Course Overview

594 Multiway Data Analysis

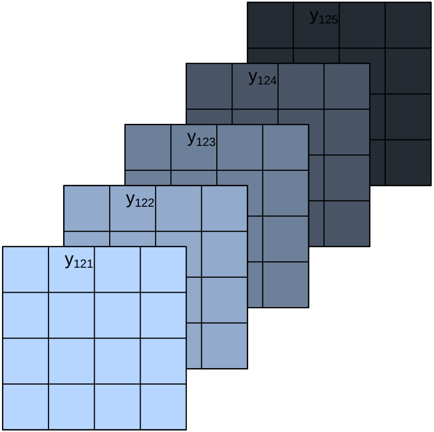
Peter Hoff

Statistics, University of Washington

Array-valued data

$y_{i,j,k} =$

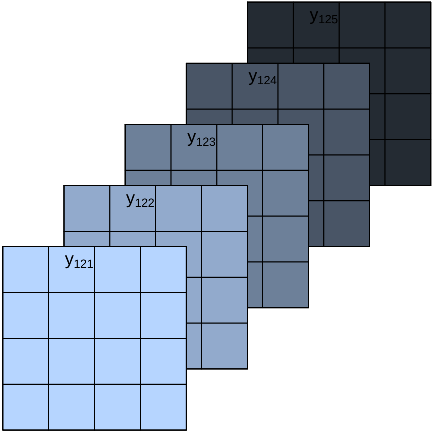
- jth measurement on ith subject under condition k (psychometrics)
- sample mean of variable i for group j in state k (cross-classified data)
- type-k relationship between i and j (multivariate relational data)
- time-k relationship between i and j (dynamic relational data)



Array-valued data

$y_{i,j,k} =$

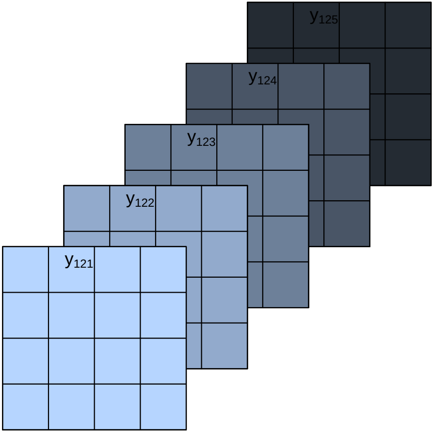
- jth measurement on ith subject under condition k (psychometrics)
- sample mean of variable i for group j in state k (cross-classified data)
- type-k relationship between i and j (multivariate relational data)
- time-k relationship between i and j (dynamic relational data)



Array-valued data

$y_{i,j,k}$ =

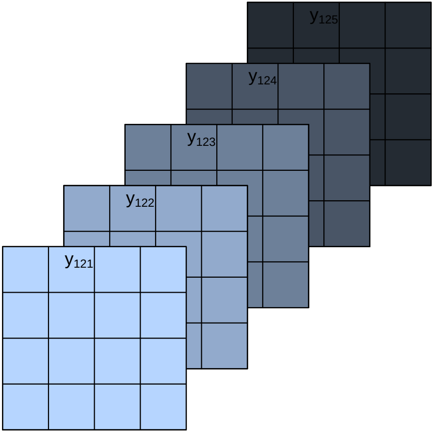
- jth measurement on ith subject under condition k (psychometrics)
- sample mean of variable i for group j in state k (cross-classified data)
- type-k relationship between i and j (multivariate relational data)
- time-k relationship between i and j (dynamic relational data)



Array-valued data

$y_{i,j,k} =$

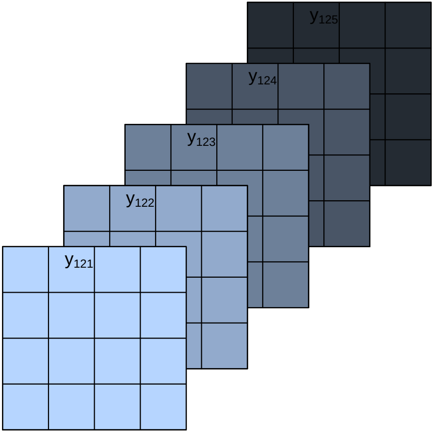
- jth measurement on ith subject under condition k (psychometrics)
- sample mean of variable i for group j in state k (cross-classified data)
- type-k relationship between i and j (multivariate relational data)
- time-k relationship between i and j (dynamic relational data)



Array-valued data

$y_{i,j,k} =$

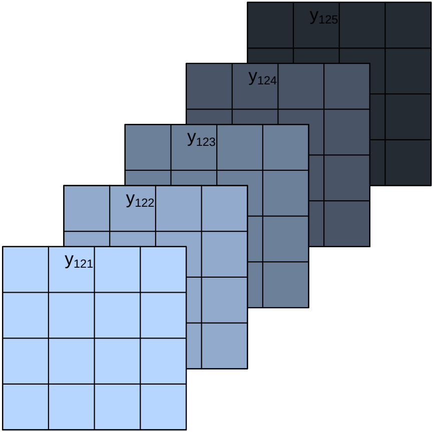
- jth measurement on ith subject under condition k (psychometrics)
- sample mean of variable i for group j in state k (cross-classified data)
- type-k relationship between i and j (multivariate relational data)
- time-k relationship between i and j (dynamic relational data)



Array-valued data

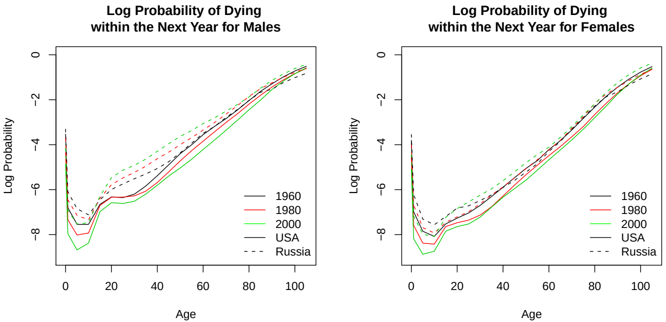
$y_{i,j,k} =$

- jth measurement on ith subject under condition k (psychometrics)
- sample mean of variable i for group j in state k (cross-classified data)
- type-k relationship between i and j (multivariate relational data)
- time-k relationship between i and j (dynamic relational data)



Mortality tables

(Joint work with Bailey Fosdick)



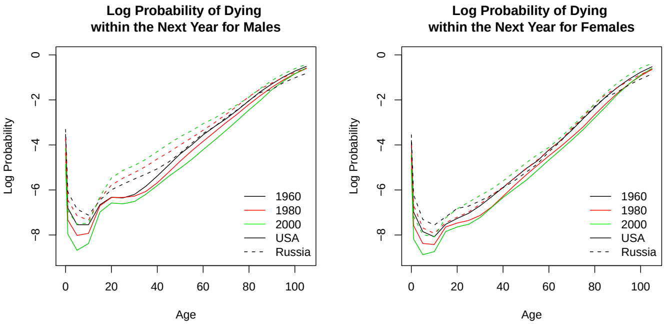
Human Mortality Database: (log) probability of dying in the next year

- 38 countries
- 23 age levels (0, 1 and then every 5 years)
- 9 times periods (1960 to 2000 every 5 years)
- 2 sexes

A $39 \times 23 \times 9 \times 2$-dimensional table.

Mortality tables

(Joint work with Bailey Fosdick)



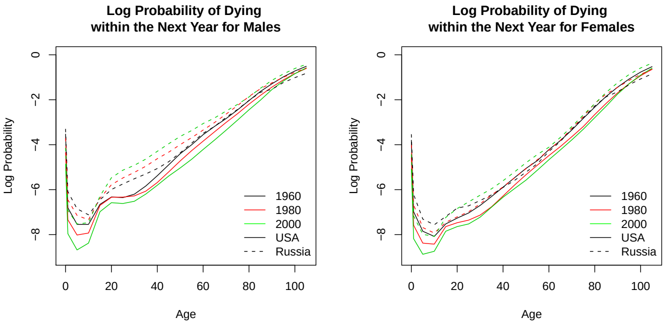
Human Mortality Database: (log) probability of dying in the next year

- 38 countries
- 23 age levels (0, 1 and then every 5 years)
- 9 times periods (1960 to 2000 every 5 years)
- 2 sexes

A $39 \times 23 \times 9 \times 2$-dimensional table.

Mortality tables

(Joint work with Bailey Fosdick)



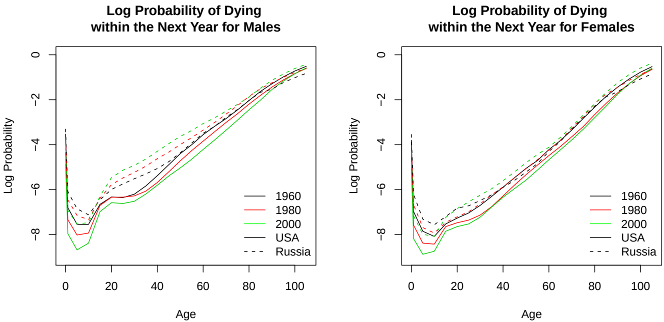
Human Mortality Database: (log) probability of dying in the next year

- 38 countries
- 23 age levels (0, 1 and then every 5 years)
- 9 times periods (1960 to 2000 every 5 years)
- 2 sexes

A $39 \times 23 \times 9 \times 2$-dimensional table.

Mortality tables

(Joint work with Bailey Fosdick)



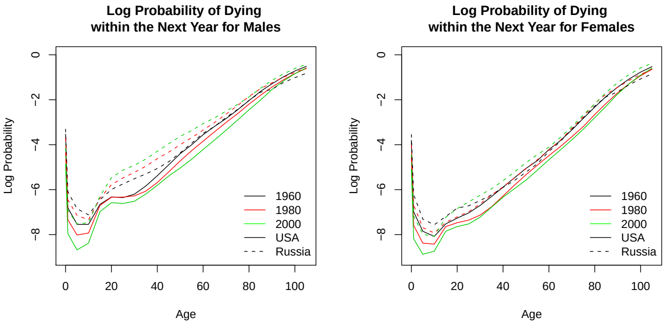
Human Mortality Database: (log) probability of dying in the next year

- 38 countries
- 23 age levels (0, 1 and then every 5 years)
- 9 times periods (1960 to 2000 every 5 years)
- 2 sexes

A $39 \times 23 \times 9 \times 2$-dimensional table.

Mortality tables

(Joint work with Bailey Fosdick)



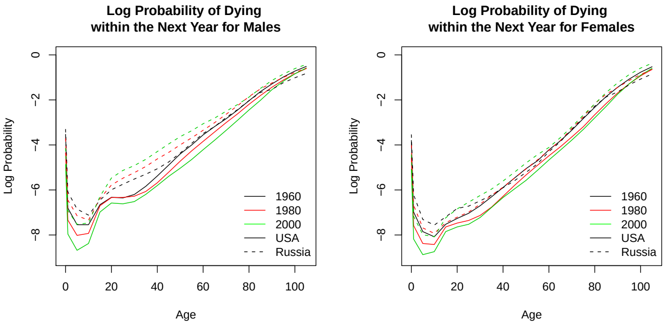
Human Mortality Database: (log) probability of dying in the next year

- 38 countries
- 23 age levels (0, 1 and then every 5 years)
- 9 times periods (1960 to 2000 every 5 years)
- 2 sexes

A $39 \times 23 \times 9 \times 2$-dimensional table.

Mortality tables

(Joint work with Bailey Fosdick)



Human Mortality Database: (log) probability of dying in the next year

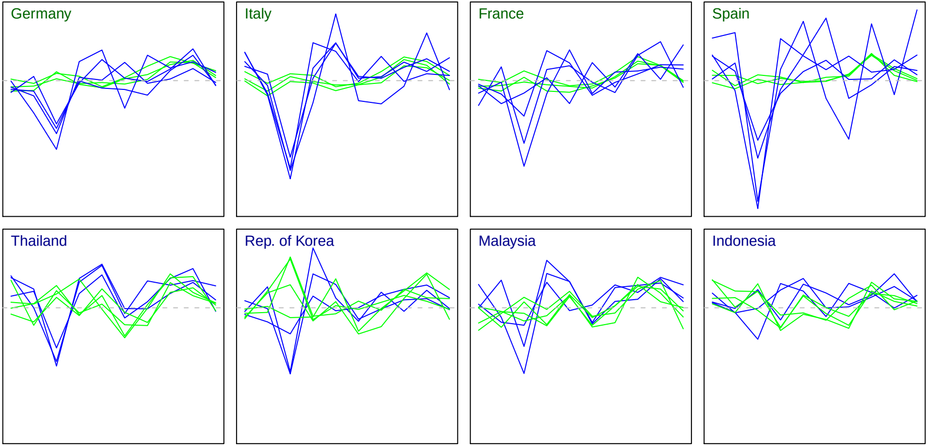
- 38 countries
- 23 age levels (0, 1 and then every 5 years)
- 9 times periods (1960 to 2000 every 5 years)
- 2 sexes

A $39 \times 23 \times 9 \times 2$-dimensional table.

Longitudinal trade

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

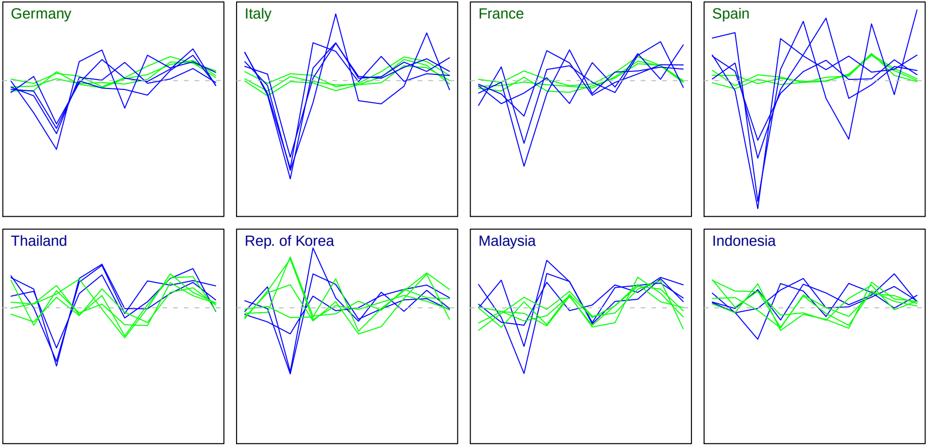
- $i \in \{1,\ldots,30\}$ indexes exporting nation
- $j \in \{1,\ldots,30\}$ indexes importing nation
- $k \in \{1,\ldots,6\}$ indexes commodity
- $l \in \{1,\ldots,10\}$ indexes year



Longitudinal trade

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

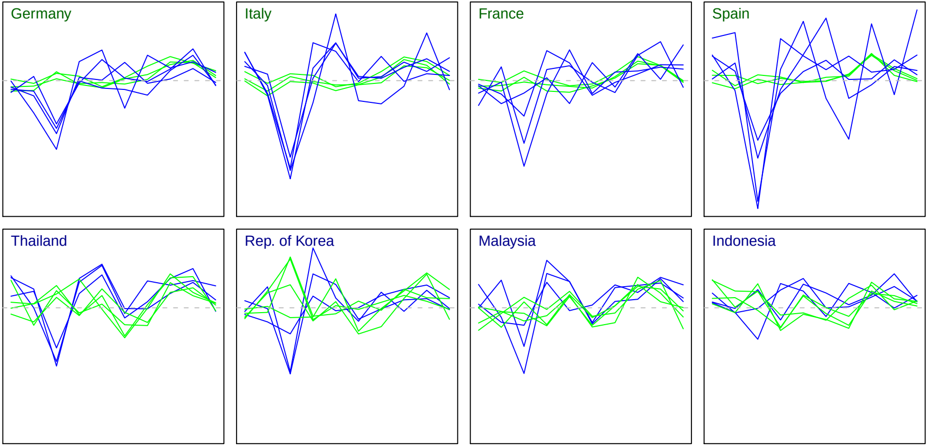
- $i \in \{1, \ldots, 30\}$ indexes exporting nation
- $j \in \{1, \ldots, 30\}$ indexes importing nation
- $k \in \{1, \ldots, 6\}$ indexes commodity
- $l \in \{1, \ldots, 10\}$ indexes year



Longitudinal trade

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

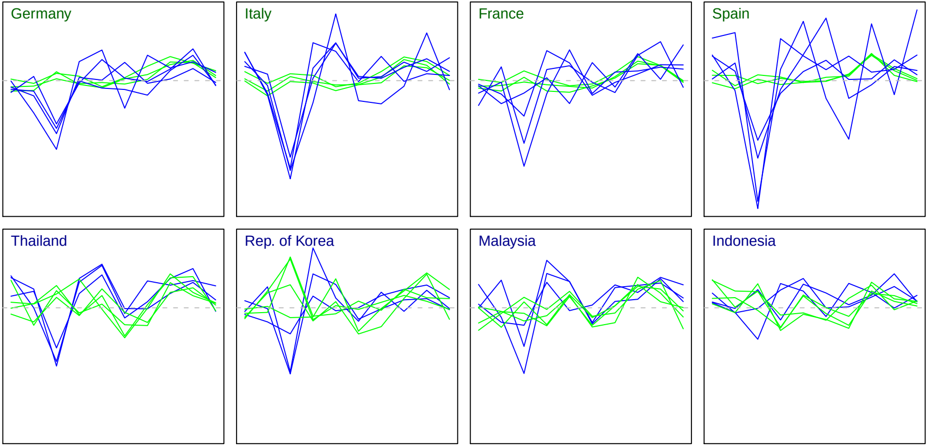
- $i \in \{1,\ldots,30\}$ indexes exporting nation
- $j \in \{1,\ldots,30\}$ indexes importing nation
- $k \in \{1,\ldots,6\}$ indexes commodity
- $l \in \{1,\ldots,10\}$ indexes year



Longitudinal trade

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

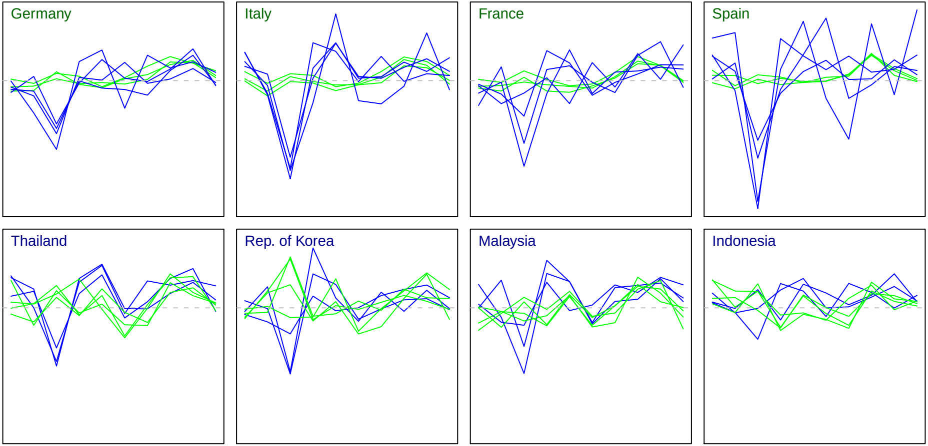
- $i \in \{1,\ldots,30\}$ indexes exporting nation
- $j \in \{1,\ldots,30\}$ indexes importing nation
- $k \in \{1,\ldots,6\}$ indexes commodity
- $l \in \{1,\ldots,10\}$ indexes year



Longitudinal trade

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

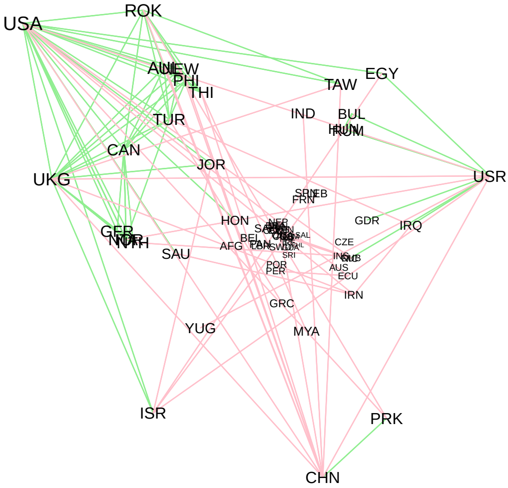
- $i \in \{1, \ldots, 30\}$ indexes exporting nation
- $j \in \{1, \ldots, 30\}$ indexes importing nation
- $k \in \{1, \ldots, 6\}$ indexes commodity
- $l \in \{1, \ldots, 10\}$ indexes year



Longitudinal network example

Cold war cooperation and conflict

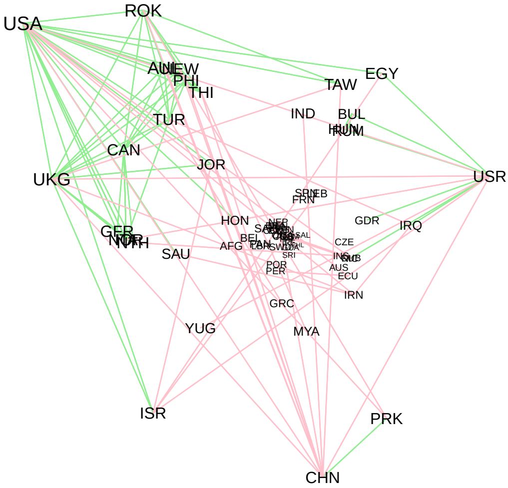
- 66 countries
- 8 years (1950,1955,...,1980,1985)
- $y_{i,j,t}$ =relation between i, j in year t
- also have data on gdp and polity



Longitudinal network example

Cold war cooperation and conflict

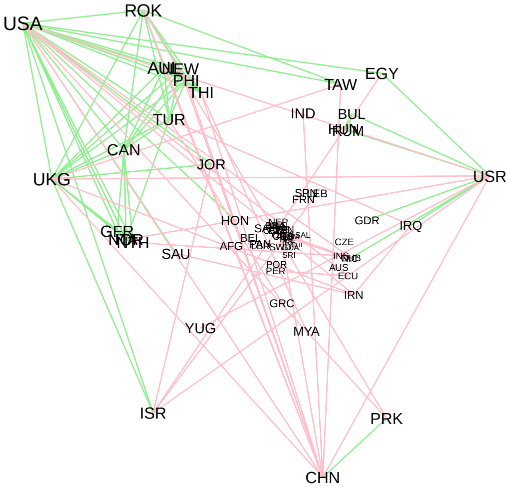
- 66 countries
- 8 years (1950,1955,...,1980,1985)
- $y_{i,j,t}$ =relation between i, j in year t
- also have data on gdp and polity



Longitudinal network example

Cold war cooperation and conflict

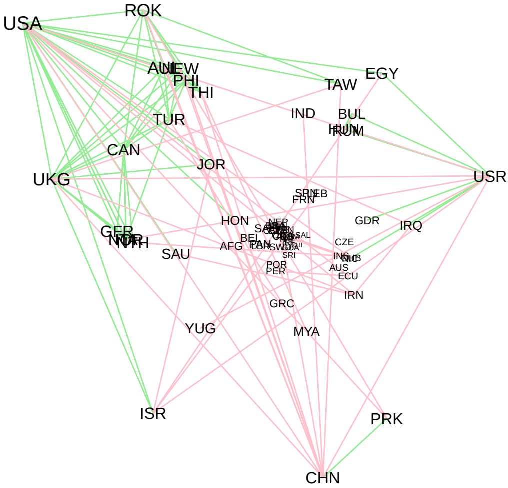
- 66 countries
- 8 years (1950,1955,...,1980,1985)
- $y_{i,j,t}$ =relation between i,j in year t
- also have data on gdp and polity



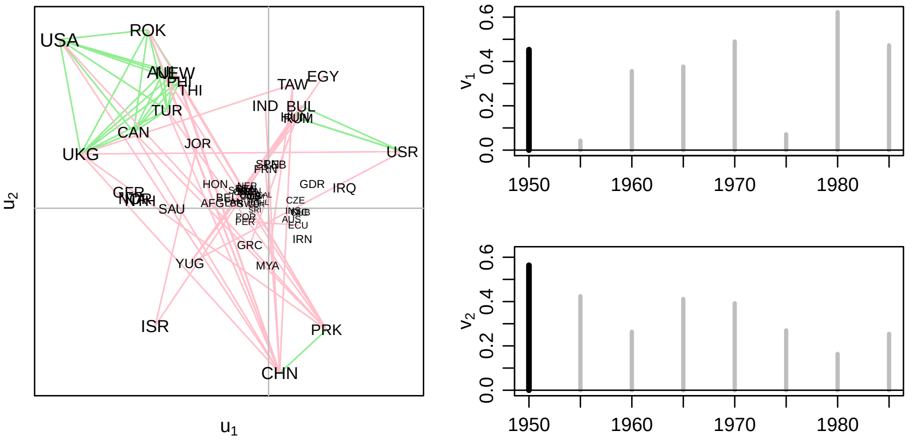
Longitudinal network example

Cold war cooperation and conflict

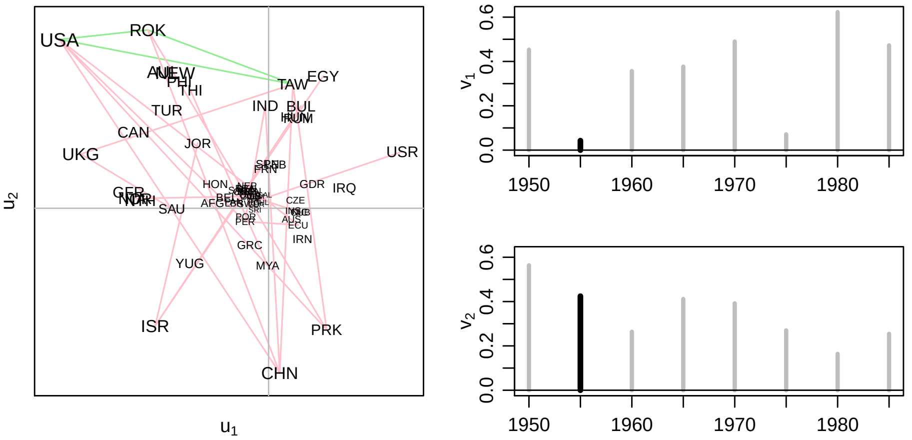
- 66 countries
- 8 years (1950,1955,. . . ,1980,1985)
- $y_{i,j,t}$ =relation between i, j in year t
- also have data on gdp and polity



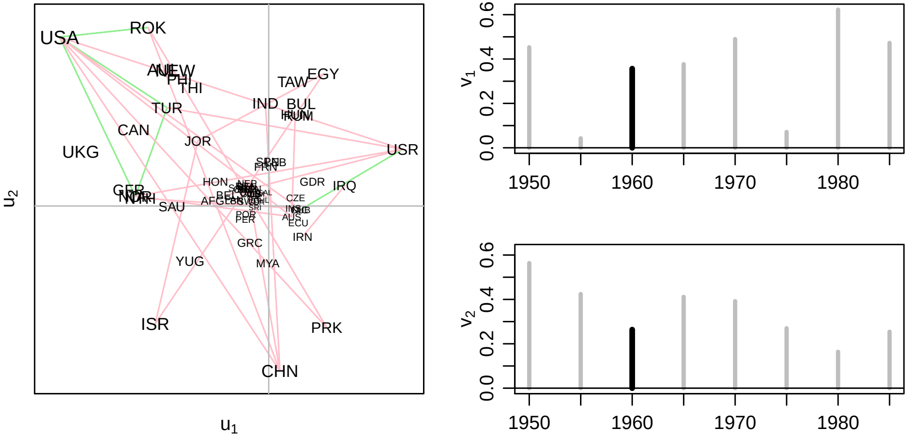
Longitudinal network example



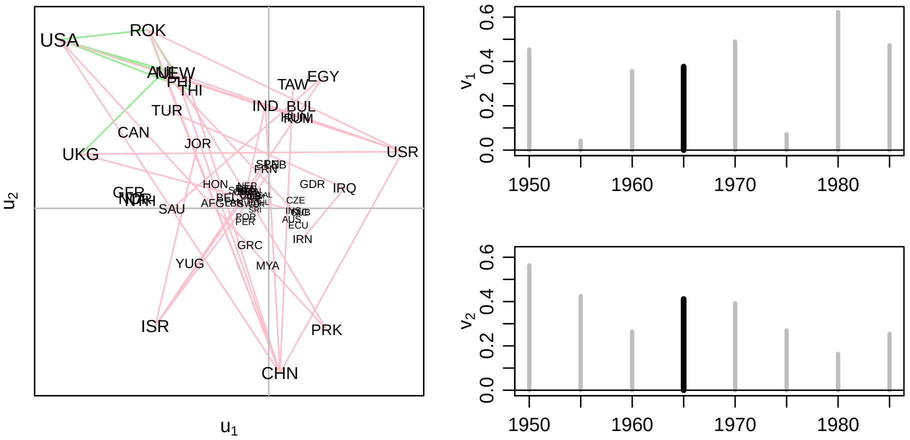
Longitudinal network example



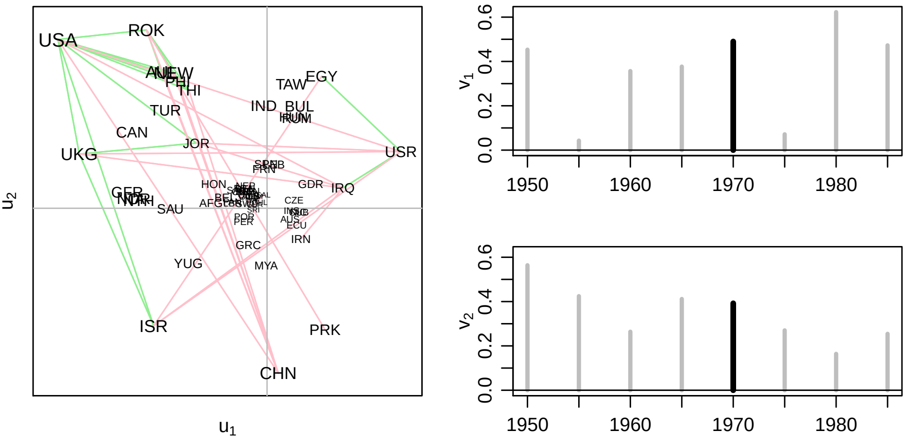
Longitudinal network example



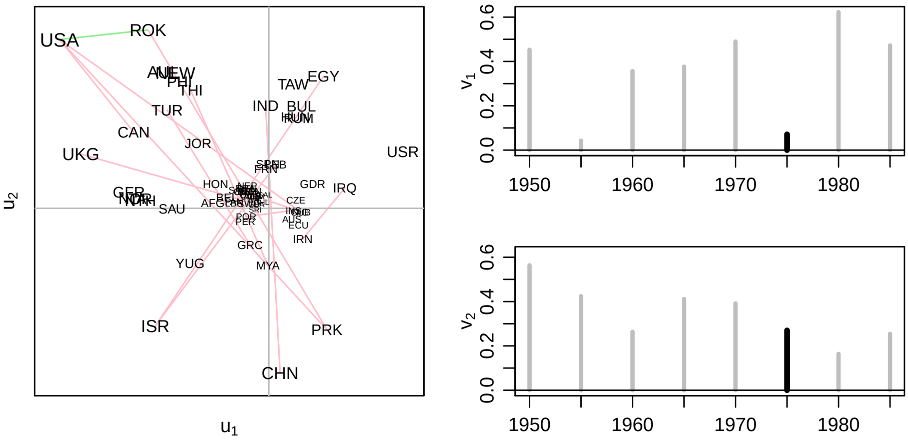
Longitudinal network example



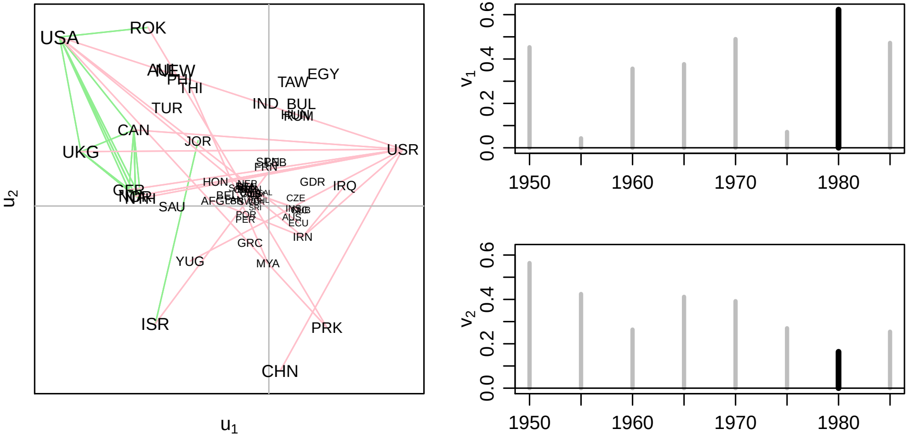
Longitudinal network example



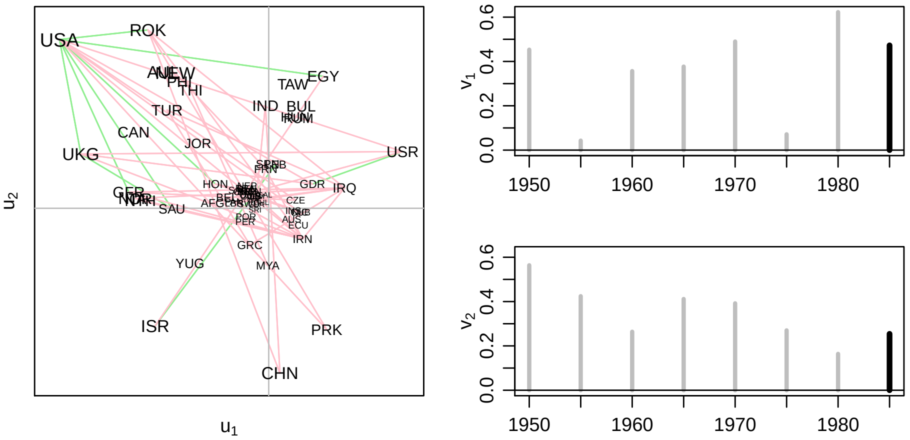
Longitudinal network example



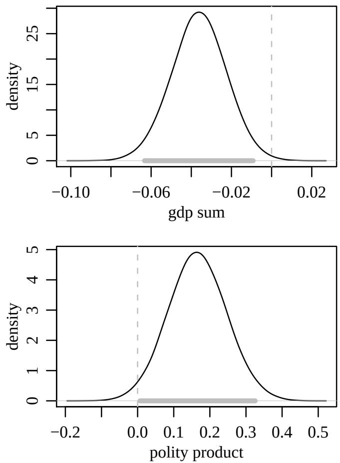
Longitudinal network example



Longitudinal network example



Longitudinal network example



Deep interactions

(Joint work with Alex Volfovsky)

Consider the usual three-factor "ANOVA decomposition" model:

$$\begin{aligned} y_{i,j,k,l} &= \mu_{j,k,l} + \epsilon_{i,j,k,l} \\ &= \mu + \Big[a_j + b_k + c_l\Big] + \Big[(ab)_{j,k} + (ac)_{j,l} + (bc)_{k,l}\Big] + \Big[(abc)_{j,k,l}\Big] + \epsilon_{i,j,k,l} \end{aligned}$$

Parameters are vectors, matrices and arrays based on three index sets.

Estimation methods:

- OLS estimation
- OLS with reduced model
- Bayes/penalized estimation

Parameters in models for contingency table data are similar.

Deep interactions

(Joint work with Alex Volfovsky)

Consider the usual three-factor "ANOVA decomposition" model:

$$
\begin{aligned}
y_{i,j,k,l} &= \mu_{j,k,l} + \epsilon_{i,j,k,l} \\
&= \mu + \left[a_j + b_k + c_l\right] + \left[(ab)_{j,k} + (ac)_{j,l} + (bc)_{k,l}\right] + \left[(abc)_{j,k,l}\right] + \epsilon_{i,j,k,l}
\end{aligned}
$$

Parameters are vectors, matrices and arrays based on three index sets.

Estimation methods:

- OLS estimation
- OLS with reduced model
- Bayes/penalized estimation

Parameters in models for contingency table data are similar.

Deep interactions

(Joint work with Alex Volfovsky)

Consider the usual three-factor "ANOVA decomposition" model:

$$\begin{aligned} y_{i,j,k,l} &= \mu_{j,k,l} + \epsilon_{i,j,k,l} \\ &= \mu + \left[a_j + b_k + c_l\right] + \left[(ab)_{j,k} + (ac)_{j,l} + (bc)_{k,l}\right] + \left[(abc)_{j,k,l}\right] + \epsilon_{i,j,k,l} \end{aligned}$$

Parameters are vectors, matrices and arrays based on three index sets.

Estimation methods:

- OLS estimation
- OLS with reduced model
- Bayes/penalized estimation

Parameters in models for contingency table data are similar.

Deep interactions

(Joint work with Alex Volfovsky)

Consider the usual three-factor "ANOVA decomposition" model:

$$
\begin{aligned}
y_{i,j,k,l} &= \mu_{j,k,l} + \epsilon_{i,j,k,l} \\
&= \mu + \left[a_j + b_k + c_l\right] + \left[(ab)_{j,k} + (ac)_{j,l} + (bc)_{k,l}\right] + \left[(abc)_{j,k,l}\right] + \epsilon_{i,j,k,l}
\end{aligned}
$$

Parameters are vectors, matrices and arrays based on three index sets.

Estimation methods:

- OLS estimation
- OLS with reduced model
- Bayes/penalized estimation

Parameters in models for contingency table data are similar.

Deep interactions

(Joint work with Alex Volfovsky)

Consider the usual three-factor "ANOVA decomposition" model:

$$\begin{aligned} y_{i,j,k,l} &= \mu_{j,k,l} + \epsilon_{i,j,k,l} \\ &= \mu + \Big[a_j + b_k + c_l\Big] + \Big[(ab)_{j,k} + (ac)_{j,l} + (bc)_{k,l}\Big] + \Big[(abc)_{j,k,l}\Big] + \epsilon_{i,j,k,l} \end{aligned}$$

Parameters are vectors, matrices and arrays based on three index sets.

Estimation methods:

- OLS estimation
- OLS with reduced model
- Bayes/penalized estimation

Parameters in models for contingency table data are similar.

Deep interactions

(Joint work with Alex Volfovsky)

Consider the usual three-factor "ANOVA decomposition" model:

$$\begin{aligned} y_{i,j,k,l} &= \mu_{j,k,l} + \epsilon_{i,j,k,l} \\ &= \mu + \left[a_j + b_k + c_l\right] + \left[(ab)_{j,k} + (ac)_{j,l} + (bc)_{k,l}\right] + \left[(abc)_{j,k,l}\right] + \epsilon_{i,j,k,l} \end{aligned}$$

Parameters are vectors, matrices and arrays based on three index sets.

Estimation methods:

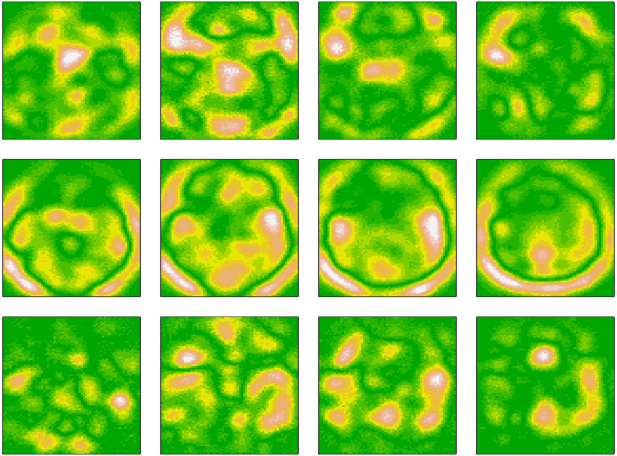
- OLS estimation
- OLS with reduced model
- Bayes/penalized estimation

Parameters in models for contingency table data are similar.

Neuropsychology imaging data

Blood flow in the brain: $\mathbf{Y} = \{y_{i,j,k,l,m,n}\}$

- (i, j, k) index spatial location
- l indexes time
- m indexes treatment/stimulus
- n indexes subject



Data models and Probability models

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{E}$$

$\mathbf{\Theta}$ contains the "main features" we hope to recover,

$\mathbf{E}$ is "patternless."

Data model:

- $\mathbf{\Theta}$ represents main features of the data
- $\mathbf{E}$ represents "residual" features
- Goal is to compactly represent/summarize/describe the data

Probability model:

- $\mathbf{\Theta}$ represents a fixed process or population parameter
- $\mathbf{E}$ represents measurement error or sample-to-sample variation
- Goal is to estimate $\mathbf{\Theta}$ and describe our estimation uncertainty

Data models and Probability models

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{E}$$

$\mathbf{\Theta}$ contains the "main features" we hope to recover,

$\mathbf{E}$ is "patternless."

Data model:

- $\mathbf{\Theta}$ represents main features of the data
- $\mathbf{E}$ represents "residual" features
- Goal is to compactly represent/summarize/describe the data

Probability model:

- $\mathbf{\Theta}$ represents a fixed process or population parameter
- $\mathbf{E}$ represents measurement error or sample-to-sample variation
- Goal is to estimate $\mathbf{\Theta}$ and describe our estimation uncertainty

Data models and Probability models

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{E}$$

$\mathbf{\Theta}$ contains the "main features" we hope to recover,

$\mathbf{E}$ is "patternless."

Data model:

- $\mathbf{\Theta}$ represents main features of the data
- $\mathbf{E}$ represents "residual" features
- Goal is to compactly represent/summarize/describe the data

Probability model:

- $\mathbf{\Theta}$ represents a fixed process or population parameter
- $\mathbf{E}$ represents measurement error or sample-to-sample variation
- Goal is to estimate $\mathbf{\Theta}$ and describe our estimation uncertainty

Data models and Probability models

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{E}$$

$\mathbf{\Theta}$ contains the "main features" we hope to recover,

$\mathbf{E}$ is "patternless."

Data model:

- $\mathbf{\Theta}$ represents main features of the data
- $\mathbf{E}$ represents "residual" features
- Goal is to compactly represent/summarize/describe the data

Probability model:

- $\mathbf{\Theta}$ represents a fixed process or population parameter
- $\mathbf{E}$ represents measurement error or sample-to-sample variation
- Goal is to estimate $\mathbf{\Theta}$ and describe our estimation uncertainty

Data models and Probability models

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{E}$$

$\mathbf{\Theta}$ contains the "main features" we hope to recover,

$\mathbf{E}$ is "patternless."

Data model:

- $\mathbf{\Theta}$ represents main features of the data
- $\mathbf{E}$ represents "residual" features
- Goal is to compactly represent/summarize/describe the data

Probability model:

- $\mathbf{\Theta}$ represents a fixed process or population parameter
- $\mathbf{E}$ represents measurement error or sample-to-sample variation
- Goal is to estimate $\mathbf{\Theta}$ and describe our estimation uncertainty

Data models and Probability models

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{E}$$

$\mathbf{\Theta}$ contains the "main features" we hope to recover,

$\mathbf{E}$ is "patternless."

Data model:

- $\mathbf{\Theta}$ represents main features of the data
- $\mathbf{E}$ represents "residual" features
- Goal is to compactly represent/summarize/describe the data

Probability model:

- $\mathbf{\Theta}$ represents a fixed process or population parameter
- $\mathbf{E}$ represents measurement error or sample-to-sample variation
- Goal is to estimate $\mathbf{\Theta}$ and describe our estimation uncertainty

Data models and Probability models

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{E}$$

$\mathbf{\Theta}$ contains the "main features" we hope to recover,

$\mathbf{E}$ is "patternless."

Data model:

- $\mathbf{\Theta}$ represents main features of the data
- $\mathbf{E}$ represents "residual" features
- Goal is to compactly represent/summarize/describe the data

Probability model:

- $\mathbf{\Theta}$ represents a fixed process or population parameter
- $\mathbf{E}$ represents measurement error or sample-to-sample variation
- Goal is to estimate $\mathbf{\Theta}$ and describe our estimation uncertainty

Mean models

"$\boldsymbol{\Theta}$ = main features" means "$\boldsymbol{\Theta}$ = low dimensional"

$$\text{dimension}(\boldsymbol{\Theta}) < \text{ambient dimension}(\boldsymbol{\Theta}) = \text{dimension}(\mathbf{Y})$$

Modeling possibilities:

- regression: $\boldsymbol{\Theta} = \boldsymbol{\Theta}(\mathbf{B}, \mathbf{X})$, $\mathbf{X}$ known.
- replication: $\boldsymbol{\Theta} = \boldsymbol{\mu} \circ \mathbf{1}$
- rank reduction/factor model: $\boldsymbol{\Theta} = \boldsymbol{\Theta}(\mathbf{A}, \mathbf{B})$

Mean models

"$\mathbf{\Theta}$ = main features" means "$\mathbf{\Theta}$ = low dimensional"

$$\text{dimension}(\mathbf{\Theta}) < \text{ambient dimension}(\mathbf{\Theta}) = \text{dimension}(\mathbf{Y})$$

Modeling possibilities:

- regression: $\mathbf{\Theta} = \mathbf{\Theta}(\mathbf{B}, \mathbf{X})$, $\mathbf{X}$ known.
- replication: $\mathbf{\Theta} = \boldsymbol{\mu} \circ \mathbf{1}$
- rank reduction/factor model: $\mathbf{\Theta} = \mathbf{\Theta}(\mathbf{A}, \mathbf{B})$

Mean models

"$\mathbf{\Theta}$ = main features" means "$\mathbf{\Theta}$ = low dimensional"

$$\text{dimension}(\mathbf{\Theta}) < \text{ambient dimension}(\mathbf{\Theta}) = \text{dimension}(\mathbf{Y})$$

Modeling possibilities:

- regression: $\mathbf{\Theta} = \mathbf{\Theta}(\mathbf{B}, \mathbf{X})$, $\mathbf{X}$ known.
- replication: $\mathbf{\Theta} = \boldsymbol{\mu} \circ \mathbf{1}$
- rank reduction/factor model: $\mathbf{\Theta} = \mathbf{\Theta}(\mathbf{A}, \mathbf{B})$

Mean models

"$\boldsymbol{\Theta}$ = main features" means "$\boldsymbol{\Theta}$ = low dimensional"

$$\text{dimension}(\boldsymbol{\Theta}) < \text{ambient dimension}(\boldsymbol{\Theta}) = \text{dimension}(\mathbf{Y})$$

Modeling possibilities:

- regression: $\boldsymbol{\Theta} = \boldsymbol{\Theta}(\mathbf{B}, \mathbf{X})$, $\mathbf{X}$ known.
- replication: $\boldsymbol{\Theta} = \boldsymbol{\mu} \circ \mathbf{1}$
- rank reduction/factor model: $\boldsymbol{\Theta} = \boldsymbol{\Theta}(\mathbf{A}, \mathbf{B})$

Mean models

"$\mathbf{\Theta}$ = main features" means "$\mathbf{\Theta}$ = low dimensional"

$$\text{dimension}(\mathbf{\Theta}) < \text{ambient dimension}(\mathbf{\Theta}) = \text{dimension}(\mathbf{Y})$$

Modeling possibilities:

- regression: $\mathbf{\Theta} = \mathbf{\Theta}(\mathbf{B}, \mathbf{X})$, $\mathbf{X}$ known.
- replication: $\mathbf{\Theta} = \boldsymbol{\mu} \circ \mathbf{1}$
- rank reduction/factor model: $\mathbf{\Theta} = \mathbf{\Theta}(\mathbf{A}, \mathbf{B})$

PC decomposition

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{E}$$

Matrix decomposition: If $\mathbf{\Theta}$ is a rank-R matrix, then

$$\theta_{i,j} = \langle \mathbf{u}_i, \mathbf{v}_j \rangle = \sum_{r=1}^{R} u_{i,r} v_{j,r} \qquad \mathbf{\Theta} = \sum_{r=1}^{R} \mathbf{u}_r \mathbf{v}_r^T = \sum_{r=1}^{R} \mathbf{u}_r \otimes \mathbf{v}_r$$

Array decomposition: If $\mathbf{\Theta}$ is a rank-R array, then

$$\theta_{i,j,k} = \langle \mathbf{u}_i, \mathbf{v}_j, w_k \rangle = \sum_{r=1}^{R} u_{i,r} v_{j,r} w_{k,r} \qquad \mathbf{\Theta} = \sum_{r=1}^{R} \mathbf{u}_r \otimes \mathbf{v}_r \otimes \mathbf{w}_r$$

(PARAFAC/CANDECOMP: Harshman[1970], Kruskal[1976,1977])

PC decomposition

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{E}$$

Matrix decomposition: If $\mathbf{\Theta}$ is a rank-R matrix, then

$$\theta_{i,j} = \langle \mathbf{u}_i, \mathbf{v}_j \rangle = \sum_{r=1}^{R} u_{i,r} v_{j,r} \qquad \mathbf{\Theta} = \sum_{r=1}^{R} \mathbf{u}_r \mathbf{v}_r^T = \sum_{r=1}^{R} \mathbf{u}_r \otimes \mathbf{v}_r$$

Array decomposition: If $\mathbf{\Theta}$ is a rank-R array, then

$$\theta_{i,j,k} = \langle \mathbf{u}_i, \mathbf{v}_j, w_k \rangle = \sum_{r=1}^{R} u_{i,r} v_{j,r} w_{k,r} \qquad \mathbf{\Theta} = \sum_{r=1}^{R} \mathbf{u}_r \otimes \mathbf{v}_r \otimes \mathbf{w}_r$$

(PARAFAC/CANDECOMP: Harshman[1970], Kruskal[1976,1977])

Tucker decomposition and HOSVD

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{E}$$

Decompose $\mathbf{\Theta}$ using the Tucker decomposition (Tucker 1964,1966):

$$\begin{aligned} \theta_{i,j,k} &= \sum_{r=1}^{R} \sum_{s=1}^{S} \sum_{t=1}^{T} z_{r,s,t} a_{i,r} b_{j,r} c_{k,r} \\ \mathbf{\Theta} &= \mathbf{Z} \times \{\mathbf{A}, \mathbf{B}, \mathbf{C}\} \end{aligned}$$

- $\mathbf{Z}$ 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 $\mathbf{\Theta}$
- "$\times$" is array-matrix multiplication (De Lathauwer et al., 2000)

Tucker decomposition and HOSVD

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{E}$$

Decompose $\mathbf{\Theta}$ using the Tucker decomposition (Tucker 1964,1966):

$$\begin{aligned} \theta_{i,j,k} &= \sum_{r=1}^{R} \sum_{s=1}^{S} \sum_{t=1}^{T} z_{r,s,t} a_{i,r} b_{j,r} c_{k,r} \\ \mathbf{\Theta} &= \mathbf{Z} \times \{\mathbf{A}, \mathbf{B}, \mathbf{C}\} \end{aligned}$$

- $\mathbf{Z}$ 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 $\mathbf{\Theta}$
- "$\times$" is array-matrix multiplication (De Lathauwer et al., 2000)

Tucker decomposition and HOSVD

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{E}$$

Decompose $\mathbf{\Theta}$ using the Tucker decomposition (Tucker 1964,1966):

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

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

- $\mathbf{Z}$ 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 $\mathbf{\Theta}$
- "$\times$" is array-matrix multiplication (De Lathauwer et al., 2000)

Tucker decomposition and HOSVD

$$\mathbf{Y} = \boldsymbol{\Theta} + \mathbf{E}$$

Decompose $\boldsymbol{\Theta}$ using the Tucker decomposition (Tucker 1964,1966):

$$\begin{aligned} \theta_{i,j,k} &= \sum_{r=1}^{R} \sum_{s=1}^{S} \sum_{t=1}^{T} z_{r,s,t} a_{i,r} b_{j,r} c_{k,r} \\ \boldsymbol{\Theta} &= \mathbf{Z} \times \{\mathbf{A}, \mathbf{B}, \mathbf{C}\} \end{aligned}$$

- $\mathbf{Z}$ 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 $\boldsymbol{\Theta}$
- "$\times$" is array-matrix multiplication (De Lathauwer et al., 2000)

Tucker decomposition and HOSVD

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{E}$$

Decompose $\mathbf{\Theta}$ using the Tucker decomposition (Tucker 1964,1966):

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

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

- $\mathbf{Z}$ 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 $\mathbf{\Theta}$
- "$\times$" is array-matrix multiplication (De Lathauwer et al., 2000)

Tucker decomposition and HOSVD

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{E}$$

Decompose $\mathbf{\Theta}$ using the Tucker decomposition (Tucker 1964,1966):

$$\begin{aligned} \theta_{i,j,k} &= \sum_{r=1}^{R} \sum_{s=1}^{S} \sum_{t=1}^{T} z_{r,s,t} a_{i,r} b_{j,r} c_{k,r} \\ \mathbf{\Theta} &= \mathbf{Z} \times \{\mathbf{A}, \mathbf{B}, \mathbf{C}\} \end{aligned}$$

- $\mathbf{Z}$ 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 $\mathbf{\Theta}$
- "$\times$" is array-matrix multiplication (De Lathauwer et al., 2000)

Tucker decomposition and HOSVD

$$\mathbf{Y} = \boldsymbol{\Theta} + \mathbf{E}$$

Decompose $\boldsymbol{\Theta}$ using the Tucker decomposition (Tucker 1964,1966):

$$\begin{aligned} \theta_{i,j,k} &= \sum_{r=1}^{R}\sum_{s=1}^{S}\sum_{t=1}^{T} z_{r,s,t} a_{i,r} b_{j,r} c_{k,r} \\ \boldsymbol{\Theta} &= \mathbf{Z} \times \{\mathbf{A}, \mathbf{B}, \mathbf{C}\} \end{aligned}$$

- $\mathbf{Z}$ 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 $\boldsymbol{\Theta}$
- "$\times$" is array-matrix multiplication (De Lathauwer et al., 2000)

Covariance modeling

Matrix-variate data:

Additive random effects

$$y_{i,j} = \theta_{i,j} + a_i + b_j + \epsilon_{i,j}$$
$$\{a_i\} \sim \text{i.i.d.} N(0, \sigma_a^2)$$
$$\{b_i\} \sim \text{i.i.d.} N(0, \sigma_b^2)$$
$$\{\epsilon_{i,j}\} \sim \text{i.i.d.} N(0, \sigma_\epsilon^2)$$

Multiplicative factors

$$y_{i,j} = \theta_{i,j} + \mathbf{u}_i^T \mathbf{v}_j + \epsilon_{i,j}$$

Multilinear transformations

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{AEB}^T$$

In this case, we have $\text{Cov}[\mathbf{Y}] = \mathbf{AA}^T \circ \mathbf{BB}^T \equiv \Sigma_1 \circ \Sigma_2$.

Covariance modeling

Matrix-variate data:

Additive random effects

$$y_{i,j} = \theta_{i,j} + a_i + b_j + \epsilon_{i,j}$$
$$\{a_i\} \sim \text{i.i.d.} N(0, \sigma_a^2)$$
$$\{b_i\} \sim \text{i.i.d.} N(0, \sigma_b^2)$$
$$\{\epsilon_{i,j}\} \sim \text{i.i.d.} N(0, \sigma_\epsilon^2)$$

Multiplicative factors

$$y_{i,j} = \theta_{i,j} + \mathbf{u}_i^T \mathbf{v}_j + \epsilon_{i,j}$$

Multilinear transformations

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{AEB}^T$$

In this case, we have $\text{Cov}[\mathbf{Y}] = \mathbf{AA}^T \circ \mathbf{BB}^T \equiv \Sigma_1 \circ \Sigma_2$.

Covariance modeling

Matrix-variate data:

Additive random effects

$$
\begin{aligned}
y_{i,j} &= \theta_{i,j} + a_i + b_j + \epsilon_{i,j} \\
\{a_i\} &\sim \text{i.i.d.} N(0, \sigma_a^2) \\
\{b_i\} &\sim \text{i.i.d.} N(0, \sigma_b^2) \\
\{\epsilon_{i,j}\} &\sim \text{i.i.d.} N(0, \sigma_\epsilon^2)
\end{aligned}
$$

Multiplicative factors

$$y_{i,j} = \theta_{i,j} + \mathbf{u}_i^T \mathbf{v}_j + \epsilon_{i,j}$$

Multilinear transformations

$$\mathbf{Y} = \mathbf{\Theta} + \mathbf{A}\mathbf{E}\mathbf{B}^T$$

In this case, we have $\text{Cov}[\mathbf{Y}] = \mathbf{A}\mathbf{A}^T \circ \mathbf{B}\mathbf{B}^T \equiv \Sigma_1 \circ \Sigma_2$.

Statistical analysis of matrix data

1. Reduced rank representations
   1.1 singular value decomposition
   1.2 ANOVA decomposition
2. Error models
   2.1 additive random effects
   2.2 multiplicative effects and the matrix normal model

Statistical analysis of matrix data

1. Reduced rank representations
   1.1 singular value decomposition
   1.2 ANOVA decomposition
2. Error models
   2.1 additive random effects
   2.2 multiplicative effects and the matrix normal model

Statistical analysis of matrix data

1. Reduced rank representations
   1.1 singular value decomposition
   1.2 ANOVA decomposition
2. Error models
   2.1 additive random effects
   2.2 multiplicative effects and the matrix normal model

Statistical analysis of matrix data

1. Reduced rank representations
   1.1 singular value decomposition
   1.2 ANOVA decomposition
2. Error models
   2.1 additive random effects
   2.2 multiplicative effects and the matrix normal model

Statistical analysis of matrix data

1. Reduced rank representations
   1.1 singular value decomposition
   1.2 ANOVA decomposition
2. Error models
   2.1 additive random effects
   2.2 multiplicative effects and the matrix normal model

Statistical analysis of matrix data

1. Reduced rank representations
   1.1 singular value decomposition
   1.2 ANOVA decomposition
2. Error models
   2.1 additive random effects
   2.2 multiplicative effects and the matrix normal model

Statistical analysis of matrix data

1. Reduced rank representations
    1.1 singular value decomposition
    1.2 ANOVA decomposition
2. Error models
    2.1 additive random effects
    2.2 multiplicative effects and the matrix normal model

Statistical analysis of matrix data

1. Reduced rank representations
    1.1 singular value decomposition
    1.2 ANOVA decomposition
2. Error models
    2.1 additive random effects
    2.2 multiplicative effects and the matrix normal model

Higher order array decompositions

1. Notions of tensor rank
2. CANDECOMP/PARAFAC
3. Higher order SVD
4. Theoretical concerns

Higher order array decompositions

1. Notions of tensor rank
2. CANDECOMP/PARAFAC
3. Higher order SVD
4. Theoretical concerns

Higher order array decompositions

1. Notions of tensor rank
2. CANDECOMP/PARAFAC
3. Higher order SVD
4. Theoretical concerns

Higher order array decompositions

1. Notions of tensor rank
2. CANDECOMP/PARAFAC
3. Higher order SVD
4. Theoretical concerns

Higher order array decompositions

1. Notions of tensor rank
2. CANDECOMP/PARAFAC
3. Higher order SVD
4. Theoretical concerns

Higher order array decompositions

1. Notions of tensor rank
2. CANDECOMP/PARAFAC
3. Higher order SVD
4. Theoretical concerns

Estimation methods

1. Alternating least squares
2. Penalized likelihoods
3. Bayesian estimation
4. Equivariant estimation

Estimation methods

1. Alternating least squares
2. Penalized likelihoods
3. Bayesian estimation
4. Equivariant estimation

Estimation methods

1. Alternating least squares
2. Penalized likelihoods
3. Bayesian estimation
4. Equivariant estimation

Estimation methods

1. Alternating least squares
2. Penalized likelihoods
3. Bayesian estimation
4. Equivariant estimation

Estimation methods

1. Alternating least squares
2. Penalized likelihoods
3. Bayesian estimation
4. Equivariant estimation

Estimation methods

1. Alternating least squares
2. Penalized likelihoods
3. Bayesian estimation
4. Equivariant estimation

Covariance models and estimation

1. Additive random effects
2. Multiplicative effects/Multiway factor models
3. Separable covariance models

Covariance models and estimation

1. Additive random effects
2. Multiplicative effects/Multiway factor models
3. Separable covariance models

Covariance models and estimation

1. Additive random effects
2. Multiplicative effects/Multiway factor models
3. Separable covariance models

Covariance models and estimation

1. Additive random effects
2. Multiplicative effects/Multiway factor models
3. Separable covariance models

Covariance models and estimation

1. Additive random effects
2. Multiplicative effects/Multiway factor models
3. Separable covariance models

Statistical concerns

1. Missing data
2. Non-normality/outliers
3. Categorical data
4. Model selection

Statistical concerns

1. Missing data
2. Non-normality/outliers
3. Categorical data
4. Model selection

Statistical concerns

1. Missing data
2. Non-normality/outliers
3. Categorical data
4. Model selection

Statistical concerns

1. Missing data
2. Non-normality/outliers
3. Categorical data
4. Model selection

Statistical concerns

1. Missing data
2. Non-normality/outliers
3. Categorical data
4. Model selection

Statistical concerns

1. Missing data
2. Non-normality/outliers
3. Categorical data
4. Model selection