A Review of Nonlinear Factor Analysis Methods and Applications

Melanie M. Wall, Division of Biostatistics, University of Minnesota
Yasuo Amemiya, IBM Thomas J. Watson Research Center

Disclaimer

We are statisticians.

What is nonlinear factor analysis?

- What is factor analysis?

- What is nonlinear?

What is nonlinear factor analysis?

What is factor analysis?

factor analysis expresses p observed variables $\mathbf{Z}$ in terms of q unobserved latent variables $\mathbf{f}$.

	latent continuous	latent discrete
observed continuous		
observed discrete		

What is nonlinearity?

discrete variables nonlinearity
relationship nonlinearity

SCOPE of this talk

	latent continuous	latent discrete
observed continuous		
observed discrete		

- continuous observed variables, continuous latent variables "factors"
- limited emphasis on exploratory/confirmatory distinction
- inclusion of structural equation models

OUTLINE

- A Google search

- A taxonomy of nonlinear factor analysis

- History, Now, and Future

  - The "difficulty factor"

  - Development of the nonlinear factor analysis - nonlinear in factors but linear in parameters

  - Development of nonlinear structural equation model - nonlinear in factors but linear in parameters

  - Development of the **general** nonlinear factor analysis model

  - Last and next 100 years?

What is nonlinear factor analysis?

Google search finds the term "nonlinear factor analysis", for example

- In economics

  A Nonlinear Factor Analysis of S&P 500 Index Option Returns

  C. Jones 2001 *JEL*

- In signal processing: blind source separation - independent component analysis

  A generic framework for blind source separation in structured nonlinear models,

  A. Taleb *IEEE on Signal Processing*, 2002

- In pattern recognition: speech recognition - neural networks

  Bayesian nonlinear state space modeling (Honkela,2000)

A taxonomy of nonlinear factor analysis models

Linear factor analysis model

$$\mathbf{Z}_i = \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{f}_i + \boldsymbol{\epsilon}_i$$

$\mathbf{Z}_i$ $p \times 1$ observable variables for individuals $i = 1 \ldots n$

$\mathbf{f}_i$ $q \times 1$ "latent" factors for each individual

$\boldsymbol{\epsilon}_i$ $p \times 1$ measurement error

A taxonomy of nonlinear factor analysis

Nonlinear factor analysis model

$$\mathbf{Z}_i = \mathbf{G}(\mathbf{f}_i) + \boldsymbol{\epsilon}_i$$

- Nonlinear in the factors, but linear in the parameters:
  $\mathbf{G}(\mathbf{f}_i) = \boldsymbol{\Lambda}\mathbf{g}(\mathbf{f}_i)$ where $\mathbf{g}(\mathbf{f}_i)$ is r dimensional
  e.g. $G_j(\mathbf{f}_i) = \lambda_1 + \lambda_2 f_i + \lambda_3 f_i^2 + \lambda_4 f_i^3$

- Nonlinear in the factors and nonlinear in the parameters:
  e.g. $G_j(\mathbf{f}_i) = \lambda_1 + \frac{\lambda_2}{1+e^{\lambda_3 - \lambda_4 f_i}}$

- No particular parametric functional form $\mathbf{G}(\mathbf{f}_i)$:
  principal curves, blind source separation, neural networks, semiparametric dynamic factor analysis.

This general formulation includes nonlinear structural equation model -
But care must be taken!

Dawn of nonlinear factor analysis
The Difficulty Factor

Referencing several papers with Ferguson (1941) as the earliest, Gibson (1960) writes of the following dilemma for factor analysis:

"When a group of tests quite homogeneous as to content but varying widely in difficulty is subjected to factor analysis, the result is that more factors than content would demand are required to reduce the residuals to a random pattern"

These additional factors are called the "difficulty factors"

Dawn of nonlinear factor analysis
Gibson's work

1951 Finishes his dissertation at the University of Chicago. "Applications of the mathematics of multiple-factor analysis to problems of latent structure analysis."

1955 Presentation - 63^{rd} Annual Convention of the APA, San Fran., CA "Nonlinear factor analysis, single factor case".

1956 Presentation - 64^{th} Annual Convention of the APA, Chicago, IL "Nonlinear factors in two dimensions"

1959 Gibson, W.A. Three multivariate models: Factor analysis, latent structure analysis, and latent profile analysis. *Pscychometrika*, 24(3), 229-252.

1960 Gibson, W.A. (1960). Nonlinear factors in two dimensions. *Psychometrika*, 25(4), 381-392.

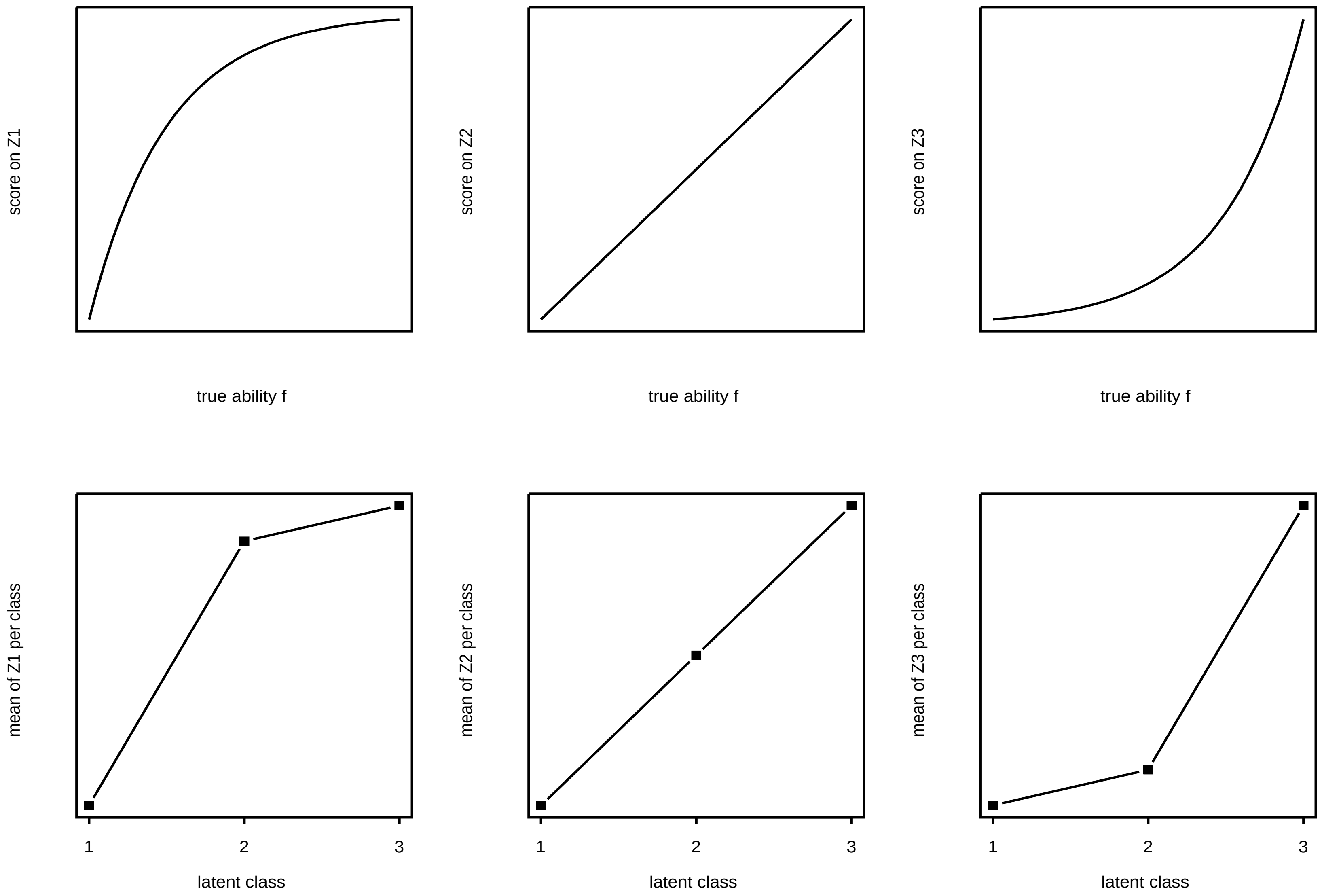
Dawn of nonlinear factor analysis

The difficulty factor and Gibson (1960)

"Coefficients of linear correlation, when applied to such data, will naturally underestimate the degree of nonlinear functional relation that exists between such tests. Implicitly, then it is nonlinear relations among tests that lead to difficulty factors. Explicitly, however, the factor model rules out, in its fundamental linear postulate, only such curvilinear relations as may exist between tests and factors."

Gibson goes on to use discrete latent factors, i.e. latent classes via profile analysis (Lazarsfeld, 1950), to avoid "inherent problem of linear factor analysis"

Gibson's method - Latent profile analysis



First parametric functional form for nonlinear factor analysis

1962 McDonald, R.P., A general approach to nonlinear factor analysis. *Psychometrika*

Unsatisfied with Gibson's "ad hoc" nature of discretizing the continuous underlying factor in order to fit a nonlinear relation between it and the observed variables, McDonald (1962) developed a nonlinear functional relationship between the underlying continuous factors and the continuous observed variables.

Given a vector $\mathbf{Z}_i$ of p observed variables for individuals $i = 1 \ldots n$

$$\begin{aligned} \mathbf{Z}_i &= \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{g}(\mathbf{f}_i) + \boldsymbol{\epsilon}_i \\ \text{where } \mathbf{g}(\mathbf{f}_i) &= (g_1(\mathbf{f}_i), g_2(\mathbf{f}_i), \ldots g_r(\mathbf{f}_i))' \\ \text{and } \mathbf{f}_i &= (f_{1i} \ldots f_{qi})' \end{aligned}$$

NOTE: Linear in the parameters, nonlinear in the factors $\mathbf{f}_i$.

Basic idea of McDonald (1962)

Fit a **linear** factor analysis with r underlying factors and then look to see if there is a **nonlinear** functional relationship between the fitted factor scores

"We obtain factor scores on the two factors and plot them, one against the other. If there is a significant curvilinear relation (in the form of a parabola) we can estimate a quadratic function relating one set of factor scores to the other; this relation can then be put into the specification equation to yield the required quadratic functions for the individual manifest variates."

McDonald's work on nonlinear factor analysis

1967 McDonald, R.P., Numerical methods for polynomial models in nonlinear factor analysis. *Psychometrika*

- Develops programs COPE (corrected polynomials etc.) and POSER for (polynomial series program)

1967 McDonald, R.P., Factor interaction in nonlinear factor analysis. *British Journal of Mathematical and Statistical Psychology*.

- Develops program FAINT (factor analysis interactions)

- He gives credit to Bartlett (1953)...

  "In a discussion of the linearity hypothesis in factor analysis, Bartlett (1953) noted that if we include the product of two existing factors as an additional term in the specification equation of linear factor analysis, it will act just like a further common factor as far as the correlation properties of the observed variables are concerned."

McDonald's work on nonlinear factor analysis

1983 Etezadi-Amoli, J., and McDonald, R.P. A second generation nonlinear factor analysis. *Psychometrika*

- Same model, different estimation...

- Uses the estimation method proposed in McDonald (1979) which treats the underlying factors as fixed and minimizes

$$
\begin{aligned}
l &= \frac{1}{2}(\log|diag\mathbf{Q}| - \log|\mathbf{Q}|) \\
\mathbf{Q} &= \frac{1}{n}\sum_{i=1}^{n}(\mathbf{Z}_i - \boldsymbol{\Lambda}\mathbf{g}(\mathbf{f}_i))(\mathbf{Z}_i - \boldsymbol{\Lambda}\mathbf{g}(\mathbf{f}_i))'
\end{aligned}
$$

  a likelihood-ratio discrepancy function.

- Now the dimension of $\mathbf{g}()$ is not restricted to be less than or equal to the number of linear factors that could be fit.

Nonlinear factor analysis - normal factors

1986 Mooijaart, A. and Bentler, P. "Random polynomial factor analysis" In *Data Analysis and Informatics*, IV (E. Diday et al., Eds.)

$$\begin{aligned} \mathbf{Z}_i &= \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{g}(f_i) + \boldsymbol{\epsilon}_i \\ \text{where } \mathbf{g}(f_i) &= (f_i, f_i^2, \ldots, f_i^r)' \end{aligned}$$

Acknowledging McDonald's work, Mooijaart, A. and Bentler (1986) estimate the model assuming the underlying factor is normally distributed rather than fixed.

Assume $f_i \sim N(0, 1)$ is normally distributed.

They point out that if $r > 1$, then $\mathbf{Z}_i$ is necessarily NON-normal

Mooijaart and Bentler (1986)

- Because $\mathbf{Z}_i$ is not normally distributed, use ADF method which has been recently developed (Brown, 1984 *BJMSP*).

- Minimize
$$\frac{1}{n}(s-\sigma(\boldsymbol{\theta}))'\mathbf{W}(s-\sigma(\boldsymbol{\theta}))$$
w.r.t $\boldsymbol{\theta}$ containing $\boldsymbol{\Lambda}$ and $\boldsymbol{\Psi} = diag(var(\epsilon_{1i}), \ldots, var(\epsilon_{pi}))$, where s is a vector of the sample second and third order cross-products and $\sigma(\boldsymbol{\theta})$ the population expectation of the cross-products.

- Assuming $f_i \sim N(0,1)$ possible to explicitly calculate the moments $E(\mathbf{g}(f_i)\mathbf{g}(f_i)')$ and $E\{(\mathbf{g}(f_i) \otimes \mathbf{g}(f_i))\mathbf{g}(f_i)'\}$
for $\mathbf{g}(f_i) = (f_i, f_i^2, \ldots, f_i^r)'$.
Treat as fixed and known.

- Provided test for overall fit and considered standard errors

Application to attitude scale
Mooijaart and Bentler (1986)

1986 Mooijaart, A. and Bentler, P. "Random polynomial factor analysis"

"Not only in analyzing test scores a nonlinear model may be important, also in analyzing attitudes this model may be the proper model. Besides finding an attitude factor an intensity-factor is also often found with linear factor analysis. Again, the reason here is that subjects on both ends of the scale (factor) are more extreme in their opinion that can be expected from a linear model."

Example

In this example attitudes about nuclear weapons were analyzed. The sample consisted of students of the University of California, Los Angeles. Data were collected in 1983. The data we analyze here are a part of a survey: UCLA longitudinal study of growth. This project was sponsored by U.S. Department of Health and Human Services. The wording of the questions is given in Table 1. Also in this Table are given the mean, standard deviation and skewness of the variables. Possible answers to the questions are "strongly disagree", "disagree", "don't know", "agree" and "strongly agree", which were coded for our purpose from 1 up to 5.

Table 1

Questions about nuclear weapons
and some sample statistics (N=405)

Wording		mean	stand.dev.	skewness
1.	I am quite concerned about how many countries have nuclear weapons.	3.90	.92	-.93
2.	It is essential for our protection that the United States produce as many nuclear weapons as possible.	2.57	1.13	-.30
3.	I feel frightened when I think of all the nuclear weapons in the world.	3.79	1.11	-.90
4.	I imagine I would survive a nuclear war.	2.26	1.00	-.27
5.	I have never really worried about nuclear war.	2.63	1.24	-.40
6.	Many people tend to overreact about the threat of a nuclear war.	2.78	1.14	-.13
7.	There are times I have felt depressed thinking about the possibility of nuclear war.	2.92	1.27	.00
8.	The world feels like a very dangerous place because of so many nuclear weapons.	3.26	1.12	-.37

Application to attitude scale
Mooijaart and Bentler (1986)

1 and 2 linear factors do not fit well, 3 linear factors do fit well.

Third degree polynomial factor model did fit well (with one factor).

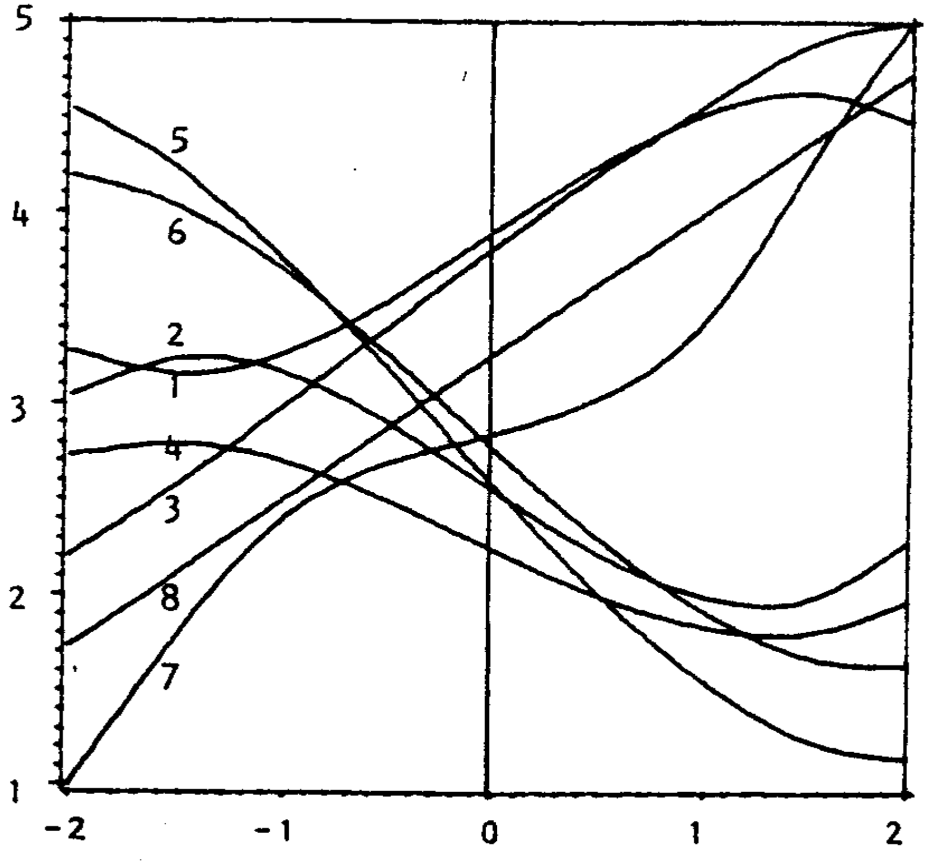


Figure 1. Solution 1.

A change of focus

1970's and 1980's - Introduction and development of LISREL and SEM

Increased focus on modeling and estimating the relationship *between* variables in which some of the variables may be measured with error.

Responding to calls (e.g. Busymeyer and Jones, 1983, *Psych Bulletin*) for methods that could handle errors of measurement in structural models with *nonlinear* relations...

Kenny and Judd (1984), *Psych Bulletin* introduced a method for the quadratic and cross product (interaction) model.

Kenny and Judd (1984)

Dealt with two specific models - interaction and quadratic

$$
\begin{aligned}
Z_{1i} &= \lambda_{11} f_{1i} + \lambda_{12} f_{2i} + \lambda_{13} f_{1i} f_{2i} + \epsilon_{1i} \\
Z_{2i} &= \lambda_{21} f_{1i} + \epsilon_{2i} \\
Z_{3i} &= f_{1i} + \epsilon_{3i} \\
Z_{4i} &= \lambda_{42} f_{2i} + \epsilon_{4i} \\
Z_{5i} &= f_{2i} + \epsilon_{5i}
\end{aligned}
$$

$$
\begin{aligned}
Z_{1i} &= \lambda_{11} f_{1i} + \lambda_{12} f_{1i}^2 + \epsilon_{1i} \\
Z_{2i} &= \lambda_{21} f_{1i} + \epsilon_{2i} \\
Z_{3i} &= f_{1i} + \epsilon_{3i}
\end{aligned}
$$

Note these are just special cases of $\mathbf{Z}_i = \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{g}(\mathbf{f}_i) + \boldsymbol{\epsilon}_i$ with some elements of $\boldsymbol{\Lambda}$ fixed to 1 or 0

No reference given to any previous nonlinear methodological work.

Basic idea of Kenny and Judd (1984)

Create new "observed variables" by taking products of existing variables and use these as additional indicators of nonlinear terms

$$\begin{pmatrix} Z_{1i} \\ Z_{2i} \\ Z_{3i} \\ Z_{4i} \\ Z_{5i} \\ Z_{2i}Z_{4i} \\ Z_{2i}Z_{5i} \\ Z_{3i}Z_{4i} \\ Z_{3i}Z_{5i} \end{pmatrix} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \lambda_{42} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda_{21}\lambda_{42} \\ 0 & 0 & \lambda_{21} \\ 0 & 0 & \lambda_{42} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f_{1i} \\ f_{2i} \\ f_{1i}f_{2i} \end{pmatrix} + \begin{pmatrix} \epsilon_{1i} \\ \epsilon_{2i} \\ \epsilon_{3i} \\ \epsilon_{4i} \\ \epsilon_{5i} \\ u_{1i} \\ u_{2i} \\ u_{3i} \\ u_{4i} \end{pmatrix},$$

where

$$\begin{aligned} u_{1i} &= \lambda_{21}f_{1i}\epsilon_{4i} + \lambda_{42}f_{2i}\epsilon_{2i} + \epsilon_{2i}\epsilon_{4i} \\ u_{2i} &= \lambda_{21}f_{1i}\epsilon_{5i} + f_{2i}\epsilon_{2i} + \epsilon_{2i}\epsilon_{5i} \\ u_{3i} &= f_{1i}\epsilon_{4i} + \lambda_{42}f_{2i}\epsilon_{3i} + \epsilon_{3i}\epsilon_{4i} \\ u_{4i} &= f_{1i}\epsilon_{5i} + f_{2i}\epsilon_{3i} + \epsilon_{3i}\epsilon_{5i} \end{aligned}$$

Basic idea of Kenny and Judd (1984)

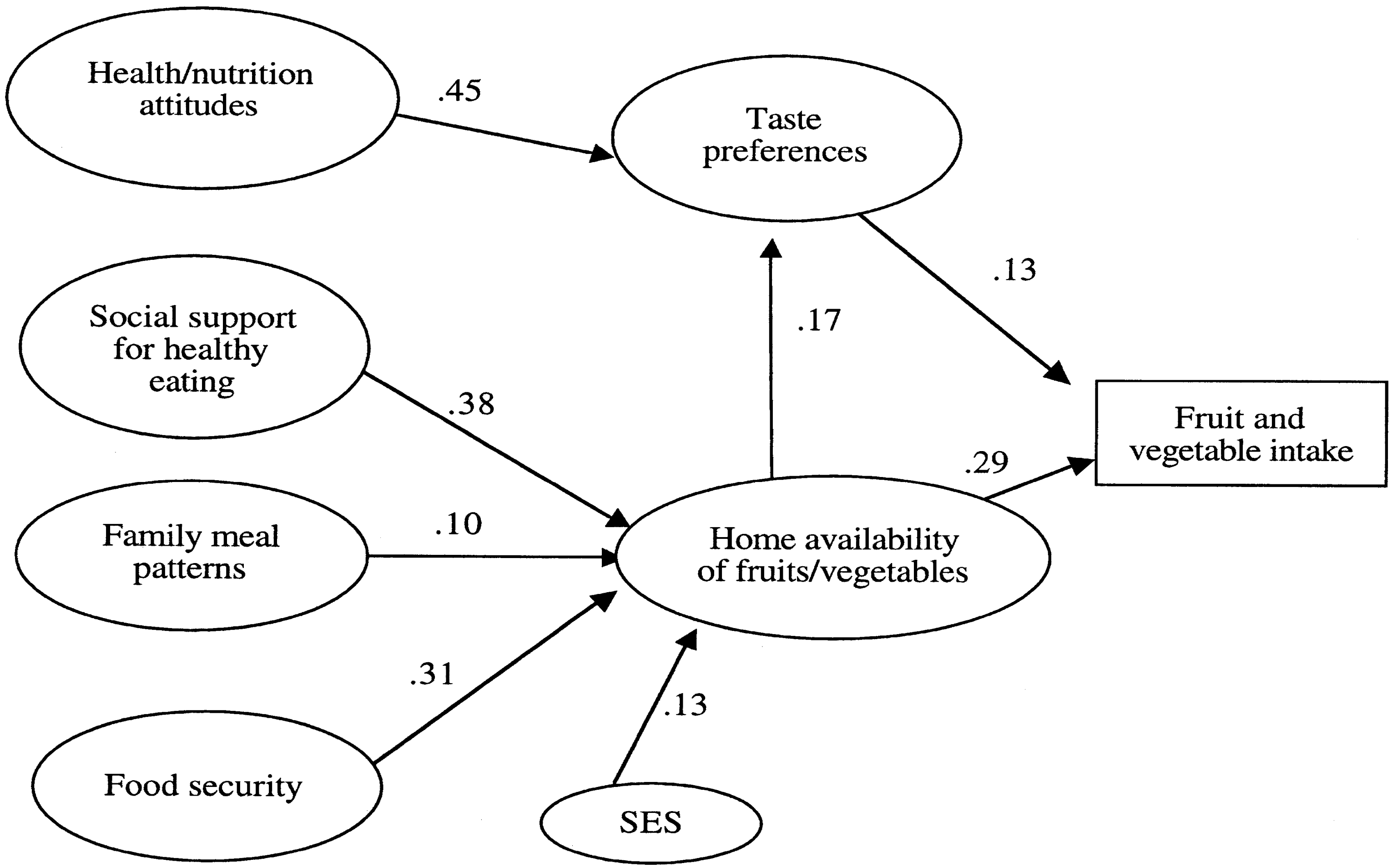
Assuming $\mathbf{f}_i$ and $\boldsymbol{\epsilon}_i$ are normally distributed, construct the covariance matrix of $(f_{1i}, f_{2i}, f_{1i}f_{2i})'$ and $(\boldsymbol{\epsilon}', \mathbf{u}')'$.

This results in many (tedious) constraints on the model covariance matrix...

BUT they can be included as model restrictions and fit in existing *linear* software programs (e.g. LISREL)

An example to emphasize the different focus

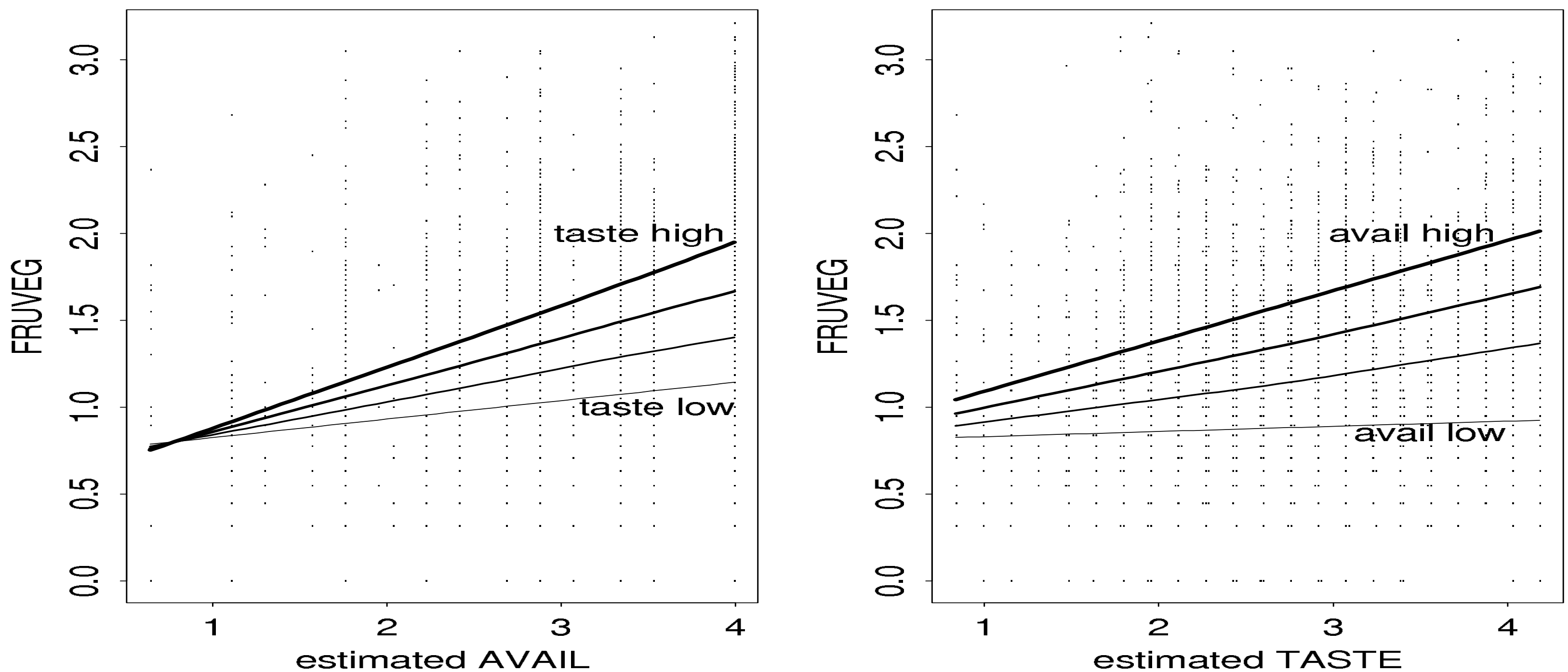
Project EAT - Neumark-Sztainer, et al. *Preventive Medicine* (2003)



Project EAT example

$$\begin{pmatrix} Z_{1i} \\ Z_{2i} \\ Z_{3i} \\ Z_{4i} \\ Z_{5i} \end{pmatrix} = \begin{pmatrix} \beta_{01} \\ \beta_{02} \\ \beta_{03} \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \\ \beta_{31} & \beta_{32} \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} AVAIL_i \\ TASTE_i \end{pmatrix} + \begin{pmatrix} \epsilon_{1i} \\ \epsilon_{2i} \\ \epsilon_{3i} \\ \epsilon_{4i} \\ \epsilon_{5i} \end{pmatrix}$$

$$FRVEG_i = \alpha_0 + \alpha_1 AVAIL_i + \alpha_2 TASTE_i + \alpha_3 AVAIL_i * TASTE_i + \zeta_i.$$



Desirable to properly test $\alpha_3 = 0$

A change of focus and an explosion in the methodological literature

The Kenny and Judd (1984) paper proved to be the spark for a flurry of methodological papers introducing new estimation methods for the "nonlinear structural equation model"

Methodological papers focused on "nonlinear structural equation models"

- Product indicator methods
- Two stage least squares
- Direct Maximum likelihood
- Bayesian methods
- Method of moments

Product indicator methods

Hayduck, LA (1987). *Structural equation modeling with LISREL: Essentials and advances*.

Ping, R.A. (1995). A parsimonious estimating technique for interaction and quadratic latent variables. *Journal of Marketing Research*

Ping, R.A. (1996). Latent variable interaction and quadratic effect estimation: A two-step technique using structural equation analysis. *Psychological Bulletin*

Ping, RA. (1996) Latent variable regression: A technique for estimating interaction and quadratic coefficients. *Multivariate Behavioral Research*

Ping, RA. (1996) Estimating latent variable interactions and quadratics: The state of this art. *Journal of Management*.

Jaccard, J., and Wan, C.K. (1995). Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: Multiple indicator and structural equation approaches. *Psychological Bulletin*.

Jaccard, J., and Wan, C.K. (1996). *LISREL approaches to interaction effects in multiple regression*, Sage.

Jöreskog, K.G., and Yang, F. (1996). Non-linear structural equation models: The Kenny-Judd model with interaction effects. In G.A. Marcoulides and R.E. Schumacker (Eds.) *Advanced Structural Equation Modeling: Issues and Techniques.*

Jöreskog, K.G., and Yang, F. (1997). Estimation of interaction models using the augmented moment matrix: Comparison of asymptotic standard errors. In W. Bandilla and F. Faulbaum (Eds.) *SoftStat '97 Advances in Statistical Software 6*

Product indicator methods (continued)

- Li, F., Harmer, P., Duncan, T., Duncan, S., Acock, A., and Boles, S. (1998). Approaches to testing interaction effects using structural equation modeling methodology. *Multivariate Behavioral Research*.

  Compared Jöreskog, K.G., and Yang, F. (1996), Jaccard and Wan (1995), and Ping (1996).

- Wall, M.M. and Amemiya, Y., (2001) Generalized appended product indicator procedure for nonlinear structural equation analysis. *J EDUC BEHAV STAT*

  drops the distributional assumption of normality for the underlying factors

Two stage least squares

Bollen, K.A. (1995) Structural equation models that are nonlinear in latent variables: A least squares estimator. *Sociological Methodology*.

Bollen, K.A. (1996) An alternative two stage least squares (2SLS) estimator for latent variable equation. *Psychometrika*.

Bollen, K.A. and Paxton, P. (1998), Interactions of latent variables in structural equation models *Structural equation modeling*.

Maximum likelihood

Klein, A., Moosbrugger, H., Schermelleh-Engel, K., Frank, D. (1997). A new approach to the estimation of latent interaction effects in structural equation models. In W. Bandilla and F. Faulbaum (Eds.) *SoftStat '97 Advances in Statistical Software 6*.

Klein, A., Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*.

Lee SY, Zhu HT. (2002) Maximum likelihood estimation of nonlinear structural equation models *PSYCHOMETRIKA*.

Lee SY, Song XY, Lee JCK. (2003) Maximum likelihood estimation of nonlinear structural equation models with ignorable missing data *J EDUC BEHAV STAT*.

Lee SY, Song XY. (2003) Maximum likelihood estimation and model comparison of nonlinear structural equation models with continuous and polytomous variables" *COMPUT STAT DATA AN*.

Bayesian modelling

Wittenberg, J., and Arminger, G. (1997). Bayesian non-linear latent variable models-specification and estimation with the program system BALAM. In W. Bandilla and F. Faulbaum (Eds.) *SoftStat '97 Advances in Statistical Software 6*.

Arminger, G and Muthén, B. (1998) A Bayesian approach to nonlinear latent varaible models using the Gibbs sampler and the Metropolis-Hastings algorithm. *Psychometrika*, 63(3), 271-300. Reference to Hayduck, YandJ96 and KJ.

Zhu HT, Lee SY. (1999) Statistical analysis of nonlinear factor analysis models. *BRIT J MATH STAT PSY*.

Lee SY, Zhu HT. (2000) Statistical analysis of nonlinear structural equation models with continuous and polytomous data. *BRIT J MATH STAT PSY*.

Song XY, Lee SY. (2002) A Bayesian approach for multigroup nonlinear factor analysis" *STRUCT EQU MODELING*.

Method of moments

Wall, M.M. and Amemiya, Y., (2000) Estimation for polynomial structural equation models. *JASA*.

Wall, M.M. and Amemiya, Y., (2003) A method of moments technique for fitting interaction effects in structural equation models. *BRIT J MATH STAT PSY*

Nonlinear structural equation model nonlinear in factors but linear in parameters

Partitioning $\mathbf{f}_i$ into *endogenous* and *exogenous* variables, i.e. $\mathbf{f}_i = (\boldsymbol{\eta}_i, \boldsymbol{\xi}_i)'$

$$\begin{aligned} \mathbf{Z}_i &= \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{f}_i + \boldsymbol{\epsilon}_i \\ \boldsymbol{\eta}_i &= \boldsymbol{\Gamma}\mathbf{g}(\boldsymbol{\xi}_i) + \boldsymbol{\delta}_i \end{aligned}$$

Note: This model is not a special case of the nonlinear factor analysis model linear in parameters presented before.

This framework is still incomplete and not general enough.

General nonlinear factor analysis model (parametric)

1993 Amemiya, Y. "On nonlinear factor analysis." Proceedings of the Social Statistics Section, ASA

1993 Amemiya, Y. "Instrumental variable estimation for nonlinear factor analysis." Multivariate Analysis: Future Directions 2

Introduced the nonlinear factor analysis which is nonlinear in the parameters and nonlinear in the factors

$$\mathbf{Z}_i = \mathbf{G}(\mathbf{f}_i, \mathbf{\Lambda}) + \boldsymbol{\epsilon}_i$$

Addressed identification and parameterization issues

Developed estimation and inference procedures

Different from nonlinear simultaneous equations

EV parameterization

$$\begin{aligned}\mathbf{Z}_i &= \mathbf{G}(\mathbf{f}_i, \boldsymbol{\Lambda}) + \boldsymbol{\epsilon}_i \\ &= \begin{pmatrix} \mathbf{g}(\mathbf{f}_i, \boldsymbol{\Lambda}) \\ \mathbf{f}_i \end{pmatrix} + \boldsymbol{\epsilon}_i\end{aligned}$$

- Simple way to express identifiable models
- Allow proper and useful interpretation of model
- Meaningful for multi group and other analyses
- Facilitate instrumental variable estimation
- Allow distributional diagnostics

Instrumental variable estimation

Building on nonlinear errors-in-variables work:
Amemiya (1985 and 1990 *Journal of Econometrics*)
Amemiya and Fuller (1988 *Annals of Statistics*)

- $\mathbf{f}_i$ - fixed or any distribution

- bias adjustment

- statistical properties

- statistical inference for parameters

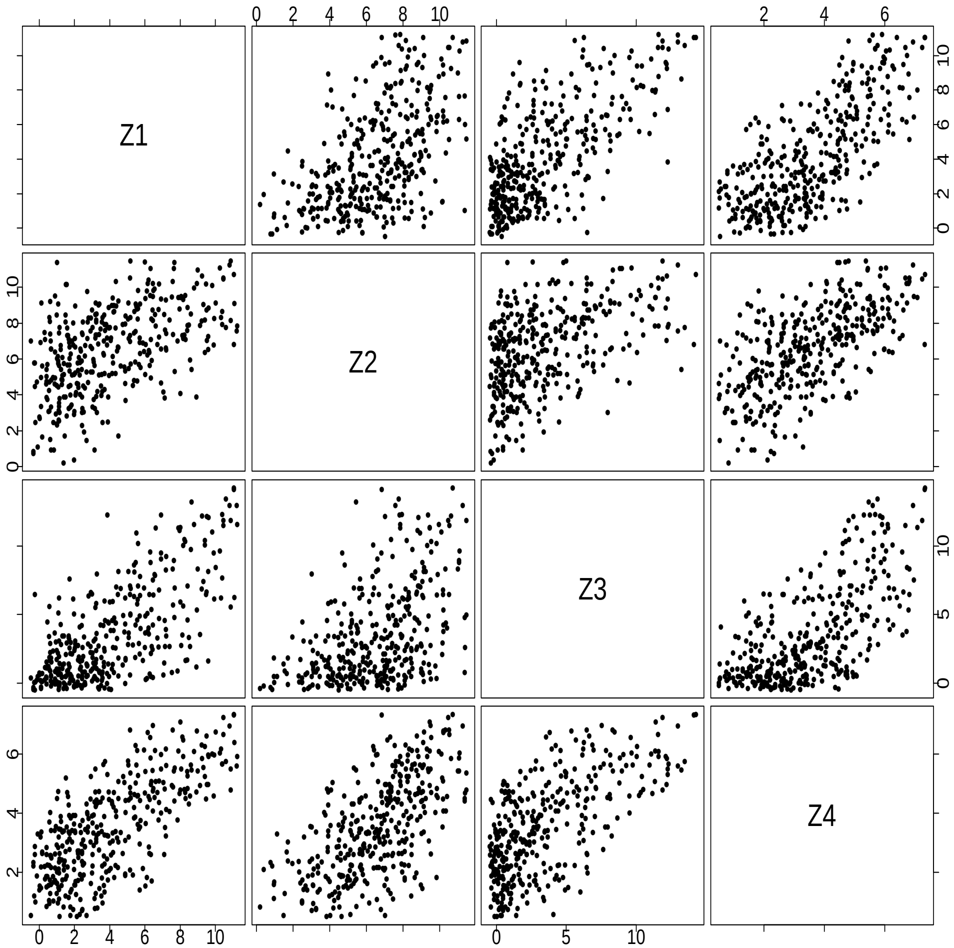
General nonlinear factor analysis 1993-2001

Culminating work:

2001 Yalcin, I. and Amemiya, Y. "Nonlinear factor analysis as a statistical method." *Statistical Science*

- Re-emphasis of EV parameterization
- Two new practical model fitting and analysis procedures based on statistical theory
- Models can be fitted/tested with # of factors condition the same as exploratory linear case
- Statistical inference procedures for parameters
- Statistical procedures for testing overall model fit (related to # of factors)
- Factor score estimation
- Graphical diagnostic methods

Vocational interest data - Dumenci



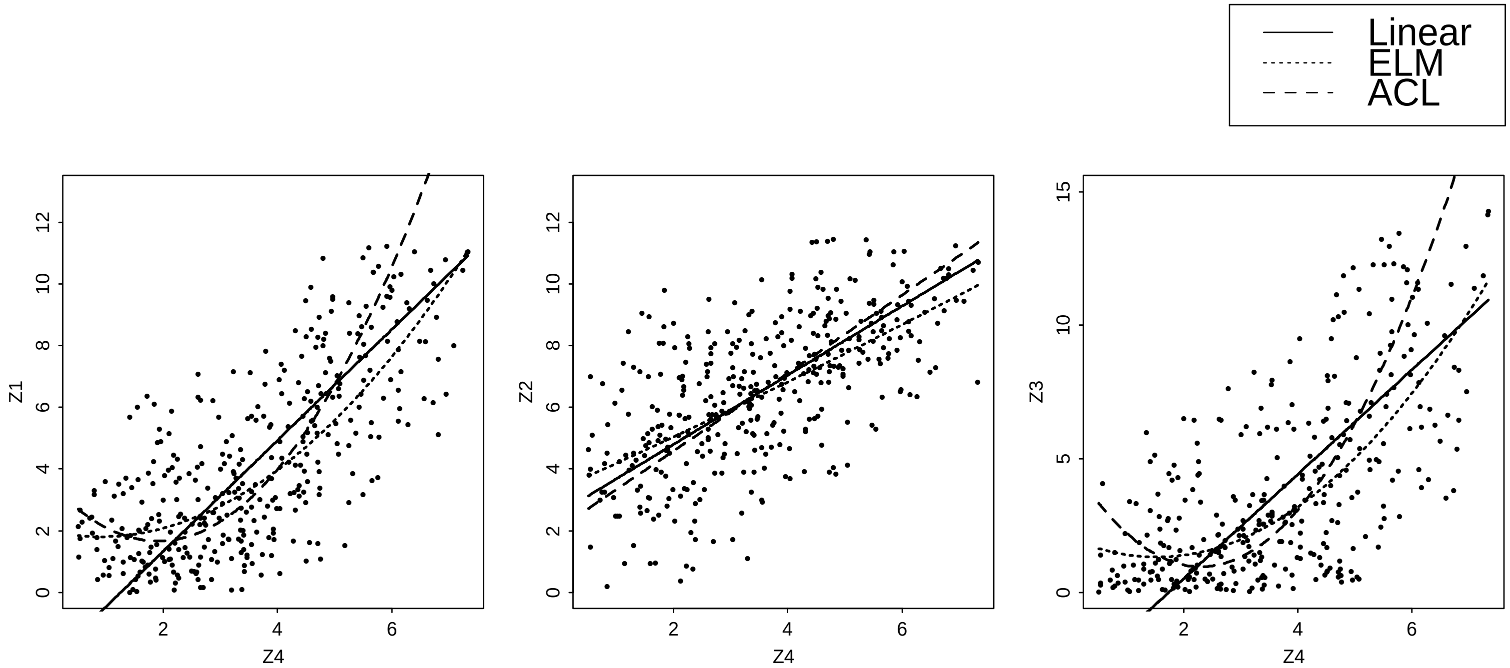
Fit of one factor linear model: $\chi^2 = 17.45$, 2 d.f., p-value=0.0002
Not possible to fit two factor linear model

Yalcin and Amemiya 2001

$$
\begin{aligned}
Z_{1i} &= \lambda_{10} + \lambda_{11} f_{1i} + \lambda_{12} f_{1i}^2 + \epsilon_{1i} \\
Z_{2i} &= \lambda_{20} + \lambda_{21} f_{1i} + \lambda_{22} f_{1i}^2 + \epsilon_{2i} \\
Z_{3i} &= \lambda_{30} + \lambda_{31} f_{1i} + \lambda_{32} f_{1i}^2 + \epsilon_{3i} \\
Z_{4i} &= f_{1i} + \epsilon_{4i}
\end{aligned}
$$

Linear one factor model $\chi^2 = 17.45$, 2 d.f., p-value=0.0002
ELM procedure $\chi^2 = 4.37$, 2 d.f., p-value=0.11
ACL procedure $\chi^2 = 3.78$, 2 d.f., p-value=0.15

Yalcin and Amemiya 2001



General nonlinear structural equation analysis 2001-2004

- Formulation of general nonlinear SEM

- Likelihood-based procedures/algorithms

- Generalized linear model formulation

2001 Amemiya, Y. and Zhao, Y. "Estimation for nonlinear structural equation system with an unspecified distribution." Proceedings of Business and Economic Statistics Section, ASA

2001 Eickhoff, J. C. and Amemiya, Y. "Latent variable modeling for mixed-type outcome variables." Social Statistics Section, ASA

2002 Amemiya, Y. and Zhao, Y. "Pseudo likelihood approach for nonliear and nonnormal structural equation analysis." Business and Economic Statistics Section, ASA

General nonlinear structural equation system

Nonlinear measurement model

$$\mathbf{Z}_i = \mathbf{G}(\mathbf{f}_i) + \boldsymbol{\epsilon}_i$$

Nonlinear structural model

$$\mathbf{H}(\mathbf{f}_i) = \boldsymbol{\delta}_i$$

EV parameterization and reduced form

Nonlinear measurement model in EV parameterization

$$\begin{aligned} \mathbf{Z}_i &= \mathbf{G}(\mathbf{f}_i) + \boldsymbol{\epsilon}_i \\ &= \begin{pmatrix} \mathbf{g}(\mathbf{f}_i, \boldsymbol{\Lambda}) \\ \mathbf{f}_i \end{pmatrix} + \boldsymbol{\epsilon}_i \end{aligned}$$

Nonlinear structural model $\mathbf{H}(\mathbf{f}_i) = \boldsymbol{\delta}_i$ in reduced form

$$\boldsymbol{\eta}_i = \mathbf{h}(\mathbf{w}_i; \boldsymbol{\beta})$$

where

$$\mathbf{f}_i = \begin{pmatrix} \boldsymbol{\eta}_i \\ \boldsymbol{\xi}_i \end{pmatrix}$$

$$\mathbf{w}_i = \begin{pmatrix} \boldsymbol{\delta}_i \\ \boldsymbol{\xi}_i \end{pmatrix}$$

Observation reduced model

$$
\begin{aligned}
\mathbf{Z}_i &= \begin{pmatrix} \mathbf{g}\left(\mathbf{h}(\mathbf{w}_i;\boldsymbol{\beta}), \boldsymbol{\xi}_i; \boldsymbol{\Lambda}\right) \\ \mathbf{h}(\mathbf{w}_i;\boldsymbol{\beta}) \\ \boldsymbol{\xi}_i \end{pmatrix} + \boldsymbol{\epsilon}_i \\
&= \mathbf{G}^*(\mathbf{w}_i, \boldsymbol{\alpha}) + \boldsymbol{\epsilon}_i
\end{aligned}
$$

General nonlinear factor analysis model
with modified EV parameterization

General nonlinear factor analysis

$$\mathbf{Z}_i = \mathbf{G}(\mathbf{f}_i) + \boldsymbol{\epsilon}_i$$

- EV parameterization
- general parametric function
- proper model fitting
- statistical inference for parameters
- test for model fit
- distributional assumption consideration

Last 100 years (statistical models)

- General factor analysis concept

- Statistical formulation of linear factor analysis model

- Linear SEM (LISREL model) development

- Models for discrete observations and factors

- Nonlinear factor analysis linear in parameters

- Nonlinear SEM linear in parameters

- General nonlinear SEM

- General nonlinear factor analysis

Last 100 years (statistical methods)

- General factor analysis concept

- Identification

  orthogonal $\Rightarrow$ EV parameterization $\Rightarrow$ multi population analysis

- Model fitting (factor, SEM, discrete variables)

  correlation $\Rightarrow$ covariance $\Rightarrow$ distribution

  moments $\Rightarrow$ minimum discrepency / likelihood $\Rightarrow$ distribution free

- Statistical inference theory development

  consistency $\Rightarrow$ likelihood theory $\Rightarrow$ asymptotic distribution free

Last 100 years (usage and applications)

Usage

- Multi group

- Longitudinal data analysis

- Time series

- Spatial data analysis

Applications

- Psychology, Sociology, Education, Public Health, Medicine, Marketing, Business, Biological sciences, Geology, etc.

Next 100 years

Statistical Models and Methods - Next 5 years

- Address distributional difficulty in general nonlinear factor analysis

- More general analysis encompassing discrete observations and factors

- Diagnostics and model checking

- Nonlinear analysis for multi group, longitudinal, time series, spatial

- Continue methodological development for subject matter analysis

Next 100 years

Usage and Applications

- More and more applications

- Usage follows subject matter theory development

Subject matter theory development

- What's next?