Orthogonal Factor Model

- **Model:** $X_1-\mu_1 = l_{11}\cdot F_1 + l_{12}F_2 + \cdots + l_{1m}F_m + \varepsilon_1$

$$X_2-\mu_2 = l_{21}F_1 + l_{22}F_2 + \cdots + l_{2m}F_m + \varepsilon_2$$

$$\vdots \qquad\qquad \vdots$$

$$X_p-\mu_p = l_{p1}F_1 + l_{p2}F_2 + \cdots + l_{pm}F_m + \varepsilon_m$$

or

matrix of factor loadings

Matrix notation: $X-\mu = L \times F + \varepsilon$ → specific factors

common factors

Assume that $E(F)=0 \quad cov(F)=E(FF')=I$

$$E(\varepsilon)=0 \quad cov(\varepsilon)=E(\varepsilon\varepsilon')=\Psi=\begin{pmatrix}\psi_1 & & & \\ & \psi_2 & & \\ & & \ddots & \\ & & & \psi_p\end{pmatrix}$$

$$(X-\mu)(X-\mu)' = (LF+\varepsilon)(LF+\varepsilon)'$$

$$=(LF)(LF)' + \varepsilon(LF)' + (LF)\varepsilon' + \varepsilon\varepsilon'$$

$$\Sigma = cov((X-\mu)(X-\mu)') = LE(FF')L' + E(\varepsilon F')L' + LE(F\varepsilon') + E(\varepsilon\varepsilon')$$

$$= LL' + \Psi$$

- **Covariance structure:**

1. $Cov(X)=LL'+\Psi$. $Var(X_i) = l_{i1}^2 + l_{i2}^2 + \cdots + l_{im}^2 + \psi_i$

$$Cov(X_i,X_k) = l_{i1}l_{k1} + l_{i2}l_{k2} + \cdots + l_{im}l_{km}$$

2. $Cov(X,F)=L \qquad Cov(X_i,F_j)=l_{ij}$

3. $Var(X_i) = \sigma_{ii} = \underbrace{l_{i1}^2 + l_{i2}^2 + \cdots + l_{im}^2}_{\text{Communality}} + \underbrace{\psi_i}_{\text{specific variance}}$

- **Ambiguity:** $X-\mu = LF+\varepsilon = LTT'F+\varepsilon = L^*F^*+\varepsilon$

where $L^*=LT$ and $F^*=T'F$

since $E(F^*)=T'E(F)=0 \quad Cov(F^*)=T'cov(F)T = T'T = I$

$$\Sigma = LL'+\Psi = LTT'L'+\Psi = (L^*)(L^*)'+\Psi$$

Principal Component Method

- $\Sigma = \lambda_1 e_1 e_1' + \lambda_2 e_2 e_2' + \cdots + \lambda_p e_p e_p'$

$$= \left(\sqrt{\lambda_1} e_1 \,\vdots\, \sqrt{\lambda_2} e_2 \,\vdots\, \cdots \,\vdots\, \sqrt{\lambda_p} e_p\right) \begin{pmatrix} \sqrt{\lambda_1} e_1' \\ \hline \sqrt{\lambda_2} e_2' \\ \vdots \\ \hline \sqrt{\lambda_p} e_p' \end{pmatrix}$$

$$\approx \underbrace{\left(\sqrt{\lambda_1} e_1 \,\vdots\, \sqrt{\lambda_2} e_2 \,\vdots\, \cdots \,\vdots\, \sqrt{\lambda_m} e_m\right)}_{\tilde{L}} \begin{pmatrix} \sqrt{\lambda_1} e_1 \\ \hline \sqrt{\lambda_2} e_2 \\ \vdots \\ \hline \sqrt{\lambda_m} e_m \end{pmatrix} + \begin{pmatrix} \psi_1 & & \\ & \psi_2 & \\ & & \ddots \\ & & & \psi_p \end{pmatrix} \rightarrow \tilde{\Psi}$$

where $\psi_i = \sigma_{ii} - \sum_{j=1}^{m} l_{ij}^2$

communalities are estimated as $\tilde{h}_i^2 = \tilde{l}_{i1}^2 + \tilde{l}_{i2}^2 + \cdots + \tilde{l}_{im}^2$

- Remarks: if X_i are not commensurate, it's desirable to work with standardized variables. $Z = \left(\frac{X_1 - \mu_1}{\sqrt{S_{11}}}, \frac{X_2 - \mu_2}{\sqrt{S_{22}}}, \cdots, \frac{X_p -}{\sqrt{S_{pp}}}\right)$

- Exercise 9.5 Sum of square entries of $(S - (\tilde{L}\tilde{L}' + \tilde{\Psi})) \le \tilde{\lambda}_{m+1} + \cdots + \tilde{\lambda}_p$

proof: $\text{diag}\{S - (\tilde{L}\tilde{L}' + \tilde{\Psi})\} = \left(S_{ii} - \sum_{j=1}^{m} l_{ij}^2 + \sum_{j=1}^{m} l_{ij}^2 - S_{ii}\right) = 0$

$$\begin{aligned} \|S - (\tilde{L}\tilde{L}' + \tilde{\Psi})\|_F^2 &= \text{off}(S - (\tilde{L}\tilde{L}' + \tilde{\Psi})) \\ &= \text{off}(S - \tilde{L}\tilde{L}') \\ &= \text{off}(\tilde{\lambda}_{m+1} \tilde{e}_{m+1} \tilde{e}_{m+1}' + \cdots + \tilde{\lambda}_p \tilde{e}_p \tilde{e}_p') \\ &\le \|\tilde{\lambda}_{m+1} \tilde{e}_{m+1} \tilde{e}_{m+1}' + \cdots + \tilde{\lambda}_p \tilde{e}_p \tilde{e}_p'\|_F^2 \\ &= \tilde{\lambda}_{m+1} + \cdots + \tilde{\lambda}_p \end{aligned}$$

Remark: proportion of total sample variance $= \dfrac{\hat{\lambda}_j}{S_{11} + S_{22} + \cdots + S_{pp}}$

$\hat{\lambda}_j = \tilde{l}_{11}^2 + \tilde{l}_{21}^2 + \cdots + \tilde{l}_{p1}^2$

Maximum Likelihood Method

- Likelihood function:

$$L(\mu,\Sigma) = (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{n}{2}} e^{-\frac{1}{2}tr\left(\Sigma^{-1}\left(\sum_{j=1}^{n}(x_j-\bar{x})(x_j-\bar{x})' + n(\bar{x}-\mu)(\bar{x}-\mu)'\right)\right)}$$

$$= (2\pi)^{-\frac{(n-1)p}{2}} |\Sigma|^{-\frac{n-1}{2}} e^{-\frac{1}{2}tr\left(\Sigma^{-1}\left(\sum_{j=1}^{n}(x_j-\bar{x})(x_j-\bar{x})'\right)\right)}$$

$$\times (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} e^{-\frac{n}{2}(\bar{x}-\mu)'\Sigma^{-1}(\bar{x}-\mu)}$$

Let $\Sigma = LL' + \Psi$, $L'\Psi^{-1}L = \Delta$ (uniqueness)

- Result 9.1. $\hat{\mu} = \bar{x}$ maximizes $L(\mu,\Sigma)$

The communalities are $\hat{h}_i^2 = \hat{l}_{i1}^2 + \hat{l}_{i2}^2 + \cdots + \hat{l}_{im}^2$, and so

Proportion of total sample variances due to jth factor $= \dfrac{\hat{l}_{1j}^2 + \hat{l}_{2j}^2 + \cdots + \hat{l}_{pj}^2}{s_{11} + s_{22} + \cdots + s_{pp}}$

Factor Scores

- factor scores: estimate of value f_j attained by F_j

- Weighted Least Squares: $X - \mu = L \cdot F + \varepsilon$

weighted sum of square errors: $\sum_{i=1}^{p} \frac{\varepsilon_i^2}{\psi_i} = \varepsilon' \psi^{-1} \varepsilon = (x - \mu - Lf)' \psi^{-1} (x - \mu - Lf)$

$$\frac{\partial J}{\partial f_j} = L \psi^{-1} (x - \mu - Lf_j) = 0 \qquad \hat{f}_j = (L' \psi^{-1} L)^{-1} L' \psi^{-1} (x - \mu)$$

$$\approx (\hat{L}' \hat{\psi}^{-1} \hat{L})^{-1} \hat{L}' \hat{\psi}^{-1} (x - \bar{x})$$

$$\min \sum_{i=1}^{n} \| X_i - LF_i \|_2^2 = \| X - LF \|_F^2 \qquad \text{assume } \psi_1 = \psi_2 = \cdots = \psi_n$$

$$\hat{f}_j = (L' L)^{-1} L' (x_j - \bar{x})$$