Final exam: Tuesday April 19 7–9 pm
Coverage: Chap. 1–6 of text + classification
- 2 (5.5"×11") aid sheets (both sides) and a calculator are allowed.

Office hrs: Next week TBA
Mon April 18 1–3pm

Structured multivariate Models

So far: $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, C)$

Graphical model: $K = C^{-1} \longrightarrow$ assume that K is sparse — many 0 entries

PCA: C can be approximated by a lower rank matrix.

Factor Analysis: $C = \underset{p\times r}{L}\, L^T + \underset{\text{diagonal}}{\psi}$

More generally: Apply multivariate thinking to univariate models

Example: Longitudinal/panel data
n subjects followed over time
- measure some (univariate) response at time pts $t_1, \cdots, t_p$

$$X_{ij} = \text{response for subject } i \text{ at time } t_j$$
$$= f(t_j) + \text{error}$$

$\underset{\sim}{X_i} = \begin{pmatrix} X_{i1} \\ \vdots \\ X_{ip} \end{pmatrix} \longrightarrow$ model dependence of observations within each subjects.

Example: Paired difference design
- 2 treatments
- n subjects

Each subject receives both treatments

Model: $\underset{\text{subject}}{X_{i}}{}_{\underset{\text{treatment}}{j}} = \underset{\text{subject effect (random)}}{A_i} + \underset{\text{treatment effect (fixed)}}{\mu_j} + \underset{\text{noise}}{\varepsilon_{ij}}$

Look at $\Delta_i = X_{i1} - X_{i2} = \underbrace{\mu_1 - \mu_2}_{\text{treatment difference}} + \underbrace{\varepsilon_{i1} - \varepsilon_{i2}}_{\text{noise}}$

For example, the test $H_0: \mu_1 = \mu_2$ or obtain a CI for $\mu_1 - \mu_2$, we use $\Delta_1, \cdots, \Delta_n$

Back to original data

$$\underset{\sim}{X_i} = \begin{pmatrix} X_{i1} \\ X_{i2} \end{pmatrix} = A_i \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \end{pmatrix}$$

Assume $A_i \sim N(0, \sigma_A^2)$ (independent)

$\underset{\sim}{\varepsilon}_i \sim N_2(\underset{\sim}{0}, \sigma_\varepsilon^2 I)$ (independent)

A_i independent of $\underset{\sim}{\varepsilon}_i$

$\Rightarrow \underset{\sim}{X}_i \sim N_2(\underset{\sim}{\mu}, C)$ where $C = \mathrm{Cov}(A_i \begin{pmatrix}1\\1\end{pmatrix}, \underset{\sim}{\varepsilon}_i) = \mathrm{Cov}(A_i \begin{pmatrix}1\\1\end{pmatrix}) + \mathrm{Cov}(\underset{\sim}{\varepsilon}_i)$

$$C = \sigma_A^2 \begin{pmatrix}1\\1\end{pmatrix}(1 \quad 1) + \sigma_\varepsilon^2 \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

$$= \begin{pmatrix}\sigma_A^2+\sigma_\varepsilon^2 & \sigma_A^2 \\ \sigma_A^2 & \sigma_A^2+\sigma_\varepsilon^2\end{pmatrix} \qquad \mathrm{Corr}(X_{i1}, X_{i2}) = \frac{\sigma_A^2}{\sigma_A^2+\sigma_\varepsilon^2}$$

Note: C has a single factor form $C = \underset{\sim}{l}\,\underset{\sim}{l}^T + \psi$

A step further: Repeated measures design

$\left.\begin{array}{l}k \text{ treatments}\\ n \text{ subjects}\end{array}\right\}$ each subject receives all k treatments

— "treatment" used loosely here — for example, we could have a response measured at different points in time.

$$\underset{\sim}{X}_i = \begin{pmatrix}X_{i1}\\ \vdots \\ X_{ik}\end{pmatrix} = A_i \begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} + \begin{pmatrix}\mu_1\\ \vdots \\ \mu_k\end{pmatrix} + \begin{pmatrix}\varepsilon_{i1}\\ \vdots\\ \varepsilon_{ik}\end{pmatrix}$$

$$= A_i \begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} + \underset{\sim}{\mu} + \underset{\sim}{\varepsilon}_i$$

Assumption: $A_i \sim N(0, \sigma_A^2)$ (independent)

$\underset{\sim}{\varepsilon}_i \sim N_k(\underset{\sim}{0}, \sigma_\varepsilon^2 I)$ (indep)

$A_i + \underset{\sim}{\varepsilon}_i$ independent

$$\underset{\sim}{X}_i \sim N(\underset{\sim}{\mu}, C) \text{ where } C = \begin{pmatrix}\sigma_A^2+\sigma_\varepsilon^2 & \sigma_A^2 & \cdots & \sigma_A^2 \\ \vdots & \ddots & & \vdots \\ \sigma_A^2 & & \cdots & \sigma_A^2+\sigma_\varepsilon^2\end{pmatrix} \leftarrow$$

C is defined by 2 parameters $\sigma_A^2, \sigma_\varepsilon^2$

Estimation: $\mu_1, \cdots, \mu_k, \sigma_\varepsilon^2, \sigma_A^2$

① Likelihood estimation of all parameters

— full likelihood

dependent on $\sigma_A^2, \sigma_\varepsilon^2$

$$L(\mu_1, \cdots, \mu_k, \sigma_A^2, \sigma_\varepsilon^2) = \prod_{i=1}^n \left\{ \frac{1}{(2\pi)^{k/2}|C|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\underset{\sim}{X}_i - \underset{\sim}{\mu})^T C^{-1} (\underset{\sim}{X}_i - \underset{\sim}{\mu})\right]\right\}$$

Restricted ML estimation (REML)

- transform data to eliminate $\mu_1, \cdots, \mu_k \rightarrow$ get estimate of $\sigma_A^2, \sigma_\varepsilon^2$
- Then estimate $\underset{\sim}{\mu}_1, \cdots, \underset{\sim}{\mu}_k$ by generalized least squares.

② Transform data to eliminate random effects $A_1, \cdots, A_n$

$$\underset{\sim}{X}_i = A_i \begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} + \underset{\sim}{\mu} + \underset{\sim}{\varepsilon}_i$$

Multiply by some $(k-1)\times k$ matrix B such that $B\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} = \underset{\sim}{0}$

$$\underbrace{B\underline{X_i}}\ = \underbrace{B\,\underline{\mu}} + \underbrace{B\,\underline{\varepsilon_i}}$$

$$X_i^* = \underline{\mu}^* + \underline{\varepsilon_i^*} \quad N_{k-1}(\underline{0},\ \sigma_\varepsilon^2 BB^T)$$

<u>Examples</u>

$$B = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & & 0 \\ 0 & 0 & 1 & -1 & \\ 0 & \cdots & & 1 & -1 \end{pmatrix} \qquad B\underline{\mu} = \begin{pmatrix} \mu_1 - \mu_2 \\ \mu_2 - \mu_3 \\ \vdots \\ \mu_{k-1} - \mu_k \end{pmatrix}$$

$$B = \begin{pmatrix} -1 & 1 & 0 & & 0 \\ -1 & 0 & 1 & & \\ \vdots & & & \ddots & \\ -1 & & 0 & & 1 \end{pmatrix} \qquad B\underline{\mu} = \begin{pmatrix} \mu_2 - \mu_1 \\ \vdots \\ \mu_k - \mu_1 \end{pmatrix}$$

<u>Advantage</u>: Eliminate random effect

<u>Disadvantage</u>: Lose information in transformation

$\longrightarrow$ can't estimate $\underline{\mu}$ itself but B instead.