Factor Analysis:

Introduction:

Similar to PC, goal is to find a few factors that describe a large number of variables.

First famous factor analysis:

From a number of students taking 3 exams in Classics (x_1), French (x_2), and English (x_3). Spearman in 1904 gave the following Correlation matrix:

$$\mathbf{S} = \begin{pmatrix} 1 & 0.83 & 0.78 \\ 0.83 & 1 & 0.67 \\ 0.78 & 0.67 & 1 \end{pmatrix}$$

Spearman believed one factor, "general ability", generates the three exams. I.e.:

$$x_1 = \lambda_1 f + u_1,\ x_2 = \lambda_2 f + u_2,\ x_3 = \lambda_3 f + u_3,$$

where, f is the underlying "common factor", λ_1, λ_2, λ_3 are the "factor loadings" and u_i's are errors.

Two sources of errors:

1) Ability in a subject not *perfectly* determined by any factor.

2) Measurement on exam not a *perfect* measure of ability in subject.

Model:

Let $\mathbf{x}$ be a random $p-$vector with $\mathrm{E}(\mathbf{x}) = \mu$ and $\mathrm{Var}(\mathbf{x}) = \mathbf{\Sigma}$. A model with k factors is given by:

$$\mathbf{x} - \mu = \mathbf{\Lambda f} + \mathbf{u},$$

where $\mathbf{\Lambda}$ is a $p \times k$ matrix of parameters and $\mathbf{f}$ and $\mathbf{u}$ are random vectors. The items in $\mathbf{f}$ are "common factors" and the items in $\mathbf{u}$ are "unique factors".

Usual Assumptions:

1) $\mathrm{E}(\mathbf{f}) = 0$

2) $\mathrm{Var}(\mathbf{f}) = \mathbf{I}$

3) $\mathrm{E}(\mathbf{u}) = 0$

4) $\mathrm{Var}(\mathbf{u}) = \mathbf{\Psi}$ ($\mathbf{\Psi} = \mathrm{diag}(\psi_{11}, ...\psi_{pp})$)

5) $\mathrm{Cov}(\mathbf{f}, \mathbf{u}) = \mathbf{0}$.

All factors are uncorrelated and common factors standardized. Componentwise:

$$x_i - \mu_i = \sum_{j=1}^{k} \lambda_{ij} f_j + u_i,$$

for $i = 1, ..., p$.

Thus:

$$\sigma_{ii} = \sum_{j=1}^{k} \lambda_{ij}^2 + \psi_{ii}.$$

Two components to $\text{Var}(\mathbf{x})$:

"communality": $h_i^2 := \sum_{j=1}^{k} \lambda_{ij}^2$

"unique variance": ψ_{ii}

Using assumptions 1)-5) have:

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}.$$

Factor Loadings $\boldsymbol{\Lambda}$ are not unique:

$$\mathbf{x} - \mu = \boldsymbol{\Lambda}\mathbf{f} + \mathbf{u}$$

implies that

$$\mathbf{x} - \mu = (\boldsymbol{\Lambda}\mathbf{G})(\mathbf{G}'\mathbf{f}) + \mathbf{u},$$