Partitioning the Variance of the Data Vectors

Note: $X_i = \lambda_{i1} f_1 + \lambda_{i2} f_2 + \cdots + \lambda_{ik} f_k + u_i$

Since the f_j's are assumed to be uncorrelated with each other and with the u_i's:

$$\text{var}(X_i) = \lambda_{i1}^2 \underbrace{\text{var}(f_1)}_{1} + \lambda_{i2}^2 \underbrace{\text{var}(f_2)}_{1} + \cdots + \lambda_{ik}^2 \underbrace{\text{var}(f_k)}_{1} + \text{var}(u_i)$$

$$= \sum_{j=1}^{k} \lambda_{ij}^2 + \text{var}(u_i)$$

$$= h_i^2 + \psi_i, \text{ where } h_i^2 = \sum_{j=1}^{k} \lambda_{ij}^2 \text{ and } \psi_i = \text{var}(u_i)$$

The *communality* h_i^2 is the variability in manifest variable x_i shared with the other variables (via the factors) and ψ_i is the specific variance, not shared with the other variables.

The (i,j) element of Σ is σ_{ij}.

Covariance of the Data Vectors

$$
\begin{aligned}
\sigma_{ij} &= \operatorname{cov}(X_i, X_j) \\
&= \operatorname{cov}(\lambda_{i1} f_1, \lambda_{j1} f_1) + \operatorname{cov}(\lambda_{i2} f_2, \lambda_{j2} f_2) + \cdots + \operatorname{cov}(\lambda_{ik} f_k, \lambda_{jk} f_k) \\
&\quad + 0 + 0 + \cdots + 0
\end{aligned}
$$

$\Big[$ since $\operatorname{cov}(f_\ell, f_m) = 0$ for all $\ell \neq m$

$\operatorname{cov}(u_i, f_\ell) = 0$ for all i, ℓ

$\operatorname{cov}(u_i, u_j) = 0$ for all $i \neq j$ $\Big]$

$$
= \lambda_{i1}\lambda_{j1} \underbrace{\operatorname{var}(f_1)}_{1} + \lambda_{i2}\lambda_{j2} \underbrace{\operatorname{var}(f_2)}_{1} + \cdots + \lambda_{ik}\lambda_{jk} \underbrace{\operatorname{var}(f_k)}_{1}
$$

$= \sum_{\ell=1}^{k} \lambda_{i\ell}\lambda_{j\ell}$. And recall $\sigma_{ii} = \operatorname{var}(X_i) = h_i^2 + \psi_i = \sum_{j=1}^{k} \lambda_{ij}^2 + \psi_i$.

Hence the population covariance matrix $\boldsymbol{\Sigma}$ for $(x_1, x_2, \ldots, x_q)$ is $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}$, where $\boldsymbol{\Psi} = diag(\psi_i)$.