconstructed (x^*) fuzzy sets for subjects 2 and 9 on item x_5

the system recorded the streaming x–y coordinates of the computer mouse. Figure 3a, b shows two empirical patterns of mouse movements. In particular, Fig. 3a represents an empirical pattern with a low imprecision/fuzziness, by contrast Fig. 3b shows a pattern with a higher level of imprecision and vagueness. Figure 3c, d shows the histograms and the associated fuzzy sets constructed using the radial positions of the x–y mouse movement coordinates of the empirical patterns. In particular, the fuzzy sets were obtained by a heuristic optimization procedure that allowed to convert histograms into fuzzy sets (Ciavolino et al. 2014). The empirical datasets are reported in the supplementary material of this article. Before running the NCFCA analysis, the datasets were first pre-processed according to the procedure described in Sect. 3.3.1, next the NCFCA algorithm was used to extract two main components ($p = 2$ with $R = 0.90$). The algorithm converged after 30 iterations only. Moreover, model accuracy and reliability were also good as indicated by the low standard errors reported in Table 1. Finally, the estimated loading matrix was orthonormalized and varimax-rotated to simplify the interpretation of the components in the NCFCA model. Figure 4 shows an example of some observed (resp. reconstructed) fuzzy sets on the fifth variable (x_5).

To identify the meaning of each extracted component, we selected the variables with loading values larger than ± 0.35 (relevant variables). The results reported in Table 1 showed that the first component depended on x_2 (-0.50) and x_5 (-0.43) and, to a less extent, on x_1 (-0.38), x_6 (-0.39) and x_7 (-0.39), whereas the second component exclusively depended on x_3 (-0.98). Therefore, taking into account the meaning of the significant variables, the first component can be understood as 'individual dimension' whereas the second component can be referred as 'interpersonal dimension' of psychology of worry.

Figure 5 shows the score plot for the NCFCA model in the two-dimensional space spanned by $\widehat{\mathbf{\Gamma}}$. A substantive interpretation of Fig. 5 can be provided by considering both the positions and sizes of the rectangles. In particular, the size of the rectangles reflected the imprecision associated with the clinical items which played a significant role in the definition of the two components. In our case, the fuzzy units were arranged into three main regions: central (units 1, 4, 5, 6, 7, 10), left-outer (units 2, 9), right-outer (units 3, 8). In particular, the individuals in the middle part of the plot showed clinical scores that were in the mean range for both the components. It is important to note that due to the mean-centering procedure, the origin of the axis represents the *average region* and therefore the units located in this area are characterized by *mean profiles*. On the contrary, the units located far away from the center of gravity of the plot show typical features which did not belong to the mean profile.

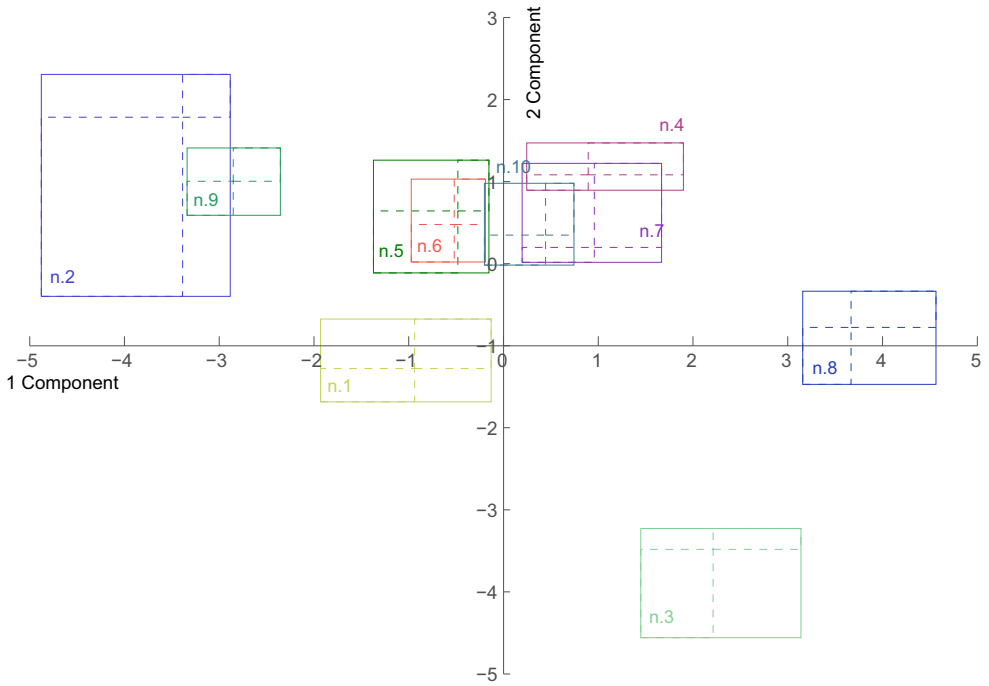

Fig. 5 *Example 1*: score plot for the first and second components (fuzzy units are consecutively numbered and represented with *different colors*) (color figure online)

Table 2 Example 2: loading matrix $\widehat{\mathbf{\Gamma}}$ with standard errors in parenthesis ($Q = 5{,}000$)

Variables	Comp. 1	Comp. 2	Comp. 3
(x_1) Degree of development of Italian university system	**−0.46** (0.07)	0.03 (0.06)	−0.25 (0.08)
(x_2) Level of uselfuness of university programs	0.23 (0.06)	−0.17 (0.07)	**−0.80** (0.10)
(x_3) Level of trust in local government	−0.26 (0.06)	0.17 (0.07)	**−0.53** (0.08)
(x_4) Level of trust in private enterprise	**−0.46** (0.07)	−0.20 (0.06)	0.03 (0.08)
(x_5) Degree of devel. of Italy	**−0.67** (0.06)	0.04 (0.06)	0.09 (0.07)
(x_6) Degree of importance of psychologists	−0.08 (0.06)	**−0.53** (0.05)	−0.05 (0.06)
(x_7) Level of uselfuness of psychologists	0.02 (0.07)	**−0.79** (0.06)	0.08 (0.09)

Loadings higher than 0.35 (in absolute value) are in boldface type

4.2 Example 2: self perception of professional roles

In this second application we studied a real dataset about psychologists' self perception of their professional role. To this end, a personality 7-item questionnaire (see Table 2) was administered to a group of 24 psychology students from the University of Trento (Italy). Data were collected by means of a computerized questionnaire based on fuzzy rating scales commonly used in human rating studies (Hesketh et al. 1988). In particular, the fuzzy scale was based on a pseudo-continuous scale implemented using a suitable graphical interface (see Fig. 6). Interestingly, the fuzzy rating scale may elicit two different scenarios: the respondent chooses a single level of the scale (Fig. 6b) or s/he selects an intermediate position which lies between the two levels (Fig. 6c).

The empirical datasets are reported in the supplementary material of this article. The NCFCA model was applied to the pre-processed data and three components ($p = 3$) were extracted using the NCFCA algorithm ($R = 0.88$). The algorithm converged after 112 iterations. Model accuracy and reliability were also good as shown by the low standard errors reported in Table 2. Like for the previous analysis, also in this second application the estimated loading matrix was orthonormalized and varimax-rotated in the NCFCA model.

The results reported in Table 2 showed that the first component was inversely related to x_1 (−0.46), x_4 (−0.46) and x_5 (−0.67). Similarly, the second component was also inversely

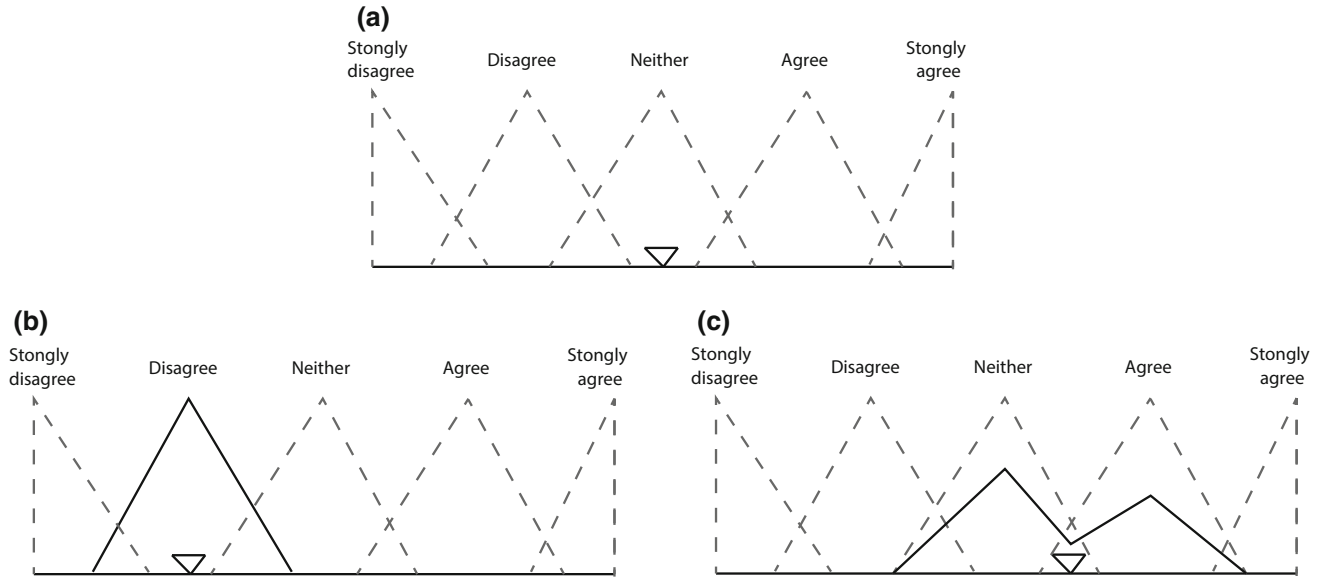

Fig. 6 Fuzzy rating scale with a pseudo-continuous line, a movable cursor (*in solid line*) and a fuzzy variable (*in dotted line*) representing the hidden fuzzy levels of the scale

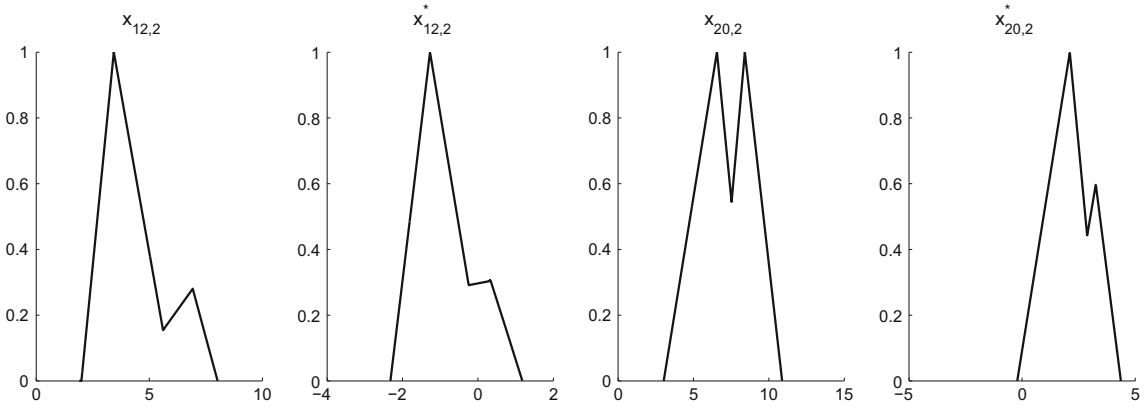

Fig. 7 *Example 2*: observed (x) and model reconstructed (x^*) fuzzy sets for subjects 12 and 20 on item x_2

related to x_6 (-0.53) and x_7 (-0.79). Finally, the third component inversely depended on x_2 (-0.80) and x_3 (-0.53). In line with these results, the first component can be interpreted as 'future perspective dimension', the second component refers to 'psychology as profession', whereas the third component can be understood as 'present dimension'. Figure 7 shows an example of some observed (resp. reconstructed) fuzzy sets on the second variable (x_2). Figures 8 and 9 show the score plots for the first vs. second and second vs. third components, respectively. In particular, the first score plot (Fig. 8) contrasts *future* and *profession*. It shows an interesting pattern in which most of the units were located on the middle part of the plot whereas only small groups of units were located in the left-outer (units 3, 14) and right-outer (units 5, 7, 13, 16) parts of the graphical representation. The second score plot (Fig. 9) contrasts *present* and *profession*. Like for the previous graphical representation, also for the second score plot most of the units are located in the middle portion of the plot. Finally, the sizes of the rectangles were generally small and similar among dimensions (in general, the students seemed to convey the same degree of imprecision in providing their responses).

4.3 Example 3: welfare and productivity of Italian regions

In this last example we tested our model on a real dataset about economic and social indicators collected by the National Institute of Statistics (ISTAT). The original dataset contained 10 socio-economic indicators referred to 20 Italian regions (see Table 3). To test our model the crisp variables were first rescaled to a common scale and next fuzzified with a suitable fuzzification procedure based on the Mamdani fuzzy system (Lalla et al. 2005). The fuzzification routine yielded the following sets for the fuzzy variables: null

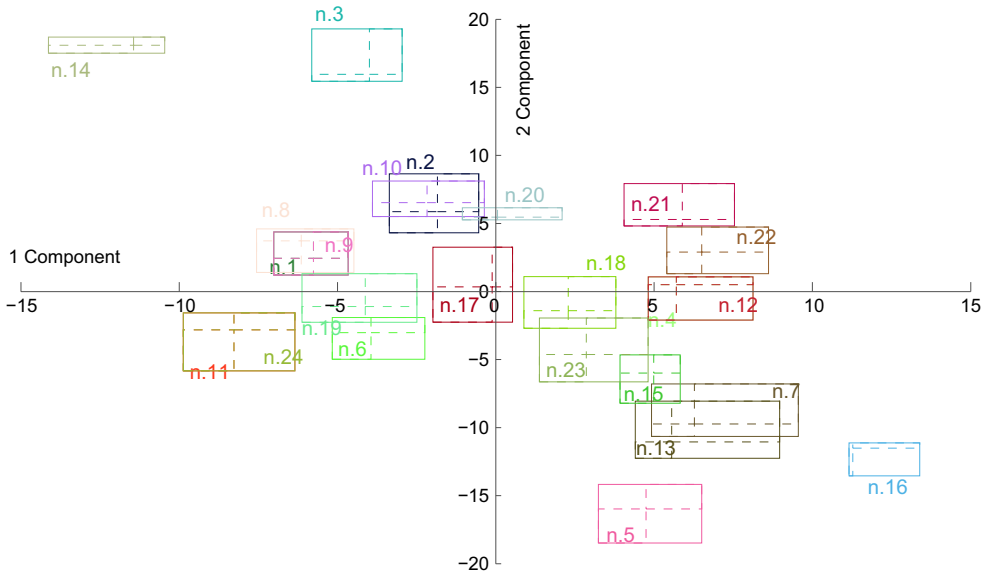

Fig. 8 *Example 2*: score plot for the first and second components (fuzzy units are consecutively numbered and represented with *different colors*) (color figure online)

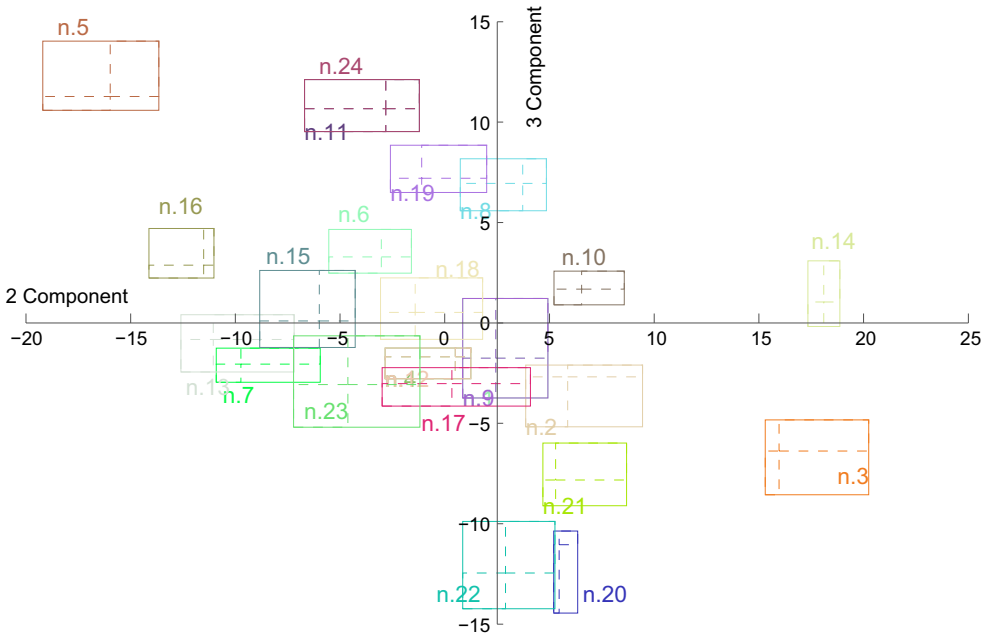

Fig. 9 *Example 2*: score plot for the second and third components (fuzzy units are consecutively numbered and represented with *different colors*) (color figure online)

(0, 0, 1.67), very-low (0, 2.13, 3.33), low (1.67, 3.87, 5.0), middle (3.33, 4.69, 6.67), high (5.0, 7.47, 8.33), very-high (6.67, 8.95, 10.0), extreme (8.33, 10.0, 11.67). The empirical datasets are reported in the supplementary material of this article. The NCFCA algorithm was applied to the preprocessed data and two components ($p = 2$) were extracted ($R = 0.90$). The algorithm converged after 20 iterations. Model accuracy and reliability were also good as shown by the low standard errors reported in Table 3. Finally. the estimated loading matrix was orthonormalized and varimax-rotated.

By inspecting Table 3, one can observe that the first component was negatively related to x_1 (−0.41), x_2 (−0.43), x_3 (−0.42), x_4 (−0.40), x_{10} (−0.41), and to a less extent to x_9 (−0.37). By contrast, the second component was positively related to x_5 (0.49), x_6 (0.54), x_7 (0.53), and x_8

Table 3 Example 3: loading matrix $\widehat{\mathbf{\Gamma}}$ with standard errors in parenthesis ($Q = 5{,}000$)

Indicators/variables	Comp. 1	Comp. 2
(x_1) Household spending	**−0.41** (0.01)	0.04 (0.02)
(x_2) Investments	**−0.43** (0.02)	0.03 (0.01)
(x_3) Income	**−0.42** (0.02)	0.01 (0.01)
(x_4) Salaries	**−0.40** (0.01)	0.05 (0.02)
(x_5) Marriage index	−0.04 (0.02)	**0.49** (0.02)
(x_6) Public education expenditures	0.04 (0.01)	**0.54** (0.01)
(x_7) Unemployment index	0.04 (0.01)	**0.53** (0.01)
(x_8) Energy consumption	−0.05 (0.02)	**0.42** (0.03)
(x_9) Public culture expenditures	**−0.37** (0.02)	−0.06 (0.04)
(x_{10}) Efficiency of health index	**−0.41** (0.02)	−0.06 (0.02)

Loadings higher than 0.35 (in absolute value) are in boldface type

(0.42). The first component can be interpreted as 'overall productivity' and the second one can be referred to 'territorial welfare'. Figure 10 shows an example of some observed (resp. reconstructed) fuzzy sets on the seventh variable (x_7). Figure 11 shows the score plot for the extracted components. A clear pattern can be read from the score plot. In particular, the southern regions (Campania, Calabria, Sicily, Sardinia, Apulia, Molise) were located in the second quadrant, whereas many of the richest northern regions (Lombardy, Veneto, Emilia Romagna, Tuscany) were in the fourth quadrant. Finally, the middle part of the plot contained central regions as well as some small northern ones (e.g., Friuli, Aosta Valley, Trentino). The regions in the fourth quadrant showed socio-economical profiles characterized by high productivity as well as a solid territorial welfare, whereas the regions in the second quadrant of the map showed the opposite pattern. In particular, the regions Campania, Sicily, Calabria, and Basilicata showed low values for productivity and territorial welfare. By considering the sizes of the projected rectangles, most of the regions were characterized by similar degree of fuzziness whereas Lombardy seemed to be the region with the higher fuzziness.

5 Conclusion and further perspectives

In this paper, we extended the component analysis approach to non-convex fuzzy data. The proposed NCFCA method allowed to reduce the dimensionality of multivariate datasets with non-convex fuzzy observations. In particular, the proposed method considered non-convexity by directly incorporating the membership values of fuzzy observations in the NCFCA model. To better illustrate the NCFCA features, we also described three real applications with non-convex fuzzy data. The empirical results suggested the important role played by this property when researchers have to deal with complex, imprecise, and vague information. Furthermore, it is straightforward to note how NCFCA can also be applied when data are represented by standard convex fuzzy features, in particular by setting $\mathbf{MU_0} = \mathbf{MU_1} = \mathbf{MU_2} = \mathbf{1}_{n\times m}$ for trapezoidal cases and $\mathbf{H} = \mathbf{I} = \mathbf{0}_{n\times m}$ together with $\mathbf{MU_0} = \mathbf{MU_1} = \mathbf{MU_2} = \mathbf{1}_{n\times m}$ for triangular cases.

However, the proposed method can potentially suffer from some limitations. For instance, in some empirical cases the piecewise-linear representation for 2-mode fuzzy data cannot be valid and therefore other representations should be preferred (e.g., quadratic or cubic 2-mode fuzzy numbers). Finally, for some datasets the unconstrained algorithm could estimate parameter values which violate the normative properties or the MCAR representation may not adequately represent the whole information provided by 2-mode fuzzy data. In particular, although the 2-mode representation can be still reasonable for modeling data from human decision-making contexts, other empirical situations may require more complex and flexible representations (e.g., k-mode fuzzy numbers, $k \geq 2$). Various possible extensions of our proposal could be considered. For example, the adoption of a constrained approach for the estimation

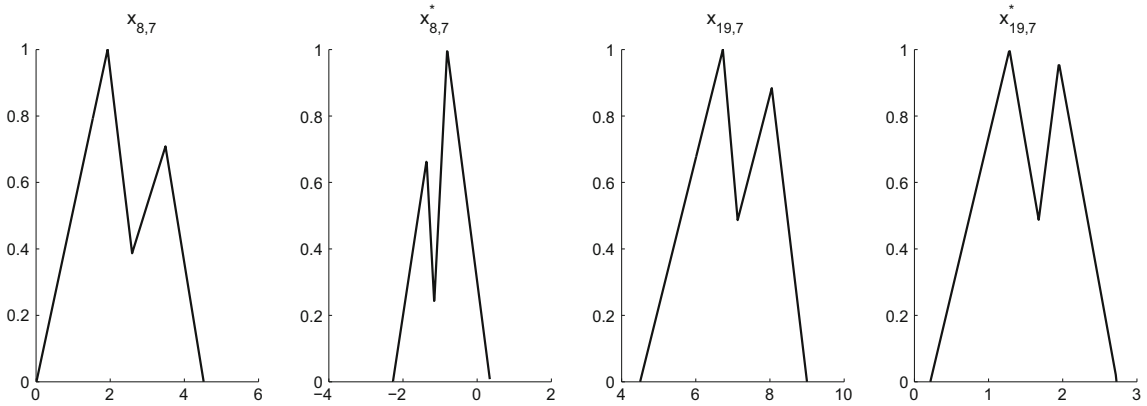

Fig. 10 *Example 3*: observed (x) and model reconstructed (x^*) fuzzy sets for subjects eight and 19 on variable x_7

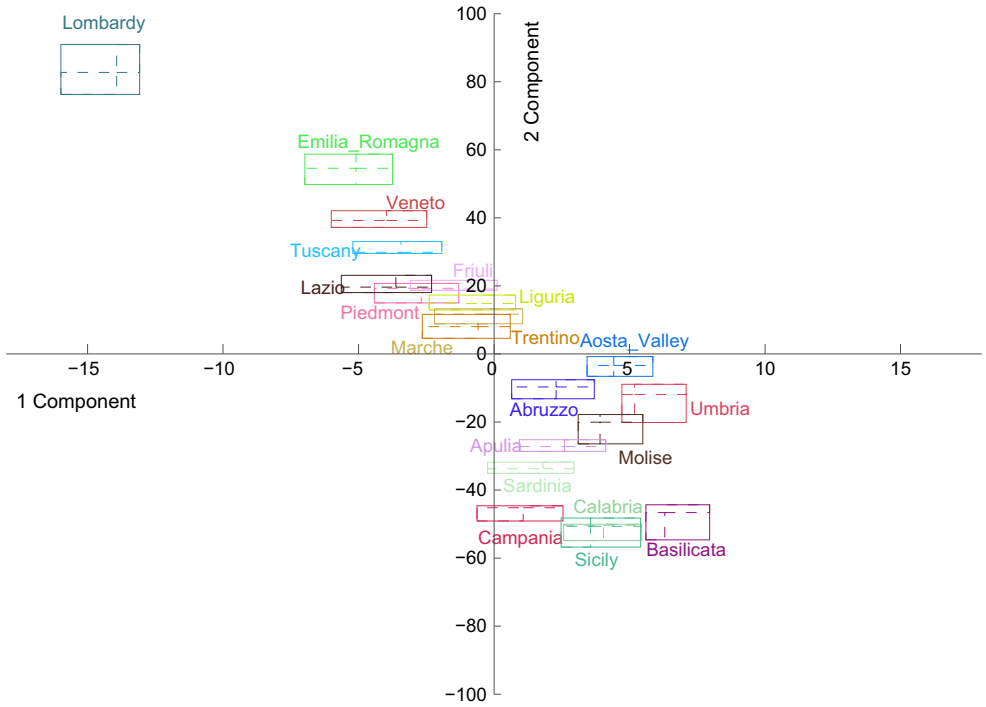

Fig. 11 *Example 3*: score plot for the first and second components (fuzzy units are named and represented with *different colors*) (color figure online)

procedure would always guarantee the consistency between the estimated parameters and the corresponding normative representations for non-convex fuzzy numbers. Moreover, a future venue of research would also consist in the improvement of the graphical representation and/or the extension of the NCFCA method to cases where data are described by means of k-mode fuzzy data beyond the ones characterized by piecewise-linear membership functions. However, this extension would necessarily require a deeper investigation of the foundational/mathematical aspects concerning the non-convexity property in fuzzy numbers. Finally, an organic framework able to deal with randomness and fuzziness simultaneously might be adopted in future developments (e.g., using Fuzzy Random Variables). This would extend our proposal beyond the semi-descriptive approach presented in this contribution.

References

Abdi H, Williams LJ (2010) Principal component analysis. Wiley Interdiscip Rev Comput Stat 2(4):433–459

van den Berg RA, Hoefsloot HC, Westerhuis JA, Smilde AK, van der Werf MJ (2006) Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genomics 7(1):142

Bloch I (1999) On fuzzy distances and their use in image processing under imprecision. Pattern Recognit 32(11):1873–1895

Bro R, Smilde AK (2003) Centering and scaling in component analysis. J Chemom 17(1):16–33

Calcagnì A, Lombardi L (2014) Dynamic fuzzy rating tracker (DYFRAT): a novel methodology for modeling real-time dynamic cognitive processes in rating scales. Appl Soft Comput 24(1):948–961

Calcagnì A, Lombardi L, Pascali E (2014) Non-convex fuzzy data and fuzzy statistics: a first descriptive approach to data analysis. Soft Comput 18(8):1575–1588

Cazes P, Chouakria A, Diday E, Schektrman Y (1997) Entension de l'analyse en composantes principales à des données de type intervalle. Revue de Stat appliquée 45(3):5–24

Ciavolino E, Salvatore S, Calcagnì A (2014) A fuzzy set theory based computational model to represent the quality of inter-rater agreement. Qual Quant 48(4):2225–2240

Coppi R, Giordani P, D'Urso P (2006) Component models for fuzzy data. Psychometrika 71(4):733–761

Diamond P (1988) Fuzzy least squares. Inf Sci 46(3):141–157

Douzal-Chouakria A, Billard L, Diday E (2011) Principal component analysis for interval-valued observations. Stat Anal Data Min 4(2):229–246

Dubois D, Prade HM, Farreny H, Martin-Clouaire R, Testemale C (1988) Possibility theory: an approach to computerized processing of uncertainty, vol 2. Plenum press, New York

Eriksson L (2006) Multi-and megavariate data analysis. MKS Umetrics AB

Facchinetti G, Pacchiarotti N (2006) Evaluations of fuzzy quantities. Fuzzy Sets Syst 157(7):892–903

Garibaldi JM, Musikasuwan S, Ozen T, John RI (2004) A case study to illustrate the use of non-convex membership functions for linguistic terms. In: Fuzzy systems, 2004. Proceedings 2004 IEEE International Conference on, IEEE, vol 3. pp 1403–1408

Giordani P (2010) Three-way analysis of imprecise data. J Multivar Anal 101(3):568–582

Giordani P, Kiers HA (2004) Principal component analysis of symmetric fuzzy data. Comput Stat Data Anal 45(3):519–548

Giordani P, Kiers HA (2007) Principal component analysis with boundary constraints. J Chemom 21(12):547–556
Greene J, Haidt J (2002) How (and where) does moral judgment work? Trends Cogn Sci 6(12):517–523
Hanss M (2005) Applied fuzzy arithmetic. Springer
Hastie T, Tibshirani R, Friedman J (2001) The elements of statistical learning theory. Springer
Hesketh T, Pryor R, Hesketh B (1988) An application of a computerized fuzzy graphic rating scale to the psychological measurement of individual differences. Int J Man Mach Stud 29(1):21–35
Irpino A, Lauro C, Verde R (2003) Visualizing symbolic data by closed shapes. In: Between data science and applied data analysis, Springer, pp 244–251
Jahanshahloo GR, Lotfi FH, Izadikhah M (2006) Extension of the topsis method for decision-making problems with fuzzy data. Appl Math Comput 181(2):1544–1551
Johnson A, Mulder B, Sijbinga A, Hulsebos L (2012) Action as a window to perception: measuring attention with mouse movements. Appl Cogn Psychol 26(5):802–809
Kaiser HF (1958) The varimax criterion for analytic rotation in factor analysis. Psychometrika 23(3):187–200
Kiers HA (1997) Techniques for rotating two or more loading matrices to optimal agreement and simple structure: a comparison and some technical details. Psychometrika 62(4):545–568
Kiers HA (2002) Setting up alternating least squares and iterative majorization algorithms for solving various matrix optimization problems. Comput Stat Data Anal 41(1):157–170
Kiers HA (2004) Bootstrap confidence intervals for three-way methods. J Chemom 18(1):22–36
Kiers HA, ten Berge JM (1989) Alternating least squares algorithms for simultaneous components analysis with equal component weight matrices in two or more populations. Psychometrika 54(3):467–473
Lalla M, Facchinetti G, Mastroleo G (2005) Ordinal scales and fuzzy set systems to measure agreement: an application to the evaluation of teaching activity. Qual Quant 38(5):577–601
Lauro CN, Palumbo F (2000) Principal component analysis of interval data: a symbolic data analysis approach. Comput Stat 15(1):73–87
Le-Rademacher J, Billard L (2012) Symbolic covariance principal component analysis and visualization for interval-valued data. J Comput Graph Stat 21(2):413–432
Magnuson JS (2005) Moving hand reveals dynamics of thought. Proc Natl Acad Sci USA 102(29):9995–9996
Meredith W, Millsap RE (1985) On component analyses. Psychometrika 50(4):495–507
Millsap RE, Meredith W (1988) Component analysis in cross-sectional and longitudinal data. Psychometrika 53(1):123–134
Reuter U (2008) Application of non-convex fuzzy variables to fuzzy structural analysis. Soft methods for handling variability and imprecision. pp 369–375
Stöber J, Joormann J (2001) A short form of the worry domains questionnaire: construction and factorial validation. Personal Individ Differ 31(4):591–598
Taheri SM (2003) Trends in fuzzy statistics. Austrian J Stat 32(3):239–257
Trefethen LN, Bau D (1997) Numerical linear algebra. Siam, p 50
Viertl R (2011) Statistical methods for fuzzy data. Wiley
Weber EU, Johnson EJ (2009) Mindful judgment and decision making. Ann Rev Psychol 60:53–85
Yang MS, Ko CH (1996) On a class of fuzzy c-numbers clustering procedures for fuzzy data. Fuzzy Sets Syst 84(1):49–60
Zimmermann HJ (2001) Fuzzy set theory-and its applications. Springer