STAT 511 – Spring 2002 – Solutions to Assignment 4

1. There is more than one way to approach the problem of deciding what linear combinations of parameters are estimable. One approach is to use the definition of estimability to directly show that $\underline{c}^T\underline{\beta}$ is estimable by finding a vector $\underline{a}$ such that $E(\underline{a}^T\underline{Y}) = \underline{c}^T\underline{\beta}$. To use this approach to show that $\underline{c}^T\underline{\beta}$ is not estimable you would have to provide an argument to show that there is no vector $\underline{a}$ for which $E(\underline{a}^T\underline{Y}) = \underline{c}^T\underline{\beta}$.

A second approach is based on Result 3.8(i) from the notes which says that $\underline{c}^T\underline{\beta}$ is estimable if and only if $\underline{c}^T$ is a linear combination of the rows of the model matrix. To use this approach you only have to work with the distinct rows of the model matrix. In this case, there are only six distinct rows. (In general you could work with any set of k linearly independent rows, where k is the rank of the model matrix.) Here the six distinct rows are

$$W = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

All vectors that are linear combinations of the rows of the model matrix are alos linear combinations of these six rows, and all such vectors have the form

$$\begin{aligned} \underline{c}^T &= a^T W \\ &= (a_1,\ a_2,\ a_3,\ a_4,\ a_5,\ a_6)W \\ &= (a_1+a_2+a_3+a_4+a_5+a_6, a_1+a_2+a_3,\ a_4+a_5+a_6,\ a_1+a_4,\ a_2+a_5,\ a_3+a_6,\ a_1,\ a_2,\ a_3,\ a_4,\ a_5,\ a_6) \end{aligned}$$

Now you have a description of all possible $\underline{c}^T$ vectors that can provide an estimable function of the parameters. To show, for example, that

$$\gamma_{11} - \gamma_{13} - \gamma_{21} + \gamma_{23} = (0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ \ -1\ \ -1\ 0\ 1)\underline{\beta}$$

is estimable we can simply pick $a_1 = 1$, $a_2 = 0$, $a_3 = -1$, $a_4 = -1$, $a_5 = 0$, $a_6 = 1$ to get

$$\underline{c}^T = \underline{a}^T W = (0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ \ -1\ \ -1\ 0\ 1)$$

To show that $\beta_1 - \beta_2$ is not estimable, we would note that we would have to have $a_1 = 0$, $a_2 = 0$, $a_3 = 0$, $a_4 = 0$, $a_5 = 0$, $a_6 = 0$ to prevent any γ_{ij} from appearing in the linear combination of parameters, but this would put zeros in the remaining positions of the $\underline{c}^T$ vector, and we are unable to produce $\beta_1 - \beta_2$.

A third approach is to make use of Result 3.8 (ii) from the course notes. It may be convenient to use this result to show that a linear combination of the parameters is not estimable. This result says that $\underline{c}^T\underline{\beta}$ is estimable if and only if $\underline{c}^T\underline{d} = 0$ for every $\underline{d}$ for which $X\underline{d} = 0$. Since

$$(\text{number of columns in } X) - rank(X) = 12 - 6 = 6$$

you would have to find a set of 6 linearly independent $\underline{d}$'s that satisfy $X\underline{d} = 0$ to completely describe the linear combinations of parameters that are either estimable or not estimable. Showing that a particular linear combination of parameters is not estimable may be handled by identifying a single appropriate $\underline{d}$ vector. For this exercise, consider

$$\begin{aligned} \underline{\beta} &= \begin{bmatrix} \mu & \alpha_1 & \alpha_2 & \beta_1 & \beta_2 & \beta_3 & \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{21} & \gamma_{22} & \gamma_{23} \end{bmatrix}^T, \\ \underline{d}_1 &= \begin{bmatrix} 3 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix}^T, \\ \underline{d}_2 &= \begin{bmatrix} 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}^T \end{aligned}$$

and

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

Note that $X\underline{d}_1 = X\underline{d}_2 = \underline{0}$.

(a) Let $\underline{c}_1 = [1, 0, ..., 0]^T$. Note that $\underline{c}_1^T\underline{\beta} = \mu$ but $\underline{c}_1^T\underline{d}_1 \neq \underline{0}$; i.e., there exists a vector $\underline{d}$ such that $X\underline{d} = \underline{0}$ but $\underline{c}_1^T\underline{\beta} \neq \underline{0}$. Thus, by Result 3.8 (ii), μ is not estimable.

(b) Let $\underline{c}_2 = [0, 0, 1, 0, ..., 0]^T$. Note that $\underline{c}_2^T\underline{\beta} = \alpha_2$ but $\underline{c}_2^T\underline{d}_1 \neq \underline{0}$. Thus, by Result 3.8 (ii), α_2 is not estimable.

(c) Let $\underline{c}_3 = [0, 0, 0, 0, 1, -1, 0, ..., 0]^T$. Note that $\underline{c}_3^T\underline{\beta} = \beta_2 - \beta_3$ but $\underline{c}_3^T\underline{d}_2 \neq \underline{0}$. Thus, by Result 3.9 (ii), $\beta_2 - \beta_3$ is not estimable.

(d) Let $\underline{c}_4 = [0, ..., 0, 1]^T$. Note that $\underline{c}_4^T\underline{\beta} = \gamma_{23}$ but $\underline{c}_4^T\underline{d}_1 \neq \underline{0}$. Thus, by Result 3.9 (ii), $\gamma_{23} = \underline{c}_4^T\underline{\beta}$ is not estimable.

(e) Let $\underline{c}_5 = [1, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1]^T$ and $\underline{a}_5 = [0, ..., 0, 1]^T$. Then, we have $\underline{a}_5^T X = \underline{c}_5^T$. Thus, by Result 3.9 (i), $\underline{c}_5^T\underline{\beta} = \mu + \alpha_2 + \beta_3 + \gamma_{23}$ is estimable. $\mu + \alpha_2 + \beta_3 + \gamma_{23}$ is the mean volume when fat 2 is used with surfactant C and the OLS estimator is $\bar{y}_{23\cdot}$.

(f) Let $\underline{c}_6 = [0, 0, 0, 0, 0, 0, 1, -1, 0, 0, 0, 0]^T$. Then, $\underline{c}_6^T\underline{\beta} = \gamma_{11} - \gamma_{12}$ but $\underline{c}_6^T\underline{d}_2 \neq \underline{0}$. Thus, by Result 3.9 (ii), $\gamma_{11} - \gamma_{12} = \underline{c}_6^T\underline{\beta}$ is not estimable.

(g) Let $\underline{c}_7 = [0, 0, 0, 0, 0, 0, 1, 0, -1, -1, 0, 1]^T$ and
$\underline{a}_7 = [1/3, 1/3, 1/3, 0, 0, 0, -1/3, -1/3, -1/3, -1/3, -1/3, -1/3, 0, 0, 1/4, 1/4, 1/4, 1/4]^T$.
Then, we have $\underline{a}_7^T X = \underline{c}_7^T$. Thus, by Result 3.9 (i), $\underline{c}_7^T\underline{\beta} = \gamma_{11} - \gamma_{13} - \gamma_{21} + \gamma_{23}$ is estimable.
$\gamma_{11} - \gamma_{13} - \gamma_{21} + \gamma_{23}$ = (difference in mean bread volumes for surfactants A and C when fat 1 is used)
-(difference in mean bread volumes for surfactants A and C when fat 2 is used)
This is an interaction contrast and the OLS estimator is $\bar{y}_{11\cdot} - \bar{y}_{13\cdot} - \bar{y}_{21\cdot} + \bar{y}_{23\cdot}$

(h) Let $\underline{c}_8 = [0, 0, 0, 0, 1, -1, 0, 1, -1, 0, 1, -1]^T$ and
$\underline{a}_8 = [0, 0, 0, 1/6, 1/6, 1/6, -1/6, -1/6, -1/6, 0, 0, 0, 1/2, 1/2, -1/4, -1/4, -1/4, -1/4]^T$.
Then, we have $\underline{a}_8^T X = \underline{c}_8^T$. Thus, by Result 3.9 (i), $\underline{c}_8^T\underline{\beta} = (\beta_2 - \beta_3) + \frac{1}{2}(\gamma_{12} + \gamma_{22} - \gamma_{13} - \gamma_{23})$ is estimable.
$(\beta_2 - \beta_3) + \frac{1}{2}(\gamma_{12} + \gamma_{22} - \gamma_{13} - \gamma_{23})$ is the difference between the mean bread volume when surfactant B is used and the mean bread volume when surfactant C is used, averaging across the fats giving equal weight to each fat.
The OLS estimator is $\frac{1}{2}\sum_{i=1}^{2}\bar{y}_{i2\cdot} - \frac{1}{2}\sum_{i=1}^{2}\bar{y}_{i3\cdot}$

(i) Let $\underline{c}_9 = [0, 0, 0, 0, 0, 0, 0, 1, -1, 0, 1, -1]^T$. Note that $\underline{c}_9^T\underline{\beta} = \gamma_{12} + \gamma_{22} - \gamma_{13} - \gamma_{23}$ but $\underline{c}_9^T\underline{d}_2 \neq \underline{0}$. Thus, by Result 3.9 (ii), $\gamma_{12} + \gamma_{22} - \gamma_{13} - \gamma_{23} = \underline{c}_9^T\underline{\beta}$ is not estimable.

Many students simply reported "estimable" or "not estimable". Assignments may be graded in a generous manner, but this will not earn full credit on an exam. You must be able to justfiy your answer using either the definition of estimability or result 3.8 or result 3.9 from the course notes to show that you are not merely quessing.

2. (a)

$$X = \begin{bmatrix} 1 & 1 & 0 & 90 \\ 1 & 1 & 0 & 95 \\ 1 & 1 & 0 & 100 \\ 1 & 1 & 0 & 105 \\ 1 & 1 & 0 & 110 \\ 1 & 0 & 1 & 90 \\ 1 & 0 & 1 & 95 \\ 1 & 0 & 1 & 100 \\ 1 & 0 & 1 & 105 \\ 1 & 0 & 1 & 110 \end{bmatrix}, \quad \underline{\beta} = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \gamma \end{bmatrix}.$$

(b) Note that for the model matrix given in (a), $X\underline{d} = 0$ if and only if $\underline{d}^T = w[-1\ 1\ 1\ 0]$ for some scalar w.

	function	estimable	Explanation
(i)	μ	No	Here $\underline{c}^T = [1\ 0\ 0\ 0]$ and $w[-1\ 1\ 1\ 1]\underline{c} = -w \neq 0$ for $w \neq 0$.
(ii)	$\mu + \alpha_2$	Yes	use $\underline{a}^T$ =[0 0 0 0 0 10 0 -9 0 0]
(iii)	β	Yes	use $\underline{a}^T$ =[-1/5 1/5 0 0 0 0 0 0 0 0]
(iv)	$\alpha_1 - \alpha_2$	Yes	use $\underline{a}^T$ =[1 0 0 0 0 -1 0 0 0 0]
(v)	$\mu + \beta T$	No	Here $\underline{c}^T = [1\ 0\ 0\ T]$ and $w[-1\ 1\ 1\ 1]\underline{c} = -w + wT \neq 0$ for $w \neq 0$
(vi)	$\mu + \alpha_1 + \beta(T-100)$	Yes	use $\underline{a}^T = [0.5 - \beta(T-100)/20, 0, 0, 0, 0.5 + \beta(T-100), 0, 0, 0, 0, 0]$

Many students had great difficulty with part (vi). You could reason in the following way to find a vector $\underline{a}$ such that $E(\underline{a}^T\underline{Y}) = \mu + \alpha_1 + \beta(T-100)$. Since this quantity involves $\mu + \alpha_1$ and does not involve $\mu + \alpha_2$ only use observations from runs with catalyst A. To estimate the slope β you will need to use observations from runs with catalyst A at two different temperatures. You could the observations from runs with catalyst A, but we will consider the observations Y_{11} and Y_{15} from runs at 90 and 100 degrees C. We need coefficients a_1 and a_5 such that

$$\mu + \alpha_1 + \beta(T-100) = E(a_1Y_{11} + a_5Y_{15}) = a_1(\mu + \alpha_1 + \beta(90-100)) + a_5(\mu + \alpha_1 + \beta(110-100)) = (a_1 + a_5)(\mu + \alpha_1) + \beta(10a_5 - 10a_1)$$

Consequently, we need $a_1 + a_5 = 1$ and $10a_5 - 10a_1 = \beta(T-100)$. Solvong these two equations, we obtain $a_1 = 0.5 - \beta(T-100)/20$ and $a_5 = 0.5 + \beta(T-100)/20$.

(c) The generalized inverse is shown in the solution to part (d).

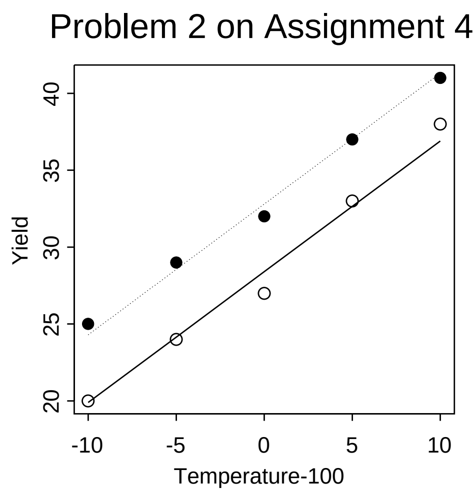
(d)
```
> X
   ..1 A B temp
1    1 1 0  -10
2    1 1 0   -5
3    1 1 0    0
4    1 1 0    5
5    1 1 0   10
6    1 0 1  -10
7    1 0 1   -5
8    1 0 1    0
9    1 0 1    5
10   1 0 1   10
>
> A <- t(X) %*% X
> G <- ginverse( A )
> G
           [,1]        [,2]        [,3]  [,4]
[1,] 0.04444444  0.02222222  0.02222222 0.000
[2,] 0.02222222  0.11111111 -0.08888889 0.000
[3,] 0.02222222 -0.08888889  0.11111111 0.000
```

```
[4,] 0.00000000  0.00000000  0.00000000 0.002
attr(, "rank"): [1] 3
>
> # (i)   AGA = A
> round( A %*% G %*% A - A, 5 )
     ..1 A B temp
 ..1   0 0 0    0
   A   0 0 0    0
   B   0 0 0    0
temp   0 0 0    0
>
> # (ii)  GAG = G
> round( G %*% A %*% G - G, 5 )
     [,1] [,2] [,3] [,4]
[1,]    0    0    0    0
[2,]    0    0    0    0
[3,]    0    0    0    0
[4,]    0    0    0    0
>
> # (iii) (AG)^T = AG
> round( t(A %*% G) - A %*% G , 5 )
      ..1 A B temp
 [1,]   0 0 0    0
 [2,]   0 0 0    0
 [3,]   0 0 0    0
 [4,]   0 0 0    0
>
> # (iv)  (GA)^T = GA
> round( t(G %*% A) - G %*% A , 5 )
     [,1] [,2] [,3] [,4]
 ..1    0    0    0    0
   A    0    0    0    0
   B    0    0    0    0
temp    0    0    0    0
```

Thus, the $(X^TX)^(-)$ generated from `ginverse()` in S-plus satisfies the four properties of the Moore-Penrose inverse.

(e)
```
> b <- ginverse( t(X) %*% X ) %*% t(X) %*% Y
> b
      [,1]
[1,] 20.40
[2,]  8.00
[3,] 12.40
[4,]  0.85
```

(f)



Based on the figure above, we can see that the observed data conform nicely to the two parallel lines corresponding to catalysts A and B (Same $\hat{\gamma} = .55$). Consequently, the observed data seem to agree with the proposed model. You could also examine residual plots.

(g) No. Note that the projection matrix $P_X = X(X^TX)^-X^T$ is invariant to the choice of generalized inverse $(X^TX)^-$. Then, the estimates for the mean yield $\underline{\hat{Y}} = P_X\underline{Y}$ are also invariant and thus unchanged even if different solutions to the normal equations are used.

3. (a)

$$(\mathbf{1}^T\mathbf{1})^{-1}\mathbf{1}^T\mathbf{Y} = (1/n)\sum_{i=1}^{2}\sum_{j=1}^{5} Y_{ij} = \bar{Y}_{..}, \qquad n = 10.$$

(b)

$$\begin{aligned}\mathbf{P_1P_1} &= \mathbf{1}(\mathbf{1}^T\mathbf{1})^{-1}\mathbf{1}^T\mathbf{1}(\mathbf{1}^T\mathbf{1})^{-1}\mathbf{1}^T = \mathbf{1}(\mathbf{1}^T\mathbf{1})^{-1}\mathbf{1}^T \\ &= \mathbf{P_1}.\end{aligned}$$

(c)

$$\hat{\mathbf{Y}} = \mathbf{P_1Y} = \mathbf{1}(\mathbf{1}^T\mathbf{1})^{-1}\mathbf{1}^T\mathbf{Y} = \mathbf{1}\bar{Y}_{..}$$

(d)

$$\begin{aligned}\mathbf{Y}^T\mathbf{Y} &= \mathbf{Y}^T(\mathbf{P_1} + \mathbf{I} - \mathbf{P_1})\mathbf{Y} \\ &= \mathbf{Y}^T\mathbf{P_1Y} + \mathbf{Y}^T(\mathbf{I} - \mathbf{P_1})\mathbf{Y} \\ &= \mathbf{Y}^T\mathbf{P_1P_1Y} + \mathbf{Y}^T(\mathbf{I} - \mathbf{P_1})(\mathbf{I} - \mathbf{P_1})\mathbf{Y} \\ &= (\mathbf{P_1Y})^T\mathbf{P_1Y} + \{(\mathbf{I} - \mathbf{P_1})\mathbf{Y}\}^T(\mathbf{I} - \mathbf{P_1})\mathbf{Y} \\ &= \hat{\mathbf{Y}}^T\hat{\mathbf{Y}} + (\mathbf{Y} - \hat{\mathbf{Y}})^T(\mathbf{Y} - \hat{\mathbf{Y}}) \\ &= \sum_{i=1}^{2}\sum_{j=1}^{5}\hat{Y_{ij}}^2 + \sum_{i=1}^{2}\sum_{j=1}^{5}(Y_{ij} - \hat{Y}_{ij})^2 \\ &= SS_{model,uncorrected} + SS_{residuals}\end{aligned}$$

(e)

$$\begin{aligned}\mathbf{Y}^T\mathbf{P_1Y} &= \mathbf{Y}^T\mathbf{P_1P_1Y} \\ &= (\mathbf{P_1Y})^T\mathbf{P_1Y} \\ &= (\mathbf{1}\bar{Y}_{..})^T\mathbf{1}\bar{Y}_{..}^T \\ &= n\bar{Y}_{..}^2\end{aligned}$$

(f)

$$\begin{aligned}\mathbf{Y}^T\mathbf{Y} &= \mathbf{Y}^T(\mathbf{P_1} + \mathbf{P_x} - \mathbf{P_1} + \mathbf{I} - \mathbf{P_x})\mathbf{Y} \\ &= \mathbf{Y}^T\mathbf{P_1}\mathbf{Y} + \mathbf{Y}^T(\mathbf{P_x} - \mathbf{P_1})\mathbf{Y} + \mathbf{Y}^T(\mathbf{I} - \mathbf{P_x})\mathbf{Y}\end{aligned}$$

(g)

$$\begin{aligned}\mathbf{Y}^T(\mathbf{P_x} - \mathbf{P_1})\mathbf{Y} &= \mathbf{Y}^T(\mathbf{I} - \mathbf{P_1} + \mathbf{P_x} - \mathbf{I})\mathbf{Y} \\ &= \mathbf{Y}^T(\mathbf{I} - \mathbf{P_1})\mathbf{Y} - \mathbf{Y}^T(\mathbf{I} - \mathbf{P_x})\mathbf{Y} \\ &= \{(\mathbf{I} - \mathbf{P_1})\mathbf{Y}\}^T(\mathbf{I} - \mathbf{P_1})\mathbf{Y} - \{(\mathbf{I} - \mathbf{P_x})\mathbf{Y}\}^T(\mathbf{I} - \mathbf{P_x})\mathbf{Y} \\ &= (\mathbf{Y} - \mathbf{1}\bar{Y_{..}})^T(\mathbf{Y} - \mathbf{1}\bar{Y_{..}}) - (\mathbf{Y} - \hat{\mathbf{Y}})^T(\mathbf{Y} - \hat{\mathbf{Y}}) \\ &= \sum_{i=1}^{2}\sum_{j=1}^{5}(Y_{ij} - \bar{Y_{..}})^2 - \sum_{i=1}^{2}\sum_{j=1}^{5}(Y_{ij} - \hat{Y}_{ij})^2 \\ &= SS_{residuals,commonmeans,model} - SS_{residuals,problem2}\end{aligned}$$