Midterm Friday March 4 1:30 -3 PM

- covers material on Ass 1& 2 (up to Feb 12 lecture notes except Factor analysis)

Extra OH: Tuesday 2-3 PM

$$\mathrm{Cov}(\underset{\sim}{F}) = I$$

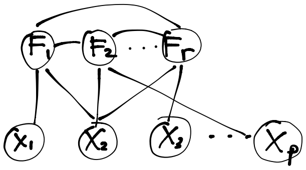
$$\mathrm{Cov}(\underset{\sim}{\varepsilon}) = \Psi = \begin{pmatrix} \psi_1 & & 0 \\ & \ddots & \\ 0 & & \psi_p \end{pmatrix}$$

Factor analysis

Model: $$\underset{\substack{\downarrow \\ \text{length } p}}{\underset{\sim}{X}} = \underset{\substack{\downarrow \\ p\times r}}{\underset{\sim}{\mu}} + L \underset{\substack{\uparrow \\ \text{factors (length } r<p)}}{\underset{\sim}{F}} + \underset{\sim}{\varepsilon}$$

$$\mathrm{Cov}(\underset{\sim}{X}) = LL^T + \Psi$$

$$\begin{pmatrix} \underset{\sim}{X} \\ \underset{\sim}{F} \end{pmatrix} \begin{matrix} \leftarrow \text{observed} \\ \leftarrow \text{unobserved} \end{matrix} \qquad \mathrm{Cov}\begin{pmatrix} \underset{\sim}{X} \\ \underset{\sim}{F} \end{pmatrix} = \begin{pmatrix} \Psi + LL^T & L \\ L^T & I \end{pmatrix}$$



— all the dependence between $\underset{\sim}{X}_1, \cdots \underset{\sim}{X}_p$ is driven by the unobserved factors $F_1, \cdots, F_r$

Question: Given data $\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n$, how to estimate L, Ψ?

Assume: r is known (specified)

Starting point: $\mathrm{Cov}(\underset{\sim}{X}) = LL^T + \Psi$

$$\text{estimated by} \quad \overset{\nearrow}{S} = \frac{1}{n-1} \sum_{i=1}^{n} (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_i - \bar{\underset{\sim}{x}})^T$$

Suppose we know L: Then $S \doteq LL^T + \Psi \leftarrow$ diagonal

$$\hat{\Psi} = \mathrm{diag}(\underbrace{S - LL^T}_{\begin{pmatrix} \ddots & 0 \\ 0 & \ddots \end{pmatrix}})$$

But, this assumes that all diagonal elements of $S - LL^T$ are positive.

Now sps we know Ψ: Then

$$LL^T \doteq S - \Psi \leftarrow \text{symmetric \& hopefully non-negative definite}$$

$$S - \Psi = V\Lambda V^T = (\underset{\sim}{v}_1 \cdots \underset{\sim}{v}_r) \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_r \end{pmatrix} \begin{pmatrix} \underset{\sim}{v}_1^T \\ \vdots \\ \underset{\sim}{v}_r^T \end{pmatrix} \quad \text{where } \lambda_1 \geq \cdots \geq \lambda_r$$

$$\doteq \underbrace{(\sqrt{\lambda_1}\,\underset{\sim}{v}_1 \quad \sqrt{\lambda_2}\,\underset{\sim}{v}_2 \quad \cdots \quad \sqrt{\lambda_r}\,\underset{\sim}{v}_r)}_{\hat{L}} \underbrace{\begin{pmatrix} \sqrt{\lambda_1}\,\underset{\sim}{v}_1^T \\ \vdots \\ \sqrt{\lambda_r}\,\underset{\sim}{v}_r^T \end{pmatrix}}_{\hat{L}^T}$$

$$C = \mathrm{Cov}(\underset{\sim}{X}) = LL^T + \Psi$$

$$\underset{S}{\uparrow} \qquad \underset{\text{known}}{\uparrow}$$

$$\overset{LL^T}{\|} \; C - \Psi = V^* \Lambda^* V^{*T}$$

$$\Lambda^* = \begin{pmatrix} \lambda_1 & & & & 0 \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & 0 & \\ 0 & & & & \ddots \; 0 \end{pmatrix}$$

Principal factors method

Algorithm:

⓪ Set $\hat{\Psi} = 0$

① Set $S - \hat{\Psi} = V\Lambda V^T$ and $\hat{L} = (\sqrt{\lambda_1}\,\underset{\sim}{v}_1 \cdots \sqrt{\lambda_r}\,\underset{\sim}{v}_r)$

② Set $\hat{\Psi} = \mathrm{diag}(S - \hat{L}\hat{L}^T)$

③ Iterate step ① + ② until convergence (?)

Example: $p=3,\ r=1,\ L=\underset{\sim}{l}=\begin{pmatrix}1\\-1\\1\end{pmatrix}$

$$S=\begin{pmatrix}3&-1&1\\-1&3&-1\\1&-1&3\end{pmatrix}=\begin{pmatrix}1\\-1\\1\end{pmatrix}(1\quad -1\quad 1)+\begin{pmatrix}2&0&0\\0&2&0\\0&0&2\end{pmatrix}$$

① $\lambda_1=5,\ \hat{\underset{\sim}{l}}=\begin{pmatrix}1.29\\-1.29\\1.29\end{pmatrix}\qquad \hat{\psi}=\begin{pmatrix}1.33&&0\\&1.33&\\0&&1.33\end{pmatrix}$

② $\lambda_1=3.67,\ \hat{\underset{\sim}{l}}=\begin{pmatrix}1.11\\-1.11\\1.11\end{pmatrix}\qquad \hat{\psi}=\begin{pmatrix}1.78&&0\\&1.78&\\0&&1.78\end{pmatrix}$

③ $\lambda_1=3.22,\ \hat{\underset{\sim}{l}}=\begin{pmatrix}1.04\\-1.04\\1.04\end{pmatrix}\qquad \hat{\psi}=\begin{pmatrix}1.93&&0\\&1.93&\\0&&1.93\end{pmatrix}$

④ $\lambda_1=3.07,\ \hat{\underset{\sim}{l}}=\begin{pmatrix}1.01\\-1.01\\1.01\end{pmatrix}\qquad \hat{\psi}=\begin{pmatrix}1.98&&0\\&1.98&\\0&&1.98\end{pmatrix}$

⑤ $\lambda_1=3.02\quad \hat{\underset{\sim}{l}}=\begin{pmatrix}1\\-1\\1\end{pmatrix}\qquad \hat{\psi}=\begin{pmatrix}2&&0\\&2&\\0&&2\end{pmatrix}$

Notes:

① This example is not typical:

- S can be exactly expressed as $\hat{\underset{\sim}{l}}\hat{\underset{\sim}{l}}^T+\psi$
- typically S has a much more complicated form
- $r=1$ greatly facilitates convergence

② For $r\geq 2$, loadings matrix is not uniquely determined.

If Q is an $r\times r$ diagonal matrix $(QQ^T=I)$

then $Cov(\underset{\sim}{X})=LL^T+\psi=\underbrace{LQ}_{L^*}(LQ)^T+\psi$

$\underset{\sim}{X}=\underset{\sim}{\mu}+L^*\underset{\sim}{F}^*+\underset{\sim}{\varepsilon}$ where $\underset{\sim}{F}^*=Q^T\underset{\sim}{F}$

What does this mean?

- in our model, we can only uniquely identify LL^T

Does this pose problems for the factor model?

Opportunity!

If $L^*=LQ=(\underset{\sim}{l}_1^*,\ \cdots\ \underset{\sim}{l}_r^*)$, we can try to choose Q to make the loadings $\underset{\sim}{l}_1^*,\cdots,\underset{\sim}{l}_r^*$ more interpretable

How to define interpretable?

- many loadings $=0$ (factors involve only subsets of variables)
- mathematical criterion (or criteria)