Psychology Seminar
Psych 406
Structural Equation Modeling
Jeffrey D. Leitzel, Ph.D.

Factor Analysis

Outline/overview

- Applications of Factor Analysis
- Types of Factor Analysis (EFA vs. CFA)
- Terminology/Concepts
  - Factor loadings
  - Communality
  - Eigenvalues
- Rotation
- Art of interpretation
- Dataset concerns
- Example(s)

Factor analysis

- widely used (and misused) multivariate technique
- salvage poorly planned and executed research
- fertile ground for "fishing expeditions"
- assumption - smaller number of dimensions underlying relations in the data

Uses of Factor Analysis

- 1. data reduction
  - large number of variables
  - reduce to smaller number of dimensions
- 2. select a subset of variables
  - composite measure
  - drop those that don't fit
- 3. multicollinearity in multiple regression
  - combine highly correlated predictors
  - create uncorrelated factors to use as predictors
- 4. scale/index construction/validation
  - have ideas about areas of domain
  - construct items to measure each
  - determine whether items selected represent coherent constructs

Simple structure

- want items in scales that represent only one factor per item
- items representing more than one factor are factorially complex
- generally drop these items during the measure construction phase

Exploratory vs. Confirmatory

- EFA: any indicator can be associated with any/all other factors
- no restrictions on loadings
- CFA: determine whether the number of factors and the loadings conform with what is expected
- do items purported to measure a factor or latent construct actually belong together?

Terminology - components vs. factors

- principal components analysis yields components
- principal axis factoring yields factors
- will use factors and components interchangeably

Principal Components Analysis

- most commonly used form of factor analysis
- seeks linear combination of variables that extracts the maximum variance
- this variance is removed and the process is repeated

Principal Axis Factoring

- same strategy
- operates only with the common variance
- seeks the smallest # of factors that can account for common variance
- PCA tries to account for common and unique variance

Factor loadings

- correlations between the items and the factors
- squared factor loading is the % of variance in that variable that can be explained by the factor
- in PCA it is labeled the component matrix, in PAF the factor matrix, with an oblique rotation called the pattern matrix.

Communality

- h^2
- squared multiple correlation for a variable using all factors as predictors
- % of variance in the variable that can be explained by all factors

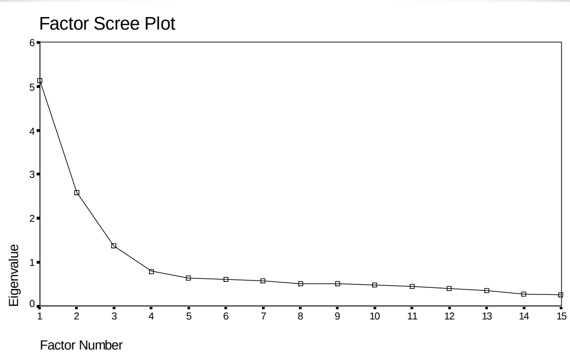
Eigenvalues

- a.k.a. characteristic roots
- reflect variance in all variables accounted for by each factor
- sum of the squared factor loadings
- Eigenvalue/# variables = proportion of variance explained by a factor

Criteria for # of factors to retain:

- 1. Kaiser criterion - keep all with eigenvalues greater than or equal to 1.0
- 2. scree test - plot components on x axis and eigenvalues on y axis
  - where plot levels off the "scree" has occurred
  - keep all factors prior to leveling
  - criticized as generally selecting too few factors
- 3. Comprehensibility - a non mathematical criterion
  - retain factors that can be reasonably interpreted
  - fit with the underlying theory
- ideally, retained factors account for 60 and preferably 75% of variance

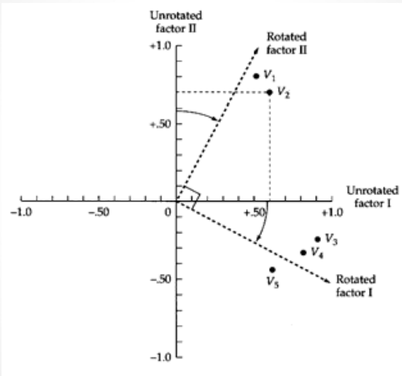
Scree test



Rotation

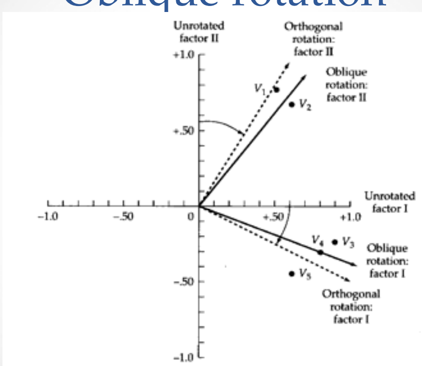
- facilitates interpretation
- unrotated solutions: variables have similar loadings on two or more factors
- makes hard to interpret which variables belong to which factor

Orthogonal rotation



Orthogonal Factor Rotation

Oblique rotation



Oblique Factor Rotation

Rotated and Unrotated Factor Loadings

	Unrotated Factor Loadings		Rotated Factor Loadings	
Variables	I	II	I	II
V1	.50	.80	.03	.94
V2	.60	.70	.16	.90
V3	.90	-.25	.95	.24
V4	.80	-.30	.84	.15
V5	.60	-.50	.76	.13

Types of rotation

- Varimax rotation
  - o most commonly used
  - o uncorrelated factors
- Oblimin
  - o an oblique rotation
  - o allows factors to be correlated
  - o does not mean they will be
- There are many others

When to use oblique rotation?

- constructs not reasonably expected to be uncorrelated
- unsure, request oblique rotation and examine factor correlation matrix, if correlations exceed .32 oblique warranted

How many…?

- …cases?
  - o many "rules" (in order of popularity)
    - 10 cases per item in the instrument
    - subjects to variables ratio of no less than 5
    - 5 times the number of variables or 100
    - minimum of 200 cases, regardless of stv ratio
- …variables?
  - o constructing a scale start with large number of items
  - o measure domains with "best indicators" want at least 3 indicators of each
  - o more indicators = greater reliability of measurement

Interpreting loadings

- minimum cut-off is .3
- .4 or below is considered weak
- .6 and above is considered strong
- moderate at all points in between
- Guidelines from Comrey and Lee (1992)
  - .71 excellent
  - .63 very good
  - .55 good
  - .45 fair
  - .32 poor

Final considerations

- Size of loadings effected by
  - homogeneity of the sample
  - restricted range
    - correlations will be lower
    - smaller loadings worth attention
- Naming factors
  - descriptive names for the factors
  - very important part of process
  - fitting findings into informational network of the field