STAT 3690 Lecture 25

zhiyanggeezhou.github.io

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca)

Mar 30, 2022

Estimating $\mathbf{L}$ and $\boldsymbol{\Psi}$ (con'd)

- ML method
  - Further assumptions
    - $\mathbf{F} \sim MVN_q(\mathbf{0}, \mathbf{I})$
    - $\mathbf{E} \sim MVN_p(\mathbf{0}, \boldsymbol{\Psi})$
    - $\mathbf{L}^\top \boldsymbol{\Psi}^{-1} \mathbf{L}$ is diagonal
  - `factanal` or `psych::fa`

- Comments on estimation of $\mathbf{L}$ and $\boldsymbol{\Psi}$
  - Other methods
  - Different statistical softwares may apply different methods
    - Have to look into help manuals to figure out what is going on for different
  - Compare the outputs of multiple estimation methods
    - For a good fit, similar answers would be reached regardless of the method

Factor rotation

- $\mathbf{L}$ is not uniquely defined: if $\mathbf{Y} - \boldsymbol{\mu} = \mathbf{LF} + \mathbf{E}$, then $\mathbf{Y} - \boldsymbol{\mu} = \widetilde{\mathbf{L}}\widetilde{\mathbf{F}} + \mathbf{E}$, where
  - $\widetilde{\mathbf{L}} = \mathbf{LP}$ and $\widetilde{\mathbf{F}} = \mathbf{P}^\top \mathbf{F}$ with $\mathbf{P}$ a $q \times q$ orthogonal matrix, i.e., $\mathbf{PP}^\top = \mathbf{I}$
- A blessing to improve interpretation: pick up a $\mathbf{P}$ such that $\widetilde{\mathbf{F}}$ is more interpretable; to ease interpretation, we want:
  - Each entry of $\mathbf{Y}$ to have large loadings for merely one common factor and negligible loadings for the others
- varimax: find rotation $\mathbf{P}$ to maximize the sum of variance of squared (scaled) loadings over all the common factors
$$\sum_{j=1}^q \left\{ \frac{1}{p} \sum_{i=1}^p \tilde{\ell}_{ij}^{*4} - \left(\frac{1}{p} \sum_{i=1}^p \tilde{\ell}_{ij}^{*2} \right)^2 \right\}$$
  - $\tilde{\ell}_{ij}^* = \tilde{\ell}_{ij} / \sum_{j=1}^q \tilde{\ell}_{ij}^2$ with $\tilde{\ell}_{ij}$ the (i,j)-th entry of $\widetilde{\mathbf{L}} = \mathbf{LP}$

- Comments on factor rotation
  - Especially useful with loadings obtained through ML
  - Sometimes used for PCA loadings

Factor scores

- Weighted least square (WLS) method
  - Given $\bar{\mathbf{Y}}$, $\widehat{\mathbf{L}}$, and $\widehat{\boldsymbol{\Psi}}$
  - For the ith observation $\mathbf{Y}_i$, to minimize $(\mathbf{Y}_i - \bar{\mathbf{Y}} - \widehat{\mathbf{L}}\mathbf{F})^\top \widehat{\boldsymbol{\Psi}}^{-1} (\mathbf{Y}_i - \bar{\mathbf{Y}} - \widehat{\mathbf{L}}\mathbf{F})$ with respect to $\mathbf{F}$
  - $\widehat{\mathbf{F}}_i = (\widehat{\mathbf{L}}^\top \widehat{\boldsymbol{\Psi}}^{-1} \widehat{\mathbf{L}})^{-1} \widehat{\mathbf{L}}^\top \widehat{\boldsymbol{\Psi}}^{-1} (\mathbf{Y}_i - \bar{\mathbf{Y}})$

- Regression method
  - Under normality $\mathbf{F} \sim MVN_p(\mathbf{0}, \mathbf{I})$ and $\mathbf{E} \sim MVN_p(\mathbf{0}, \boldsymbol{\Psi})$
    * $[\mathbf{Y}^\top - \boldsymbol{\mu}^\top, \mathbf{F}^\top]^\top$ is of zero mean and normally distributed with covariance matrix

$$\begin{bmatrix} \mathbf{L}\mathbf{L}^\top + \Psi & \mathbf{L} \\ \mathbf{L}^\top & \mathbf{I} \end{bmatrix}$$

    * Hence $\mathbf{F} \mid \mathbf{Y}$ is normally distributed with mean $\mathbf{L}^\top(\mathbf{L}\mathbf{L}^\top + \boldsymbol{\Psi})^{-1}(\mathbf{Y} - \boldsymbol{\mu})$ and covariance matrix $\mathbf{I} - \mathbf{L}^\top(\mathbf{L}\mathbf{L}^\top + \boldsymbol{\Psi})^{-1}\mathbf{L}$
  - Given $\bar{\mathbf{Y}}$, $\widehat{\mathbf{L}}$, and $\widehat{\boldsymbol{\Psi}}$,

$$\widehat{\mathbf{F}}_i = \widehat{\mathbf{L}}^\top (\widehat{\mathbf{L}}\widehat{\mathbf{L}}^\top + \widehat{\boldsymbol{\Psi}})^{-1} (\mathbf{Y}_i - \bar{\mathbf{Y}})$$

    * Sometimes replace $\widehat{\mathbf{L}}\widehat{\mathbf{L}}^\top + \widehat{\boldsymbol{\Psi}}$ with $\mathbf{S}$

- Comments on factor scores
  - More methods available
  - No uniformly superior way