University of Virginia
Curry School of Education

Application of Generalized Structured Component Analysis to Item Response Theory

Ryoo, J., Jung, K., Hwang, H., Meyer, J. P., Molfese, V., & Brown, E. T.

Modern Modeling Methods (5/25/2016)

Overview

- Introduction to generalized structured component analysis (GSCA)
  - As a structural equation model
- GSCA
  - Model specification, estimation, & evaluation
- Application
  - Item response theory in educational research
- Research topics

Intro to GSCA

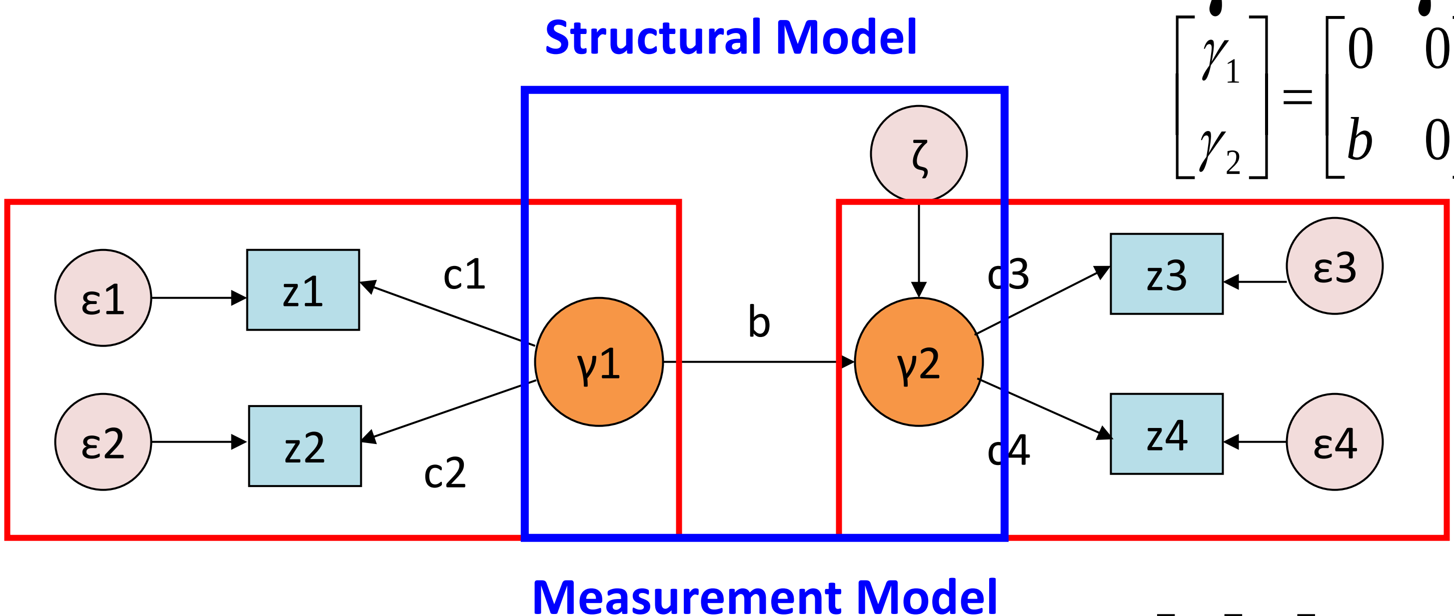
- Structural equation modeling (SEM)
  - SEM has been used for the analysis of interdependencies among observed variables and underlying constructs, often called latent variables

- Components in SEM (using LISREL Model)
  - Measurement model
  - Structural model

Components in SEM (LISREL Model)

$$\boldsymbol{\gamma} = \mathbf{B}\boldsymbol{\gamma} + \boldsymbol{\zeta}$$

$$\begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ b & 0 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}$$



$$\mathbf{z} = \mathbf{C}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} c_1 & 0 \\ c_2 & 0 \\ 0 & c_3 \\ 0 & c_4 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}$$

Intro to GSCA

- To approaches to SEM

SEM

Factor-based approach

Component-based approach

Latents ≈ Factors

- Covariance Structure Analysis (Jőreskog, 1970)

Latents ≈ Components

- PLS Path Modeling (Wold, 1982)
- **GSCA (Hwang & Takane, 2004)**

Intro to GSCA

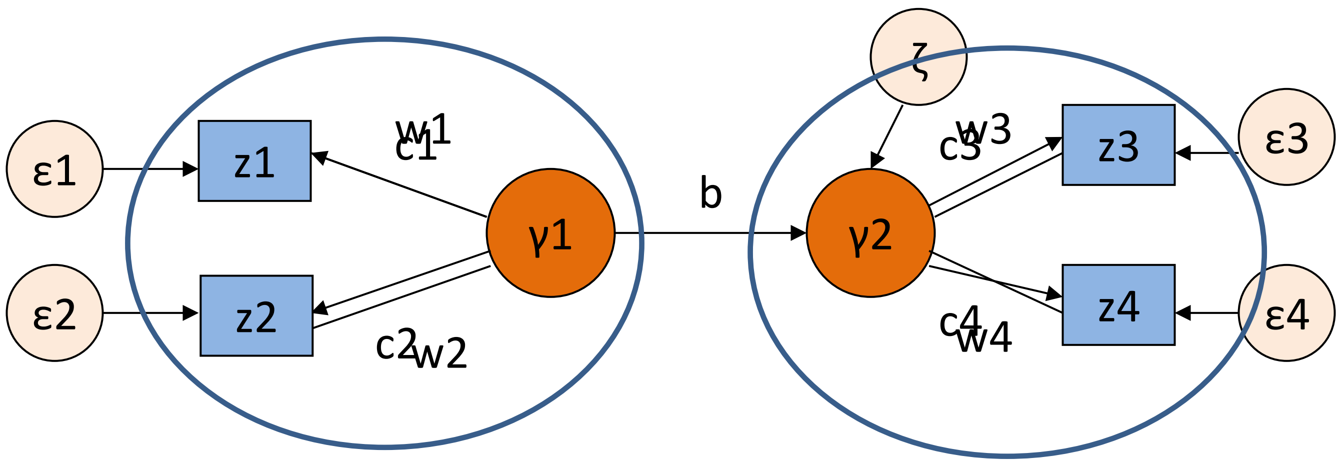
- Similarities and dissimilarities

		Factor-based	Component-based	
		CSA	PLS	GSCA
Model Specification	Latent variables	Factors (Random)	Components (Fixed)	
	Parameters	Loadings, Path coefficients, Error variances, Factor variances and/or means	Loadings, Path coefficients, Component weights	
Parameter Estimation	Input data	Covariance/ Correlation	Individual-level raw data	
	Estimation method	ML (mainly)	Least squares	
	Global optimization function	Yes	No	Yes

Intro to GSCA

- GSCA (Hwang & Takane, 2004)
  - Utilizes least square estimation method
  - Computes a composite component score using weights
- GSCA consists of three models
  - A measurement model: $z = C'\gamma + \epsilon$
  - A structural model: $\gamma = B'\gamma + \zeta$
  - A weighted relation model: $\gamma = W'z$

A Weighted Relation Model in GSCA



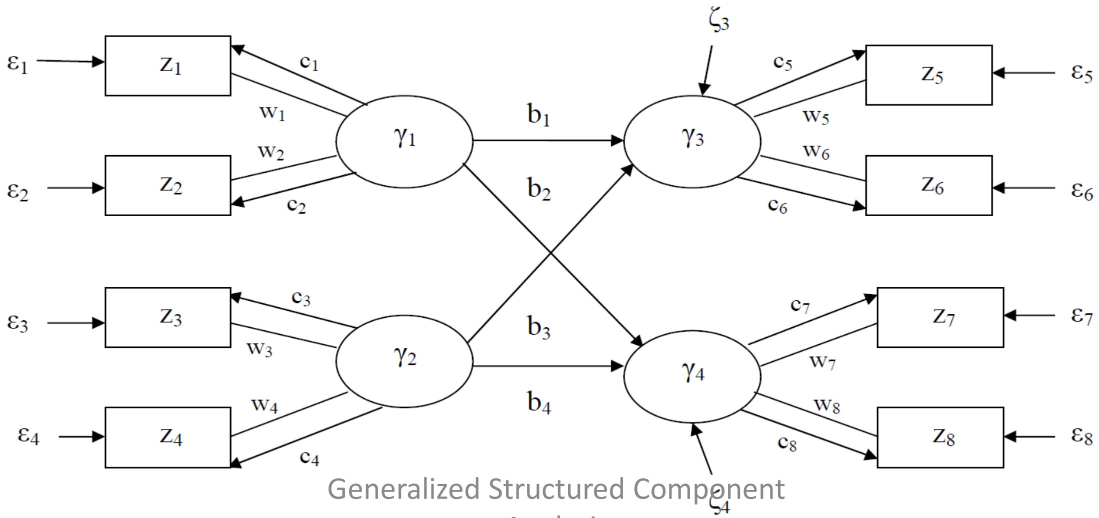
$\gamma_1 = z_1 w_1 + z_2 w_2$

$\gamma_2 = z_3 w_3 + z_4 w_4$

$$\boldsymbol{\gamma} = \mathbf{Wz} \qquad \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} w_1 & w_2 & 0 & 0 \\ 0 & 0 & w_3 & w_4 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix}$$

GSCA

- Model specification
  - GSCA consists of three models
    - A measurement model: $z = C'\gamma + \epsilon$
    - A structural model: $\gamma = B'\gamma + \zeta$
    - A weighted relation model: $\gamma = W'z$



GSCA

- Features in GSCA
  - No model identification problems and improper solutions
  - No rigid distributional assumptions
  - Stable parameter estimates even in small samples

GSCA

- Advantages over ML-based SEM & SEM with PLS
  - Avoid improper solutions by replacing factors with linear composites of observed variables (Same as in partial least squares (PLS; Wold, 1966, 1973, 1982))
  - Address the global optimization problem (Mulaik, 1972), which is an additional feature that PLS does not have

GSCA

- Parameter estimates
  - Least Square Criterion

$$\Phi = \sum_{j=1}^{J} \mathrm{SS}(E_j) = \sum_{j=1}^{J} \mathrm{tr}(E_j{}'E_j) \quad , \text{ where } E_j = [E_{Mj}, E_{Sj}]$$

  - Alternating Least Square Algorithm (de Leeuw, Young, & Takane, 1976)

  - Bootstrap method (Efron, 1982)

GSCA

- Model evaluation
  - Overall model fit measures (using variances)
    - $FIT = 1 - \frac{SS(ZV - ZWA)}{SS(ZV)} = \frac{1}{T}\sum_{t=1}^{T} R_t^2$ (Henseler, 2012)
      - Indicates the proportion of the total variance explained by a given particular model specification (Similar as the R-squared)
      - Can be used in model comparison with Bootstrapping standard errors or confidence intervals of the difference in FIT
    - $AFIT = 1 - (1 - FIT)\frac{d_0}{d_1}$ (Hwang et al., 2007)
      - Adjusted FIT(Similar as the adjusted R-squared)

GSCA

- Model evaluation
  - Overall model fit measures (using covariances)
    - $GFI = 1 - \frac{trace\left[\left(S - \hat{\Sigma}\right)^2\right]}{trace(S^2)}$ (Jöreskog & Sörbom, 1986)
      - Cut-off = higher than 0.9 (McDonal & Ho, 2002)
    - $SRMR = \sqrt{2 \sum_{j=1}^{J} \sum_{q=1}^{j} \frac{\left[\frac{s_{jq} - \hat{\sigma}_{jq}}{s_{jj} s_{qq}}\right]^2}{J(J+1)}}$ (Hwang, 2008)
      - Cut-off = less than 0.08 (Hu & Bentler, 1999)

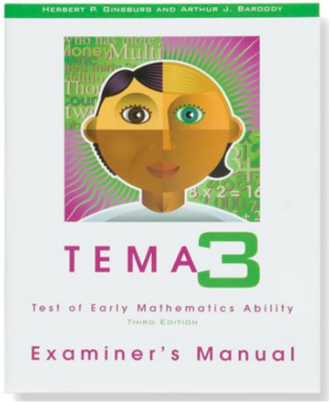
GSCA

- Model evaluation
  - Local model fit measures
    - $FIT_M = 1 - \frac{SS(Z - ZWC)}{SS(Z)}$
    - $FIT_S = 1 - \frac{SS(ZW - ZWB)}{SS(ZW)}$
  - Composite reliability (Werts et al., 1974)
    - $\rho_p = \frac{\left(\sum_{j=1}^{Jp} c_{pj}\right)^2}{\left(\sum_{j=1}^{Jp} c_{pj}\right)^2 + \sum_{j=1}^{Jp}(1 - c_{pj}^2)}$

GSCA

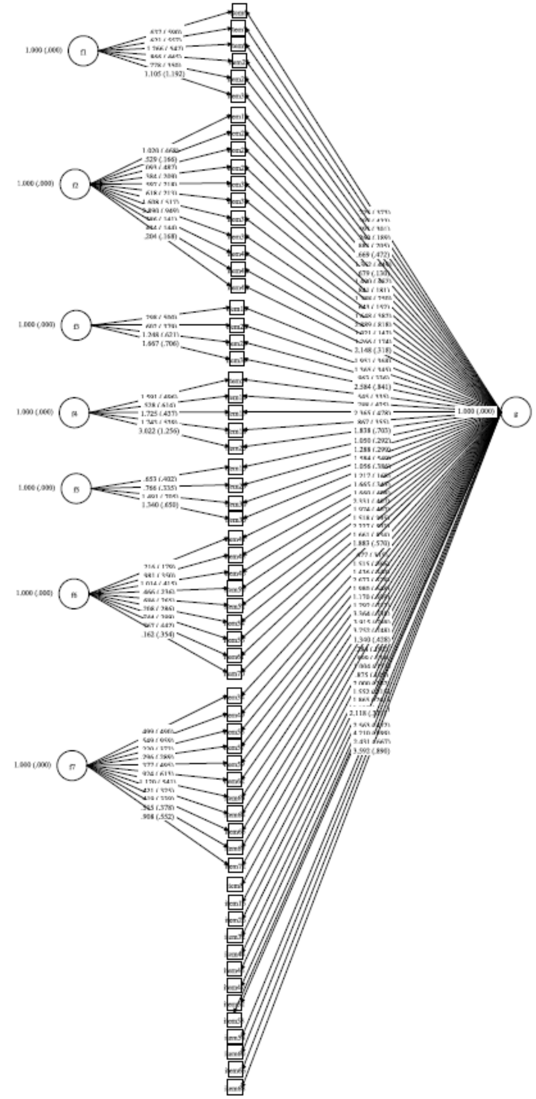
- Applicability of GSCA
  - Nonlinear GSCA (NL-GSCA; Hwang & Takane, 2010) for non-normal distribution in SEM
  - Fuzzy clusterwise GSCA for group-level heterogeneity such as mixture modeling, latent class/transition analysis, clustering, or classification in ML-based SEM (Hwang, DeSarbo, & Takane, 2007)
  - Longitudinal and time series data analysis (Jung et al., 2012)

Application



- Data
  - Math achievement data from Test of Early Mathematics Ability – 3 (TEMA-3)
    - Measure of math concepts, processes, and knowledge skills for children ages from 3 years to 8 years
  - Participants: 389 children from state-funded and/or Head Start pre-kindergarten classrooms
    - 182 boys (46.7%) at the beginning of data collection
    - Average age of 54.46 month (47 to 59 months, SD=3.47)

Bi-factor model for TEMA-3 (Ryoo, et al., 2015)

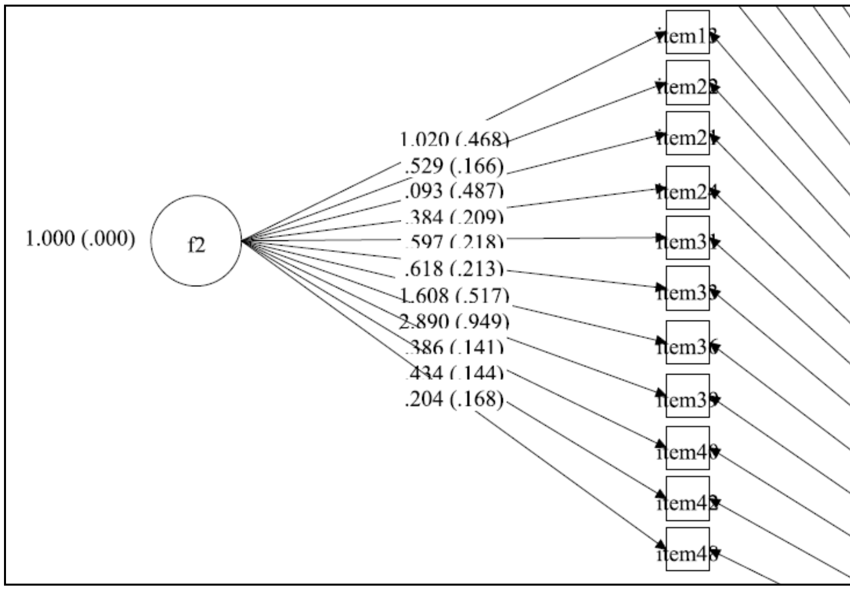


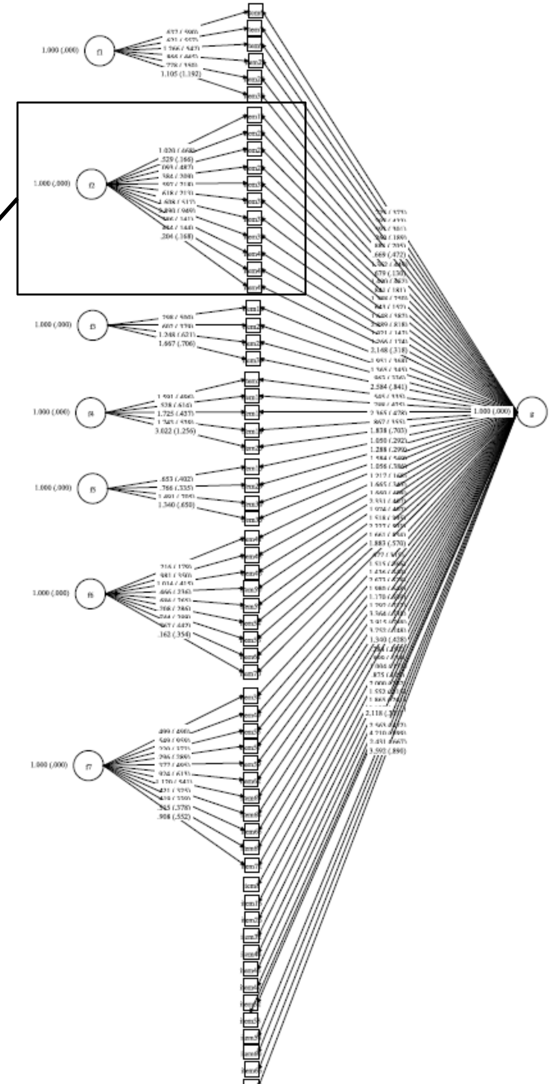
where ‘f1’ is representing Counting objects, ‘f2’ is Verbal counting, ‘f3’ is Numerical comparison, ‘f4’ is Set construction, ‘f5’ is Numeral literacy, ‘f6’ is Number facts, and ‘f7’ is Calculation.

Application

- Data
  - Longitudinal study for three years
    - Fall and Spring of Pre-K, Spring of K, and Spring of 1^{st} grade
  - Two sub-datasets were used in this study
    - Data1: Whole group of 294 at Spring of 1^{st} grade
    - Data2: Its subgroup of 50% randomly selected children (N=147)

Bi-factor model for TEMA-3 (Ryoo, et al., 2015)
- Verbal counting factor (f2) -





where ‘f1’ is representing Counting objects, ‘f2’ is Verbal counting, ‘f3’ is Numerical comparison, ‘f4’ is Set construction, ‘f5’ is Numeral literacy, ‘f6’ is Number facts, and ‘f7’ is calculation.

Application

- To item response theory (IRT)
  - The two parameter logistic (2PL) model

$$\pi_i = \frac{1}{1 + exp(-\beta_{1i}(z - \beta_{2i}))}$$

    - Difficulty parameter (β_{2i})
    - Discrimination parameter (β_{1i})
  - 2PL model also produces examinee's ability score (θ_j) where j denotes jth examinee

Application

- Ideas behind this application
  - Maximum likelihood estimate (MLE) used in IRT provides unbiased estimates when (1) sample data are large and (2) multivariate normality assumption are met
  - What if we have small sample data for ML-based IRT or if multivariate normality assumption are not met
    - Biased estimates and not efficient estimates

Application

- Alternatives
  - Small sample issue: non-parametric IRT
  - Multivariate normality issue: Bayesian IRT
  - But, we still have question about generalizing to big data (e.g., fMRI brain-imaging data (Jung, et al., 2012)
- GSCA accounts for both small sample and computer intensity for big data

Application

- GSCA accounts for both small sample and computer intensity for big data
  - Least square estimate (LS) provides unbiased point estimates regardless to distributional assumptions
    - Bootstrap estimates for interval estimates and/or hypothesis testing
  - LS estimation is efficient and computationally faster for both small and large samples

Application

- GSCA accounts for both small sample and computer intensity for big data
  - (Known) LS cannot be used for estimation for other distributional assumptions like binomial and multinomial – Really?
    - Not really (Hwang & Takane, 2010)

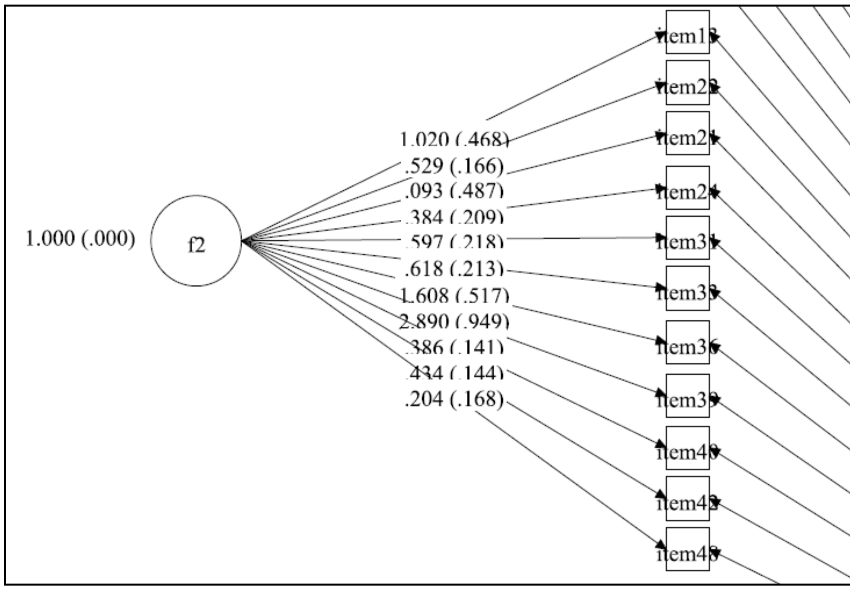
Application

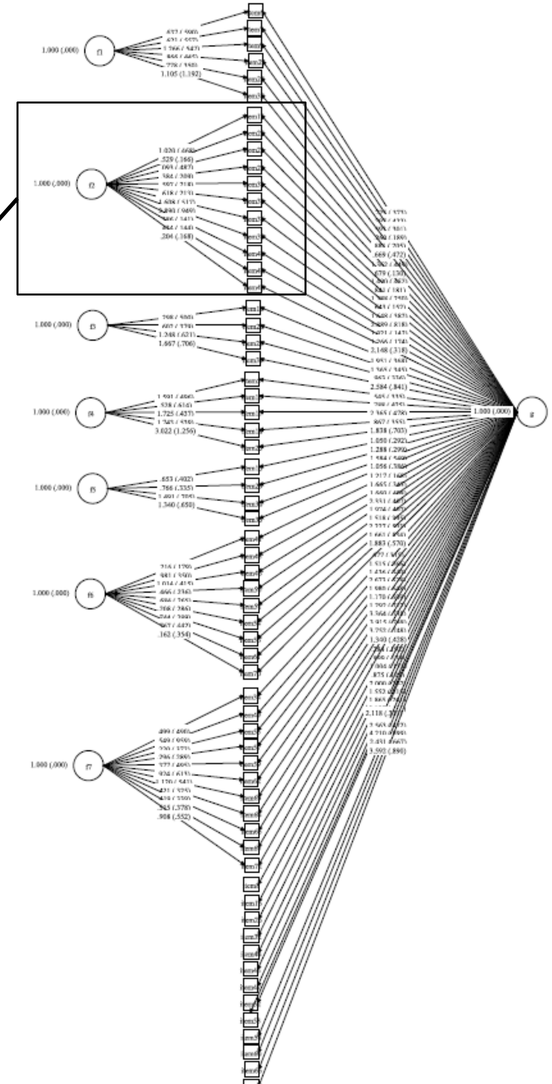
- Nonlinear GSCA (Hwang & Takane, 2010)
  - Applying GSCA to qualitative data such as nominal and categorical data
  - How? Resolve the linearity issue afflicting LS methods by applying the optimal scaling method (Kruskal, 1964a,b; McDonald, 2000; Young, 1981)

Application

- Component-based IRT (CB-IRT)
  - Application of nonlinear GSCA to IRT
  - Estimation procedure
    - Phase one
      - Updating model parameters including loadings and weights
    - Phase two (optimal scaling phase)
      - Step 1: Updating the model prediction $\hat{s}_j$ corresponding to s_j for fixed parameters from Phase one
      - Step 2: Obtaining the optimally transformed data s_j such that it is as close to $\hat{s}_j$ as possible in the LS sense

Bi-factor model for TEMA-3 (Ryoo, et al., 2015)
- Verbal counting factor (f2) -

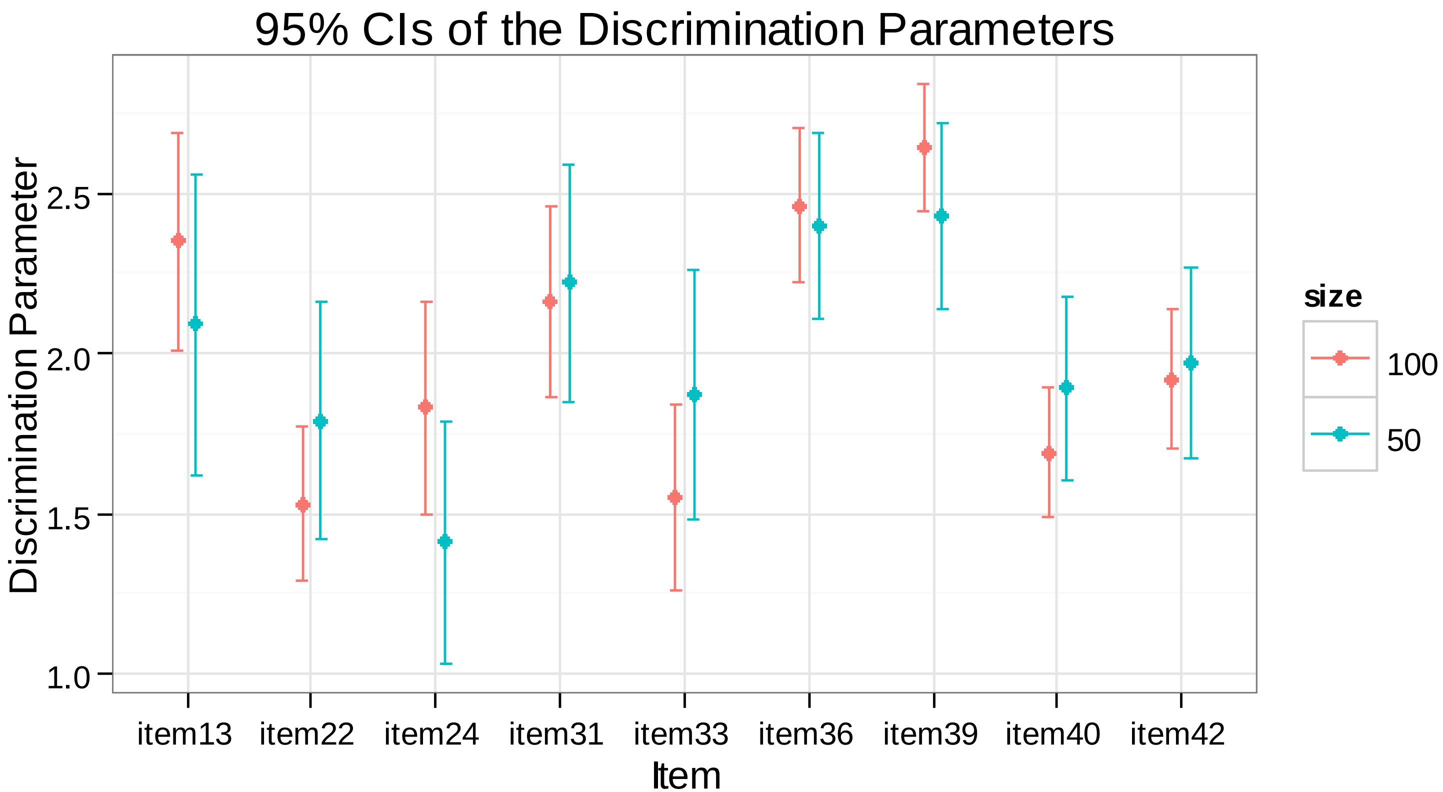




where 'f1' is representing Counting objects, 'f2' is Verbal counting, 'f3' is Numerical comparison, 'f4' is Set construction, 'f5' is Numeral literacy, 'f6' is Number facts, and 'f7' is calculation.

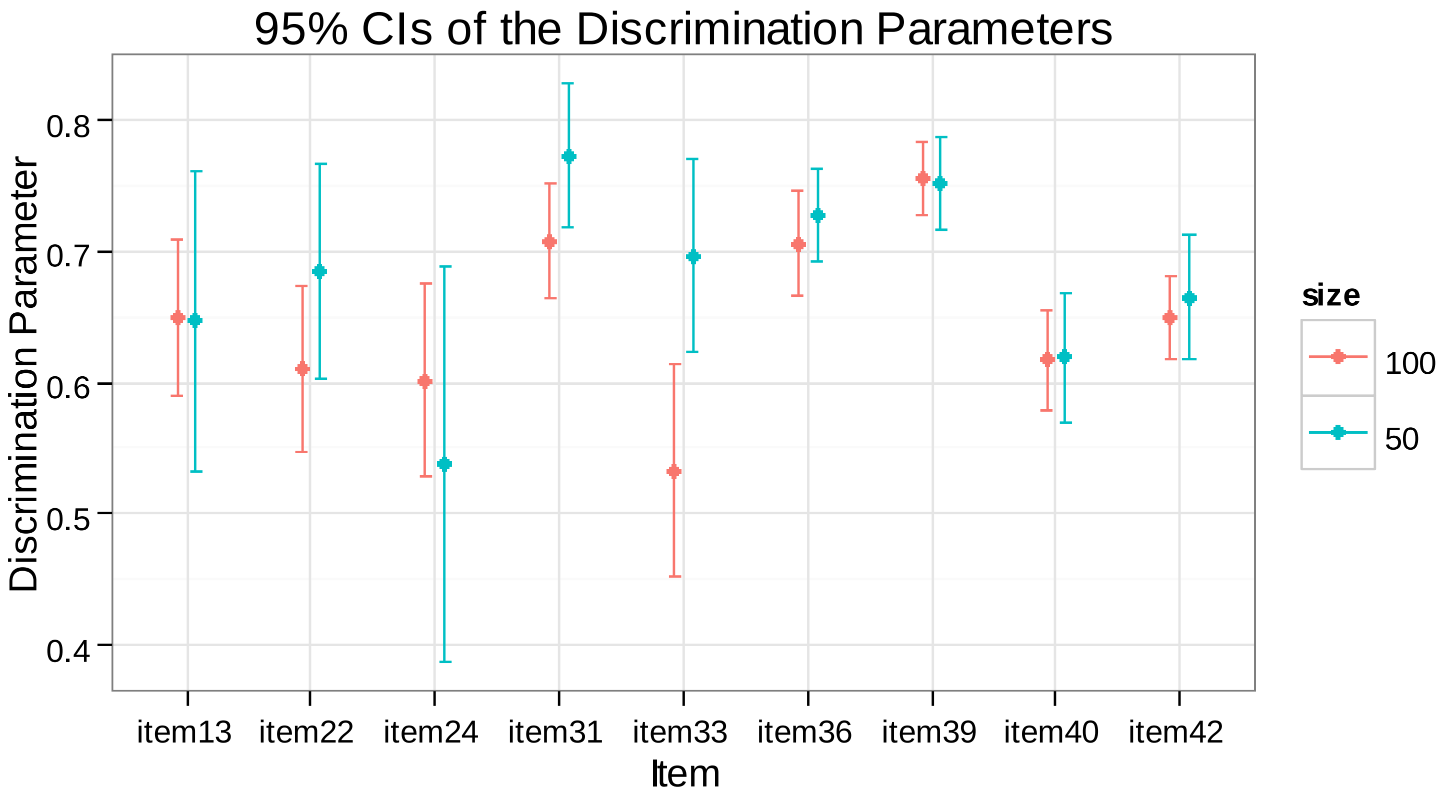
Application

- Result of discrimination in 2PL model



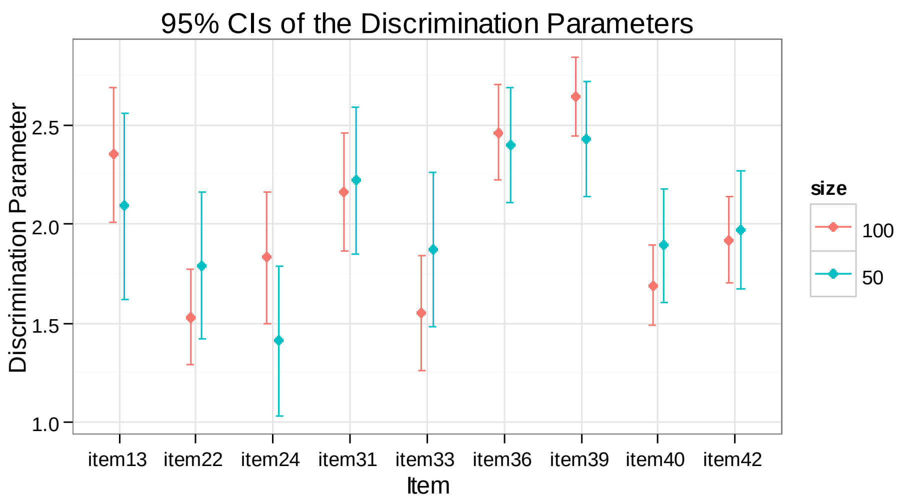
Application

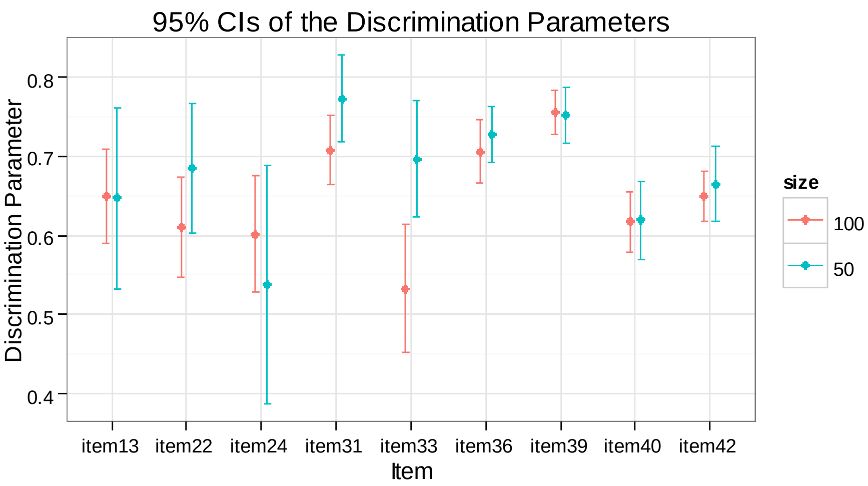
- Result of discrimination in CB-IRT model



Application

- Results
  - Estimates in CB-IRT over different sample sizes are relatively closer (Right)
  - 2PL provides relatively more consistent SEs (Left)





Application

- Ongoing research
  - Interpretable composite scores comparable to ability in ML-based IRT
  - Proper model comparison tools that can be used in differential item functioning, equating, and linking in IRT literature

Research topics in GSCA

- Model evaluation
  - Confirmatory Tetrad Analysis (CTA) for model comparison (Bollen & Ting, 1993)
- Application GSCA to longitudinal/multilevel data analysis
  - Multilevel latent class/transition analysis
  - Dynamic SEM

Reference

Bollen, K. A., & Ting, K-F. (1993). Confirmatory Tetrad Analysis. *Sociological Methodology, 23*, 147-175.

Hwang, H., DeSarbo, S. W., & Takane, Y. (2007). Fuzzy clusterwise generalized structured component analysis. *Psychometrika, 72*, 181-198.

Hwang, H. & Takane, Y. (2004). Generalized structured component analysis. *Psychometrika, 69*, 81-99.

Hwang, H. & Takane, Y. (2010). Nonlinear generalized structured component analysis. *Behaviormetrika, 34*, 95-109.

Jung, K. Takane, Y., Hwang, H., & Woodward, T. S. (2012). Dynamics generalized structured component analysis with applications to the analysis of effective connectivity in functional neuroimaging data. *Psychometrika, 77*, 827-848.

McDonald, R. P. (2000). *Test theory: A unified treatment*. Mahwah, New Jersey: Lawrence Erlbaum Associates, INC.

Mulaik, S. A. (1972). *The foundations of factor analysis*. New York: McGraw-Hill Book Company.

Rigdon, S. E. & Tsutakawa, R. K. (1983). Parameter estimation in latent trait models. *Psychometrika, 48*, 567-574.

Ryoo, J. H., Molfese, V., Brown, E. t., Karp, K. S., Welch, G., & Bovaird, J. A. (2015). Examining factor structures on the Test of Early Mathematics ability – 3: A longitudinal approach. Learning and Individual Differences, 41, 21-29.

Wold, H. (1966). Estimation of principal components and related methods by iterative least squares. In P. R. Krishnaiah (Ed.), *Multivariate Analysis* (pp. 391-420). New York: Academic Press.

Wold, H. (1973). Nonlinear iterative partial least squares (NIPALS) modeling: some current developments. In P. R. Krishnaiah (Ed.), *Multivariate Analysis* (pp. 383-487). New York: Academic Press.

Wold, H. (1982). Soft modeling: the basic design and some extensions. In K. G. Jöreskog & H. Wold (Eds.), *Systems under Indirect Observations II* (pp. 1-54). Amsterdam: North Holland.

Young, F. W. (1981). Quantitative analysis of qualitative data. *Psychometrika, 46*, 347-388.