Statistics 312/612, fall 2016
Homework # 7
Due: Wednesday 2 November

For this homework I want you to determine all the orthogonal subspaces of $\mathbb{R}^{96}$ that correspond to the calculations in the paper of Eden and Fisher (1927). You can get clues by making numerical calculations, but I am asking for more than numerical confirmation.

Remember that $*$ means componentwise multiplication in **R**. For example $c(1,4,9)* c(2,3,4) = (2,12,36)$.

```
EF <- read.csv("EF.csv")
```

You will notice that I have used Helmert contrasts, so that the fit `lm(grain~block+treat,EF)` corresponds to estimation for the model

$$\mathbb{E}(grain \mid fert = i, conc = j, cat = k) = \mu + \phi_i + \gamma_j + \kappa_k,$$

with $\phi_M + \phi_S = 0$, and so on.

First create a bunch of dummy variables and differences of dummies:

```
fM <- (EF$fert == "M")+0; fS <- (EF$fert == "S")+0
a2 <- (EF$amount == "d")+0; a1 <- (EF$amount == "s")+0
tE <- (EF$time == "E")+0; tL<- (EF$time == "L")+0
notreat <-(EF$treat == "none")+0
ff <- fM - fS; aa <- a2 - a1; tt <- tL -tE

DUM <- cbind(notreat,fM,fS,a2,a1,tL,tE, ff,aa,tt)
B <- outer(EF$block, levels(EF$block), "==") + 0
dimnames(B)[[2]] <- levels(EF$block)
nfat <- cbind(n= sqrt(3)*(notreat - 1/3),
    ff , aa,tt, fa=ff*aa,ft=ff*tt,at=aa*tt,fat=ff*aa*tt)/8
# As a check:
#t(DUM) %*% DUM
#t(B) %*% DUM
```

[1] Here is the useful part of the anova table for the fit `lm(grain~block+treat,EF)`:

```
##           Df    Sum Sq   Mean Sq  F value       Pr(>F)
## block      7 2286.4395 326.63421 7.961986 2.617151e-07
## treat      8  387.0135  48.37669 1.179223 3.219950e-01
## Residuals 80 3281.9370  41.02421       NA           NA
```

The component labelled "block" comes comes from the subspace $\mathcal{X}_B$ of the eight-dimensional subspace spanned by the columns of B that is orthogonal to $\mathbb{1}_{96}$. The component labelled "treat" comes from the subspace $\mathcal{X}_{treat}$ of the nine-dimensional space

$$\mathcal{X}_9 = \text{span}\{notreat, fM * a2 * tE, fM * a2 * tL, \ldots, fS * a1 * tL\}$$

that is orthogonal to $\mathbb{1}_{96}$. Every vector in the model space $\mathcal{X}$, which has dimension 16, is a sum of a part in span$(\mathbb{1})$, a part in $\mathcal{X}_B$, and a part in $\mathcal{X}_{treat}$.

(i) (10 points) Explain why all the columns of `DUM` belong to $\mathcal{X}_9$. (Here and for subsequent questions I want a mathematical argument, not just a numerical check. You do not need to argue separately for all 10 columns, as long as it is clear that your idea works more generally.)

(ii) (5 points) Explain why $fM + fS = a2 + a1 = tE + tL = \mathbb{1}_{96} - notreat$.

(iii) (10 points) Explain why all the columns of `nfat` are orthogonal to all the columns of `B`. Hint: Inner products between dummy variables reduce to counting.

(iv) (10 points) Explain why the columns of `nfat` form an orthonormal basis for $\mathcal{X}_{treat}$.

(v) (20 points) Explain how the columns of `nfat` are related to the eight "contrasts" listed in Table III.

[2] E&F$_{559}$ spoke of "two independent estimates of error", which suggests that they were working with two orthogonal subspaces of $\mathcal{X}^{\perp}$.

(i) (10 points) The vector `notreat` can be written as a sum of unit vectors $n_1 + \cdots + n_{32}$ that indicate where the untreated plots lie. For example, n_1, n_2, n_3, n_4 identitfy the four untreated plots in block I. Show that the vectors $n_i - B_I * notreat/4$ span a three-dimension space that is orthogonal to $\mathcal{X}$.

(ii) (10 points) Identify the 24-dimensional subspace of $\mathcal{X}^{\perp}$ that contributed the 24 degrees of freedom in Table IV.

(iii) (extra credit) Find an orthogonal basis for 56-dimensional subspace that contributed the 56 degrees of freedom in Table IV. Make sure you explain why it is orthogonal to the subspace in the previous quesion. Hint:

```
anova(lm (grain ~ block * treat,EF))

## Analysis of Variance Table
##
## Response: grain
##             Df  Sum Sq Mean Sq F value    Pr(>F)
## block        7 2286.44  326.63 10.1385 7.354e-06 ***
## treat        8  387.01   48.38  1.5016    0.2087
## block:treat 56 2508.72   44.80  1.3905    0.1892
## Residuals   24  773.21   32.22
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

[3] (extra credit) Explain where Table VI came from. Is it asserting that some differences are significant? If so, please explain.

References

Eden, T. and R. A. Fisher (1927, 10). Studies in crop variation IV: The experimental determination of the value of top dressings with cereals. *The Journal of Agricultural Science 17*(4), 548–562.