Notations

- Y_{ij} is the response for the ith subject at the jth occasion
- X_{ij} is the predictor(s) at time t_{ij}
- $j = 1, \cdots, n_i,\ i = 1, \cdots, m$
- $\boldsymbol{Y}_i = (Y_{i1}, \cdots, Y_{in_i})$
- $\mathbb{E}(Y_{ij}) = \mu_{ij}$, $\mathbb{E}(\boldsymbol{Y}_i) = \boldsymbol{\mu}_i$
- $\text{var}(\boldsymbol{Y}_i) = \begin{pmatrix} \text{var}(Y_{i1}) & \text{cov}(Y_{i1}, Y_{i2}) & \cdots & \text{cov}(Y_{i1}, Y_{in_i}) \\ & \text{var}(Y_{i2}) & \cdots & \text{cov}(Y_{i2}, Y_{in_i}) \\ & & \cdots & \\ & & \cdots & \text{var}(Y_{in_i}) \end{pmatrix}$

Approaches to LDA (Overview)

For continuous responses:

- Marginal Models

$$\mathbb{E}(Y_{ij}) = X_{ij}\beta, \ \underline{\mathrm{var}(\boldsymbol{Y}_i) = V_i}$$

- Mixed Effects Models

$$\mathbb{E}(Y_{ij}|\beta_i) = X_{ij}\beta_i, \ \beta_i = \beta + U_i$$

- Transition Models

$$\mathbb{E}(Y_{ij}|Y_{i,j-1}, \cdots, Y_{i1}, X_{ij})$$

Marginal Model for Linear Regression

Consider a simple linear model (e.g., for CD4+ example)

$$Y_{ij} = \beta_0 + \beta_1 t_{ij} + \varepsilon_{ij}$$

- Mean part: $\mathbb{E}(Y_{ij}) = \beta_0 + \beta_1 t_{ij}$
- Variance part: $\text{var}(\boldsymbol{Y}_i) = \text{var}(\boldsymbol{\varepsilon}_i)$
  More often, we focus on the correlation structure

$$\text{corr}(\boldsymbol{Y}_i) = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1n_i} \\ \rho_{21} & 1 & \cdots & \rho_{2n_i} \\ & & \cdots & \\ \rho_{n_i1} & \rho_{n_i2} & \cdots & 1 \end{pmatrix}$$

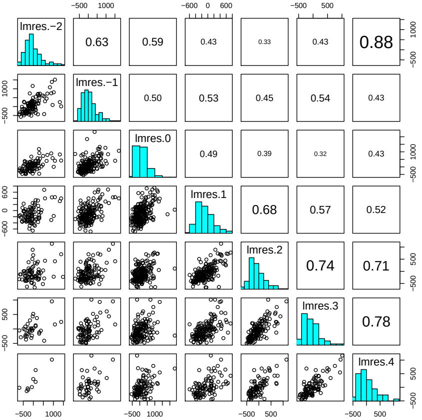

Figure: Correlations of residuals for repeated measurements of CD4+ counts

About the Correlation

Empirical observations about the nature of the correlation among repeated measures:

- correlations are usually positive
- decrease with increasing time separation
- rarely approach zero
- approach one if a pair of repeated measures are taken very closely in time

Modeling Covariance Structure

Assume a **balanced design**, where number and timing of the repeated measurements are the same for all individuals. We have $t_{ij} = t_j$, $j = 1, \cdots, n$. The covariance of the response variable $\boldsymbol{Y}$ (length $m \times n = N$) is:

$$\text{cov}(\boldsymbol{Y}) = \begin{pmatrix} \Sigma_1 & 0 & \cdots & 0 \\ 0 & \Sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_m \end{pmatrix}.$$

We further assume the covariance matrices for different subjects are the same ($\Sigma = \Sigma_i$).

Covariance Pattern Models

Compound Symmetry (or Exchangeable)

Assume variance is constant across visits (say σ^2), and correlation between any two visits are constant (say ρ).

$$\mathrm{cov}(\boldsymbol{Y}_i) = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \rho & \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \cdots & 1 \end{pmatrix}$$

- Parsimonious: two parameters regardless of number of visits per subject
- Strong assumptions about variance and correlation are usually not valid for longitudinal data

Toeplitz

Assume variance is constant across visits and $\text{corr}(Y_{ij}, Y_{i,j+k}) = \rho_k$.

$$\text{cov}(\boldsymbol{Y}_i) = \sigma^2 \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{n-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \cdots & 1 \end{pmatrix}$$

- Assume correlation among responses at adjacent measurements is constant.
- Only suitable for measurements made at equal intervals of time.
- Toeplitz covariance has n parameters (1 variance and $n-1$ correlation parameters)

Autoregressive

A special case of Toeplitz with $\mathrm{corr}(Y_{ij}, Y_{i,j+k}) = \rho^k$.

$$\mathrm{cov}(\boldsymbol{Y}_i) = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{pmatrix}$$

- Only 2 parameters, regardless of the number of visits
- Only suitable for measurements made at equal intervals of time

- **Banded**
  - Assume correlation is 0 beyond some specified interval.
  - Can be combined with the previous patterns.
  - This is a very strong assumption about how quickly the correlation decays to 0 with increasing time separation
- **Exponential**
  - A generalization of autoregressive pattern
  - Suitable for unevenly spaced measurements
  - Let $\{t_{i1}, \cdots, t_{in}\}$ denote the observation times for the ith individual. The correlation is $\mathrm{corr}(Y_{ij}, Y_{ik}) = \rho^{|t_{ij} - t_{ik}|}$.
  - Correlation decreases exponentially with the time separations between them.

General Linear Model

Assume all subjects are independent. Consider the general linear model:

$$\boldsymbol{Y}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{\varepsilon}_i$$

where $\boldsymbol{\varepsilon}_i \sim MVN(\mathbf{0}, \Sigma_i)$. We further assume Σ_is are the same for different subjects, and have a parametric or unstructured pattern.

- Use OLS to estimate $\boldsymbol{\beta}$
- Use weighted least squares (WLS)

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{W} \boldsymbol{Y}$$

- If $\boldsymbol{W} = \text{blkdiag}(\boldsymbol{\Sigma})^{-1}$, more efficient than OLS
- Maximum likelihood estimate

Restricted Maximum Likelihood

- Assume the covariance is unknown. One can use maximum likelihood approach to estimate $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$.
- However, $\hat{\boldsymbol{\Sigma}}_{MLE}$ is typically biased
- For example, in LM, $\widehat{\sigma^2} = RSS/n$ is based
- Restricted Maximum Likelihood (REML) is used to correct the bias of MLE
- Strictly speaking, REML is for the variance components. Once estimated, it is plugged back to the WLS estimator to get the "REML estimate" of $\boldsymbol{\beta}$.

Additional Topics

- How to select the most appropriate covariance pattern?
- How to account for mis-specification in inference?
- How to perform REML?
- More in the course of Longitudinal Data Analysis

Example

Opposites Naming:

- 35 people completed an inventory that assesses their performance on a timed cognitive task "opposites naming."
- Each person completed 4 tests, along with a baseline assessment of cognitive skill.
- Question: whether opposites-naming skill increases with time; whether the skill increases more rapidly among individuals with stronger cognitive skills.