LINEAR MIXED MODELS SUMMARY

SHRAVAN VASISHTH (VASISHTH@UNI-POTSDAM.DE)

These notes summarize the lecture notes from the Linear Modelling course at Sheffield's School of Mathematics and Statistics, MSc degree programme. The original notes were written by Dr. Jeremy Oakley. This summary is completely derived from these notes and from other MSc sources. Any errors are most probably mine.

Everything is in matrix form unless a lower case letter with a subscript (such as x_i) is used (even there, I might deviate from this convention if I need to index sub-matrices; it's best to look at the context to decide what is meant).

1. Some basic types of linear mixed model and their variance components

1.1. Varying intercepts model.

```
> library(lme4)
> fm1<-lmer(wear~material+(1|Subject),BHHshoes)
> ranef(fm1)
$Subject
   (Intercept)
1      2.74820
2     -2.32081
3      0.21369
4      3.39425
5      0.41248
6     -4.30866
7     -1.17780
8      0.21369
9     -1.77415
10     2.59911
```

The model is:

$$Y_{ijk} = \beta_j + b_i + \epsilon_{ijk} \tag{1}$$

$i = 1, \ldots, 10$ is subject id, $j = 1, 2$ is the factor level, k is the number of replicates (here 1). $b_i \sim N(0, \sigma_b^2), \epsilon_{ijk} \sim N(0, \sigma^2)$.

The general form for any model in this case is:

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \sim N\left(\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, V\right) \tag{2}$$

where $V = \begin{pmatrix} \sigma_b^2 + \sigma^2 & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 + \sigma^2 \end{pmatrix} = \begin{pmatrix} \sigma_b^2 + \sigma^2 & \rho\sigma_b\sigma_b \\ \rho\sigma_b\sigma_b & \sigma_b^2 + \sigma^2 \end{pmatrix}$.

We can recover these variance components as follows:

```
> VarCorr(fm1)
```

```
$Subject
            (Intercept)
(Intercept)      6.1009
attr(,"stddev")
(Intercept)
       2.47
attr(,"correlation")
            (Intercept)
(Intercept)           1

attr(,"sc")
[1] 0.27376
```

$\hat{V}$ is therefore:

$$\begin{pmatrix} \hat{\sigma}_b^2 + \hat{\sigma}^2 & \hat{\rho}\hat{\sigma}_b\hat{\sigma}_b \\ \hat{\rho}\hat{\sigma}_b\hat{\sigma}_b & \hat{\sigma}_b^2 + \hat{\sigma}^2 \end{pmatrix} = \begin{pmatrix} 2.47^2 + 0.27376^2 & 2.47^2 \\ 2.47^2 & 2.47^2 + 0.27376^2 \end{pmatrix} \tag{3}$$

Note: $\hat{\rho} = 1$ because the off-diagonal is $2.47^2 = 1 \times 2.47 \times 2.47$. But this correlation is not estimated in the varying intercepts model.

1.2. **Varying intercepts and slopes (with correlation).**

```
> fm2<-lmer(wear~material+(1+material|Subject),BHHshoes)
> ranef(fm2)
```

```
$Subject
  (Intercept)  materialB
1     2.71318  0.0752088
2    -2.28776 -0.0634150
3     0.21003  0.0058217
4     3.34435  0.0927024
5     0.41145  0.0114069
6    -4.25374 -0.1179130
7    -1.16241 -0.0322216
8     0.21137  0.0058593
```

```
9     -1.74925 -0.0484882
10     2.56278  0.0710387
> VarCorr(fm2)
$Subject
            (Intercept) materialB
(Intercept)     5.93634 0.1645525
materialB       0.16455 0.0045617
attr(,"stddev")
(Intercept)   materialB
    2.43646     0.06754
attr(,"correlation")
            (Intercept) materialB
(Intercept)     1.00000   0.99996
materialB       0.99996   1.00000

attr(,"sc")
[1] 0.26956
```

The model is

$$Y_{ijk} = \beta_j + b_{ij} + \epsilon_{ijk} \tag{4}$$

$b_{ij} \sim N(0, \sigma_b)$. The variance σ_b must be a 2×2 matrix:

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \tag{5}$$

We can recover this from the random effects:

```
> var(ranef(fm2)$Subject)
            (Intercept) materialB
(Intercept)     5.90124 0.1635798
materialB       0.16358 0.0045344
```

Note that $1 \times \sqrt{5.90124} \times \sqrt{0.0045344} = 0.16358$, which is how we get that $\hat{\rho} = 1$.

The general form for the model is:

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \sim N\left(\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, V \right) \tag{6}$$

where

$$V = \begin{pmatrix} \sigma_{b,A}^2 + \sigma^2 & \rho\sigma_{b,A}\sigma_{b,B} \\ \rho\sigma_{b,A}\sigma_{b,B} & \sigma_{b,B}^2 + \sigma^2 \end{pmatrix} \tag{7}$$

And that's equal to (see VarCorr output above):

$$\begin{pmatrix} 5.93634 + 0.07266 & \rho\sigma_{b,A}\sigma_{b,B} = 0.1645525 \\ 0.1645525 & 0.0045617 + 0.07266 \end{pmatrix} \tag{8}$$

Note that $\hat{\rho}$ is shown in VarCorr output (at the bottom) and can be computed since $\frac{Covar}{\sigma_a \times \sigma_b} = \rho$ and we know all the quantities on the LHS:

```
> 0.1645525/(sqrt(5.93634)*sqrt(0.0045617))
[1] 0.99996
```

How to recover, from V, the correlation of 1 in the lmer random effects output of fm2? Is that 1 supposed to represent 0.99996?

1.3. No varying intercepts, only slopes for each level.

```
> fm3<-lmer(wear~material-1 + (material-1|Subject),BHHshoes)
> ranef(fm3)
$Subject
   materialA materialB
1    2.71318   2.78838
2   -2.28776  -2.35117
3    0.21003   0.21585
4    3.34435   3.43705
5    0.41145   0.42286
6   -4.25374  -4.37165
7   -1.16241  -1.19463
8    0.21137   0.21723
9   -1.74925  -1.79774
10   2.56278   2.63382
```

The model is

$$Y_{ijk} = \beta_j + b_{ij} + \epsilon_{ijk} \tag{9}$$

The random effects are:
$b_{ij} = \begin{pmatrix} b_{i1} \\ b_{i12} \end{pmatrix} \sim N(0, \sigma_b^2)$, where $\sigma_b^2 = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$.
We can recover these values from:

```
> var(ranef(fm3)$Subject)
          materialA materialB
materialA    5.9012    6.0648
materialB    6.0648    6.2329
```

$\hat{\rho}$ is 1 because $1 \times \sqrt{5.9012} * \sqrt{6.2329} = 6.0648$.
Here, V is

$$V = \begin{pmatrix} \sigma_{b,A}^2 + \sigma^2 & \rho\sigma_{b,A}\sigma_{b,B} \\ \rho\sigma_{b,A}\sigma_{b,B} & \sigma_{b,B}^2 + \sigma^2 \end{pmatrix} \tag{10}$$

Note that the interpretation of the random effects is different from fm2: here, a random effect is computed for each material separately.

From the VarCorr output, we have $\hat{V}$:

$$\begin{pmatrix} 5.9363 + 0.26956^2 & 1 \times 2.4365 \times 2.5040 \\ 1 \times 2.4365 \times 2.5040 & 6.27 + 0.26956^2 \end{pmatrix} \tag{11}$$

One insight is that V can be derived from the random effects variance components, and the error term's variance component:

$$V = \begin{pmatrix} \sigma_{b,A}^2 & \rho\sigma_{b,A}\sigma_{b,B} \\ \rho\sigma_{b,A}\sigma_{b,B} & \sigma_{b,B}^2 \end{pmatrix} + \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \tag{12}$$

1.4. **Nested models (e.g., Worker/Machine).** The model is:

$$Y_{ijk} = \beta_j + b_i + b_{ij} + \epsilon_{ijk} \tag{13}$$

Here, we force force all random effects to be independent. Observations between workers are independent, but observations on the same worker are correlated.

$b_i \sim N(0, \sigma_1^2), b_{ij} \sim N(0, \sigma_2^2)$, and $\epsilon \sim N(0, \sigma^2)$. i is Worker, j is machine, and k is replicate.

```
> fm1<-lmer(score~Machine-1+(1|Worker/Machine),
  data=Machines)
```

The variance components in fm1:

Comp.	Groups	Name	Var
$\hat{\sigma}_2^2$	Machine:Worker	(Int)	13.909
$\hat{\sigma}_1^2$	Worker	(Int)	22.858
$\hat{\sigma}^2$	Res		0.925

Number of obs: 54, groups: Machine:Worker, 18; Worker, 6.

For observations on Worker i,

$$Var(Y_{ijk}) = \sigma_1^2 + \sigma_2^2 + \sigma^2 \tag{14}$$

Variance between machines within workers:

$$Covar(Y_{ijk}, Y_{ijk'}) = \sigma_1^2 + \sigma_2^2 \tag{15}$$

Variance between workers:

$$Covar(Y_{ijk}, Y_{ij'k'}) = \sigma_1^2 \tag{16}$$

Note:

1. $\hat{\sigma}_1^2$ all observations have the same variance;
2. $\hat{\sigma}_2^2$: the covariance between observations corresponding to the same worker using different machines is the same, for any pair of machines.

```
> ranef(fm1)
$`Machine:Worker`    $Worker
    (Intercept)
A:6      1.91609     6 -7.514666
A:2      1.55253     2 -1.375925
```

In this model, the sum of the random effects for Worker 1 on Machine A is
$s_1 = b_1 + b_{11}$

```
> ranef(fm1)
...
     $`Machine:Worker`       $Worker
        (Intercept)
s1 = A:1    -0.75012 +       1.044598 = 0.29448
```

and for Worker 1 on machine B,
$s_2 = b_1 + b_{21}$.

```
> ranef(fm1)
...
    $`Machine:Worker`       $Worker
     (Intercept)
s2 = B:1     1.50002 +       1.044598 =  2.5446
```

For all Workers and machines, we can obtain these random effects s from this matrix:

```
> mat<-matrix(
     unlist(ranef(fm1)$`Machine:Worker`),6,3)
> +
     matrix(unlist(ranef(fm1)$Worker),6,3)
          [,1]      [,2]      [,3]
[1,] -7.514666 -7.514666 -7.514666
[2,] -1.375925 -1.375925 -1.375925
[3,] -0.059823 -0.059823 -0.059823
[4,]  1.044598  1.044598  1.044598
[5,]  5.361045  5.361045  5.361045
[6,]  2.544771  2.544771  2.544771
> mat
         A        B        C
6  1.91609 -8.97590  2.48677
2  1.55253  0.60682 -2.99667
4 -1.03937  2.41736 -1.41440
```

```
1 -0.75012  1.50002 -0.11421
3  1.77775  2.29952 -0.81481
5 -3.45687  2.15218  2.85331
```

Using lmer, we have b_i and b_{ij} independent, but s_1 and s_2 are correlated via the common term b_1. We can recover the correlations between machine through the vcov matrix of the random effects (BLUPs) (**but note that we never see this in the lmer output—what's the significance of the fact that these are correlated?**):

```
> var(mat)
        A       B       C
A  4.5670 -4.6492 -1.9288
B -4.6492 19.7897 -4.6925
C -1.9288 -4.6925  5.1966
```

1.5. **Varying intercepts and slopes (no correlation).**

$$Y_{ijk} = \beta_j + b_{ij} + \epsilon_{ijk} \tag{17}$$

```
> fm3<-lmer(score~Machine-1+
              (Machine-1|Worker),
            data=Machines)
> ranef(fm3)
$Worker
  MachineA  MachineB MachineC
6 -5.59160 -16.58381  -5.0305
2  0.18387  -0.80332  -4.2823
4 -1.02388   2.32846  -1.4144
1  0.31199   2.55323   0.9304
3  6.96922   7.77935   4.4733
5 -0.84961   4.72610   5.3235
```

The random effects for Worker 1 on Machine A is
$s_1 = b_{11} = 0.31199$
and for Worker 1 on Machine B,
$s_2 = b_{12} = 2.55323$.

The 'Machine independent' Worker random effect (varying intercept) b_i has been dropped. We have b_{11} correlated with b_{12}. We can see this when we recover the (co-)variances between machines from the random effects:

```
> var(ranef(fm3)$Worker)
         MachineA MachineB
MachineA   16.347   28.239
MachineB   28.239   74.093
MachineC   11.146   29.181
         MachineC
```

```
MachineA   11.146
MachineB   29.181
MachineC   18.972
```

Also, the variances for each machine (16, 74, 18) are also allowed to be different. Here are the variance components:

Comp.	Groups	Name	Variance	Corr$_{1,\cdot}$	Corr$_{2,\cdot}$
$\hat{\sigma}^2_{j=1}$	Worker	A	16.640		
$\hat{\sigma}^2_{j=2}$		B	74.395	$\hat{\rho}_{1,2}$ 0.803	
$\hat{\sigma}^2_{j=3}$		C	19.268	$\hat{\rho}_{1,3}$ 0.623	$\hat{\rho}_{2,3}$ 0.771
$\hat{\sigma}^2$	Res		0.925		

$$Var(Y_{ijk}) = \sigma_j^2 + \sigma^2 \tag{18}$$

$$Covar(Y_{ijk}, Y_{ijk'}) = \sigma_j^2 \tag{19}$$

$$Covar(Y_{ijk}, Y_{ij'k'}) = \rho_{j,j'}\sigma_j\sigma_{j'} \tag{20}$$

Note that the BLUPs' vcov matrix reflects the estimated values:

```
> diag(var(ranef(fm3)$Worker))

MachineA MachineB MachineC
  16.347   74.093   18.972

> cor(ranef(fm3)$Worker)

         MachineA MachineB
MachineA  1.00000  0.81141
MachineB  0.81141  1.00000
MachineC  0.63292  0.77832
         MachineC
MachineA  0.63292
MachineB  0.77832
MachineC  1.00000

> # look at the fm3 output
> ## (the random effects table)
```

1. $\hat{\sigma}_j^2$ the variance of an observation depends on the machine being used;
2. $\rho_{j,j'}\sigma_j\sigma_{j'}$ the covariance between observations corresponding to the same worker using different machines is different, for different pairs of machines.

```
> var(ranef(fm3)$Worker)
         MachineA MachineB MachineC
MachineA   16.347   28.239   11.146
MachineB   28.239   74.093   29.181
MachineC   11.146   29.181   18.972
```

$$\begin{pmatrix} \sigma_A^2 & Cov_{A,B} & Cov_{A,C} \\ & \sigma_B^2 & Cov_{B,C} \\ & & \sigma_C^2 \end{pmatrix} \tag{21}$$

Note that, for given machines j and j', say A, B:
$Covar(Y_{ijk}, Y_{ij'k'}) = Cov_{A,B} = 28.239 \approx \rho_{A,B}\sigma_A\sigma_B = .803\times\sqrt{16.347}\times\sqrt{74.093} =$ 27.946.

1.6. **Comparing fm1 and fm3.** The sum of fm1's (Worker/Machine) ranefs ($b_{ij} + b_i$) are roughly the same as fm3's (Machine-1| Worker) random effects b_{ij} for each machine. **In other words, the random effect b_i is folded into b_{ij} in fm3.**

```
> #fm1's ranefs summed up are
> ## roughly the same as the fm3 ranefs:
> matrix(unlist(ranef(fm1)$`Machine:Worker`),6,3) +
   matrix(unlist(ranef(fm1)$Worker),6,3)
          [,1]      [,2]     [,3]
[1,] -5.59858 -16.49057 -5.02789
[2,]  0.17661  -0.76911 -4.37259
[3,] -1.09920   2.35754 -1.47422
[4,]  0.29448   2.54462  0.93039
[5,]  7.13879   7.66056  4.54624
[6,] -0.91210   4.69695  5.39808
> ranef(fm3)
$Worker
  MachineA  MachineB MachineC
6 -5.59160 -16.58381  -5.0305
2  0.18387  -0.80332  -4.2823
4 -1.02388   2.32846  -1.4144
1  0.31199   2.55323   0.9304
3  6.96922   7.77935   4.4733
5 -0.84961   4.72610   5.3235
```

2. How the random effects are 'predicted' when using the ranef() command (section 4.4.3).

In linear mixed models, we fit models like these (the Ware-Laird formulation–see Pinheiro and Bates 2000, for example):

$$Y = X\beta + Zu + \epsilon \tag{22}$$

Let $u \sim N(0, \sigma_u^2)$, and this is independent from $\epsilon \sim N(0, \sigma^2)$.

Given Y, the "minimum mean square error predictor" of u is the conditional expectation:

$$\hat{u} = E(u \mid Y) \tag{23}$$

We can find $E(u \mid Y)$ as follows. We write the joint distribution of Y and u as:

$$\begin{pmatrix} Y \\ u \end{pmatrix} = N\left(\begin{pmatrix} X\beta \\ 0 \end{pmatrix}, \begin{pmatrix} V_Y & C_{Y,u} \\ C_{u,Y} & V_u \end{pmatrix} \right) \tag{24}$$

$V_Y, C_{Y,u}, C_{u,Y}, V_u$ are the various variance-covariance matrices. It is a fact (need to track this down) that

$$u \mid Y \sim N(C_{u,Y} V_Y^{-1}(Y - X\beta)), Y_u - C_{u,Y} V_Y^{-1} C_{Y,u}) \tag{25}$$

This apparently allows you to derive the BLUPs:

$$\hat{u} = C_{u,Y} V_Y^{-1}(Y - X\beta)) \tag{26}$$

Substituting $\hat{\beta}$ for β, we get:

$$BLUP(u) = \hat{u}(\hat{\beta}) = C_{u,Y} V_Y^{-1}(Y - X\hat{\beta})) \tag{27}$$

Here's an example with R:

```
> # Calculate the predicted random effects by hand for the ergoStool data
> (fm1<-lmer(effort~Type-1 + (1|Subject),ergoStool))
Linear mixed model fit by REML
Formula: effort ~ Type - 1 + (1 | Subject)
   Data: ergoStool
 AIC BIC logLik deviance REMLdev
 133 143  -60.6      122     121
Random effects:
 Groups   Name        Variance
 Subject  (Intercept) 1.78
 Residual             1.21
 Std.Dev.
 1.33
 1.10
Number of obs: 36, groups: Subject, 9

Fixed effects:
       Estimate Std. Error t value
TypeT1    8.556      0.576    14.8
```

```
TypeT2   12.444      0.576    21.6
TypeT3   10.778      0.576    18.7
TypeT4    9.222      0.576    16.0

Correlation of Fixed Effects:
       TypeT1 TypeT2 TypeT3
TypeT2 0.595
TypeT3 0.595  0.595
TypeT4 0.595  0.595  0.595
> ## Here are the BLUPs we will estimate by hand:
> ranef(fm1)
$Subject
  (Intercept)
1  1.7088e+00
2  1.7088e+00
3  4.2720e-01
4 -8.5439e-01
5 -1.4952e+00
6 -1.3546e-14
7  4.2720e-01
8 -1.7088e+00
9 -2.1360e-01
> ## this gives us all the variance components:
> VarCorr(fm1)
$Subject
            (Intercept)
(Intercept)      1.7755
attr(,"stddev")
(Intercept)
     1.3325
attr(,"correlation")
            (Intercept)
(Intercept)           1

attr(,"sc")
[1] 1.1003
> # First, calculate the predicted random effect for subject 1:
>
> ## The variance for the random effect subject is the term C_{u,Y}:
> covar.u.y<-VarCorr(fm1)$Subject[1]
> # Estimated covariance between u_1 and Y_1
```

```
> ## make up a var-covar matrix from this:
> (cov.u.Y<-matrix(covar.u.y,1,4))
       [,1]   [,2]   [,3]   [,4]
[1,] 1.7755 1.7755 1.7755 1.7755
> # Estimated variance matrix for Y_1
> (V.Y<-matrix(1.7755,4,4)+diag(1.2106,4,4))
       [,1]   [,2]   [,3]   [,4]
[1,] 2.9861 1.7755 1.7755 1.7755
[2,] 1.7755 2.9861 1.7755 1.7755
[3,] 1.7755 1.7755 2.9861 1.7755
[4,] 1.7755 1.7755 1.7755 2.9861
> # Extract observations for subject 1
> (Y<-matrix(ergoStool$effort[1:4],4,1))
     [,1]
[1,]   12
[2,]   15
[3,]   12
[4,]   10
> # Estimated fixed effects
> (beta.hat<-matrix(fixef(fm1),4,1))
        [,1]
[1,]  8.5556
[2,] 12.4444
[3,] 10.7778
[4,]  9.2222
> # Predicted random effect
> cov.u.Y %*% solve(V.Y)%*%(Y-beta.hat)
       [,1]
[1,] 1.7087
> # Compare with ranef command
> ranef(fm1)$Subject[1,1]
[1] 1.7088
> # Calculate predicted random effects for all subjects
> t(cov.u.Y %*% solve(V.Y)%*%(matrix(ergoStool$effort,4,9)-matrix(fixef(fm1),4
             [,1]
 [1,]  1.7087e+00
 [2,]  1.7087e+00
 [3,]  4.2717e-01
 [4,] -8.5435e-01
 [5,] -1.4951e+00
```

```
 [6,] -1.3906e-14
 [7,]  4.2717e-01
 [8,] -1.7087e+00
 [9,] -2.1359e-01
> ranef(fm1)
$Subject
  (Intercept)
1  1.7088e+00
2  1.7088e+00
3  4.2720e-01
4 -8.5439e-01
5 -1.4952e+00
6 -1.3546e-14
7  4.2720e-01
8 -1.7088e+00
9 -2.1360e-01
```

3. Correlations of fixed effects

For an ordinary linear model, the covariance matrix (from which we can get the correlation matrix) of $\hat{beta}$ is

$$\sigma^2 \times (X^T X)^{-1}. \tag{28}$$

For a mixed effects model, the standard deviations (standard errors) and correlations for the fixed effects estimators are listed at the end of the lmer output.

```
> lm.full<-lmer(wear~material-1+(1|Subject), data = BHHshoes)
```

The estimated correlation between $\hat{beta}_1$ and $\hat{beta}_2$ is 0.988. In this case, we have simple forms for the parameter estimators:

$$\hat{\beta}_1 = (Y_{1,1} + Y_{2,1} + \cdots + Y_{10,1})/10 \tag{29}$$

$$\hat{\beta}_2 = (Y_{1,2} + Y_{2,2} + \cdots + Y_{10,2})/10 \tag{30}$$

```
> b1.vals<-subset(BHHshoes,material=="A")$wear
> b2.vals<-subset(BHHshoes,material=="B")$wear
> vcovmatrix<-var(cbind(b1.vals,b2.vals))
> covar<-vcovmatrix[1,2]
> sds<-sqrt(diag(vcovmatrix))
> covar/(sds[1]*sds[2])
b1.vals
0.98823
```

```
> #cf:
> covar/((0.786*sqrt(10))^2)
[1] 0.98752
```

In a regular linear model version, we would have had:

```
> fm.lm<-lm(wear~material-1,BHHshoes)
> X<-model.matrix(fm.lm)
> 2.49^2*solve(t(X)%*%X)
          materialA materialB
materialA   0.62001   0.00000
materialB   0.00000   0.62001
```

because $Var(\hat{\beta}) = \hat{\sigma}^2(X^TX)^{-1}$.

From this, see if you can work out the covariance, and where the estimated correlation comes from, using the remainder of the lmer output above.

```
> b1.diffs<-b1.vals-mean(b1.vals)
> b2.diffs<-b2.vals-mean(b2.vals)
> b1.diffs<-b1.vals-mean(BHHshoes$wear)
> b2.diffs<-b2.vals-mean(BHHshoes$wear)
> covar<-t(b1.diffs)%*%b2.diffs
> b1.sd<-sd(b1.vals)
> b2.sd<-sd(b2.vals)
> corr<-covar/(b1.sd*b2.sd)
```

How does this work for multiple factors?

```
> m1<-lmer(effort~Type-1+(1|Subject),ergoStool)
> T1.vals<-subset(ergoStool,Type=="T1")$effort
> T2.vals<-subset(ergoStool,Type=="T2")$effort
> T3.vals<-subset(ergoStool,Type=="T3")$effort
> T4.vals<-subset(ergoStool,Type=="T4")$effort
> vals<-cbind(T1.vals,T2.vals,T3.vals,T4.vals)
> ## compute variance covariance matrix:
> vcovmat<-var(vals)
> ## get sd's of each level:
> sds<-sqrt(diag(vcovmat))
> ## T1.T2 correlation, the sds come from the model fit:
> 1.7222/(1.728*1.728)
[1] 0.57676
```

Note: Not sure if the above is correct (the case of multiple levels in a factor).

4. σ_b^2 describes both between-block variance, and within block covariance

Consider the following model:

(31) $$Y_{ij} = b_i + e_{ij},$$

with $b_i \sim N(0, \sigma_b^2)$, e_{ij} $N(0, \sigma^2)$.

Now try this in R (corresponding to $\sigma = 1, \sigma_b = 100, i = 1, 2, 3$ and $j = 1, 2, 3$):

```
> block<-gl(3,3)
> ## very small within group:
> eij<-rnorm(9,0,1)
> ## very high between group variance:
> ei<-rnorm(3,0,100)
> y<-rep(ei,each=3)+eij
> plot(block,y)
> fm1<-lm(y~1)
> aggregated<-tapply(y,block,mean)
> agg.data<-data.frame(means=aggregated,block=factor(1:3))
> fm1a<-lm(y~1,agg.data)
> fm3<-lmer(y~1+(1|block))
> a<-y[c(1,4,7)]
> b<-y[c(1,4,7)+1]
> c<-y[c(1,4,7)+2]
> (cov(a,b)+cov(a,c)+cov(b,c))/3
[1] 5664.2
> ## more like what we experience:
> block<-gl(3,3)
> ## large within group:
> eij<-rnorm(9,0,100)
> ## small between group:
> ei<-rnorm(3,0,1)
> y<-rep(ei,each=3)+eij
> plot(block,y)
> fm1<-lm(y~1)
> aggregated<-tapply(y,block,mean)
> agg.data<-data.frame(means=aggregated,block=factor(1:3))
> fm1a<-lm(y~1,agg.data)
> fm3<-lmer(y~1+(1|block))
```

Perhaps it's just worth remembering that a variance is a covariance of a random variable with itself, and then consider the model formulation. If we have

(32) $$Y_{ij} = \mu + b_i + \epsilon_{ij}$$

where i is the group, j is the replication, if we *define* $b_i \sim N(0, \sigma_b^2)$, and refer to σ_b^2 as the between group variance, then we must have

(33)

$$
\begin{aligned}
Cov(Y_{i1}, Y_{i2}) =& Cov(\mu + b_i + \epsilon_{i1}, \mu + b_i + \epsilon_{i2}) \\
=& \underset{\substack{\uparrow \\ =0}}{Cov(\mu, \mu)} + \underset{\substack{\uparrow \\ =0}}{Cov(\mu, b_i)} + \underset{\substack{\uparrow \\ =0}}{Cov(\mu, \epsilon_{i2})} + \underset{\substack{\uparrow \\ =0}}{Cov(b_i, \mu)} + \underset{\substack{\uparrow \\ +ve}}{Cov(b_i, b_i)} \dots \\
=& Cov(b_i, b_i) = Var(b_i) = \sigma_b^2
\end{aligned}
$$