Three-way component models for imprecise data

Paolo Giordani
paolo.giordani@uniroma1.it

Sapienza University of Rome

TRICAP, Vall de Núria – Spain, June 15th, 2009
## Outline

1. **Uncertainty**
   - Some clarifications

2. **Single valued case**
   - Membership functions
   - Data
   - Models

3. **Interval valued case**
   - Membership functions
   - Data
   - Intermezzo: two-way case
   - Models
   - Application: Air pollution in Rome

4. **Fuzzy valued case**
   - Membership functions
   - Data
   - Intermezzo: two-way case
   - Models
   - Application: Italian bank branches performance

5. **Conclusion**

6. **References**
Outline

1. Uncertainty
   - Some clarifications

2. Single valued case
   - Membership functions
   - Data
   - Models

3. Interval valued case
   - Membership functions
   - Data
   - Intermezzo: two-way case
   - Models
   - Application: Air pollution in Rome

4. Fuzzy valued case
   - Membership functions
   - Data
   - Intermezzo: two-way case
   - Models
   - Application: Italian bank branches performance

5. Conclusion

6. References
So far as the laws of mathematics refer to reality, they are not certain. And so far they are certain, they do not refer to reality. *Albert Einstein (Geometry and Experience)*

- **UNCERTAINTY**: the state of being uncertain
- **UNCERTAIN**: something not clearly known
Some clarifications

- randomness (uncertainty about the (precise) outcome of a (random) mechanism \[probabilistic uncertainty\])
- imprecision (uncertainty concerning the placement of an outcome in a given class \[non probabilistic uncertainty\])
  - vagueness (outcome expressed in linguistic terms)
  - partial or total ignorance concerning an outcome
  - “granularity” of an outcome with reference to the way it is defined and used in the analysis (e.g. the age of a person may be described in terms of 5-year intervals, or just as “young” “middle age”, “old”; a different amount of uncertainty of imprecision is associated to each of these “granulations”)
  - calibration of the measurement device registering the outcome
- see Klir (2006)
Some clarifications

Examples of imprecise outcomes

- about 3
- temperature in Rome
- I feel good
- wait me for 5 minutes

✓ every example can be suitably handled in order to manage the associated imprecision
Outline

1. Uncertainty
   - Some clarifications

2. Single valued case
   - Membership functions
   - Data
   - Models

3. Interval valued case
   - Membership functions
   - Data
   - Intermezzo: two-way case
   - Models
   - Application: Air pollution in Rome

4. Fuzzy valued case
   - Membership functions
   - Data
   - Intermezzo: two-way case
   - Models
   - Application: Italian bank branches performance

5. Conclusion

6. References
How to handle the previous examples: practical point of view
How to handle the previous examples: theoretical point of view

- $U \rightarrow$ universe (of elements)
- $X \subseteq U \rightarrow$ set (of elements) – attribute
- $\mu_X(x) \rightarrow$ membership function (of $x$ to $X$, $\forall x \in U$):

$$
\mu_X(x) = \begin{cases} 
1 & x = m \\
0 & \text{otherwise}
\end{cases}
$$

$m$: midpoint
the available data array $X$ contains single valued information with regard to a set of $I$ observation units on which $J$ variables have been recorded at $K$ different occasions

- array $X$ $\rightarrow$ supermatrix $X_a \ (I \times JK)$

where

$$X_a = \begin{bmatrix} X_1 & \cdots & X_k & \cdots & X_K \end{bmatrix}$$

with

$$X_k = \begin{bmatrix} m_{ijk} \end{bmatrix} \ (I \times J)$$

✓ the CANDECOMP/PARAFAC and Tucker3 models can be applied
the CANDECOMP/PARAFAC (CP) model (Carroll & Chang, 1970; Harshman, 1970) can be formalized as follows ($S$ components):

$$X_a = A (C \odot B)' + E_a = A I_a (C' \otimes B') + E_a$$

where:

- $A$ ($I \times S$) $\rightarrow$ component matrix for the observation unit mode
- $B$ ($J \times S$) $\rightarrow$ component matrix for the variable mode
- $C$ ($K \times S$) $\rightarrow$ component matrix for the occasion mode
- $I_a$ ($S \times S^2$) $\rightarrow$ supermatrix (from identity array $I$ ($S \times S \times S$))
- $E_a$ ($I \times JK$) $\rightarrow$ supermatrix (from error array $E$ ($I \times J \times K$))
the **Tucker3 (T3)** model (Tucker, 1966) can be formalized as follows ($P$, $Q$ and $R$ components for the observation unit, variable and occasion modes, respectively):

$$X_a = A G_a (C' \otimes B') + E_a$$

where:

- **$A$** ($I \times P$) → component matrix for the observation unit mode
- **$B$** ($J \times Q$) → component matrix for the variable mode
- **$C$** ($K \times R$) → component matrix for the occasion mode
- **$G_a$** ($P \times QR$) → supermatrix (from core array $G$ ($P \times Q \times R$))
- **$E_a$** ($I \times JK$) → supermatrix (from error array $E$ ($I \times J \times K$))
About CP and T3

- **CP** solution is **unique** (under mild conditions)
- **T3** solution suffers from rotational **indeterminacy** (if $S$, $T$ and $U$ are nonsingular matrices, then
  \[ AG_a(C' \otimes B') = \hat{A}\hat{G}_a(\hat{C'} \otimes \hat{B'}) \text{ with } \hat{A} = AS, \hat{B} = BT, \]
  \[ \hat{C} = CU \text{ and } \hat{G}_a = S^{-1}G_a(U^{-1} \otimes T^{-1}) \])
- **CP** as constrained version of **T3** (imposing $G = I$ with $P = Q = R = S$)
- **ALS** algorithm for determining the optimal solution
- **goodness of fit** by means of $1 - \left( \frac{\|E_a\|^2}{\|X_a\|^2} \right)$
- the entities of the observation unit mode can be represented as **points** into the $P$- or $S$-dimensional subspace of $\mathbb{R}^{JK}$ spanned by $F_a = (C \otimes B)G'_a$ (for T3) or $F_a = (C \otimes B)l'_a$ (for CP) using the rows of $A$ as coordinates
Outline

1. Uncertainty
   - Some clarifications

2. Single valued case
   - Membership functions
   - Data
   - Models

3. Interval valued case
   - Membership functions
   - Data
   - Intermezzo: two-way case
   - Models
   - Application: Air pollution in Rome

4. Fuzzy valued case
   - Membership functions
   - Data
   - Intermezzo: two-way case
   - Models
   - Application: Italian bank branches performance

5. Conclusion

6. References
How to handle the previous examples: practical point of view
How to handle the previous examples: theoretical point of view

- $U$        $\rightarrow$ universe (of elements)
- $X \subseteq U$  $\rightarrow$ set (of elements) – attribute
- $\mu_X(x)$  $\rightarrow$ membership function (of $x$ to $X$, $\forall x \in U$):
  $$\mu_X(x) = \begin{cases} 
1 & m_1 \leq x \leq m_2 \\
0 & \text{otherwise}
\end{cases}$$

$m_1$: left mode
$m_2$: right mode
the available data array $\mathbf{X}$ contains interval valued information with regard to a set of $I$ observation units on which $J$ variables have been recorded at $K$ different occasions.

- array $\mathbf{X} \longrightarrow$ supermatrix $\mathbf{X}_a \ (I \times JK)$

where

$$\mathbf{X}_a = \begin{bmatrix} \mathbf{X}_1 & \cdots & \mathbf{X}_k & \cdots & \mathbf{X}_K \end{bmatrix}$$

with

$$\mathbf{X}_k = \begin{bmatrix} (m_{1ijk}, m_{2ijk}) \end{bmatrix}, (I \times J)$$

✓ the CP and T3 models cannot be applied (!)
**Graphical representation**

In the **single valued** case, each observation unit $i$ ($i = 1, \ldots, I$) can be represented as a point in $\mathbb{R}^{JK}$.

In the **interval valued** case, each observation unit $i$ ($i = 1, \ldots, I$) can be represented as a hyperrectangle in $\mathbb{R}^{JK}$ with $2^{JK}$ vertices.

As a particular case, when $JK = 2$, each observation unit is a rectangle. For instance, if $J = 2$ and $K = 1$, we have $m_{ijk} = \left( \frac{m_{1ijk} + m_{2ijk}}{2} \right)$. 

![Diagram](image-url)
Midpoint Principal Component Analysis (M-PCA)

Cazes et al. (1997), Giordani & Kiers (2006)

compute **midpoints** (contained in $X_{M-PCA}$ ($I \times J$)) and perform PCA on $X_{M-PCA}$

$X_{M-PCA} = AF' + E$, where $A$ and $F$ ($F'F = I$) are the component score and loading matrices, respectively, $E$ is the error matrix

\[
X = \begin{bmatrix}
[m_{111}, m_{211}] & \cdots & [m_{11J}, m_{21J}] \\
\vdots & \ddots & \vdots \\
[m_{1I1}, m_{2I1}] & \cdots & [m_{1IJ}, m_{2IJ}]
\end{bmatrix}
\]

\[
X_{M-PCA} = M = \begin{bmatrix}
m_{11} & \cdots & m_{1J} \\
\vdots & \ddots & \vdots \\
m_{I1} & \cdots & m_{IJ}
\end{bmatrix}
\]

✓ very simple!
Intermezzo: two-way case

**Vertices Principal Component Analysis (V-PCA)**

Cazes et al. (1997), Giordani & Kiers (2006)

compute vertices (contained in $X_{V-PCA}$ ($I2^J \times J$)) and perform PCA on $X_{V-PCA}$ ($X_{V-PCA} = AF' + E$)

\[
X \xrightarrow{\text{V-PCA transformation}} X_{V-PCA} = \begin{bmatrix}
    1_{X_{V-PCA}} \\
    \vdots \\
    i_{X_{V-PCA}} \\
    \vdots \\
    I_{X_{V-PCA}}
\end{bmatrix}
\]

where, if $J = 2$, $iX_{V-PCA} = \begin{bmatrix}
m_{1i1} & m_{1i2} \\
m_{1i1} & m_{2i2} \\
m_{2i1} & m_{1i2} \\
m_{2i1} & m_{2i2}
\end{bmatrix}$

✓ curse of dimensionality! $X_{V-PCA}$ contains the $I$ matrices $iX_{V-PCA}$ ($2^J \times J$) stacked one below each other. If $I = 12$ and $J = 8$, $X_{V-PCA}$ has 3072 rows (and 8 columns)
in case of $S$ components, the entities of the observation unit mode can be represented as **$S$-dimensional hyperrectangles** (rectangles if $S = 2$) into the $S$-dimensional subspace of $\mathbb{R}^J$ spanned by $\mathbf{F}$ with coordinates given by ($i = 1, \ldots, I$; $s = 1, \ldots, S$)

$$a_{M1i s} = \sum_{j: f_{js} < 0} m_{2ij} f_{js} + \sum_{j: f_{js} > 0} m_{1ij} f_{js}$$

$$a_{M2i s} = \sum_{j: f_{js} < 0} m_{1ij} f_{js} + \sum_{j: f_{js} > 0} m_{2ij} f_{js}$$

**Note:** for every observation unit, the projection of all the vertices does not define an $S$-dimensional hyperrectangle. By means of the above formulas, this problem is solved, however, by representing a generic observation unit on each axis as the segment which includes all the projections.
Comparison between M-PCA and V-PCA (+ V-PCA shortcut)...

- **M-PCA** is simpler than **V-PCA**...
- ...but **M-PCA** does not consider the width of the intervals when extracting the components.
- the plotting procedure can be performed using the loadings (the scores can also be unknown)

**Idea**

in order to find the loadings, use the **cross-product matrix** $X'_{V-PCA} X_{V-PCA}$, which has order $(J \times J)$. In fact, it is well known that the **component loadings** (in $F$) are obtained as **eigenvectors** of the cross-product matrix.
...Comparison between M-PCA and V-PCA (+ V-PCA shortcut)

- **V-PCA cross-product matrix** \((X'_{V-PCA}X_{V-PCA})\):

\[
2^{J-2} \begin{bmatrix}
2 \sum_{i=1}^{I} (m_{1i1}^2 + m_{2i1}^2) & \cdots & \sum_{i=1}^{I} (m_{1i1} + m_{2i1})(m_{1iJ} + m_{2iJ}) \\
\vdots & \ddots & \vdots \\
\sum_{i=1}^{I} (m_{1iJ} + m_{2iJ})(m_{1i1} + m_{2i1}) & \cdots & 2 \sum_{i=1}^{I} (m_{1iJ}^2 + m_{2iJ}^2)
\end{bmatrix}
\]

- **M-PCA cross-product matrix** \((X'_{M-PCA}X_{M-PCA})\):

\[
2^{-2} \begin{bmatrix}
\sum_{i=1}^{I} (m_{1i1}^2 + m_{2i1}^2) & \cdots & \sum_{i=1}^{I} (m_{1i1} + m_{2i1})(m_{1iJ} + m_{2iJ}) \\
\vdots & \ddots & \vdots \\
\sum_{i=1}^{I} (m_{1iJ} + m_{2iJ})(m_{1i1} + m_{2i1}) & \cdots & \sum_{i=1}^{I} (m_{1iJ}^2 + m_{2iJ}^2)
\end{bmatrix}
\]

✓ hence the only essential difference between the cross-product matrices is in the diagonal elements
Three-way Midpoint Principal Component Analysis (3M-PCA)

Giordani & Kiers (2004a)

compute midpoints (in array $X_{3M-PCA} = [m_{ijk}] (I \times J \times K)$)
perform T3 or CP on $X_{3M-PCA}$

- the same properties of standard T3 and CP are guaranteed
- the plotting procedure of the entities of the observation unit mode (as low-dimensional hyperrectangles) can be done by using the two-way formulas provided that $F$ is replaced by $F_a = (C \otimes B) G'_a$ (for T3) or $F_a = (C \otimes B) I'_a$ (for CP)

✓ straightforward extension!
Three-way Vertices Principal Component Analysis (3V-PCA) [1]

Giordani & Kiers (2004a)

computing matrix $X_{3V-PCA}$ is **unfeasible** (*Curse of dimensionality!*)

- if we apply the $V$-PCA transformation to one row of the supermatrix $X_a$, we obtain the matrix of order $(2^{JK} \times JK)$ whose rows refer exactly to all the $2^{JK}$ vertices of the hyperrectangle representing each observation unit
- it also follows that the **transformed data matrix**, say $X_{3V-PCA}$, for the full matrix $X_a$ has order $(I2^{JK} \times JK)$
- the simple case with $I = J = 3$ and $K = 4$ leads to a huge $X_{3V-PCA}$ ($12288 \times 12$)

**Idea**

use the $V$-PCA shortcut (modified to the three-way case).
For the analysis of $X_{3V-PCA}$ we can use modified algorithms proposed for handling three-way arrays in which $I \gg \max(J, K)$.
The basis of such algorithms is that, in the iterative part of the procedure, only the **cross-products** of $X_{3V-PCA}$ are needed
Three-way Vertices Principal Component Analysis (3V-PCA) [2]

3V-PCA cross-product matrix $X'_{3V-PCA}X_{3V-PCA} = [X_{kk'}]$

$$X_{kk'} = kX'_{V-PCA}k'X_{3V-PCA}$$

where $kX'_{V-PCA}$ and $k'X'_{V-PCA}$ denote the matrices resulting from the V-PCA transformation applied on $X_k$ and $X_{k'}$ respectively $(k, k' = 1, \ldots, K)$

$$X_{kk'} =$$

$$2^{J-2} \begin{bmatrix}
\sum_{i=1}^{I} (m_{1i1k} + m_{2i1k}) (m_{1i1k'} + m_{2i1k'}) & \cdots & \sum_{i=1}^{I} (m_{1i1k} + m_{2i1k}) (m_{1iJk'} + m_{2iJk'}) \\
\vdots & \ddots & \vdots \\
\sum_{i=1}^{I} (m_{1i1k} + m_{2i1k}) (m_{1i1k'} + m_{2i1k'}) & \cdots & \sum_{i=1}^{I} (m_{1i1k} + m_{2i1k}) (m_{1i1k'} + m_{2i1k'})
\end{bmatrix}$$

whereas, if $k = k'$, the diagonal elements are replaced by $2 \sum_{i=1}^{I} (m_{1ijk}^2 + m_{2ijk}^2)$ $(j = 1, \ldots, J; k, = 1, \ldots, K)$

**Note:** as for the two-way case, the 3V-PCA and 3M-PCA cross-product matrices differ only in the diagonal elements
Three-way Vertices Principal Component Analysis (3V-PCA) [3]

using Kiers & Harshman (1997)
the three-way analysis of the supermatrix $X_{3V-PCA}$ and that of a
supermatrix, say $W$, which has exactly the same cross-products as
$X_{3V-PCA}$, give essentially the same matrices $B$, $C$ and $G_a$

- decompose $X'_{3V-PCA}X_{3V-PCA}$, e.g. eigendecomposition ($K$
  and $D^2$ contain the eigenvectors and the eigenvalues,
  respectively):

$$X'_{3V-PCA}X_{3V-PCA} = KD^2K'$$

- set $W' = KD$
- perform T3 or CP on $W$

✓ the same properties and plotting procedure of 3M-PCA also hold
for 3V-PCA
Application: Air pollution in Rome

Problem

- The air pollution in Rome is monitored by 12 testing stations.
- In research by the municipality of Rome, the testing stations are classified into four groups:
  - A: Lower pollution urban station (park)
  - B: Urban stations (residential areas)
  - C: Urban stations (high traffic areas)
  - D: Suburban stations (areas indirectly exposed to vehicular pollution)
Data array

- three-way analysis on:
  - i $l = 7$ testing stations
    - A Ada
    - B Magnagrecia, Preneste
    - C Fermi, Francia
    - D Castel di Guido, Tenuta del Cavaliere
  - ii $j = 3$ pollutants (NO, NO$_2$ and O$_3$)
  - iii $k = 90$ days (January, 1st, – March, 31st, 1999)

- $m_1$ and $m_2$ $\rightarrow$ min and max values daily registered

- we apply 3M-PCA and 3V-PCA using T3

- preprocessing: centering across the observation units and scaling within the variables
choice of $P$, $Q$ and $R$ + simplicity

- 3M-PCA and 3V-PCA with $P = Q = 2$ and $R = 1$
- solution with good fit (73.17%) and easily interpretable
- rotations to simplicity
  - since $R = 1$, T3 written as $X_a = (C' \otimes AG_aB') + E_a$
  - since $P = Q = 2$, $\text{rank}(AG_aB') = 2$
  - SVD of $AG_aB'$: $AG_aB' = PDQ'$
  - $\hat{A} = P_2D_2$ and $\hat{B} = Q_2$ (first two columns)
  - varimax on $\hat{B}$ (compensating it into $\hat{A}$) $\rightarrow B^\nu$ and $A^\nu$
  - $AG_aB' = A^\nu B^\nu = A^\nu IB^\nu$ (the rotated core is $G_a^\nu = I$)
Interpretation [1]

- **Component matrix for the variables** *(V-PCA within parentheses)*

<table>
<thead>
<tr>
<th>pollutant</th>
<th>C1</th>
<th>C2</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
<td>-0.10</td>
<td>-0.07</td>
</tr>
<tr>
<td>NO₂</td>
<td>0.08</td>
<td>0.04</td>
</tr>
<tr>
<td>O₃</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

✓ **C1:** secondary pollutants; **C2:** primary pollutants

- **Component matrix for the occasions**

✓ **C1:** measure of air pollution during each day
Interpretation [2]

- Component matrix for the observation units

\[ G^v_a = I \]  \hspace{1cm} \text{one-to-one relation between the observation unit and variable components}

↓ midpoint coordinates ↓

<table>
<thead>
<tr>
<th>testing station</th>
<th>C1</th>
<th>C2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ada</td>
<td>-3.41</td>
<td>-10.24</td>
</tr>
<tr>
<td>Fermi</td>
<td>-7.94</td>
<td>16.16</td>
</tr>
<tr>
<td>Francia</td>
<td>-9.30</td>
<td>6.04</td>
</tr>
<tr>
<td>Magnagrecia</td>
<td>-3.55</td>
<td>11.48</td>
</tr>
<tr>
<td>Preneste</td>
<td>4.59</td>
<td>2.74</td>
</tr>
<tr>
<td>Castel di Guido</td>
<td>16.41</td>
<td>-16.53</td>
</tr>
<tr>
<td>Tenuta del Cavaliere</td>
<td>3.20</td>
<td>-9.65</td>
</tr>
</tbody>
</table>

✓ high left side: most polluted stations (‘B’ and ‘C’)
✓ low right side: less polluted stations (‘A’ and ‘D’)
✓ size of the rectangles reflects pollutant daily variations
Outline

1. Uncertainty
   - Some clarifications

2. Single valued case
   - Membership functions
   - Data
   - Models

3. Interval valued case
   - Membership functions
   - Data
   - Intermezzo: two-way case
   - Models
   - Application: Air pollution in Rome

4. Fuzzy valued case
   - Membership functions
   - Data
   - Intermezzo: two-way case
   - Models
   - Application: Italian bank branches performance

5. Conclusion

6. References
How to handle the previous examples: practical point of view

- About 3
- Temperature in Rome
- I feel good (0−10 scale)
- Wait me for 5 minutes
How to handle the previous examples: theoretical point of view

- **U** → universe (of elements)
- **X ⊆ U** → set (of elements) – attribute
- **μ_X(x)** → membership function (of x to X, ∀x ∈ U):

\[
μ_X(x) = \begin{cases} 
L \left( \frac{m_1 - x}{l} \right) & m_1 - l ≤ x < m_1 \\
1 & m_1 ≤ x ≤ m_2 \\
R \left( \frac{x - m_2}{r} \right) & m_2 < x ≤ m_2 + r \\
0 & \text{otherwise}
\end{cases}
\]

- **m_1**: left mode
- **m_2**: right mode
- **l (> 0)**: left spread
- **r (> 0)**: right spread
Trapezoidal fuzzy datum

- \( L \left( \frac{m_1 - x}{l} \right) \) and \( R \left( \frac{x - m_2}{r} \right) \) \rightarrow \) decreasing shape functions from \( \mathbb{R}^+ \) to \([0, 1]\) (fulfilling specific requirements)

- almost always

\[
L (w) = R (w) = \begin{cases} 
1 - w^q & 0 \leq w \leq 1 \\
0 & \text{otherwise}
\end{cases}
\]

with \( q > 0 \)

- if \( q = 1 \) (this is usually the case), \( \tilde{X} \equiv (m_1, m_2, l, r) \) is a trapezoidal fuzzy datum:

\[
\mu_{\tilde{X}}(x) = \begin{cases} 
1 - \frac{m_1 - x}{l} & m_1 - l \leq x < m_1 \\
1 & m_1 \leq x \leq m_2 \\
1 - \frac{x - m_2}{r} & m_2 < x \leq m_2 + r \\
0 & \text{otherwise}
\end{cases}
\]
Comparison among single, interval and fuzzy data
Some further comments \cite{1}

- symbol ‘\(\sim\)’ denotes fuzziness
- \(\mu_{\tilde{X}}(x) \in [0, 1]\) (in the previous cases \(\mu_X(x) \in \{0, 1\}\))
- what means \(\mu_{\tilde{X}}(x) = 0.8\)?

It expresses the degree of truth of \(x \in U\) to \(X\)

Example
\(U = \{0, \ldots, 10\}\), what is the subset \(X\) of ‘small’ integers?

- Person n.1 may reasonably state that \(X = \{0, 1\}\)
- Person n.2 may argue that also 2 is small, even if to some extent. For instance, in a scale from 0 to 1, 2 is small with a grade equal to 0.8

0.8 is not a measure of randomness. It stems from the (non-probabilistic) uncertainty about classifying 2 as a ‘small’ integer (degree of truth of ‘small’).
Some further comments [2]

- can $\mu_{\tilde{X}}(x)$ be interpreted as a probability?
  - i. it does not follow the laws of probability (e.g., $\sum_{x \in U} \mu_{\tilde{X}}(x) > 1$)
  - ii. membership function as frequentist conditional probability (Loginov, 1966) [deeply criticized]
  - iii. membership function as subjectivistic conditional probability exploiting the de Finetti theory (Coletti & Scozzafava, 2002)

- is probability theory (or fuzzy set theory) the best tool for solving problems involving uncertainty (imprecision and/or randomness) [much debated question]?
  - i. three lines of thought (Yes, No, Complementary)
  - ii. the debate is far from reaching a conclusion, even inside a specific line of thought (Laviolette et al., 1995; Singpurwalla & Booker, 2004)
the available data array $\tilde{X}$ contains fuzzy valued information with regard to a set of $I$ observation units on which $J$ variables have been recorded at $K$ different occasions

- array $\tilde{X} \rightarrow$ supermatrix $\tilde{X}_a (I \times JK)$

where

$$\tilde{X}_a = [ \tilde{X}_1 \ \cdots \ \tilde{X}_k \ \cdots \ \tilde{X}_K ]$$

with

$$\tilde{X}_k = [ \tilde{x}_{ijk} \equiv (m_{1ijk}, m_{2ijk}, l_{ijk}, r_{ijk}) ] , (I \times J)$$

✓ once again, the CP and T3 models cannot be applied (!)
(Single valued) arrays and graphical representation

starting from the fuzzy valued array $\tilde{X}$, we can define:

- left mode (single valued) array $M_1$ (and $M_{1a}$)
- right mode (single valued) array $M_2$ (and $M_{2a}$)
- left spread (single valued) array $L$ (and $L_a$)
- right spread (single valued) array $R$ (and $R_a$)

as in the interval valued case, in the fuzzy valued one, each observation unit $i$ ($i = 1, \ldots, I$) can be represented as a hyperrectangle in $\mathbb{R}^{JK}$ with $2^{JK}$ vertices

for each pair of $j$ and $k$ ($j = 1, \ldots, J; k = 1, \ldots, K$), the bounds of the hyperrectangle are given by $m_{1ijk} - l_{ijk}$ and $m_{2ijk} + r_{ijk}$ ($i = 1, \ldots, I$)
Intermezzo: two-way case

Principal Component Analysis for Fuzzy valued data (PCA-F)

Giordani & Kiers (2004b), Coppi et al. (2006)

Aim

to find component matrices such that the dissimilarity measure between observed and modelled data is minimized

But ...

need for a dissimilarity measure between observed \( \tilde{X} \equiv (M_1, M_2, L, R) \) and estimated \( \tilde{X}^* \equiv (M_1^*, M_2^*, L^*, R^*) \) fuzzy valued matrices

Idea

since each observation unit can be represented as a hyperrectangle in \( \mathbb{R}^J \), compute the dissimilarity measure by comparing all the \( 2^J \) vertices
Intermezzo: two-way case

PCA-F: (Squared) dissimilarity measure for fuzzy valued matrices [1]

\[
d^2 \left( \tilde{X}, \tilde{X}^* \right) = \sum_{v=1}^{2^J} \left\| \left[ (M_1 - LY) H_L^v + (M_2 + RZ) H_R^v \right] - \left[ (M_1^* - L^*Y) H_L^v + (M_2^* + R^*Z) H_R^v \right] \right\|^2
\]

where

- **Y** and **Z** → diagonal matrices of order \( J \) which help us to suitably take into account the decreasing shape functions \( L_j(w) \) and \( R_j(w) \) for the \( j \)-th fuzzy variable \( (j = 1, \ldots, J) \)

- **H_L^v** and **H_R^v** → diagonal matrices of order \( J \) which help us to describe every vertex of the hyperrectangles, \( v = 1, \ldots, 2^J \)
Intermezzo: two-way case

PCA-F: (Squared) dissimilarity measure for fuzzy valued matrices [2]

about \( Y \) and \( Z \):

\[
y_j = \int_{\mathbb{R}} L_j(w) \, dw = \int_{0}^{1} (1 - w^{q_j}) \, dw \\
z_j = \int_{\mathbb{R}} R_j(w) \, dw = \int_{0}^{1} (1 - w^{q_j}) \, dw
\]

- they **scale the two spreads** (yielding hyperrectangles with shorter sides covering the most ‘essential’ part of the fuzzy valued data)
- they **are lower than one** whenever the importance of the points decreases as they are farther from the center
- they **take low or high values** according to whether the membership function values are high, respectively, only close to the modes or in almost the entire interval \([m_1 - l, m_2 + r]\)
- \( q_j = \{1/2, 1, 2\} \rightarrow y_j = z_j = \{1/3, 1/2, 2/3\} \), respectively
- \( \mu_{\tilde{\tilde{X}}_j}(m_1 - y_j l) > \mu_{\tilde{\tilde{X}}_j}(m_1 - l) = 0 \) and \( \mu_{\tilde{\tilde{X}}_j}(m_2 + z_j r) > \mu_{\tilde{\tilde{X}}_j}(m_2 + r) = 0 \)
Intermezzo: two-way case

PCA-F: (Squared) dissimilarity measure for fuzzy valued matrices [3]

about \( H^L_v \) and \( H^R_v \) \((v = 1, \ldots, 2^J)\):

- their diagonal elements are the rows of matrices \( H^L \) and \( H^R \) of order \((2^J \times J)\)

- \( H^L \) contains all the possible \( J \)-dimensional vectors of 0 and 1

- \( H^R \) has elements equal to 0 and 1 but switched places with respect to \( H^L \)

for instance, if \( J = 2 \):

\[
H^L = \begin{bmatrix}
1 & 1 \\
1 & 0 \\
0 & 1 \\
0 & 0 \\
\end{bmatrix}, \quad H^R = \begin{bmatrix}
0 & 0 \\
0 & 1 \\
1 & 0 \\
1 & 1 \\
\end{bmatrix}.
\]

for instance, the last rows refer to the vertex of the upper bounds:

\[
(M_1 - L) H^L_4 + (M_2 + R) H^R_4 = (M_1 - L) 0_2 + (M_2 + R) I_2 = M_2 + R
\]
PCA-F: (Squared) dissimilarity measure for fuzzy valued matrices [4]

it can be shown that $d^2(\tilde{X}, \tilde{X}^*)$ can be simplified

\[ \text{[hint: } H^L_H^R = 0 \text{ and } \sum_{v=1}^{2^J} tr(YH^L_v) = 2^J \sum_{v=1}^{2^J} tr(YH^R_v) = 2^{J-1} tr(Y), \forall Y (J \times J)\]

\[2^J \text{ terms}\]

\[d^2 \left( \tilde{X}, \tilde{X}^* \right) = \sum_{v=1}^{2^J} \left\| \left[ (M_1 - LY) H^L_v + (M_2 + RZ) H^R_v \right] - \left[ (M_1^* - L^*Y) H^L_v + (M_2^* + R^*Z) H^R_v \right] \right\|^2\]

\[\downarrow \text{ up to constant } 2^{J-1}\]

\[6 \text{ terms}\]

\[d^2 \left( \tilde{X}, \tilde{X}^* \right) = \left\| M_1 - M_1^* \right\|^2 + \left\| M_2 - M_2^* \right\|^2 - 2 tr \left[ (M_1 - M_1^*) (LY - L^*Y) \right] + 2 tr \left[ (M_2 - M_2^*) (RZ - R^*Z) \right] + \left\| LY - L^*Y \right\|^2 + \left\| RZ - R^*Z \right\|^2\]
PCA-F: Model

the PCA-F model can be formalized as follows ($v = 1, \ldots, 2^J$)

- $(M_1 - L)H_v^L + (M_2 + R)H_v^R = (M_1^* - L^*)H_v^L + (M_2^* + R^*)H_v^R + E$
- $M_1^* = A_{M_1}F'$
- $M_2^* = A_{M_2}F'$
- $L^* = A_LF'$
- $R^* = A_RF'$

where

- $A_{M_1}, A_{M_2}$ → component score matrices for the left modes, right modes, left spreads and right spreads, respectively
- $A_L, A_R$ → component loading matrix

**Note:** PCA-F may resemble SCA-P (Kiers & ten Berge, 1994) (Simultaneous Component Analysis with invariant Pattern)
PCA-F: Comments

- optimal solution by ALS algorithm
- goodness of fit index obtained comparing the explained sum of squares to the observed one
- PCA-F admits rotations
- plotting procedure of the observation units as low-dimensional hyperrectangles ($F'F=I$) as follows ($i = 1, \ldots, I$)
  - $A_{M_1}, A_{M_2}, A_L, A_R$, give the coordinates of $M_1, M_2, L, R$
  - $A_{M_1i}, A_{M_2i}, A_{Li}, A_{Ri}$ are diagonal matrices with the $i$-th row of $A_{M_1}, A_{M_2}, A_L, A_R$, in their diagonals
  - Vertices of the $S$-dimensional hyperrectangle given by
    \[
    A_i = \left[ S^H L (A_{M_1i} - A_{Li}) + S^H R (A_{M_2i} + A_{Ri}) \right]
    \]
    where $S^H L$ and $S^H R$ help us to describe all the vertices (same structure of $H_L$ and $H_R$ but order ($2^S \times S$) [$S$ components])
- A variant can be considered for plotting only the most ‘essential’ part of the fuzzy valued data
PCA-F: Application [1]

student perceptions and opinions on Iraq war, threat of terrorism in Italy and peace in the world

**questions** ($J = 13$):

- Q1 Justification of the Iraq war
- Q2 Consequences of the Iraq war
- Q3 Threat of terrorism in Italy
- Q4 Italian humanitarian aid
- Q5 Reinforcement of Italian humanitarian aid
- Q6 Sending of Italian troops
- Q7 Threat to peace in the world: Afghanistan
- Q8 Threat to peace in the world: China
- Q9 Threat to peace in the world: North Korea
- Q10 Threat to peace in the world: Iran
- Q11 Threat to peace in the world: Israel
- Q12 Threat to peace in the world: Syria
- Q13 Threat to peace in the world: United States

- modified version of the survey ‘Flash Eurobarometer 151’ (European Commission, October, 2003)

- data on $I = 24$ students collected on December, 10th-11st, 2003
PCA-F: Application [2]

fuzzification process

Q1, Q2, Q7-Q13 qualitative ordinal scales
- fuzzification based on suggestions provided by various sources in the literature (Herrera et al., 1997; Sii et al., 2001)

Q3-Q6 visual analogue scale (segment of length 1)
- fuzzification as follows:
  \[ m_1 = m_2 = m = \text{equal to the distance between the left bound and the filled mark} \]
  \[ l = r = \begin{cases} 
  0.125 & 0.125 \leq m \leq 0.875 \\
  m & m < 0.125 \\
  1 - m & m > 0.875 
\end{cases} \]

DK/NA do not know / not applicable
- fuzzification by means of a uniform distribution (a sort of non-informative prior distribution in a Bayesian probabilistic approach) with \( m_1 = 0, m_2 = 1, l = r = 0 \)
## PCA-F: Application [3]

choice of $S$ + simplicity + preprocessing

- preprocessing by centering the modes (no artificial differences among the variables)
- PCA-F with $S = 3$ components
- solution with high fit (81.49%) well capturing the variable information (71.25% if $S = 2$ and 87.87% if $S = 4$)
- maximal simplicity by varimax rotating the component loading matrix (compensating this rotation into the component score matrix)
PCA-F: Application [4]

Component loading matrix

<table>
<thead>
<tr>
<th>Question</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Justification of the Iraq war (Q1)</td>
<td>0.33</td>
<td>0.21</td>
<td>-0.04</td>
</tr>
<tr>
<td>Consequences of the Iraq war (Q2)</td>
<td>0.08</td>
<td>0.02</td>
<td>0.23</td>
</tr>
<tr>
<td>Threat of terrorism in Italy (Q3)</td>
<td>-0.03</td>
<td>0.41</td>
<td>0.21</td>
</tr>
<tr>
<td>Italian humanitarian aid (Q4)</td>
<td>0.49</td>
<td>-0.19</td>
<td>0.10</td>
</tr>
<tr>
<td>Reinforcement of Italian humanitarian aid (Q5)</td>
<td>0.45</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Sending of Italian troops (Q6)</td>
<td>0.42</td>
<td>0.26</td>
<td>-0.19</td>
</tr>
<tr>
<td>Threat to peace in the world: Afghanistan (Q7)</td>
<td>-0.01</td>
<td>0.70</td>
<td>-0.10</td>
</tr>
<tr>
<td>Threat to peace in the world: China (Q8)</td>
<td>-0.17</td>
<td>-0.06</td>
<td>0.46</td>
</tr>
<tr>
<td>Threat to peace in the world: North Korea (Q9)</td>
<td>0.03</td>
<td>-0.07</td>
<td>0.49</td>
</tr>
<tr>
<td>Threat to peace in the world: Iran (Q10)</td>
<td>0.32</td>
<td>0.13</td>
<td>0.28</td>
</tr>
<tr>
<td>Threat to peace in the world: Israel (Q11)</td>
<td>-0.15</td>
<td>0.37</td>
<td>0.20</td>
</tr>
<tr>
<td>Threat to peace in the world: Syria (Q12)</td>
<td>0.21</td>
<td>-0.06</td>
<td>0.49</td>
</tr>
<tr>
<td>Threat to peace in the world: United States (Q13)</td>
<td>-0.27</td>
<td>0.18</td>
<td>0.20</td>
</tr>
</tbody>
</table>

✓ C1: role of Italy in Iraq
✓ C2: threat of terrorism in Italy
✓ C3: threat to peace in the world
PCA-F: Application [5]

plot of the students (components 1 & 2)

✓ positions and sizes consistent with the component interpretation
✓ a lot of DK / NA for Student n.17
✓ smaller sizes for students expressing radical positions
Three-way Principal Component Analysis for fuzzy valued data (3PCA-F)

Aim

to find parameter matrices and arrays such that the dissimilarity measure between observed and modelled data is minimized

But . . .

need for a dissimilarity measure between observed \( \tilde{X} \equiv (M_1, M_2, L, R) \) and estimated \( \tilde{X}^* \equiv (M_1^*, M_2^*, L^*, R^*) \) fuzzy valued arrays

Idea

since each observation unit can be represented as a hyperrectangle in \( \mathbb{R}^{JK} \), compute the dissimilarity measure by comparing all the \( 2^{JK} \) vertices
3PCA-F: (Squared) dissimilarity measure for fuzzy valued arrays

6–term simplified version

\[
d^2 \left( \tilde{X}, \tilde{X}^* \right) = \|M_{1a} - M_{1a}^*\|^2 + \|M_{2a} - M_{2a}^*\|^2 \\
-2tr \left[ (M_{1a} - M_{1a}^*) (L_a Y - L_a^* Y) \right] \\
+2tr \left[ (M_{2a} - M_{2a}^*) (R_a Z - R_a^* Z) \right] \\
+ \|L_a Y - L_a^* Y\|^2 + \|R_a Z - R_a^* Z\|^2
\]

where

- **Y** and **Z** \(\rightarrow\) diagonal matrices of order \(JK\) which help us to suitably take into account the decreasing shape functions \(L_{jk}(w)\) and \(R_{jk}(w)\) for the \(j\)-th fuzzy variable \((j = 1, \ldots, J)\) at the \(k\)-th occasion \((k = 1, \ldots, K)\)

- **H^L_v** and **H^L_v** \(\rightarrow\) diagonal matrices of order \(JK\) which help us to describe every vertex of the hyperrectangles, \(v = 1, \ldots, 2^{JK}\)
3-PCAF: Model

the PCA-F model (T3-based) can be formalized as follows

\((v = 1, \ldots, 2^{JK})\)

1. \((M_{1a} - L_a) H_v^L + (M_{2a} + R_a) H_v^R = (M_{1a}^* - L_a^*) H_v^L + (M_{2a}^* + R_a^*) H_v^R + E_a\)
2. \(M_{1a}^* = A_{M_1} G_a (C' \otimes B')\)
3. \(M_{2a}^* = A_{M_2} G_a (C' \otimes B')\)
4. \(L_a^* = A_L G_a (C' \otimes B')\)
5. \(R_a^* = A_R G_a (C' \otimes B')\)

where

- \(A_{M_1}, A_{M_2}\) → component matrices for the observation unit mode concerning the left modes, right modes, left spreads and right spreads, respectively
- \(A_L, A_R\) → component matrix for the variable mode
- \(B\) → component matrix for the occasion mode
- \(G_a\) → supermatrix resulting from core array \(G\)

Note: the CP-based version can be obtained replacing \(G_a\) with \(I_a\)
3-PCAF: model (Graphical interpretation)

\( i \)-th estimated hyperrectangle \((\text{trivial case with } J=1, K=2 \text{ and } P=Q=R=1)\)

- modes and vertices share the same \( B \) and \( C \) (and \( G \) in the T3 case)
- \( B \) and \( C \) (and \( G \)) found by making a sort of compromise between modes and spreads information
- different components for the observation unit mode (every entity of the fuzzy valued array is associated to a specific component matrix)
3PCA-F: Comments

- optimal solution by **ALS** algorithm with respect to $A_{M_1}, A_{M_2}, A_L, A_R, B, C$ and $G$
- **goodness of fit** index obtained comparing the explained sum of squares to the observed one
- 3PCA-F admits **rotations** (only in the T3 case)
- plotting procedure of the observation units as **low-dimensional hyperrectangles** as in PCA-F replacing $F$ with $F_a = (C \otimes B) G'_a$ (or $F_a = (C \otimes B) I'_a$ in the CP case)
- special cases:
  
  i. $L = R = 0$ \[\rightarrow\] three-way model for **interval valued** arrays (optimal solution found by performing ordinary three-way component model on $[M_{1a} \quad M_{2a}]$)
  
  ii. $L = R = 0$, $M_1 = M_2$ \[\rightarrow\] ordinary three-way component model for **single valued** arrays
3PCA-F: Imprecision + randomness

- $\tilde{X}$ is a **random sample** (with respect to the observation unit mode)
- it is interesting to assess the **uncertainty due to sampling** for $B$, $C$ (and $G$)
- possible way to do it by means of **resampling techniques** such as (non-parametric) bootstrap (following Kiers, 2004)
  - i. ML-based three-way methods exist (e.g., Bentler & Lee, 1978) in the single valued case
  - ii. they only analyze derived (covariance matrices) rather than original data, limiting the scope of the analysis
  - iii. they tend to have more problems in fitting them (to be sometimes instable, not to converge, etc.)

- able to determine probabilistic uncertainty estimates in terms of **percentile intervals** ($pi$'s) or **standard errors** ($se$'s)
let $B$, $C$ and $G_a$ be the optimal parameter matrices resulting from PCA-F applied to $\tilde{X}$

3PCA-F: Bootstrap procedure

i. generate a bootstrap sample of size $I$, say $\tilde{X}^b$,

ii. perform 3PCA-F on $\tilde{X}^b$ (same preprocessing and $P$, $Q$ and $R$); let $B^b$, $C^b$ and $G^b_a$ be the optimal parameter matrices

iii. repeat steps (i) and (ii) $B$ times to get $B^b$, $C^b$ and $G^b_a$ ($b = 1, \ldots, B$)

iv. compute the pi or se estimate of every element of every parameter matrix; for instance,

$$\hat{\text{se}}(c_{kr}) = \sqrt{\frac{1}{B} \sum_{b=1}^{B} \left( c^b_{kr} - \frac{1}{B} \sum_{n=1}^{B} c^n_{kr} \right)^2}$$
3PCA-F: ... Bootstrap (T3 case)

But ... 

attention must be paid to the non-uniqueness property (!)

i rotate $B^b$ and $C^b$ so that they resemble as much as possible $B$ and $C$ by $T^b$ and $U^b$ such that $\|B^bT^b - B\|^2$ and $\|C^bU^b - C\|^2$ are minimized

ii compensate these rotations into the core $(G^b_a \left(U^{b^{-1}} \otimes T^{b^{-1}}\right)'$)

iii (assuming that $G^b_a$ denotes the so-obtained rotated core), transform $G^b$ so that it becomes optimally similar to $G$ by $S^b$ such that $\|S^bG^b_a - G_a\|^2$ is minimized

iv compensate this rotation into the component matrices for the observation unit mode (in practice, this is not done since we do not use them)

Note: attention also in the CP case due to possible rescaling and joint permutation of the columns (Tucker congruence coefficient)
3PCA-F: Simulation experiment [1]

comparison between the performances of T3 on either $\frac{m_1 + m_2}{2}$ or $(m_1 - l) + (m_2 + r)$ and T3-based 3PCA-F

research questions:
- does the use of fuzzy valued data improve the quality of the results in terms of recovery of the parameter matrices and array?
- does the use of fuzzy valued data improve the quality of the measures of statistical validity provided by the bootstrap procedure?

construction of simulated data:
- randomly generated fuzzy population data with size $N = 1000$ and known T3 model structure and different:
  - numbers of variables and occasions (from $J = 10 = 10$ to $J = 40 = 40$)
  - numbers of components (from $P = Q = R = 2$ to $P = Q = R = 4$)
  - widths of the fuzzy data
  - kinds of core array (with or without zeroing an half of its elements)
  - levels of added noise
- from each of the above defined populations, random samples with size $I = 10$ or $I = 40$
- for each combination of all the design variables, three replications
3PCA-F: Simulation experiment [2]

results (recovery):
- Taking into account the rotational freedom, use of the Proportion of Recovery (PR) index; e.g., for $C$: 
  \[ PR_C = 1 - \frac{\|C - C^*\|^2}{\|C^*\|^2} \]
  where the obtained matrix $C$ is such that it resembles as much as possible the known in advance matrix $C^*$
- 3PCA-F worked better, or at least equally well, than T3’s (average PR values 0.865 and 0.846 or 0.861, respectively)

results (statistical validity):
- Taking into account the rotational freedom, use of the average se value ($B = 500$ bootstrap samples)
- Sample parameter matrices from 3PCA-F were rather accurate estimates of the population parameter matrices while those from T3’s were a little less accurate (average $\hat{se}$ values 0.033 and 0.040 or 0.039, respectively)

✓ the performance of 3PCA-F was the best one, if compared with those of T3’s
Problem

Bank indicators pertaining to the Italian branches of an important Italian bank are collected during years 2000 – 2005.

- A situation of **partial ignorance** occurs: due to privacy reasons, the information is imprecise.
- We do not exactly know the scores corresponding to each and every bank branch.
- Partial information with respect to some pre-specified geographical areas (provinces, regions or sub-regions) is available ($j$-th variable at $k$-th occasion):
  1. Average value $\mu_{ijk}$
  2. Min value $x_{ijk}$
  3. Max value $\bar{x}_{ijk}$
  4. Standard deviation $\sigma_{ijk}$
- Previous values computed with respect to the bank branches located in the geographical area at hand.
Application: Italian bank branches performance

Data array

- **three-way analysis on:**
  1. \( I = 52 \) geographical areas
  2. \( J = 7 \) bank indicators
  3. \( K = 6 \) years (2000 – 2005)

- **generic fuzzy valued datum:**

  \[
  \begin{align*}
  m_1 &= \mu_{ijk} - \sigma_{ijk} \\
  m_2 &= \mu_{ijk} + \sigma_{ijk} \\
  l &= (\mu_{ijk} - \sigma_{ijk}) - x_{ijk} \\
  r &= x_{ijk} - (\mu_{ijk} + \sigma_{ijk})
  \end{align*}
  \]

- **application** of 3PCA-F using T3 (CP too restrictive)

- **preprocessing:** centering across the observation units and scaling within the variables
Choice of $P$, $Q$ and $R$ ($P + Q + R \leq 9$) + simplicity

- **3PCA-F** with $P = Q = 3$ and $R = 2$

<table>
<thead>
<tr>
<th>$P$</th>
<th>$Q$</th>
<th>$R$</th>
<th>$P + Q + R$</th>
<th>Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>53.42%</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>66.89%</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>68.23%</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td>70.54%</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td>8</td>
<td>75.34%</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
<td>9</td>
<td>77.22%</td>
</tr>
</tbody>
</table>

- **Rotations to simplicity**
  1. **$B$ and $C$** first orthonormalized and then varimax rotated to simple structure
  2. Rotations compensated into the core matrix
  3. Resulting core $G_a$ (row-wise) orthonormalized (helpful in order to plot the observation units) and (row-wise) varimax rotated
  4. Rotations finally compensated into the component matrices for the observation unit mode

← parsimonious solution but tends to oversimplify and is not well interpretable
← optimal solution covering the most important structural aspects in the data
Interpretation [1]

- **Component matrix for the variables (\( \hat{s}e \) within parentheses with \( B = 1000 \))**

<table>
<thead>
<tr>
<th>Bank indicator</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loans / Deposits</td>
<td>0.46 (0.03)</td>
<td>0.48 (0.05)</td>
<td>0.09 (0.06)</td>
</tr>
<tr>
<td>Bad Debts / Loans</td>
<td>-0.76 (0.03)</td>
<td>0.03 (0.04)</td>
<td>0.05 (0.04)</td>
</tr>
<tr>
<td>Financial Intermediation and Sundry Incomes / Total Assets</td>
<td>0.14 (0.03)</td>
<td>0.57 (0.05)</td>
<td>0.03 (0.06)</td>
</tr>
<tr>
<td>Net Interest Income / Total Assets</td>
<td>-0.30 (0.03)</td>
<td>0.34 (0.04)</td>
<td>0.13 (0.05)</td>
</tr>
<tr>
<td>Administrative Costs / Total Assets</td>
<td>-0.30 (0.02)</td>
<td>0.57 (0.04)</td>
<td>-0.17 (0.04)</td>
</tr>
<tr>
<td>Return On Equity – ROE</td>
<td>-0.00 (0.01)</td>
<td>0.04 (0.02)</td>
<td>0.64 (0.03)</td>
</tr>
<tr>
<td>Return On Assets – ROA</td>
<td>-0.03 (0.01)</td>
<td>-0.05 (0.03)</td>
<td>0.73 (0.03)</td>
</tr>
</tbody>
</table>

- \( \checkmark \) C1: low credit risk; C2: stock exchange trading and credit; C3: branches performance
- \( \hat{s}e \)'s small (especially for C1); accurate estimate of the corresponding population parameters for the variable mode

- **Component matrix for the occasions (\( \hat{s}e \) within parentheses with \( B = 1000 \))**

<table>
<thead>
<tr>
<th>Year</th>
<th>C1</th>
<th>C2</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>0.48 (0.10)</td>
<td>0.20 (0.12)</td>
</tr>
<tr>
<td>2001</td>
<td>0.08 (0.10)</td>
<td>0.64 (0.13)</td>
</tr>
<tr>
<td>2002</td>
<td>-0.07 (0.15)</td>
<td>0.73 (0.19)</td>
</tr>
<tr>
<td>2003</td>
<td>0.55 (0.08)</td>
<td>-0.07 (0.10)</td>
</tr>
<tr>
<td>2004</td>
<td>0.51 (0.06)</td>
<td>-0.09 (0.07)</td>
</tr>
<tr>
<td>2005</td>
<td>0.43 (0.07)</td>
<td>-0.02 (0.08)</td>
</tr>
</tbody>
</table>

- \( \checkmark \) C1: ‘standard’ years; C2: Twin Towers attack and its consequences
- \( \hat{s}e \)'s slightly bigger than those for \( B \); little less accurate estimate
**Interpretation [2]**

- core matrix $G_a$ (see within parentheses with $B = 1000$)

<table>
<thead>
<tr>
<th></th>
<th>$q = 1$</th>
<th>$q = 2$</th>
<th>$q = 3$</th>
<th>$q = 1$</th>
<th>$q = 2$</th>
<th>$q = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p = 1$</td>
<td>0.76 (0.03)</td>
<td>-0.03 (0.03)</td>
<td>0.07 (0.04)</td>
<td>0.64 (0.03)</td>
<td>0.01 (0.04)</td>
<td>-0.01 (0.02)</td>
</tr>
<tr>
<td>$p = 2$</td>
<td>-0.01 (0.02)</td>
<td>0.05 (0.05)</td>
<td>-0.27 (0.10)</td>
<td>0.03 (0.03)</td>
<td>0.06 (0.08)</td>
<td>-0.96 (0.06)</td>
</tr>
<tr>
<td>$p = 3$</td>
<td>-0.00 (0.01)</td>
<td>0.72 (0.03)</td>
<td>0.36 (0.06)</td>
<td>-0.02 (0.02)</td>
<td>0.59 (0.05)</td>
<td>0.03 (0.02)</td>
</tr>
</tbody>
</table>

- it allows us to make a summary description of the components for the bank branches:
  - $p = 1$ $a_{i1}$’s high if branches present a low credit risk during all the years under investigation ($g_{111}$ & $g_{112}$)
  - $p = 2$ $a_{i2}$’s high if branches have poor performance increasing their pains due to the Twin Towers attack ($g_{231}$ & $g_{232}$)
  - $p = 3$ $a_{i3}$’s high if branches follow a corporate policy related to high stock exchange trading and credit ($g_{321}$ & $g_{322}$) and if have good performance during the ‘standard’ years ($g_{331}$)

- very stable core elements well detecting all the strong population interactions among the different components (except for $g_{231}$)
Interpretation [3]

low dimensional plot of the bank branches (left side: first 26; right side: last 26)

✓ (positions) C1 also interpreted as duality between Northern and Southern Italy (branches from Northern (Southern) Italy characterized by low (high) credit risk)

✓ (sizes) biggest rectangles for ‘Rome’ and ‘Sicily (others)’ because their bank branches have very heterogeneous values
Outline

1. Uncertainty
   - Some clarifications

2. Single valued case
   - Membership functions
   - Data
   - Models

3. Interval valued case
   - Membership functions
   - Data
   - Intermezzo: two-way case
   - Models
   - Application: Air pollution in Rome

4. Fuzzy valued case
   - Membership functions
   - Data
   - Intermezzo: two-way case
   - Models
   - Application: Italian bank branches performance

5. Conclusion

6. References
Conclusion

- component models for uncertain data
- imprecision managed by interval or fuzzy valued data
- randomness managed by non parametric bootstrap

future research: component models ...

... based on fuzzy / interval arithmetics
... with fuzzy component matrix for the observation unit
mode \( \tilde{A} \) (fuzzy case)
... able to manage both imprecision and randomness analyzing indirect data (based on the concept of fuzzy random variables, e.g., Puri & Ralescu, 1986)
... able to manage both imprecision and randomness analyzing direct data (starting point: random single valued case)
References [1]

References [2]

References [3]

Appendix

Zoom of Figure ‘about 3’
Zoom of Figure ‘temperature in Rome’
Zoom of Figure ‘I feel good’

I feel good (0–10 scale)
Zoom of Figure ‘wait me for 5 minutes’
Appendix

Zoom of Figure ‘membership function’

single valued datum

\[ \mu_x(x) \]

\[ U \]

\[ m \]
Zoom of Figure ‘about 3’
Appendix

Zoom of Figure ‘temperature in Rome’
Zoom of Figure ‘I feel good’
Zoom of Figure ‘wait me for 5 minutes’
interval valued datum

$\mu_x(x)$

$m_1$

$m_2$

$U$
Zoom of Figure ‘about 3’

\[ \mu_{\text{about 3}}(x) \]

\[ m_1 - l \quad m_1: m_2 \quad m_1 + r \]

\[ 0.0 \quad 0.2 \quad 0.4 \quad 0.6 \quad 0.8 \quad 1.0 \]

\[ 0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \]

\[ U \]
Zoom of Figure ‘temperature in Rome’
Zoom of Figure ‘I feel good’

I feel good (0–10 scale)
Zoom of Figure ‘wait me for 5 minutes’
Appendix

Zoom of Figure ‘membership function’

fuzzy valued datum

\[ \mu_{\tilde{x}}(x) \]

\[ m_1 \quad m_1 - l \quad m_2 \quad m_2 + r \]

\[ U \]
Appendix 1. Questionnaire

Q1: Today, to what extent do you agree that the military intervention of the United States and their allies in Iraq was justified?

- Null
- Extremely Low
- Very Low
- Low
- Medium
- High
- Very High
- Extremely High
- Total
- DK/NA

Q2: What would you say about the consequences of the war in Iraq with respect to the international scene?

- Negligible
- Minor
- Moderate
- Severe
- Catastrophic
- DK/NA

Q3: How would you evaluate the threat of terrorism in Italy?*

Absolutely weak
Absolutely strong

Tell me your opinion about each of the following propositions concerning the after-war in Iraq:

Q4: The participation of humanitarian aid from Italy to Iraq.

Absolutely unfavorable
Absolutely favorable

Q5: The reinforcement of humanitarian aid from Italy to Iraq.

Absolutely unfavorable
Absolutely favorable

Q6: The sending of Italian troops in order to maintain peace in Iraq.

Absolutely unfavorable
Absolutely favorable

* Make a vertical mark along the segment in the position where you wish to place your opinion.

For each of the following countries, tell me in your opinion, the degree to which its policy presents a threat to peace in the world?

<table>
<thead>
<tr>
<th></th>
<th>Afghanistan</th>
<th>Low</th>
<th>Possible</th>
<th>Substantial</th>
<th>High</th>
<th>DK/NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q8</td>
<td>China</td>
<td>Low</td>
<td>Possible</td>
<td>Substantial</td>
<td>High</td>
<td>DK/NA</td>
</tr>
<tr>
<td>Q9</td>
<td>North Korea</td>
<td>Low</td>
<td>Possible</td>
<td>Substantial</td>
<td>High</td>
<td>DK/NA</td>
</tr>
<tr>
<td>Q10</td>
<td>Iran</td>
<td>Low</td>
<td>Possible</td>
<td>Substantial</td>
<td>High</td>
<td>DK/NA</td>
</tr>
<tr>
<td>Q11</td>
<td>Israel</td>
<td>Low</td>
<td>Possible</td>
<td>Substantial</td>
<td>High</td>
<td>DK/NA</td>
</tr>
<tr>
<td>Q12</td>
<td>Syria</td>
<td>Low</td>
<td>Possible</td>
<td>Substantial</td>
<td>High</td>
<td>DK/NA</td>
</tr>
<tr>
<td>Q13</td>
<td>United States</td>
<td>Low</td>
<td>Possible</td>
<td>Substantial</td>
<td>High</td>
<td>DK/NA</td>
</tr>
</tbody>
</table>

* Make a vertical mark along the segment in the position where you wish to place your opinion.