1. Return to the (mixed model) thermocouple calibration problem of HW8. In HW8 you computed likelihood values for two different vectors of variance components. Here let's use an optimizer in R and find both the MLE and the REML estimates of \( (\sigma^2_y, \sigma^2_\delta, \sigma^2) \).

a) Add the MASS and nml2 packages to your R session. Then type

```r
> I3<-diag(c(1,1,1))
> J3<-matrix(c(1,1,1,1,1,1,1,1,1),3,3)
> M<-matrix(c(0,0,0,0,1,4,0,4,16),3,3)
> X<-matrix(c(rep(1,9),rep(c(0,1,4),3)),9,2)
> Y<-matrix(c(99.8,108.1,136.0,100.3,109.5,137.7,98.3,110.1,142.2),9,1)
```

Using these, you can create a negative profile loglikelihood function for \( (\sigma^2_y, \sigma^2_\delta, \sigma^2) \) as below. (We're using a negative profile loglikelihood here because the R optimizer seeks to minimize rather than maximize a function.)

```r
> minuslLstar<-function(s2,Y)
+ { 
+  temp0<-kronecker(I3,((s2[1]*J3)+(s2[2]*M)+(s2[3]*(I3))))
+  temp1<-ginv(temp0)
+  temp2<-X%*%ginv(t(X))%^%temp1%^%X%^%t(X)%^%temp1%^%Y
+  temp3<=-(Y-temp2)
+  (.5*(log(det(temp0))))+ 
+  (4.5*log(2*pi))+ 
+  (.5*(t(temp3)%^%temp1%^%temp3))
+ }
```

Evaluate this function for the two sets of variance components in parts c) and d) of the problem from HW8 as below (and verify that your answers are consistent with what you got on the previous homework).

```r
> minuslLstar(c(1,1,1),Y)
> minuslLstar(c(1,1,.25),Y)
```

I did some "hunt and peck"/"by hand" optimization of this function, and came up with what we'll use as starting values of \( \sigma^2_y = .07, \sigma^2_\delta = .45, \) and \( \sigma^2 = .37 \). Use the R function optim to find better values (MLEs) by typing

```r
> optim(c(.07,.45,.37),minuslLstar,Y=Y,hessian=TRUE)
```

This will set in motion an iterative optimization procedure with the starting value above. This call produces a number of interesting summaries, including a matrix of second partials of the objective function evaluated at the optimum (the hessian). What are the MLEs?

b) Now consider seeking REML estimates of the variance components. Compute the projection matrices \( P_x \) and \( N = I - P_x \). Notice that since \( (I - P_x)X = 0 \), every row of \( N \) is (after transposing) in \( C(X')^\perp \) and so any set of 7 rows of \( N \) that make a \( 7 \times 9 \) matrix, say \( B \), of rank 7
will serve to create the vector of "error contrasts" $\mathbf{r} = \mathbf{BY}$ which one uses to find REML estimates here. Verify that the first 7 rows of $\mathbf{N}$ will work in the present situation by typing

```
> qr(B)$rank
```

Note that by construction, $\mathbf{r} \sim \text{MVN}_r\left(\mathbf{0}, \mathbf{B}\left(\begin{pmatrix} \sigma_\gamma^2, \\ \sigma_\delta^2, \\ \sigma_\sigma^2 \end{pmatrix}\right)\mathbf{B}'\right)$ and "REML" is maximum likelihood for the variance components based on $\mathbf{r}$. You may create a negative loglikelihood function here as

```
> minuslLstar2<-function(s2,Y,B){
+ temp0<-kronecker(I3, ((s2[1]*J3)+(s2[2]*M)+(s2[3]*(I3))))
+ temp1<-B%*%temp0%*%t(B)
+ temp2<-ginv(temp1)
+ temp3<-B%*%Y
+ (.5*(log(det(temp1)))) +
+ (3.5*log(2*pi)) +
+ (.5*(t(temp3)%*%temp2%*%temp3))
+ }
```

Then you may find REML estimates of the variance components by optimizing this function. Use the starting values from part a) and type

```
> optim(c(.07,.45,.37), minuslLstar2, Y=Y, B=B, hessian=TRUE)
```

How do your estimates here compare to the MLEs you found in part a)? (Statistical folklore says that usually REML estimates are bigger than MLEs.)

c) The (unavailable) BLUE of any (vector of) estimable function(s) $\mathbf{C}\beta$ is the generalized least squares estimator $\mathbf{C}\hat{\beta} = \mathbf{C}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}$. Use first maximum likelihood and then REML estimates of the variance components to estimate $\mathbf{V}$ and then produce realizable estimates of the vector of fixed effects $(\alpha, \beta)'$ as $\mathbf{C}\hat{\beta} = \mathbf{C}\left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{Y}$ for the $\mathbf{C} = \mathbf{I}$ case.

d) As on panel 811 of Koehler's notes, $\mathbf{C}\beta$ has covariance matrix $\mathbf{C}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{C}'$ and so the covariance matrix of $\mathbf{C}\beta$ from part c) might be estimated as $\mathbf{C}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{C}'$. Use this fact to produce very approximate standard errors for your estimates of fixed effects in part c).

e) If $\hat{\mathbf{Y}}^*(\sigma^2)$ is the (unavailable) generalized least squares estimate of the mean of $\mathbf{Y}$, the BLUP of the vector of random effects is

$$
\hat{\mathbf{u}} = \mathbf{GZ}\mathbf{V}^{-1}\left(\mathbf{Y} - \hat{\mathbf{Y}}^*(\sigma^2)\right)
= \mathbf{GZ}\mathbf{V}^{-1}\left(\mathbf{I} - \mathbf{X}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{V}^{-1}\right)\mathbf{Y}
$$
Letting $B = X\left(X'V^{-1}X\right)^{-1}X'V^{-1}$ and $P = V^{-1}(I-B)$, this is

$$\hat{u} = GZ'V^{-1}(I-B)Y = GZ'PY$$

Estimates of variance components lead not only to an estimate of $V$, but also to estimates of $G$ and $P$. Use first maximum likelihood and then REML estimates of variance components to approximate the BLUP as

$$\hat{u} = GZ'PY$$

f) As the BLUP $\hat{u}$ is an (admittedly somewhat unpleasant) matrix times $Y$, it is possible to work out a prediction variance for it,

$$\text{Var}(\hat{u} - u) = G - GZ'PZG$$

Once again, estimates of variance components lead to very approximate standard errors of the available predictions $\hat{u}$ based on diagonal elements of this matrix. Find these using first maximum likelihood and then REML estimates of variance components.

g) The handout on BLUPs discusses prediction of a quantity

$$l = c'\beta + s'u$$

for an estimable $c'\beta$. In light of that material, consider again the original context of calibration of the thermocouples. Use first MLEs of variance components and then REML estimates to find predictors for

$$\alpha_i = \alpha + \gamma_i \text{ and } \beta_i = \beta + \delta_i$$

and standard errors for your predictions. (These are the intercept and slope for the calibration equation for the $1^\text{st}$ thermocouple.)

2. Let's use the R mixed model routine to attempt some of what we've already done "by hand" with this calibration problem. With the \texttt{nlme} package added to your R session, make a file called (say) \texttt{thermocouples.txt} that looks like

```
y group temp
99.8 1 0
108.1 1 1
136.0 1 4
100.3 2 0
109.5 2 1
137.7 2 4
98.3 3 0
110.1 3 1
142.2 3 4
```

and read it into your R session as

```r
c <- read.table("thermocouples.txt ", header=T)
```

Then prepare the data table for use in the fitting of a mixed model by typing

```r
gd <- groupedData(y~temp|group, data=thermo)
```

Then issue commands that will do model fitting via maximum likelihood and REML
> fm1 <- lme(y~temp, data=gd, random=~1+temp | group, method="ML")
> fm2 <- lme(y~temp, data=gd, random=~1+temp | group, method="REML")

As it turns out, if you examine the fitted model objects fm1 and fm2 you will come to the conclusion that something other than what you've found to this point has been computed. This is because these calls allow each \((\gamma_i, \delta_i)\) pair to have a general covariance matrix, and the model we've been using specified that \(\text{Cov} \left( \gamma_i, \delta_i \right) = 0\). So we need to modify the fitting to agree with our modeling. Type

> fm3 <- update(fm1, random=pdDiag(~1+temp), method="ML")
> fm4 <- update(fm2, random=pdDiag(~1+temp), method="REML")

I believe that these calls create fitted model objects that correspond to the analysis we've done thus far. You may see some of what these calls have created by typing

> summary(fm3)
> summary(fm4)

But much more than this is available. Type

> ?nlmeObject

to see exactly what is available. Then apply all of the functions random.effects(), fixed.effects(), predict(), coefficients(), and intervals() to both fitted model objects. What parts of what these calls produce have you calculated in Problem 1 of this assignment and on HW8?

What do the results of the intervals() calls suggest about what this very small data set provides in terms of precision of estimation of the variance components? (You should have learned in a basic course that sample sizes required to really pin down variances or standard deviations are typically huge. Notice, for example, that we have only 3 thermocouples ... a sample of size 3 for estimating \(\sigma_\gamma^2\) and \(\sigma_\delta^2\).)