The Common Factor Model

Measurement Methods

Lecture 15
Chapter 9
Today’s Class

• Common Factor Model
  – Multiple factors with a single test

• ML Estimation Methods

• New fit indices because of ML Estimation method

• Example
Upcoming Schedule
Our Upcoming Schedule

Note: THE CLASS PROJECT IS DUE ON 5/16

<table>
<thead>
<tr>
<th>Date</th>
<th>Chapter and Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>3/30</td>
<td>Chapter 7: Reliability – Applications</td>
</tr>
<tr>
<td>4/4</td>
<td>Computer Lab session</td>
</tr>
<tr>
<td>4/6, 4/11</td>
<td>No Class</td>
</tr>
<tr>
<td>4/13</td>
<td>Chapter 8: Prediction and Multiple Regression</td>
</tr>
<tr>
<td>4/18, 4/20</td>
<td>Chapter 9: The Common Factor Model</td>
</tr>
<tr>
<td>4/25</td>
<td>Chapter 10: Validity</td>
</tr>
<tr>
<td>4/27</td>
<td>Chapter 11: Classical Item Analysis</td>
</tr>
<tr>
<td>5/2</td>
<td>Chapter 12: Item Response Models</td>
</tr>
<tr>
<td>5/4</td>
<td>Chapter 13: Properties of Item Response Models</td>
</tr>
<tr>
<td>5/9</td>
<td>Chapter 14: Multidimensional Item Response Models</td>
</tr>
<tr>
<td>5/11</td>
<td>Review</td>
</tr>
<tr>
<td>5/16</td>
<td>Final Exam (1:30pm – 4:00pm)</td>
</tr>
</tbody>
</table>
The Single Factor Model

• Spearman single-factor model assumes items form a homogenous set and they measure just one attribute in common.

• This set of items is also described as unidimensional.

• Thurstone generalized this model to more than one factor – multiple factor model.
The Common Factor Model

• The common factor model is a useful device for checking on the structure of the tests we choose to create.

• It is applicable for studying relations between variables that might be:
  – scores on tests or subtests containing enough items to assume linear relations or joint normal distributions.
  – quantitative item scores for which these assumptions will hold approximately.
  – binary item scores, for which the model is a linear approximation to the one we need.
    • For this last point we will need methods introduced in Ch. 12-14.
Common Factor Model Specifics

• Common factor model can also be applied on subtest scores but the component items should first be examined for content homogeneity.

• The item sums substitute for attributes at a first level of generality.

• If relationships among these composites can be explained by common factors, these factors are at a second level of abstraction from the responses to individual items.
Multiple Factors with the Common Factor Model
The Model

- A $p$-common factor model is represented as (for item $j$):

$$X_j = \lambda_{j1} F_1 + \lambda_{j2} F_2 + \ldots + \lambda_{jp} F_p + E_j$$

- $E_j$ is uncorrelated with common factors.
- Common factors, by definition, are what the test scores measure in common.
- The residuals are what they measure uniquely.
Partial Correlation

• Another way to say this is that the partial correlation between any pairs of tests, conditioned on values of the $p$ factors, is zero:

$$\rho_{X_jX_k|F_1F_2\ldots F_p} = 0, \quad j \neq k$$

• That is, in a subpopulation in which the factor scores take fixed values $F_1 = f_1, \ldots, F_p = f_p$, then test scores are linearly uncorrelated.
Conditional Independence

• This is a form of *the principle of conditional independence*.

• Local independence is an important concept in psychometrics.

• Many different models use local independence.  
  – So familiarize yourself with it now…
Predicted Covariances of the Common Factor Model

• Assume the factors are standardized (i.e. with mean zero and variance one).

• The model-predicted covariance between items i and j is then:

$$\sigma_{jk} = \sum_s \sum_t \lambda_{js} \lambda_{kt} \phi_{st}$$

Here, $\phi_{st}$ is correlation between common factors s and t.
Predicted Variances of the Common Factor Model

• The model-predicted variance for an item \( j \) is then:

\[
\sigma_{jj} = \sum_{s} \sum_{t} \lambda_{js} \lambda_{jt} \phi_{st} + \psi_{j}^{2}
\]
Additional Estimation Methods
Estimation Methods

• Two estimation methods are usually used to obtain parameter estimates in common factor model:
  
  – ULS (unweighted least squares – really GLS is more common).
  
  – ML (maximum likelihood – by far the most common method).
Unweighted Least Squares

- ULS principle:
  
  - minimize the sum of squared differences between the sample statistics (sample covariance or correlation) and the corresponding fitted values from the model (fitted covariance or correlation).
Maximum Likelihood

• **ML principle:**
  – Parameter estimates are chosen that maximize the likelihood (probability) of the sample observations.

• **A useful equivalent is to minimize a criterion based on the ratio of two probabilities**
  – The probability of the sample data if the restrictive hypothesis is true.
  – The probability of the data if they could come from a population not subject to any restrictions.
Model Evaluation Procedures
(From ML Estimation)
Chi-squared Test of Model Fit

- Usually, models are evaluated based on chi-squared test or goodness-of-fit indices.

**Chi-squared test**

- Chi-squared is equal to zero if the model fits perfectly.
- It is positive if the model fits imperfectly.
Goodness of Fit Indices

Goodness-of-fit indices

• Based on the non-centrality Parameter,

\[ d = \frac{\chi^2 - df}{n} \]

• Where \( df \) is degree of freedom of the model.
• \( n \) is the sample size.
• This quantity is an unbiased estimate of the error of approximation of the model to the population.
• If the model is perfect fit, \( d = 0 \).
Other Measures – McDonald’s Goodness of Fit Index

- McDonald's Goodness of Fit Index,

\[ M_c = e^{-(1/2)d} \]

- Which is scaled to lie between zero and one, where one represents perfect fit.

- Rod used to tell his classes that he came to dislike this measure of model fit.
Other Measures - RMSEA

- Root Mean Squared Error of Approximation, RMSEA:

\[ RMSEA = \sqrt{d / df} \]

- The RMSEA is zero for perfect fit.

- A conventional "rule of thumb" is that the model is acceptable when RMSEA < .05.
Analysis of Residual Matrix

- Examination of the discrepancy or residual matrix is one way to evaluate model fit (difference between observed and fitted covariance or correlation matrix):
  - A reasonable rule of thumb is that the fit is acceptable, because it cannot be improved on by adding parameters that are nontrivial,
  - If the absolute discrepancies are less than .1 in analysis of a sample correlation matrix.
Factor Analysis Example
Confirmatory Factor Analysis

- As an example, consider the data given on p. 502 of Johnson and Wichern:

  Lawley and Maxwell present the sample correlation matrix of examinee scores for six subject areas and 220 male students.

- The subject tests are:
  - Gaelic.
  - English.
  - History.
  - Arithmetic.
  - Algebra.
  - Geometry.
Confirmatory Factor Analysis

- It seems plausible that these subjects should load onto one of two types of ability: verbal and mathematical.

- If we were to specify what the pattern of loadings would look like, the Factor Loading Matrix might look like:

\[
\Lambda = \begin{bmatrix}
\lambda_{11} & 0 \\
\lambda_{21} & 0 \\
\lambda_{31} & 0 \\
0 & \lambda_{42} \\
0 & \lambda_{52} \\
0 & \lambda_{62} \\
\uparrow & \uparrow \\
\text{Verbal} & \text{Math} \\
\text{Ability} & \text{Ability}
\end{bmatrix}
\]

- \( \rightsquigarrow \) Gaelic
- \( \rightsquigarrow \) English
- \( \rightsquigarrow \) History
- \( \rightsquigarrow \) Arithmetic
- \( \rightsquigarrow \) Algebra
- \( \rightsquigarrow \) Geometry
Confirmatory Factor Analysis

The model-predicted covariance matrix would then be:

\[ \Sigma = \Lambda \Phi \Lambda' + \Psi \]

Where:
- \( \Phi \) is the factor correlation matrix (here it is size 2 \( \times \) 2).
- \( \Psi \) is a diagonal matrix of unique variances.

Specifically:

\[
\Sigma = \begin{bmatrix}
\lambda_{11}^2 + \psi_1 & \lambda_{11} \lambda_{21} & \lambda_{11} \lambda_{31} & \lambda_{11} \phi_{12} \lambda_{42} & \lambda_{11} \phi_{12} \lambda_{52} & \lambda_{11} \phi_{12} \lambda_{62} \\
\lambda_{11} \lambda_{21} & \lambda_{21}^2 + \psi_2 & \lambda_{21} \lambda_{31} & \lambda_{21} \phi_{12} \lambda_{42} & \lambda_{21} \phi_{12} \lambda_{52} & \lambda_{21} \phi_{12} \lambda_{62} \\
\lambda_{11} \lambda_{31} & \lambda_{21} \lambda_{31} & \lambda_{31}^2 + \psi_3 & \lambda_{31} \phi_{12} \lambda_{42} & \lambda_{31} \phi_{12} \lambda_{52} & \lambda_{31} \phi_{12} \lambda_{62} \\
\lambda_{11} \phi_{12} \lambda_{42} & \lambda_{21} \phi_{12} \lambda_{42} & \lambda_{31} \phi_{12} \lambda_{42} & \lambda_{42}^2 + \psi_4 & \lambda_{42} \lambda_{52} & \lambda_{42} \lambda_{62} \\
\lambda_{11} \phi_{12} \lambda_{52} & \lambda_{21} \phi_{12} \lambda_{52} & \lambda_{31} \phi_{12} \lambda_{52} & \lambda_{42} \lambda_{52} & \lambda_{52}^2 + \psi_5 & \lambda_{52} \lambda_{62} \\
\lambda_{11} \phi_{12} \lambda_{62} & \lambda_{21} \phi_{12} \lambda_{62} & \lambda_{31} \phi_{12} \lambda_{62} & \lambda_{42} \lambda_{62} & \lambda_{52} \lambda_{62} & \lambda_{62}^2 + \psi_6
\end{bmatrix}
\]
Confirmatory Factor Analysis

- Using an optimization routine (and some type of criterion function, such as ML), the parameter estimates that minimize the function are found.

- To assess the fit of the model, the predicted covariance matrix is subtracted from the observed covariance matrix, and the residuals are summarized into fit statistics.

- Based on the goodness-of-fit of the model, the result is taken as-is, or modifications are made to the structure.

- CFA is a measurement model - the factors are measured by the data.

- SEM is a model for the covariance between the factors.
*SAS Example;

data examscores (type=corr);
  input _type_ $ _name_ $ gaelic english history arith algebra geometry;
cards;
n  .220 . . . . .
corr gaelic 1.00 . . . .
corr english .439 1.00 . . . .
corr history .410 .351 1.00 . .
corr arith .288 .354 .164 1.00 .
corr algebra .329 .320 .190 .595 1.00 .
corr geometry .245 .329 .181 .470 .464 1.00 ;
run;

proc calis data=examscores residual;
  lineqs
      gaelic  = lambda11 f1 + e1,
      english = lambda21 f1 + e2,
      history = lambda31 f1 + e3,
      arith   = lambda41 f2 + e4,
      algebra = lambda51 f2 + e5,
      geometry = lambda61 f2 + e6;
  std
      f1=1,
      f2=1,
      e1-e6=psi1-psi6;
  cov
      f1 f2=phi1;
run;
Wrapping Up

• The Common Factor Model is a general way to estimate more than a single factor in a test.

• This is a practical way to do test development.

• The ML estimation method is the most common way to estimate model parameters.

• The RMSEA is perhaps the most often cited fit index.
Next Time

• More from Chapter 9
  – How to use the Common Factor Model for exploratory analyses.
  – What to do about Covariance or Correlation Matrix estimation.
  – Differing types of models (independent clusters and hierarchical factors).
  – Another example.