L19: Factor models

1. Factor models

   (1) Definitions

   For random vector $X \in R^p$, $X = \mu + LF + \epsilon$ is a factor model where

   (i) Factor $F \sim (0, I_q)$ and random error $\epsilon \sim (0, \Psi)$ with $\Psi = \text{diag}(\psi_1, ..., \psi_p)$ are uncorrelated.

   (ii) Loading matrix $L \in R^{p\times q}$ is non-random.

   (iii) "=" means the parameters (mean vector and covaraince matrix) of two sides are equal.

   Comments: $q \le p$ so that F is simpler than X.

   $X_i = l_{i1}F_1 + \cdots + l_{iq}F_q$, i.e., the ith row of L loads the factor F to X_i.

   l_{ij} loads F_j to X_i.

   (2) Essence of the model: $\text{Cov}(X) = LL' + \Psi$

   Model$\Longrightarrow$ $\text{Cov}(X) = LL' + \Psi$. If $\text{Cov}(X) = LL' + \Psi$ by defining F, ϵ and L, one has the model.

   (3) On the existence and uniqueness

   For $X \sim (\mu, \Sigma)$ a trivial factor model with $q = p$, $L = \Sigma^{1/2}$ and $\Psi = 0$ exists. But with specified $q < p$, factor model may not exist (Ex1). With fixed q factor model may not be unique (Ex2).

   Ex1: Ex9.2 on p486. $X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$ has $\text{Cov}(X) = \Sigma = \begin{pmatrix} 1 & 0.9 & 0.7 \\ 0.9 & 1 & 0.4 \\ 0.7 & 0.4 & 1 \end{pmatrix}$. With $q = 1$ factor model $X - \mu = LF + \epsilon$ does not exist.

   Proof. If the model exists, then $\Sigma = LL' + \Psi$. So $\begin{pmatrix} 1 & 0.9 & 0.7 \\ 0.9 & 1 & 0.4 \\ 0.7 & 0.4 & 1 \end{pmatrix} = \begin{pmatrix} l_1^2 + \psi_1 & l_1 l_2 & l_1 l_3 \\ l_2 l_1 & l_2^2 + \psi_2 & l_2 l_3 \\ l_3 l_1 & l_3 l_2 & l_3^2 + \psi_3 \end{pmatrix}$.

   $\psi_1 = \sigma_1^2 - l_1^2 = 1 - \frac{0.9\times 0.7\times 0.4}{0.4\times 04} = -\frac{23}{40} < 0$ that contradicts with $\psi_1 = \text{var}(\epsilon_1) \ge 0$.

   Ex2: For $\Sigma = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ with $L = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix}$ and $\Psi = \text{diag}(0, 0, 1)$, $\Sigma = LL' + \Psi$.

   Let $L_* = \begin{pmatrix} 0 & 0 \\ \sqrt{2} & 0 \\ 0 & 1 \end{pmatrix}$ and $\Psi_* = \text{diag}(3, 0, 0)$. Then $\Sigma = L_* L_*' + \Psi_*$.

   So for $X \in R^3$ with $\text{Cov}(X) = \Sigma$, the factor model with $q = 2$ is not unique.

2. Factor analysis

   (1) Analysis table

X_i	$\text{var}(X_i)$	F_1	$\cdots$	F_q	h_i^2	ψ_i
X_1	σ_1^2	l_{11}^2	$\cdots$	l_{1q}^2	h_1^2	ψ_1
X_2	σ_2^2	l_{21}^2	$\cdots$	l_{2q}^2	h_2^2	ψ_2
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
X_p	σ_p^2	l_{p1}^2	$\cdots$	l_{pq}^2	h_p^2	ψ_p
	$\sum_i \sigma_i^2$	f_1^2	$\cdots$	f_q^2	$\text{tr}(LL')$	$\text{tr}(\Psi)$

   (2) Communalities

   $X - \mu = LF + \epsilon \Longrightarrow X_i - \mu_i = l_{i1}F_1 + \cdots + l_{iq}F_q + \psi_i \Longrightarrow \text{var}(X_i) = l_{i1}^2 + \cdots + l_{iq}^2 + \psi_i = h_i^2 + \psi_i$

   $h_i^2 = l_{i1}^2 + \cdots + l_{iq}^2$, called the ith communality, is the contribution of F to the variance of X_i.

   So $\sum_i h_i^2 = \text{tr}(LL')$ is the contribution of F to the total variances in X, $\sum_i \sigma_i^2 = \text{tr}(\Sigma)$.

   (3) f_j^2, $j = 1, ..., q$

   l_{ij}^2 in h_i^2 is $\text{var}(l_{ij}F_j)$, the contribution of F_j to $\text{var}(X_i)$.

   Thus $f_j^2 = \sum_i l_{ij}^2$ is the contribution of F_j to the total variances in X, $\sum_i \sigma_i^2$.

   $\sum_j f_j^2 = \text{tr}(L'L)$ is the contribution of F to the total variances in X.

Ex3: p484 Ex9.1

With $\Sigma = \begin{pmatrix} 19 & 30 & 2 & 12 \\ 30 & 57 & 5 & 23 \\ 2 & 5 & 38 & 47 \\ 12 & 23 & 47 & 68 \end{pmatrix}$ and $L = \begin{pmatrix} 4 & 1 \\ 7 & 2 \\ -1 & 6 \\ 1 & 8 \end{pmatrix}$,

X_i	σ_i^2	F_1	F_2	h_i^2	ψ_i
X_1	19	16	1	17	2
X_2	57	49	4	53	4
X_3	38	1	36	37	1
X_4	68	1	64	65	3
	182	67	105	172	10

(a) The proportion of total variances in X explained by LF is $\frac{\text{tr}(LL')}{\text{tr}(\Sigma)} = \frac{172}{182} = 94.51\%$.

(b) The contribution from F_2 to $\text{var}(X_3)$ is 36.

(c) The contribution from F to $\text{var}(X_3)$ is $h_3^2 = 37$.

(d) The contribution from F_2 to the total variances in X is 105.

3. Factor model for standardized X

(1) Standardized X
For $X \sim (\mu, \Sigma)$ with $V = \text{diag}(\Sigma)$, $Z = V^{-1/2}(X - \mu) \sim (0, \rho)$ is standardized X.

(2) Converting factor model for X to that for Z
$X - \mu = LF + \epsilon \Longrightarrow Z = L_z F + \epsilon_z$ where $L_z = V^{-1/2}L$ and $\epsilon_z = V^{-1/2}\epsilon \sim \left(0,\, V^{-1/2}\Psi V^{-1/2}\right)$
with analysis table,

Z_i	ρ_{ii}	F_1	$\cdots$	F_q	h_{zi}^2	ψ_{zi}
Z_1	1	l_{11}^2/σ_1^2	$\cdots$	l_{1q}^2/σ_1^2	h_1^2/σ_1^2	ψ_1/σ_1^2
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Z_p	1	l_{p1}^2/σ_p^2	$\cdots$	l_{pq}^2/σ_p^2	h_p^2/σ_p^2	ψ_p/σ_p^2
	p	f_{z1}^2	$\cdots$	f_{zq}^2	$\text{tr}(L_z L_z')$	$\text{tr}(\Psi_z)$

L20 Estimation in factor models

1. Converting models for X and for Z

   (1) From model for X to that for Z
   If $X-\mu = LF+\epsilon$ has analysis table

X_i	$\text{var}(X_i)$	F_1	$\cdots$	F_q	h_i^2	ψ_i
X_1	σ_1^2	l_{11}^2	$\cdots$	l_{1q}^2	h_1^2	ψ_1
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
X_p	σ_p^2	l_{p1}^2	$\cdots$	l_{pq}^2	h_p^2	ψ_p
	$\sum_i \sigma_i^2$	f_1^2	$\cdots$	f_q^2	$\text{tr}(LL')$	$\text{tr}(\Psi)$

   then $Z = V^{-1/2}(X-\mu) = (V^{-1/2}L)F + (V^{-1/2}\epsilon) = L_zF+\epsilon_z$ has analysis table

Z_i	ρ_{ii}	F_1	$\cdots$	F_q	h_{zi}^2	ψ_{zi}
Z_1	1	l_{11}^2/σ_1^2	$\cdots$	l_{1q}^2/σ_1^2	h_1^2/σ_1^2	ψ_1/σ_1^2
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Z_p	1	l_{p1}^2/σ_p^2	$\cdots$	l_{pq}^2/σ_p^2	h_p^2/σ_p^2	ψ_p/σ_p^2
	p	f_{1z}^2	$\cdots$	f_{qz}^2	$\text{tr}(L_zL_z')$	$\text{tr}(\Psi_z)$

   (2) From model for Z to that for X
   If $Z = V^{-1/2}(X-\mu) = L_zF+\epsilon_z$ has analysis table

Zi	ρ_{ii}	F_1	$\cdots$	F_q	h_{zi}^2	ψ_{zi}
Z_1	1	l_{z11}^2	$\cdots$	l_{z1q}^2	h_{z1}^2	ψ_{z1}
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Z_p	1	l_{zp1}^2	$\cdots$	l_{zpq}^2	h_{zp}^2	ψ_{zp}
	p	f_{1z}^2	$\cdots$	f_{qz}^2	$\text{tr}(L_zL_z')$	$\text{tr}(\Psi_z)$

   then $X-\mu = V^{1/2}Z = (V^{1/2}L_z)F + (V^{1/2}\epsilon_z) = LF+\epsilon$ has analysis table

X_i	σ_i^2	F_1	$\cdots$	F_q	h_{zi}^2	ψ_{zi}
X_1	$1\cdot\sigma_1^2$	$l_{z11}^2\cdot\sigma_1^2$	$\cdots$	$l_{z1q}^2\cdot\sigma_1^2$	$h_{z1}^2\cdot\sigma_1^2$	$\psi_{z1}\cdot\sigma_1^2$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
X_p	$1\cdot\sigma_p^2$	$l_{xp1}^2\cdot\sigma_p^2$	$\cdots$	$l_{zpq}^2\cdot\sigma_p^2$	$h_{zp}^2\cdot\sigma_p^2$	$\psi_{zp}\cdot\sigma_p^2$
	$\text{tr}(\Sigma)$	f_1^2	$\cdots$	f_q^2	$\text{tr}(LL')$	$\text{tr}(\Psi)$

   Besides items in the table for Z we need σ_i^2 to get the table for X.

2. Estimators in $X-\mu = LF+\epsilon$

   (1) Estimated σ_i^2
   From a sample from X we obtain sample covariance matrix S. σ_i^2 is estimated by $\widehat{\sigma}_i^2 = s_i^2$.

   (2) Estimated L
   With specified q. let $S = P\Lambda P' = (P_I,\, P_{II})\begin{pmatrix}\Lambda_I & 0\\ 0 & \Lambda_{II}\end{pmatrix}(P_I,\, P_{II})' = P_I\Lambda_I P_I' + P_{II}\Lambda_{II}P_{II}$ where $\Lambda = \text{diag}(\lambda_1, .., \lambda_p)$, $\lambda_1 \geq \cdots \geq \lambda_p > 0$ and $\Lambda_I \in R^{q\times q}$.
   LL', the main part of Σ, is estimated by $P_I\Lambda P_I'$, the main part of S. So L is estimated by $\widehat{L} = P_I\lambda_I^{1/2} = \left(P_1\sqrt{\lambda_1}, ..., P_q\sqrt{\lambda_q}\right)$.

   (3) Estimated ψ_i: $\quad \psi_i = \sigma_i^2 - h_i^2$. So $\widehat{\psi}_i = s_i^2 - \widehat{h}_i^2$.

   (4) Estimated Σ: $\quad \Sigma$ is estimated by $\widehat{\Sigma} = \widehat{L}\widehat{L}' + \widehat{\Psi}$.

   (5) Analysis table
   $\widehat{L} = P_I\Lambda_I^{1/2} = (\sqrt{\lambda_1}P_1, ..., \sqrt{\lambda_q}P_q)$.

X_i	$\widehat{\sigma}_i^2$	F_1	$\cdots$	F_q	$\widehat{h}_i^2$	$\widehat{\psi}_i$
X_1	s_i^2	$\widehat{l}_{11}^2$	$\cdots$	$\widehat{l}_{1q}^2$	$\widehat{h}_1^2$	$s_i^2 - \widehat{\psi}_1$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
X_p	s_p^2	$\widehat{l}_{p1}^2$	$\cdots$	$\widehat{l}_{pq}^2$	$\widehat{h}_p^2$	$s_p^2 - \widehat{\psi}_p$
	$\sum_{i=1}^p \lambda_i$	$\widehat{f}_1^2 = \lambda_1$	$\cdots$	$\widehat{f}_q^2 = \lambda_q$	$\sum_{i=1}^q \lambda_i$	$\sum_{i=q+1}^p \lambda_i$

$\widehat{L} = P_I \Lambda_I^{1/2} = (\sqrt{\lambda_1} P_1, ..., \sqrt{\lambda_q} P_q)$, $\widehat{L}_z = \widehat{V}^{-1/2} \widehat{L}$.

Z_i	$\widehat{\rho}_{ii}$	F_1	$\cdots$	F_q	$\widehat{h}_{zi}^2$	$\widehat{\psi}_{zi}$
Z_1	1	$\widehat{l}_{11}^2/s_1^2$	$\cdots$	$\widehat{l}_{1q}^2/s_1^2$	$\widehat{h}_1^2/s_1^2$	$\widehat{\psi}_1/s_1^2$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Z_p	1	$\widehat{l}_{p1}^2/s_p^2$	$\cdots$	$\widehat{l}_{pq}^2/s_p^2$	$\widehat{h}_p^2/s_p^2$	$\widehat{\psi}_p/s_p^2$
	p	$\widehat{f}_{1z}^2$	$\cdots$	$\widehat{f}_{qz}^2$	$\mathrm{tr}(\widehat{L}_z \widehat{L}_z')$	$\mathrm{tr}(\widehat{\Psi}_z)$

3. SAS

(1) SAS code

```
data a;
   infile "D\ex.txt";
   input x1 x2 x3;
proc factor nfactor=2 COV;
   var x1 x2 x3;
   run;
```

(2) SAS output

Eigenvalue of Covariance			
Total $\lambda_1 + \lambda_2 + \lambda_3$		Average $\frac{\lambda_1+\lambda_2+\lambda_3}{3}$	
Eigenvalue	Difference	Proportion	Cumulative
λ_1	$\lambda_1 - \lambda_2$	$\lambda_1/\sum_i \lambda_i$	$\lambda_1/\sum_i \lambda_i$
λ_2	$\lambda_2 - \lambda_3$	$\lambda_2/\sum_i \lambda_i$	$(\lambda_1 + \lambda_2)/\sum_i \lambda_i$
λ_3		$\lambda_3/\sum_i \lambda_i$	1

$\widehat{L}_z = \widehat{V}^{-1/2} \widehat{L}$

Factor Pattern	
Factor 1	Factor 2
$\widehat{l}_{11}/s_1$	$\widehat{l}_{12}/s_1$
$\widehat{l}_{21}/s_2$	$\widehat{l}_{22}/s_2$
$\widehat{l}_{31}/s_3$	$\widehat{l}_{32}/s_3$

Variance explained by each factor		
Factor	Weighted	Unweighted
Factor 1	$\widehat{f}_1^2$	$\widehat{f}_{1z}^2$
Factor 2	$\widehat{f}_2^2$	$\widehat{f}_{2z}^2$

Communality		
Weighted: $\sum_i \widehat{h}_i^2$	Unweighted: $\sum_i \widehat{h}_{zi}^2$	
Variable	Communality	Weight
X_1	$\widehat{h}_{z1}^2 = \widehat{h}_1^2/s_1^2$	s_1^2
X_2	$\widehat{h}_{z2}^2 = \widehat{h}_2^2/s_2^2$	s_2^2
X_3	$\widehat{h}_{z3}^2 = \widehat{h}_3^2/s_3^2$	s_3^2