Simplicity transformations for three-way arrays with symmetric slices

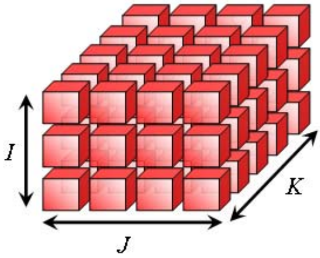
Jorge Tendeiro

University of Groningen

12 December 2007 / IOPS Conference

Outline

Definition



Idea

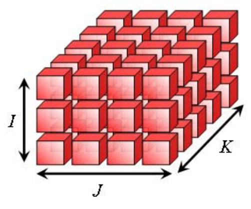
- three-way arrays: generalize matrix structure to 3D
- loaf-of-bread structure

Examples of three-way data

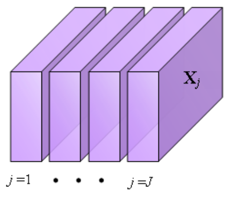
- different anxiety measures, different circumstances, various subjects
- sales of different products, in different shops, in different weeks
- job requirements for various jobs, according to various job analysts

SLICES of a three-way array

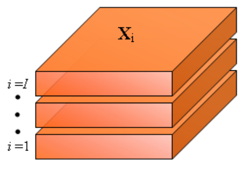
Three-way array



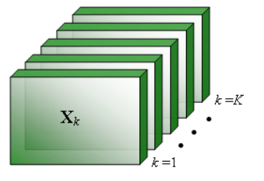
Lateral slices ($\mathbf{X}_j$)



Horizontal slices ($\mathbf{X}_i$)

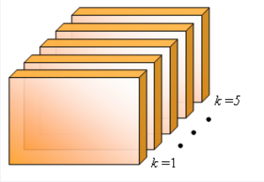


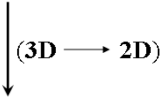
Frontal slices ($\mathbf{X}_k$)



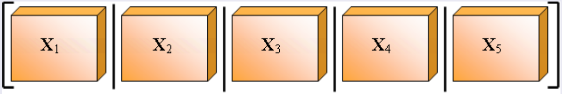
Unfolding a three-way array

Frontal slices ($\mathbf{X}_k$)





Matricizing $\underline{\mathbf{X}}$



PCA

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

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

▸ ...

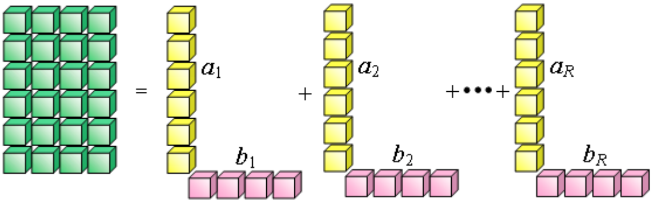
- x_{ij} = score of subject i on variable j
- a_{ir} = score of subject i on component r
- b_{jr} = loading of variable j on component r
- e_{ij} = residual error

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



CANDECOMP/PARAFAC (CP)

$\underline{\mathbf{X}}$: array of order $I \times J \times K$ (I=subjects, J=variables, K=situations)

Goal: find components for subjects, variables and situations.

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

PCA

...

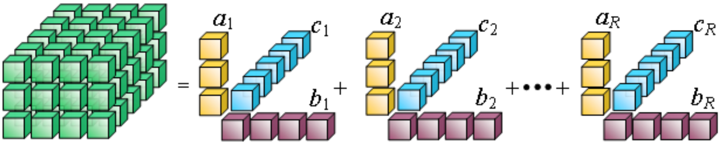
- x_{ijk} = score of subject i on variable j on situation k
- a_{ir} = score of subject i on component r
- b_{jr} = loading of variable j on component r
- c_{kr} = loading of situation k on component r
- e_{ijk} = residual error

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
...

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



Tucker3

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

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

▸ CP

- 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 **G**, order $P \times Q \times R$)
- e_{ijk} = residual error

Tucker3 – other formulations

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

▸ CP

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

Formula using unfolded notation

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

$$\mathbf{X} = \mathbf{AG}(\mathbf{C}' \otimes \mathbf{B}')$$

Tucker3 – seeing CP as particular situation

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

$$\underline{\mathbf{G}} = \left[\begin{array}{cccc|cccc|c|cccc} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & & 0 & 0 & \cdots & 1 \end{array}\right]$$

- only interactions between corresponding components are accounted for in CP

Tucker3 – freedom of rotation

PCA's freedom of rotation (motivation)

S nonsingular

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

Tucker3's freedom of rotation

S nonsingular

$$\begin{aligned}\mathbf{X} &= \mathbf{A}\mathbf{G}(\mathbf{C}' \otimes \mathbf{B}') \\ &= (\mathbf{A}\mathbf{S})((\mathbf{S})^{-1}\mathbf{G})(\mathbf{C}' \otimes \mathbf{B}')\end{aligned}$$

- same applies to **B** and **C**

Tucker3 – illustration (Kiers & Van Mechelen (2001))

X=data set of...

- 6 individuals: Anne, Bert, Claus, Dolly, Edna, Frances
- 5 response variables: emotional, sensitive, caring, thorough, accurate
- 4 different situations: doing an exam, giving a speech, family picnic, meeting a new date

Component matrix **A**

Individual	Femininity	Masculinity
Anne	1.0	0.0
Bert	0.0	1.0
Claus	0.0	1.0
Dolly	1.0	0.0
Edna	0.5	0.5
Frances	1.0	0.0

Tucker3 – illustration (Kiers & Van Mechelen (2001))

Component matrix **B**

Response	Emotionality	Conscientiousness
Emotional	1.0	0.0
Sensitive	1.0	0.0
Caring	0.6	0.4
Thorough	0.0	1.0
Accurate	0.0	1.0

Component matrix **C**

Situation	Performance situations	Social situations
Doing an exam	1.0	0.0
Giving a speech	0.8	0.2
Family picnic	0.0	1.0
Meeting a new date	0.3	1.2

Tucker3 – illustration (Kiers & Van Mechelen (2001))

Core array <u>**G**</u>

	Performance situations	
	Emotionality	Conscientiousness
Femininity	0.0	3.0
Masculinity	0.0	2.0

	Social situations	
	Emotionality	Conscientiousness
Femininity	3.0	0.0
Masculinity	1.0	1.0

Simplify three-way arrays

Goal

S, **T**, **U**=?: $\mathbf{H} = \mathbf{SX}(\mathbf{U} \otimes \mathbf{T})$

$\downarrow$

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

Why?

Statistical reasons:

- Tucker3: simpler core $\underline{\mathbf{G}} \Longrightarrow$ usually simpler interpretation
- constrained Tucker3: distinguish between tautology and non-trivial model

Mathematical reasons:

- typical rank, maximal rank

Some examples (I-II)

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

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

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

Some examples (II-II)

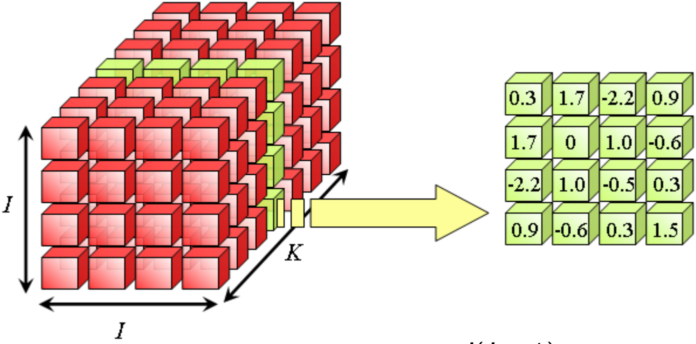
$\underline{\mathbf{X}}$ of order $P \times Q \times R$, $P = QR - 1$

Murakami, Ten Berge & Kiers (1998)
Example: $\underline{\mathbf{X}}$ of order $5 \times 3 \times 2$

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

Our goal: simplifying arrays with SYMMETRIC slices

Example: set of correlation matrices over time



Number of symmetric slices: $K = 1, \ldots, \underbrace{\frac{I(I+1)}{2}}_{K_{\max}}$.

Some results proven

Simplification achieved for:

- $3 \times 3 \times K$ when $K = 1, 2, 4, 5, 6$
- $4 \times 4 \times K$ when $K = 1, 2, 8, 9, 10$
- $I \times I \times 1$
- $I \times I \times (K_{\text{max}} - 1)$
- $I \times I \times K_{\text{max}}$

Example: symmetric slice array $3 \times 3 \times 4$

▸ ...

$$\left[\begin{array}{ccc|ccc|ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & \mu_1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu_2 & 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right]$$

Some results proven

Maximal simplicity

- proved for all $3 \times 3 \times K$ presented
- simulations using SIMPLIMAX (Kiers, 1998) seem to confirm maximal simplicity for the targets deduced for $4 \times 4 \times K$ (ongoing)

Typical rank

Rules-of-thumb were deduced concerning inspection of typical rank for $3 \times 3 \times K$, $K \neq 3$ (completion of Ten Berge, Sidiropoulos & Rocci, 2004)

- example ▸ $3 \times 3 \times 4$: rank is 4 iff $\mu_1, \mu_2 > 0$, otherwise is 5

Conclusions, developments

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

Developments

- extend results to other orders
- if possible, use procedures to address issues like: maximal simplicity, typical rank