Mean and covariance models for relational arrays

Peter Hoff

Statistics, Biostatistics and the CSSS
University of Washington
Outline

Introduction and examples

Models for multiway mean structure

Models for separable covariance arrays
Array-valued data

$y_{i,j,k} =$

- $j$th measurement on $i$th subject under condition $k$ (psychometrics)
- type-$k$ relationship between $i$ and $j$ (relational data/network)
- relationship between $i$ and $j$ at time $t$ (dynamic network)
Array-valued data

\[ y_{i,j,k} = \]

- jth measurement on ith subject under condition k (psychometrics)
- type-k relationship between i and j (relational data/network)
- relationship between i and j at time t (dynamic network)
Array-valued data

$y_{i,j,k} =$

- $j$th measurement on $i$th subject under condition $k$ (psychometrics)
- type-$k$ relationship between $i$ and $j$ (relational data/network)
- relationship between $i$ and $j$ at time $t$ (dynamic network)
Dynamic network example

Cold war cooperation and conflict

- 66 countries
- 8 years (1950, 1955, …, 1980, 1985)
- $y_{i,j,t} =$ relation between $i,j$ in year $t$
- also have data on gdp polity

$Y$ is a $66 \times 66 \times 8$ data array

Model for the mean structure of the data:

$$Y \sim M + E$$

Task: Estimate a low-rank $M$
Dynamic network example

Cold war cooperation and conflict

- 66 countries
- 8 years (1950, 1955, ..., 1980, 1985)
- \( y_{i,j,t} \) = relation between \( i, j \) in year \( t \)
- also have data on gdp polity

\( Y \) is a \( 66 \times 66 \times 8 \) data array

Model for the mean structure of the data:

\[ Y \sim M + E \]

Task: Estimate a low-rank \( M \)
Dynamic network example

Cold war cooperation and conflict

- 66 countries
- 8 years (1950, 1955, ..., 1980, 1985)
- $y_{i,j,t} =$ relation between $i,j$ in year $t$
- also have data on gdp polity

$Y$ is a $66 \times 66 \times 8$ data array

Model for the mean structure of the data:

$Y \sim M + E$

Task: Estimate a low-rank $M$
Multivariate relational data example

Yearly change in log exports (2000 dollars) : \( \mathbf{Y} = \{ y_{i,j,k,l} \} \in \mathbb{R}^{30 \times 30 \times 6 \times 10} \)

- \( i \in \{1, \ldots, 30\} \) indexes exporting nation
- \( j \in \{1, \ldots, 30\} \) indexes importing nation
- \( k \in \{1, \ldots, 6\} \) indexes commodity
- \( l \in \{1, \ldots, 10\} \) indexes year

“Replications” over time: \( \mathbf{Y} = \{ \mathbf{Y}_1, \ldots, \mathbf{Y}_{10} \} \)

\[ \mathbf{Y}_t = \mathbf{M} + \mathbf{E}_t \]

- \( \mathbf{M} \in \mathbb{R}^{30 \times 30 \times 6} \), constant over time;
- \( \mathbf{E}_t \in \mathbb{R}^{30 \times 30 \times 6} \), changing over time.

How should the covariance among \( \{ \mathbf{E}_1, \ldots, \mathbf{E}_{10} \} \) be described?
Multivariate relational data example

Yearly change in log exports (2000 dollars) : \( Y = \{ y_{i,j,k,l} \} \in \mathbb{R}^{30 \times 30 \times 6 \times 10} \)

- \( i \in \{1, \ldots, 30\} \) indexes exporting nation
- \( j \in \{1, \ldots, 30\} \) indexes importing nation
- \( k \in \{1, \ldots, 6\} \) indexes commodity
- \( l \in \{1, \ldots, 10\} \) indexes year

“Replications” over time: \( Y = \{ Y_1, \ldots, Y_{10} \} \)

\[
Y_t = M + E_t
\]

- \( M \in \mathbb{R}^{30 \times 30 \times 6} \), constant over time;
- \( E_t \in \mathbb{R}^{30 \times 30 \times 6} \), changing over time.

How should the covariance among \( \{ E_1, \ldots, E_{10} \} \) be described?
Reduced rank models for mean structure

\[ Y = \Theta + E \]

\( \Theta \) contains the “main features” we hope to recover,

\( E \) is “patternless.”

**Matrix decomposition:** If \( \Theta \) is a rank-\( R \) matrix, then

\[
\theta_{i,j} = \langle u_i, v_j \rangle = \sum_{r=1}^{R} u_{i,r} v_{j,r} \quad \Theta = \sum_{r=1}^{R} u_r v_r^T = \sum_{r=1}^{R} u_r \circ v_r
\]

**Array decomposition:** If \( \Theta \) is a rank-\( R \) array, then

\[
\theta_{i,j,k} = \langle u_i, v_j, w_k \rangle = \sum_{r=1}^{R} u_{i,r} v_{j,r} w_{k,r} \quad \Theta = \sum_{r=1}^{R} u_r \circ v_r \circ w_r
\]

Reduced rank models for mean structure

\[ Y = \Theta + E \]

\( \Theta \) contains the “main features” we hope to recover, 
\( E \) is “patternless.”

**Matrix decomposition:** If \( \Theta \) is a rank-\( R \) matrix, then

\[
\theta_{i,j} = \langle u_i, v_j \rangle = \sum_{r=1}^{R} u_{i,r} v_{j,r} \\
\Theta = \sum_{r=1}^{R} u_r v_r^T = \sum_{r=1}^{R} u_r \circ v_r
\]

**Array decomposition:** If \( \Theta \) is a rank-\( R \) array, then

\[
\theta_{i,j,k} = \langle u_i, v_j, w_k \rangle = \sum_{r=1}^{R} u_{i,r} v_{j,r} w_{k,r} \\
\Theta = \sum_{r=1}^{R} u_r \circ v_r \circ w_r
\]

A model-based approach

For a $K$-way array $Y$,

$$Y = \Theta + E$$

$$\Theta = \sum_{r=1}^{R} u_r^{(1)} \circ \cdots \circ u_r^{(K)} \equiv \langle U^{(1)}, \ldots, U^{(K)} \rangle$$

$$u_r^{(k)} \sim \text{multivariate normal}(\mu_k, \Psi_k),$$

with $\{\mu_k, \Psi_k, k = 1, \ldots, K\}$ to be estimated.

Some motivation:

- shrinkage: $\Theta$ contains lots of parameters.
- hierarchical: covariance among columns of $U^{(k)}$ is identifiable.
- estimation: $p(Y|U^{(1)}, \ldots, U^{(K)})$ multimodal, MCMC “stochastic search”
- adaptability: incorporate reduced rank arrays as a model component
  - multilinear predictor in a GLM
  - multilinear effects for regression parameters
A model-based approach

For a $K$-way array $Y$,

$$
Y = \Theta + E
$$

$$
\Theta = \sum_{r=1}^{R} u^{(1)}_r \circ \cdots \circ u^{(K)}_r \equiv \langle U^{(1)}, \ldots, U^{(K)} \rangle
$$

$$
u^{(k)}_1, \ldots, u^{(k)}_{m_k} \overset{iid}{\sim} \text{multivariate normal}(\mu_k, \Psi_k),
$$

with \{\mu_k, \Psi_k, k = 1, \ldots, K\} to be estimated.

Some motivation:

- shrinkage: $\Theta$ contains lots of parameters.
- hierarchical: covariance among columns of $U^{(k)}$ is identifiable.
- estimation: $p(Y|U^{(1)}, \ldots, U^{(K)})$ multimodal, MCMC “stochastic search”
- adaptability: incorporate reduced rank arrays as a model component
  - multilinear predictor in a GLM
  - multilinear effects for regression parameters
A model-based approach

For a $K$-way array $Y$,

\[
Y = \Theta + E
\]

\[
\Theta = \sum_{r=1}^{R} u_r^{(1)} \circ \cdots \circ u_r^{(K)} \equiv \langle U^{(1)}, \ldots, U^{(K)} \rangle
\]

\[
u_1^{(k)}, \ldots, u_{m_k}^{(k)} \overset{iid}{\sim} \text{multivariate normal}(\mu_k, \Psi_k),
\]

with $\{\mu_k, \Psi_k, k = 1, \ldots, K\}$ to be estimated.

Some motivation:

- shrinkage: $\Theta$ contains lots of parameters.
- hierarchical: covariance among columns of $U^{(k)}$ is identifiable.
- estimation: $p(Y|U^{(1)}, \ldots, U^{(K)})$ multimodal, MCMC “stochastic search”
- adaptability: incorporate reduced rank arrays as a model component
  - multilinear predictor in a GLM
  - multilinear effects for regression parameters
A model-based approach

For a $K$-way array $Y$,

$$
Y = \Theta + E
$$

$$
\Theta = \sum_{r=1}^{R} u_r^{(1)} \circ \cdots \circ u_r^{(K)} \equiv \langle U^{(1)}, \ldots, U^{(K)} \rangle
$$

$$
\mu^{(k)}_1, \ldots, \mu^{(k)}_{m_k} \sim \text{iid multivariate normal}(\mu_k, \Psi_k),
$$

with $\{\mu_k, \Psi_k, k = 1, \ldots, K\}$ to be estimated.

**Some motivation:**

- shrinkage: $\Theta$ contains lots of parameters.
- hierarchical: covariance among columns of $U^{(k)}$ is identifiable.
- estimation: $p(Y|U^{(1)}, \ldots, U^{(K)})$ multimodal, MCMC “stochastic search”
- adaptability: incorporate reduced rank arrays as a model component
  - multilinear predictor in a GLM
  - multilinear effects for regression parameters
Simulation study

\[ K = 3 \ , \ R = 4 \ , \ (m_1, m_2, m_3) = (10, 8, 6) \]

1. Generate \( \mathbf{M} \), a random array of roughly full rank
2. Set \( \Theta = \text{ALS}_4(\mathbf{M}) \)
3. Set \( \mathbf{Y} = \Theta + \mathbf{E}, \ \{e_{i,j,k}\} \overset{iid}{\sim} \text{normal}(0, \nu(\Theta)/4) \).

For each of 100 such simulated datasets, we obtain \( \hat{\Theta}_{\text{LS}} \) and \( \hat{\Theta}_{\text{HB}} \).

Questions: How well do \( \hat{\Theta}_{\text{LS}} \) and \( \hat{\Theta}_{\text{HB}} \)
- recover the “truth” \( \Theta \)?
- represent \( \mathbf{Y} \)?
Simulation study

\[ K = 3 \quad , \quad R = 4 \quad , \quad (m_1, m_2, m_3) = (10, 8, 6) \]

1. Generate \( \mathbf{M} \), a random array of roughly full rank
2. Set \( \Theta = \text{ALS}_4(\mathbf{M}) \)
3. Set \( \mathbf{Y} = \Theta + \mathbf{E} \), \( \{e_{i,j,k}\} \overset{iid}{\sim} \text{normal}(0, \nu(\Theta)/4) \).

For each of 100 such simulated datasets, we obtain \( \hat{\Theta}_{\text{LS}} \) and \( \hat{\Theta}_{\text{HB}} \).

Questions: How well do \( \hat{\Theta}_{\text{LS}} \) and \( \hat{\Theta}_{\text{HB}} \)

- recover the “truth” \( \Theta \)?
- represent \( \mathbf{Y} \)?
Simulation study: known rank

- Bayesian MSE vs. least squares MSE
- Bayesian RSS vs. least squares RSS

- Black points: Mode
- Gray points: Mean
Simulation study: misspecified rank

![Box plots comparing log RSS and log MSE for least squares and hierarchical Bayes models with different assumed ranks.](image)

**Least squares**
- Log RSS: The box plots show the distribution of log RSS for different assumed ranks. The error bars indicate the variability around the median.
- Log MSE: Similar box plots for log MSE, showing variability and central tendency.

**Hierarchical Bayes**
- Log RSS: Box plots for log RSS, with error bars indicating spread.
- Log MSE: Box plots for log MSE, showing distribution and variability.

The plots suggest that misspecifying the rank can lead to increased error in both RSS and MSE, with hierarchical Bayes models generally performing better in terms of log MSE compared to least squares.
Mean structure for dynamic networks

\[
\{ Y_1, \ldots, Y_T \} = Y = M + E
\]

\[
M = \sum u_r \circ v_r \circ w_r
\]

What if each \( Y_t \) is symmetric?

\[
M = \sum u_r \circ u_r \circ w_r, \quad u_r \in \mathbb{R}^n, w_r \in \mathbb{R}^T
\]

(INDSCAL (Carroll and Chang, 1970))

Under this model,

\[
M_t = U \Lambda_t U^T
\]

\[
M = \{ M_1, \ldots, M_T \}, \quad \Lambda_t = \text{diag}(w_{1,t}, \ldots, w_{R,t})
\]
Mean structure for dynamic networks

\[
\{Y_1, \ldots, Y_T\} = Y = M + E
\]

\[
M = \sum u_r \circ v_r \circ w_r
\]

What if each \(Y_t\) is symmetric?

\[
M = \sum_r u_r \circ u_r \circ w_r, \quad u_r \in \mathbb{R}^n, w_r \in \mathbb{R}^T
\]

(INDSCAL (Carroll and Chang, 1970))

Under this Model,

\[
M_t = U \Lambda_t U^T
\]

\[
M = \{M_1, \ldots, M_T\}, \quad \Lambda_t = \text{diag}(w_{1,t}, \ldots, w_{R,t})
\]
Dynamic network example

- $y_{i,j,t} \in \{-5, -4, \ldots, +1, +2\}$, the level of military conflict/cooperation
- $x_{i,j,t,1} = \log \text{gdp}_i + \log \text{gdp}_j$, the sum of the log gdps of the two countries;
- $x_{i,j,t,2} = (\log \text{gdp}_i) \times (\log \text{gdp}_j)$, the product of the log gdps;
- $x_{i,j,t,3} = \text{polity}_i \times \text{polity}_j$, where $\text{polity}_i \in \{-1, 0, +1\}$;
- $x_{i,j,t,4} = (\text{polity}_i > 0) \times (\text{polity}_j > 0)$.

Model:

$$
\begin{align*}
  y_{i,j,t} &= f(z_{i,j,t}, c_{-5}, \ldots, c_{+2}) = \max\{y : z_{i,j,t} > c_y\} \\
  z_{i,j,t} &= \beta^T x_{i,j,t} + u_i^T \Lambda_t u_j + e_{i,j,t} \\
  \tilde{Z} &= \{z_{i,j,t} - \beta^T x_{i,j,t}\} = U \Lambda_t U^T + E
\end{align*}
$$
Longitudinal network example
Longitudinal network example
Longitudinal network example
Longitudinal network example
Introduction and examples

Models for multiway mean structure

Longitudinal network example
Longitudinal network example
Longitudinal network example
Covariance structure of multiple relational arrays

Yearly change in log exports (2000 dollars) : \( \mathbf{Y} = \{y_{i,j,k,l}\} \in \mathbb{R}^{30 \times 30 \times 6 \times 10} \)

- \( i \in \{1, \ldots, 30\} \) indexes exporting nation
- \( j \in \{1, \ldots, 30\} \) indexes importing nation
- \( k \in \{1, \ldots, 6\} \) indexes commodity
- \( l \in \{1, \ldots, 10\} \) indexes year

“Replications” over time: \( \mathbf{Y} = \{\mathbf{Y}_1, \ldots, \mathbf{Y}_{10}\} \)

\[
\mathbf{Y}_t = \mathbf{M} + \mathbf{E}_t
\]

- \( \mathbf{M} \in \mathbb{R}^{30 \times 30 \times 6} \), constant over time;
- \( \mathbf{E}_t \in \mathbb{R}^{30 \times 30 \times 6} \), changing over time.

How should the covariance among \( \{\mathbf{E}_1, \ldots, \mathbf{E}_{10}\} \) be described?
Separable covariance via Tucker products

\[
Y = \Theta + E
\]

Decompose \( \Theta \) using the Tucker decomposition (Tucker 1964, 1966):

\[
\theta_{i,j,k} = \sum_{r=1}^{R} \sum_{s=1}^{S} \sum_{t=1}^{T} z_{r,s,t} a_{i,r} b_{j,r} c_{k,r}
\]

\[
\Theta = \mathbf{Z} \times \{ \mathbf{A}, \mathbf{B}, \mathbf{C} \}
\]

- \( \mathbf{Z} \) is the \( R \times S \times T \) core array
- \( \mathbf{A}, \mathbf{B}, \mathbf{C} \) are \( R \times m_1, S \times m_2, T \times m_3 \) matrices.
- \( R, S \) and \( T \) are the 1-rank, 2-rank and 3-rank of \( \Theta \)
- “\( \times \)” is array-matrix multiplication (De Lathauwer et al., 2000)
Separable covariance via Tucker products

Multivariate normal model:

\[ z = \{z_j : j = 1, \ldots, m\} \quad \text{iid} \quad \sim \quad \text{normal}(0, 1) \]
\[ y = \mu + Az \quad \sim \quad \text{multivariate normal}(\mu, \Sigma = AA^T) \]

Matrix normal model:

\[ Z = \{z_{i,j}\}_{i=1,j=1}^{m_1,m_2} \quad \text{iid} \quad \sim \quad \text{normal}(0, 1) \]
\[ Y = M + AZB^T \quad \sim \quad \text{matrix normal}(M, \Sigma_1 = AA^T, \Sigma_2 = BB^T) \]

**NOTE:** \( AZB^T = Z \times \{A, B\} \)

Array normal model:

\[ Z = \{z_{i,j,k}\}_{i=1,j=1,k=1}^{m_1,m_2,m_3} \quad \text{iid} \quad \sim \quad \text{normal}(0, 1) \]
\[ Y = M + Z \times \{A, B, C\} \quad \sim \quad \text{array normal}(M, \Sigma_1 = AA^T, \Sigma_2 = BB^T, \Sigma_3 = CC^T) \]
Separable covariance via Tucker products

Multivariate normal model:

\[ z = \{z_j : j = 1, \ldots, m\} \sim iid \sim \text{normal}(0, 1) \]
\[ y = \mu + Az \sim \text{multivariate normal}(\mu, \Sigma = AA^T) \]

Matrix normal model:

\[ Z = \{z_{i,j}\}_{i=1,j=1}^{m_1,m_2} \sim iid \sim \text{normal}(0, 1) \]
\[ Y = M + AZB^T \sim \text{matrix normal}(M, \Sigma_1 = AA^T, \Sigma_2 = BB^T) \]

**NOTE:** \(AZB^T = Z \times \{A, B\}\)

Array normal model:

\[ Z = \{z_{i,j,k}\}_{i=1,j=1,k=1}^{m_1,m_2,m_3} \sim iid \sim \text{normal}(0, 1) \]
\[ Y = M + Z \times \{A, B, C\} \sim \text{array normal}(M, \Sigma_1 = AA^T, \Sigma_2 = BB^T, \Sigma_3 = CC^T) \]
**Separable covariance via Tucker products**

**Multivariate normal model:**

\[
\mathbf{z} = \{z_j : j = 1, \ldots, m\} \overset{iid}{\sim} \text{normal}(0, 1)
\]

\[
\mathbf{y} = \mu + \mathbf{A}\mathbf{z} \sim \text{multivariate normal}(\mu, \Sigma = \mathbf{A}\mathbf{A}^T)
\]

**Matrix normal model:**

\[
\mathbf{Z} = \{z_{i,j}\}_{i=1,j=1}^{m_1,m_2} \overset{iid}{\sim} \text{normal}(0, 1)
\]

\[
\mathbf{Y} = \mathbf{M} + \mathbf{A}\mathbf{Z}\mathbf{B}^T \sim \text{matrix normal}(\mathbf{M}, \Sigma_1 = \mathbf{A}\mathbf{A}^T, \Sigma_2 = \mathbf{B}\mathbf{B}^T)
\]

**NOTE:** $\mathbf{A}\mathbf{Z}\mathbf{B}^T = \mathbf{Z} \times \{\mathbf{A}, \mathbf{B}\}$

**Array normal model:**

\[
\mathbf{Z} = \{z_{i,j,k}\}_{i=1,j=1,k=1}^{m_1,m_2,m_3} \overset{iid}{\sim} \text{normal}(0, 1)
\]

\[
\mathbf{Y} = \mathbf{M} + \mathbf{Z} \times \{\mathbf{A}, \mathbf{B}, \mathbf{C}\} \sim \text{array normal}(\mathbf{M}, \Sigma_1 = \mathbf{A}\mathbf{A}^T, \Sigma_2 = \mathbf{B}\mathbf{B}^T, \Sigma_3 = \mathbf{C}\mathbf{C}^T)
\]
Separable covariance structure

For the matrix normal model:

\[
\begin{align*}
\text{Cov}[\mathbf{Y}] &= \Sigma_1 \circ \Sigma_2 \\
\text{Cov}[\text{vec}(\mathbf{Y})] &= \Sigma_2 \otimes \Sigma_1 \\
E[\mathbf{Y}\mathbf{Y}^T] &= \Sigma_1 \times \text{tr}(\Sigma_2) \\
E[\mathbf{Y}^T\mathbf{Y}] &= \Sigma_2 \times \text{tr}(\Sigma_1)
\end{align*}
\]

For the array normal model:

\[
\begin{align*}
\text{Cov}[\mathbf{Y}] &= \Sigma_1 \circ \Sigma_2 \circ \Sigma_3 \\
\text{Cov}[\text{vec}(\mathbf{Y})] &= \Sigma_K \otimes \cdots \otimes \Sigma_1 \\
E[\mathbf{Y}(k)\mathbf{Y}(k)^T] &= \Sigma_k \times \prod_{j \neq k} \text{tr}(\Sigma_j)
\end{align*}
\]
International trade example

Yearly change in log exports (2000 dollars)  :  \( Y = \{y_{i,j,k,l}\} \in \mathbb{R}^{30 \times 30 \times 6 \times 7} \)

- \( i \in \{1, \ldots, 30\} \) indexes exporting nation
- \( j \in \{1, \ldots, 30\} \) indexes importing nation
- \( k \in \{1, \ldots, 6\} \) indexes commodity
- \( l \in \{1, \ldots, 10\} \) indexes year

Full “cell means” model:

\[ y_{i,j,k,l} = \mu_{i,j,k} + e_{i,j,k,l} \]

Let \( E = \{e_{i,j,k,l}\} \)

- iid error model: \( E \sim \text{array normal}(0, I, I, I, \sigma^2 I) \)
- vector normal error model: \( E \sim \text{array normal}(0, I, I, \Sigma_3, I) \)
- matrix normal error model: \( E \sim \text{array normal}(0, I, I, \Sigma_3, \Sigma_4) \)
- array normal model: \( E \sim \text{array normal}(0, \Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4) \)
International trade example

Yearly change in log exports (2000 dollars) : \( \mathbf{Y} = \{y_{i,j,k,l}\} \in \mathbb{R}^{30\times30\times6\times7} \)

- \( i \in \{1, \ldots, 30\} \) indexes exporting nation
- \( j \in \{1, \ldots, 30\} \) indexes importing nation
- \( k \in \{1, \ldots, 6\} \) indexes commodity
- \( l \in \{1, \ldots, 10\} \) indexes year

Full “cell means” model:

\[
y_{i,j,k,l} = \mu_{i,j,k} + e_{i,j,k,l}
\]

Let \( \mathbf{E} = \{e_{i,j,k,l}\} \)

- iid error model: \( \mathbf{E} \sim \text{array normal}(0, \mathbf{I}, \mathbf{I}, \mathbf{I}, \sigma^2 \mathbf{I}) \)
- vector normal error model: \( \mathbf{E} \sim \text{array normal}(0, \mathbf{I}, \mathbf{I}, \mathbf{\Sigma}_3, \mathbf{I}) \)
- matrix normal error model: \( \mathbf{E} \sim \text{array normal}(0, \mathbf{I}, \mathbf{I}, \mathbf{\Sigma}_3, \mathbf{\Sigma}_4) \)
- array normal model: \( \mathbf{E} \sim \text{array normal}(0, \mathbf{\Sigma}_1, \mathbf{\Sigma}_2, \mathbf{\Sigma}_3, \mathbf{\Sigma}_4) \)
International trade example

Yearly change in log exports (2000 dollars) : $Y = \{y_{i,j,k,l}\} \in \mathbb{R}^{30 \times 30 \times 6 \times 7}$

- $i \in \{1, \ldots, 30\}$ indexes exporting nation
- $j \in \{1, \ldots, 30\}$ indexes importing nation
- $k \in \{1, \ldots, 6\}$ indexes commodity
- $l \in \{1, \ldots, 10\}$ indexes year

Full “cell means” model:

$$y_{i,j,k,l} = \mu_{i,j,k} + e_{i,j,k,l}$$

Let $E = \{e_{i,j,k,l}\}$

- iid error model: $E \sim \text{array normal}(0, I, I, I, \sigma^2 I)$
- vector normal error model: $E \sim \text{array normal}(0, I, I, \Sigma_3, I)$
- matrix normal error model: $E \sim \text{array normal}(0, I, I, \Sigma_3, \Sigma_4)$
- array normal model: $E \sim \text{array normal}(0, \Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4)$
International trade example

Yearly change in log exports (2000 dollars) : \( Y = \{ y_{i,j,k,l} \} \in \mathbb{R}^{30 \times 30 \times 6 \times 7} \)

- \( i \in \{1, \ldots, 30\} \) indexes exporting nation
- \( j \in \{1, \ldots, 30\} \) indexes importing nation
- \( k \in \{1, \ldots, 6\} \) indexes commodity
- \( l \in \{1, \ldots, 10\} \) indexes year

Full “cell means” model:

\[
y_{i,j,k,l} = \mu_{i,j,k} + e_{i,j,k,l}
\]

Let \( E = \{ e_{i,j,k,l} \} \)

- iid error model: \( E \sim \text{array normal}(0, I, I, I, \sigma^2 I) \)
- vector normal error model: \( E \sim \text{array normal}(0, I, I, \Sigma_3, I) \)
- matrix normal error model: \( E \sim \text{array normal}(0, I, I, \Sigma_3, \Sigma_4) \)
- array normal model: \( E \sim \text{array normal}(0, \Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4) \)
International trade example

Model comparison:

**reduced:** array normal\((0, I, I, \Sigma_3, \Sigma_4)\)

**full:** array normal\((0, \Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4)\)
Introduction and examples

Models for multiway mean structure

Models for separable covariance arrays

International trade example

Graphs showing trade relationships among countries.
Discussion

- Scientific studies increasingly involve data with multiway array structure
  often this structure is unrecognized
  array structure may be present in data, latent data, or parameters

- Model-based versions of array decompositions offer several benefits
  applicability of multiway methods is broadened
  parameter estimates have reduced MSE

- Applicability to dynamic, multivariate relational data
  reduced-rank methods can capture mean structure
  array-matrix transformations can model covariance
• Scientific studies increasingly involve data with multiway array structure
  often this structure is unrecognized
  array structure may be present in data, latent data, or parameters

• Model-based versions of array decompositions offer several benefits
  applicability of multiway methods is broadened
  parameter estimates have reduced MSE

• Applicability to dynamic, multivariate relational data
  reduced-rank methods can capture mean structure
  array-matrix transformations can model covariance
Discussion

• Scientific studies increasingly involve data with multiway array structure
  often this structure is unrecognized
  array structure may be present in data, latent data, or parameters

• Model-based versions of array decompositions offer several benefits
  applicability of multiway methods is broadened
  parameter estimates have reduced MSE

• Applicability to dynamic, multivariate relational data
  reduced-rank methods can capture mean structure
  array-matrix transformations can model covariance