We’ve already seen that for longitudinal and cluster correlated data:

\[
\text{Var}[Y_i] = \text{Var} \begin{bmatrix}
Y_{i1} \\
Y_{i2} \\
\vdots \\
Y_{iJ}
\end{bmatrix}
\]

so that: 
\[
\text{Cov}[Y_{ij}, Y_{il}] = \sigma_{jl} \quad \forall \ i \\
\text{Var}[Y_{ij}] = \sigma_j^2 \quad \forall \ i
\]

We will now let \( \Sigma_i \) denote \( \text{Var}[Y_i] \).

It follows that:

\[
\text{Var}[\varepsilon_i] = \text{Var} \begin{bmatrix}
\varepsilon_{i1} \\
\varepsilon_{i2} \\
\vdots \\
\varepsilon_{iJ}
\end{bmatrix} = \Sigma_i
\]

Why?
Consider the variance of our model: 
\( Y_i = X_i \alpha + \varepsilon_i \)

Thus:
\[
\varepsilon_i \sim F(0, \Sigma_i) \quad i = 1, 2, \ldots n.
\]

Note (on the previous page) that there are no \( i \) subscripts in the elements of \( \Sigma_i \) anywhere. (This assumption can be relaxed.) What does that mean?
Putting this all together, we can now write down a General Linear Model for cluster \( i \):

\[
Y_i = X_i \alpha + \varepsilon_i \quad i = 1, \ldots, n
\]

\( \varepsilon_i \sim F(0, \Sigma_i) \)

or, written equivalently,

\[
Y_i \sim F(X_i \alpha, \Sigma_i)
\]

for all \( i \).

**Note:** Repeated Measures ANOVA is a special case of a GLM. It is a GLM with a group, time, and group by time structure in \( X \) and with \( \Sigma \) having a compound symmetry structure.

Clearly we now have to specify two things before we can fit any models:

1. a structure for \( \Sigma_i \) (SAS calls it \( R \))
2. a form for \( X_i \alpha \).

What structures can \( \Sigma_i \) take? Let’s look at examples assuming \( J = 4 \) like for the Contact Time Study:

- **INDEPENDENT** with equal variances and zero correlations
  \[
  \Sigma_i = \begin{bmatrix}
  \sigma^2 & 0 & 0 & 0 \\
  0 & \sigma^2 & 0 & 0 \\
  0 & 0 & \sigma^2 & 0 \\
  0 & 0 & 0 & \sigma^2
  \end{bmatrix} = \sigma^2 I
  \]

- **COMPOUND SYMMETRY** [exchangeable, equicorrelation] with equal variances and constant correlations for every lag
  \[
  \Sigma_i = \tau^2 \begin{bmatrix}
  1 & \rho & \rho & \rho \\
  \rho & 1 & \rho & \rho \\
  \rho & \rho & 1 & \rho \\
  \rho & \rho & \rho & 1
  \end{bmatrix}
  \]

- **FIRST ORDER AUTOREGRESSIVE** [ar(1)] with equal variances, constant correlations within lag, and decaying correlations across lags
  \[
  \Sigma_i = \tau^2 \begin{bmatrix}
  1 & \rho^2 & \rho^3 \\
  \rho & 1 & \rho \\
  \rho^2 & \rho & 1 \\
  \rho^3 & \rho^2 & \rho
  \end{bmatrix}
  \]

- **BANDED MAIN DIAGONAL** (more general than Weiss considers) with one band:
  \[
  \Sigma_i = \begin{bmatrix}
  \sigma_{11} & 0 & 0 & 0 \\
  \sigma_{21} & \sigma_{22} & 0 & 0 \\
  \sigma_{31} & \sigma_{32} & \sigma_{33} & 0 \\
  \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44}
  \end{bmatrix}
  \]

  with two bands:
  \[
  \Sigma_i = \begin{bmatrix}
  \sigma_{11} & 0 & 0 & 0 \\
  \sigma_{21} & \sigma_{22} & 0 & 0 \\
  \sigma_{31} & \sigma_{32} & \sigma_{33} & 0 \\
  \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44}
  \end{bmatrix}
  \]

  both with unequal variances and with covariances for a limited number of lags only.
TOEPLITZ with equal variances and constant correlations within lag

\[ \Sigma_i = \tau^2 \begin{bmatrix} 1 & \rho_1 & 1 \\ \rho_2 & \rho_1 & 1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix} \]

UNSTRUCTURED

\[ \Sigma_i = \begin{bmatrix} \sigma_{11} & \sigma_{21} & \sigma_{31} & \sigma_{41} \\ \sigma_{21} & \sigma_{22} & \sigma_{32} & \sigma_{42} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{34} & \sigma_{44} \end{bmatrix} \]

How do we choose a structure?

- consider several structures which seem reasonable according to the science
- examine covariance EDA plots
- fit a model separately for each structure and compare.

Always keep in mind, which structures make sense for longitudinal data:

- with equally spaced time points?
- with unequally spaced time points?

and which structures make sense for non-longitudinal, cluster correlated data?

Aside: How do we get from the covariance matrix to the correlation matrix?

Recall the basic formula for calculating a correlation between two random variables \( Z_1 \) and \( Z_2 \):

\[ \text{Corr}[Z_1, Z_2] = \frac{\text{Cov}[Z_1, Z_2]}{\sqrt{\text{Var}[Z_1] \text{Var}[Z_2]}} \]

which implies that

\[ \text{Cov}[Z_1, Z_2] = \sqrt{\text{Var}[Z_1] \text{Corr}[Z_1, Z_2]} \sqrt{\text{Var}[Z_2]} \]

so now we just need to translate this into matrix terms.

\[ \text{Var}[\varepsilon] = \Sigma_i = \begin{bmatrix} \sqrt{\sigma_{11}} & \sqrt{\sigma_{21}} & \sqrt{\sigma_{31}} & \sqrt{\sigma_{41}} \\ \sqrt{\sigma_{21}} & \sqrt{\sigma_{22}} & \sqrt{\sigma_{32}} & \sqrt{\sigma_{42}} \\ \sqrt{\sigma_{31}} & \sqrt{\sigma_{32}} & \sqrt{\sigma_{33}} & \sqrt{\sigma_{34}} \\ \sqrt{\sigma_{41}} & \sqrt{\sigma_{42}} & \sqrt{\sigma_{34}} & \sqrt{\sigma_{44}} \end{bmatrix} \]

standard deviations correlations standard deviations

As long as you have an intuitive understanding of what a correlation means, then don’t worry about these formulas. Just remember that a correlation is a covariance that has been re-scaled by the variances.
How do we next incorporate the idea that clusters are independent of each other? Let’s consider all data for all clusters, all together, with \( \text{Var}[Y_i] = \Sigma_i \):

\[
\text{Var}[Y] = \begin{bmatrix} \text{Var}[Y_1] & \text{Cov}[Y_2, Y_1] & \cdots & \text{Cov}[Y_n, Y_1] \\ \text{Cov}[Y_2, Y_1] & \text{Var}[Y_2] & \cdots & \text{Cov}[Y_n, Y_2] \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}[Y_n, Y_1] & \text{Cov}[Y_n, Y_2] & \cdots & \text{Var}[Y_n] \end{bmatrix}
\]

\[
= \begin{bmatrix} \Sigma_1 & 0 & \cdots & 0 \\ 0 & \Sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_n \end{bmatrix}
\]

\( \varepsilon_i \sim F(0, \Sigma_i) \) for \( i = 1, \ldots, n \)

\( \text{this last matrix is "block-diagonal"} \)

The full model structure is thus:

\[
Y = X\alpha + \varepsilon \quad \text{with}
\]

\[
\varepsilon \sim F \left( \begin{bmatrix} \Sigma_1 & 0 & \cdots & 0 \\ 0 & \Sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_n \end{bmatrix} \right)
\]

or, written equivalently,

\[
Y \sim F \left( \begin{bmatrix} \Sigma_1 & 0 & \cdots & 0 \\ 0 & \Sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_n \end{bmatrix} \right)
\]

\( \Sigma \) is said to have a block diagonal structure with each \( \Sigma_i \) as a 'block' lined up along the diagonal.