Estimation in multi-level multivariate model

Katarzyna Filipiak[1] and Daniel Klein[2]

[1]Institute of Mathematics, Poznań University of Technology, Poland

[2]Faculty of Science, P.J. Šafárik University in Košice, Slovakia

MET 2019

Vector of observations

n – the number of independent objects, for which we measure a feature

Vector of observations

n – the number of independent objects, for which we measure a feature

$$\mathbf{y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}$$

Vector of observations

n – the number of independent objects, for which we measure a feature

$$\mathbf{y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \quad \sim \quad N_n(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n)$$

Vector of observations

n – the number of independent objects, for which we measure a feature

$$\mathbf{y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \quad \sim \quad N_n(\boldsymbol{\mu}, \, \sigma^2 \mathbf{I}_n)$$

$\boldsymbol{\mu}$ – mean vector

Vector of observations

n – the number of independent objects, for which we measure a feature

$$\mathbf{y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \quad \sim \quad N_n(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n)$$

$\boldsymbol{\mu}$ – mean vector: $\boldsymbol{\mu} = \mu \mathbf{1}_n$

Vector of observations

n – the number of independent objects, for which we measure a feature

$$\mathbf{y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \quad \sim \quad N_n(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n)$$

$\boldsymbol{\mu}$ – mean vector: $\boldsymbol{\mu} = \mu \mathbf{1}_n$

$$\boldsymbol{\mu} = \mathbf{A}\boldsymbol{\gamma}$$

Vector of observations

n – the number of independent objects, for which we measure a feature

$$\mathbf{y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \quad \sim \quad N_n(\boldsymbol{\mu},\, \sigma^2 \mathbf{I}_n)$$

$\boldsymbol{\mu}$ – mean vector: $\quad \boldsymbol{\mu} = \mu \mathbf{1}_n$

$$\boldsymbol{\mu} = \mathbf{A}\boldsymbol{\gamma}, \quad \mathbf{A} : n \times n_1$$

Vector of observations

n – the number of independent objects, for which we measure a feature

$$\mathbf{y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \quad \sim \quad N_n(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n)$$

$\boldsymbol{\mu}$ – mean vector: $\quad \boldsymbol{\mu} = \mu \mathbf{1}_n$

$$\boldsymbol{\mu} = \mathbf{A}\boldsymbol{\gamma}, \quad \mathbf{A} : n \times n_1$$

σ^2 – variance of observations

Matrix of observations

n – number of independent objects
p – number of observed characteristics

Matrix of observations

n – number of independent objects
p – number of observed characteristics

$$\mathbf{y}_i = \begin{pmatrix} Y_{1i} \\ Y_{2i} \\ \vdots \\ Y_{ni} \end{pmatrix}$$

Matrix of observations

n – number of independent objects
p – number of observed characteristics

$$\mathbf{y}_i = \begin{pmatrix} Y_{1i} \\ Y_{2i} \\ \vdots \\ Y_{ni} \end{pmatrix} \quad \sim \quad N_n(\mu,\ \sigma^2\mathbf{I}_n)$$

Matrix of observations

n – number of independent objects
p – number of observed characteristics

$$\mathbf{Y} = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{np} \end{pmatrix}$$

Matrix of observations

n – number of independent objects
p – number of observed characteristics

$$\mathbf{Y} = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{np} \end{pmatrix} \quad \sim \quad N_{n,p}(\mathbf{M}, \mathbf{I}_n, \mathbf{\Psi})$$

Matrix of observations

n – number of independent objects

p – number of observed characteristics

$$\mathbf{Y} = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{np} \end{pmatrix} \quad \sim \quad N_{n,p}(\mathbf{M}, \mathbf{I}_n, \mathbf{\Psi})$$

$\mathbf{M}$ – mean matrix

Matrix of observations

n – number of independent objects
p – number of observed characteristics

$$\mathbf{Y} = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{np} \end{pmatrix} \quad \sim \quad N_{n,p}(\mathbf{M}, \mathbf{I}_n, \mathbf{\Psi})$$

$\mathbf{M}$ – mean matrix
$\mathbf{\Psi}$ – dispersion matrix of characteristics (the same for each object)

Matrix of observations

n – number of independent objects

p – number of observed characteristics

$$\mathbf{Y} = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{np} \end{pmatrix} \quad \sim \quad N_{n,p}(\mathbf{M}, \mathbf{I}_n, \mathbf{\Psi})$$

Matrix of observations

n – number of independent objects
p – number of observed characteristics

$$\mathbf{Y} = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{np} \end{pmatrix} \quad \sim \quad N_{n,p}(\mathbf{M}, \mathbf{I}_n, \mathbf{\Psi})$$

$$\mathbf{M} = (\mu_1 \mathbf{1}_n, \ldots, \mu_p \mathbf{1}_n)$$

Matrix of observations

n – number of independent objects
p – number of observed characteristics

$$\mathbf{Y} = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{np} \end{pmatrix} \quad \sim \quad N_{n,p}(\mathbf{M}, \mathbf{I}_n, \mathbf{\Psi})$$

$$\mathbf{M} = (\mu_1 \mathbf{1}_n, \ldots, \mu_p \mathbf{1}_n) = \boldsymbol{\mu}' \otimes \mathbf{1}_n$$

Matrix of observations

n – number of independent objects

p – number of observed characteristics

$$\mathbf{Y} = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{np} \end{pmatrix} \quad \sim \quad N_{n,p}(\mathbf{M}, \mathbf{I}_n, \mathbf{\Psi})$$

$$\mathbf{M} = (\mu_1 \mathbf{1}_n, \ldots, \mu_p \mathbf{1}_n) = \boldsymbol{\mu}' \otimes \mathbf{1}_n$$

$$\mathbf{M} = \mathbf{A\Gamma}$$

Matrix of observations

n – number of independent objects
p – number of observed characteristics

$$\mathbf{Y} = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{np} \end{pmatrix} \quad \sim \quad N_{n,p}(\mathbf{M}, \mathbf{I}_n, \mathbf{\Psi})$$

$$\mathbf{M} = (\mu_1 \mathbf{1}_n, \ldots, \mu_p \mathbf{1}_n) = \boldsymbol{\mu}' \otimes \mathbf{1}_n$$

$$\mathbf{M} = \mathbf{A}\mathbf{\Gamma}, \qquad\qquad \mathbf{A} : n \times n_1$$

Matrix of observations

n – number of independent objects
p – number of observed characteristics

$$\mathbf{Y} = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{np} \end{pmatrix} \quad \sim \quad N_{n,p}(\mathbf{M}, \mathbf{I}_n, \mathbf{\Psi})$$

$$\mathbf{M} = (\mu_1 \mathbf{1}_n, \ldots, \mu_p \mathbf{1}_n) = \boldsymbol{\mu}' \otimes \mathbf{1}_n$$

$$\mathbf{M} = \mathbf{A\Gamma}, \qquad \mathbf{A} : n \times n_1$$

$$\mathbf{M} = \mathbf{AXB}'$$

Matrix of observations

n – number of independent objects
p – number of observed characteristics

$$\mathbf{Y} = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{np} \end{pmatrix} \quad \sim \quad N_{n,p}(\mathbf{M}, \mathbf{I}_n, \mathbf{\Psi})$$

$$\mathbf{M} = (\mu_1 \mathbf{1}_n, \ldots, \mu_p \mathbf{1}_n) = \boldsymbol{\mu}' \otimes \mathbf{1}_n$$

$$\mathbf{M} = \mathbf{A\Gamma}, \qquad \mathbf{A} : n \times n_1$$

$$\mathbf{M} = \mathbf{AXB}', \qquad \mathbf{A} : n \times n_1, \quad \mathbf{B} : p \times p_1$$

Matrix of observations - vectorization

n – number of independent objects
p – number of observed characteristics

$$\mathrm{vec}\mathbf{Y} = (\mathbf{y}_1', \ldots, \mathbf{y}_p')'$$

Matrix of observations - vectorization

n – number of independent objects
p – number of observed characteristics

$$\text{vec}\mathbf{Y} = (\mathbf{y}_1', \ldots, \mathbf{y}_p')' \quad \sim \quad N_{np}(\text{vec}\mathbf{M}, \; \boldsymbol{\Psi} \otimes \mathbf{I}_n)$$

Matrix of observations - vectorization

n – number of independent objects

p – number of observed characteristics

$$\mathrm{vec}\mathbf{Y} = (\mathbf{y}_1', \ldots, \mathbf{y}_p')' \quad \sim \quad N_{np}(\mathrm{vec}\mathbf{M},\ \mathbf{\Psi} \otimes \mathbf{I}_n)$$

$$\mathbf{M} = \boldsymbol{\mu}' \otimes \mathbf{1}_n$$

Matrix of observations - vectorization

n – number of independent objects
p – number of observed characteristics

$$\text{vec}\mathbf{Y} = (\mathbf{y}_1', \ldots, \mathbf{y}_p')' \quad \sim \quad N_{np}(\text{vec}\mathbf{M},\ \boldsymbol{\Psi} \otimes \mathbf{I}_n)$$

$\mathbf{M} = \boldsymbol{\mu}' \otimes \mathbf{1}_n$: $\qquad \text{vec}\mathbf{M} = \boldsymbol{\mu} \otimes \mathbf{1}_n$

Matrix of observations - vectorization

n – number of independent objects

p – number of observed characteristics

$$\mathrm{vec}\mathbf{Y} = (\mathbf{y}_1', \dots, \mathbf{y}_p')' \quad \sim \quad N_{np}(\mathrm{vec}\mathbf{M}, \, \boldsymbol{\Psi} \otimes \mathbf{I}_n)$$

$\mathbf{M} = \boldsymbol{\mu}' \otimes \mathbf{1}_n$: $\quad \mathrm{vec}\mathbf{M} = \boldsymbol{\mu} \otimes \mathbf{1}_n$

$\mathbf{M} = \mathbf{A}\boldsymbol{\Gamma}$

Matrix of observations - vectorization

n – number of independent objects
p – number of observed characteristics

$$\mathrm{vec}\mathbf{Y} = (\mathbf{y}_1', \ldots, \mathbf{y}_p')' \quad \sim \quad N_{np}(\mathrm{vec}\mathbf{M},\ \boldsymbol{\Psi} \otimes \mathbf{I}_n)$$

$\mathbf{M} = \boldsymbol{\mu}' \otimes \mathbf{1}_n$: $\qquad \mathrm{vec}\mathbf{M} = \boldsymbol{\mu} \otimes \mathbf{1}_n$

$\mathbf{M} = \mathbf{A}\boldsymbol{\Gamma}$: $\qquad \mathrm{vec}\mathbf{M} = (\mathbf{I}_p \otimes \mathbf{A})\mathrm{vec}\boldsymbol{\Gamma}$

Matrix of observations - vectorization

n – number of independent objects
p – number of observed characteristics

$$\mathrm{vec}\mathbf{Y} = (\mathbf{y}_1', \ldots, \mathbf{y}_p')' \quad \sim \quad N_{np}(\mathrm{vec}\mathbf{M}, \, \boldsymbol{\Psi} \otimes \mathbf{I}_n)$$

$\mathbf{M} = \boldsymbol{\mu}' \otimes \mathbf{1}_n$: $\qquad \mathrm{vec}\mathbf{M} = \boldsymbol{\mu} \otimes \mathbf{1}_n$

$\mathbf{M} = \mathbf{A}\boldsymbol{\Gamma}$: $\qquad \mathrm{vec}\mathbf{M} = (\mathbf{I}_p \otimes \mathbf{A})\mathrm{vec}\boldsymbol{\Gamma}$

$\mathbf{M} = \mathbf{AXB}$

Matrix of observations - vectorization

n – number of independent objects
p – number of observed characteristics

$$\text{vec}\mathbf{Y} = (\mathbf{y}_1', \ldots, \mathbf{y}_p')' \quad \sim \quad N_{np}(\text{vec}\mathbf{M}, \, \boldsymbol{\Psi} \otimes \mathbf{I}_n)$$

$\mathbf{M} = \boldsymbol{\mu}' \otimes \mathbf{1}_n$: $\quad \text{vec}\mathbf{M} = \boldsymbol{\mu} \otimes \mathbf{1}_n$

$\mathbf{M} = \mathbf{A}\boldsymbol{\Gamma}$: $\quad \text{vec}\mathbf{M} = (\mathbf{I}_p \otimes \mathbf{A})\text{vec}\boldsymbol{\Gamma}$

$\mathbf{M} = \mathbf{AXB}$: $\quad \text{vec}\mathbf{M} = (\mathbf{B} \otimes \mathbf{A})\text{vec}\mathbf{X}$

Tensor of observations

n – number of independent objects
p – number of observed characteristics
q – number of time points

Tensor of observations

n – number of independent objects
p – number of observed characteristics
q – number of time points

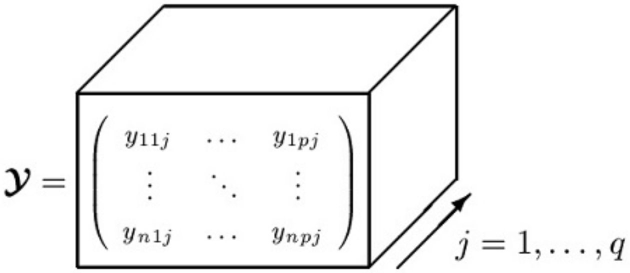
$$\mathbf{Y}_j = \begin{pmatrix} Y_{11j} & Y_{12j} & \cdots & Y_{1pj} \\ Y_{21j} & Y_{12j} & \cdots & Y_{2pj} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1j} & Y_{n2j} & \cdots & Y_{npj} \end{pmatrix}$$

Tensor of observations

n – number of independent objects
p – number of observed characteristics
q – number of time points

$$\mathbf{Y}_j = \begin{pmatrix} Y_{11j} & Y_{12j} & \cdots & Y_{1pj} \\ Y_{21j} & Y_{12j} & \cdots & Y_{2pj} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1j} & Y_{n2j} & \cdots & Y_{npj} \end{pmatrix} \quad \sim \quad N_{n,p}(\mathbf{M}, \mathbf{I}_n, \mathbf{\Psi})$$

Tensor of observations



Visualization of a third-order tensor $\mathcal{Y} \in \mathbb{R}^{n \times p \times q}$

Tensor of observations

n – number of independent objects
p – number of observed characteristics
q – number of time points

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\mathcal{M}, \mathbf{I}_n, \mathbf{\Psi}, \mathbf{\Sigma})$$

Tensor of observations

n – number of independent objects
p – number of observed characteristics
q – number of time points

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\mathcal{M},\, \mathbf{I}_n, \mathbf{\Psi}, \mathbf{\Sigma})$$

$\mathcal{M}$ – tensor of means

Tensor of observations

n – number of independent objects
p – number of observed characteristics
q – number of time points

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\mathcal{M}, \mathbf{I}_n, \mathbf{\Psi}, \mathbf{\Sigma})$$

$\mathcal{M}$ – tensor of means

$\mathbf{\Psi}$ – dispersion matrix of characteristics (the same for each object and each time point)

Tensor of observations

n – number of independent objects
p – number of observed characteristics
q – number of time points

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\mathcal{M}, \mathbf{I}_n, \mathbf{\Psi}, \mathbf{\Sigma})$$

$\mathcal{M}$ – tensor of means

$\mathbf{\Psi}$ – dispersion matrix of characteristics (the same for each object and each time point)

$\mathbf{\Sigma}$ – dispersion matrix of time points (the same for each object and each characteristic)

Tensor of observations

n – number of independent objects
p – number of observed characteristics
q – number of time points

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\mathcal{M}, \mathbf{I}_n, \mathbf{\Psi}, \mathbf{\Sigma})$$

Tensor of observations

n – number of independent objects
p – number of observed characteristics
q – number of time points

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\mathcal{M},\, \mathbf{I}_n, \boldsymbol{\Psi}, \boldsymbol{\Sigma})$$

$$\mathcal{M} = [\![\mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$$

Tensor of observations

n – number of independent objects
p – number of observed characteristics
q – number of time points

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\mathcal{M},\, \mathbf{I}_n, \boldsymbol{\Psi}, \boldsymbol{\Sigma})$$

$$\mathcal{M} = [\![\mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$$

$$\mathbf{A} : n \times n_1, \;\; \mathbf{B} : p \times p_1, \;\; \mathbf{C} : q \times q_1$$

Tucker operator $[\![\mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$

Tucker operator $[\![\mathcal{X}; \mathrm{A}, \mathrm{B}, \mathrm{C}]\!]$

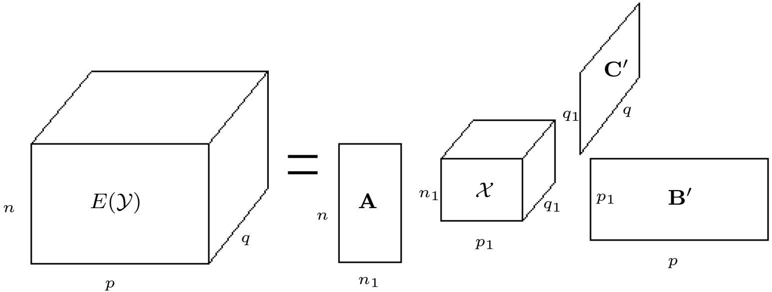
$[\![\mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$ – trilinear tensor product $\mathcal{X}$ from each of three "sides" respectively by matrices $\mathbf{A} \in \mathbb{R}^{n \times n_1}$, $\mathbf{B} \in \mathbb{R}^{p \times p_1}$, $\mathbf{C} \in \mathbb{R}^{q \times q_1}$:

Tucker operator $[\![\mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$

$[\![\mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$ – trilinear tensor product $\mathcal{X}$ from each of three "sides" respectively by matrices $\mathbf{A} \in \mathbb{R}^{n \times n_1}$, $\mathbf{B} \in \mathbb{R}^{p \times p_1}$, $\mathbf{C} \in \mathbb{R}^{q \times q_1}$:

$$([\![\mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!])_{kij} = \sum_{\alpha=1}^{n_1} \sum_{\beta=1}^{p_1} \sum_{\gamma=1}^{q_1} a_{k\alpha} b_{i\beta} c_{j\gamma} x_{\alpha\beta\gamma}$$

Vizualization of Tucker operator for third-order tensor $\mathcal{Y} \in \mathbb{R}^{n\times p\times q}$



Tensor of observations - vectorization

n – number of independent objects
p – number of observed characteristics
q – number of time points

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\mathcal{M}, \mathbf{I}_n, \boldsymbol{\Psi}, \boldsymbol{\Sigma})$$

Tensor of observations - vectorization

n – number of independent objects
p – number of observed characteristics
q – number of time points

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\mathcal{M}, \mathbf{I}_n, \mathbf{\Psi}, \mathbf{\Sigma})$$

$$\mathrm{vec}\mathcal{Y} \quad \sim \quad N_{npq}(\mathrm{vec}\mathcal{M}, \mathbf{\Sigma} \otimes \mathbf{\Psi} \otimes \mathbf{I}_n)$$

Tensor of observations - vectorization

n – number of independent objects
p – number of observed characteristics
q – number of time points

$$\mathcal{Y} \sim N_{n,p,q}(\mathcal{M}, \mathbf{I}_n, \mathbf{\Psi}, \mathbf{\Sigma})$$

$$\mathrm{vec}\mathcal{Y} \sim N_{npq}(\mathrm{vec}\mathcal{M}, \mathbf{\Sigma} \otimes \mathbf{\Psi} \otimes \mathbf{I}_n)$$

$$\mathrm{vec}\mathcal{Y} \sim N_{npq}((\mathbf{C} \otimes \mathbf{B} \otimes \mathbf{A})\mathrm{vec}\mathcal{X}, \mathbf{\Sigma} \otimes \mathbf{\Psi} \otimes \mathbf{I}_n)$$

Tensor of observations - vectorization

n – number of independent objects
p – number of observed characteristics
q – number of time points

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\mathcal{M},\, \mathbf{I}_n,\, \mathbf{\Psi},\, \mathbf{\Sigma})$$

$$\mathrm{vec}\mathcal{Y} \quad \sim \quad N_{npq}(\mathrm{vec}\mathcal{M},\, \mathbf{\Sigma} \otimes \mathbf{\Psi} \otimes \mathbf{I}_n)$$

$$\mathrm{vec}\mathcal{Y} \sim N_{npq}((\mathbf{C} \otimes \mathbf{B} \otimes \mathbf{A})\mathrm{vec}\mathcal{X},\, \mathbf{\Sigma} \otimes \mathbf{\Psi} \otimes \mathbf{I}_n)$$

$$\mathrm{vec}\mathcal{Y} \sim N_{npq}((\mathbf{C} \otimes \mathbf{B} \otimes \mathbf{A})\mathrm{vec}\mathcal{X},\, \mathbf{\Omega} \otimes \mathbf{I}_n)$$

Vectorization

Definition [Kolda and Bader (2009)]

For the thir-order tensor $\mathcal{Y} = (y_{kij})$, $k = 1, \ldots, n$, $i = 1, \ldots, p$, $j = 1, \ldots, q$, **vectorization** is defined as

$$\mathrm{vec}\mathcal{Y} = \sum_{k=1}^{n} \sum_{i=1}^{p} \sum_{j=1}^{q} y_{kij}\, \mathbf{e}_{j,q} \otimes \mathbf{e}_{i,p} \otimes \mathbf{e}_{k,n},$$

with $\mathbf{e}_{r,s}$ being r^{th} column of identity matrix $\mathbf{I}_s$.

Example

Variable: temperature

Example

Variable: temperature

n - number of lakes

Example

Variable: temperature

n - number of lakes

p - number of depths, in which temperature is measured

Example

Variable: temperature

n - number of lakes

p - number of depths, in which temperature is measured

q - number of time points, in which temperature is measured

Example

Assumptions:

- polynomial trend of $p_1 - 1$ order for the depth

Example

Assumptions:

- polynomial trend of $p_1 - 1$ order for the depth
- polynomial trend of $q_1 - 1$ order for time

Example

Assumptions:

- polynomial trend of $p_1 - 1$ order for the depth
- polynomial trend of $q_1 - 1$ order for time

Model (Srivastava et al., 2009): $E(\mathbf{Y}_i) = \mathbf{BXC}'$

Example

Assumptions:

- polynomial trend of $p_1 - 1$ order for the depth
- polynomial trend of $q_1 - 1$ order for time

Model (Srivastava et al., 2009): $E(\mathbf{Y}_i) = \mathbf{BXC}'$

$$\mathrm{E}(\mathcal{Y}) = [\![\mathcal{X}; \mathbf{1}_n, \mathbf{B}, \mathbf{C}]\!], \quad \mathcal{X}: 1 \times p_1 \times q_1$$

Example

Assumptions:

- polynomial trend of $p_1 - 1$ order for the depth
- polynomial trend of $q_1 - 1$ order for time

Model (Srivastava et al., 2009): $E(\mathbf{Y}_i) = \mathbf{BXC}'$

$$\mathrm{E}(\mathcal{Y}) = [\![\mathcal{X}; \mathbf{1}_n, \mathbf{B}, \mathbf{C}]\!], \quad \mathcal{X}: \ 1 \times p_1 \times q_1$$

$$\mathbf{B} = \begin{pmatrix} 1 & b_1 & \cdots & b_1^{p_1-1} \\ 1 & b_2 & \cdots & b_2^{p_1-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & b_p & \cdots & b_p^{p_1-1} \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & c_1 & \cdots & c_1^{q_1-1} \\ 1 & c_2 & \cdots & c_2^{q_1-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & c_q & \cdots & c_q^{q_1-1} \end{pmatrix}$$

Example

Assumptions:

- polynomial trend of $p_1 - 1$ order for the depth
- polynomial trend of $q_1 - 1$ order for time

Example

Assumptions:

- polynomial trend of $p_1 - 1$ order for the depth
- polynomial trend of $q_1 - 1$ order for time
- additional trend between lakes, which depends on location (north/south of Sweden)

Example

Assumptions:

- polynomial trend of $p_1 - 1$ order for the depth
- polynomial trend of $q_1 - 1$ order for time
- additional trend between lakes, which depends on location (north/south of Sweden)

Model:

$$\mathrm{E}(\mathcal{Y}) = [\![\mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] , \quad \mathcal{X} : \; n_1 \times p_1 \times q_1$$

Example

Assumptions:

- polynomial trend of $p_1 - 1$ order for the depth
- polynomial trend of $q_1 - 1$ order for time
- additional trend between lakes, which depends on location (north/south of Sweden)

Model:

$$\mathrm{E}(\mathcal{Y}) = [\![\mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!], \quad \mathcal{X} : n_1 \times p_1 \times q_1$$

$$\mathbf{A} = \mathrm{diag}\left(\mathbf{1}_{u_1}, \mathbf{1}_{u_2}, \ldots, \mathbf{1}_{u_{n_1}}\right), \quad \sum_{i=1}^{n_1} u_i = n$$

Multi-level multivariate model

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket, \mathbf{I}_n, \mathbf{\Psi}, \mathbf{\Sigma})$$

Multi-level multivariate model

$$\text{vec}\mathcal{Y} \sim N_{npq}((\mathbf{C} \otimes \mathbf{B} \otimes \mathbf{A})\text{vec}\mathcal{X}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi} \otimes \mathbf{I}_n)$$

Multi-level multivariate model

$$\mathrm{vec}\mathcal{Y} \quad \sim \quad N_{npq}((\mathbf{C} \otimes \mathbf{B} \otimes \mathbf{A})\mathrm{vec}\mathcal{X}, \ \mathbf{\Sigma} \otimes \mathbf{\Psi} \otimes \mathbf{I}_n)$$

$\mathbf{\Sigma} \in \mathbb{R}^{q \times q}$ – unknown matrix (p.d.)

Multi-level multivariate model

$$\mathrm{vec}\mathcal{Y} \quad \sim \quad N_{npq}((\mathbf{C} \otimes \mathbf{B} \otimes \mathbf{A})\mathrm{vec}\mathcal{X}, \, \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi} \otimes \mathbf{I}_n)$$

$\boldsymbol{\Sigma} \in \mathbb{R}^{q \times q}$ – unknown matrix (p.d.)

$\boldsymbol{\Psi} \in \mathbb{R}^{p \times p}$, $\Psi_{11} = 1$ – known or unknown matrix (p.d.)

Multi-level multivariate model

$$\text{vec}\mathcal{Y} \sim N_{npq}((\mathbf{C} \otimes \mathbf{B} \otimes \mathbf{A})\text{vec}\mathcal{X}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi} \otimes \mathbf{I}_n)$$

$\boldsymbol{\Sigma} \in \mathbb{R}^{q\times q}$ – unknown matrix (p.d.)

$\boldsymbol{\Psi} \in \mathbb{R}^{p\times p}$, $\Psi_{11} = 1$ – known or unknown matrix (p.d.)

$$\text{vec}\mathcal{Y} \sim N_{npq}((\mathbf{C} \otimes \mathbf{B} \otimes \mathbf{A})\text{vec}\mathcal{X}, \boldsymbol{\Omega} \otimes \mathbf{I}_n)$$

Multi-level multivariate model

$$\mathrm{vec}\mathcal{Y} \sim N_{npq}((\mathbf{C} \otimes \mathbf{B} \otimes \mathbf{A})\mathrm{vec}\mathcal{X}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi} \otimes \mathbf{I}_n)$$

$\boldsymbol{\Sigma} \in \mathbb{R}^{q\times q}$ – unknown matrix (p.d.)

$\boldsymbol{\Psi} \in \mathbb{R}^{p\times p}$, $\Psi_{11} = 1$ – known or unknown matrix (p.d.)

$$\mathrm{vec}\mathcal{Y} \sim N_{npq}((\mathbf{C} \otimes \mathbf{B} \otimes \mathbf{A})\mathrm{vec}\mathcal{X}, \boldsymbol{\Omega} \otimes \mathbf{I}_n)$$

$\boldsymbol{\Omega} \in \mathbb{R}^{pq\times pq}$ – unknown matrix (p.d.)

Partial trace operator

$\mathbf{R} = (\mathbf{R}_{ij}) \in \mathbb{R}^{np \times np}$

Partial trace operator

$$\mathbf{R} = (\mathbf{R}_{ij}) \in \mathbb{R}^{np \times np}, \qquad \mathbf{R}_{ij} \in \mathbb{R}^{n \times n}$$

Partial trace operator

$$\mathbf{R} = (\mathbf{R}_{ij}) \in \mathbb{R}^{np \times np}, \qquad \mathbf{R}_{ij} \in \mathbb{R}^{n \times n}$$

$$\mathrm{Tr}_n \mathbf{R} = (\mathrm{tr} \mathbf{R}_{ij})$$

Partial trace operator

$$\mathbf{R} = (\mathbf{R}_{ij}) \in \mathbb{R}^{np \times np}, \qquad \mathbf{R}_{ij} \in \mathbb{R}^{n \times n}$$

$$\mathrm{Tr}_n \mathbf{R} = (\mathrm{tr} \mathbf{R}_{ij}) \in \mathbb{R}^{p \times p}$$

Partial trace operator

$$\mathbf{R} = (\mathbf{R}_{ij}) \in \mathbb{R}^{np \times np}, \qquad \mathbf{R}_{ij} \in \mathbb{R}^{n \times n}$$

$$\mathrm{Tr}_n \mathbf{R} = (\mathrm{tr}\mathbf{R}_{ij}) \in \mathbb{R}^{p \times p}$$

The properties of partial trace operator – see

Filipiak, Klein and Vojtková (2018)

Notation

$\mathbf{G} : m_1 \times m_2$

Notation

$\mathbf{G} : m_1 \times m_2$

$\mathcal{R}(\mathbf{G})$ - column space of $\mathbf{G}$
$\mathcal{R}^{\perp}(\mathbf{G})$ - orthogonal complement of $\mathcal{R}(\mathbf{G})$

Notation

$\mathbf{G} : m_1 \times m_2$

$\mathcal{R}(\mathbf{G})$ - column space of $\mathbf{G}$
$\mathcal{R}^{\perp}(\mathbf{G})$ - orthogonal complement of $\mathcal{R}(\mathbf{G})$

$\mathbf{P}_G = \mathbf{G}(\mathbf{G}'\mathbf{G})^{-}\mathbf{G}'$ - orthogonal projection matrix onto $\mathcal{R}(\mathbf{G})$
$\mathbf{Q}_G = \mathbf{I}_{m_1} - \mathbf{P}_G$ - orthogonal projection matrix onto $\mathcal{R}^{\perp}(\mathbf{G})$

Notation

$\mathbf{G} : m_1 \times m_2$

$\mathcal{R}(\mathbf{G})$ - column space of $\mathbf{G}$
$\mathcal{R}^{\perp}(\mathbf{G})$ - orthogonal complement of $\mathcal{R}(\mathbf{G})$

$\mathbf{P}_G = \mathbf{G}(\mathbf{G}'\mathbf{G})^{-}\mathbf{G}'$ - orthogonal projection matrix onto $\mathcal{R}(\mathbf{G})$
$\mathbf{Q}_G = \mathbf{I}_{m_1} - \mathbf{P}_G$ - orthogonal projection matrix onto $\mathcal{R}^{\perp}(\mathbf{G})$

$\mathbf{H} : m_1 \times m_1$ - arbitrary symmetric matrix p.d.

$\mathbf{P}_{G;H} = \mathbf{G}(\mathbf{G}'\mathbf{H}\mathbf{G})^{-}\mathbf{G}'\mathbf{H}$
$\mathbf{Q}_{G;H} = \mathbf{I} - \mathbf{P}_{G;H}$

Notation

$\mathbf{G} : m_1 \times m_2$

$\mathcal{R}(\mathbf{G})$ - column space of $\mathbf{G}$
$\mathcal{R}^{\perp}(\mathbf{G})$ - orthogonal complement of $\mathcal{R}(\mathbf{G})$

$\mathbf{P}_G = \mathbf{G}(\mathbf{G}'\mathbf{G})^{-}\mathbf{G}'$ - orthogonal projection matrix onto $\mathcal{R}(\mathbf{G})$
$\mathbf{Q}_G = \mathbf{I}_{m_1} - \mathbf{P}_G$ - orthogonal projection matrix onto $\mathcal{R}^{\perp}(\mathbf{G})$

$\mathbf{H} : m_1 \times m_1$ - arbitrary symmetric matrix p.d.

$\mathbf{P}_{G;H} = \mathbf{G}(\mathbf{G}'\mathbf{H}\mathbf{G})^{-}\mathbf{G}'\mathbf{H}$
$\mathbf{Q}_{G;H} = \mathbf{I} - \mathbf{P}_{G;H}$

$\mathbf{G}^{o}$ arbitrary matrix spanning $\mathcal{R}^{\perp}(\mathbf{G})$

Notation

$\mathbf{G} : m_1 \times m_2$

$\mathcal{R}(\mathbf{G})$ - column space of $\mathbf{G}$
$\mathcal{R}^{\perp}(\mathbf{G})$ - orthogonal complement of $\mathcal{R}(\mathbf{G})$

$\mathbf{P}_G = \mathbf{G}(\mathbf{G}'\mathbf{G})^{-}\mathbf{G}'$ - orthogonal projection matrix onto $\mathcal{R}(\mathbf{G})$
$\mathbf{Q}_G = \mathbf{I}_{m_1} - \mathbf{P}_G$ - orthogonal projection matrix onto $\mathcal{R}^{\perp}(\mathbf{G})$

$\mathbf{H} : m_1 \times m_1$ - arbitrary symmetric matrix p.d.

$\mathbf{P}_{G;H} = \mathbf{G}(\mathbf{G}'\mathbf{H}\mathbf{G})^{-}\mathbf{G}'\mathbf{H}$
$\mathbf{Q}_{G;H} = \mathbf{I} - \mathbf{P}_{G;H}$

$\mathbf{G}^o$ arbitrary matrix spanning $\mathcal{R}^{\perp}(\mathbf{G})$

Properties: $\mathbf{Q}_{HG} = \mathbf{P}_{H^{-1}G^o}, \quad \mathbf{Q}_{G;H} = \mathbf{P}'_{G^o;H^{-1}}$

$\mathrm{D}(\mathcal{Y}) = \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi} \otimes \mathbf{I}_n$, $\boldsymbol{\Psi}$ - known

Theorem

The maximum likelihood estimators of unknown parameters in the multi-level multivariate model with known $\boldsymbol{\Psi}$ have the forms

$$\begin{aligned}\widehat{\mathcal{X}} &= [\![\mathcal{Y}; (\mathbf{A}'\mathbf{A})^{-}\mathbf{A}', (\mathbf{B}'\boldsymbol{\Psi}^{-1}\mathbf{B})^{-}\mathbf{B}'\boldsymbol{\Psi}^{-1}, (\mathbf{C}'\mathbf{S}_1^{-1}\mathbf{C})^{-}\mathbf{C}'\mathbf{S}_1^{-1}]\!] + \mathcal{Q} \\ &\quad - [\![\mathcal{Q}; (\mathbf{A}'\mathbf{A})^{-}\mathbf{A}'\mathbf{A}, (\mathbf{B}'\boldsymbol{\Psi}^{-1}\mathbf{B})^{-}\mathbf{B}'\boldsymbol{\Psi}^{-1}\mathbf{B}, (\mathbf{C}'\mathbf{S}_1^{-1}\mathbf{C})^{-}\mathbf{C}'\mathbf{S}_1^{-1}\mathbf{C}]\!] \\ np\widehat{\boldsymbol{\Sigma}} &= \mathbf{S}_1 + \mathbf{Q}_{C;S_1^{-1}}\mathbf{S}_2\mathbf{Q}'_{C;S_1^{-1}}\end{aligned}$$

with

$$\begin{aligned}\mathbf{S}_1 &= \mathrm{Tr}_{np}\left[\{\mathbf{I}_q \otimes (\boldsymbol{\Psi}^{-1/2} \otimes \mathbf{I}_n)\mathbf{Q}_{(\Psi^{-1/2}B\otimes A)}(\boldsymbol{\Psi}^{-1/2} \otimes \mathbf{I}_n)\}\mathrm{vec}\mathcal{Y}\mathrm{vec}'\mathcal{Y}\right] \\ \mathbf{S}_2 &= \mathrm{Tr}_{np}\left[\{\mathbf{I}_q \otimes (\boldsymbol{\Psi}^{-1/2} \otimes \mathbf{I}_n)\mathbf{P}_{(\Psi^{-1/2}B\otimes A)}(\boldsymbol{\Psi}^{-1/2} \otimes \mathbf{I}_n)\}\mathrm{vec}\mathcal{Y}\mathrm{vec}'\mathcal{Y}\right]\end{aligned}$$

being p.d. matrices and $\mathcal{Q}$ is an arbitrary third-order tensor.

$\mathrm{D}(\mathcal{Y}) = \mathbf{\Sigma} \otimes \mathbf{\Psi} \otimes \mathbf{I}_n$, $\mathbf{\Psi}$ - known

Theorem

The maximum likelihood estimators of unknown parameters in the multi-level multivariate model with known $\mathbf{\Psi}$ have the forms

$$[\![\widehat{\mathcal{X}}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] = [\![\mathcal{Y}; \mathbf{P}_A, \mathbf{P}_{B;\Psi^{-1}}, \mathbf{P}_{C;S_1^{-1}}]\!]$$

$$np\widehat{\mathbf{\Sigma}} = \mathbf{S}_1 + \mathbf{Q}_{C;S_1^{-1}}\mathbf{S}_2\mathbf{Q}'_{C;S_1^{-1}}$$

with

$$\mathbf{S}_1 = \mathrm{Tr}_{np}\left[\{\mathbf{I}_q \otimes (\mathbf{\Psi}^{-1/2} \otimes \mathbf{I}_n)\mathbf{Q}_{(\Psi^{-1/2}B\otimes A)}(\mathbf{\Psi}^{-1/2} \otimes \mathbf{I}_n)\}\mathrm{vec}\mathcal{Y}\mathrm{vec}'\mathcal{Y}\right]$$

$$\mathbf{S}_2 = \mathrm{Tr}_{np}\left[\{\mathbf{I}_q \otimes (\mathbf{\Psi}^{-1/2} \otimes \mathbf{I}_n)\mathbf{P}_{(\Psi^{-1/2}B\otimes A)}(\mathbf{\Psi}^{-1/2} \otimes \mathbf{I}_n)\}\mathrm{vec}\mathcal{Y}\mathrm{vec}'\mathcal{Y}\right]$$

being p.d. matrices.

$\mathrm{D}(\mathcal{Y}) = \mathbf{\Sigma} \otimes \mathbf{\Psi} \otimes \mathbf{I}_n$, $\mathbf{\Psi}$ - unknown

Theorem

The maximum likelihood estimators of unknown parameters in the multi-level multivariate model with unknown $\mathbf{\Psi}$ *such that* $\psi_{11} = 1$ *have the forms*

$$\begin{aligned}
\widehat{\mathcal{X}} &= [\![\mathcal{Y}; (\mathbf{A}'\mathbf{A})^{-}\mathbf{A}', (\mathbf{B}'\widehat{\mathbf{\Psi}}^{-1}\mathbf{B})^{-}\mathbf{B}'\widehat{\mathbf{\Psi}}^{-1}, (\mathbf{C}'\mathbf{S}_3^{-1}\mathbf{C})^{-}\mathbf{C}'\mathbf{S}_3^{-1}]\!] + \mathcal{Q} \\
&\quad - [\![\mathcal{Q}; (\mathbf{A}'\mathbf{A})^{-}\mathbf{A}'\mathbf{A}, (\mathbf{B}'\widehat{\mathbf{\Psi}}^{-1}\mathbf{B})^{-}\mathbf{B}'\widehat{\mathbf{\Psi}}^{-1}\mathbf{B}, (\mathbf{C}'\mathbf{S}_3^{-1}\mathbf{C})^{-}\mathbf{C}'\mathbf{S}_3^{-1}\mathbf{C}]\!] \\
np\widehat{\mathbf{\Sigma}} &= \mathbf{S}_3 + \mathbf{Q}_{C;S_3^{-1}}\mathbf{S}_4\mathbf{Q}'_{C;S_3^{-1}} \\
nq\widehat{\mathbf{\Psi}} &= \mathbf{S}_5
\end{aligned}$$

with

$$\begin{aligned}
\mathbf{S}_3 &= \mathrm{Tr}_{np}[\{\mathbf{I}_q \otimes (\widehat{\mathbf{\Psi}}^{-1/2} \otimes \mathbf{I}_n)\mathbf{Q}_{(\widehat{\Psi}^{-1/2}B\otimes A)}(\widehat{\mathbf{\Psi}}^{-1/2} \otimes \mathbf{I}_n)\}\mathrm{vec}\mathcal{Y}\mathrm{vec}'\mathcal{Y}] \\
\mathbf{S}_4 &= \mathrm{Tr}_{np}[\{\mathbf{I}_q \otimes (\widehat{\mathbf{\Psi}}^{-1/2} \otimes \mathbf{I}_n)\mathbf{P}_{(\widehat{\Psi}^{-1/2}B\otimes A)}(\widehat{\mathbf{\Psi}}^{-1/2} \otimes \mathbf{I}_n)\}\mathrm{vec}\mathcal{Y}\mathrm{vec}'\mathcal{Y}]
\end{aligned}$$

being p.d. matrices and $\mathbf{S}_5 = \sum_{k=1}^{q}\sum_{\ell=1}^{q} \mathrm{Tr}_n(s_{k\ell}\widehat{\mathbf{\Gamma}}_{k\ell})$, *where* $s_{k\ell}$ *and* $\widehat{\mathbf{\Gamma}}_{k\ell}$ *are respectively* $(k,\ell)^{\mathrm{th}}$ *element of* $\widehat{\mathbf{\Sigma}}^{-1}$ *and* $np \times np$ *block of a matrix*

$$\mathrm{vec}(\mathcal{Y} - [\![\widehat{\mathcal{X}}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!])\mathrm{vec}'(\mathcal{Y} - [\![\widehat{\mathcal{X}}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!]),$$

while $\mathcal{Q}$ *is an arbitrary tensor of third-order.*

$\mathrm{D}(\mathcal{Y}) = \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi} \otimes \mathbf{I}_n$, $\boldsymbol{\Psi}$ - unknown

Theorem

The maximum likelihood estimators of unknown parameters in the multi-level multivariate model with unknown $\boldsymbol{\Psi}$ *such that* $\psi_{11} = 1$ *have the forms*

$$\begin{aligned} [\![\widehat{\mathcal{X}}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] &= [\![\mathcal{Y}; \mathbf{P}_A, \mathbf{P}_{B;\widehat{\boldsymbol{\Psi}}^{-1}}, \mathbf{P}_{C;S_3^{-1}}]\!] \\ np\widehat{\boldsymbol{\Sigma}} &= \mathbf{S}_3 + \mathbf{Q}_{C;S_3^{-1}}\mathbf{S}_4\mathbf{Q}'_{C;S_3^{-1}} \\ nq\widehat{\boldsymbol{\Psi}} &= \sum_{k=1}^{q}\sum_{\ell=1}^{q} \mathrm{Tr}_n\left[\left(\widehat{\boldsymbol{\Sigma}}^{-1}\right)_{k\ell}\widehat{\boldsymbol{\Gamma}}_{k\ell}\right] \end{aligned}$$

with

$$\begin{aligned} \mathbf{S}_3 &= \mathrm{Tr}_{np}[\{\mathbf{I}_q \otimes (\widehat{\boldsymbol{\Psi}}^{-1/2} \otimes \mathbf{I}_n)\mathbf{Q}_{(\widehat{\Psi}^{-1/2}B\otimes A)}(\widehat{\boldsymbol{\Psi}}^{-1/2} \otimes \mathbf{I}_n)\}\mathrm{vec}\mathcal{Y}\mathrm{vec}'\mathcal{Y}] \\ \mathbf{S}_4 &= \mathrm{Tr}_{np}[\{\mathbf{I}_q \otimes (\widehat{\boldsymbol{\Psi}}^{-1/2} \otimes \mathbf{I}_n)\mathbf{P}_{(\widehat{\Psi}^{-1/2}B\otimes A)}(\widehat{\boldsymbol{\Psi}}^{-1/2} \otimes \mathbf{I}_n)\}\mathrm{vec}\mathcal{Y}\mathrm{vec}'\mathcal{Y}] \end{aligned}$$

being p.d. matrices and $\widehat{\boldsymbol{\Gamma}}_{k\ell}$ *is* $(k, \ell)^{th}$ *block of order* np *of a matrix*

$$\widehat{\boldsymbol{\Gamma}} = \mathrm{vec}(\mathcal{Y} - [\![\widehat{\mathcal{X}}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!])\mathrm{vec}'(\mathcal{Y} - [\![\widehat{\mathcal{X}}; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!]).$$

Example

Variable: temperature

Example

Variable: temperature

$n = 17$ – number of lakes

Example

Variable: temperature

$n = 17$ – number of lakes

$u_1 = 7$ – north region (above 60^{th} parallel)

$u_2 = 10$ – south region (under 60^{th} parallel)

Example

Variable: temperature

$n = 17$ – number of lakes

$u_1 = 7$ – north region (above 60^{th} parallel)

$u_2 = 10$ – south region (under 60^{th} parallel)

$p = 3$ – number of depths ($0.5\,\text{m}, 5\,\text{m}, 15\,\text{m}$)

Example

Variable: temperature

$n = 17$ – number of lakes

$u_1 = 7$ – north region (above 60^{th} parallel)

$u_2 = 10$ – south region (under 60^{th} parallel)

$p = 3$ – number of depths ($0.5\,\text{m}, 5\,\text{m}, 15\,\text{m}$)

$q = 3$ – number of time points (years $1990, 2000, 2009$)

Example

Depth 0.5				Depth 5				Depth 15			
Lake	1990	2000	2009	Lake	1990	2000	2009	Lake	1990	2000	2009
1	7.7	5.1	7.4	1	7.6	5.0	7.2	1	6.9	5.0	6.8
2	15.7	11.8	11.2	2	9.9	11.8	11.0	2	8.5	12.0	10.5
3	13.2	9.4	10.8	3	11.3	9.4	10.0	3	10.1	9.3	8.5
4	15.6	12.7	12.1	4	13.4	12.7	11.8	4	12.6	12.6	11.4
5	15.0	13.0	12.9	5	14.7	12.8	12.4	5	8.9	9.6	8.2
6	18.5	16.5	19.6	6	13.7	11.5	14.5	6	5.2	4.3	4.7
7	15.6	13.5	13.3	7	9.9	12.7	12.0	7	6.4	6.3	7.6
8	15.7	18.5	19.0	8	15.6	14.9	14.6	8	14.9	13.8	13.0
9	19.2	19.0	18.3	9	15.9	17.2	16.0	9	5.6	4.9	3.8
10	19.0	15.7	16.3	10	11.0	9.5	8.8	10	6.0	4.7	3.8
11	18.4	16.7	17.6	11	11.8	12.6	11.8	11	4.8	5.4	4.1
12	17.8	15.0	15.9	12	17.1	13.5	14.3	12	8.0	4.2	4.0
13	20.2	15.7	14.6	13	14.2	15.6	13.7	13	6.2	5.0	4.6
14	18.7	16.6	15.9	14	16.4	12.5	9.7	14	8.4	6.6	6.0
15	17.5	14.3	15.9	15	16.0	14.1	14.4	15	9.2	5.8	5.7
16	18.5	16.1	15.1	16	14.5	15.5	9.5	16	8.4	5.8	5.8
17	16.2	15.1	14.2	17	16.2	15.0	13.3	17	6.9	5.2	5.3

Example

$$\mathbf{A} = \begin{pmatrix} \mathbf{1}_7 & \mathbf{0}_7 \\ \mathbf{0}_{10} & \mathbf{1}_{10} \end{pmatrix}$$

Example

$$\mathbf{A} = \begin{pmatrix} \mathbf{1}_7 & \mathbf{0}_7 \\ \mathbf{0}_{10} & \mathbf{1}_{10} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 0.5 \\ 1 & 5 \\ 1 & 15 \end{pmatrix}$$

Example

$$\mathbf{A} = \begin{pmatrix} \mathbf{1}_7 & \mathbf{0}_7 \\ \mathbf{0}_{10} & \mathbf{1}_{10} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 0.5 \\ 1 & 5 \\ 1 & 15 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 0 \\ 1 & 10 \\ 1 & 19 \end{pmatrix}$$

Example

$$\mathbf{A} = \begin{pmatrix} \mathbf{1}_7 & \mathbf{0}_7 \\ \mathbf{0}_{10} & \mathbf{1}_{10} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 0.5 \\ 1 & 5 \\ 1 & 15 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 0 \\ 1 & 10 \\ 1 & 19 \end{pmatrix}$$

$\widehat{\mathcal{X}}$: $\widehat{\mathbf{X}}_{ij1}$ - front slice, $\widehat{\mathbf{X}}_{ij2}$ - back slice

Example

$$\mathbf{A} = \begin{pmatrix} \mathbf{1}_7 & \mathbf{0}_7 \\ \mathbf{0}_{10} & \mathbf{1}_{10} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 0.5 \\ 1 & 5 \\ 1 & 15 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 0 \\ 1 & 10 \\ 1 & 19 \end{pmatrix}$$

$\widehat{\mathcal{X}}$: $\widehat{\mathbf{X}}_{ij1}$ - front slice, $\widehat{\mathbf{X}}_{ij2}$ - back slice

$$\left(\widehat{\mathbf{X}}_{ij1} \mid \widehat{\mathbf{X}}_{ij2}\right) = \left(\begin{array}{cc|cc} 13.949 & -0.370 & -0.054 & 0.003 \\ 18.348 & -0.708 & -0.092 & -0.001 \end{array}\right)$$

$$\widehat{\mathbf{\Psi}} = \begin{pmatrix} 1.000 & 0.381 & 0.015 \\ 0.381 & 0.988 & 0.185 \\ 0.015 & 0.185 & 0.574 \end{pmatrix}, \qquad \widehat{\mathbf{\Sigma}} = \begin{pmatrix} 7.000 & 5.543 & 5.255 \\ 5.543 & 8.527 & 6.585 \\ 5.255 & 6.585 & 6.992 \end{pmatrix}$$

Example

$$\left(\widehat{\mathbf{X}}_{ij1} \mid \widehat{\mathbf{X}}_{ij2}\right) = \left(\begin{array}{cc|cc} 13.949 & -0.370 & -0.054 & 0.003 \\ 18.348 & -0.708 & -0.092 & -0.001 \end{array}\right)$$

$\hat{x}_{111} = 13.949$ – intercept for northern lakes (mean temperature at the beginning of experiment, in 1990, at the depth 0 m)

$\hat{x}_{121} = -0.37$ – linear trend for depth for northern lakes

$\hat{x}_{112} = -0.054$ – linear trend for time for northern lakes

$\hat{x}_{122} = 0.003$ – interaction between depth and time for northern lakes

Example

$$\left(\widehat{\mathbf{X}}_{ij1} \mid \widehat{\mathbf{X}}_{ij2}\right) = \left(\begin{array}{cc|cc} 13.949 & -0.370 & -0.054 & 0.003 \\ 18.348 & -0.708 & -0.092 & -0.001 \end{array}\right)$$

$\hat{x}_{111} = 13.949$ – intercept for northern lakes (mean temperature at the beginning of experiment, in 1990, at the depth 0 m)

$\hat{x}_{121} = -0.37$ – linear trend for depth for northern lakes

$\hat{x}_{112} = -0.054$ – linear trend for time for northern lakes

$\hat{x}_{122} = 0.003$ – interaction between depth and time for northern lakes

Regression (d - depth, t - time):

$$y = 13.949 - 0.37\,d - 0.054\,t + 0.003\,t\,d$$

Example

$$\left(\widehat{\mathbf{X}}_{ij1} \mid \widehat{\mathbf{X}}_{ij2}\right) = \left(\begin{array}{cc|cc} 13.949 & -0.370 & -0.054 & 0.003 \\ 18.348 & -0.708 & -0.092 & -0.001 \end{array}\right)$$

$\hat{x}_{211} = 18.348$ – intercept for southern lakes (mean temperature at the beginning of experiment in 1990, at the depth 0 m)

$\hat{x}_{221} = -0.708$ – linear trend for depth for southern lakes

$\hat{x}_{212} = -0.092$ – linear trend for time for southern lakes

$\hat{x}_{222} = -0.001$ – interaction between depth and time for southern lakes

Example

$$\left(\widehat{\mathbf{X}}_{ij1} \mid \widehat{\mathbf{X}}_{ij2}\right) = \left(\begin{array}{cc|cc} 13.949 & -0.370 & -0.054 & 0.003 \\ 18.348 & -0.708 & -0.092 & -0.001 \end{array}\right)$$

$\hat{x}_{211} = 18.348$ – intercept for southern lakes (mean temperature at the beginning of experiment in 1990, at the depth 0 m)

$\hat{x}_{221} = -0.708$ – linear trend for depth for southern lakes

$\hat{x}_{212} = -0.092$ – linear trend for time for southern lakes

$\hat{x}_{222} = -0.001$ – interaction between depth and time for southern lakes

Regression (d - depth, t - time):

$$y = 18.348 - 0.708\, d - 0.092\, t - 0.001\, t\, d$$

Matricization

Model

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket \,, \mathbf{I}_n, \mathbf{\Psi}, \mathbf{\Sigma})$$

Matricization

Model

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket , \mathbf{I}_n, \boldsymbol{\Psi}, \boldsymbol{\Sigma})$$

$$\mathbf{Y}^{(1)} \sim N_{np,q}((\mathbf{B} \otimes \mathbf{A})\mathbf{X}^{(1)}\mathbf{C}', \boldsymbol{\Psi} \otimes \mathbf{I}_n, \boldsymbol{\Sigma})$$

Matricization

Model

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket, \mathbf{I}_n, \mathbf{\Psi}, \mathbf{\Sigma})$$

$$\mathbf{Y}^{(1)} \sim N_{np,q}((\mathbf{B} \otimes \mathbf{A})\mathbf{X}^{(1)}\mathbf{C}', \mathbf{\Psi} \otimes \mathbf{I}_n, \mathbf{\Sigma})$$
$$\mathbf{Y}^{(2)} \sim N_{n,pq}(\mathbf{A}\mathbf{X}^{(2)}(\mathbf{C}' \otimes \mathbf{B}'), \mathbf{I}_n, \mathbf{\Sigma} \otimes \mathbf{\Psi})$$

Matricization

Model

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket, \mathbf{I}_n, \mathbf{\Psi}, \mathbf{\Sigma})$$

$$\begin{aligned}
\mathbf{Y}^{(1)} &\sim N_{np,q}((\mathbf{B} \otimes \mathbf{A})\mathbf{X}^{(1)}\mathbf{C}', \ \mathbf{\Psi} \otimes \mathbf{I}_n, \ \mathbf{\Sigma}) \\
\mathbf{Y}^{(2)} &\sim N_{n,pq}(\mathbf{A}\mathbf{X}^{(2)}(\mathbf{C}' \otimes \mathbf{B}'), \ \mathbf{I}_n, \ \mathbf{\Sigma} \otimes \mathbf{\Psi}) \\
\mathbf{Y}^{(3)} &\sim N_{nq,p}((\mathbf{C} \otimes \mathbf{A})\mathbf{X}^{(3)}\mathbf{B}', \ \mathbf{\Sigma} \otimes \mathbf{I}_n, \ \mathbf{\Psi})
\end{aligned}$$

Matricization

Model

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket, \mathbf{I}_n, \boldsymbol{\Psi}, \boldsymbol{\Sigma})$$

$$\mathbf{Y}^{(1)} \sim N_{np,q}((\mathbf{B} \otimes \mathbf{A})\mathbf{X}^{(1)}\mathbf{C}', \boldsymbol{\Psi} \otimes \mathbf{I}_n, \boldsymbol{\Sigma})$$
$$\mathbf{Y}^{(2)} \sim N_{n,pq}(\mathbf{A}\mathbf{X}^{(2)}(\mathbf{C}' \otimes \mathbf{B}'), \mathbf{I}_n, \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi})$$
$$\mathbf{Y}^{(3)} \sim N_{nq,p}((\mathbf{C} \otimes \mathbf{A})\mathbf{X}^{(3)}\mathbf{B}', \boldsymbol{\Sigma} \otimes \mathbf{I}_n, \boldsymbol{\Psi})$$

$$\mathbf{X}^{(1)} : n_1 p_1 \times q_1$$

Matricization

Model

$$\mathcal{Y} \sim N_{n,p,q}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket, \mathbf{I}_n, \mathbf{\Psi}, \mathbf{\Sigma})$$

$$\mathbf{Y}^{(1)} \sim N_{np,q}((\mathbf{B} \otimes \mathbf{A})\mathbf{X}^{(1)}\mathbf{C}', \mathbf{\Psi} \otimes \mathbf{I}_n, \mathbf{\Sigma})$$
$$\mathbf{Y}^{(2)} \sim N_{n,pq}(\mathbf{A}\mathbf{X}^{(2)}(\mathbf{C}' \otimes \mathbf{B}'), \mathbf{I}_n, \mathbf{\Sigma} \otimes \mathbf{\Psi})$$
$$\mathbf{Y}^{(3)} \sim N_{nq,p}((\mathbf{C} \otimes \mathbf{A})\mathbf{X}^{(3)}\mathbf{B}', \mathbf{\Sigma} \otimes \mathbf{I}_n, \mathbf{\Psi})$$

$\mathbf{X}^{(1)} : n_1 p_1 \times q_1$

$\mathbf{X}^{(2)} : n_1 \times p_1 q_1$

Matricization

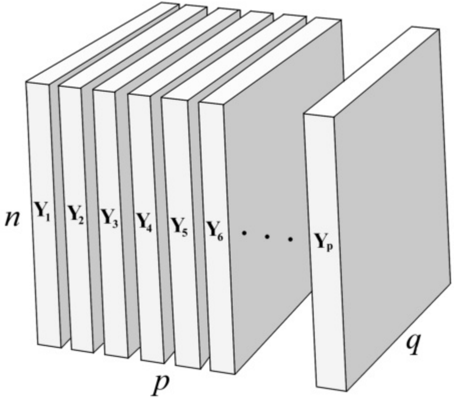
Model

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket, \mathbf{I}_n, \boldsymbol{\Psi}, \boldsymbol{\Sigma})$$

$$\mathbf{Y}^{(1)} \sim N_{np,q}((\mathbf{B} \otimes \mathbf{A})\mathbf{X}^{(1)}\mathbf{C}', \boldsymbol{\Psi} \otimes \mathbf{I}_n, \boldsymbol{\Sigma})$$
$$\mathbf{Y}^{(2)} \sim N_{n,pq}(\mathbf{A}\mathbf{X}^{(2)}(\mathbf{C}' \otimes \mathbf{B}'), \mathbf{I}_n, \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi})$$
$$\mathbf{Y}^{(3)} \sim N_{nq,p}((\mathbf{C} \otimes \mathbf{A})\mathbf{X}^{(3)}\mathbf{B}', \boldsymbol{\Sigma} \otimes \mathbf{I}_n, \boldsymbol{\Psi})$$

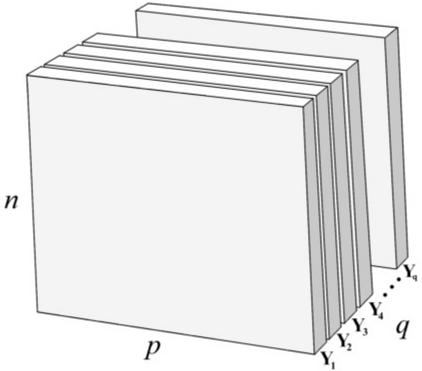
$$\mathbf{X}^{(1)} : n_1 p_1 \times q_1$$
$$\mathbf{X}^{(2)} : n_1 \times p_1 q_1$$
$$\mathbf{X}^{(3)} : n_1 q_1 \times p_1$$

Observations - matricization



Matricization: $\mathbf{Y}_i \in \mathbb{R}^{n \times q}$ written one under the other – we obtain $\mathbf{Y}^{(1)} \in \mathbb{R}^{np \times q}$

Observations - matricization



Matricization: $\mathbf{Y}_i \in \mathbb{R}^{n\times p}$ written:

one next to the other – we obtain $\mathbf{Y}^{(2)} \in \mathbb{R}^{n\times pq}$

one under the other – we obtain $\mathbf{Y}^{(3)} \in \mathbb{R}^{nq\times p}$

Matricization

Model

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket, \mathbf{I}_n, \mathbf{\Psi}, \mathbf{\Sigma})$$

Matricization

Model

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket, \mathbf{I}_n, \mathbf{\Psi}, \mathbf{\Sigma})$$

$$\mathbf{Y}^{(1)} \sim N_{np,q}((\mathbf{B} \otimes \mathbf{A})\mathbf{X}^{(1)}\mathbf{C}', \ \mathbf{\Psi} \otimes \mathbf{I}_n, \mathbf{\Sigma})$$
$$\mathbf{Y}^{(2)} \sim N_{n,pq}(\mathbf{A}\mathbf{X}^{(2)}(\mathbf{C}' \otimes \mathbf{B}'), \ \mathbf{I}_n, \mathbf{\Sigma} \otimes \mathbf{\Psi})$$
$$\mathbf{Y}^{(3)} \sim N_{nq,p}((\mathbf{C} \otimes \mathbf{A})\mathbf{X}^{(3)}\mathbf{B}', \ \mathbf{\Sigma} \otimes \mathbf{I}_n, \mathbf{\Psi})$$

Matricization

Model

$$\mathcal{Y} \quad \sim \quad N_{n,p,q}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket, \mathbf{I}_n, \boldsymbol{\Psi}, \boldsymbol{\Sigma})$$

$$\begin{aligned}
\mathbf{Y}^{(1)} &\sim N_{np,q}((\mathbf{B} \otimes \mathbf{A})\mathbf{X}^{(1)}\mathbf{C}', \ \boldsymbol{\Psi} \otimes \mathbf{I}_n, \boldsymbol{\Sigma}) \\
\mathbf{Y}^{(2)} &\sim N_{n,pq}(\mathbf{A}\mathbf{X}^{(2)}(\mathbf{C}' \otimes \mathbf{B}'), \ \mathbf{I}_n, \ \ \boldsymbol{\Omega}) \\
\mathbf{Y}^{(3)} &\sim N_{nq,p}((\mathbf{C} \otimes \mathbf{A})\mathbf{X}^{(3)}\mathbf{B}', \ \boldsymbol{\Sigma} \otimes \mathbf{I}_n, \boldsymbol{\Psi})
\end{aligned}$$

$\mathrm{D}(\mathcal{Y}) = \mathbf{\Omega} \otimes \mathbf{I}_n$

Theorem

The maximum likelihood estimator of unknown parameters under the multi-level multivariate model $\mathbf{Y}^{(2)}$ *with dispersion matrix* $\mathrm{D}(\mathcal{Y}) = \mathbf{\Omega} \otimes \mathbf{I}_n$ *have the form*

$$\begin{aligned} \widehat{\mathbf{X}}^{(2)} &= (\mathbf{A}'\mathbf{A})^{-}\mathbf{A}'\mathbf{Y}^{(2)}\mathbf{S}_3^{-1}(\mathbf{C}\otimes\mathbf{B})\{(\mathbf{C}'\otimes\mathbf{B}')\mathbf{S}_3^{-1}(\mathbf{C}\otimes\mathbf{B})\}^{-} \\ &\quad +(\mathbf{A}')^{o}\mathbf{Z}_1 + \mathbf{A}'\mathbf{Z}_2(\mathbf{C}'\otimes\mathbf{B}')^{o'} \\ n\widehat{\mathbf{\Omega}} &= \mathbf{S}_3 + \mathbf{Q}_{C\otimes B;S_3^{-1}}\mathbf{Y}^{(2)'}\mathbf{P}_A\mathbf{Y}^{(2)}\mathbf{Q}'_{C\otimes B;S_3^{-1}} \end{aligned}$$

with $\mathbf{S}_3 = \mathbf{Y}^{(2)'}\mathbf{Q}_A\mathbf{Y}^{(2)}$ *being p.d. matrix and* $\mathbf{Z}_1$, $\mathbf{Z}_2$ *are arbitrary matrices.*

Tensor of order $k > 3$

$$\mathcal{Y} \sim N_{n,p_{k-1},\ldots,p_1}\left(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}_{k-1}, \ldots, \mathbf{B}_1 \rrbracket, \mathbf{I}_n, \mathbf{\Psi}_{k-1}, \ldots, \mathbf{\Psi}_1\right)$$

Tensor of order $k > 3$

$$\mathcal{Y} \sim N_{n,p_{k-1},\ldots,p_1}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}_{k-1}, \ldots, \mathbf{B}_1 \rrbracket, \mathbf{I}_n, \mathbf{\Psi}_{k-1}, \ldots, \mathbf{\Psi}_1)$$

$\mathbf{A} \in \mathbb{R}^{n \times n^*}$, $\mathbf{B}_i \in \mathbb{R}^{p_i \times p_i^*}$ – known design matrices

Tensor of order $k > 3$

$$\mathcal{Y} \sim N_{n,p_{k-1},\dots,p_1}\left(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}_{k-1}, \dots, \mathbf{B}_1 \rrbracket, \mathbf{I}_n, \mathbf{\Psi}_{k-1}, \dots, \mathbf{\Psi}_1\right)$$

$\mathbf{A} \in \mathbb{R}^{n \times n^*}$, $\mathbf{B}_i \in \mathbb{R}^{p_i \times p_i^*}$ – known design matrices

$\mathcal{X} \in \mathbb{R}^{n^* \times p_{k-1}^* \times \dots \times p_1^*}$ – tensor of order k

Tensor of order $k > 3$

$$\mathcal{Y} \sim N_{n,p_{k-1},\ldots,p_1}\left(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}_{k-1}, \ldots, \mathbf{B}_1 \rrbracket, \mathbf{I}_n, \mathbf{\Psi}_{k-1}, \ldots, \mathbf{\Psi}_1\right)$$

$\mathbf{A} \in \mathbb{R}^{n \times n^*}$, $\mathbf{B}_i \in \mathbb{R}^{p_i \times p_i^*}$ – known design matrices

$\mathcal{X} \in \mathbb{R}^{n^* \times p_{k-1}^* \times \cdots \times p_1^*}$ – tensor of order k

$\mathbf{\Psi}_i \in \mathbb{R}^{p_i \times p_i}$ – known/unknown matrices p.d.

Tensor of order $k > 3$

$$\mathcal{Y} \sim N_{n,p_{k-1},\ldots,p_1}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}_{k-1}, \ldots, \mathbf{B}_1 \rrbracket, \mathbf{I}_n, \mathbf{\Psi}_{k-1}, \ldots, \mathbf{\Psi}_1)$$

Tensor of order $k > 3$

$$\mathcal{Y} \sim N_{n,p_{k-1},\dots,p_1}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}_{k-1}, \dots, \mathbf{B}_1 \rrbracket, \mathbf{I}_n, \mathbf{\Psi}_{k-1}, \dots, \mathbf{\Psi}_1)$$

$$\mathrm{D}(\mathcal{Y}) = \mathbf{\Psi}_1 \otimes \cdots \otimes \mathbf{\Psi}_{k-1} \otimes \mathbf{I}_n, \qquad \mathbf{\Psi}_1 \text{ unknown}$$

Tensor of order $k > 3$

$$\mathcal{Y} \sim N_{n,p_{k-1},\ldots,p_1}\left(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}_{k-1}, \ldots, \mathbf{B}_1 \rrbracket, \mathbf{I}_n, \boldsymbol{\Psi}_{k-1}, \ldots, \boldsymbol{\Psi}_1\right)$$

$$\mathrm{D}(\mathcal{Y}) = \boldsymbol{\Psi}_1 \otimes \cdots \otimes \boldsymbol{\Psi}_{k-1} \otimes \mathbf{I}_n, \quad \boldsymbol{\Psi}_1 \text{ unknown}$$

$$\mathrm{D}(\mathcal{Y}) = \boldsymbol{\Psi}_1 \otimes \cdots \otimes \boldsymbol{\Psi}_{k-1} \otimes \mathbf{I}_n, \quad \boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2 \text{ unknown}$$

Tensor of order $k > 3$

$$\mathcal{Y} \sim N_{n,p_{k-1},\dots,p_1}\left(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}_{k-1}, \dots, \mathbf{B}_1 \rrbracket, \mathbf{I}_n, \boldsymbol{\Psi}_{k-1}, \dots, \boldsymbol{\Psi}_1\right)$$

$$\mathrm{D}(\mathcal{Y}) = \boldsymbol{\Psi}_1 \otimes \cdots \otimes \boldsymbol{\Psi}_{k-1} \otimes \mathbf{I}_n, \quad \boldsymbol{\Psi}_1 \text{ unknown}$$

$$\mathrm{D}(\mathcal{Y}) = \boldsymbol{\Psi}_1 \otimes \cdots \otimes \boldsymbol{\Psi}_{k-1} \otimes \mathbf{I}_n, \quad \boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2 \text{ unknown}$$

$$\mathrm{D}(\mathcal{Y}) = \boldsymbol{\Omega} \otimes \mathbf{I}_n, \quad \boldsymbol{\Omega} \text{ unknown}$$

Tensor of order $k > 3$

$$\mathcal{Y} \sim N_{n,p_{k-1},\ldots,p_1}\left(\left[\!\left[\mathcal{X}; \mathbf{A}, \mathbf{B}_{k-1}, \ldots, \mathbf{B}_1\right]\!\right], \mathbf{I}_n, \mathbf{\Psi}_{k-1}, \ldots, \mathbf{\Psi}_1\right)$$

Tensor of order $k > 3$

$$\mathcal{Y} \sim N_{n,p_{k-1},\dots,p_1}(\llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}_{k-1}, \dots, \mathbf{B}_1 \rrbracket, \mathbf{I}_n, \mathbf{\Psi}_{k-1}, \dots, \mathbf{\Psi}_1)$$

Matricization:

$$\mathbf{Y} \sim N_{np_{k-1}\dots p_2, p_1}\left(\left(\bigotimes_{i=2}^{k-1} \mathbf{B}_i \otimes \mathbf{A}\right) \mathbf{X}\mathbf{B}_1^\top, \bigotimes_{i=2}^{k-1} \mathbf{\Psi}_i \otimes \mathbf{I}_n, \mathbf{\Psi}_1\right)$$

Tensor of order $k > 3$

Theorem

The maximum likelihood estimators of unknown parameters under the kth-level multivariate model, $k > 3$, with the dispersion matrix $\mathrm{D}(\mathcal{Y}) = \boldsymbol{\Psi}_1 \otimes \cdots \otimes \boldsymbol{\Psi}_{k-1} \otimes \mathbf{I}_n$, $\boldsymbol{\Psi}_1$ unknown, have the forms

$$\begin{aligned}
\widehat{\mathbf{X}} &= (\mathbf{B}'\boldsymbol{\Psi}^{-1}\mathbf{B} \otimes \mathbf{A}'\mathbf{A})^{-}(\mathbf{B}'\boldsymbol{\Psi}^{-1} \otimes \mathbf{A}')\mathbf{Y}^{(1)}\mathbf{S}_4^{-1}\mathbf{B}_1(\mathbf{B}_1'\mathbf{S}_4^{-1}\mathbf{B}_1)^{-} \\
&\quad +(\mathbf{B}'\boldsymbol{\Psi}^{-1/2} \otimes \mathbf{A}')^{o}\mathbf{Z}_1 + (\mathbf{B}'\boldsymbol{\Psi}^{-1/2} \otimes \mathbf{A}')\mathbf{Z}_2\mathbf{B}_1^{o\prime} \\
np\widehat{\boldsymbol{\Psi}}_1 &= \{\mathbf{Y} - (\mathbf{B} \otimes \mathbf{A})\widehat{\mathbf{X}}\mathbf{B}_1'\}'(\boldsymbol{\Psi}^{-1} \otimes \mathbf{I}_n)\{\mathbf{Y} - (\mathbf{B} \otimes \mathbf{A})\widehat{\mathbf{X}}\mathbf{B}_1'\}
\end{aligned}$$

with $\mathbf{S}_4 = \mathbf{Y}'(\boldsymbol{\Psi}^{-1/2} \otimes \mathbf{I}_n)\mathbf{Q}_{(\Psi^{-1/2}B\otimes A)}(\boldsymbol{\Psi}^{-1/2} \otimes \mathbf{I}_n)\mathbf{Y}$, $\mathbf{B} = \bigotimes_{i=2}^{k-1}\mathbf{B}_i$, $\boldsymbol{\Psi} = \bigotimes_{i=2}^{k-1}\boldsymbol{\Psi}_i$ *and* $p = p_2 \cdots p_{k-1}$ *is p.d. matrix and* $\mathbf{Z}_1$, $\mathbf{Z}_2$ *are arbitrary matrices.*

Tensor of order $k > 3$

Theorem

The maximum likelihood estimators of unknown parameters under the kth-level multivariate model, $k > 3$, with the dispersion matrix $\mathrm{D}(\mathcal{Y}) = \boldsymbol{\Psi}_1 \otimes \cdots \otimes \boldsymbol{\Psi}_{k-1} \otimes \mathbf{I}_n$, $\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2$ unknown, have the forms

$$
\begin{aligned}
\widehat{\mathbf{X}} &= (\mathbf{B}'\widehat{\boldsymbol{\Psi}}^{-1}\mathbf{B} \otimes \mathbf{A}'\mathbf{A})^{-}(\mathbf{B}'\widehat{\boldsymbol{\Psi}}^{-1} \otimes \mathbf{A}')\mathbf{Y}\mathbf{S}_5^{-1}\mathbf{B}_1(\mathbf{B}_1'\mathbf{S}_5^{-1}\mathbf{B}_1)^{-} \\
&\quad + (\mathbf{B}'\widehat{\boldsymbol{\Psi}}^{-1/2} \otimes \mathbf{A}')^{o}\mathbf{Z}_1 + (\mathbf{B}'\widehat{\boldsymbol{\Psi}}^{-1/2} \otimes \mathbf{A}')\mathbf{Z}_2\mathbf{B}_1^{\prime o\prime} \\
np\widehat{\boldsymbol{\Psi}}_1 &= \{\mathbf{Y} - (\mathbf{B} \otimes \mathbf{A})\widehat{\mathbf{X}}\mathbf{B}_1'\}'(\widehat{\boldsymbol{\Psi}}^{-1} \otimes \mathbf{I}_n)\{\mathbf{Y} - (\mathbf{B} \otimes \mathbf{A})\widehat{\mathbf{X}}\mathbf{B}_1'\} \\
nq\widehat{\boldsymbol{\Psi}}_2 &= \mathrm{Tr}_h\left[\{\mathbf{Y} - (\mathbf{B} \otimes \mathbf{A})\widehat{\mathbf{X}}\mathbf{B}_1'\}\widehat{\boldsymbol{\Psi}}_1^{-1}\{\mathbf{Y} - (\mathbf{B} \otimes \mathbf{A})\widehat{\mathbf{X}}\mathbf{B}_1'\}'\right]
\end{aligned}
$$

with $\mathbf{B} = \bigotimes_{i=2}^{k-1} \mathbf{B}_i$, $\widehat{\boldsymbol{\Psi}} = \widehat{\boldsymbol{\Psi}}_2 \otimes \bigotimes_{i=3}^{k-1} \boldsymbol{\Psi}_i$, $p = p_2 \cdots p_{k-1}$, $q = p_1 p_3 \cdots p_{k-1}$ *and* $h = p_3 \cdots p_{k-1} n$, *where matrix* $\mathbf{S}_5 = \mathbf{Y}'(\widehat{\boldsymbol{\Psi}}^{-1/2} \otimes \mathbf{I}_n)\mathbf{Q}_{(\widehat{\boldsymbol{\Psi}}^{-1/2}B \otimes A)}(\widehat{\boldsymbol{\Psi}}^{-1/2} \otimes \mathbf{I}_n)\mathbf{Y}$ *is p.d., and* $\mathbf{Z}_1$, $\mathbf{Z}_2$ *are arbitrary matrices.*

Tensor of order $k > 3$

Theorem

The maximum likelihood estimators of unknown parameters under the kth-level multivariate model, $k > 3$, with the dispersion matrix $\mathrm{D}(\mathcal{Y}) = \mathbf{\Omega} \otimes \mathbf{I}_n$ have the forms

$$\begin{aligned}
\widehat{\mathbf{X}} &= (\mathbf{A}'\mathbf{A})^{-}\mathbf{A}'\mathbf{Y}\mathbf{S}_3^{-1}(\mathbf{C} \otimes \mathbf{B})\{(\mathbf{C}' \otimes \mathbf{B}')\mathbf{S}_3^{-1}(\mathbf{C} \otimes \mathbf{B})\}^{-} \\
&\quad +(\mathbf{A}')^{o}\mathbf{Z}_1 + \mathbf{A}'\mathbf{Z}_2(\mathbf{C}' \otimes \mathbf{B}')^{o'}, \\
n\widehat{\mathbf{\Omega}} &= \mathbf{S}_3 + \mathbf{Q}_{C \otimes B; S_3^{-1}}\mathbf{Y}'\mathbf{P}_A\mathbf{Y}\mathbf{Q}'_{C \otimes B; S_3^{-1}},
\end{aligned}$$

with $\mathbf{C} \otimes \mathbf{B} = \bigotimes_{i=1}^{k} \mathbf{B}_i$, $\mathbf{S}_3 = \mathbf{Y}'\mathbf{Q}_A\mathbf{Y}$ *is p.d. matrix and* $\mathbf{Z}_1$, $\mathbf{Z}_2$ *are arbitrary matrices.*

Main references

- Filipiak, K. and D. Klein (2017). Estimation of parameters under a generalized growth curve model. *J. Multivariate Anal. 158*, 73–86.
- Filipiak, K., D. Klein, E. Vojtková (2018). The properties of partial trace and block trace operators of partitioned matrices. *Electronic J. Linear Algebra 33*, 3–15.
- Kolda, T. G. and B. W. Bader (2009). Tensor decompositions and applications. *SIAM Review* 51(3), 455–500.
- Magnus, J. R. and H. Neudecker (1986). Symmetry, 0-1 matrices and Jacobians. *Econometric Theory* 2, 157–190.
- Ohlson, M., M. R. Ahmad and D. von Rosen (2013). The multilinear normal distribution: Introduction and some basic properties. *J. Multivariate Anal.* 113, 37–47.
- Savas, B. and L.-H. Lim (2008). Best multilinear rank approximation of tensors with quasi-Newton methods on Grassmannians. Linköping University Report, LITH-MAT-R–2008-01–SE.
- Srivastava, M., T. von Rosen and D. von Rosen (2009). Estimation and testing in general multivariate linear models with Kronecker product covariance structure. *Sankhyã* 71-A, 137–163.