Stat 571

11-26-13

(1)

Simple Structure (Thurstone)

1.) Each row should have at least one zero.

2.) Each column should have at least m zeros

↑

factors

3.) For each pair of columns, there should be several variables loading on exactly 1 of the columns, and very few variables loading on both columns.

(2)

Rotation methods

Varimax rotation maximizes

$$V = \frac{1}{p} \sum_{j=1}^{m} \left[\sum_{i=1}^{p} \tilde{l}_{ij}^{4} - \frac{\left(\sum_{i=1}^{p} \tilde{l}_{ij}^{2} \right)^2}{p} \right]$$

where $\tilde{l}_{ij} = \frac{l_{ij}}{h_i}$ ← $(ij)^{th}$ entry of L; h_i ← i^{th} communality

Note that this maximizes the variance of the squared loadings.

③

Quartimax rotation minimizes

$$\sum_{i=1}^{p} \sum_{j \neq k} l_{ij}^2 l_{ik}^2$$

Note: this would be exactly 0 if every row contains a 0.

Minimizing $\sum_{i=1}^{p} \sum_{j \neq k} l_{ij}^2 l_{ik}^2$ is equivalent to maximizing $\sum_{i=1}^{p} \sum_{j=1}^{m} l_{ij}^4$

④

Steps to follow:

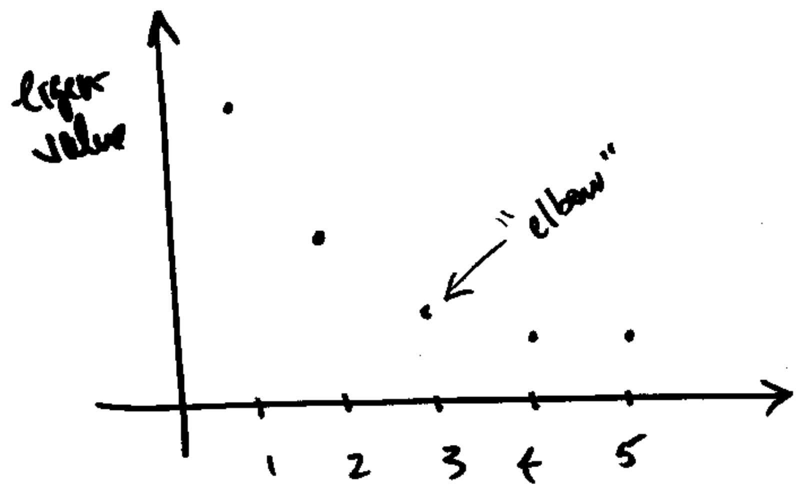
0. Decide whether to use S or R.
1. Extract the factors by PC and at least one other method (e.g. max. likelihood)
2. Decide how many factors to keep — scree plot
3. Decide on a rotation

Look at all of the solutions & choose the one with simplest structure.

4. Try to name the factors

⑤

Scree plot is a plot of the eigenvalues



⑥

Factor scores: - usually found using regression techniques
- represent the estimated values for each observation, on these latent variables

---

Discriminant Analysis (Fisher, 1936)

Goal: to describe in as few dimensions as possible the features of objects from several known populations (Discrimination)

Goal 2: to sort new objects into one of the existing populations (Classification) ⑦

Fisher's method for 2 populations

Let π_1 and π_2 be the names of the 2 populations.

Let $\vec{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_p \end{bmatrix}$ be a random vector.

Let $\vec{\mu}_1 = E[\vec{X} \mid \pi_1]$ and $\vec{\mu}_2 = E[\vec{X} \mid \pi_2]$

⑧

Assume that $\Sigma_1 = \Sigma_2 = \Sigma$

Let Y be a random variable, $Y = \vec{\ell}\,'\vec{X}$

Fisher's approach: find $\vec{\ell}$ so that $E[Y \mid \pi_1]$ and $E[Y \mid \pi_2]$ are as different as possible.

We maximized $\dfrac{\left(E[Y \mid \pi_1] - E[Y \mid \pi_2]\right)^2}{\sigma_Y^2}$.

(9)

$$Y = \vec{\ell}'\vec{X}$$

$$E[Y|\pi_1] = E[\vec{\ell}'\vec{X}|\pi_1]$$
$$= \vec{\ell}' E[\vec{X}|\pi_1] = \vec{\ell}'\vec{\mu}_1$$

$$E[Y|\pi_2] = \vec{\ell}'\vec{\mu}_2$$

$$\sigma_Y^2 = \mathrm{Var}(\vec{\ell}'\vec{X}) = \vec{\ell}'\Sigma\vec{\ell}$$

$$\frac{(\vec{\ell}'\vec{\mu}_1 - \vec{\ell}'\vec{\mu}_2)^2}{\vec{\ell}'\Sigma\vec{\ell}} = \frac{[\vec{\ell}'(\vec{\mu}_1 - \vec{\mu}_2)]^2}{\vec{\ell}'\Sigma\vec{\ell}}$$

(10)

$$= \frac{(\vec{\ell}'\vec{\delta})^2}{\vec{\ell}'\Sigma\vec{\ell}}$$

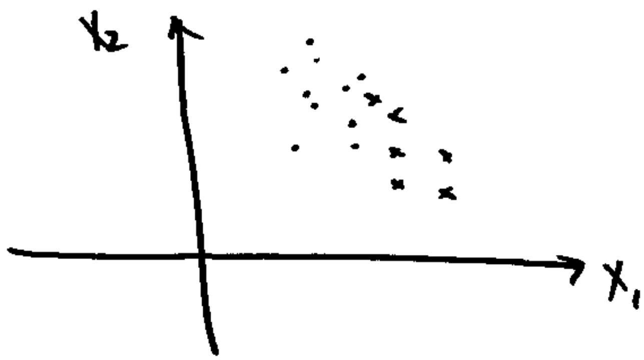
Recall: $$\max_{\vec{x} \neq \vec{0}} \frac{(\vec{x}'\vec{d})^2}{\vec{x}'B\vec{x}} = \vec{d}'B^{-1}\vec{d},$$

achieved when $\vec{x} = cB^{-1}\vec{d}$

So the maximum is $\vec{\delta}'\Sigma^{-1}\vec{\delta}$, achieved when $\vec{\ell} = c\Sigma^{-1}\vec{\delta}$

Then
$$Y = \vec{\lambda}'\vec{X}$$
$$= c\,\vec{\delta}'\,\Sigma^{-1}\vec{X}$$
$$= c(\vec{\mu}_1 - \vec{\mu}_2)'\,\Sigma^{-1}\vec{X}$$

Defn: This is Fisher's Linear Discriminant Function



HW #7
#9.19 (a,b)