Exploratory Factor Analysis: rotation

Psychology 588: Covariance structure and factor models

Rotational indeterminacy

- Given an initial (orthogonal) solution (i.e., $\boldsymbol{\Phi} = \mathbf{I}$), there exist infinite pairs of "rotated" factor loading and score matrices such that all have exactly identical fit since

$$\mathbf{x} - \boldsymbol{\delta} = \boldsymbol{\Lambda}\boldsymbol{\xi} = \boldsymbol{\Lambda}\mathbf{T}^{-1}\mathbf{T}\boldsymbol{\xi} = \tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{\xi}}$$

subject to "rows" of $\mathbf{T}$ having unit-norm (sum of squares = 1) so that the total variances of $\boldsymbol{\xi}$ and $\tilde{\boldsymbol{\xi}}$ are the same, i.e., $\mathrm{tr}(\boldsymbol{\xi}\boldsymbol{\xi}') = \mathrm{tr}\left(\tilde{\boldsymbol{\xi}}\tilde{\boldsymbol{\xi}}'\right)$ --- rotation preserves the VAF collectively by the n factors, or equivalently, rotation doesn't change communalities

- If $\mathbf{T}\mathbf{T}' = \mathbf{I}$, $\tilde{\boldsymbol{\Lambda}} = \boldsymbol{\Lambda}\mathbf{T}'$ --- orthogonal (or rigid) rotation which preserves the angles between the initial factors, and so the initial orthogonal factors rotated rigidly to orthogonal factors

- Note that the rotation matrix is defined for "factor score" matrix ξ, not for the loading matrix $\mathbf{\Lambda}$ --- makes no difference in the orthogonal case, but should be clear in the "oblique" case

- Rotational indeterminacy shown in the covariance structure:

$$\mathbf{\Sigma} - \mathbf{\Theta} = \mathbf{\Lambda}\mathbf{\Lambda}' = \mathbf{\Lambda}\mathbf{T}^{-1}\mathbf{T}\boldsymbol{\xi}\boldsymbol{\xi}'\mathbf{T}'\mathbf{T}'^{-1}\mathbf{\Lambda}' = \tilde{\mathbf{\Lambda}}\mathbf{T}\mathbf{T}'\tilde{\mathbf{\Lambda}}' = \tilde{\mathbf{\Lambda}}\mathbf{\Phi}\tilde{\mathbf{\Lambda}}'$$

$$\tilde{\mathbf{\Lambda}} = \mathbf{\Lambda}\mathbf{T}^{-1}, \quad \mathbf{\Lambda} = \tilde{\mathbf{\Lambda}}\mathbf{T}, \quad \mathbf{\Phi} = \mathbf{T}\mathbf{T}'$$

If rows of $\mathbf{T}$ are orthogonal, angles between rotated factors remain orthogonal; otherwise, the angle between factors i and j is defined by

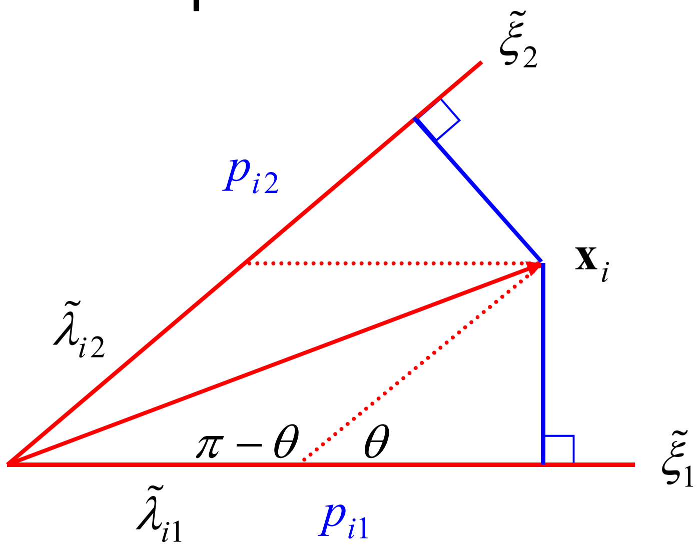
$$\phi_{ij} = \sum_{k=1}^{n} t_{ik} t_{jk} = \cos\theta_{ij}, \quad i \neq j = 1,\ldots,n$$

- Covariances between $\mathbf{x}$ and $\tilde{\boldsymbol{\xi}}$ (a.k.a., “factor structure” matrix $\mathbf{P}$, as compared to “factor pattern” matrix $\tilde{\boldsymbol{\Lambda}}$):

$$E\left(\mathbf{x}\tilde{\boldsymbol{\xi}}'\right) = E\left(\left(\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{\xi}} + \boldsymbol{\delta}\right)\tilde{\boldsymbol{\xi}}'\right) = \tilde{\boldsymbol{\Lambda}}\boldsymbol{\Phi}$$

- Communality is <u>invariant</u> over rotation and represented generally in an oblique system as:

$$h_i^2 = \sum_{k=1}^{n} \tilde{\lambda}_k^2 + 2\sum_{j=1}^{n-1}\sum_{j<k=2}^{n} \tilde{\lambda}_{ij}\tilde{\lambda}_{ik}\phi_{jk}$$



- Could absolute factor loadings in an oblique system be greater than 1 even when a correlation matrix is analyzed (so that communality is upper bounded by 1)?

Simple structure

- To resolve the rotational indeterminacy problem, Thurstone (1947) suggested, so called "simple structure" to be sought to attain n distinctive (not necessarily orthogonal) factors that are likely to be substantively meaningful; satisfying the following conditions:

1. Each row (variable) of the loading matrix should have at least one zero
2. Each column (factor) should have at least n zeros
3. Every pair of columns should have several rows where entries "vanish" in one column but not in the other

$$\begin{bmatrix} 0 & .2 & .2 & .2 \\ 0 & .9 & .2 & .2 \\ 0 & .9 & .2 & .9 \\ 0 & .9 & .9 & .9 \\ .2 & 0 & .2 & .2 \\ .2 & 0 & .9 & .2 \\ .9 & 0 & .9 & .2 \\ .9 & 0 & .9 & .9 \\ .2 & .2 & 0 & .2 \\ .2 & .2 & 0 & .9 \\ .2 & .9 & 0 & .9 \\ .9 & .9 & 0 & .9 \\ .2 & .2 & .2 & 0 \\ .9 & .2 & .2 & 0 \\ .9 & .2 & .9 & 0 \\ .9 & .9 & .9 & 0 \end{bmatrix}$$

4. For every pair of columns, a large proportion of rows have vanishing entries in both columns if $n > 3$

5. Every pair of columns should have only a small number of rows with non-vanishing entries in both columns

- These conditions are, though providing conceptual grounds for rotational criteria, more intuitive than numerically manageable, in that the simple structure is not defined as a scalar-valued optimization function

$$\begin{bmatrix} 0 & .2 & .2 & .2 \\ 0 & .9 & .2 & .2 \\ 0 & .9 & .2 & .9 \\ 0 & .9 & .9 & .9 \\ .2 & 0 & .2 & .2 \\ .2 & 0 & .9 & .2 \\ .9 & 0 & .9 & .2 \\ .9 & 0 & .9 & .9 \\ .2 & .2 & 0 & .2 \\ .2 & .2 & 0 & .9 \\ .2 & .9 & 0 & .9 \\ .9 & .9 & 0 & .9 \\ .2 & .2 & .2 & 0 \\ .9 & .2 & .2 & 0 \\ .9 & .2 & .9 & 0 \\ .9 & .9 & .9 & 0 \end{bmatrix}$$

Graphical rotation with small n

- When $n = 2$, it's possible to find new axes in the factor space that go through n distinctive clusters of variables, or that many variables vanish on one factor and stand out on the other

- When $n = 3$, the graphical rotation should alternate across all pairs of factors and iterate until a satisfactory loading matrix is obtained --- since rotating two axes in a 2-dimensional subspace will change coordinates on the other axis unless it's orthogonal to the $(n-2)$ dimensional subspace

- With $n > 3$, this graphical process becomes very tedious and needs some subjective judgment; and so a quantitative criterion was of a great demand

Analytic rotations

- Analytic rotations operationally define the simple structure by a "simplicity function"; unfortunately, there are quite many

Quartimax (Carroll, 1953) seeks an $n \times n$ nonsingular orthogonal rotation matrix $\mathbf{T}$ so as to maximize the overall variance of all squared loadings with its simplicity function defined as:

$$f_Q = \frac{1}{nq}\sum_{k=1}^{n}\sum_{j=1}^{q}\left(\tilde{\lambda}_{jk}^2\right)^2 - \left(\lambda_M^2\right)^2,$$

$$\lambda_M^2 = \frac{1}{nq}\sum_{k=1}^{n}\sum_{j=1}^{q}\tilde{\lambda}_{jk}^2, \qquad \mathbf{TT}' = \mathbf{I}, \quad \text{and so} \quad \tilde{\mathbf{\Lambda}} = \mathbf{\Lambda T}'$$

- Quartimax tends to yield a "g" factor and $n-1$ small factors

Varimax (Kaiser, 1958) seeks an $n \times n$ nonsingular orthogonal rotation matrix $\mathbf{T}$ so that the sum of variances of squared loadings within columns are maximized with its simplicity function defined as:

$$f_V = \frac{1}{q^2} \sum_{k=1}^{n} \left[q \sum_{j=1}^{q} \left(\tilde{\lambda}_{jk}^2 \right)^2 - \left(\sum_{j=1}^{q} \tilde{\lambda}_{jk}^2 \right)^2 \right]$$

- A normalized version uses adjusted squared loadings for communality sizes to make an even contribution of variables to the function; also true for all other analytic rotations

$$\overset{*}{\lambda}{}_{jk}^2 = \tilde{\lambda}_{jk}^2 \big/ h_j^2$$

- Varimax tends to produce rather evenly sized factors, compared to Quartimax

Oblimin (Carroll, 1960) minimizes sum of covariances between all paired columns of squared elements of the structure matrix $\mathbf{P}$, with the following simplicity function; $\mathbf{TT'} \neq \mathbf{I}$

$$f_O = \sum_{k>k'=1}^{n} \left[q \sum_{j=1}^{q} p_{jk}^2 p_{jk'}^2 - \gamma \left(\sum_{j=1}^{q} p_{jk}^2 \right) \left(\sum_{j=1}^{q} p_{jk'}^2 \right) \right]$$

where γ is a user-provided parameter and ranges $[0,1]$, which controls degree of obliqueness as:

- Quartimin: $\gamma = 0$ --- most oblique
- Biquartimin: $\gamma = 0.5$ --- less oblique; recommended by author
- Covarimin: $\gamma = 1$ --- least oblique; this criterion is equivalent to the Varimax criterion if $\mathbf{T}$ is orthogonal

Direct Oblimin (Jennrich & Sampson, 1966) is equivalent to Oblimin, but the simplicity function is minimized directly with factor loadings (i.e., pattern matrix), instead of elements of the structure matrix as:

$$f_{DO} = \sum_{k>k'=1}^{n} \left[\sum_{j=1}^{q} \tilde{\lambda}_{jk}^2 \tilde{\lambda}_{jk'}^2 - \frac{\delta}{q} \left(\sum_{j=1}^{p} \tilde{\lambda}_{jk}^2 \right) \left(\sum_{j=1}^{p} \tilde{\lambda}_{jk'}^2 \right) \right]$$

where δ is a user-provided parameter, which controls degree of obliqueness as for Oblimin, but its range is not bounded

- With $\delta = 0,$ the solution is most oblique (maybe called "Direct Quartimin); the loss function f_{DO} approaches $-\infty$ iff $\delta > 0.8$ and the solution becomes less oblique with negative and smaller δ

Procrustes rotation --- as an alternative to analytic rotation, if an ideal loading pattern is known, the initial loading matrix may be rotated maximally toward the ideal (named after the inn keeper in Greek Myth who chopped or stretched his customer to fit to his bed)

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

- Or, an ideal pattern might be enforced via constraints, as in confirmatory FA

$$\begin{bmatrix} \lambda_{11} & 0 & 0 \\ \lambda_{21} & 0 & 0 \\ \lambda_{31} & 0 & 0 \\ 0 & \lambda_{42} & 0 \\ 0 & \lambda_{52} & 0 \\ 0 & \lambda_{62} & 0 \\ 0 & 0 & \lambda_{73} \\ 0 & 0 & \lambda_{83} \\ 0 & 0 & \lambda_{93} \end{bmatrix}$$

- Given a target $\mathbf{\Lambda}_T$, the least-square solution for the rotation matrix $\mathbf{T}$ is found as --- assuming zero column means of $\mathbf{\Lambda}_T$ and $\mathbf{\Lambda}$, and $\mathrm{tr}(\mathbf{\Lambda}_T'\mathbf{\Lambda}_T) = \mathrm{tr}(\mathbf{\Lambda}'\mathbf{\Lambda})$

$$\mathbf{C} = \mathbf{\Lambda}_T'\mathbf{\Lambda} = \mathbf{USV}', \quad \mathbf{U}'\mathbf{U} = \mathbf{V}'\mathbf{V} = \mathbf{I}_n$$

$$\tilde{\mathbf{\Lambda}} = \mathbf{\Lambda T}, \quad \mathbf{T} = \mathbf{UV}'$$

- Procrustes rotation may be called orthogonal target rotation to be specific; in comparison, procrustes analysis refers to a linear transformation, consisting of translation and overall scaling (a.k.a., dilation) as well as the orthogonal target rotation

- Some “sign flipping” may be included in the rotation matrix in addition to literal “rotation”

Promax --- (a) an orthogonal $\mathbf{T}$ is found first (e.g., by Varimax); (b) entries of $\mathbf{T}$ are raised to k-th power (typically with $k = 4$ as default) to exaggerate the distinction between small and large loadings; (c) the raised loading matrix is used as a target to rotate obliquely the original loading matrix

- 2 factors extracted from the national track data and rotated by Varimax (black) and by Direct Oblimin (red)