Parameter Identifiability Rules[1]

Note: All the rules listed here assume that errors are independent of exogenous variables that are not errors, and that the variables have been centered to have expected value zero.

1. **Parameter Count Rule:** If a model has more parameters than covariance structure equations, the parameter vector can be identifiable on at most a set of volume zero in the parameter space. This applies to all models.

2. *Latent variable model*: $\mathbf{y}_i = \boldsymbol{\beta}\mathbf{y}_i + \boldsymbol{\Gamma}\mathbf{x}_i + \boldsymbol{\epsilon}_i$ Here, identifiability means that the parameters involved are functions of $cov(\mathbf{F}_i) = \boldsymbol{\Phi}$.

   (a) **Regression Rule:** If no endogenous variables influence other endogenous variables, the model parameters are identifiable.

   (b) **Acyclic Rule**: Parameters of the Latent Variable Model are identifiable if the model is acyclic (no feedback loops through straight arrows) and the following conditions hold.

   - Organize the variables that are not error terms into sets. Set 0 consists of all the exogenous variables.
   - For $j = 1, \ldots, k$, each endogenous variable in set j is influenced by at least one variable in set $j - 1$, and also possibly by variables in earlier sets.
   - Error terms for the variables in a set may have non-zero covariances. All other covariances between error terms are zero. These conditions are satisfied if $\boldsymbol{\Psi}$ is diagonal.

3. *Measurement model* (Factor analysis) In these rules, latent variables that are not error terms are described as "factors."

   Definition: A *reference variable* for a latent variable is an observable variable that is a function only of that latent variable and an error term. The factor loading is non-zero.

   (a) **Double Measurement Rule**: Parameters of the double measurement model are identifiable. All factor loadings equal one. Correlated measurement errors are allowed within sets of measurements, but not between sets.

   (b) **Three-Variable Rule**: For a factor analysis model with a single factor, the parameters will be identifiable provided that

   - There are at least three reference variables.
   - Either the factor is standardized and the sign one factor loading is known, or else at least one factor loading equals one.
   - Errors are independent of one another and of the factors.

   (c) **Four-variable 2-factor Rule**: The parameters of a factor analysis model are identifiable provided

   - There are four observable variables and two factors.
   - There are two reference variables for each factor.
   - For each factor, either the variance equals one and the sign of one factor loading is known, or the factor loading of at least one reference variable is equal to one.
   - The two factors have non-zero covariance.
   - Errors are independent of one another and of the factors.

[1] This handout was prepared by Jerry Brunner, Department of Statistical Sciences, University of Toronto. It is licensed under a Creative Commons Attribution - ShareAlike 3.0 Unported License. Use any part of it as you like and share the result freely. The LaTeX source code is available from the course website: http://www.utstat.toronto.edu/brunner/oldclass/431s23

(d) **Two-variable Addition Rule**: A factor with just two reference variables may be added to a measurement model whose parameters are identifiable, and the parameters of the combined model will be identifiable provided

- The errors for the two additional variables are independent of one another and of the error terms already in the model.
- Either the variance of the additional factor equals one and the sign of one factor loading is known, or the factor loading of at least one reference variable is equal to one.
- In the existing model with identifiable parameters,
  - There is at least one reference variable for each factor, and
  - At least one factor has a non-zero covariance with the new factor.

(e) **Vector 3-variable Rule**: Let $\mathbf{d}_1 = \mathbf{F} + \mathbf{e}_1$, $\mathbf{d}_2 = \mathbf{\Lambda}_2\mathbf{F} + \mathbf{e}_2$ and $\mathbf{d}_3 = \mathbf{\Lambda}_3\mathbf{F} + \mathbf{e}_3$, where $\mathbf{F}$, $\mathbf{d}_1$ and $\mathbf{d}_2$ and $\mathbf{d}_3$ are all $p \times 1$, $\mathbf{\Lambda}_2$ and $\mathbf{\Lambda}_3$ have inverses, and $\mathbf{F}$ and the $\mathbf{e}_j$ are all independent. Then all the parameters are identifiable.

(f) **Factor Model Combination Rule**: Suppose there are two factor analysis models A and B with identifiable parameters. Parameters of the combined model will be identifiable provided

- Every factor has at least one reference variable.
- The error terms of the reference variables in set A have zero covariance with the error terms of the reference variables in set in B.

The combined model includes covariances between factors, and also potentially covariances between error terms from the two sets, except as noted above.

(g) **Extra Variables Rule**: A set of observable variables may be added to a measurement model whose parameters are *already identifiable*, provided

- There is a reference variable for each factor in the existing model.
- Error terms of the additional variables have zero covariance with the error terms of the reference variables in the existing model.

Under these conditions,

- Straight arrows with factor loadings on them may point from each existing factor to each new variable.
- Error terms for the new set of variables may have non-zero covariances with each other, and with the error terms in the original model that do not belong to the reference variables.
- You don't need to include all such links.

(h) **Equivalence Rule**: For a factor analysis model with at least one reference variable for each factor, suppose that surrogate models are obtained by either standardizing the factors, or by setting the factor loading of a reference variable equal to one for each factor. Then the parameters of one surrogate model are identifiable if and only if the parameters of the other surrogate model are identifiable.

(i) **Error-Free Rule**: A vector of observable variables may be added to the factors of a measurement model whose parameters are identifiable, provided that the new observable variables are independent of the errors in the measurement model, and there is at least one reference variable for each factor in the existing model. Parameters of the combined moel are identifiable.

4. **Two-Step Rule**: This applies to models with both a measurement component and a latent variable component, including the full two-stage structural equation model.

   1: Consider the latent variable model as a model for observed variables. Check identifiability (usually using the Regression Rule and the Acyclic Rule).

   2: Consider the measurement model as a factor analysis model, ignoring the structure of $cov(\mathbf{F}_i)$. Check identifiability.

   If both identification checks are successful, the parameters of the combined model are identifiable.