Estimating Latent Trends in Multivariate Longitudinal Data via Parafac2 with Functional and Structural Constraints

Nathaniel E. Helwig

Assistant Professor of Psychology and Statistics
University of Minnesota (Twin Cities)

Modern Modeling Methods Conference
University of Connecticut
May 23, 2017

Outline of Talk

1) Multiway Data Analysis
- Bilinear Models
- Multiway Extensions

2) Multiway R Package
- Overview of Package
- Constraint Options

3) Multiway Constraints
- Functional Constraints
- Structural Constraints

4) Simulation Study
- Design & Analyses
- Results

5) U.S. Alcohol Consumption
- Data & Analyses
- Results

Multiway Data Analysis

Bilinear Model Form (PCA, FA, ICA)

Let $\mathbf{X} = \{x_{ij}\}_{I \times J}$ where x_{ij} is i-th subject's observed value on the j-th variable.

A bilinear model assumes that

$$\mathbf{X} = \mathbf{AB}' + \mathbf{E} \qquad \longleftrightarrow \qquad x_{ij} = \sum_{r=1}^{R} a_{ir} b_{jr} + e_{ij} \tag{1}$$

where

- $\mathbf{A} = \{a_{ir}\}_{I \times R}$ with a_{ir} denoting the weight (score) of the i-th subject on the r-th factor/component
- $\mathbf{B} = \{b_{jr}\}_{J \times R}$ with b_{jr} denoting the weight (loading) of the j-th variable on the r-th factor/component
- $\mathbf{E} = \{e_{ij}\}_{I \times J}$ with e_{ij} denoting the error term corresponding to x_{ij}

Rotational Indeterminacy Problem

Suppose that $\mathbf{R}$ is an $R \times R$ orthogonal rotation matrix.

- $\mathbf{R}'\mathbf{R} = \mathbf{R}\mathbf{R}' = \mathbf{I}_R$ (identity matrix)

Rotational indeterminacy problem of the bilinear model:

$$\mathbf{A}\mathbf{B}' = \tilde{\mathbf{A}}\tilde{\mathbf{B}}'$$

where $\tilde{\mathbf{A}} = \mathbf{A}\mathbf{R}$ and $\tilde{\mathbf{B}} = \mathbf{B}\mathbf{R}$.

Need to make some assumptions to solve the rotational indeterminacy.

- PCA *assumes* components are orthogonal and explain maximal variance
- ICA *assumes* components are statistically independent

Two-Way versus Three-Way Arrays

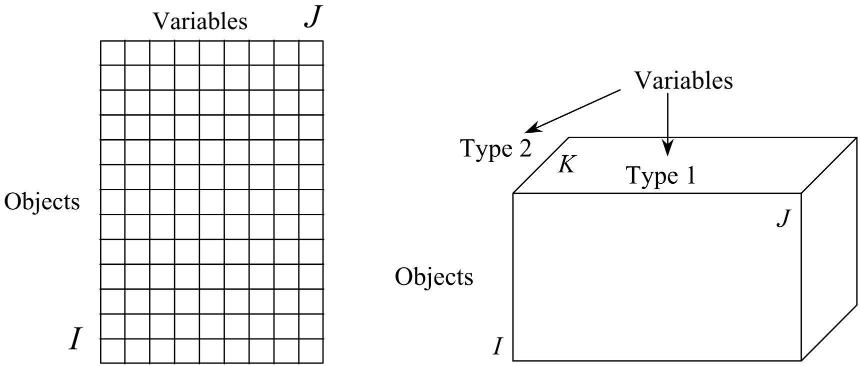

Figure 1: Visualization of 2-way and 3-way arrays from Smilde, Bro, and Geladi (2004).

Talking about Tensors

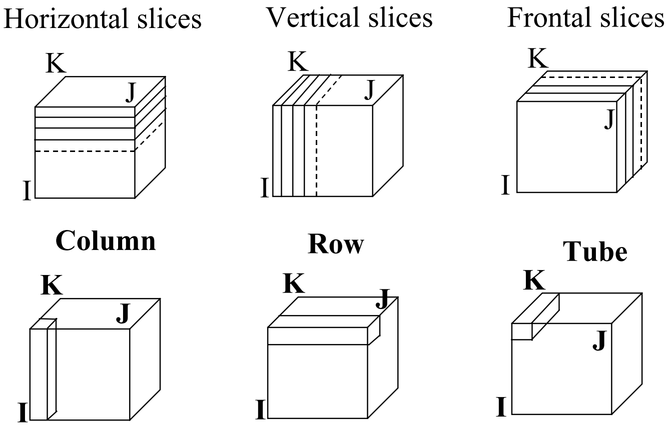

Figure 2: Visualization of three-way array partitions from Smilde, Bro, and Geladi (2004).

"The Covariation Chart" from Cattell (1952)

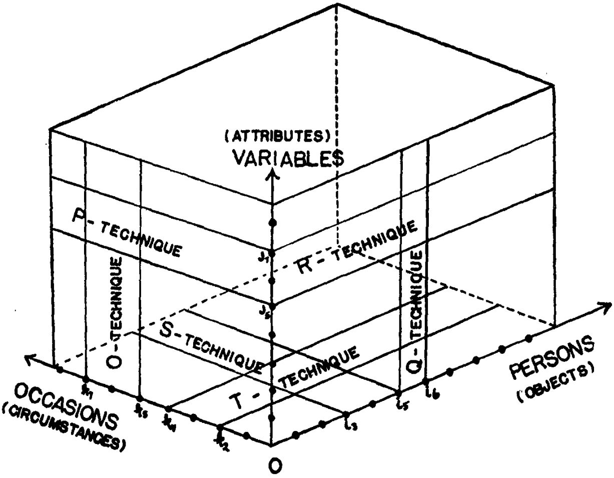

Figure 3: The first illustration of a three-way array.

Tucker's (1966) Three-Way Factor Analysis Model

$$x_{ijk} = \sum_{r=1}^{R} \sum_{s=1}^{S} \sum_{t=1}^{T} g_{rst} a_{ir} b_{js} c_{kt} + e_{ijk}$$

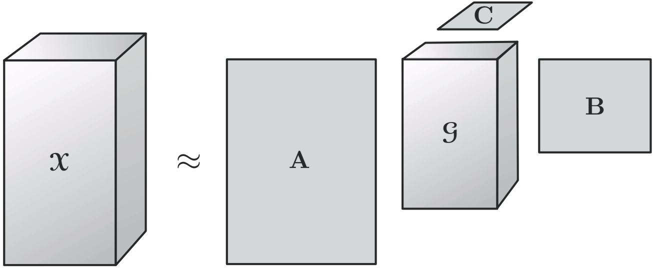

Figure 4: Visualization of Tucker3 structure from Kolda and Bader (2009).

A Popular Model for Three-Way Data

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

Name	Proposed by
Polyadic form of a tensor	Hitchcock, 1927 [105]
PARAFAC (parallel factors)	Harshman, 1970 [90]
CANDECOMP or CAND (canonical decomposition)	Carroll and Chang, 1970 [38]
Topographic components model	Möcks, 1988 [166]
CP (CANDECOMP/PARAFAC)	Kiers, 2000 [122]

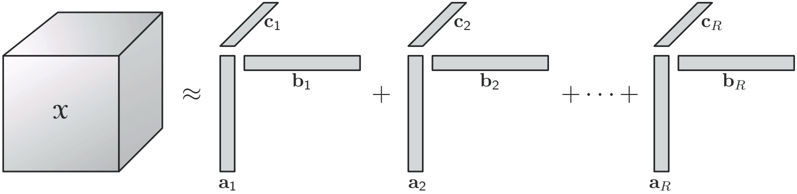

Figure 5: Visualization of trilinear structure from Kolda and Bader (2009).

Harshman's Parafac and Parafac2 Models

Note that the Parafac model (Harshman, 1970) can be written as

$$\mathbf{X}_k = \mathbf{A}\mathbf{C}_k\mathbf{B}' + \mathbf{E}_k$$

where $\mathbf{X}_k = \{x_{ij(k)}\}_{I\times J}$, $\mathbf{C}_k = \mathrm{diag}(c_{k1}, \ldots, c_{kR})$, and $\mathbf{E}_k = \{e_{ij(k)}\}_{I\times J}$.

The Parafac2 model (Harshman, 1972) is more general and can be written as

$$\mathbf{X}_k = \mathbf{A}_k\mathbf{C}_k\mathbf{B}' + \mathbf{E}_k \qquad \text{subject to } \mathbf{A}_k'\mathbf{A}_k = \mathbf{\Phi}$$

where $\mathbf{X}_k = \{x_{ij(k)}\}_{I_k\times J}$, $\mathbf{E}_k = \{e_{ij(k)}\}_{I_k\times J}$, and $\mathbf{\Phi} = \mathbf{A}_k'\mathbf{A}_k$ is the common Mode A cross-product matrix.

Intrinsic Axis Property of Parafac and Parafac2

Parafac and Parafac2 can provide essentially unique solutions.

- No rotational indeterminacy (unlike PCA, ICA, FA, Tucker, etc.)
- Data determines factor configuration and orientation

Parafac uniqueness: If $(\mathbf{A}, \mathbf{B}, \mathbf{C}_k)$ and $(\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}}_k)$ have the same fit, then:

- $\tilde{\mathbf{A}} = \mathbf{A}\mathbf{P}\mathbf{S}_a$ and $\tilde{\mathbf{B}} = \mathbf{B}\mathbf{P}\mathbf{S}_b$ and $\tilde{\mathbf{C}}_k = \mathbf{P}'\mathbf{C}_k\mathbf{P}\mathbf{S}_c$
- $\mathbf{P}$ is an $R \times R$ permutation matrix
- $\mathbf{S}_a$, $\mathbf{S}_b$, and $\mathbf{S}_c$ are diagonal and satisfy $\mathbf{S}_a\mathbf{S}_b\mathbf{S}_c = \mathbf{I}_R$

Parafac2 uniqueness: If $(\mathbf{A}_k, \mathbf{B}, \mathbf{C}_k)$ and $(\tilde{\mathbf{A}}_k, \tilde{\mathbf{B}}, \tilde{\mathbf{C}}_k)$ have the same fit, then:

- $\tilde{\mathbf{A}}_k = z_k\mathbf{A}_k\mathbf{P}\mathbf{S}_a$ and $\tilde{\mathbf{B}} = \mathbf{B}\mathbf{P}\mathbf{S}_b$ and $\tilde{\mathbf{C}}_k = z_k\mathbf{P}'\mathbf{C}_k\mathbf{P}\mathbf{S}_c$
- $\mathbf{P}$ and $\{\mathbf{S}_a, \mathbf{S}_b, \mathbf{S}_c\}$ have the same interpretation
- $z_k \in \{-1, 1\}$ is due to the special sign indeterminacy (see Helwig, 2013)

Simultaneous Component Analysis Models

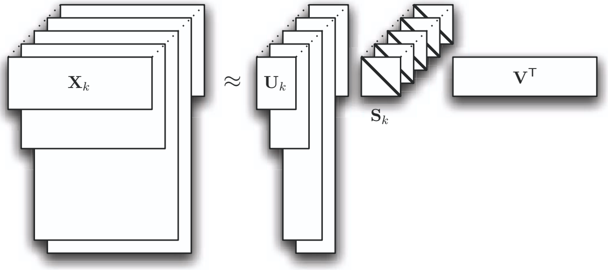

Figure 6: Visualization of SCA structure from Kolda and Bader (2009).

Four versions of SCA which assume different cross-product structures for Mode A weights (Timmerman & Kiers, 2003).

Alternating Least Squares Estimation

Weight matrices are typically estimated via alternating least squares (ALS):

1. Initialize all weight matrices
2. Update each weight matrix given others
3. Repeat Step 2 until convergence

Above algorithm will converge to a locally optimal solution, which depends on the weight matrices initialized in Step 1 of the ALS algorithm.

Should try many random starts of the above ALS algorithm to increase chance of obtaining the globally optimal solution.

Multiway R Package

Fitting Multiway Models in R

`multiway` (Helwig, 2017) is an `R` package (R Core Team, 2017) for fitting multiway models via ALS with optional constraints.

Fit models include:

- Individual Differences Scaling (`indscal`)
- Parallel Factor Analysis 1 (`parafac`)
- Parallel Factor Analysis 2 (`parafac2`)
- Simultaneous Component Analysis (`sca`)
- Tucker Factor Analysis (`tucker`)

Parafac and Tucker models are implemented for 3-way and 4-way data.

Example Syntax for Fitting Parafac Model

```
> # fit Parafac model (unconstrained)
> pfac <- parafac(X, nfac=3)
> pfac

3-way Parafac with 3 factors

Constraints:
    A    B    C
 none none none

Fit Information:
  SSE = 950.1628
  R^2 = 0.5150726
  GCV = 0.2078407
  EDF = 219

Converged: TRUE (17 iterations)
```

Flexible Parafac and Parafac2 Fitting

`parafac` and `parafac2` allow the user to:

- Fix a mode's weights using `fixed` arguments
- Constrain structure of a mode's weights using `struc` arguments

`const` argument allows user to set constraints for each mode's weights:

(0) Unconstrained (default)
(1) Orthogonal
(2) Non-negative
(3) Unimodal
(4) Monotonic
(5) Periodic
(6) Smooth

```
> # fit Parafac model (non-negativity on Modes B and C)
> pfacNN <- parafac(X, nfac=3, const=c(0,2,2))
> pfacNN

3-way Parafac with 3 factors

Constraints:
    A           B           C
 none nonnegative nonnegative

Fit Information:
  SSE = 950.0538
  R^2 = 0.5151283
  GCV = 0.2078168
  EDF = 219

Converged: TRUE (11 iterations)
```

Multiway Constraints

Multiway Models with Functional Weights

Parafac Model: $\mathbf{X}_k = \mathbf{A}\mathbf{C}_k\mathbf{B}' + \mathbf{E}_k$ for $k = 1, \ldots, K$

Assume that Mode A is the functional mode:

$$\mathbf{A} = \begin{pmatrix} \eta_1(1) & \eta_2(1) & \cdots & \eta_R(1) \\ \eta_1(2) & \eta_2(2) & \cdots & \eta_R(2) \\ \vdots & \vdots & \ddots & \vdots \\ \eta_1(I) & \eta_2(I) & \cdots & \eta_R(I) \end{pmatrix}$$

where $\eta_r(\cdot)$ is the r-th component function.

Letting $\{f_1, \ldots, f_\nu\}$ denote a set of known basis functions

$$\eta_r(i) = \sum_{\ell=1}^{\nu} f_\ell(i)\alpha_{\ell r}$$

where $\boldsymbol{\alpha}_r = (\alpha_{1r}, \ldots, \alpha_{\ell r})'$ are the unknown basis function coefficients.

Polynomial Functions

If $\eta_r(\cdot)$ is a polynomial function of degree $\nu - 1$, then

$$\eta_r(i) = \sum_{\ell=1}^{\nu} i^{\ell-1} \alpha_{\ell r}$$

and the Parafac model can be written as

$$\mathbf{X}_k = \mathbf{A}\mathbf{C}_k\mathbf{B}' + \mathbf{E}_k = \mathbf{F}\boldsymbol{\alpha}\mathbf{C}_k\mathbf{B}' + \mathbf{E}_k$$

where

$$\mathbf{F} = \begin{pmatrix} 1 & i & i^2 & \cdots & i^{\nu-1} \\ 1 & i & i^2 & \cdots & i^{\nu-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & i & i^2 & \cdots & i^{\nu-1} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\alpha} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1R} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2R} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{\nu 1} & \alpha_{\nu 2} & \cdots & \alpha_{\nu R} \end{pmatrix}$$

Polynomial Splines

For smooth functions of an unknown form, we can use polynomial splines.

- Piecewise polynomial functions that join at "knots"
- Formed by taking a linear combination of basis functions: $\mathbf{A} = \mathbf{F}\boldsymbol{\alpha}$
- Can adjust polynomial degree and degrees of freedom (# of knots)

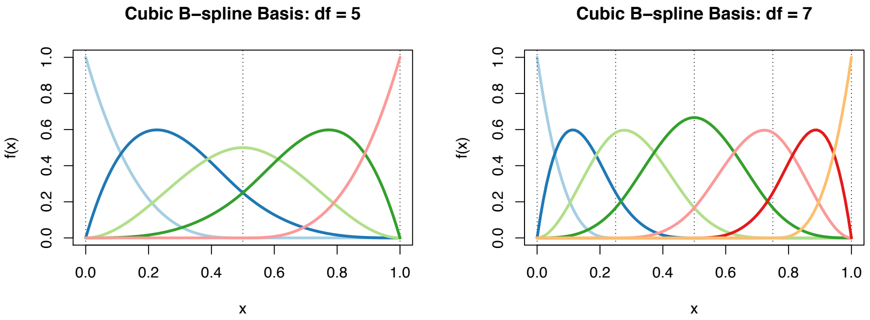

Figure 7: Cubic B-spline basis with 5 (left) and 7 (right) degrees of freedom.

Benefit of Functional Constraints

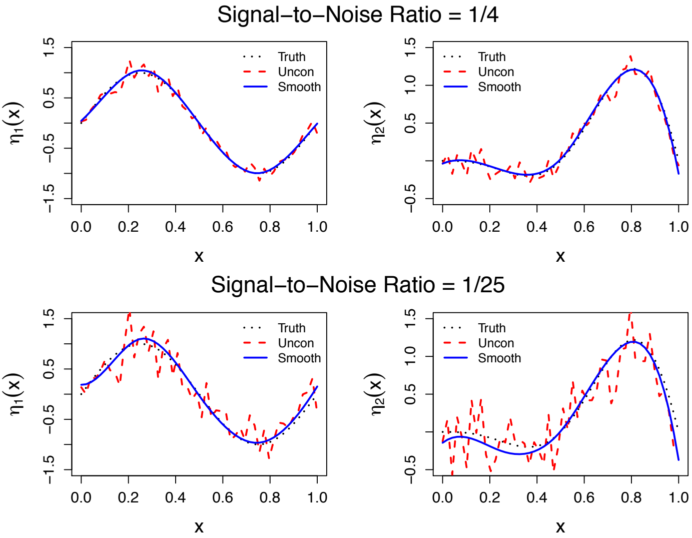

Figure 8: Results from fitting the 3-way Parafac model to a $50 \times 20 \times 10$ tensor with functional Mode A and $R = 2$ factors. SNR = $\|\mathbf{X} - \mathbf{E}\|^2 / \|\mathbf{E}\|^2$

Functional Constraints in Parafac2

Parafac2 Model: $\mathbf{X}_k = \mathbf{A}_k\mathbf{C}_k\mathbf{B}' + \mathbf{E}_k$ subject to $\mathbf{A}_k'\mathbf{A}_k = \mathbf{\Phi}$ for all k

Assume that Mode A is the functional mode:

$$\mathbf{A}_k = \begin{pmatrix} \eta_{1k}(1) & \eta_{2k}(1) & \cdots & \eta_{Rk}(1) \\ \eta_{1k}(2) & \eta_{2k}(2) & \cdots & \eta_{Rk}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \eta_{1k}(I_k) & \eta_{2k}(I_k) & \cdots & \eta_{Rk}(I_k) \end{pmatrix}$$

where $\eta_{rk}(\cdot)$ is the r-th component function for the k-th level of Mode C.

Need component functions to satisfy: $\sum_{i=1}^{I_k} \eta_{rk}(i)\eta_{sk}(i) = \phi_{rs}$ for all k.

- $\mathbf{A}_k = \mathbf{F}_k\boldsymbol{\alpha}_k$ subject to $\boldsymbol{\alpha}_k'\mathbf{F}_k'\mathbf{F}_k\boldsymbol{\alpha}_k = \mathbf{\Phi}$
- Modified ALS algorithm to update $\boldsymbol{\alpha}_k$ matrices (see Helwig, 2016).

Multiway Models with Structured Weights

Parafac Model: $\mathbf{X}_k = \mathbf{A}\mathbf{C}_k\mathbf{B}' + \mathbf{E}_k$ for $k = 1, \ldots, K$

Suppose $\mathbf{B} = \{b_{jr}\}_{J \times R}$ where b_{jr} is weight of j-th variable on r-th factor, and each variable is an indicator for one or more factors.

Need to constrain the weights such that $b_{jr} = 0$ if the j-th observed variable is not an indicator for the r-th factor.

Constrained columnwise update of $\mathbf{B}$ in ALS algorithm (see Helwig, 2016).

Example Structures for Multiway Weights

Table 1: Possible weight structures with $R = 2$ factors.

	Unstructured		Discrete		Overlapping	
	$r = 1$	$r = 2$	$r = 1$	$r = 2$	$r = 1$	$r = 2$
$j = 1$	$*$	$*$	$*$	0	$*$	0
$j = 2$	$*$	$*$	$*$	0	$*$	0
$j = 3$	$*$	$*$	$*$	0	$*$	$*$
$j = 4$	$*$	$*$	0	$*$	$*$	$*$
$j = 5$	$*$	$*$	0	$*$	0	$*$
$j = 6$	$*$	$*$	0	$*$	0	$*$

Note. An entry of "$*$" denotes a non-zero factor loading.

The classic ALS algorithm corresponds to the "Unstructured" weights.

Simulation Study

Parafac2 Simulation Design

Generate data from Parafac2 model with…

$I_k = 44$ measurements on $J = 12$ variables from $K = 51$ units of observation

- Variables 1–4 are indicators for factor 1
- Variables 5–8 are indicators for factor 2
- Variables 9–12 are indicators for factor 3

Latent functions have $\nu = 10$ degrees of freedom and crossproduct matrix

$$\boldsymbol{\Phi} = \begin{pmatrix} 1 & 0.6 & -0.3 \\ 0.6 & 1 & -0.3 \\ -0.3 & -0.3 & 1 \end{pmatrix}$$

Parafac2 Simulation Analyses

Compare four different algorithms:

1. ALS unconstrained
2. ALS with only functional constraints
3. ALS with only structural constraints
4. ALS with both functional and structural constraints

Examine four different SNRs: $\{1/2, 1, 2, 4\}$

- SNR = $\|\mathbf{X} - \mathbf{E}\|^2 / \|\mathbf{E}\|^2$

Use 100 random starts of each algorithm in each condition.

Parafac2 Simulation Results: Parameter Recovery

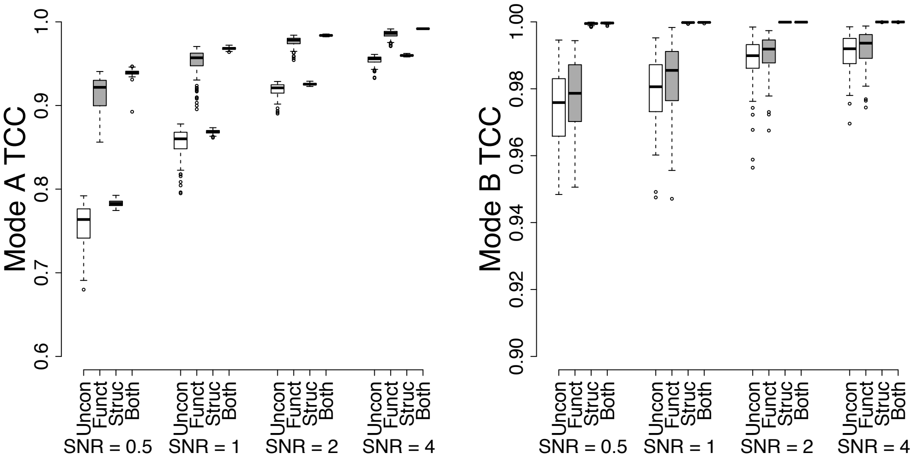

Figure 9: Tucker congruence coefficient (TCC) between true and estimated weights.

Parafac2 Simulation Results: Algorithm Convergence

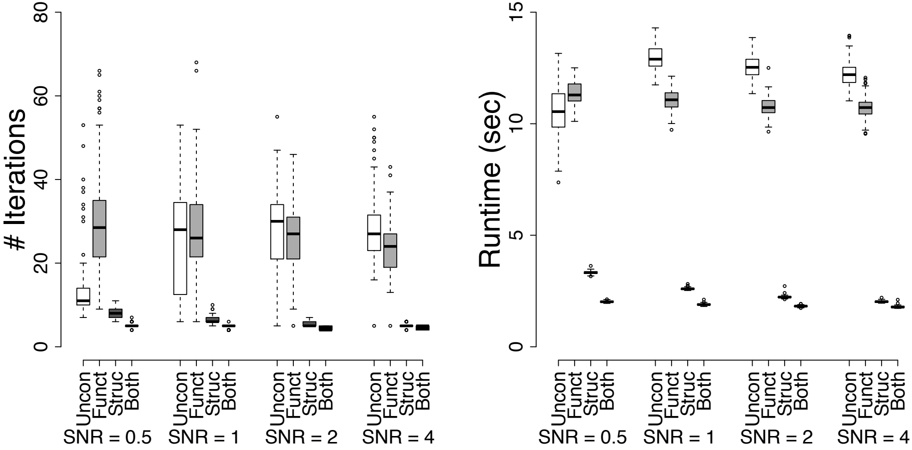

Figure 10: Number of iterations (left) and runtime (right) of the ALS algorithm.

Summary of Simulation Findings

Parameter Recovery:

- Functional constraints improve recovery of Mode A weights
- Structural constraints improve recovery of Mode B weights

Algorithm Convergence:

- Functional constraints have little effect on convergence
- Structural constraints lead to faster convergence

Functional and structural constraints combined show improvements over using either method alone.

United States Alcohol Consumption Example

NIAAA Alcohol Consumption Data from 1970–2013

Yearly consumption data from the 50 United States and the District of Columbia for three types of alcoholic beverages: beer, spirits, and wine.

```
> library(multiway)
> data("USalcohol")
> head(USalcohol)
  year   state region    type beverage ethanol pop14 pop21
1 1970 Alabama  South Spirits     3863 1738.35  2499  2020
2 1970 Alabama  South    Wine     1412  225.92  2499  2020
3 1970 Alabama  South    Beer    33098 1489.41  2499  2020
5 1970  Alaska   West Spirits      945  425.25   205   165
6 1970  Alaska   West    Wine      470   75.20   205   165
7 1970  Alaska   West    Beer     5372  241.74   205   165
```

Data were obtained from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) Surveillance Report #102*

* https://pubs.niaaa.nih.gov/publications/surveillance102/pcyr19702013.txt

Data Tensor: Years $\times$ Variables $\times$ States

Create a tensor of the form 44 years $\times$ 6 variables $\times$ 51 states

- Years: 1970 – 2013
- Variables: Beer (Bev. & Eth.), Spirits (Bev. & Eth.), Wine (Bev. & Eth.)
- States: 50 United States and District of Columbia

Bev. = gallons of beverage consumed per capita age 21+
Eth. = gallons of ethanol consumed per capita age 21+

```
> Xbev <- with(USalcohol,
+              tapply(beverage/pop21, list(year, type, state), c))
> Xeth <- with(USalcohol,
+              tapply(ethanol/pop21, list(year, type, state), c))
> X <- array(0, dim=c(44, 6, 51))
> X[, c(1,3,5) ,] <- Xbev
> X[, c(2,4,6) ,] <- Xeth
```

Assumed Model for Data Tensor

$\mathbf{X}_k$ denotes the 44 time points $\times$ 6 variables data matrix for the k-th state.

Assumed model for the observed data matrices:

$$\mathbf{X}_k = \mathbf{1}\boldsymbol{\mu}_k' + \mathbf{A}_k\mathbf{C}_k\mathbf{B}'\boldsymbol{\Sigma} + \mathbf{E}_k\boldsymbol{\Sigma} \qquad \text{subject to } \mathbf{A}_k'\mathbf{A}_k = \boldsymbol{\Phi}$$

where

- $\boldsymbol{\mu}_k$ is the k-th state's unknown mean vector
- $\boldsymbol{\Sigma} = \text{diag}(\sigma_1, \ldots, \sigma_J)$ is an unknown scaling matrix with $\sigma_j > 0$
- Other terms are Parafac2 weight matrices

Preprocessing the Data Tensor

Before fitting the Parafac2 model, we need to preprocess the data:

1. Center variables across time w/in states (to remove $\boldsymbol{\mu}_k$)
2. Scale variables across time and states (to remove $\boldsymbol{\Sigma}$)

```
> # center each variable across time (within state)
> Xc <- ncenter(X, mode=1)

> # scale each variable (across time and states)
> Xs <- nscale(Xc, mode=2, ssnew=44*51)
```

Now the centered and scaled data tensor `Xs` is ready for Parafac2 fitting.

Plot the Subtracted Means

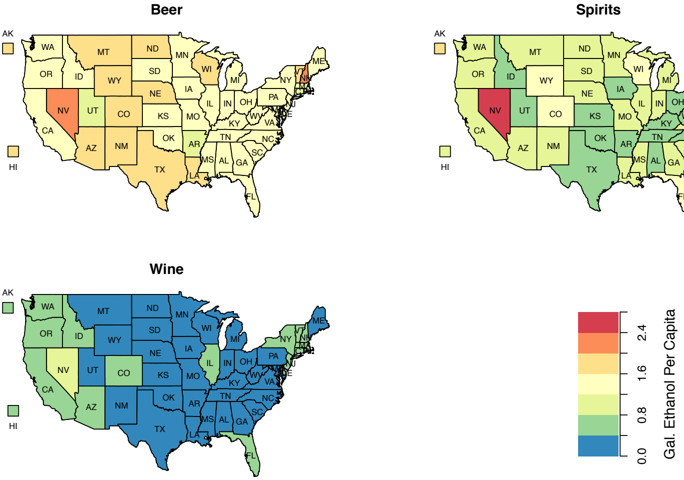

Figure 11: Mean gallons of ethanol consumed per capita for each beverage type.

Plot the Standardized Data

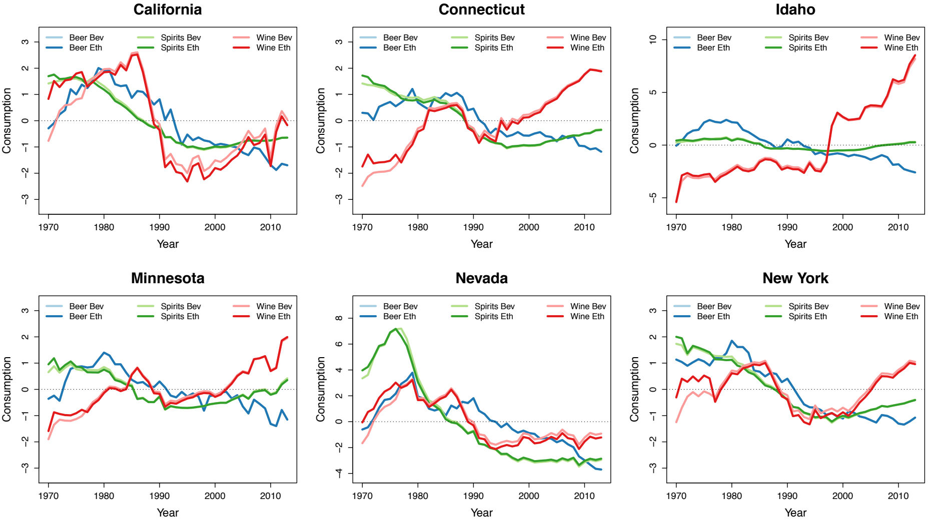

Figure 12: Standardized data for a sample of six states.

Create the Structure Matrix

Create the structure constraint matrix (for the `Bstruc` argument):

```
> Bstruc <- matrix(c(T, T, F, F, F, F,
+                    F, F, T, T, F, F,
+                    F, F, F, F, T, T), nrow=6, ncol=3)
> rownames(Bstruc) <- dnames[[2]]
> colnames(Bstruc) <- paste0("factor",1:3)
> Bstruc
            factor1 factor2 factor3
Beer.bev       TRUE   FALSE   FALSE
Beer.eth       TRUE   FALSE   FALSE
Spirits.bev   FALSE    TRUE   FALSE
Spirits.eth   FALSE    TRUE   FALSE
Wine.bev      FALSE   FALSE    TRUE
Wine.eth      FALSE   FALSE    TRUE
```

Fitting the Parafac Model

Start by fitting the Parafac model to the data:

```
> set.seed(1)
> pfac <- parafac(Xs, nfac=3, nstart=100, const=c(6,0,0), Bstruc=Bstruc)
> pfac

3-way Parafac with 3 factors

Constraints:
      A         B    C
 smooth structure none

Fit Information:
  SSE = 3583.515
  R^2 = 0.7338447
  GCV = 0.2731702
  EDF = 174

Converged: TRUE (4 iterations)
```

Fitting the Parafac2 Model

Now fit the Parafac2 model to the data:

```
> set.seed(1)
> pfac2 <- parafac2(Xs, nfac=3, nstart=100, const=c(6,0,0), Bstruc=Bstruc)
> pfac2

3-way Parafac2 with 3 factors

Constraints:
      A          B    C
 smooth structure none

Fit Information:
  SSE = 1939.451
  R^2 = 0.8559528
  GCV = 0.1660572
  EDF = 924

Converged: TRUE (9 iterations)
```

Average Parafac and Parafac2 Functional Factor Scores

Parafac model essentially captures the average pattern across states.

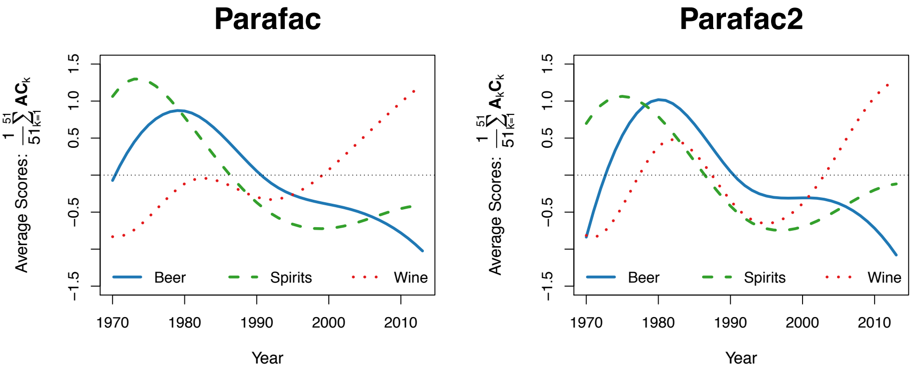

Figure 13: Average (across states) factor scores for Parafac and Parafac2 models.

Parafac2 Functional Factor Scores

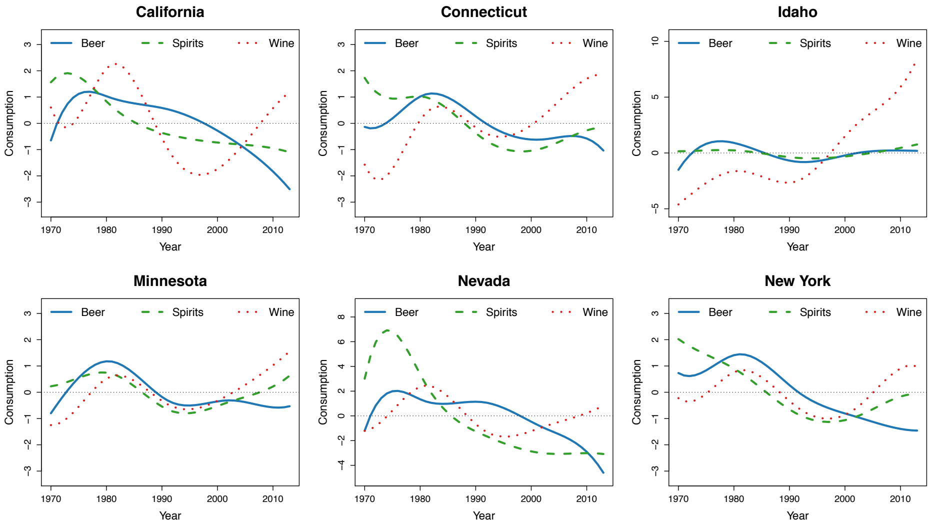

Figure 14: Parafac2 functional factor scores for a sample of six states.

Check Factor Correlations and Loadings

```
> # correlation matrix
> round(pfac2$Phi, 3)
        factor1 factor2 factor3
factor1   1.000   0.643   0.184
factor2   0.643   1.000   0.420
factor3   0.184   0.420   1.000

> # factor loadings
> round(pfac2$B, 3)
            factor1 factor2 factor3
Beer.bev          1   0.000   0.000
Beer.eth          1   0.000   0.000
Spirits.bev       0   1.000   0.000
Spirits.eth       0   0.997   0.000
Wine.bev          0   0.000   1.000
Wine.eth          0   0.000   0.984
```

Parafac2 Mode C Weights (Factor Score SDs)

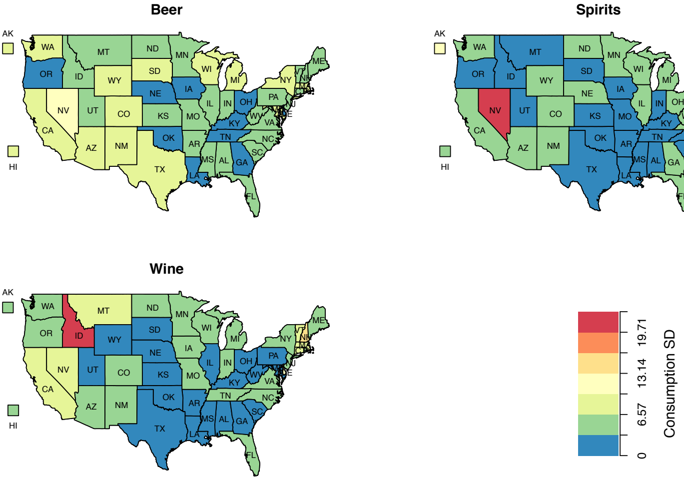

Figure 15: Absolute value of Parafac2 Mode C weights, which are factor score SDs.

Parafac2 Predicted Gallons of Ethanol Consumed per Capita

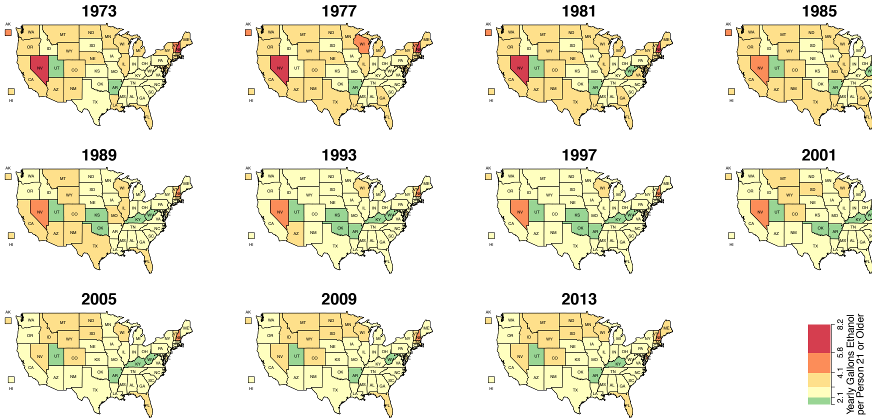

Figure 16: Parafac2 predicted gallons of ethanol consumed per capita for each state.

Summary of Results

Beer and spirits consumption is decreasing; wine consumption is increasing.

- US consumes most of its alcohol in the form of beer
- Trends suggest wine is becoming more popular in many states

State-specific longitudinal differences in alcohol consumption trends exist.

- Some states show little variation in their consumption
- Other states have large fluctuations in particular beverage types

Parafac2 is a powerful model for analyzing multivariate longitudinal data.

- Constraints can improve estimation/interpretation
- Flexible estimates of individual differences in latent trends

References

Cattell, R. B. (1952). The three basis factor-analytic research designs — their interrelations and derivatives. *Psychological Bulletin 49*, 499–520.

Harshman, R. A. (1970). Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis. *UCLA Working Papers in Phonetics 16*, 1–84.

Harshman, R. A. (1972). PARAFAC2: Mathematical and technical notes. *UCLA Working Papers in Phonetics 22*, 30–44.

Helwig, N. E. (2013). The special sign indeterminacy of the direct-fitting Parafac2 model: Some implications, cautions, and recommendations for Simultaneous Component Analysis. *Psychometrika 78*, 725–739.

Helwig, N. E. (2016). Estimating latent trends in multivariate longitudinal data via Parafac2 with functional and structural constraints. *Biometrical Journal [Advance Online Version]*, 1–21.

Helwig, N. E. (2017). *multiway: Component Models for Multi-Way Data*. R package version 1.0-3.

Kolda, T. G. and B. W. Bader (2009). Tensor decompositions and applications. *SIAM Review 51*, 455–500.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Smilde, A., R. Bro, and P. Geladi (2004). *Multi-way Analysis with Applications in the Chemical Sciences*. Chicester, UK: Wiley.

Timmerman, M. E. and H. A. L. Kiers (2003). Four simultaneous component analysis models for the analysis of multivariate time series from more than one subject to model intraindividual and interindividual differences. *Psychometrika 68*, 105–121.

Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika 31*, 279–311.