# STAT 511 HW#1 SPRING 2009

**PROBLEM 1.**

a.

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ y_{10} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 1 & 4 & 0.5 & 0 \\ 1 & 5 & 1.5 & 0 \\ 1 & 6 & 2.5 & 0 \\ 1 & 7 & 3.5 & 0 \\ 1 & 8 & 4.5 & 4 \\ 1 & 9 & 5.5 & 13.5 \\ 1 & 10 & 6.5 & 25 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \end{bmatrix}$$

b.

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{122} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & x_{1111} & x_{2111} \\ 1 & 1 & 0 & 1 & 0 & x_{1112} & x_{2112} \\ 1 & 1 & 0 & 0 & 1 & x_{1121} & x_{2121} \\ 1 & 1 & 0 & 0 & 1 & x_{1122} & x_{2122} \\ 1 & 0 & 1 & 1 & 0 & x_{1211} & x_{2211} \\ 1 & 0 & 1 & 1 & 0 & x_{1212} & x_{2212} \\ 1 & 0 & 1 & 0 & 1 & x_{1221} & x_{2221} \\ 1 & 0 & 1 & 0 & 1 & x_{1222} & x_{2222} \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \gamma_1 \\ \gamma_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_{111} \\ \varepsilon_{112} \\ \varepsilon_{121} \\ \varepsilon_{122} \\ \varepsilon_{211} \\ \varepsilon_{212} \\ \varepsilon_{221} \\ \varepsilon_{222} \end{bmatrix}$$

**PROBLEM 2.**

1. $C_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ and $C_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

2. $C_1{}^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$, $C_2{}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $(C_1{}^{-1})' = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$, $(C_2{}^{-1})' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

3. and 4. $A_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

5. $A_1{}' = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $A_2{}' = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

```
A<-matrix(c(1,1,0,1,0,1,1,0,1),3,3,byrow=TRUE)
ginv(A)
         [,1]      [,2]      [,3]
[1,]  0.3333333  0.1666667  0.1666667
[2,]  0.6666667 -0.1666667 -0.1666667
[3,] -0.3333333  0.3333333  0.3333333
```

**PROBLEM 3.**

$$X = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

```
X<-matrix(c(rep(1,9),rep(0,7),rep(1,2),rep(0,7) ,1,rep(0,7),rep(1,2)),7,5,byrow=FALSE)
```

1.

$$M = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \end{bmatrix}$$

Using R Create M, 5 matrices representing each column of X, and w a vector perpendicular to C(X)

```
M<-
matrix(c(rep(c(rep(0.5,2),rep(0,5)),2),rep(c(0,0,0.5,0.5,0,0,0),2),c(0,0,0,0,1,0,0),rep(c(
rep(0,5),rep(0.5,2)),2)),7,7,byrow=TRUE)
C1<-matrix(X[,1],7,1,byrow=TRUE)
C2<- matrix(X[,2] ,7,1,byrow=TRUE)
C3<- matrix(X[,3] ,7,1,byrow=TRUE)
C4<- matrix(X[,4] ,7,1,byrow=TRUE)
C5<- matrix(X[,5] ,7,1,byrow=TRUE)
w<-matrix(c(1,-1,1,-1,0,1,-1),7,1,byrow=TRUE)
```

Check that w is perpendicular to C(X) by

t(Ci)%*%w=0 for i=1,2,…,5

Using definition B.31 (Christensen) we can prove that M is a perpendicular projection operator onto C(X)

To show that is a projection operator multiply M by each column of X and you should get the column of X as

M%*%Ci=Ci for i=1,2,…,5

To show that is perpendicular to C(X) multiply M by w, and you should get a 0 vector

M%*%w=0

By Proposition B.32 (Christensen) C(M)=C(X)

By Theorem B.33 (Christensen) M is the perpendicular projection operator onto C(M) if $MM = M$ and $M' = M$

```
M%*%M
t(M)
M%*%M
```

```
     [,1] [,2] [,3] [,4] [,5] [,6] [,7]
[1,]  0.5  0.5  0.0  0.0    0  0.0  0.0
[2,]  0.5  0.5  0.0  0.0    0  0.0  0.0
[3,]  0.0  0.0  0.5  0.5    0  0.0  0.0
[4,]  0.0  0.0  0.5  0.5    0  0.0  0.0
[5,]  0.0  0.0  0.0  0.0    1  0.0  0.0
[6,]  0.0  0.0  0.0  0.0    0  0.5  0.5
[7,]  0.0  0.0  0.0  0.0    0  0.5  0.5
t(M)
     [,1] [,2] [,3] [,4] [,5] [,6] [,7]
[1,]  0.5  0.5  0.0  0.0    0  0.0  0.0
[2,]  0.5  0.5  0.0  0.0    0  0.0  0.0
[3,]  0.0  0.0  0.5  0.5    0  0.0  0.0
[4,]  0.0  0.0  0.5  0.5    0  0.0  0.0
[5,]  0.0  0.0  0.0  0.0    1  0.0  0.0
[6,]  0.0  0.0  0.0  0.0    0  0.5  0.5
[7,]  0.0  0.0  0.0  0.0    0  0.5  0.5
```

2.

$P_X = X(X'X)^{-}X'$

```
PX<-X%*%ginv(t(X)%*%X)%*%t(X)
PX
     [,1] [,2] [,3] [,4] [,5] [,6] [,7]
[1,]  0.5  0.5  0.0  0.0    0  0.0  0.0
[2,]  0.5  0.5  0.0  0.0    0  0.0  0.0
[3,]  0.0  0.0  0.5  0.5    0  0.0  0.0
[4,]  0.0  0.0  0.5  0.5    0  0.0  0.0
[5,]  0.0  0.0  0.0  0.0    1  0.0  0.0
[6,]  0.0  0.0  0.0  0.0    0  0.5  0.5
[7,]  0.0  0.0  0.0  0.0    0  0.5  0.5
```

**PROBLEM 4.**

```
Y<-matrix(c(3,1,4,6,3,5,4),7,1,byrow=TRUE)
```

a.

$\hat{Y} = \begin{bmatrix} 2 \\ 2 \\ 5 \\ 5 \\ 3 \\ 4.5 \\ 4.5 \end{bmatrix}$ Pick $b_1 = \begin{bmatrix} 2 \\ 0 \\ 3 \\ 1 \\ 2.5 \end{bmatrix}$ and $b_2 = \begin{bmatrix} 1 \\ 1 \\ 4 \\ 2 \\ 3.5 \end{bmatrix}$. Then $Xb_1 = Xb_2 = \hat{Y}$

b.

```
Yhat<-PX%*%Y
Yhat
    [,1]
[1,]  2.0
[2,]  2.0
[3,]  5.0
[4,]  5.0
[5,]  3.0
[6,]  4.5
[7,]  4.5

Y-Yhat
              [,1]
[1,]  1.000000e+00
[2,] -1.000000e+00
[3,] -1.000000e+00
[4,]  1.000000e+00
[5,] -2.664535e-15
[6,]  5.000000e-01
[7,] -5.000000e-01

t(Yhat)%*%( Y-Yhat)
  -9.769963e-15

t(Y)%*%Y
 112

t(Yhat)%*%Yhat
  107.5

t(Y-Yhat)%*%(Y-Yhat)
  4.5
```

**PROBLEM 5.**

a) What is the marginal distribution of $y_3$ ?

$y_3 \sim N(0,4)$

b) What is the (marginal joint) distribution of $y_1$ and $y_3$ ?

$(y_1, y_3) \sim MVN_2(\mu_{13}, \Sigma_{13})$ where $\mu_{13} = (1,0)$ and $\Sigma_{13} = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$

c) What is the conditional distribution of $y_3$ given that $y_1 = 2$ ?

$y_3 | y_1 = 2 \sim N(\frac{1}{4}, \frac{15}{4})$

d) What is the conditional distribution of $y_3$ given that $y_1 = 2$ and $y_2 = -1$?

$y_3|y_1 = 2, y_2 = -1 \sim N(-2,3)$

e) ) What is the conditional distribution of $y_3$ and $y_1 = 2$ given that $y_2 = -1$?

$y_3, y_1|y_2 = -1 \sim MVN_2\left((-1,-2), \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}\right)$

f) What are the correlations $\rho_{12}$, $\rho_{13}$, and $\rho_{23}$?

$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$ then

$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \sigma_2} = \frac{2}{\sqrt{4}\sqrt{4}} = \frac{1}{2}$

$\rho_{13} = \frac{\sigma_{13}}{\sigma_1 \sigma_3} = \frac{1}{\sqrt{4}\sqrt{4}} = \frac{1}{4}$

$\rho_{23} = \frac{\sigma_{23}}{\sigma_2 \sigma_3} = \frac{2}{\sqrt{4}\sqrt{4}} = \frac{1}{2}$

g) What is the joint distribution of $u = y_1 - 2y_2 - y_3$ and $v = y_1 - y_2 + 3$?

$(u, v) \sim MVN_2\left((-5,1), \begin{bmatrix} 22 & 7 \\ 7 & 4 \end{bmatrix}\right)$

**PROBLEM 6.**

```
sigma<-matrix(c(4,2,1,2,4,2,1,2,4),3,3,byrow=TRUE)
eigen.sigma<-eigen(sigma)
     $values
     [1] 7.372281 3.000000 1.627719

     $vectors
             [,1]          [,2]       [,3]
     [1,] -0.5417743  7.071068e-01  0.4544013
     [2,] -0.6426206  4.157915e-17 -0.7661846
     [3,] -0.5417743 -7.071068e-01  0.4544013
U<-eigen.sigma$vectors
```

$D^{-\frac{1}{2}}$ : Dsqrt<-solve(diag(sqrt(eigen.sigma$values)))

The inverse square root matrix for sigma is:

```
W<-U%*%Dsqrt%*%t(U)
W
           [,1]       [,2]        [,3]
[1,]  0.55861904 -0.1446625 -0.01873123
[2,] -0.14466250  0.6122191 -0.14466250
[3,] -0.01873123 -0.1446625  0.55861904
```

##Check that $WW = \Sigma^{-1}$

```
W%*%W
              [,1]       [,2]          [,3]
[1,]  3.333333e-01 -0.1666667  4.471148e-17
[2,] -1.666667e-01  0.4166667 -1.666667e-01
[3,]  5.919913e-17 -0.1666667  3.333333e-01
> solve(sigma)
```

```
           [,1]       [,2]       [,3]
[1,]  0.3333333 -0.1666667  0.0000000
[2,] -0.1666667  0.4166667 -0.1666667
[3,]  0.0000000 -0.1666667  0.3333333
```

**PROBLEM 7.**

```
A<-matrix(c(4,4.001,4.001,4.002),2,2,byrow=TRUE)
B<-matrix(c(4,4.001,4.001,4.002001),2,2,byrow=TRUE)
 det(A)
   -1e-06
 det(B)
    3e-06
solve(A)
        [,1]     [,2]
[1,] -4002000  4001000
[2,]  4001000 -4000000
 solve(B)
        [,1]     [,2]
[1,]  1334000 -1333667
[2,] -1333667  1333333
```

**PROBLEM 8.**

```
PtX<-t(X)%*%ginv(X%*%t(X))%*%X
```

$\tau_2$ *not estimable*

```
  c1<-c(0,0,1,0,0)
  PtX%*%c1
       [,1]
  [1,]  0.2
  [2,] -0.2
  [3,]  0.8
  [4,] -0.2
  [5,] -0.2
```

$\tau_1 - \tau_3$ *estimable*

```
c2<-c(0,1,0,-1,0)
PtX%*%c2
              [,1]
[1,] -1.665335e-16
[2,]  1.000000e+00
[3,] -2.220446e-16
[4,] -1.000000e+00
```

```
[5,] -2.775558e-16
t(c2)%*%ginv(t(X)%*%X)%*%t(X)
0.5  0.5 -2.914335e-16 -2.914335e-16   -1 -6.938894e-17 -6.938894e-17
```

$\mu + \frac{1}{3}(\tau_1 + \tau_2 + \tau_3)$ *estimable*

```
c3<-c(1,1/3,1/3,1/3,0)
PtX%*%c3
             [,1]
[1,] 1.000000e+00
[2,] 3.333333e-01
[3,] 3.333333e-01
[4,] 3.333333e-01
[5,] 3.699840e-18
t(c3)%*%ginv(t(X)%*%X)%*%t(X)
0.1666667 0.1666667 0.1666667 0.1666667 0.3333333 5.551115e-17 5.551115e-17
```

$\mu + \tau_1 + \tau_2 - \tau_3$ *estimable*

```
c4<-c(1,1,1,-1,0)
PtX%*%c4
              [,1]
[1,]  1.000000e+00
[2,]  1.000000e+00
[3,]  1.000000e+00
[4,] -1.000000e+00
[5,] -5.551115e-17
t(c4)%*%ginv(t(X)%*%X)%*%t(X)
0.5  0.5  0.5  0.5   -1 -1.387779e-17 -1.387779e-17
```

$\tau_1 - \frac{1}{2}(\tau_2 + \tau_3)$ *estimable*

```
c5<-c(0,1,-1/2,-1/2,0)
PtX%*%c5
              [,1]
[1,] -2.220446e-16
[2,]  1.000000e+00
[3,] -5.000000e-01
[4,] -5.000000e-01
[5,] -3.885781e-16
t(c5)%*%ginv(t(X)%*%X)%*%t(X)
0.5  0.5 -0.25 -0.25 -0.5 -6.245005e-17 -6.245005e-17
```

$(\tau_1 - \tau_2) - (\tau_3 - \tau_4)$ *estimable*

```
c6<-c(0,1,-1,-1,1)
PtX%*%c6
              [,1]
[1,] -4.996004e-16
[2,]  1.000000e+00
[3,] -1.000000e+00
[4,] -1.000000e+00
[5,]  1.000000e+00
t(c6)%*%ginv(t(X)%*%X)%*%t(X)
0.5  0.5 -0.5 -0.5   -1  0.5  0.5
```

**PROBLEM 9.** $P_X = X(X'X)^{-}X'$

```
project<-function(X)
{
    return(X%*%ginv(t(X)%*%X)%*%t(X))
}

project(X)
     [,1] [,2] [,3] [,4] [,5] [,6] [,7]
[1,]  0.5  0.5  0.0  0.0    0  0.0  0.0
[2,]  0.5  0.5  0.0  0.0    0  0.0  0.0
[3,]  0.0  0.0  0.5  0.5    0  0.0  0.0
[4,]  0.0  0.0  0.5  0.5    0  0.0  0.0
[5,]  0.0  0.0  0.0  0.0    1  0.0  0.0
[6,]  0.0  0.0  0.0  0.0    0  0.5  0.5
[7,]  0.0  0.0  0.0  0.0    0  0.5  0.5

project(t(X))
     [,1] [,2] [,3] [,4] [,5]
[1,]  0.8  0.2  0.2  0.2  0.2
[2,]  0.2  0.8 -0.2 -0.2 -0.2
[3,]  0.2 -0.2  0.8 -0.2 -0.2
[4,]  0.2 -0.2 -0.2  0.8 -0.2
[5,]  0.2 -0.2 -0.2 -0.2  0.8
```

**PROBLEM 10.**

a. Create matrix X and vectors ci (i=1…8). Then determine which ci's are estimable (ie $P_{X'}c_i = c_i$)

```
X<-matrix(c(1,1,0,1,0,1,0,0,0,1,1,0,1,0,1,0,0,0,1,1,0,0,1,0,1,0,0,1,1,0,0,1,0,1,0,0,1,0,1,1,0,0,0,1,0,
1,0,1,1,0,0,0,1,0,1,0,1,0,1,0,0,0,1, 1,0,1,0,1,0,0,0,1),8,9,byrow=T)
```

$\alpha_1$ *not estimable*

```
c1<-matrix((c(0,1,0,0,0,0,0,0,0)),9,1,byrow=T)
project(t(X))%*%c1
           [,1]
 [1,]  0.2222222
 [2,]  0.4444444
 [3,] -0.2222222
 [4,]  0.1111111
 [5,]  0.1111111
 [6,]  0.2222222
 [7,]  0.2222222
 [8,] -0.1111111
 [9,] -0.1111111
```

$\alpha\beta_{11}$ *not estimable*

```
c2<-matrix((c(0,0,0,0,0,1,0,0,0)),9,1,byrow=T)
project(t(X))%*%c2
           [,1]
 [1,]  0.1111111
 [2,]  0.2222222
 [3,] -0.1111111
 [4,]  0.2222222
 [5,] -0.1111111
 [6,]  0.4444444
 [7,] -0.2222222
 [8,] -0.2222222
 [9,]  0.1111111
```

$\alpha_1 - \alpha_2$ *not estimable*

```
c3<-matrix((c(0,1,-1,0,0,0,0,0,0)),9,1,byrow=T)
project(t(X))%*%c3
              [,1]
 [1,] -2.775558e-17
 [2,]  6.666667e-01
 [3,] -6.666667e-01
 [4,] -1.387779e-16
 [5,]  8.326673e-17
 [6,]  3.333333e-01
 [7,]  3.333333e-01
 [8,] -3.333333e-01
 [9,] -3.333333e-01
```

$\beta_1 - \beta_2 + \alpha\beta_{11} - \alpha\beta_{12}$ *estimable*

```
c4<-matrix((c(0,0,0,1,-1,1,-1,0,0)),9,1,byrow=T)
project(t(X))%*%c4
      [,1]
 [1,]    0
 [2,]    0
 [3,]    0
 [4,]    1
 [5,]   -1
 [6,]    1
 [7,]   -1
 [8,]    0
 [9,]    0
t(c4)%*%ginv(t(X)%*%X)%*%t(X)
  0.5  0.5 -0.5 -0.5    0    0    0    0
```

$\mu + \alpha_1 + \beta_1 + \alpha\beta_{11}$ *estimable*

```
c5<-matrix((c(1,1,0,1,0,1,0,0,0)),9,1,byrow=T)
project(t(X))%*%c5
      [,1]
 [1,]    1
 [2,]    1
 [3,]    0
 [4,]    1
 [5,]    0
 [6,]    1
 [7,]    0
 [8,]    0
 [9,]    0
t(c5)%*%ginv(t(X)%*%X)%*%t(X)
  0.5  0.5    0    0    0    0    0    0
```

$\alpha\beta_{12} - \alpha\beta_{11} - (\alpha\beta_{22} - \alpha\beta_{12})$ *estimable*

```
c6<-matrix((c(0,0,0,0,0,-1,1,1,-1)),9,1,byrow=T)
project(t(X))%*%c6
      [,1]
 [1,]    0
 [2,]    0
 [3,]    0
 [4,]    0
 [5,]    0
 [6,]   -1
 [7,]    1
```

```
    [8,]   1
    [9,]  -1
t(c6)%*%ginv(t(X)%*%X)%*%t(X)
  -0.5 -0.5  0.5  0.5  0.5  0.5 -0.5 -0.5
```

b. The first row of C correspond to $\alpha_1 - \alpha_2$ which is not estimable, therefore C is not testable

**PROBLEM 11.**

$H_0: E[y_{111}] = E[y_{112}] = E[y_{121}]$

$$H_0: \mu + \alpha_1 + \beta_1 + \gamma_1 x_{1111} + \gamma_2 x_{2111} = \mu + \alpha_1 + \beta_1 + \gamma_1 x_{1112} + \gamma_2 x_{2112}$$
$$= \mu + \alpha_1 + \beta_2 + \gamma_1 x_{1121} + \gamma_2 x_{2121}$$

$$H_0: \mu + \alpha_1 + \beta_1 + \gamma_1 x_{1111} + \gamma_2 x_{2111} - (\mu + \alpha_1 + \beta_1 + \gamma_1 x_{1112} + \gamma_2 x_{2112)}$$
$$= \mu + \alpha_1 + \beta_1 + \gamma_1 x_{1112} + \gamma_2 x_{2112} - (\mu + \alpha_1 + \beta_2 + \gamma_1 x_{1121} + \gamma_2 x_{2121}) = 0$$

$$H_0: \gamma_1(x_{1111} - x_{1112}) + \gamma_2(x_{2111} - x_{2112}) = \beta_1 - \beta_2 + \gamma_1(x_{1112} - x_{1121}) + \gamma_2(x_{2112} - x_{2121}) = 0$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & x_{1111} - x_{1112} & x_{2111} - x_{2112} \\ 0 & 0 & 0 & 1 & -1 & x_{1112} - x_{1121} & x_{2112} - x_{2121} \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$