MODEL SELECTION IN FACTOR ANALYSIS

ZHINING WANG

Notation is presented in Table **??**.

What is factor analysis and why is it important?

Factor analysis is a mathematical model which tries to use fewer underlying factors to explain the correlation between a large set of observed variables(Mardia, Kent, and Bibby 1979). It provides a useful tool for exploring the covariance structure among observable variables (Hirose and Yamamoto 2015). One of the major assumptions that factor analytic model stands on is that it is impossible for us to observe those underlying factors directly. This assumption is especially suited to subjects like psychology where we cannot observe exactly some concept like how intelligent our subjects are(Mardia, Kent, and Bibby 1979).

Suppose we have a observable random vector $\mathbf{y} \in \mathbb{R}^p$ with mean $\mathbb{E}[\mathbf{y}] = \mu$ and variance $\mathbb{V}[\mathbf{y}] = \Sigma$. Then a d-order factor analysis model for $\mathbf{y}$ can be given by

$$\mathbf{y} = \Lambda \mathbf{f} + \mu + \epsilon, \tag{1}$$

where $\Lambda \in \mathbb{R}^{p \times d}$ is called *loading matrix*, we call $\mathbf{f} \in \mathbb{R}^d$ as *common vectors* and $\epsilon \in \mathbb{R}^p$ is *unique factors*. To make the model well-defined, we may assume

$$\mathbb{E}[\mathbf{f}] = \mathbf{0}_d, \mathbb{V}[\mathbf{f}] = \mathbf{I}_{d\times d}, \mathbb{E}[\epsilon] = \mathbf{0}_p, \mathbb{V}[\epsilon] =: \Psi = \mathrm{diag}(\psi_{11}, \dots, \psi_{pp})$$

and also the independence between any elements from $\mathbf{f}$ and ϵ separately, i.e.

$$\mathbf{Cov}[\mathbf{f}_i, \epsilon_j] = 0, \text{for all } i \in \{1, 2, \dots, d\} \text{ and } j \in \{1, 2, \dots, p\}$$

Straightforwardly, the covariance of observable vector $\mathbf{y}$ can be modelled by

$$\mathbb{V}[\mathbf{y}] = \Lambda\Lambda^\top + \Psi \tag{2}$$

Indeterminacy of the loading matrix

One can easily see that if our factor analytic model is given by (1), then it can also be modelled as

$$\mathbf{y} = (\Lambda \mathbf{M})(\mathbf{M}^\top \mathbf{f}) + \mu + \epsilon$$

where the matrix $\mathbf{M}$ is orthogonal and simultaneously the variance of $\mathbf{y}$ given by (2) still holds, since

$$\mathbb{V}[\mathbf{y}] = (\Lambda \mathbf{M}\mathbf{M}^\top)\mathbb{V}[\mathbf{f}](\Lambda \mathbf{M}\mathbf{M}^\top)^\top + \Psi = \Lambda\Lambda^\top + \Psi.$$

Date: 2024-03-13.