

Estymacja nieznanymi parametrów w wielopoziomowych modelach wielowymiarowych

Katarzyna Filipiak

Instytut Matematyki, Politechnika Poznańska

19.04.2017

Obserwacje wektorowe

n – liczba niezależnych obiektów, na których obserwowana jest pojedyncza zmienna (cecha)

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

$\boldsymbol{\mu}$ – wektor średnich: $\boldsymbol{\mu} = \mu \mathbf{1}_n$
 $\boldsymbol{\mu} = \mathbf{A}\boldsymbol{\gamma}$, $\mathbf{A} : n \times n_1$

σ^2 – wariancja obserwacji

Obserwacje macierzowe

n – liczba niezależnych obiektów

p – liczba obserwowanych zmiennych (cech)

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

Obserwacje macierzowe

n – liczba niezależnych obiektów

p – liczba obserwowanych zmiennych (cech)

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

\mathbf{M} - macierz średnich

Ψ - macierz wariancji-kowariancji cech dla danego obiektu (taka sama dla wszystkich obiektów)

Obserwacje macierzowe

n – liczba niezależnych obiektów

p – liczba obserwowanych zmiennych (cech)

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

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

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

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

Obserwacje macierzowe - wektoryzacja

n – liczba niezależnych obiektów

p – liczba obserwowanych zmiennych (cech)

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

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

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

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

Obserwacje tensorowe

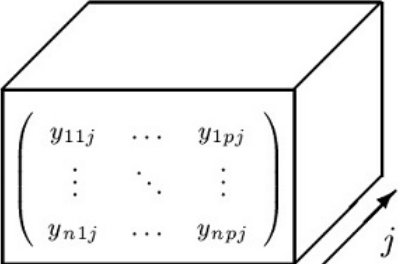
n – liczba niezależnych obiektów

p – liczba obserwowanych zmiennych (cech)

q – liczba punktów czasowych

$$\mathbf{Y}_j = \begin{pmatrix} Y_{11j} & Y_{12j} & \cdots & Y_{1pj} \\ Y_{21j} & Y_{22j} & \cdots & Y_{2pj} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1j} & Y_{n2j} & \cdots & Y_{npj} \end{pmatrix} \sim N_{n,p}(\mathbf{M}, \mathbf{I}_n, \Psi)$$

Obserwacje tensorowe

$$\mathcal{Y} = \begin{pmatrix} y_{11j} & \cdots & y_{1pj} \\ \vdots & \ddots & \vdots \\ y_{n1j} & \cdots & y_{npj} \end{pmatrix}$$


$j = 1, \dots, q$

Wizualizacja tensora trzeciego rzędu $\mathcal{Y} \in \mathbb{R}^{n \times p \times q}$

Obserwacje tensorowe

n – liczba niezależnych obiektów

p – liczba obserwowanych zmiennych (cech)

q – liczba punktów czasowych

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

\mathcal{M} - tensor średnich

Ψ - macierz wariancji-kowariancji cech dla danego obiektu (taka sama dla wszystkich obiektów i punktów czasowych)

Σ - macierz wariancji-kowariancji punktów czasowych dla danego obiektu (taka sama dla wszystkich obiektów i wszystkich cech)

Obserwacje tensorowe

n – liczba niezależnych obiektów

p – liczba obserwowanych zmiennych (cech)

q – liczba punktów czasowych

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

$$\mathcal{M} = \llbracket \mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket,$$

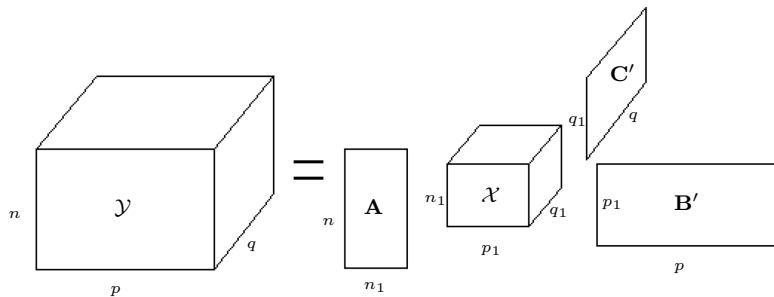
$$\mathbf{A} : n \times n_1, \quad \mathbf{B} : p \times p_1, \quad \mathbf{C} : q \times q_1$$

Operator Tuckera $[[\mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C}]]$

$[[\mathcal{X}; \mathbf{A}, \mathbf{B}, \mathbf{C}]]$ – potrójnie-liniowe mnożenie tensora \mathcal{X} z każdej z trzech "stron" odpowiednio przez macierze $\mathbf{A} \in \mathbb{R}^{n \times n_1}$, $\mathbf{B} \in \mathbb{R}^{p \times p_1}$, $\mathbf{C} \in \mathbb{R}^{q \times q_1}$:

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

Wizualizacja operatora Tuckera dla tensora trzeciego rzędu $\mathcal{Y} \in \mathbb{R}^{n \times p \times q}$



Obserwacje tensorowe - wektoryzacja

n – liczba niezależnych obiektów

p – liczba obserwowanych zmiennych (cech)

q – liczba punktów czasowych

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

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

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

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

Definicja [Kolda i Bader (2009)]

Dla tensora trzeciego rzędu $\mathcal{Y} = (y_{kij})$, $k = 1, \dots, n$,
 $i = 1, \dots, p$, $j = 1, \dots, q$, bf wektoryzowanie
definiowane jest jako

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

gdzie $\mathbf{e}_{r,s}$ jest r -tą kolumną macierzy jednostkowej \mathbf{I}_s .

Przykład

Obserwowana zmienna: **temperatura**

n - liczba jezior

p - liczba głębokości, na których obserwowana jest temperatura

q - liczba punktów czasowych, w których obserwowana jest temperatura

Przykład

Założenia:

- wielomianowy trend stopnia $p_1 - 1$ dla głębokości
- wielomianowy trend stopnia $q_1 - 1$ dla czasu

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

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

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

Przykład

Założenia:

- wielomianowy trend stopnia $p_1 - 1$ dla głębokości
- wielomianowy trend stopnia $q_1 - 1$ dla czasu
- trend między jeziorami w zależności od lokalizacji (regionu)

Model:

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

$$\mathbf{A} = \text{diag}(\mathbf{1}_{u_1}, \mathbf{1}_{u_2}, \dots, \mathbf{1}_{u_{n_1}}), \quad \sum_{i=1}^{n_1} u_i = n$$

Uogólniony model krzywych wzrostu

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

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

$\boldsymbol{\Sigma} \in \mathbb{R}^{q \times q}$ – **nieznana** macierz p.d.

$\boldsymbol{\Psi} \in \mathbb{R}^{p \times p}$, $\Psi_{11} = 1$ – **znana** lub **nieznana** macierz p.d.

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

$\boldsymbol{\Omega} \in \mathbb{R}^{pq \times pq}$ – **nieznana** macierz p.d.

Macierz śladów częściowych

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

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

Lemat

Niech \mathbf{R} będzie macierzą podzieloną jak wyżej oraz niech \mathbf{G} będzie dowolną macierzą stopnia p . Wówczas

- (i) $(\mathbf{I}_{p^2} \otimes \text{vec}' \mathbf{I}_n)(\mathbf{I}_p \otimes \mathbf{K}_{p,n} \otimes \mathbf{I}_n) \text{vec} \mathbf{R} = \text{vec}(\text{Tr}_n \mathbf{R});$
- (ii) $(\text{vec}' \mathbf{I}_n \otimes \mathbf{I}_p)(\mathbf{I}_n \otimes \mathbf{K}_{n,p} \mathbf{R} \mathbf{K}_{p,n})(\text{vec} \mathbf{I}_n \otimes \mathbf{I}_p) = \text{Tr}_n \mathbf{R};$
- (iii) $\text{Tr}_n[(\mathbf{G} \otimes \mathbf{I}_n) \mathbf{R}] = \mathbf{G} \cdot \text{Tr}_n \mathbf{R}.$

Notacja

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

$\mathcal{R}(\mathbf{G})$ - przestrzeń kolumn macierzy \mathbf{G}

$\mathcal{R}^\perp(\mathbf{G})$ - dopełnienie ortogonalne $\mathcal{R}(\mathbf{G})$

$\mathbf{P}_G = \mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'$ - macierz rzutu ortogonalnego na $\mathcal{R}(\mathbf{G})$

$\mathbf{Q}_G = \mathbf{I}_{m_1} - \mathbf{P}_G$ - macierz rzutu ortogonalnego na $\mathcal{R}^\perp(\mathbf{G})$

$\mathbf{H} : m_1 \times m_1$ - dowolna symetryczna macierz p.d.

$$\mathbf{P}_{G;H} = \mathbf{G}(\mathbf{G}'\mathbf{H}\mathbf{G})^{-1}\mathbf{G}'\mathbf{H}$$

$$\mathbf{Q}_{G;H} = \mathbf{I} - \mathbf{P}_{G;H}$$

\mathbf{G}° dowolna macierz rozpinająca $\mathcal{R}^\perp(\mathbf{G})$

Własności: $\mathbf{Q}_{HG} = \mathbf{P}_{H^{-1}\mathbf{G}^\circ}$, $\mathbf{Q}_{G;H} = \mathbf{P}'_{\mathbf{G}^\circ;H^{-1}}$

$$D(\mathcal{Y}) = \Sigma \otimes \Psi \otimes \mathbf{I}_n, \quad \Psi - \text{znane}$$

Twierdzenie

Estymatorami największej wiarygodności nieznanych parametrów w uogólnionym modelu krzywych wzrostu ze znaną macierzą Ψ są

$$\hat{\chi} = \left[\mathcal{Y}; (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}', (\mathbf{B}'\Psi^{-1}\mathbf{B})^{-1}\mathbf{B}'\Psi^{-1}, (\mathbf{C}'\mathbf{S}_1^{-1}\mathbf{C})^{-1}\mathbf{C}'\mathbf{S}_1^{-1} \right] + \mathcal{Q} \\ - \left[\mathcal{Q}; (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{A}, (\mathbf{B}'\Psi^{-1}\mathbf{B})^{-1}\mathbf{B}'\Psi^{-1}\mathbf{B}, (\mathbf{C}'\mathbf{S}_1^{-1}\mathbf{C})^{-1}\mathbf{C}'\mathbf{S}_1^{-1}\mathbf{C} \right],$$

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

gdzie

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

są macierzami p.d. oraz \mathcal{Q} jest dowolnym tensorem trzeciego rzędu.

$$D(\mathcal{Y}) = \Sigma \otimes \Psi \otimes \mathbf{I}_n, \quad \Psi - \text{znane}$$

Twierdzenie

Estymatorami największej wiarygodności nieznanymi parametrów w uogólnionym modelu krzywych wzrostu ze znaną macierzą Ψ są

$$\begin{aligned} [\hat{\chi}; \mathbf{A}, \mathbf{B}, \mathbf{C}] &= [\mathcal{Y}; \mathbf{P}_A, \mathbf{P}_{B; \Psi^{-1}}, \mathbf{P}_{C; S_1^{-1}}] \\ np\hat{\Sigma} &= \mathbf{S}_1 + \mathbf{Q}_{C; S_1^{-1}} \mathbf{S}_2 \mathbf{Q}'_{C; S_1^{-1}}, \end{aligned}$$

gdzie

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

są macierzami p.d.

$$D(\mathcal{Y}) = \Sigma \otimes \Psi \otimes \mathbf{I}_n, \quad \Psi - \text{nieznane}$$

Twierdzenie

Estymatorami największej wiarygodności nieznanymi parametrów w uogólnionym modelu krzywych wzrostu z nieznaną macierzą Ψ taką, że $\psi_{11} = 1$, są

$$\hat{\chi} = [\mathcal{Y}; (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}', (\mathbf{B}'\hat{\Psi}^{-1}\mathbf{B})^{-1}\mathbf{B}'\hat{\Psi}^{-1}, (\mathbf{C}'\mathbf{S}_3^{-1}\mathbf{C})^{-1}\mathbf{C}'\mathbf{S}_3^{-1}] + \mathcal{Q} \\ - [\mathcal{Q}; (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{A}, (\mathbf{B}'\hat{\Psi}^{-1}\mathbf{B})^{-1}\mathbf{B}'\hat{\Psi}^{-1}\mathbf{B}, (\mathbf{C}'\mathbf{S}_3^{-1}\mathbf{C})^{-1}\mathbf{C}'\mathbf{S}_3^{-1}\mathbf{C}],$$

$$np\hat{\Sigma} = \mathbf{S}_3 + \mathbf{Q}_{C;S_3^{-1}}\mathbf{S}_4\mathbf{Q}'_{C;S_3^{-1}},$$

$$nq\hat{\Psi} = \mathbf{S}_5,$$

gdzie

$$\mathbf{S}_3 = \text{Tr}_{np}[\{\mathbf{I}_q \otimes (\hat{\Psi}^{-1/2} \otimes \mathbf{I}_n)\mathbf{Q}_{(\hat{\Psi}^{-1/2}B \otimes A)}(\hat{\Psi}^{-1/2} \otimes \mathbf{I}_n)\}\text{vec}\mathcal{Y}\text{vec}'\mathcal{Y}],$$

$$\mathbf{S}_4 = \text{Tr}_{np}[\{\mathbf{I}_q \otimes (\hat{\Psi}^{-1/2} \otimes \mathbf{I}_n)\mathbf{P}_{(\hat{\Psi}^{-1/2}B \otimes A)}(\hat{\Psi}^{-1/2} \otimes \mathbf{I}_n)\}\text{vec}\mathcal{Y}\text{vec}'\mathcal{Y}]$$

są macierzami p.d. oraz $\mathbf{S}_5 = \sum_{k=1}^q \sum_{\ell=1}^q \text{Tr}_n(s_{k\ell}\hat{\Gamma}_{k\ell})$, gdzie $s_{k\ell}$ i $\hat{\Gamma}_{k\ell}$ są odpowiednio (k, ℓ) -tym elementem macierzy $\hat{\Sigma}^{-1}$ i $np \times np$ blokiem macierzy $\text{vec}(\mathcal{Y} - [\hat{\chi}; \mathbf{A}, \mathbf{B}, \mathbf{C}])\text{vec}'(\mathcal{Y} - [\hat{\chi}; \mathbf{A}, \mathbf{B}, \mathbf{C}])$,

natomiast \mathcal{Q} jest dowolnym tensorem trzeciego rzędu.

$$D(\mathcal{Y}) = \Sigma \otimes \Psi \otimes \mathbf{I}_n, \quad \Psi - \text{nieznane}$$

Twierdzenie

Estymatorami największej wiarygodności nieznanymi parametrami w uogólnionym modelu krzywych wzrostu z nieznaną macierzą Ψ taką, że $\psi_{11} = 1$, są

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

gdzie

$$\mathbf{S}_3 = \text{Tr}_{np} \left\{ \left[\mathbf{I}_q \otimes \left(\hat{\Psi}^{-1/2} \otimes \mathbf{I}_n \right) \mathbf{Q}_{\left(\hat{\Psi}^{-1/2} B \otimes A \right)} \left(\hat{\Psi}^{-1/2} \otimes \mathbf{I}_n \right) \right] \text{vec} \mathcal{Y} \text{vec}' \mathcal{Y} \right\},$$

$$\mathbf{S}_4 = \text{Tr}_{np} \left\{ \left[\mathbf{I}_q \otimes \left(\hat{\Psi}^{-1/2} \otimes \mathbf{I}_n \right) \mathbf{P}_{\left(\hat{\Psi}^{-1/2} B \otimes A \right)} \left(\hat{\Psi}^{-1/2} \otimes \mathbf{I}_n \right) \right] \text{vec} \mathcal{Y} \text{vec}' \mathcal{Y} \right\}$$

są macierzami p.d. oraz $\hat{\Gamma}_{k\ell}$ jest (k, ℓ) -tym blokiem wymiaru $np \times np$ macierzy

$$\hat{\Gamma} = \text{vec}(\mathcal{Y} - [\hat{\mathcal{X}}; \mathbf{A}, \mathbf{B}, \mathbf{C}]) \text{vec}'(\mathcal{Y} - [\hat{\mathcal{X}}; \mathbf{A}, \mathbf{B}, \mathbf{C}]).$$

Przykład

Obserwowana zmienna: **temperatura**

$n = 17$ - liczba jezior

$u_1 = 7$ - północ (powyżej równoleżnika 60°)

$u_2 = 10$ - południe (poniżej równoleżnika 60°)

$p = 3$ - liczba głębokości, na których obserwowana jest temperatura (0.5 m, 5 m, 15 m)

$q = 3$ - liczba punktów czasowych, w których obserwowana jest temperatura (lata 1990, 2000, 2009)

Przykład

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

Przykład

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

$\hat{\mathcal{X}}$: $\hat{\mathbf{X}}_{ij1}$ - macierz "przednia", $\hat{\mathbf{X}}_{ij2}$ - macierz "tylna"

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

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

Przykład

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

$\hat{x}_{111} = 13.949$ – punkt przecięcia z osiami układu dla północnych jezior (średnia temperatura na początku eksperymentu w 1990, na głębokości 0 m);

$\hat{x}_{121} = -0.37$ – liniowy trend głębokości dla północnych jezior;

$\hat{x}_{112} = -0.054$ – liniowy trend czasu dla północnych jezior;

$\hat{x}_{122} = 0.003$ – interakcja między głębokością i czasem dla północnych jezior

Regresja (d - głębokość, t - czas):

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

Przykład

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

$\hat{x}_{211} = 18.348$ – punkt przecięcia z osiami układu dla południowych jezior (średnia temperatura na początku eksperymentu w 1990, na głębokości 0 m);

$\hat{x}_{221} = -0.708$ – liniowy trend głębokości dla południowych jezior;

$\hat{x}_{212} = -0.092$ – liniowy trend czasu dla południowych jezior;

$\hat{x}_{222} = -0.001$ – interakcja między głębokością i czasem dla południowych jezior

Regresja (d - głębokość, t - czas):

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

Model

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

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

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

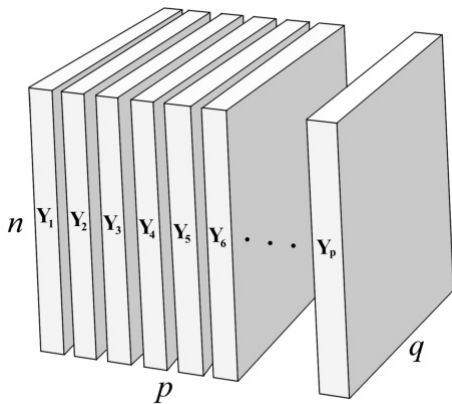
$$\mathbf{Y}^{(3)} \sim N_{nq,p}((\mathbf{C} \otimes \mathbf{A})\mathbf{X}^{(3)}\mathbf{B}', \Sigma \otimes \mathbf{I}_n, \Psi)$$

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

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

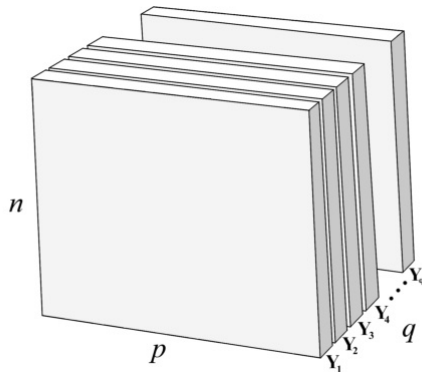
$$\mathbf{X}^{(3)} : n_1 q_1 \times p_1$$

Obserwacje - macierzowanie



Macierzowanie: $\mathbf{Y}_i \in \mathbb{R}^{n \times q}$ zapisane jedna pod drugą –
otrzymujemy $\mathbf{Y}^{(1)} \in \mathbb{R}^{np \times q}$

Obserwacje - macierzowanie



Macierzowanie: $\mathbf{Y}_i \in \mathbb{R}^{n \times p}$ zapisane:
jedna obok drugiej – otrzymujemy $\mathbf{Y}^{(2)} \in \mathbb{R}^{n \times pq}$
jedna pod drugą – otrzymujemy $\mathbf{Y}^{(3)} \in \mathbb{R}^{nq \times p}$

Model

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

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

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

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

Model

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

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

$$\mathbf{Y}^{(2)} \sim N_{n,pq}(\mathbf{A}\mathbf{X}^{(2)}(\mathbf{C}' \otimes \mathbf{B}'), \mathbf{I}_n, \Omega)$$

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

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

Twierdzenie

Estymatorami największej wiarygodności nieznanymi parametrów w modelu $\mathbf{Y}^{(2)}$ z macierzą wariancji-kowariancji $D(\mathcal{Y}) = \Omega \otimes \mathbf{I}_n$ są

$$\widehat{\mathbf{X}}^{(2)} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}^{(2)}\mathbf{S}_3^{-1}(\mathbf{C} \otimes \mathbf{B})\{(\mathbf{C}' \otimes \mathbf{B}')\mathbf{S}_3^{-1}(\mathbf{C} \otimes \mathbf{B})\}^{-1} + (\mathbf{A}')^{-1}\mathbf{Z}_1 + \mathbf{A}'\mathbf{Z}_2(\mathbf{C}' \otimes \mathbf{B}')^{-1},$$

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

gdzie $\mathbf{S}_3 = \mathbf{Y}^{(2)'} \mathbf{Q}_A \mathbf{Y}^{(2)}$ jest macierzą p.d. oraz $\mathbf{Z}_1, \mathbf{Z}_2$ są dowolnymi macierzami.

Tensor rzędu $k > 3$

$$\mathcal{Y} \sim N_{n, p_{k-1}, \dots, p_1}([\mathcal{X}; \mathbf{A}, \mathbf{B}_{k-1}, \dots, \mathbf{B}_1], \mathbf{I}_n, \Psi_{k-1}, \dots, \Psi_1)$$

$\mathbf{A} \in \mathbb{R}^{n \times n^*}$, $\mathbf{B}_i \in \mathbb{R}^{p_i \times p_i^*}$ – znane macierze układu,

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

$\Psi_i \in \mathbb{R}^{p_i \times p_i}$ – znane/nieznane macierze p.d.

Tensor rzędu $k > 3$

$$\mathcal{Y} \sim N_{n,p_{k-1},\dots,p_1}([\mathcal{X}; \mathbf{A}, \mathbf{B}_{k-1}, \dots, \mathbf{B}_1], \mathbf{I}_n, \Psi_{k-1}, \dots, \Psi_1)$$

$$D(\mathcal{Y}) = \Psi_1 \otimes \dots \otimes \Psi_{k-1} \otimes \mathbf{I}_n, \quad \Psi_1 \text{ nieznane}$$

$$D(\mathcal{Y}) = \Psi_1 \otimes \dots \otimes \Psi_{k-1} \otimes \mathbf{I}_n, \quad \Psi_1, \Psi_2 \text{ nieznane,}$$

$$D(\mathcal{Y}) = \Omega \otimes \mathbf{I}_n, \quad \Omega \text{ nieznane.}$$

Tensor rzędu $k > 3$

$$\mathcal{Y} \sim N_{n, p_{k-1}, \dots, p_1}([\mathcal{X}; \mathbf{A}, \mathbf{B}_{k-1}, \dots, \mathbf{B}_1], \mathbf{I}_n, \Psi_{k-1}, \dots, \Psi_1)$$

Macierzowanie:

$$\mathbf{Y} \sim N_{np_{k-1} \dots p_2, p_1} \left(\left(\bigotimes_{i=2}^{k-1} \mathbf{B}_i \otimes \mathbf{A} \right) \mathbf{X} \mathbf{B}_1^\top, \bigotimes_{i=2}^{k-1} \Psi_i \otimes \mathbf{I}_n, \Psi_1 \right)$$

Tensor rzędu $k > 3$

Twierdzenie

Estymatorami największej wiarygodności nieznanymi parametrów w uogólnionym modelu krzywych wzrostu rzędu $k > 3$ z macierzą wariancji-kowariancji

$D(\mathcal{Y}) = \Psi_1 \otimes \cdots \otimes \Psi_{k-1} \otimes \mathbf{I}_n$, Ψ_1 nieznanne, są

$$\hat{\mathbf{X}} = (\mathbf{B}'\Psi^{-1}\mathbf{B} \otimes \mathbf{A}'\mathbf{A})^{-1}(\mathbf{B}'\Psi^{-1} \otimes \mathbf{A}')\mathbf{Y}^{(1)}\mathbf{S}_4^{-1}\mathbf{B}_1(\mathbf{B}'_1\mathbf{S}_4^{-1}\mathbf{B}_1)^{-1} \\ + (\mathbf{B}'\Psi^{-1/2} \otimes \mathbf{A}')\mathbf{Z}_1 + (\mathbf{B}'\Psi^{-1/2} \otimes \mathbf{A}')\mathbf{Z}_2\mathbf{B}'_1{}',$$

$$np\hat{\Psi}_1 = \{\mathbf{Y} - (\mathbf{B} \otimes \mathbf{A})\hat{\mathbf{X}}\mathbf{B}'_1\}'(\Psi^{-1} \otimes \mathbf{I}_n)\{\mathbf{Y} - (\mathbf{B} \otimes \mathbf{A})\hat{\mathbf{X}}\mathbf{B}'_1\},$$

gdzie $\mathbf{S}_4 = \mathbf{Y}'(\Psi^{-1/2} \otimes \mathbf{I}_n)\mathbf{Q}_{(\Psi^{-1/2}\mathbf{B} \otimes \mathbf{A})}(\Psi^{-1/2} \otimes \mathbf{I}_n)\mathbf{Y}$ z $\mathbf{B} = \bigotimes_{i=2}^{k-1} \mathbf{B}_i$,
 $\Psi = \bigotimes_{i=2}^{k-1} \Psi_i$ i $p = p_2 \cdots p_{k-1}$ jest macierzą p.d. oraz $\mathbf{Z}_1, \mathbf{Z}_2$ są dowolnymi macierzami.

Tensor rzędu $k > 3$

Twierdzenie

Estymatorami największej wiarygodności nieznanymi parametrów w uogólnionym modelu krzywych wzrostu rzędu $k > 3$ z macierzą wariancji-kowariancji

$D(\mathcal{Y}) = \Psi_1 \otimes \cdots \otimes \Psi_{k-1} \otimes \mathbf{I}_n$, Ψ_1, Ψ_2 nieznanne, są

$$\hat{\mathbf{X}} = (\mathbf{B}'\hat{\Psi}^{-1}\mathbf{B} \otimes \mathbf{A}'\mathbf{A})^{-1}(\mathbf{B}'\hat{\Psi}^{-1} \otimes \mathbf{A}')\mathbf{Y}\mathbf{S}_5^{-1}\mathbf{B}_1(\mathbf{B}'_1\mathbf{S}_5^{-1}\mathbf{B}_1)^{-1} \\ + (\mathbf{B}'\hat{\Psi}^{-1/2} \otimes \mathbf{A}')\mathbf{Z}_1 + (\mathbf{B}'\hat{\Psi}^{-1/2} \otimes \mathbf{A}')\mathbf{Z}_2\mathbf{B}'_1\mathbf{o}' ,$$

$$n_p\hat{\Psi}_1 = \{\mathbf{Y} - (\mathbf{B} \otimes \mathbf{A})\hat{\mathbf{X}}\mathbf{B}'_1\}'(\hat{\Psi}^{-1} \otimes \mathbf{I}_n)\{\mathbf{Y} - (\mathbf{B} \otimes \mathbf{A})\hat{\mathbf{X}}\mathbf{B}'_1\},$$

$$n_q\hat{\Psi}_2 = \text{Tr}_h \left[\{\mathbf{Y} - (\mathbf{B} \otimes \mathbf{A})\hat{\mathbf{X}}\mathbf{B}'_1\}\hat{\Psi}_1^{-1}\{\mathbf{Y} - (\mathbf{B} \otimes \mathbf{A})\hat{\mathbf{X}}\mathbf{B}'_1\}' \right]$$

gdzie $\mathbf{B} = \bigotimes_{i=2}^{k-1} \mathbf{B}_i$, $\hat{\Psi} = \hat{\Psi}_2 \otimes \bigotimes_{i=3}^{k-1} \Psi_i$, $p = p_2 \cdots p_{k-1}$, $q = p_1 p_3 \cdots p_{k-1}$ i
 $h = p_3 \cdots p_{k-1} n$, macierz $\mathbf{S}_5 = \mathbf{Y}'(\hat{\Psi}^{-1/2} \otimes \mathbf{I}_n)\mathbf{Q}_{(\hat{\Psi}^{-1/2} \otimes \mathbf{I}_n)}(\hat{\Psi}^{-1/2} \otimes \mathbf{I}_n)\mathbf{Y}$
jest macierzą p.d. oraz $\mathbf{Z}_1, \mathbf{Z}_2$ są dowolnymi macierzami.

Tensor rzędu $k > 3$

Twierdzenie

Estymatorami największej wiarogodności nieznanych parametrów w uogólnionym modelu krzywych wzrostu rzędu $k > 3$ z macierzą wariancji-kowariancji $D(\mathcal{Y}) = \Omega \otimes \mathbf{I}_n$ są

$$\widehat{\mathbf{X}} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}\mathbf{S}_3^{-1}(\mathbf{C} \otimes \mathbf{B})\{(\mathbf{C}' \otimes \mathbf{B}')\mathbf{S}_3^{-1}(\mathbf{C} \otimes \mathbf{B})\}^{-1} + (\mathbf{A}')^o\mathbf{Z}_1 + \mathbf{A}'\mathbf{Z}_2(\mathbf{C}' \otimes \mathbf{B}')^o,$$

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

gdzie $\mathbf{C} \otimes \mathbf{B} = \bigotimes_{i=1}^k \mathbf{B}_i$, $\mathbf{S}_3 = \mathbf{Y}'\mathbf{Q}_A\mathbf{Y}$ jest macierzą p.d. oraz \mathbf{Z}_1 , \mathbf{Z}_2 są dowolnymi macierzami.

Podstawowa literatura

- Filipiak, K. and D. Klein (2017). Estimation of parameters under a generalized growth curve model. *J. Multivariate Anal.*, DOI: 10.1016/j.jmva.2017.04.005.
- 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. *Sankhyā* 71-A, 137–163.