Simplicity transformations for three-way arrays with symmetric slices

Jorge Tendeiro

University of Groningen

17 June 2011 / Classification Society Annual Meeting

Outline

- Introducing three-way arrays
- Methods to analyze three-way arrays
- Simplifying three-way arrays
- Maximal simplicity
- 5 Example of application: typical rank
- 6 Conclusions. Considerations. Developments

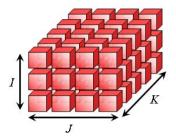
Introducing three-way arrays

- Methods to analyze three-way arrays
- Simplifying three-way arrays
- Maximal simplicity

Example of application: typical rank

6 Conclusions, Considerations, Developments

Definition



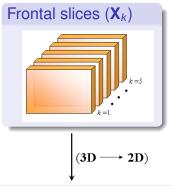
Idea

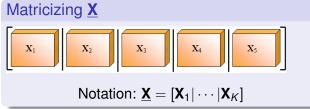
- three-way arrays: generalize matrix structure to 3D
- loaf-of-bread structure

Examples of three-way data

- different anxiety measures, different circumstances, various subjects
- sales of different products, in different shops, in different weeks
- job requirements for various jobs, according to various job analysts

Unfolding a three-way array





Introducing three-way arrays

Methods to analyze three-way arrays

- Simplifying three-way arrays
- Maximal simplicity

Example of application: typical rank

Conclusions, Considerations, Developments

PCA

X : matrix of order $I \times J$ (I=subjects, J=variables) Goal: representation of variables in low-space dimension.

$$x_{ij} = \sum_{r=1}^{R} a_{ir} b_{jr} + e_{ij}$$

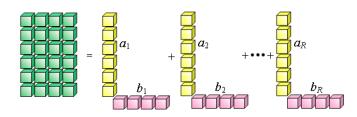
- x_{ij} = score of subject i on variable j
- a_{ir} = score of subject i on component r
- b_{jr} = loading of variable j on component r
- e_{ij} = residual error

PCA – other formulation

$$\mathbf{X} = \sum_{r=1}^{R} (\mathbf{a}_r \circ \mathbf{b}_r) + \mathbf{E}$$



- $\mathbf{a}_r \circ \mathbf{b}_r$: rank-1 matrix
- PCA decomposes X as a sum of rank-1 matrices
- rank(\mathbf{X}): minimum R such that $\mathbf{E} \equiv \mathbf{0}$



CANDECOMP/PARAFAC (CP)

 \underline{X} : array of order $I \times J \times K$ (I=subjects, J=variables, K=situations) Goal: find components for subjects, variables and situations.

$$x_{ijk} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr} + e_{ijk},$$

- x_{ijk} = score of subject i on variable j on situation k
- a_{ir} = score of subject i on component r
- b_{jr} = loading of variable j on component r
- c_{kr} = loading of situation k on component r
- e_{ijk} = residual error

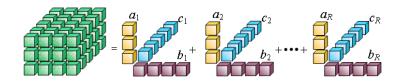
CP – other formulation

$$\underline{\mathbf{X}} = \sum_{r=1}^{R} (\mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r) + \underline{\mathbf{E}}$$

▶ PCA

D...

- $\mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$: rank-1 array
- CP decomposes <u>X</u> as a sum of rank-1 arrays
- rank(\mathbf{X}): minimum R such that $\mathbf{E} \equiv \mathbf{0}$



Tucker3

 $\underline{\mathbf{X}}$: array of order $I \times J \times K$ (I=subjects, J=variables, K=situations) Goal: find components for subjects, variables and situations.

$$x_{ijk} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} g_{pqr} \left(a_{ip} b_{jq} c_{kr} \right) + e_{ijk},$$

- x_{ijk} = score of subject i on variable j on situation k
- a_{ip} = score of subject i on component p
- b_{jq} = loading of variable j on component q
- c_{kr} = loading of situation k on component r
- g_{pqr} = weight (core array \mathbf{G} , order $P \times Q \times R$)
- e_{iik} = residual error

Tucker3 – other formulations

$$oldsymbol{\underline{X}} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} g_{pqr} \left(\mathbf{a}_{p} \circ \mathbf{b}_{q} \circ \mathbf{c}_{r}
ight) + \underline{\mathbf{E}}$$

- $\mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$: rank-1 array
- Tucker3 decomposes X as a sum of rank-1 arrays
- rank(X)≤ PQR (usually rank(X) ≪ PQR)

Formula using unfolded notation

$$\underline{\mathbf{X}} (I \times J \times K) \longrightarrow \mathbf{X} = [\mathbf{X}_1 | \mathbf{X}_2 | \cdots | \mathbf{X}_K]$$
 (fitted part) $\mathbf{G} (P \times Q \times R) \longrightarrow \mathbf{G} = [\mathbf{G}_1 | \mathbf{G}_2 | \cdots | \mathbf{G}_R]$

$$X - \Delta G(C' \otimes B')$$

Tucker3 – seeing CP as particular situation

 Tucker3 reduces to Candecomp/Parafac when the core array has a super-diagonal form:

 only interactions between corresponding components are accounted for in CP

Tucker3 – freedom of rotation

PCA's freedom of rotation (motivation)

S nonsingular

$$\mathbf{X} = \mathbf{A}\mathbf{B}'$$
 $= (\mathbf{A}\mathbf{S})(\mathbf{S}^{-1}\mathbf{B}')$

Tucker3's freedom of rotation

S, T, U nonsingular

$$\begin{array}{cccc} \textbf{A} & \longrightarrow & \textbf{A}(\textbf{S}')^{-1} \\ \textbf{B} & \longrightarrow & \textbf{B}(\textbf{T}')^{-1} \\ \textbf{C} & \longrightarrow & \textbf{C}(\textbf{U}')^{-1} \\ \textbf{G}_a = [\textbf{G}_1|\cdots|\textbf{G}_R] & \longrightarrow & \textbf{S}'\textbf{G}_a(\textbf{U}\otimes\textbf{T}) \end{array}$$

Tucker transformation

- Simplifying three-way arrays

Goal

Find suitable linear combinations of frontal (and/or lateral and/or horizontal) slices that allow transforming $\underline{\mathbf{X}}$ into an "equivalent" array with many zero entries.

Formally: S, T, U=?:
$$H = SX(U \otimes T)$$

- many zero entries = few nonzero entries
- weight of $\underline{\mathbf{H}} = \#$ nonzero entries of $\underline{\mathbf{H}}$

Goal

Find suitable linear combinations of frontal (and/or lateral and/or horizontal) slices that allow transforming $\underline{\mathbf{X}}$ into an "equivalent" array with many zero entries.

Formally: S, T, U=?:
$$H = SX(U \otimes T)$$

- many zero entries = few nonzero entries
- weight of **H** = # nonzero entries of **H**



Facilitate interpretation of 3PCA decompositions

Example: rotate **G** so that several entries become zero

less interactions of components to account for during interpretion of 3PCA

- Constrained 3PCA: distinguish between tautologies and non-trivial models
- Mathematical applications: typical rank, maximal rank



Facilitate interpretation of 3PCA decompositions

Example: rotate **G** so that several entries become zero

less interactions of components to account for during interpretion of 3PCA

- Constrained 3PCA: distinguish between tautologies and non-trivial models
- Mathematical applications: typical rank, maximal rank



Facilitate interpretation of 3PCA decompositions

Example: rotate **G** so that several entries become zero

less interactions of components to account for during interpretion of 3PCA

- Constrained 3PCA: distinguish between tautologies and non-trivial models
- Mathematical applications: typical rank, maximal rank

Some examples (I-III)

- Cohen (1974, 1975), MacCallum (1976), Kroonenberg (1983): "diagonalize" frontal slices of \mathbf{G} (P = Q)
- Kiers (1992): "super-diagonalize" \mathbf{G} (P = Q = R)
- Kiers (1998): SIMPLIMAX $\underline{\mathbf{G}} \longrightarrow \text{minimize ssq} (m \text{ smallest elements})$

$$\underline{\mathbf{X}}$$
 of order $P \times Q \times R$, $P = QR$

Example: $\underline{\mathbf{X}}$ of order $6 \times 3 \times 2$

$$\underline{\boldsymbol{X}} \, \longrightarrow \left[\begin{array}{ccc|ccc|c} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] = \boldsymbol{X}^{-1}\boldsymbol{X}(\boldsymbol{I}_2 \otimes \boldsymbol{I}_3)$$

Some examples (II-III)

X of order
$$P \times Q \times R$$
, $P = QR - 1$

Murakami, Ten Berge & Kiers (1998)

Example: **X** of order $5 \times 3 \times 2$

$$\underline{\mathbf{X}} \, \longrightarrow \left[\begin{array}{ccc|ccc|c} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \mu_1 & 0 & 0 & 0 & \mu_2 & 0 \end{array} \right]$$

Some examples (III-III)

X of order $P \times Q \times 2$

P > Q: Ten Berge & Kiers (1999)

$$\underline{\mathbf{X}} \longrightarrow \left[\begin{array}{c|c} \mathbf{I}_Q & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I}_Q \end{array} \right]$$

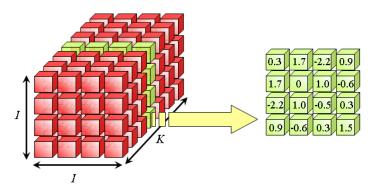
• P = Q: Rocci & Ten Berge (2002) Example: $\underline{\mathbf{X}} = [\mathbf{X}_1 | \mathbf{X}_2]$ of order $3 \times 3 \times 2$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \mu_1 & 0 & 0 & \mu_2 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \mu \\ 0 & 0 & 1 & 0 & -\mu & 0 \end{bmatrix}$$

$$(\mathbf{X}_1^{-1}\mathbf{X}_2 \text{ has real eigs.}) \qquad (\mathbf{X}_1^{-1}\mathbf{X}_2 \text{ has complex eigs.})$$

Our goal: simplifying arrays with SYMMETRIC slices

Example: set of similarity matrices over time



Symmetric-slice arrays

- $\underline{\mathbf{X}} = [\mathbf{X}_1 | \cdots | \mathbf{X}_K]$: order $I \times I \times K$
 - ▶ <u>assume</u>: \mathbf{X} is randomly sampled from a continuous distribution with symmetry constraint (\mathbf{X}_k symmetric, $\forall k$)
 - ightharpoonup slices \mathbf{X}_k linearly independent
 - ▶ number of slices: $K = 1, 2, ..., \underbrace{\frac{I(I+1)}{2}}_{K_{\text{max}}}$
- symmetry-preserving transformation of X
 - ▶ $S_{I \times I}$, $U_{K \times K}$ nonsingular

$$\mathbf{H}_{I} = \mathbf{S}'\left(\sum_{k} u_{kl}\mathbf{X}_{k}\right)\mathbf{S}, \quad I = 1, 2, \dots, K$$

- ▶ GOAL: introduce as many zeros in **H** as possible
- Orthogonal Complement Method: "symmetric" version

Symmetric slice $I \times I \times K_{max}$ arrays

- $\{\text{frontal slices}\} = \text{basis for the space of symmetric } I \times I \text{ matrices}$
- simple basis for the same space (Rocci & Ten Berge(1994)): (notation: \mathbf{e}_i = column i of \mathbf{I}_i)

$$\mathbf{e}_i \mathbf{e}_i', \quad i = 1, \dots, I$$

 $\mathbf{e}_i \mathbf{e}_i' + \mathbf{e}_j \mathbf{e}_i', \quad 1 \leqslant i < j \leqslant I$

Example: I = 3

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

frontal slice mix suffices

Symmetric slice $3 \times 3 \times K$ arrays

- $K_{\text{max}} = 6$, so K = 1, 2, 3, 4, 5, 6
- $3 \times 3 \times 6$: done (K_{max} situation)
- 3 × 3 × 1: use EVD

$$\underline{\boldsymbol{X}} \longrightarrow \left[\begin{array}{ccc} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{array} \right]; \text{ if } d_2d_3 < 0: \left[\begin{array}{ccc} d_1 & 0 & 0 \\ 0 & 0 & 2d_2 \\ 0 & 2d_2 & 0 \end{array} \right]$$

• $3 \times 3 \times 5$ = orthogonal complement of $3 \times 3 \times 1$

Conclusion for $3 \times 3 \times 5$:

- weight 10 is always possible
- if $\underline{\mathbf{X}}^c$ has eigenvalues of both signs then weight 9 is possible

Symmetric slice $3 \times 3 \times K$ arrays

- 3 × 3 × 2: see EVD(**X**₁⁻¹**X**₂)
 - real eigenvalues

$$\underline{\mathbf{X}} \longrightarrow \left[\begin{array}{ccc|c} 0 & 0 & 0 & \beta & 0 & 0 \\ 0 & \alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right]; \text{ also: } \left[\begin{array}{ccc|c} -\alpha & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array} \right]$$

complex eigenvalues

$$\underline{\mathbf{X}} \longrightarrow \left[\begin{array}{ccc|ccc} -\alpha & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 & 1 & 0 \end{array} \right]$$

Conclusion for $3 \times 3 \times 2$:

- weight 5 is always possible
- if $X_1^{-1}X_2$ has real eigenvalues then weight 4 is possible

Symmetric slice $3 \times 3 \times K$ arrays

• $3 \times 3 \times 4 = \text{orthogonal complement of } 3 \times 3 \times 2$

$$\begin{bmatrix}
1 & 0 & 0 & | & 1 & 0 & 0 & | & 0 & 1 & 0 & | & 0 & 0 & 1 \\
0 & \alpha & 0 & | & 0 & 0 & 0 & | & 1 & 0 & 0 & | & 0 & 0 & 0 \\
0 & 0 & 0 & | & 0 & \delta \alpha & | & 0 & 0 & 0 & | & 1 & 0 & 0
\end{bmatrix}$$



 $(\delta = 1/-1)$ in real/complex case)

Conclusion for $3 \times 3 \times 4$:

- weight 8 is always possible
- 3 × 3 × 3: still open!
 - when a $3 \times 3 \times 3$ array has an orthogonal complement, it is also $3 \times 3 \times 3 \dots$
 - simulation: a weight 9 pattern seems to be possible almost 90% of the times
 - to be continued (...)

Symmetric slice $4 \times 4 \times K$ arrays

- $K_{\text{max}} = 10$, so $K = 1, 2, \dots, 8, 9, 10$
- $4 \times 4 \times 10$: done (K_{max} situation)
- 4 × 4 × 1: use EVD

in general:
$$\underline{\mathbf{X}}$$
 \longrightarrow
$$\begin{bmatrix} d_1 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_4 \end{bmatrix}$$
if $d_1, d_2, d_3 > 0$

$$d_4 < 0 \longrightarrow \begin{bmatrix} d_1 & 0 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_4 \end{bmatrix}$$
if $d_1, d_3 > 0$

$$d_2, d_4 < 0 \longrightarrow \begin{bmatrix} 0 & 2d_1 & 0 & 0 \\ 2d_1 & 0 & 0 & 0 \\ 0 & 0 & 2d_3 & 0 \end{bmatrix}$$

$$= \text{orthogonal complement of } 4 \times 4 \times 1$$

- $4 \times 4 \times 9$ = orthogonal complement of $4 \times 4 \times 1$
 - weight 18 is always possible
 - depending on the signs of eigs(X^c) we can have weight 17 or 16

Symmetric slice $4 \times 4 \times K$ arrays

- $4 \times 4 \times 2$: see EVD($\mathbf{X}_1^{-1}\mathbf{X}_2$)
 - real eigenvalues: weight 6

$$\left[\begin{array}{ccc|ccc|ccc|ccc|ccc|ccc|ccc|} 0 & 0 & 0 & 0 & \gamma & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta & 0 & 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

one pair of complex eigenvalues: weight 7

$$\begin{bmatrix} \alpha & 0 & 0 & 0 & | & \gamma & 0 & 0 & 0 \\ 0 & \beta & 0 & 0 & | & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & | & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & | & 0 & 0 & 1 & 0 \end{bmatrix}$$

two pairs of complex eigenvalues: weight 8

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \gamma & 0 & 0 \\ 0 & -1 & 0 & 0 & \gamma & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Symmetric slice $4 \times 4 \times K$ arrays

- $4 \times 4 \times 8$ = orthogonal complement of $4 \times 4 \times 2$
 - any symmetric slice 4 x 4 x 8 array can almost surely be simplified into one out of two weight 18 arrays

Example: one of the targets

Introducing three-way arrays

Methods to analyze three-way arrays

- Simplifying three-way arrays
- Maximal simplicity

Example of application: typical rank

6 Conclusions. Considerations. Developments

Maximal simplicity

<u>Question</u>: can simpler targets be found for the cases previously presented?

Answer

- $3 \times 3 \times K$ for K = 1, 2, 4, 5, 6: NO (proved)
- $4 \times 4 \times K$ for K = 8, 9: NO(?) (simulation)

Maximal simplicity

<u>Question</u>: can simpler targets be found for the cases previously presented?

Answer:

- $3 \times 3 \times K$ for K = 1, 2, 4, 5, 6: NO (proved)
- $4 \times 4 \times K$ for K = 8, 9: NO(?) (simulation)

Introducing three-way arrays

- Methods to analyze three-way arrays
- Simplifying three-way arrays
- Maximal simplicity

5 Example of application: typical rank

6 Conclusions. Considerations. Developments

Example of application: typical rank

- $\underline{\mathbf{X}}$: symmetric slice $3 \times 3 \times 4$ array
- Ten Berge et al. (2004)

typical rank (
$$\underline{\mathbf{X}}$$
)= $\{4,5\}$

rank=4?, rank=5?
 Check if roots of a certain fourth degree polynomial are real and distinct.

Using $\underbrace{\circ 3 \times 3 \times 4 \text{ simple form}}$, and applying the same reasoning as in Ten Berge et al. (2004), we conclude that:

- rank ($\underline{\mathbf{X}}$)=4 <u>iif</u> $\delta = 1$ and $\alpha > 0$ (and rank is 5 otherwise)
- a CP decomposition is now straightforward

Example: rank=4

$$\mathbf{A} = \left[\begin{array}{cccc} 1 & 1 & 1 & 1 & 0.5 &$$

Example of application: typical rank

- $\underline{\mathbf{X}}$: symmetric slice $3 \times 3 \times 4$ array
- Ten Berge et al. (2004)

typical rank (
$$\underline{\mathbf{X}}$$
)= $\{4,5\}$

 rank=4?, rank=5?
 Check if roots of a certain fourth degree polynomial are real and distinct.

Using $3 \times 3 \times 4$ simple form, and applying the same reasoning as in Ten Berge et al. (2004), we conclude that:

- rank ($\underline{\mathbf{X}}$)=4 $\underline{\text{iif}}$ $\delta=1$ and $\alpha>0$ (and rank is 5 otherwise)
- a CP decomposition is now straightforward

Example: rank=4

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0.5 &$$

Conclusions

- simplification achieved for some types of arrays with symmetric frontal slices; closed form rotation matrices available
- maximal simplicity achieved (mathematically proved or empirically verified via SIMPLIMAX)
- typical rank considerations come as nice follow-ups

Considerations

- 3PCA core arrays are not "randomly sampled from a continuous distribution", but do behave as if they were
- valid contribution for Matrix Theory: simultaneous reduction of more than a pair of matrices to sparse forms is scarce

Developments

- extend results to other orders
- address issues like: maximal simplicity, typical rank

Conclusions

- simplification achieved for some types of arrays with symmetric frontal slices; closed form rotation matrices available
- maximal simplicity achieved (mathematically proved or empirically verified via SIMPLIMAX)
- typical rank considerations come as nice follow-ups

Considerations

- 3PCA core arrays are not "randomly sampled from a continuous distribution", but do behave as if they were
- valid contribution for Matrix Theory: simultaneous reduction of more than a pair of matrices to sparse forms is scarce

Developments

- extend results to other orders
- address issues like: maximal simplicity, typical rank

Conclusions

- simplification achieved for some types of arrays with symmetric frontal slices; closed form rotation matrices available
- maximal simplicity achieved (mathematically proved or empirically verified via SIMPLIMAX)
- typical rank considerations come as nice follow-ups

Considerations

- 3PCA core arrays are not "randomly sampled from a continuous distribution", but do behave as if they were
- valid contribution for Matrix Theory: simultaneous reduction of more than a pair of matrices to sparse forms is scarce

Developments

- extend results to other orders
- address issues like: maximal simplicity, typical rank