Matrix Formulation of LMER Models

2017-11-02

Linear Mixed-Effect Model: Linear Growth

Consider two linear mixed-effects models. The first model is the model which includes a linear fixed-effect of time and a random-effect of intercept only. Writing out the equation for that model:

$$Y_{ij} = \beta_0 + \beta_1(t_{ij}) + b_{0i} + \epsilon_{ij}$$

where t_{ij} is the value of the time predictor for Subject i at time point j. The second model includes a linear fixed-effect of time, again t_{ij}, and random-effects for both intercept and slope. We would write the equation for this model as:

$$Y_{ij} = \beta_0 + \beta_1(t_{ij}) + b_{0i} + b_{1i}(t_{ij}) + \epsilon_{ij}$$

Let's also consider that we are fitting these models to a data set in which wach subject has four time points. The model-equation repressnts a system of equations for each subject; one equation for each subject/time-point combination. Since our data set includes four time points for each subject, the model-equations represent a sytem of four equations for each subject. The system of equations for the first subject ($i = 1$) for the first model are:

$$Y_{1,1} = \beta_0 + \beta_1(t_{1,1}) + b_{0i} + \epsilon_{i1}$$
$$Y_{1,2} = \beta_0 + \beta_1(t_{1,2}) + b_{0i} + \epsilon_{i2}$$
$$Y_{1,3} = \beta_0 + \beta_1(t_{1,3}) + b_{0i} + \epsilon_{i3}$$
$$Y_{1,4} = \beta_0 + \beta_1(t_{1,4}) + b_{0i} + \epsilon_{i4}$$

And, the system of equations for the first subject for the second model are:

$$Y_{1,1} = \beta_0 + \beta_1(t_{1,1}) + b_{0i} + b_{1i}(t_{1,1}) + \epsilon_{i1}$$
$$Y_{1,2} = \beta_0 + \beta_1(t_{1,2}) + b_{0i} + b_{1i}(t_{1,2}) + \epsilon_{i2}$$
$$Y_{1,3} = \beta_0 + \beta_1(t_{1,3}) + b_{0i} + b_{1i}(t_{1,3}) + \epsilon_{i3}$$
$$Y_{1,4} = \beta_0 + \beta_1(t_{1,4}) + b_{0i} + b_{1i}(t_{1,4}) + \epsilon_{i4}$$

These systems of equations would look identical for the other subjects, except that the i-value would change to acknowledge that we are writing these for a different subject. We can also express these same equations using matrices. Here is the system of equations for the first model written more generally for Subject i using matrices:

$$\begin{bmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ Y_{i4} \end{bmatrix} = \begin{bmatrix} 1 & t_{i1} \\ 1 & t_{i2} \\ 1 & t_{i3} \\ 1 & t_{i4} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} b_{0i} \end{bmatrix} + \begin{bmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \epsilon_{i3} \\ \epsilon_{i4} \end{bmatrix}$$

And for the second model:

$$\begin{bmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ Y_{i4} \end{bmatrix} = \begin{bmatrix} 1 & t_{i1} \\ 1 & t_{i2} \\ 1 & t_{i3} \\ 1 & t_{i4} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} 1 & t_{i1} \\ 1 & t_{i2} \\ 1 & t_{i3} \\ 1 & t_{i4} \end{bmatrix} \begin{bmatrix} b_{0i} \\ b_{1i} \end{bmatrix} + \begin{bmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \epsilon_{i3} \\ \epsilon_{i4} \end{bmatrix}$$

Both of these matrix expressions have the same general form. A vector of the outcomes is the sum of two matrix products and a vector of level–1 errors. The first matrix product is the product of the design matrix for the fixed-effects post-multiplied by the vector of fixed-effects. The second matrix product is the product of the design matrix for the random-effects post-multiplied by the vector of random-effects. This is referred to as the *general form* for the linear mixed-effects model. We can express this general form using the bold matrix notation as,

$$\mathbf{Y_{ij}} = \mathbf{X_i}\beta + \mathbf{Z_i}\mathbf{b_i} + \epsilon_{ij}$$

In this equation,

- $\mathbf{X_i}$ is the design matrix for the fixed-effects for the ith individual
- $\mathbf{Z_i}$ is the design matrix for the random-effects for the ith individual
- β is a vector of the fixed-effects
- $\mathbf{b_i}$ is a vector of the random-effects for the ith individual, and
- ϵ_{ij} is a vector of the level-1 errors for the ith individual

Note that the dimensions of each of these vectors and matrices depends on the model that was specified.

Example: Vocabulary Data

Consider again the vocabulary data that we used in previous notes. (The importing and preparation of these data is not shown.) Here we show the first two subjects' data.

```
## # A tibble: 8 x 4
##       id grade score female
##    <int> <dbl> <dbl>  <int>
## 1     1     0  1.75      0
## 2     1     1  2.60      0
## 3     1     2  3.76      0
## 4     1     3  3.68      0
## 5     2     0  0.90      0
## 6     2     1  2.47      0
## 7     2     2  2.44      0
## 8     2     3  3.43      0
```

Here are the systems of equations for the first model (linear fixed–effect of time and intercept random–effect) shown for both subjects. Note that we use the data values for Y_{ij} and t_{ij}:

$$1.75 = \beta_0 + \beta_1(0) + b_{0,1} + \epsilon_{1,1}$$
$$2.60 = \beta_0 + \beta_1(1) + b_{0,1} + \epsilon_{1,2}$$
$$3.76 = \beta_0 + \beta_1(2) + b_{0,1} + \epsilon_{1,3}$$
$$3.68 = \beta_0 + \beta_1(3) + b_{0,1} + \epsilon_{1,4}$$

$$0.90 = \beta_0 + \beta_1(0) + b_{0,2} + \epsilon_{2,1}$$
$$2.47 = \beta_0 + \beta_1(1) + b_{0,2} + \epsilon_{2,2}$$
$$2.44 = \beta_0 + \beta_1(2) + b_{0,2} + \epsilon_{2,3}$$
$$3.43 = \beta_0 + \beta_1(3) + b_{0,2} + \epsilon_{2,4}$$

These equations would be written out for each subject.

Once we fit the model and get the estimates for the fixed–effects (β_0, β_1) and the random–effects (b_0 for Subject 1 and 2), we can obtain the predicted values; substitute these values into the equations, and drop the error term. Fitting this model:

```
lmer.1 = lmer(score ~ 1 + grade + (1 | id), data = vocab_long, REML = FALSE)

# Obtain fixed-effecta
fixef(lmer.1)

## (Intercept)       grade
##   1.4132969   0.7466562

# Obtain random-effects (for first two subjects)
head(ranef(lmer.1)$id)

##    (Intercept)
## 1    0.3869787
## 2   -0.2085977
## 3   -1.3390251
## 4    1.2161145
## 5   -3.6512630
## 6   -1.7547608
```

$$\hat{Score}_{1,1} = 1.41 + 0.75(0) + 0.39 = 1.80$$
$$\hat{Score}_{1,2} = 1.41 + 0.75(1) + 0.39 = 2.55$$
$$\hat{Score}_{1,3} = 1.41 + 0.75(2) + 0.39 = 3.30$$
$$\hat{Score}_{1,4} = 1.41 + 0.75(3) + 0.39 = 4.05$$

$$\hat{Score}_{2,1} = 1.41 + 0.75(0) + -0.21 = 1.20$$
$$\hat{Score}_{2,2} = 1.41 + 0.75(1) + -0.21 = 1.95$$
$$\hat{Score}_{2,3} = 1.41 + 0.75(2) + -0.21 = 2.70$$
$$\hat{Score}_{2,4} = 1.41 + 0.75(3) + -0.21 = 3.45$$

These are the same values we get from the `predict()` function when we ask for random-effects to be included in the model. (When the `newdata=` argument is not included the `predict()` function predicts for the data that was used to fit the model.)

```
vocab_long %>%
  mutate( yhat = predict(lmer.1) )
```

```
## # A tibble: 8 x 5
##       id grade score female      yhat
##    <int> <dbl> <dbl>  <int>     <dbl>
## 1     1     0  1.75      0  1.800276
## 2     1     1  2.60      0  2.546932
## 3     1     2  3.76      0  3.293588
## 4     1     3  3.68      0  4.040244
## 5     2     0  0.90      0  1.204699
## 6     2     1  2.47      0  1.951355
## 7     2     2  2.44      0  2.698012
## 8     2     3  3.43      0  3.444668
```

If we want the predicted values from the fixed–effects part of the model only, we drop the random–effects part of the predicted model:

$$\hat{\text{Score}}_{1,1} = 1.41 + 0.75(0) = 1.41$$
$$\hat{\text{Score}}_{1,2} = 1.41 + 0.75(1) = 2.16$$
$$\hat{\text{Score}}_{1,3} = 1.41 + 0.75(2) = 2.91$$
$$\hat{\text{Score}}_{1,4} = 1.41 + 0.75(3) = 3.66$$

$$\overline{\hat{\text{Score}}_{2,1} = 1.41 + 0.75(0) = 1.41}$$
$$\hat{\text{Score}}_{2,2} = 1.41 + 0.75(1) = 2.16$$
$$\hat{\text{Score}}_{2,3} = 1.41 + 0.75(2) = 2.91$$
$$\hat{\text{Score}}_{2,4} = 1.41 + 0.75(3) = 3.66$$

The predictions from the fixed–effects are the exact same for each individual (recall they represent the average model). These are the same values we get from the `predict()` function when we set the random–effects to `NA`.

```
vocab_long %>%
  mutate( yhat = predict(lmer.1, re.form = NA) )
```

```
## # A tibble: 8 x 5
##       id grade score female      yhat
##    <int> <dbl> <dbl>  <int>     <dbl>
## 1     1     0  1.75      0  1.413297
## 2     1     1  2.60      0  2.159953
## 3     1     2  3.76      0  2.906609
## 4     1     3  3.68      0  3.653266
## 5     2     0  0.90      0  1.413297
## 6     2     1  2.47      0  2.159953
## 7     2     2  2.44      0  2.906609
## 8     2     3  3.43      0  3.653266
```

Using Matrix Algebra to Obtain the Predicted Values

We can also use matrix algebra to obtain the predicted values. The general form for the predicted values is:

$$\hat{\mathbf{Y}}_{ij} = \mathbf{X}_i\hat{\beta} + \mathbf{Z}_i\hat{\mathbf{b}}_i$$

To do this we need (1) the design matrix for the fixed-effects, (2) the design mtrix for the random-effects, (3) a column vector of the estimated fixed-effects, and (4) a column vector of the estimated random-effects. Below, we input the design matrices for the first subject and the two column vectors based on the first subject:

```
# Design matrix of the fixed-effects
X = matrix(data = c(1, 1, 1, 1, 0, 1, 2, 3), ncol = 2)
X
```

```
##      [,1] [,2]
## [1,]    1    0
## [2,]    1    1
## [3,]    1    2
## [4,]    1    3
```

```
# Design matrix of the random-effects
Z = matrix(data = c(1, 1, 1, 1), ncol = 1)
Z
```

```
##      [,1]
## [1,]    1
## [2,]    1
## [3,]    1
## [4,]    1
```

```
# Column vector of the estimated fixed-effects
B = matrix(data = c(1.41, 0.75), ncol = 1)
B
```

```
##      [,1]
## [1,] 1.41
## [2,] 0.75
```

```
# Column vector of the estimated random-effects
b = matrix(data = c(0.39), ncol = 1)
b
```

```
##      [,1]
## [1,] 0.39
```

Then the matrix algebra to get the predicted values would be

```
X %*% B + Z %*% b
```

```
##      [,1]
## [1,] 1.80
## [2,] 2.55
## [3,] 3.30
## [4,] 4.05
```

To do this for the second subject, we would simply use the design matrices and estimated effects associated with that subject. We could also compute the predicted values from the fixed-effects part of the model by simply dropping the random-effects portion (the second product). In practice, one would likely just use the predict() function. However, it helps us understand how that function actually computes the predicted values.

Modeled Variance–Covariance Matrix of the Repeated Measures

The mixed-effects model will try to model the variance–covariance matrix of the repeated measures. The actual variance–covariance matrix of the repeated measures is

```r
Y = vocab %>%
  select(t8:t11) %>%
  cov(.)

Y
```

```
##          t8       t9      t10      t11
## t8  3.553340 3.184662 3.540069 2.840132
## t9  3.184662 4.346827 3.547852 3.039358
## t10 3.540069 3.547852 4.703507 3.388601
## t11 2.840132 3.039358 3.388601 3.707515
```

We call the variance–covariance matrix for a single subject $\mathbf{V_i}$. This is a $n_i \times n_i$ matrix where n_i is the number of repeated measures for Subject i. We will consider the second model (linear fixed-effects of time and RE for both intercept and slope) in this section.

```r
lmer.2 = lmer(score ~ 1 + grade + (1 + grade | id), data = vocab_long, REML = FALSE)
```

For the data set we have been considering, there were four time points, so V_i is a 4×4 matrix,

$$\mathbf{V_i} = \begin{bmatrix} \mathrm{Var}(Y_{i1}) & \mathrm{Cov}(Y_{i1}, Y_{i2}) & \mathrm{Cov}(Y_{i1}, Y_{i3}) & \mathrm{Cov}(Y_{i1}, Y_{i4})) \\ \mathrm{Cov}(Y_{i2}, Y_{i1}) & \mathrm{Var}(Y_{i2}) & \mathrm{Cov}(Y_{i2}, Y_{i3}) & \mathrm{Cov}(Y_{i2}, Y_{i4}) \\ \mathrm{Cov}(Y_{i3}, Y_{i1}) & \mathrm{Cov}(Y_{i3}, Y_{i2}) & \mathrm{Var}(Y_{i3}) & \mathrm{Cov}(Y_{i3}, Y_{i4}) \\ \mathrm{Cov}(Y_{i4}, Y_{i1}) & \mathrm{Cov}(Y_{i4}, Y_{i2}) & \mathrm{Cov}(Y_{i4}, Y_{i3}) & \mathrm{Var}(Y_{i4}) \end{bmatrix}$$

This is a symmetric matrix that indexes the variation at each time point (variances) and the dependency between time points (covariances). We can decompose this matrix into two sources: between-subjects variation ($\mathbf{B_i}$) and within-subjects variation ($\mathbf{W_i}$). Then

$$\mathbf{V_i} = \mathbf{B_i} + \mathbf{W_i}$$

Within-Subjects Variance–Covariance Matrix

The within-subjects matrix, $\mathbf{W_i}$, can be expressed as

$$\mathbf{W_i} = \sigma_\epsilon^2 \mathbf{I_i}$$

where σ_ϵ^2 is the variance of the level-1 errors, and $\mathbf{I_i}$ is a $n_i \times n_i$ identity matrix. For our student with four time points,

$$\mathbf{W_i} = \sigma_\epsilon^2 \mathbf{I_i}$$
$$= \sigma_\epsilon^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \sigma_\epsilon^2 & 0 & 0 & 0 \\ 0 & \sigma_\epsilon^2 & 0 & 0 \\ 0 & 0 & \sigma_\epsilon^2 & 0 \\ 0 & 0 & 0 & \sigma_\epsilon^2 \end{bmatrix}$$

This matrix emphasizes the assumptions about the level-1 errors. Namely that they are independent with constant variance.

Between-Subjects Variance–Covariance Matrix

The between-subjects matrix, $\mathbf{B_i}$, can be expressed as

$$\mathbf{B_i} = \mathbf{Z_i G Z}_i^\mathsf{T}$$

where $\mathbf{Z_i}$ is the design matrix of the random-effects (and $\mathbf{Z}_i^\mathsf{T}$ is the transpose of this matrix) and $\mathbf{G}$ is the variance–covariance matrix of the random-effects defined as,

$$\mathbf{G} = \begin{bmatrix} \text{Var}(b_0) & \text{Cov}(b_0, b_1) \\ \text{Cov}(b_1, b_0) & \text{Var}(b_1) \end{bmatrix}$$

Substituting the between- and within-subjects matrices back into the variance–covariance formulation of the repeated measures, we get

$$\mathbf{V_i} = \mathbf{B_i} + \mathbf{W_i}$$
$$= \mathbf{Z_i G Z}_i^\mathsf{T} + \sigma_\epsilon^2 \mathbf{I_i}$$

Substituting in the matrices for the model that includes a random-effect of intercept and slope,

$$\mathbf{V_i} = \mathbf{Z_i G Z}_i^\mathsf{T} + \sigma_\epsilon^2 \mathbf{I_i}$$

$$= \begin{bmatrix} 1 & X_{i1} \\ 1 & X_{i2} \\ 1 & X_{i3} \\ 1 & X_{i4} \end{bmatrix} \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ X_{i1} & X_{i2} & X_{i3} & X_{i4} \end{bmatrix} + \begin{bmatrix} \sigma_\epsilon^2 & 0 & 0 & 0 \\ 0 & \sigma_\epsilon^2 & 0 & 0 \\ 0 & 0 & \sigma_\epsilon^2 & 0 \\ 0 & 0 & 0 & \sigma_\epsilon^2 \end{bmatrix}$$

and carrying out the matrix algebra we find each variance (element on the diagonal) has a similar mathematical form. The variance for time point 1 is:

$$\text{Var}(Y_{i1}) = \sigma_0^2 + 2(X_{i1})(\sigma_{01}) + X_{i1}^2(\sigma_1^2) + \sigma_\epsilon^2$$

This implies that the variance at each time point is a function of the time predictor and the squared time predictor, the variances and covariances of the random-effects and the level-1 error variance.

The covariance at each time point also all have the same mathematical function. For example the covariance between time point 1 and time point 2 is

$$\text{Cov}(Y_{i1}, Y_{i2}) = \sigma_0^2 + X_{i1}(\sigma_{01}) + X_{i2}(\sigma_{10}) + \sigma_1^2(X_{i1})(X_{i2})$$

This implies that the covariance between time points is a function of the time predictor, the variances and covariances of the random-effects. Let's see this in action for our data set. We will need (1) the estimate of the variance for the level-1 errors ($\hat{\sigma}_\epsilon^2$), (2) the design matrix for the random effects ($\mathbf{Z_i}$); and (3) the variance–covariance matrix of the random-effects ($\mathbf{G}$).

```
# Obtain variance estimate for the level-1 error
summary(lmer.2)$sigma ^ 2
```

```
## [1] 0.896332
```

```
# Obtain variance-covariance matrix of the random-effects
VarCorr(lmer.2)$id[1:2, 1:2]
```

```
##             (Intercept)        grade
## (Intercept)  3.13762586 1.535409e-02
## grade        0.01535409 7.513581e-05
```

The design matrix for the random-effects is

$$\mathbf{Z_i} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$$

So the estimated within-subjects variance–covariance matrix is

$$\hat{\mathbf{W}}_\mathbf{i} = \hat{\sigma}_\epsilon^2 \mathbf{I_i}$$

$$= 0.896 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.896 & 0 & 0 & 0 \\ 0 & 0.896 & 0 & 0 \\ 0 & 0 & 0.896 & 0 \\ 0 & 0 & 0 & 0.896 \end{bmatrix}$$

```
# Error variance
sigma2error = 0.89633196
```

```
# Compute estimated within-subjects variance-covariance matrix
W = sigma2error * diag(4)
W
```

```
##          [,1]     [,2]     [,3]     [,4]
## [1,] 0.896332 0.000000 0.000000 0.000000
## [2,] 0.000000 0.896332 0.000000 0.000000
## [3,] 0.000000 0.000000 0.896332 0.000000
## [4,] 0.000000 0.000000 0.000000 0.896332
```

The estimated between-subjects variance–covariance matrix is

$$\hat{\mathbf{B}}_\mathbf{i} = \mathbf{Z_i}\hat{\mathbf{G}}\mathbf{Z}_i^\mathsf{T}$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 3.138 & 0.0154 \\ 0.0154 & 0.00008 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix}$$

```
# Enter design matrix for random-effects
Z = matrix(data = c(1, 1, 1, 1, 0, 1, 2, 3), ncol = 2)
Z
```

```
##      [,1] [,2]
## [1,]    1    0
## [2,]    1    1
## [3,]    1    2
## [4,]    1    3
```

```
# Enter variance-covariance matrix for the random-effects
G = matrix(data = c(3.13762586, 0.01535409, 0.01535409, 0.00007513581) , ncol = 2)
G
```

```
##            [,1]         [,2]
## [1,] 3.13762586 1.535409e-02
## [2,] 0.01535409 7.513581e-05
```

```
# Compute the estiated between-subjects variance-covariance matrix
B = Z %*% G %*% t(Z)
B
```

```
##          [,1]     [,2]     [,3]     [,4]
## [1,] 3.137626 3.152980 3.168334 3.183688
## [2,] 3.152980 3.168409 3.183838 3.199268
## [3,] 3.168334 3.183838 3.199343 3.214847
## [4,] 3.183688 3.199268 3.214847 3.230427
```

The model-based estimate for the variance–covariance matrix of the repeated measures is then

```
V = B + W
V
```

```
##          [,1]     [,2]     [,3]     [,4]
## [1,] 4.033958 3.152980 3.168334 3.183688
## [2,] 3.152980 4.064741 3.183838 3.199268
## [3,] 3.168334 3.183838 4.095675 3.214847
## [4,] 3.183688 3.199268 3.214847 4.126759
```

This matrix is the model-based estimate for the actual variances and covariances among the repeated measures in the data. Note that the dependency in the data (the non-zero covariances) are a function of the between-subjects variance–covariance matrix. (In the within-subjects variance–covariance matrix, the off-diagonal elements are all zero.) In the between-subjects variance–covariance matrix, it is the inclusion of the random-effects that allows us to model the non-zero covariances (dependency) between the repeated measures. This is why we include a random-effect(s) to appropriately model data that are non–independent.

Standardizing the Modeled Variance–Covariance Matrix of the Repeated Measures

To further understand the dependency among the repeated measures, it is helpful to convert the modeled variance–covariance matrix of the repeated measures into a correlation matrix. To do this, we convert using,

$$\mathbf{V_i^*} = \mathbf{D_i V_i D_i}$$

where $\mathbf{D_i}$ is a $n_i \times n_i$ diagonal matrix with elements equal to

$$\frac{1}{\sqrt{\text{Var}(Y_{ij})}}$$

For our example with four time points,

$$\mathbf{D_i} = \begin{bmatrix} \frac{1}{\sqrt{\text{Var}(Y_{i1})}} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{\text{Var}(Y_{i2})}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{\text{Var}(Y_{i3})}} & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{\text{Var}(Y_{i4})}} \end{bmatrix}$$

In our specific example,

```r
# Create the diagonal matrix based on the variances
D = diag( 1/sqrt(diag(V)) )
D
```

```
##           [,1]      [,2]      [,3]     [,4]
## [1,] 0.4978911 0.0000000 0.0000000 0.000000
## [2,] 0.0000000 0.4960021 0.0000000 0.000000
## [3,] 0.0000000 0.0000000 0.4941255 0.000000
## [4,] 0.0000000 0.0000000 0.0000000 0.492261
```

```r
# Compute the standardized V_i
D %*% V %*% D
```

```
##           [,1]      [,2]      [,3]      [,4]
## [1,] 1.0000000 0.7786443 0.7794757 0.7802977
## [2,] 0.7786443 1.0000000 0.7803184 0.7811413
## [3,] 0.7794757 0.7803184 1.0000000 0.7819754
## [4,] 0.7802977 0.7811413 0.7819754 1.0000000
```

These are the model–based estimated correlations between the repeated measures. This shows a pattern of correlation similar to what is in the actual data.

```r
cov2cor(Y)
```

```
##            t8        t9       t10       t11
## t8  1.0000000 0.8103242 0.8659294 0.7824899
## t9  0.8103242 1.0000000 0.7846364 0.7571016
## t10 0.8659294 0.7846364 1.0000000 0.8114619
## t11 0.7824899 0.7571016 0.8114619 1.0000000
```

The correlations (variances/covariances) are not perfectly reproduced. The model–based correlations seem a little lower than the correlations in the data. Remember, we are modeling these correlations. We don't expect perfect replication of the data from a model. One question we can ask is whether the model–based correlations are lower because the level-1 errors are not independent. If that were the case, the off-diagonal elements of $\mathbf{W_i}$ would not be zero and the model–estimated correlations would, in turn, be higher. To determine if this is the case, or if the differences are just due to chance, we need to evaluate the model's assumptions.

References

Long, J. D. (2012). *Longitudinal data analysis for the behavioral sciences using R*. Thousand Oaks, CA: Sage Publications, Inc.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.