1 Repeated Measures General Linear Models

All of the GLMs that we have considered this far had one and only one dependent variable. When a GLM has more than one dependent variable, then the model is said to be *multivariate*. Hence, one encounters the multivariate t-test, multivariate regression, multivariate analysis of variance or MANOVA, and multivariate analysis of covariance or MANCOVA. In fact, all of the afore-mentioned techniques may be subsumed under the general rubric of multivariate GLMs. The general case of the multivariate GLM is too advanced for our purposes here, so the reader is referred to standard textbooks in multivariate analysis. One specific subset of the multivariate GLM, however, is often encountered in neuroscience research—repeated measures.

The repeated measures model had its origins in ANOVA, so the term *repeated measures* ANOVA is often used to refer to the generic design. In fact, the independent variables in repeated measures can be any combination of categorical and continuous variables, so in strict terms, we should really refer to a repeated measures GLM instead of repeated measures ANOVA.

Traditionally, repeated measures GLM was developed in the context of measuring a single phenotype over time or across different situations in the same observations. Examples include yearly measurements of body-mass index, memory testing at five-year intervals in the elderly, venous levels of corticosterone at different time points after a stressor, learning trials, improvement at different time points after a therapy, crossover designs in clinical trails, and signal detection under various levels of background complexity. The key that distinguishes a repeated measures design from all the GLMs considered thus far is that *measurement is on the same observational units*. Hence, the same rat, mouse, person or cell culture is measured over time or across situations.

Current statistical technology for analyzing repeated measures is both blessed and cursed. On the positive side, many of the early assumptions—e.g., that time intervals be evenly spaced—have been relaxed, allowing the investigator greater flexibility in study design. Of particular importance is the many different forms now allowed for the error covariance matrix. (Do not worry if you do not know what that is; we cover it below in Section X.X.) Today, it is not even necessary that the same phenotype be measured. As a result, repeated measures analysis has usurped territory once occupied by MANOVA.

The curse is that the current flexibility in repeated measures has yet to be embodied fully into simple point-and-click statistical interfaces. Furthermore, the seductive nature of the point-and-click often tempts one into an inappropriate analysis. Hence, it is particularly important to know the appropriate theory behind repeated measures analysis in order to do it right.

1.1 Why use repeated measures?

The two major reasons to use repeated measures are: (1) to increase statistical power; and (2) to learn something about the form of the response over time or across situations.
1.2 An example

Figure 1.1 lists six cases from a study comparing a Control (Group = 1 in Figure 1.1) to a Treatment group (Group = 2). After the treatment was administered, rats were given three trials in which acoustic startle (AS) and ultrasonic vocalizations (USV) were measured. Because the same rat was used in all three trials, the trials are not statistically independent of one another. Hence, it is inappropriate to reorganize the data as in Figure 1.2 and perform a GLM with Group and Trial as the independent variables\(^1\).

Figure 1.1 Example of a data set used for repeated measures analysis.

<table>
<thead>
<tr>
<th>Rat</th>
<th>Group</th>
<th>USV1</th>
<th>USV2</th>
<th>USV3</th>
<th>AS1</th>
<th>AS2</th>
<th>AS3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>8.4</td>
<td>7.0</td>
<td>8.1</td>
<td>583.3</td>
<td>342.5</td>
<td>789.4</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>7.1</td>
<td>5.8</td>
<td>7.7</td>
<td>1019.1</td>
<td>821.6</td>
<td>356.8</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>6.4</td>
<td>8.2</td>
<td>8.4</td>
<td>888.1</td>
<td>710.1</td>
<td>577.0</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>8.9</td>
<td>7.3</td>
<td>5.2</td>
<td>771.3</td>
<td>729.5</td>
<td>820.7</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>8.0</td>
<td>9.4</td>
<td>3.9</td>
<td>909.7</td>
<td>754.5</td>
<td>870.1</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>10.0</td>
<td>9.4</td>
<td>5.6</td>
<td>848.0</td>
<td>740.0</td>
<td>414.7</td>
</tr>
</tbody>
</table>

Figure 1.2 A rearranged data set showing the first two observations.

<table>
<thead>
<tr>
<th>Rat</th>
<th>Group</th>
<th>Trial</th>
<th>AS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>583.3</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>342.5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td>789.4</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1019.1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>821.6</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
<td>356.8</td>
</tr>
</tbody>
</table>

There are several reasons why the data in Figure 1.2 cannot be analyzed with ordinary regression or ANOVA, but the easiest one to understand is sample size. The original data set—i.e., the one organized as in Figure 1.1—contains 30 rats. Rearranging the data into the form of Figure 1.2 gives three rows for each rat. Hence, the GLM will be fooled into thinking that there are 90 independent observations and will base all of its statistics and significance on that number. The result is a greater likelihood of false positive (Type I) errors.

Why should one do repeated measures in the first place? After all, could one not perform individual t-tests on the dependent variables? The answer is that a repeated measures GLM

\(^1\) Note that certain software may require the data arrangement in Figure 1.2 for a repeated measures analysis. Our comment is directed to ordinary regression, ANOVA and GLM procedures using this data arrangement.
usually increases statistical power. Hence, fewer subjects are required and lab efficiency is increased.

Let us illustrate with acoustic startle responses of these data—i.e., variables AS1, AS2, and AS3 in Figure 1.1. Figure 1.3 presents a plot of the means of these three variables for the Control and the Treatment groups. Straight forward t tests on these three variables gives significance only for AS2 (t = -2.16, df = 28, p = .04). Hence, the results of the study would be termed equivocal. A repeated measures GLM, however, finds a significant overall difference between the Control and the Treatment groups (F = 4.65, df = (1, 28), p = .04).

**Figure 1.3 Mean startle (+/- 1 SEM) for control and experimental groups.**

![Graph showing mean startle response over trials for control and treatment groups.](image)

1.3 Repeated measures terminology: Between-subjects factors and within-subjects factors

The logic of the experimental design for the data in Figure 1.3 says that there are two ANOVA factors. The first is Group (Control versus Treatment) and the second is Trial (with
three levels, Trial 1, Trial 2, and Trial 3). Just looking at Figure 1.3 one would not know whether the experiment used different rats for Trials 1, 2, and 3 or the same rats for all three trials. This is the crucial distinction, however, for repeated measures analysis.

A repeated measures analysis divides the independent variables in a GLM (both ANOVA factors and continuous independent variables) into two types. The first type is referred to as between-subjects independent variables. Between-subjects independent variables differentiate the rows of the data matrix. If we examine the two factors of Group and Trial and examine which of these differentiate one row from another in Figure 1.1, then the answer is Group. Some rows in Figure 1.1 have a value of Group=1 whereas other rows have a value of Group=2. Hence, ANOVA factor group is a between-subjects factor.

The second type is called within-subject independent variables. Within-subjects independent variables differentiate the columns of a data matrix. If we examine the columns of Figure 1.1 (restricting our inquiry to those related to startle), then the three relevant columns are AS1, AS2, and AS3. The differences between these columns reflect the ANOVA factor that we will call Trial. Hence, Trial is a within-subjects factor. Note that we could also treat Trial as a discrete, continuous variable with numeric values of, say, 1, 2, and 3. In this case, Trial would still be regarded as a within-subjects independent variable.

In implementing a repeated measures GLM, one must always first identify the between-subjects factors and the within-subjects factors. The reason for this is given later (Section X.X). Right now, let us assume that we have identified Group (with two levels) as a between-subjects ANOVA factor and Trial (with three levels) as a within-subjects ANOVA factor.

1.4 The logic of repeated measures

The logic of a repeated measures GLM is identical to the logic of any GLM. The only difference lies in the mechanics of calculating the SS, and hence, the MS, F ratios, and p values. If we ignore for the moment the distinction of between and within factors, then a perusal of Figure 1.3 tells us that an ANOVA will give a main effect for Group, a main effect for Trial, and an interaction between Group and Trial. The main effect for Group asks whether, averaging over Trials, the Control group differs from the treatment group. The main effect for Trial asks whether, averaging over Group, the means of startle response for Trial1, Trial2, and Trial3 differ. And the interaction asks whether the observed means deviate significantly from the predictions made by the main effects of Group and the main effects of Trial.

A repeated measures ANOVA asks the same three questions—i.e., the main effects of Group, the main effects of Trial, and the interaction between Group and Trial. The difference in terms of output is that the repeated measures GLM labels these effects according to whether they are between-subjects or within-subjects effects. The only extra piece of information is that any interaction that contains a within-subjects independent variable is included among the within-subjects effects. Hence, if a repeated measures GLM had the between-subjects independent variables of A and B and within-subjects effects of C and D, then the interaction terms A*C, A*D, B*C, B*D, A*B*C, A*B*D, and A*B*C*D would all be included in the within-subjects effects.

---

2 Because repeated measures was initially developed within the context of ANOVA, the term between-subjects factors is also used even if the independent variable is continuous.
The output from a repeated measures GLM on the acoustic startle data is given in Figure 1.4. Note that the terms in the model are subdivided into between-subjects effects and within-subjects effects, each of which has its own error term. The interpretation of these effects is the same as it would be in an ordinary GLM. Hence, the significant effect for Group means that the overall mean for the Treatment group in Figure 1.3 is significantly higher than the overall mean for the Controls. The significant effect for Trial implies that the three Trial means differ from one another more than sampling error would predict. Finally, the Group*Time interaction is not significant. This implies that the two profile means in Figure 1.3 have the same shape.

Figure 1.4 Repeated measures output on acoustic startle.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>1</td>
<td>654148.827</td>
<td>654148.827</td>
<td>4.65</td>
<td>0.0398</td>
</tr>
<tr>
<td>Error</td>
<td>28</td>
<td>3939696.782</td>
<td>140703.457</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial</td>
<td>2</td>
<td>699007.238</td>
<td>349503.619</td>
<td>8.80</td>
<td>0.0005</td>
</tr>
<tr>
<td>Trial*Group</td>
<td>2</td>
<td>22577.363</td>
<td>11288.681</td>
<td>0.28</td>
<td>0.7536</td>
</tr>
<tr>
<td>Error(Trial)</td>
<td>56</td>
<td>2222930.319</td>
<td>39695.184</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1.5 The key concept in repeated measures: The error covariance matrix

The central difference between repeated measures and ordinary GLM is correlated error in the dependent variables. Let us examine a simple example to see what correlated errors are. Suppose height and weight are measured on a number of men and women. The independent variable in this case is Sex and the two dependent variables for a repeated measures analysis are Height and Weight. The predicted value for height for every female in the sample is just the mean height for females, and the predicted value for weight for every female is the mean weight for females. Assume that these female means are, respectively, 66 inches and 132 pounds. Let
us take a hypothetical woman from that sample, say, Hermione. According to the general linear model, we have the following two equations for Hermione

\[
Y_{\text{Hermione, Height}} = \bar{Y}_{\text{Female, Height}} + \text{Error}_{\text{Hermione, Height}} = 66 + \text{Error}_{\text{Hermione, Height}}
\]

\[
Y_{\text{Hermione, Weight}} = \bar{Y}_{\text{Female, Weight}} + \text{Error}_{\text{Hermione, Weight}} = 132 + \text{Error}_{\text{Hermione, Weight}}
\]

Let’s assume that Hermione is 71 inches tall. That means that her Error for height is 5 inches. Given that she is 5 inches taller than the average woman, she will probably weigh more than the average woman. Hence, her error for weight is likely to be positive. Consider a second woman, Wilma, who is 5 feet tall. Wilma’s error for height is -6 inches. Again, there is a good chance that she will also have a negative error on weight. In sum, if we calculated the errors for height and weight across the entire sample of females, we would find that errors for height are positively correlated with errors for weight.

If we did straight forward t-tests, ANOVAs, or regressions on these data—one for height and one for weight—the error terms from the analyses provide two estimates of the population variances, error variance for height and error variance for weight. Let us denote these as \( \sigma_{\text{Error, height}}^2 \) and Error! Objects cannot be created from editing field codes.. A repeated measures analysis also computes these two error variances, but adds the covariance between the errors for height and weight. The result is now a covariance matrix for error that has the following form

\[
\begin{pmatrix}
\sigma_{\text{Error, Height}}^2 & \text{cov}(\text{Error}_{\text{Height}}, \text{Error}_{\text{Weight}}) \\
\text{cov}(\text{Error}_{\text{Height}}, \text{Error}_{\text{Weight}}) & \sigma_{\text{Error, Weight}}^2
\end{pmatrix}
\]

A repeated measures analysis uses this whole matrix as the error for the within-subjects independent variables in the model. Hence, the difference between a univariate ANOVA and a repeated measures ANOVA is that the univariate ANOVA uses an estimate of the population variance as an error term while the repeated measure uses a whole covariance matrix—variances and covariances included.

Much of the increase in statistical power from repeated measures lies in specifying a form for the error covariance matrix. In the next three sections, we examine three different classes of the form that the error covariance matrix can take.

### 1.5.1 The error covariance matrix: Compound symmetry

In the initial development of repeated measures, it was assumed that the error covariance matrix had a form known as compound symmetry. In this form, all of the variances are equal to one another and all of the covariances are equal to one another. To see this form, recall the formula for a correlation between two variables which we call \( X \) and \( Y \):

\[
\rho_{XY} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}.
\]

Hence, the equation for a covariance is

\[
\text{cov}(X, Y) = \rho_{XY} \sigma_X \sigma_Y.
\]

When all the variances are the same, all of the standard deviations are the same, so \( \sigma_X = \sigma_Y \) and their product is \( \sigma_X \sigma_Y = \sigma^2 \), or the variance. Equation X.X now becomes

\[
\text{cov}(X, Y) = \rho_{XY} \sigma^2.
\]
It follows from this equation, that if all of the covariances are the same, then all of the correlations will be the same. Hence, any covariance will simply equal \( \rho \sigma^2 \). Thus an error covariance matrix with compound symmetry will have the following pattern:
\[
\begin{pmatrix}
\sigma^2 & \rho \sigma^2 & \rho \sigma^2 & \rho \sigma^2 \\
\rho \sigma^2 & \sigma^2 & \rho \sigma^2 & \rho \sigma^2 \\
\rho \sigma^2 & \rho \sigma^2 & \sigma^2 & \rho \sigma^2 \\
\rho \sigma^2 & \rho \sigma^2 & \rho \sigma^2 & \sigma^2 \\
\end{pmatrix}
\]

It is important to remember that \( \hat{\sigma}^2 \) in this matrix is the error variance and that \( \rho \) is the correlation for errors, not the correlation between the raw variables. Note also that this matrix is theoretical; because of sampling error, the observed error variances will never be exactly equal, nor will the observed covariances be equal.

With this assumption at hand, then mathematical transformations can be performed on the error covariance matrix to arrive at a single error term that can be used to test the within-subjects independent variables in the GLM. These tests are sometimes referred to as univariate tests. (Using the whole covariance matrix generates multivariate tests). We can see the effects of this in the output for acoustic startle previously given in Figure 1.4. There is one error term for the between-subjects independent variable Group and a second error term for the two within-subjects factors—Trial and the Trial*Group interaction. The error term for the within-subjects factors is derived from the error covariance matrix under the assumption that is has a specific form.

1.5.2 The error covariance matrix: Huhyn-Feldt conditions

More recent developments in repeated measures have demonstrated that compound symmetry is not the only form that the error covariance matrix can take. (Despite this fact, the myth persists among some old-time users of repeated measures that compound symmetry is required for analysis.) Huyhn and Feldt (1970) outlined the specific conditions for the covariance matrix in repeated measures. They also provided an approximation to the \( p \) level of the \( F \) statistic that takes into account the degree to which the observed error covariance matrix departs from the required form. A second type of correction to the \( p \) value was proposed by Greenhouse and Geisser (1959).

The details of these corrections are too complicate to pursue here, but statistical packages that implement this approach should print out the Huyhn-Feldt and/or the Greenhouse-Geisser adjustments to the \( p \) values in repeated measures analysis. They may also print out statistics called “Huyhn-Feldt \( \varepsilon \)”(i.e., epsilon) and “Greenhouse-Geiser \( \varepsilon \).” Both of these take on values close to 1 when the actual error covariance matrix has the form required for repeated measures. In this case, the adjusted \( p \) values will be very similar to the ordinary \( p \) value for the \( F \) statistic. The more that \( \varepsilon \) departs from 1, the more the observed error covariance matrix deviates from the required form, and hence, the greater the adjustment to the \( p \) levels. Logically then, the Huyhn-Feldt and the Greenhouse-Geisser \( p \) values should be interpreted in place of the \( p \) values for the observed \( F \) statistic.

Why? Well, if the assumptions of the structure of the error covariance are met—or are very close to being met—then all three \( p \) values will be similar. So the choice of Huyhn-Feldt/Greenhouse-Geisser \( p \) values over the ordinary \( p \) value for the \( F \) statistic will be inconsequential. On the other hand, if the assumptions about the error covariance matrix are not
met, then the ordinary \( p \) value for \( F \) will be biased. In this case, the Huynh-Feldt/Greenhouse-Geisser estimates will be more accurate than the ordinary \( p \) values. Hence, the Huynh-Feldt/Greenhouse-Geisser adjusted \( p \) values will, in all cases, be just as appropriate (when the assumptions are met) or more appropriate (when the assumptions are not met) than the ordinary \( p \) value for the \( F \) statistic.

Return to the output in Figure 1.4. The Huynh-Feldt \( \varepsilon \) rounds off to 1.0 and the Greenhouse-Geisser \( \varepsilon \) is .90. Hence, we are assured that the observed error covariance matrix is very close to having the required form for repeated measures. The \( p \) values for the within-subjects factor Trial are .0005 (regular \( F \)), .0008 (Greenhouse-Geisser adjustment) and .0005 (Huynh-Feldt adjustment). There is no substantive difference here. The respective \( p \) values for the Trial*Group interaction are .75, .73, and .75. Again, there is no real substantive difference.

Recall that both the Huynh-Feldt and the Greenhouse-Geisser corrections are approximations. The greater that the observed error covariance matrix departs from Huynh-Feldt conditions, the worse the approximation. Hence, these adjustments can be misleading.

1.5.3 The error covariance matrix: Modern approaches

Contemporary models for repeated measures have moved beyond the mere corrections to the \( p \) value for an \( F \) statistic offered by Huynh and Feldt or by Greenhouse and Geisser. In the software that implements these models, a user can choose among several alternative forms for the error covariance matrix, and in some cases even specify a user’s own form. Hence, one can tell the software to assume compound symmetry, to assume a matrix conforming to Huynh-Feldt conditions, or to assume a number of other matrices that do not meet the Huynh-Feldt conditions. We mention some useful forms for the error covariance matrix.

1.5.3.1 The error covariance matrix: Modern approaches: No form

Some software allows a user to specify that the error covariance matrix has no form. In other words, all of the variances and all of the correlations are permitted to differ in no patterned way

\[
\begin{pmatrix}
\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 & \rho_{14}\sigma_1\sigma_4 \\
\rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 & \rho_{24}\sigma_2\sigma_4 \\
\rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2 & \rho_{34}\sigma_3\sigma_4 \\
\rho_{14}\sigma_1\sigma_4 & \rho_{24}\sigma_2\sigma_4 & \rho_{34}\sigma_3\sigma_4 & \sigma_4^2
\end{pmatrix}
\]

In effect, this type of matrix is the general case for a multivariate GLM, so this option can perform MANOVA or MANCOVA or multivariate regression. This approach has the advantage of being atheoretical about the form of the error covariance, but at the cost of losing some statistical power.

It is useful to perform an initial analysis using an error covariance matrix with no form just to visually inspect the covariance and correlation matrix. The visual inspection can then aid one in selecting a specific form for the error covariance matrix. This approach has the added advantage that if the effects in the model are significant, then they should also be significant for any patterned form of the error covariance matrix. Hence, if the model positing no form is significant, then it is unnecessary to perform additional repeated measures analyses.
1.5.3.2 The error covariance matrix: Modern approaches: Different variances, same correlation

In compound symmetry, all variances in the error covariance matrix are the same and all correlations are the same. It is possible to relax the assumption that the variances are equal while retaining the assumption that the correlations are equal. Hence, the covariance between any two dependent variables will equal

\[ \text{cov}(Y_1, Y_2) = \rho \sigma_{Y_1} \sigma_{Y_2}. \]

Note that although the correlation is the same for all dependent variables, the covariances will not be the same. Why? Because the standard deviations for the variables will differ. Let \( \sigma_i^2 \) denote the variance at the ith time point; hence the standard deviation at the ith time point will be \( \sigma_i \). Then overall form of this matrix is

\[
\begin{pmatrix}
\sigma_1^2 & \rho \sigma_1 \sigma_2 & \rho \sigma_1 \sigma_3 & \rho \sigma_1 \sigma_4 \\
\rho \sigma_2 \sigma_1 & \sigma_2^2 & \rho \sigma_2 \sigma_3 & \rho \sigma_2 \sigma_4 \\
\rho \sigma_3 \sigma_1 & \rho \sigma_3 \sigma_2 & \sigma_3^2 & \rho \sigma_3 \sigma_4 \\
\rho \sigma_4 \sigma_1 & \rho \sigma_4 \sigma_2 & \rho \sigma_4 \sigma_3 & \sigma_4^2
\end{pmatrix}
\]

1.5.3.3 The error covariance matrix: Modern approaches: Measurement over Time

Two forms that do not conform to the Huyn-Feldt conditions are very useful for within-subjects independent variables that involve measurements over time. These are the autoregressive form and the Toeplitz form. Number the time points as 1, 2, 3, \ldots k where k is the last time point. In a Toeplitz matrix, the correlation between any two adjacent time points will be the same. Let us denote this correlation as \( \rho_t \). Hence, the correlation between time 1 and time 2 will equal \( \rho_t \) as will the correlation between times 2 and 3, between times 3 and 4, and so on. The correlation between times separated by two units will all equal \( \rho_2 \). Hence, \( \rho_2 \) will equal the correlation between times 1 and 3, between times 2 and 4, between times 3 and 5, etc.. Similarly, the correlation between time points separated by three time units—e.g., times 1 and 4, times 2 and 5, times 3 and 6, etc.—will all equal \( \rho_3 \). Hence, the Toeplitz form for the error covariance matrix involving four time points is

\[
\begin{pmatrix}
\sigma^2 & \rho_1 \sigma^2 & \rho_2 \sigma^2 & \rho_3 \sigma^2 \\
\rho_1 \sigma^2 & \sigma^2 & \rho_1 \sigma^2 & \rho_2 \sigma^2 \\
\rho_2 \sigma^2 & \rho_1 \sigma^2 & \sigma^2 & \rho_1 \sigma^2 \\
\rho_3 \sigma^2 & \rho_2 \sigma^2 & \rho_3 \sigma^2 & \sigma^2
\end{pmatrix}
\]

An autoregressive process assumes that residuals (i.e., prediction errors) at any time point influence the residuals at the next time point but not any subsequent time points\(^2\). Figure 1.5 depicts the process. Here, \( E_t \) denotes the prediction errors at the ith time point, and \( \beta \) is the regression coefficient between adjacent time points.

---

\(^2\) Strictly speaking, the described process is a first-order autoregressive process. Higher-order autoregressive process may also be modeled, but they will not be discussed here.
Figure 1.5 An autoregressive process.

\[ E_1 \xrightarrow{\beta} E_2 \xrightarrow{\beta} E_3 \xrightarrow{\beta} \ldots \xrightarrow{\beta} E_k \]

If \( \rho \) is the correlation between prediction errors at adjacent time points, then in an autoregressive model \( \rho^2 \) will be the correlation between time points separated by two units—i.e., times 1 and 3, times 2 and 4, etc.. The correlation between time points separated by three units will equal \( \rho^3 \). In general, the correlation between time points separated by \( j \) units will equal \( \rho^j \). Hence, an autoregressive error covariance matrix has the form

\[
\begin{pmatrix}
\sigma^2 & \rho \sigma^2 & \rho^2 \sigma^2 & \rho^3 \sigma^2 \\
\rho \sigma^2 & \sigma^2 & \rho \sigma^2 & \rho^2 \sigma^2 \\
\rho^2 \sigma^2 & \rho \sigma^2 & \sigma^2 & \rho \sigma^2 \\
\rho^3 \sigma^2 & \rho^2 \sigma^2 & \rho \sigma^2 & \sigma^2 \\
\end{pmatrix}
\]

Note that both the Toeplitz and the autoregressive form given above assume that the variance is the same at each time point. This assumption is not necessary, and one can allow the variances to differ at each time point. Let \( \sigma_i^2 \) denote the variance at the \( i \)th time point; hence the standard deviation at the \( i \)th time point will be \( \sigma_i \). A Toeplitz matrix with difference variances will then have the form

\[
\begin{pmatrix}
\sigma_1^2 & \rho_1 \sigma_1 \sigma_2 & \rho_2 \sigma_1 \sigma_3 & \rho_3 \sigma_1 \sigma_4 \\
\rho_1 \sigma_1 \sigma_2 & \sigma_2^2 & \rho_2 \sigma_2 \sigma_3 & \rho_3 \sigma_2 \sigma_4 \\
\rho_2 \sigma_1 \sigma_3 & \rho_1 \sigma_2 \sigma_3 & \sigma_3^2 & \rho_2 \sigma_3 \sigma_4 \\
\rho_3 \sigma_1 \sigma_4 & \rho_2 \sigma_2 \sigma_4 & \rho_1 \sigma_3 \sigma_4 & \sigma_4^2 \\
\end{pmatrix}
\]

1.5.3.4 The error covariance matrix: Modern approaches: Measurement over Space

Many studies in neuroscience involve taking slices of a nucleus in the brain for staining, autoradiography, or immunohistochemistry. Imagine a study that takes four coronal slices from the dorsal hippocampus in each rat and assays each section for a substance. It could be that the level in the anterior-most section is more highly correlated with the adjacent section than it is with the posterior-most section. Hence, there error covariance matrix can have spatial pattern that is technically referred to as spatial autocorrelation.

Depending on the problem at hand, the autoregressive and the Toeplitz matrices may be appropriate for some data. Many software routines for repeated measures may contain options for other patterns because the generic problem of spatial autocorrelation has been widely researched in geography, geology, ecology, and many other fields.

Let us recap these principles from a historical perspective, albeit an oversimplified one. When repeated measures analysis was initially developed, the mathematics of the procedure
were based on the assumption of compound symmetry for the error covariance matrix. Later, Huynh and Feldt effectively said, "No! There are several other conditions for the error covariance matrix that meet these assumptions needed for repeated measures. And we provide an approximate correction to the $p$ value of an $F$ statistic under these conditions." The current state of the art allows a user to choose among several different forms of the error covariance matrix.

Hence, there are three different approaches to repeated measures: (1) the classic approach; (2) the Huynh-Feldt/Greenhouse-Geiser approach; and (3) the modern approach. Often, the three approaches agree, particularly if the effect sizes for the repeated measures factors are large.

When to use repeated measures

Assumptions of repeated measures

Homogeneity of the error covariance matrices

Recall that the GLM with a single dependent assumes that the residuals are normally distributed and have a constant variance for all values of the independent variables. If an independent variable is an ANOVA factor, then the assumption of a constant variance implies that the variances within each level of the ANOVA factor are within sampling error of one another—i.e., the assumption of homogeneity of variance.

In repeated measures designs, it is assumed that the residuals are normally distributed and have a constant covariance matrix for all values of the independent variables. In terms of an ANOVA factor, this implies that the covariance matrices within each level of the ANOVA factor differ only as a function of sampling error. This is the assumption of homogeneity of covariance.

The customary test for homogeneity of covariance is a procedure developed by Box (1949) which gives a $\chi^2$ test statistic. If the $p$ value is less than a preset alpha level, then the assumption of homogeneity is rejected. Unfortunately, this test is very sensitive to departures from normality and is likely to give misleadingly low $p$ values in cases where the error covariance matrices within each group are not far from being homogeneous. Furthermore, the small sample properties of this test have not been rigorously investigated.

References

