Covariance Structure
Approach to Within-cases

Using SAS proc mixed
Advantages

• Straightforward: It’s familiar univariate regression
• Just MSE is different, because of correlated observations
• Nicer treatment of missing data (valid if missing at random)
• Can have time-varying covariates
• Flexible modeling of non-independence within cases
• Can accommodate more factor levels than cases (with assumptions)
Usual covariance matrix of $Y_1, \ldots, Y_n$

\[
\begin{pmatrix}
\sigma^2 & 0 & 0 & \ldots & 0 & 0 \\
0 & \sigma^2 & 0 & \ldots & 0 & 0 \\
0 & 0 & \sigma^2 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & 0 \\
0 & 0 & 0 & \ldots & \sigma^2 & 0 \\
0 & 0 & 0 & \ldots & 0 & \sigma^2 \\
\end{pmatrix}
\]
In the covariance structure approach

• There are $n$ “subjects.”
• There are $k$ (“repeated”) measurements per subject
• There are $nk$ cases: $n$ blocks of $k$ rows
• Data are multivariate normal (dimension $nk$)
• Familiar regression model for the vector of means
• Special structure for the variance-covariance matrix: not just a diagonal matrix with $\sigma^2$ on the main diagonal
Structure of the variance-covariance matrix

- Covariance matrix of the data has a **block diagonal** structure: \( n \times n \) matrix of little \( k \times k \) variance-covariance matrices (partitioned matrix)
- Off diagonal matrices are all zeros -- no correlation between data from different cases
- Matrices on the main diagonal are all the same (equal variance assumption)
Block Diagonal Covariance Matrix

\[
\begin{bmatrix}
\Sigma & 0 & \cdots & 0 & 0 \\
0 & \Sigma & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & 0 \\
0 & 0 & \cdots & \Sigma & 0 \\
0 & 0 & \cdots & 0 & \Sigma \\
\end{bmatrix}
\]

\(\Sigma\) is the matrix of variances and covariances of the data from a single subject.
\[ \Sigma \text{ may have different structures} \]

- May be unknown

\[ \Sigma = \begin{bmatrix}
\sigma_1^2 & \sigma_{1,2} & \sigma_{1,3} & \sigma_{1,4} \\
\sigma_{2,1} & \sigma_2^2 & \sigma_{2,3} & \sigma_{2,4} \\
\sigma_{3,1} & \sigma_{3,2} & \sigma_3^2 & \sigma_{3,4} \\
\sigma_{4,1} & \sigma_{4,2} & \sigma_{4,3} & \sigma_4^2
\end{bmatrix} = \begin{bmatrix}
\sigma_1^2 & \sigma_{1,2} & \sigma_{1,3} & \sigma_{1,4} \\
\sigma_2^2 & \sigma_{2,3} & \sigma_{2,4} & \\
\sigma_3^2 & \sigma_{3,4} & \\
\sigma_4^2 & 
\end{bmatrix} \]

- May be something else
Compound Symmetry

• Why are data from the same case correlated?
• Because each case makes its own contribution -- add a (random) quantity that is different for each case
• Some monkeys are just better at solving puzzles
Model for Case $i$ in Within-Cases Condition $j$

$Y_{ij} = \beta_0 + \beta_1 x_{ij1} + \beta_2 x_{ij2} + \cdots + \beta_{p-1} x_{ij(p-1)} + \delta_i + \epsilon_{ij}$, where

- Everything is independent for $i = 1, \ldots, n$
- $\epsilon_{ij}$ and $\delta_i$ are independent
- $\epsilon_{ij} \sim N(0, \sigma^2)$ and $\delta_i \sim N(0, \sigma_\delta^2)$
- $\epsilon_{ij}$ and $\epsilon_{i\ell}$ are independent for $j \neq \ell$

This implies

- $Var(Y_{ij}) = \sigma_\delta^2 + \sigma^2$
- $Cov(Y_{ij}, Y_{i\ell}) = \sigma_\delta^2$ for $j \neq \ell$
Compound Symmetry

\[ \Sigma = \begin{bmatrix}
\sigma^2 + \sigma_1 & \sigma_1 & \sigma_1 & \sigma_1 \\
\sigma_1 & \sigma^2 + \sigma_1 & \sigma_1 & \sigma_1 \\
\sigma_1 & \sigma_1 & \sigma^2 + \sigma_1 & \sigma_1 \\
\sigma_1 & \sigma_1 & \sigma_1 & \sigma^2 + \sigma_1 \\
\end{bmatrix} \]

Fewer parameters to estimate
Available covariance structures include

- Unknown: type=un
- Compound symmetry: type=cs
- Variance components: type=vc
- First-order autoregressive: type=ar(1)
- Spatial autocorrelation: covariance is a function of Euclidian distance
- Factor analysis
- Many others
Why not always assume covariance structure unknown?

- No reason why not, if you have enough data.
- When number of unknown parameters is large relative to sample size, variances of estimators are large => confidence intervals wide, tests weak.
- In some studies, there can be more treatment conditions than cases, and unique estimates of parameters don’t even exist.
- There is always a tradeoff between assumptions and amount of data.
First-order autoregressive time series

\[ \Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix} \]

- Usually much bigger matrix
- Could have a handful of cases measured at hundreds of time points
- Or even just one “case,” say a company
Eating Norm Study

- Two free meals at the psych lab
- One with another student, one alone
- But it’s not really another student. It’s a “confederate.”
- Confederate either eats a lot or a little.
- Dine with the confederate first, or second.
- DV is how much you eat. They weigh it.
- Covariates: How long since you ate, and how hungry you are. It can be different each time.
Variables

- Amount subject eats: DV
- Amount confederate eats (between)
- Eat alone or with confederate (within)
- Eat with confederate first, or second (between)
- Reported time since ate (covariate)
- Reported hunger (covariate)

- Notice these are *time-varying covariates*
Just can’t do it with the multivariate approach

\[
\mu = \begin{bmatrix}
\mu_1 \\
\mu_2 \\
\vdots \\
\mu_k
\end{bmatrix} = \begin{bmatrix}
E[Y_1|X=x] \\
E[Y_2|X=x] \\
\vdots \\
E[Y_k|X=x]
\end{bmatrix} = \begin{bmatrix}
\beta_{0,1} + \beta_{1,1}x_1 + \cdots + \beta_{p-1,1}x_{p-1} \\
\beta_{0,2} + \beta_{1,2}x_1 + \cdots + \beta_{p-1,2}x_{p-1} \\
\vdots \\
\beta_{0,k} + \beta_{1,k}x_1 + \cdots + \beta_{p-1,k}x_{p-1}
\end{bmatrix}
\]