General Models for Covariance Structures

James H. Steiger

Department of Psychology and Human Development
Vanderbilt University

## General Models for Covariance Structures

Introduction

- One subject we skipped over in our discussion of iteration is how the model matrix function $\mathbf{M}(\boldsymbol{\theta})$ is produced in an efficient way.
- We've already seen how one might construct such a general function "constructor" in the simple case of confirmatory factor analysis.
- One has a list of the elements of $\boldsymbol{\theta}$, and a list of where they are located in the matrices $\boldsymbol{\Lambda}$, $\boldsymbol{\Phi}$ (if the factors are correlated), and $\boldsymbol{\Theta}_\delta$. One then simply constructs $\mathbf{M}(\boldsymbol{\theta}) = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{\Theta}_\delta$.
- But what about more general cases?
- The model has to be manageable, and at the same time have enough generality to be useful across a wide range of circumstances.
- A key factor in early days of mainframe computing was whether derivatives of the model function could be calculated efficiently along with the model function itself.
- A misspecified and/or large model could easily exhaust an entire departmental computer budget estimating one model!

The LISREL Model

- Consider the structural equation model shown on the next slide.
- This, model, known as the "Wheaton model" after the author of the article in which it appeared, is an example of the classic "LInear Structural RELations Model" (LISREL model) often attributed to Karl Jöreskog, the developer of LISREL, the earliest and best-known general purpose structural equation modeling program.
- Note how the model sandwiches a linear regression model (on latent variables) between two confirmatory factor analysis models.

The LISREL Model
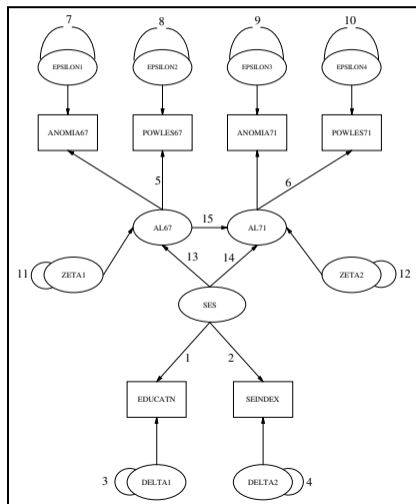
The LISREL Model

- The LISREL model breaks the larger model down into these 3 pieces.
- Perhaps the key equation is the structural equation model, which relates latent variables.

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta} \tag{1}$$

  The endogenous, or dependent latent variables are collected in the vector $\boldsymbol{\eta}$, while the exogenous, or "independent" latent variables are in $\boldsymbol{\xi}$.
- $\mathbf{B}$ and $\boldsymbol{\Gamma}$ are coefficient matrices, while $\boldsymbol{\zeta}$ is a random vector of residuals, sometimes called "errors in equations" or "disturbance terms."
- The elements of $\mathbf{B}$ and $\boldsymbol{\Gamma}$ represent path coefficients for directed relationships among latent variables. It is assumed in general that $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$ are uncorrelated, and that $\mathbf{I} - \mathbf{B}$ is of full rank.

The LISREL Model

- Because usually $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ are not observed without error, there are also factor model (or "measurement model") equations to account for measurement of these latent variables through manifest variables. The "measurement models" for the two sets of latent variables are

$$\mathbf{y} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\epsilon} \qquad (2)$$

and

$$\mathbf{x} = \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta} \qquad (3)$$

with the standard factor analytic assumptions.

The LISREL Model

- We can partition $\boldsymbol{\Sigma}$ as

$$\boldsymbol{\Sigma} = \left[\begin{array}{cc} \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Sigma}_{xx} \end{array} \right] \tag{4}$$

- You can verify that

$$\boldsymbol{\Sigma}_{yy} = \boldsymbol{\Lambda}_y \mathbf{A} \left(\boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Gamma}' + \boldsymbol{\Psi} \right) \mathbf{A}'\boldsymbol{\Lambda}_y{}' + \boldsymbol{\Theta}_\epsilon \tag{5}$$

$$\boldsymbol{\Sigma}_{xx} = \boldsymbol{\Lambda}_x \boldsymbol{\Phi} \boldsymbol{\Lambda}_x' + \boldsymbol{\Theta}_\delta \tag{6}$$

$$\boldsymbol{\Sigma}_{xy} = \boldsymbol{\Lambda}_x \boldsymbol{\Phi} \boldsymbol{\Gamma}' \mathbf{A}' \boldsymbol{\Lambda}_y{}' \tag{7}$$

- $\boldsymbol{\Phi}, \boldsymbol{\Psi}, \boldsymbol{\Theta}_\epsilon$, and $\boldsymbol{\Theta}_\delta$ are the covariance matrices for $\boldsymbol{\xi}$, $\boldsymbol{\zeta}$, $\boldsymbol{\epsilon}$, and $\boldsymbol{\delta}$ respectively.

with $\mathbf{A} = (\mathbf{I} - \mathbf{B})^{-1}$.

The LISREL Model

- Setting up a model in LISREL involved defining each model matrix separately and "telling" the program which elements of each matrix were free parameters or fixed values (often zero).
- Despite the substantial difficulty of learning to use the program, and the fact that some models were not directly specifiable in the LISREL syntax, the program became very popular.
- Those who mastered the program achieved special status, and became known as LISRELites.
- A typical structural equation modeling course spent about a month mastering the LISREL syntax, and usually terminated after, with great fanfare, students managed to execute a couple of really simple models.
- To give you some idea of what was involved, let's determine the model matrices for the Wheaton model. Executing the model on LISREL would involve an additional step of mastering the model syntax and using it to specify the matrices.

The Wheaton Model

- But before we begin, let me ask you a question: Why do `ANOMIA67` and `ANOMIA71` each have a fixed factor loading of 1?
- Let's invoke the "pipeline metaphor."
- Imagine that, instead of a fixed value of 1, there is a free parameter on the path from `AL67` to `ANOMIA67`, and that you are a little person standing between `ANOMIA67` and `POWLES67`, observing the "flow" of numbers arriving at those two boxes.
- Now imagine that I were to suddenly double every number arriving at `AL67` from `SES` and `ZETA1`. What effect would that have on the standard deviation of `AL67`?
- Could I adjust the path coefficients so that the numbers arriving at `ANOMIA67` and `POWLES67` are the same as they were before? How?

The Wheaton Model

- The fixed value of 1 is required to identify the model coefficients.
- I (Steiger, 2002) call the fixed value of 1 a "unit loading identification constraint."
- Some authors of SEM textbooks state that such a loading "fixes the scale" of the latent variable at that of the manifest "indicator."
- Certainly fixing the loading at 1 fixes the variance of the latent variable, but exactly what it fixes it to is something of an accident.
- Suppose now that one has a fixed loading of 1. What would happen in the diagram if you changed it to 3?

The Wheaton Model
Classic LISREL Specification for the Wheaton Model

Below is the standard LISREL input to analyze the Wheaton model.

```
Stability of Alienation, Model B (Uncorrelated Errors)
DA NI=6 NO=932
LA
ANOMIA67 POWLES67 ANOMIA71 POWLES71 EDUCATN SEINDEX
CM FI=EX64.COV
MO NY=4 NX=2 NE=2 NK=1 BE=SD TE=SY
LE
ALIEN67 ALIEN71
LK
SES
FR LY(2,1) LY(4,2) LX(2,1)
VA 1 LY(1,1) LY(3,2) LX(1,1)
PD
OU
```

The Wheaton Model
SIMPLIS Specification for the Wheaton Model

- The LISREL model was needlessly complex as well as being conceptually more unwieldy and less flexible than other models.
- Each manifest variable had to be typed as either "endogenous" or "exogenous" depending on whether the latent variables loading on it were endogenous or exogenous.
- Variables had to be ordered in the correlation matrix according to their status.
- Unfortunately, changing a model might change the status of a variable and require substantial work reordering variables and revising the model specification.
- When McArdle introduced the RAM model specification, quickly followed by the EQS (Bentler-Weeks) and EzPATH models, LISREL responded by introducing a simplified command language called SIMPLIS.
- The next slide shows how the Wheaton model is specified with SIMPLIS.

The Wheaton Model

SIMPLIS Specification for the Wheaton Model

```
Stability of Alienation
Observed Variables
   ANOMIA67   POWLES67   ANOMIA71   POWLES71   EDUCATN   SEINDEX
Covariance Matrix
   11.834
    6.947      9.364
    6.819      5.091     12.532
    4.783      5.028      7.495      9.986
   -3.839     -3.889     -3.841     -3.625     9.610
   -2.190     -1.883     -2.175     -1.878     3.552     4.503
Sample Size 932
Latent Variables  Al67 AL71 SES
Relationships
    ANOMIA67 POWLES67 = AL67
    ANOMIA71 POWLES71 = AL71
    EDUC SEI = SES
    AL67 = SES
    AL71 = AL67 SES
Path Diagram
End of Problem
```

MPlus Specification of the Wheaton Model

- Mplus uses a command language that is fairly efficient but highly specified.
- The next slide shows Mplus input for the Wheaton model.
- The key model specification constructs in Mplus are BY (for specifying loadings), ON (for specifying regression relationships), and WITH (for specifying covariances).

MPlus Specification of the Wheaton Model

```
TITLE: Wheaton Model
DATA: FILE IS WheatonCov.dat;
TYPE IS COVARIANCE;
NOBSERVATIONS = 932;
VARIABLE: NAMES ARE ANOMIA67 POWLES67 ANOMIA71 POWLES71 EDUCATN SEINDEX;
MODEL: AL67 BY ANOMIA67 POWLES67;
AL71 BY ANOMIA71 POWLES71;
SES BY EDUCATN SEINDEX;
AL71 ON SES AL67;
AL67 ON SES;
```

McArdle's RAM Model

- Around 1978, J. J. McArdle proposed his Reticular Action Model (RAM) system for structural equation modeling.
- Although it presented some computational difficulties for larger models, the RAM system was so conceptually straightforward that it caused researchers to examine the LISREL model in a much more critical light.
- This examination led to alternative models that allowed for much more flexible specification and estimation.

McArdle's RAM Model

- Around the same time, McDonald and Fraser had proposed the COSAN model.
- Let Σ be a population variance-covariance matrix for a set of manifest variables.
- The COSAN model (McDonald, 1978) holds if Σ may be expressed as

$$\Sigma = \mathbf{F}_1 \mathbf{F}_2 \ldots \mathbf{F}_k \mathbf{P} \mathbf{F}_k{}' \ldots \mathbf{F}_2{}' \mathbf{F}_1{}'$$

  where $\mathbf{P}$ is symmetric and Gramian, and any of the elements of any $\mathbf{F}$ matrix or $\mathbf{P}$ may be constrained under the model to be a function of the others, or to be specified numerical values.
- As a powerful additional option, any square $\mathbf{F}$ matrix may be specified to be the inverse of a patterned matrix. This "patterned inverse" option is critical for applications to path analysis.
- A COSAN model with k $\mathbf{F}$ matrices is referred to as "a COSAN model of order k."

McArdle's RAM Model

- Suppose you have a path diagram that is complete.
- Let $\mathbf{v}$ be a $(p + n) \times 1$ random vector of p manifest variables and n latent variables in the path model, possibly partitioned into manifest and latent variables subsets in $\mathbf{m}$ and $\mathbf{l}$, respectively, in which case

$$\mathbf{v} = \left[\frac{\mathbf{m}}{\mathbf{l}} \right] \tag{8}$$

McArdle's RAM Model

- For simplicity assume all variables have zero means.
- Let $\mathbf{F}$ be a matrix of multiple regression weights for predicting each variable in $\mathbf{v}$ from the $p + n - 1$ other variables in $\mathbf{v}$.
- $\mathbf{F}$ will have all diagonal elements equal to zero. In general, some elements of $\mathbf{F}$ may be constrained by hypothesis to be equal to each other, or to specified numerical values (often zero).
- Let $\mathbf{r}$ be a vector of latent exogenous variables, including residuals. The path model may then be written

$$\mathbf{v} = \mathbf{F}\mathbf{v} + \mathbf{r} \tag{9}$$

McArdle's RAM Model

- In path models, all endogenous variables are perfectly predicted through the arrows leading to them.
- Since endogenous variables are dependent variables in one or more linear equations, their variances and covariances can be determined from the variances and covariances of the variables with arrows pointing to them.
- Ultimately, the variances and covariances of all endogenous variables are explained by a knowledge of the linear equation set up and the variances and covariances of exogenous variables in the system.
- Consequently, elements of **r** corresponding to endogenous variables in **v** will be null.

McArdle's RAM Model

- The matrix $\mathbf{F}$ contains the regression coefficients normally placed along the arrows in a path diagram. The path coefficient from v_j to v_i is $f_{i,j}$. If a variable v_i is exogenous, i.e., has no arrow pointing to it, then row i of $\mathbf{F}$ will be null, and $r_i = v_i$.
- Hence, the non-null elements of the variance covariance matrix of $\mathbf{r}$ will be the coefficients in the "undirected" relationships in the path diagram.
- Define $\mathbf{P} = E(\mathbf{rr}')$, $\mathbf{W} = E(\mathbf{vv}')$, and $\boldsymbol{\Sigma} = E(\mathbf{mm}')$,
- We can "filter" the manifest variables out of $\mathbf{v}$ with a filter matrix

$$\mathbf{J} = \left[\begin{array}{c|c} \mathbf{I} & \mathbf{0} \end{array} \right]$$

- Since $\mathbf{m} = \mathbf{Jv}$, by the law of linear combinations, we have

$$\boldsymbol{\Sigma} = E(\mathbf{mm}') = \mathbf{J}E(\mathbf{vv}')\mathbf{J}' = \mathbf{JWJ}'$$

McArdle's RAM Model

- It is easy to show (C.P.) that

$$\mathbf{W} = (\mathbf{I} - \mathbf{F})^{-1}\mathbf{P}(\mathbf{I} - \mathbf{F})^{-1\prime} \tag{10}$$

and consequently

$$\mathbf{\Sigma} = \mathbf{J}(\mathbf{I} - \mathbf{F})^{-1}\mathbf{P}(\mathbf{I} - \mathbf{F})^{-1\prime}\mathbf{J}' \tag{11}$$

McArdle's RAM Model

- This shows that any structural equation model expressible in a standard path diagram may be expressed as a COSAN model of order 2, i.e.,

$$\boldsymbol{\Sigma} = \mathbf{F}_1 \mathbf{F}_2 \mathbf{P} \mathbf{F}_2{}' \mathbf{F}_1{}'$$

where

$$\mathbf{F}_1 = \mathbf{J} = \begin{bmatrix} \mathbf{I} & | & \mathbf{0} \end{bmatrix}$$

and

$$\mathbf{F}_2 = (\mathbf{F} - \mathbf{I})^{-1} = \mathbf{B}^{-1}$$

- Since McDonald and Fraser already had in place the machinery to analyze any COSAN model, the implications of this result were profound.
- McArdle had developed an incredibly simple approach to analyzing a structural equation model.

McArdle's RAM Model

McArdle's rules:

1. For convenience, order the manifest variables in the vector **m**, and the latent variables in the vector **l**. The path model is then tested as a COSAN model of order 2, in which

2.
$$\mathbf{F}_1 = \begin{bmatrix} \mathbf{I} & | & \mathbf{0} \end{bmatrix}$$

   where **I** is of order $p \times p$ and **0** is of order $p \times n$.

3. $\mathbf{F}_2$ is the inverse of a square matrix **B** of directed relationships. **B** is constructed from the path diagram as follows. Set all diagonal entries of **B** to -1. Examine the path diagram for arrows. For each arrow pointing from v_j to v_i, record its path coefficient in position $b_{i,j}$ of matrix **B**.

4. **P**, a symmetric matrix, contains coefficients for undirected paths between variables v_i and v_j recorded in positions $\mathbf{P}_{i,j}$ and $\mathbf{P}_{j,i}$.

McArdle's RAM Model

- McArdle's system was easy to implement. One simply took a spreadsheet with all variable names on both the marginal rows and columns, went through the path diagram, and transcribed all the coefficients both fixed and free into the respective positions in **B** and **P**.
- If a variable was eliminated from a model, one simply deleted the corresponding row and column.
- Unfortunately, the RAM model matrices got very large rather quickly.
- The model was quickly modified when it was realized that any factor residual variable could be removed from the list of variables, and its variances and covariances placed in P (on the respective manifest variable) without changing the results.
- This reduced the dimension of the model matrices. Further computational enhancements were developed by Mels (1989) in his M.S. thesis.

The User-Friendly Revolution

- Around that time, Steiger proposed a simplified, character-based method for conveying path diagrams, that led to the truly user-friendly program EzPATH, distributed by Systat.
- With EzPath, the user specified a path model with a series of commands similar to a path diagram. In the followup version of the program, called Sepath, the user could construct a typical confirmatory factor or SEM model with a few clicks of a mouse.
- The program now read the path diagram, figured out for itself which variables were manifest, which were latent, which were exogenous, and which were endogenous.
- EzPATH also featured an improved model that used much smaller matrices than RAM.
- Within a short time, there were numerous structural equation modeling programs, including AMOS, EQS, Mplus, RAMONA, Mx, sem, and, of course, LISREL.

The User-Friendly Revolution

- Now that a structural equation model could be specified and tested in minutes, rather than hours, and on a user's desktop rather than at the university mainframe center, people began to notice things.