Durham/Newcastle Statistics Graduate Course
Design of Experiments
March 2012

Outline solutions

Example Sheet 1

Q1 Prove the results concerning projections quoted in the lecture, namely

(i) For any vectors z and y we have that $P_T z$ and $(I - P_T)y$ are orthogonal, as are $(I-P_T)z$ and $P_T y$, so $z^T P_T^T (I-P_T)y = 0$ and also $z^T(I-P_T^T)P_T y = 0$. Subtracting these we get $z^T(P_T^T - P_T)y = 0$ for all y and z, so $P_T^T = P_T$.

(ii) For any y, $P_T y \in V_T$, so $P_T^2 y = P_T(P_T y) = P_T y$, i.e. $P_T^2 = P_T$.

(iii) From above any eigenvalue must be real and as λ must satisfy $\lambda^2 = \lambda$, $\lambda \in \{0, 1\}$.

(iv) $\mathsf{tr}(P_T) = \sum \lambda$. This is just the number of eigenvalues equal to 1. Now the space spanned by the eigenvectors with eigenvalue 1 is V_T, hence the result.

Q2 This is elementary: $\mathsf{var}(c^T\hat{\tau}_E) = \sigma^2 c^T R^{-1} c$.

Q3 An experiment to compare T treatments is designed in one of two ways.

(i) Take the average of all pairwise comparisons, i.e. the average over all $i \neq j$ of $\sigma^2(r_i^{-1}+r_j^{-1})$. This is just $(T-1)^{-1}\sigma^2\sum r_i^{-1}$. Minimising this subject to $\sum r_j = N$ gives $r_j = \text{constant} = N/T$.

(ii) Here the average to be minimised is $\sigma^2(T-1)^{-1}\sum(r_1^{-1}+r_j^{-1})$, subject to $\sum r_j = N$. Ignoring σ^2 this is $r_1^{-1} + (T-1)^{-1}\sum_{j>1} r_j^{-1}$. Differentiating the Lagrangian $r_1^{-1} + (T-1)^{-1}\sum_{j>1} r_j^{-1} + \lambda\sum r_j$ gives

$$\begin{aligned} r_1^{-2} &= \lambda \\ r_j^{-2} &= (T-1)\lambda \end{aligned}$$

Hence the replications are such that $r_1 = r_j\sqrt{T-1}$, and $r_1(1+\sqrt{T-1}) = N$

Q4 From lecture $\hat{\tau}_j = r_j^{-1}u_j^T P_T y$ where u_j is the element of V_T which is 1 when treatment j is applied and 0 otherwise. The space V_T is spanned by the vectors $< u_1, \ldots, u_T >$. Writing $U = (u_1, \ldots, u_T)$, so $U^T U = R$ where $R = \text{diag}(r_j)$. Now $u_j^T U = (0, \ldots, 1, \ldots, 0)$ where the 1 is in position j. Hence, using this and the fact that $P_T = U(U^T U)^{-1}U^T$, $\hat{\tau}_j = (0, \ldots, 1, \ldots, 0)R^{-1}U^T y$. Now $U^T y$ is a $T \times 1$ vector with i^{th} element equal to $y(i)$, the sum of the y values given the i^{th} treatment. So $\hat{\tau}_j = r_j^{-1}y(j)$.

If the vector subspace W is the range of a matrix X the orthogonal projection onto the range of X is well know to be $X(X^T X)^{-1}X^T$. Now assuming this formula, show that the $\hat{\tau}_j$ derived in lectures are of a familiar and sensible form.

Example Sheet 2

Q1 We can start by defining the outputs in y, the treatment design matrix, X, and the function `proj` which maps a matrix to its associated projection matrix.

```
> y
 [1]  7.8 15.3 18.5  6.2 11.3 24.2 11.2 17.4 29.8  1.8  3.3  7.1
> X
      t0 t1 t2
 [1,]  1  0  0
 [2,]  0  1  0
 [3,]  0  0  1
 [4,]  1  0  0
 [5,]  0  1  0
 [6,]  0  0  1
 [7,]  1  0  0
 [8,]  0  1  0
 [9,]  0  0  1
[10,]  1  0  0
[11,]  0  1  0
[12,]  0  0  1
> proj
function(X){
proj<-X%*%ginv(t(X)%*%X)%*%t(X)
proj
}
>
```

We can form the projection onto V_T, P_T

```
 4*proj(X)
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12]
 [1,]    1    0    0    1    0    0    1    0    0     1     0     0
 [2,]    0    1    0    0    1    0    0    1    0     0     1     0
 [3,]    0    0    1    0    0    1    0    0    1     0     0     1
 [4,]    1    0    0    1    0    0    1    0    0     1     0     0
 [5,]    0    1    0    0    1    0    0    1    0     0     1     0
 [6,]    0    0    1    0    0    1    0    0    1     0     0     1
 [7,]    1    0    0    1    0    0    1    0    0     1     0     0
 [8,]    0    1    0    0    1    0    0    1    0     0     1     0
 [9,]    0    0    1    0    0    1    0    0    1     0     0     1
[10,]    1    0    0    1    0    0    1    0    0     1     0     0
[11,]    0    1    0    0    1    0    0    1    0     0     1     0
[12,]    0    0    1    0    0    1    0    0    1     0     0     1
```

(This is multiplied by 4 simply to make the printout more compact).
We can check the symmetry and idempotence as follows

```
 PT-t(PT)
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12]
 [1,]    0    0    0    0    0    0    0    0    0     0     0     0
 [2,]    0    0    0    0    0    0    0    0    0     0     0     0
 [3,]    0    0    0    0    0    0    0    0    0     0     0     0
 [4,]    0    0    0    0    0    0    0    0    0     0     0     0
 [5,]    0    0    0    0    0    0    0    0    0     0     0     0
 [6,]    0    0    0    0    0    0    0    0    0     0     0     0
 [7,]    0    0    0    0    0    0    0    0    0     0     0     0
 [8,]    0    0    0    0    0    0    0    0    0     0     0     0
 [9,]    0    0    0    0    0    0    0    0    0     0     0     0
[10,]    0    0    0    0    0    0    0    0    0     0     0     0
[11,]    0    0    0    0    0    0    0    0    0     0     0     0
[12,]    0    0    0    0    0    0    0    0    0     0     0     0
> PT%*%PT-PT
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12]
 [1,]    0    0    0    0    0    0    0    0    0     0     0     0
 [2,]    0    0    0    0    0    0    0    0    0     0     0     0
 [3,]    0    0    0    0    0    0    0    0    0     0     0     0
 [4,]    0    0    0    0    0    0    0    0    0     0     0     0
 [5,]    0    0    0    0    0    0    0    0    0     0     0     0
 [6,]    0    0    0    0    0    0    0    0    0     0     0     0
 [7,]    0    0    0    0    0    0    0    0    0     0     0     0
 [8,]    0    0    0    0    0    0    0    0    0     0     0     0
 [9,]    0    0    0    0    0    0    0    0    0     0     0     0
[10,]    0    0    0    0    0    0    0    0    0     0     0     0
[11,]    0    0    0    0    0    0    0    0    0     0     0     0
[12,]    0    0    0    0    0    0    0    0    0     0     0     0
```

The sums of squares can be found as

```
 P0<-proj(rep(1,12))
> PE<-diag(12)-PT
> PW<-PT-P0
> c(t(y)%*%P0%*%y,t(y)%*%PW%*%y,t(y)%*%PE%*%y)
[1] 1973.7675  351.8450  444.1175
> sum(y^2)
[1] 2769.73
> sum(c(t(y)%*%P0%*%y,t(y)%*%PW%*%y,t(y)%*%PE%*%y))
[1] 2769.73
```

The treatment estimates can be found as follows,

```
 t0
 [1] 1 0 0 1 0 0 1 0 0 1 0 0
> t(t0)%*%PT%*%y/sum(t0)
     [,1]
[1,] 6.75
> mean(y[Rx==0])
[1] 6.75,
```

and similarly for the other treatments. Note also that the length of $P_T y$ is $\sqrt{y^T P_T y}$:

```
> t(y)%*%PT%*%y
         [,1]
[1,] 2325.612
> sum((PT%*%y)^2)
[1] 2325.612
```

Q2 Things such as the rank and eigenvalues of the following can be found as follows (this is just for P_E, the values for the other matrices are obtained in the same way.

```
 eigen(PE)$values
 [1] 1.000000e+00 1.000000e+00 1.000000e+00 1.000000e+00 1.000000e+00
 [6] 1.000000e+00 1.000000e+00 1.000000e+00 1.000000e+00 2.288202e-15
[11] 2.071571e-15 8.881784e-16
> qr(PE)$rank
[1] 9
```

Q3 Analysis using '`lm`' is:

```
analysis<-lm(y~factor(Rx))
> anova(analysis)
Analysis of Variance Table

Response: y
           Df Sum Sq Mean Sq F value Pr(>F)
factor(Rx)  2 351.84  175.92  3.5651 0.0724 .
Residuals   9 444.12   49.35
```

The sums of squares are the same as those found in Q1. The following shows the estimate of τ_1, (given as the '`(Intercept)`' because of the default identifiability constraint used by R).

```
Call:
lm(formula = y ~ factor(Rx))
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    6.750      3.512   1.922   0.0868 .
factor(Rx)1    5.075      4.967   1.022   0.3336
factor(Rx)2   13.150      4.967   2.647   0.0266 *
```

Q4 As an example compute $P_T P_W$ and $P_T P_E$.

```
12*PT%*%PW
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12]
 [1,]    2   -1   -1    2   -1   -1    2   -1   -1     2    -1    -1
 [2,]   -1    2   -1   -1    2   -1   -1    2   -1    -1     2    -1
 [3,]   -1   -1    2   -1   -1    2   -1   -1    2    -1    -1     2
 [4,]    2   -1   -1    2   -1   -1    2   -1   -1     2    -1    -1
 [5,]   -1    2   -1   -1    2   -1   -1    2   -1    -1     2    -1
 [6,]   -1   -1    2   -1   -1    2   -1   -1    2    -1    -1     2
 [7,]    2   -1   -1    2   -1   -1    2   -1   -1     2    -1    -1
 [8,]   -1    2   -1   -1    2   -1   -1    2   -1    -1     2    -1
 [9,]   -1   -1    2   -1   -1    2   -1   -1    2    -1    -1     2
[10,]    2   -1   -1    2   -1   -1    2   -1   -1     2    -1    -1
[11,]   -1    2   -1   -1    2   -1   -1    2   -1    -1     2    -1
[12,]   -1   -1    2   -1   -1    2   -1   -1    2    -1    -1     2
```

Notice that this coincides with $12P_W$. Also

```
PT%*%PE
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12]
 [1,]    0    0    0    0    0    0    0    0    0     0     0     0
 [2,]    0    0    0    0    0    0    0    0    0     0     0     0
 [3,]    0    0    0    0    0    0    0    0    0     0     0     0
 [4,]    0    0    0    0    0    0    0    0    0     0     0     0
 [5,]    0    0    0    0    0    0    0    0    0     0     0     0
 [6,]    0    0    0    0    0    0    0    0    0     0     0     0
 [7,]    0    0    0    0    0    0    0    0    0     0     0     0
 [8,]    0    0    0    0    0    0    0    0    0     0     0     0
 [9,]    0    0    0    0    0    0    0    0    0     0     0     0
[10,]    0    0    0    0    0    0    0    0    0     0     0     0
[11,]    0    0    0    0    0    0    0    0    0     0     0     0
[12,]    0    0    0    0    0    0    0    0    0     0     0     0
```

Example Sheet 3

Q1 The inclusion of the block effects can be done as follows.

```
> Blocks
 [1] 1 1 1 2 2 2 3 3 3 4 4 4
> XB
      B1 B2 B3 B4
 [1,]  1  0  0  0
 [2,]  1  0  0  0
 [3,]  1  0  0  0
 [4,]  0  1  0  0
 [5,]  0  1  0  0
 [6,]  0  1  0  0
 [7,]  0  0  1  0
 [8,]  0  0  1  0
 [9,]  0  0  1  0
[10,]  0  0  0  1
[11,]  0  0  0  1
[12,]  0  0  0  1
> PB<-proj(XB)
> PBp<-PB-P0
> PE<-diag(12)-P0-PW-PBp
> c(t(y)%*%P0%*%y,t(y)%*%PW%*%y,t(y)%*%PBp%*%y,t(y)%*%PE%*%y)
[1] 1973.7675  351.8450  369.1825   74.9350
> sum(c(t(y)%*%P0%*%y,t(y)%*%PW%*%y,t(y)%*%PBp%*%y,t(y)%*%PE%*%y))
[1] 2769.73
> sum(y^2)
[1] 2769.73
```

Q2 The orthogonality of $W_T = V_T \cap V_0^{\perp}$ and $W_B = V_B \cap V_0^{\perp}$ can be seen from $P_W P_{Bperp}$, which is

```
 round(PW%*%PBp,6)
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12]
 [1,]    0    0    0    0    0    0    0    0    0     0     0     0
 [2,]    0    0    0    0    0    0    0    0    0     0     0     0
 [3,]    0    0    0    0    0    0    0    0    0     0     0     0
 [4,]    0    0    0    0    0    0    0    0    0     0     0     0
 [5,]    0    0    0    0    0    0    0    0    0     0     0     0
 [6,]    0    0    0    0    0    0    0    0    0     0     0     0
 [7,]    0    0    0    0    0    0    0    0    0     0     0     0
 [8,]    0    0    0    0    0    0    0    0    0     0     0     0
 [9,]    0    0    0    0    0    0    0    0    0     0     0     0
[10,]    0    0    0    0    0    0    0    0    0     0     0     0
```

```
[11,]    0    0    0    0    0    0    0    0    0     0     0     0
[12,]    0    0    0    0    0    0    0    0    0     0     0     0
```

The same result follows for $P_W P_B$. However V_T and V_B are not orthogonal, as the following shows:

```
 PB%*%PT*12
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12]
 [1,]    1    1    1    1    1    1    1    1    1     1     1     1
 [2,]    1    1    1    1    1    1    1    1    1     1     1     1
 [3,]    1    1    1    1    1    1    1    1    1     1     1     1
 [4,]    1    1    1    1    1    1    1    1    1     1     1     1
 [5,]    1    1    1    1    1    1    1    1    1     1     1     1
 [6,]    1    1    1    1    1    1    1    1    1     1     1     1
 [7,]    1    1    1    1    1    1    1    1    1     1     1     1
 [8,]    1    1    1    1    1    1    1    1    1     1     1     1
 [9,]    1    1    1    1    1    1    1    1    1     1     1     1
[10,]    1    1    1    1    1    1    1    1    1     1     1     1
[11,]    1    1    1    1    1    1    1    1    1     1     1     1
[12,]    1    1    1    1    1    1    1    1    1     1     1     1
```

This follows because $P_B P_T = P_B(P0 + P_W) = P_B P0 + 0$. As $1 \in V_B$, we have $P_B 1 = 1$ and $P0 = 1(1^T 1)^{-1} 1^T$ so $P_B P0 = P0$.

Q3 The analysis of the data using ''lm'' is:

```
> anova(lm(y~factor(Rx)+factor(Blocks)))
Analysis of Variance Table

Response: y
               Df Sum Sq Mean Sq F value   Pr(>F)
factor(Rx)      2 351.84  175.92 14.0860 0.005413 **
factor(Blocks)  3 369.18  123.06  9.8534 0.009819 **
Residuals       6  74.93   12.49
```

which agrees with the results in Q1. However, note that $r_j^{-1} u_j P_T y$ does not agree with the estimates from ''lm'.

```
 t(t0)%*%PT%*%y/sum(t0)
     [,1]
[1,] 6.75
> t(t1)%*%PT%*%y/sum(t1)
       [,1]
[1,] 11.825
> t(t2)%*%PT%*%y/sum(t2)
```

```
     [,1]
[1,] 19.9

summary(lm(y~factor(Rx)+factor(Blocks)))

Call:
lm(formula = y ~ factor(Rx) + factor(Blocks))

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      7.79167    2.49892   3.118  0.02064 *
factor(Rx)1      5.07500    2.49892   2.031  0.08856 .
factor(Rx)2     13.15000    2.49892   5.262  0.00190 **
factor(Blocks)2  0.03333    2.88550   0.012  0.99116
factor(Blocks)3  5.60000    2.88550   1.941  0.10034
factor(Blocks)4 -9.80000    2.88550  -3.396  0.01456 *
```

but note that differences $\hat{\tau}_i - \hat{\tau}_j$ are consistent. Note also that the residual sum of squares is much reduced because of the assignment of the between-patient differences to the Blocks sum of squares.

Q4 The sum of squares is simply $y^T P_{\text{cont}} y$ where P_{cont} is the projection matrix generated by contrast vector c. Of course $c \in \mathbb{R}^N$ so we use $c = (-1, 0, 1, -1, 0, 1, -1, 0, 1, -1, 0, 1)^T$. We get the Sum of Squares to be 345.85.

Q5 This is easy because the default parameterisation of R uses the constraint $\tau_1 = 0$, so 13.15 is the estimate of $\tau_3 - \tau_1$ and the 95% confidence interval is

```
 lm(y~factor(Rx)+factor(Blocks))$coefficients[3]+c(-1,1)*2.49892
[1] 10.65108 15.64892
```

Q6 Treatment by block incidence matrix $A = T^T B$, which record how many times each treatment occurs in each block. Also AA^T is a $t \times t$ matrix which records how often pairs of treatments occur together in the same block. Here these are

```
 t(T)%*%B
   b1 b2 b3 b4 b5 b6 b7
t1  1  0  0  0  1  0  1
t2  1  1  0  0  0  1  0
t3  0  1  1  0  0  0  1
t4  1  0  1  1  0  0  0
t5  0  1  0  1  1  0  0
t6  0  0  1  0  1  1  0
> TB<-t(T)%*%B
> TB%*%t(TB)
   t1 t2 t3 t4 t5 t6
```

```
t1  3  1  1  1  1  1
t2  1  3  1  1  1  1
t3  1  1  3  1  1  1
t4  1  1  1  3  1  1
t5  1  1  1  1  3  1
t6  1  1  1  1  1  3
```

For this design $P_B P_T$ is not $P0$ due to the non-orthogonal nature of the design.

Q7 The key elements of the output that you need to work with are in

```
summary(analysis)

Call:
lm(formula = y ~ factor(Rx) + factor(Site) + factor(Subjects))

Residuals:
    Min      1Q  Median      3Q     Max
-1.2333 -0.4000 -0.0500  0.4375  1.1500

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        7.28333    0.54016  13.484 1.69e-11 ***
factor(Rx)2        0.21667    0.46780   0.463  0.64825
factor(Rx)3        0.33333    0.46780   0.713  0.48435
factor(Rx)4        0.36667    0.46780   0.784  0.44233
factor(Rx)5        0.16667    0.46780   0.356  0.72536
factor(Rx)6        0.11667    0.46780   0.249  0.80560
factor(Site)2     -0.60000    0.46780  -1.283  0.21430
factor(Site)3     -0.83333    0.46780  -1.781  0.09004 .
factor(Site)4      0.03333    0.46780   0.071  0.94390
factor(Site)5     -0.45000    0.46780  -0.962  0.34756
factor(Site)6     -0.65000    0.46780  -1.389  0.17996
factor(Subjects)2  1.55000    0.46780   3.313  0.00347 **
factor(Subjects)3  0.01667    0.46780   0.036  0.97193
factor(Subjects)4 -0.26667    0.46780  -0.570  0.57499
factor(Subjects)5  0.05000    0.46780   0.107  0.91595
factor(Subjects)6  0.45000    0.46780   0.962  0.34756
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.8102 on 20 degrees of freedom
Multiple R-squared: 0.5675,     Adjusted R-squared: 0.2432
F-statistic:  1.75 on 15 and 20 DF,  p-value: 0.1205
```

```
> anova(analysis)
Analysis of Variance Table

Response: y
                 Df  Sum Sq Mean Sq F value  Pr(>F)
factor(Rx)        5  0.5633 0.11267  0.1716 0.97013
factor(Site)      5  3.8333 0.76667  1.1678 0.35919
factor(Subjects)  5 12.8333 2.56667  3.9096 0.01235 *
Residuals        20 13.1300 0.65650
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```