STAT 511 Solution to HW2 Spring 2008

1. (a) The matrix A is

$$\begin{pmatrix} 0.5 & 0.5 & -0.5 & -0.5 & -0.5 & -0.5 & 0.5 & 0.5 \\ 0.25 & 0.25 & 0.25 & 0.25 & -0.25 & -0.25 & -0.25 & -0.25 \end{pmatrix}$$

   (b) Since we know that $\text{H}_0 : C\beta = 0$ is equivalent to $H_0 : EY \in C(X) \cap C(A')^{\perp}$, it is enough to show that $C(X_0) = C(X) \cap C(A')^{\perp}$. It can be verified that $P_{A'}X_0 = 0$, so $C(X_0) \subset C(A')^{\perp}$ and clearly $C(X_0) \subset C(X)$. This means $C(X_0) \subset C(X) \cap C(A')^{\perp}$. To prove $(C(X) \cap C(A')^{\perp}) \subset C(X_0)$, it is enough to show that the $1^{st}, 4^{th}, 5^{th}$ columns belong to $C(A')^{\perp}$, but the other columns of X do not belong to $C(A')^{\perp}$. We know $1^{st}, 4^{th}, 5^{th}$ columns belong to $C(A')^{\perp}$ from the fact that $P_{A'}X_0 = 0$ and we can verify that $P_{A'}x_i \neq 0$ for $i = 2, 3, 6, 7, 8$, where x_i is the ith column of X. Then $C(X_0) = C(X) \cap C(A')^{\perp}$.

   (c) $\text{H}_0 : C\beta = 0$ tests the hypothesis of "B main effects only".

2. (a) Under Gauss-Markov assumption, we know $EY = X\beta$ and $\text{Var}(Y) = \sigma^2 I$. Then $E(Y - \hat{Y}) = E(I - P_X)Y = E(I - P_X)X\beta = 0$ and $\text{Var}(Y - \hat{Y}) = (I - P_X)\text{Var}(Y)(I - P_X) = \sigma^2(I - P_X)$.

$$\begin{aligned} Cov(\hat{Y}, Y - \hat{Y}) &= Cov(P_X Y, (I - P_X)Y) = P_X Var(Y)(I - P_X)' \\ &= \sigma^2 P_X (I - P_X)' = 0. \end{aligned}$$

   Thus, there is no correlation between $\hat{Y}$ and $Y - \hat{Y}$.

   (b) $E(Y - \hat{Y})'(Y - \hat{Y}) = E(Y'(I - P_X)Y) = tr((I - P_X)\sigma^2 I) + (X\beta)'(I - P_X)(X\beta) = \sigma^2 \text{tr}(I - P_X) = \sigma^2(n - \text{rank}(X))$.

3. (a) An appropriate estimate of EY is $\hat{Y} = X(X^\tau V_i^{-1} X)^-(X^\tau V_i^{-1} Y)$, $i = 1, 2$. The estimate of $C\beta$ is $\hat{C\beta} = C(X^\tau V_i^{-1} X)^-(X^\tau V_i^{-1} Y)$.
   For V_1, the two estimates are $\hat{Y} = (1.8, 1.8, 4, 6, 3.5, 3.5, 3.5)$ and $\hat{C\beta} = (1.8, 4, 6, 3.5)$.
   For V_2, the two estimates are $\hat{Y} = (2, 2, 4, 6, 3.6129, 3.6129, 3.6129)$ and $\hat{C\beta} = (2, 4, 6, 3.6129)$.

   (b) The covariance matrix for GLS estimate of EY is

$$Var(\hat{Y}_{GLS}) = X(X^\tau V_i^{-1} X)^- X^\tau V_i^{-1} X (X^\tau V_i^{-1} X)^- X^\tau = X(X^\tau V_i^{-1} X)^- X^\tau.$$

   For OLS estimate of EY is $Var(\hat{Y}_{OLS}) = X(X^\tau X)^- X^\tau V_i X (X^\tau X)^- X^\tau$.
   For V_1 we can get the variance for $\hat{Y}_{GLS}$ and $\hat{Y}_{OLS}$, respectively, as

$$\begin{pmatrix} 0.8 & 0.8 & 0 & 0 & 0 & 0 & 0 \\ 0.8 & 0.8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2/3 & 2/3 & 2/3 \\ 0 & 0 & 0 & 0 & 2/3 & 2/3 & 2/3 \\ 0 & 0 & 0 & 0 & 2/3 & 2/3 & 2/3 \end{pmatrix} \begin{pmatrix} 1.25 & 1.25 & 0 & 0 & 0 & 0 & 0 \\ 1.25 & 1.25 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

For V_2 we can get the variance for $\hat{Y}_{GLS}$ and $\hat{Y}_{OLS}$, respectively, as

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.387 & 0.387 & 0.387 \\ 0 & 0 & 0 & 0 & 0.387 & 0.387 & 0.387 \\ 0 & 0 & 0 & 0 & 0.387 & 0.387 & 0.387 \end{pmatrix} \begin{pmatrix} 1.75 & 1.75 & 0 & 0 & 0 & 0 & 0 \\ 1.75 & 1.75 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.778 & 0.778 & 0.778 \\ 0 & 0 & 0 & 0 & 0.778 & 0.778 & 0.778 \\ 0 & 0 & 0 & 0 & 0.778 & 0.778 & 0.778 \end{pmatrix}.$$

Similarly, for $C\beta$ we can get the similar formula for the variance of the estimates. For V_1, we have the variance of $\hat{C\beta}_{GLS}$ and $\hat{C\beta}_{OLS}$, respectively, as

$$\begin{pmatrix} 0.8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2/3 \end{pmatrix} \begin{pmatrix} 1.25 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

For V_2, we get the variance of $\hat{C\beta}_{GLS}$ and $\hat{C\beta}_{OLS}$, respectively, as

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0.387 \end{pmatrix} \begin{pmatrix} 1.75 & 0 & 0 & 0 \\ 0 & 4 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0.778 \end{pmatrix}.$$

4. (a) Use lm in R, we get

```
> X <- matrix(c(1,1,rep(c(rep(0,7),1),3),1,1),7,4)
> Y<-c(2,1,4,6,3,5,4)
> V1<-c(1,4,4,1,1,4,4)
> fit<-lm(Y~X-1,weights=1/V1)
> fit$coefficients
 X1  X2  X3  X4
1.8 4.0 6.0 3.5
```

Then the estimate for 4 cell means are $(1.8, 4, 6, 3.5)$.

(b) Use lm.gls() in R, we get

```
> V2<-diag(c(1,4,4,1,1,4,4))
> V2[1,2]<-V2[2,1]<-1
> V2[3,4]<-V2[4,3]<-V2[5,6]<-V2[6,5]<--1
> fit<-lm.gls(Y~X-1, W=solve(V2))
> fit$coefficient
      X1       X2       X3       X4
2.000000 4.000000 6.000000 3.612903
```

Then the estimate for 4 cell means as $(2, 4, 6, 3.6129)$.