STAT 714 hw 6

Cochran's theorem, ANOVA

Do problem 5.23 from Monahan. In addition:

1. Let V be a subspace in $\mathbb{R}^n$ and let $\mathbf{u} \in V^\perp$. Show that $\text{proj}_V \mathbf{u} = \mathbf{0}$.

2. Let $Z_1, \dots, Z_q \overset{\text{ind}}{\sim} \text{Normal}(0,1)$ and let $\mu_1, \dots, \mu_q$ be real numbers. Derive the mean and variance of the random variable $W = \sum_{j=1}^q (Z_j + \mu_j)^2$.

This has mean

$$\mathbb{E}\sum_{j=1}^q (Z_j + \mu_j)^2 = \sum_{j=1}^q (\mathbb{E}Z_j^2 + 2\mathbb{E}Z_j\mu_j + \mu_j^2) = q + \sum_{j=1}^q \mu_j^2$$

and variance

$$\begin{aligned}
\text{Var}\sum_{j=1}^q (Z_j + \mu_j)^2 &= \sum_{j=1}^q \text{Var}(Z_j + \mu_j)^2 \\
&= \sum_{j=1}^q [\mathbb{E}(Z_j + \mu_j)^4 - (\mathbb{E}(Z_j + \mu_j)^2)^2] \\
&= \sum_{j=1}^q [\mathbb{E}Z_j^4 + 3\mathbb{E}Z_j^3\mu_j + 6\mathbb{E}Z_j^2\mu_j^2 + 3\mathbb{E}Z_j\mu_j^3 + \mu_j^4 - (\mathbb{E}Z_j^2 + 2\mathbb{E}Z_j\mu_j + \mu_j^2)^2] \\
&= \sum_{j=1}^q [3 + 6\mu_j^2 + \mu_j^4 - (1 + \mu_j^2)^2] \\
&= 2q + 4\sum_{j=1}^q \mu_j^2
\end{aligned}$$

3. Let $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$ for $i = 1, \dots, a$, $j = 1, \dots, b$ and $k = 1, \dots, n$, where $\varepsilon_{ijk} \overset{\text{ind}}{\sim} \text{Normal}(0, \sigma^2)$. Suppose the α_i and β_j represent *main effects* of factors A and B, which have a and b treatment levels, respectively, and the $(\alpha\beta)_{ij}$ represent *interaction effects* between A and B, which we will denote by AB. Note that the number of observations n is the same for all combinations of treatment levels; such a setup is called a *balanced design*.

   (a) Suppose $a = 2$ and $b = 2$.

      i. Give the design matrix $\mathbf{X}$ and the vector of parameters $\mathbf{b}$ in the matrix representation $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$ of the model.

We have

$$\mathbf{X} = \begin{bmatrix}
\mathbf{1}_n & \mathbf{1}_n & & \mathbf{1}_n & & \mathbf{1}_n & & & \\
\mathbf{1}_n & \mathbf{1}_n & & & \mathbf{1}_n & & \mathbf{1}_n & & \\
\mathbf{1}_n & & \mathbf{1}_n & \mathbf{1}_n & & & & \mathbf{1}_n & \\
\mathbf{1}_n & & \mathbf{1}_n & & \mathbf{1}_n & & & & \mathbf{1}_n
\end{bmatrix}$$

and

$$\mathbf{b} = [\mu\ \alpha_1\ \alpha_2\ \beta_1\ \beta_2\ (\alpha\beta)_{11}\ (\alpha\beta)_{12}\ (\alpha\beta)_{21}\ (\alpha\beta)_{22}]^T.$$

ii. Let $\bar{\mu}_{i.} = (1/2)\sum_{j=1}^{2}(\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij})$ for $i = 1, 2$ and $\bar{\mu}_{.j} = (1/2)\sum_{i=1}^{2}(\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij})$ for $j = 1, 2$. Check whether these contrasts are estimable in the model $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$.

Start with $\bar{\mu}_{1.} = \mu + \alpha_1 + (\beta_1 + \beta_2)/2 + ((\alpha\beta)_{11} + (\alpha\beta)_{12})/2$. This can be written as $\mathbf{c}^T\mathbf{b}$, where $\mathbf{c} = [1\ 1\ 0\ 1/2\ 1/2\ 1/2\ 1/2\ 0\ 0]^T$. Row reducing the following augmented matrix shows that the contrast is estimable:

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1/2 \\ 0 & 1 & 0 & 1 & 1/2 \\ 1 & 0 & 0 & 0 & 1/2 \\ 0 & 1 & 0 & 0 & 1/2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1/2 \\ 0 & 1 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

We just need to add the first two rows of $\mathbf{X}$ and divide by 2. We likewise find that the other contrasts are estimable.

iii. Write down the matrix $\mathbf{C}$ such that $\mathbf{Cb} = \mathbf{0}$ imposes the constraints

$$\sum_{i=1}^{a}\alpha_i = 0, \quad \sum_{j=1}^{b}\beta_j = 0, \quad \text{and} \quad \sum_{i=1}^{a}(\alpha\beta)_{ij} = 0 \text{ for all } j \text{ and } \sum_{j=1}^{b}(\alpha\beta)_{ij} = 0 \text{ for all } i. \tag{1}$$

The matrix is

$$\mathbf{C} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

iv. Give the matrix $\begin{bmatrix}\mathbf{X}^T\mathbf{X} \\ \mathbf{C}\end{bmatrix}$ and the vector $\begin{bmatrix}\mathbf{X}^T\mathbf{y} \\ \mathbf{0}\end{bmatrix}$.

We have

$$\begin{bmatrix} \mathbf{X}^T\mathbf{X} \\ \mathbf{C} \end{bmatrix} = \begin{bmatrix} 4n & 2n & 2n & 2n & 2n & n & n & n & n \\ 2n & 2n & & n & n & n & n & & \\ 2n & & 2n & n & n & & & n & n \\ 2n & n & n & 2n & & n & & n & \\ 2n & n & n & & 2n & & n & & n \\ n & n & & n & & n & & & \\ n & n & & & n & & n & & \\ n & & n & n & & & & n & \\ n & & n & & n & & & & n \\ & 1 & 1 & & & & & & \\ & & & 1 & 1 & & & & \\ & & & & & 1 & 1 & & \\ & & & & & & & 1 & 1 \\ & & & & & 1 & & 1 & \\ & & & & & & 1 & & 1 \end{bmatrix} \text{ and } \begin{bmatrix} \mathbf{X}^T\mathbf{y} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} 4n\bar{y}_{...} \\ 2n\bar{y}_{1..} \\ 2n\bar{y}_{2..} \\ 2n\bar{y}_{.1.} \\ 2ny_{.2.} \\ n\bar{y}_{11.} \\ n\bar{y}_{12.} \\ n\bar{y}_{21.} \\ n\bar{y}_{22.} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

v. Under the constraint, give the least-squares estimators of all the parameters

$$\mu, \alpha_1, \alpha_2, \beta_1, \beta_2, (\alpha\beta)_{11}, (\alpha\beta)_{12}, (\alpha\beta)_{21}, (\alpha\beta)_{22}$$

in terms of the response values Y_{ijk}.

We may write

$$\begin{bmatrix} \mathbf{X}^T\mathbf{X} & \mathbf{X}^T\mathbf{y} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \sim \begin{bmatrix} 4n & & & & & & & & & 4n\bar{y}_{...} \\ 2n & 2n & & & & & & & & 2n\bar{y}_{1..} \\ 2n & & 2n & & & & & & & 2n\bar{y}_{2..} \\ 2n & & & 2n & & & & & & 2n\bar{y}_{.1.} \\ 2n & & & & 2n & & & & & 2n\bar{y}_{.2.} \\ n & n & & n & & n & & & & n\bar{y}_{11.} \\ n & n & & & n & & n & & & n\bar{y}_{12.} \\ n & & n & n & & & & n & & n\bar{y}_{21.} \\ n & & n & & n & & & & n & n\bar{y}_{22.} \\ & 1 & 1 & & & & & & & 0 \\ & & & 1 & 1 & & & & & 0 \\ & & & & & 1 & 1 & & & 0 \\ & & & & & & & 1 & 1 & 0 \\ & & & & & 1 & & 1 & & 0 \\ & & & & & & 1 & & 1 & 0 \end{bmatrix}$$

The first row gives $\hat{\mu} = \bar{y}_{...}$. Then the next four rows give $\hat{\alpha}_1 = \bar{y}_{1..} - \bar{y}_{...}$, $\hat{\alpha}_2 = \bar{y}_{2..} - \bar{y}_{...}$, $\hat{\beta}_1 = \bar{y}_{.1.} - \bar{y}_{...}$, and $\hat{\beta}_2 = \bar{y}_{.2.} - \bar{y}_{...}$. Then the next four rows give $\widehat{(\alpha\beta)}_{ij} = \bar{y}_{i..} - \bar{y}_{...} - \hat{\alpha}_i - \hat{\beta}_j$.

So we obtain

$$\hat{\mathbf{b}} = \begin{bmatrix} \bar{y}_{...} \\ \bar{y}_{1..} - \bar{y}_{...} \\ \bar{y}_{2..} - \bar{y}_{...} \\ \bar{y}_{.1.} - \bar{y}_{...} \\ \bar{y}_{.2.} - \bar{y}_{...} \\ \bar{y}_{11.} - (\bar{y}_{1..} + \bar{y}_{.1.} - \bar{y}_{...}) \\ \bar{y}_{12.} - (\bar{y}_{1.} + \bar{y}_{.2.} - \bar{y}_{...}) \\ \bar{y}_{21.} - (\bar{y}_{2.} + \bar{y}_{.1.} - \bar{y}_{...}) \\ \bar{y}_{22.} - (\bar{y}_{2.} + \bar{y}_{.2.} - \bar{y}_{...}) \end{bmatrix}$$

vi. Give the least-squares estimators of the contrasts $\bar{\mu}_{1.}$, $\bar{\mu}_{2.}$, $\bar{\mu}_{.1}$, and $\bar{\mu}_{.2}$.

Plugging the values of the constrained least-squares solution $\hat{\mathbf{b}}$ into the expressions for the contrasts, we obtain the least-squares estimators

$$\hat{\bar{\mu}}_{1.} = (\bar{y}_{11.} + \bar{y}_{12.})/2, \quad \hat{\bar{\mu}}_{2.} = (\bar{y}_{21.} + \bar{y}_{22.})/2, \quad \hat{\bar{\mu}}_{.1} = (\bar{y}_{11.} + \bar{y}_{21.})/2, \quad \hat{\bar{\mu}}_{.2} = (\bar{y}_{12.} + \bar{y}_{22.})/2.$$

vii. Give the vector $\mathbf{P_X y}$ in terms of the values Y_{ijk}.

Using the fact that $\mathbf{P_X y} = \mathbf{X}\hat{\mathbf{b}}$, we see that the entries of $\mathbf{P_X y}$ will be given by $\hat{y}_{ijk} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + (\widehat{\alpha\beta})_{ij} = \bar{y}_{ij.}$ for all i, j, k. Therefore we have

$$\mathbf{P_X y} = \begin{bmatrix} \mathbf{1}_n \bar{y}_{11.} \\ \mathbf{1}_n \bar{y}_{12.} \\ \mathbf{1}_n \bar{y}_{21.} \\ \mathbf{1}_n \bar{y}_{22.} \end{bmatrix}.$$

viii. Make a complete ANOVA table with the sums of squares, degrees of freedom, and noncentrality parameter corresponding to each effect in the model. Use the sequential sum of squares idea based on Cochran's theorem. Give the noncentrality parameters in terms of the model parameters μ, α_i, and β_j. Create your table in this form (like the table on pg 115 of [2]):

Source	SS	df	ϕ
Mean			
A			
B			
AB			
Error			

We obtain

Source	SS	df	ϕ
Mean	$4n(\bar{y}_{...})^2$	1	$4n(\mu + \bar{\alpha} + \bar{\beta} + (\overline{\alpha\beta}))^2/\sigma^2$
A	$2n\sum_{i=1}^2(\bar{y}_{i..} - \bar{y}_{...})^2$	1	$2n\sum_{i=1}^2(\alpha_i - \bar{\alpha} + (\overline{\alpha\beta})_{i.} - (\overline{\alpha\beta}))^2/\sigma^2$
B	$2n\sum_{j=1}^2(\bar{y}_{.j.} - \bar{y}_{...})^2$	1	$2n\sum_{j=1}^2(\beta_j - \bar{\beta} + (\overline{\alpha\beta})_{.j} - (\overline{\alpha\beta}))^2/\sigma^2$
AB	$n\sum_{i=1}^2\sum_{j=1}^2(\bar{y}_{ij.} - (\bar{y}_{i..} + \bar{y}_{.j.} - \bar{y}_{...}))^2$	1	$n\sum_{i=1}^2\sum_{j=1}^2((\alpha\beta)_{ij} - (\overline{\alpha\beta}))^2/\sigma^2$
Error	$\sum_{i=1}^2\sum_{j=1}^2\sum_{k=1}^n(y_{ijk} - \bar{y}_{ij.})^2$	$4n-4$	0

where $\bar{\alpha} = (\alpha_1 + \alpha_2)/2$, $\bar{\beta} = (\beta_1 + \beta_2)/2$, $(\overline{\alpha\beta})_{i.} = (1/2)\sum_{j=1}^2(\alpha\beta)_{ij}$ for $i = 1, 2$, $(\overline{\alpha\beta})_{.j} = (1/2)\sum_{i=1}^2(\alpha\beta)_{ij}$ for $j = 1, 2$, and $(\overline{\alpha\beta}) = (1/4)\sum_{i=1}^2\sum_{j=1}^2(\alpha\beta)_{ij}$. We obtain the above by partitioning $\mathbf{X}$ as

$$\mathbf{X} = \left[\begin{array}{c|cc|cc|cccc} \mathbf{1}_n & \mathbf{1}_n & & \mathbf{1}_n & & \mathbf{1}_n & & & \\ \mathbf{1}_n & \mathbf{1}_n & & & \mathbf{1}_n & & \mathbf{1}_n & & \\ \mathbf{1}_n & & \mathbf{1}_n & \mathbf{1}_n & & & & \mathbf{1}_n & \\ \mathbf{1}_n & & \mathbf{1}_n & & \mathbf{1}_n & & & & \mathbf{1}_n \end{array}\right] = [\ \mathbf{X}_0 \mid \mathbf{X}_1 \mid \mathbf{X}_2 \mid \mathbf{X}_3\ ]$$

and defining

$$\mathbf{P}_0 = \mathbf{P}_{\mathbf{X}_0}, \quad \mathbf{P}_1 = \mathbf{P}_{[\mathbf{X}_0\ \mathbf{X}_1]}, \quad \mathbf{P}_2 = \mathbf{P}_{[\mathbf{X}_0\ \mathbf{X}_1\ \mathbf{X}_2]}, \quad \text{and } \mathbf{P}_3 = \mathbf{P}_{\mathbf{X}}.$$

Then we find

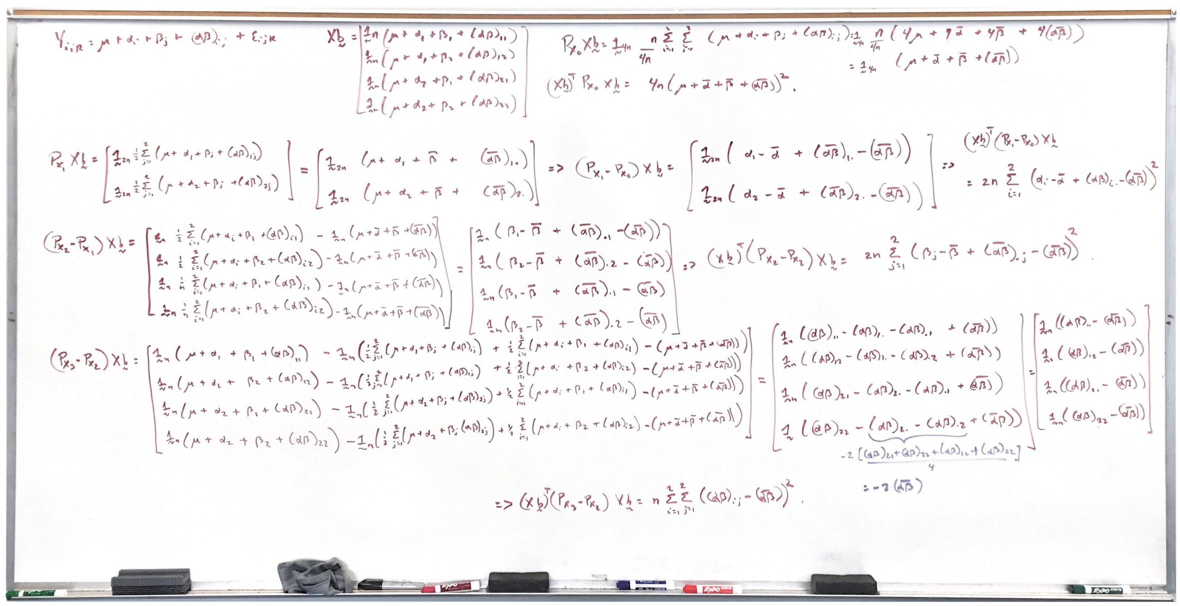
$$\mathbf{P}_0\mathbf{y} = \begin{bmatrix} \mathbf{1}_n\bar{y}_{...} \\ \mathbf{1}_n\bar{y}_{...} \\ \mathbf{1}_n\bar{y}_{...} \\ \mathbf{1}_n\bar{y}_{...} \end{bmatrix}, \quad \mathbf{P}_1\mathbf{y} = \begin{bmatrix} \mathbf{1}_n(\bar{y}_{...} + (\bar{y}_{1..} - \bar{y}_{...})) \\ \mathbf{1}_n(\bar{y}_{...} + (\bar{y}_{1..} - \bar{y}_{...})) \\ \mathbf{1}_n(\bar{y}_{...} + (\bar{y}_{2..} - \bar{y}_{...})) \\ \mathbf{1}_n(\bar{y}_{...} + (\bar{y}_{2..} - \bar{y}_{...})) \end{bmatrix}$$

and

$$\mathbf{P}_2\mathbf{y} = \begin{bmatrix} \mathbf{1}_n(\bar{y}_{...} + (\bar{y}_{1..} - \bar{y}_{...}) + (\bar{y}_{.1.} - \bar{y}_{...})) \\ \mathbf{1}_n(\bar{y}_{...} + (\bar{y}_{1..} - \bar{y}_{...}) + (\bar{y}_{.2.} - \bar{y}_{...})) \\ \mathbf{1}_n(\bar{y}_{...} + (\bar{y}_{2..} - \bar{y}_{...}) + (\bar{y}_{.1.} - \bar{y}_{...})) \\ \mathbf{1}_n(\bar{y}_{...} + (\bar{y}_{2..} - \bar{y}_{...}) + (\bar{y}_{.2.} - \bar{y}_{...})) \end{bmatrix}, \quad \mathbf{P}_3\mathbf{y} = \begin{bmatrix} \mathbf{1}_n\bar{y}_{11.} \\ \mathbf{1}_n\bar{y}_{12.} \\ \mathbf{1}_n\bar{y}_{21.} \\ \mathbf{1}_n\bar{y}_{22.} \end{bmatrix}.$$

A way to obtain $\mathbf{P}_1\mathbf{y}$ is to find $\hat{\mathbf{b}}_1$ satisfying $[\mathbf{X}_0\ \mathbf{X}_1]^T[\mathbf{X}_0\ \mathbf{X}_1]\hat{\mathbf{b}}_1 = [\mathbf{X}_0\ \mathbf{X}_1]^T\mathbf{y}$. Then $\mathbf{P}_1\mathbf{y} = [\mathbf{X}_0\ \mathbf{X}_1]\hat{\mathbf{b}}_1$. We may obtain $\mathbf{P}_2\mathbf{y}$ similarly, and we have already used this process to find $\mathbf{P}_3\mathbf{y}$.

To find the noncentrality parameter values, one finds $\mathbf{P}_0\mathbf{X}\mathbf{b}$, $\mathbf{P}_1\mathbf{X}\mathbf{b}$, $\mathbf{P}_2\mathbf{X}\mathbf{b}$, and $\mathbf{P}_3\mathbf{X}\mathbf{b}$ by replacing y_{ijk} with $\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$.



ix. Give $\hat{\sigma}^2$ in terms of the response values Y_{ijk}.

Note that rank $\mathbf{X} = 4$, so we have

$$\hat{\sigma}^2 = \frac{\|(\mathbf{I} - \mathbf{P_X})\mathbf{y}\|^2}{4n-4} = \frac{1}{4n-4}\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{K}(y_{ijk} - \bar{y}_{ij.})^2.$$

(b) Now give the ANOVA table for any $a \geq 2$ and $b \geq 2$ (you do not need to work this out step-by-step; you may just "extrapolate" from your work in the first part).

The ANOVA table becomes

Source	SS	df	ϕ
Mean	$abn(\bar{y}_{...})^2$	1	$abn(\mu + \bar{\alpha} + \bar{\beta} + (\overline{\alpha\beta}))^2/\sigma^2$
A	$bn\sum_{i=1}^{a}(\bar{y}_{i..} - \bar{y}_{...})^2$	$a-1$	$bn\sum_{i=1}^{a}(\alpha_i - \bar{\alpha} + (\overline{\alpha\beta})_{i.} - (\overline{\alpha\beta}))^2/\sigma^2$
B	$an\sum_{j=1}^{2}(\bar{y}_{.j.} - \bar{y}_{...})^2$	$b-1$	$an\sum_{j=1}^{b}(\beta_j - \bar{\beta} + (\overline{\alpha\beta})_{.j} - (\overline{\alpha\beta}))^2/\sigma^2$
AB	$n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij.} - (\bar{y}_{i.} + \bar{y}_{.j} - \bar{y}_{...}))^2$	$(a-1)(b-1)$	$n\sum_{i=1}^{a}\sum_{j=1}^{b}((\alpha\beta)_{ij} - (\overline{\alpha\beta}))^2/\sigma^2$
Error	$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.})^2$	$ab(n-1)$	0

We obtain the degrees of freedom values by partitioning $\mathbf{X}$ as we did before and defining

$$\mathbf{P}_0 = \mathbf{P}_{\mathbf{X}_0}, \quad \mathbf{P}_1 = \mathbf{P}_{[\mathbf{X}_0\ \mathbf{X}_1]}, \quad \mathbf{P}_2 = \mathbf{P}_{[\mathbf{X}_0\ \mathbf{X}_1\ \mathbf{X}_2]}, \quad \text{and } \mathbf{P}_3 = \mathbf{P}_{\mathbf{X}}.$$

Then

$$\text{tr}(\mathbf{P}_0) = 1, \quad \text{tr}(\mathbf{P}_1) = a, \quad \text{tr}(\mathbf{P}_2) = b, \quad \text{tr}(\mathbf{P}_3) = ab.$$

(c) Use the data in the table, which is scanned from [1]. Let factor A be the "Compaction Method" and factor B be the "Aggregate Type".

Table 6.3 Tensile strength (psi) of asphaltic concrete specimens for two aggregate types with each of four compaction methods

Aggregate Type	*Compaction Method*			
		Kneading		
	Static	*Regular*	*Low*	*Very Low*
Basalt	68	126	93	56
	63	128	101	59
	65	133	98	57
Silicious	71	107	63	40
	66	110	60	41
	66	116	59	44

Source: A. M. Al-Marshed (1981), Compaction effects on asphaltic concrete durability. M.S. thesis, Civil Engineering, University of Arizona.

For the following you may use R, but you may NOT use any built-in functions for fitting linear models! If you use R, provide your code.

i. Obtain the values of the least-squares estimators of μ, the α_i, the β_j, and the $(\alpha\beta)_{ij}$ under the constraints in (1).

```r
y <- c(68,63,65,71,66,66,126,128,133,107,110,
       116,93,101,98,63,60,59,56,59,57,40,41,44)
# 1 = static, 2 = regular, 3 = low, 4 = very low
A <- c(1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3,4,4,4,4,4,4)
# 1 = basalt, 2 = silicious
B <- c(1,1,1,2,2,2,1,1,1,2,2,2,1,1,1,2,2,2,1,1,1,2,2,2)
data <- cbind(y,A,B)

a <- length(unique(A))
b <- length(unique(B))
N <- length(y)
n <- 3
X0 <- matrix(1,N,1)
X1 <- matrix(0,N,a)
X2 <- matrix(0,N,b)
for(i in 1:N){
  X1[i,A[i]] <- 1
  X2[i,B[i]] <- 1
}

X3 <- diag(a*b) %x% rep(1,n)
X <- cbind(X0,X1,X2,X3)

Cmat <- rbind(c(0,rep(1,a),rep(0,b + a*b)),
          c(0,rep(0,a),rep(1,b),rep(0,a*b)),
          cbind(matrix(0,a,a + b + 1),diag(a) %x% t(rep(1,b))),
          cbind(matrix(0,b,a + b + 1),t(rep(1,a)) %x% diag(b)))

bhat <- solve(t(X) %*% X + t(Cmat) %*% Cmat) %*% t(X) %*% y
bhat
##              [,1]
##  [1,]  78.7500000
##  [2,] -12.2500000
##  [3,]  41.2500000
##  [4,]   0.2500000
##  [5,] -29.2500000
##  [6,]   8.5000000
##  [7,]  -8.5000000
##  [8,]  -9.6666667
##  [9,]   9.6666667
## [10,]   0.5000000
## [11,]  -0.5000000
## [12,]   9.8333333
## [13,]  -9.8333333
## [14,]  -0.6666667
## [15,]   0.6666667
```

ii. Give the sums of squares corresponding to the mean, factor A, factor B, the interaction AB, and the error term. Give the degrees of freedom corresponding to each sum of squares.

```
P0y <- X0 %*% solve(t(X0)%*% X0) %*% t(X0) %*% y
P1y <- X1 %*% solve(t(X1) %*% X1) %*% t(X1) %*% y
W012 <- cbind(X1[,-1],X2)
P2y <- W012 %*% solve(t(W012) %*% W012) %*% t(W012) %*% y
P3y <- X %*% bhat

SSmean <- sum(P0y^2)
SSA <- sum((P1y - P0y)^2)
SSB <- sum((P2y - P1y)^2)
SSAB <- sum((P3y - P2y)^2)
Error <- sum( (y - P3y)^2)

dfA <- a - 1
dfB <- b - 1
dfAB <- (a - 1)*(b - 1)
dfError <- N - (a*b)

SSmean
## [1] 148837.5
SSA
## [1] 16243.5
SSB
## [1] 1734
SSAB
## [1] 1145
Error
## [1] 152
dfA
## [1] 3
dfB
## [1] 1
dfAB
## [1] 3
dfError
## [1] 16
```

iii. Give $\hat{\sigma}^2$.

```
Error / dfError
## [1] 9.5
```

(d) Now consider the same data set with some observations removed so that the design is *unbalanced.*

Table 6.3 Tensile strength (psi) of asphaltic concrete specimens for two aggregate types with each of four compaction methods

	Compaction Method			
		Kneading		
Aggregate Type	*Static*	*Regular*	*Low*	*Very Low*
Basalt	68	126	93	56
	~~63~~	128	101	59
	65	133	~~98~~	57
Silicious	71	107	63	40
	66	~~110~~	60	~~41~~
	66	~~116~~	59	44

Source: A. M. Al-Marshed (1981), Compaction effects on asphaltic concrete durability. M.S. thesis, Civil Engineering, University of Arizona.

Without using any built-in linear models functions in R, obtain the values of the least-squares estimators of μ, the α_i, the β_j, and the $(\alpha\beta)_{ij}$ under the constraints

$$\sum_{i=1}^{a} n_{i.}\alpha_i = 0, \quad \sum_{j=1}^{b} n_{.j}\beta_j = 0, \quad \text{and} \quad \sum_{i=1}^{a} n_{ij}(\alpha\beta)_{ij} = 0 \;\forall\; j \text{ and } \sum_{j=1}^{b} n_{ij}(\alpha\beta)_{ij} = 0 \;\forall\; i.$$

```r
rmv <- c(2,11,12,15,23) # remove the corresponding rows of X and y
X0 <- as.matrix(X0[-rmv,])
X1 <- X1[-rmv,]
X2 <- X2[-rmv,]
X3 <- X3[-rmv,]
y <- y[-rmv]

Cmat <- rbind(c(0,rep(1,a),rep(0,b + a*b)),
              c(0,rep(0,a),rep(1,b),rep(0,a*b)),
              cbind(matrix(0,a,a + b + 1),diag(a) %x% t(rep(1,b))),
              cbind(matrix(0,b,a + b + 1),t(rep(1,a)) %x% diag(b)))


nA <- apply(X1,2,sum)
nB <- apply(X2,2,sum)
nAB <- apply(X3,2,sum)
Cmat <- Cmat %*% diag(c(1,nA,nB,nAB))

X <- cbind(X0,X1,X2,X3)

bhat <- solve(t(X) %*% X + t(Cmat) %*% Cmat) %*% t(X) %*% y
bhat
##              [,1]
##  [1,]  76.9473684
##  [2,]  -7.5085299
##  [3,]  42.5880218
##  [4,]   0.4914701
##  [5,] -27.0533575
##  [6,]   8.3956443
##  [7,]  -9.3284936
##  [8,] -11.3344828
##  [9,]   7.5563218
## [10,]   1.0689655
## [11,]  -3.2068966
## [12,]  11.1655172
## [13,]  -7.4436782
## [14,]  -0.9563218
## [15,]   1.4344828
```

References

[1] R. O. Kuehl. *Design of Experiments: Statistical Principles of Research Design and Analysis*. Duxbury/Thomson Learning, 2000. Google-Books-ID: mIV2QgAACAAJ.

[2] John F Monahan. *A primer on linear models*. CRC Press, 2008.

5.23 Let

$$\begin{bmatrix} \underset{\sim}{X}_1 \\ \underset{\sim}{X}_2 \end{bmatrix} \sim \text{Normal}\left(\begin{bmatrix} \underset{\sim}{\mu}_1 \\ \underset{\sim}{\mu}_2 \end{bmatrix}, \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \right).$$

Show that

$$\underset{\sim}{X}_1 \mid \underset{\sim}{X}_2 = \underset{\sim}{x}_2 \sim \text{Normal}\left(\underset{\sim}{\mu}_1 + V_{12} V_{22}^{-1} (\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2),\ V_{11} - V_{12} V_{22}^{-1} V_{21} \right).$$

__Solution:__

$$\frac{f(\underset{\sim}{x}_1, \underset{\sim}{x}_2)}{f_{\underset{\sim}{X}_2}(\underset{\sim}{x}_2)} = \frac{(2\pi)^{-\frac{p_1+p_2}{2}} \begin{vmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{vmatrix}^{-1/2} \exp\left[-\frac{1}{2} \begin{pmatrix} \underset{\sim}{x}_1 - \underset{\sim}{\mu}_1 \\ \underset{\sim}{x}_2 - \underset{\sim}{\mu}_2 \end{pmatrix}^T \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}^{-1} \begin{pmatrix} \underset{\sim}{x}_1 - \underset{\sim}{\mu}_1 \\ \underset{\sim}{x}_2 - \underset{\sim}{\mu}_2 \end{pmatrix} \right]}{(2\pi)^{-\frac{p_2}{2}} |V_{22}|^{-1/2} \exp\left[-\frac{1}{2} (\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2)^T V_{22}^{-1} (\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2) \right]}$$

We have

$$\begin{vmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{vmatrix} = |V_{22}| \, \left| V_{11} - V_{12} V_{22}^{-1} V_{21} \right| \qquad (*)$$

and

$$\begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}^{-1} = \begin{bmatrix} (V_{11} - V_{12} V_{22}^{-1} V_{21})^{-1} & -F^{-1} V_{12} V_{22}^{-1} \\ -V_{22}^{-1} V_{21} F^{-1} & V_{22}^{-1} + V_{22}^{-1} V_{21} F^{-1} V_{12} V_{22}^{-1} \end{bmatrix}.$$

Let $\underset{\sim}{a} = \underset{\sim}{x}_1 - \underset{\sim}{\mu}_1$ and $\underset{\sim}{b} = \underset{\sim}{x}_2 - \underset{\sim}{\mu}_2$. Then we have

$$\begin{pmatrix} \underset{\sim}{a} \\ \underset{\sim}{b} \end{pmatrix}^T \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}^{-1} \begin{pmatrix} \underset{\sim}{a} \\ \underset{\sim}{b} \end{pmatrix} - \underset{\sim}{b}^T V_{22}^{-1} \underset{\sim}{b}$$

$$= \underset{\sim}{a}^T (V_{11} - V_{12} V_{22}^{-1} V_{21})^{-1} \underset{\sim}{a} + 2 \underset{\sim}{a}^T F^{-1} V_{12} V_{22}^{-1} \underset{\sim}{b}$$

$$+ \underset{\sim}{b}^T (V_{22}^{-1} + V_{22}^{-1} V_{21} F^{-1} V_{12} V_{22}^{-1}) \underset{\sim}{b} - \underset{\sim}{b}^T V_{22}^{-1} \underset{\sim}{b}$$

$$= \underset{\sim}{a}^T \left(V_{11} - V_{12} V_{22}^{-1} V_{21}\right)^{-1} \underset{\sim}{a} + 2 \underset{\sim}{a}^T F^{-1} V_{12} V_{22}^{-1} \underset{\sim}{b}$$

$$+ \underset{\sim}{b}^T V_{22}^{-1} V_{21} F^{-1} V_{12} V_{22}^{-1} \underset{\sim}{b}$$

$$= \left(\underset{\sim}{a} - V_{12} V_{22}^{-1} \underset{\sim}{b}\right)^T \left(V_{11} - V_{12} V_{22}^{-1} V_{21}\right)^{-1} \left(\underset{\sim}{a} - V_{12} V_{22}^{-1} \underset{\sim}{b}\right)$$

$$+ \underset{\sim}{b}^T V_{22}^{-T} V_{21} \left(V_{11} - V_{12} V_{22}^{-1} V_{21}\right)^{-1} V_{12} V_{22}^{-1} \underset{\sim}{b}$$

$$-2 \underset{\sim}{a}^T \left(V_{11} - V_{12} V_{22}^{-1} V_{21}\right)^{-1} V_{12} V_{22}^{-1} \underset{\sim}{b} + 2 \underset{\sim}{a}^T F^{-1} V_{12} V_{22}^{-1} \underset{\sim}{b}$$

$V_{12} V_{22}^{-1} (\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2)$

$$+ \underset{\sim}{b}^T V_{22}^{-1} V_{21} F^{-1} V_{12} V_{22}^{-1} \underset{\sim}{b}$$

$$= \left(\underset{\sim}{a} - V_{12} V_{22}^{-1} \underset{\sim}{b}\right)^T \left(V_{11} - V_{12} V_{22}^{-1} V_{21}\right)^{-1} \left(\underset{\sim}{a} - V_{12} V_{22}^{-1} \underset{\sim}{b}\right)$$

$$= \left(\underset{\sim}{x}_1 - \underset{\sim}{\mu}_1 - V_{12} V_{22}^{-1} (\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2)\right)^T \left(V_{11} - V_{12} V_{22}^{-1} V_{21}\right)^{-1} \left(\underset{\sim}{x}_1 - \underset{\sim}{\mu}_2 - V_{12} V_{22}^{-T} (\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2)\right)$$

So we have

$$\frac{f(\underset{\sim}{x}_1, \underset{\sim}{x}_2)}{f_{\underset{\sim}{X}_2}(\underset{\sim}{x}_2)} = (2\pi)^{-\frac{p_1}{2}} \left| V_{11} - V_{12} V_{22}^{-1} V_{21} \right|$$

$$\exp\left[-\frac{1}{2} \left(\underset{\sim}{x}_1 - (\underset{\sim}{\mu}_1 + V_{12} V_{22}^{-1} (\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2))\right)^T \left(V_{11} - V_{12} V_{22}^{-1} V_{21}\right)^{-1} \left(\underset{\sim}{x}_1 - (\underset{\sim}{\mu}_1 + V_{12} V_{22}^{-1} (\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2))\right) \right],$$

giving

$$\underset{\sim}{X}_1 \mid \underset{\sim}{X}_2 = \underset{\sim}{x}_2 \sim \text{Normal}\left(\underset{\sim}{\mu}_1 + V_{12} V_{22}^{-1} (\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2),\ V_{11} - V_{12} V_{22}^{-1} V_{21} \right).$$

1. Let V be a subspace of $\mathbb{R}^n$ and let $u \in V^\perp$. Show that $\text{proj}_V u = 0$.

Solution: Let $\{v_1, \dots, v_p\}$ be an orthogonal basis for V.

Recall that $u \in V^\perp$ means u is orthogonal to every vector in V.

So $u \cdot v_j = 0$ for all $j = 1, \dots, p$.

Therefore we have

$$\text{proj}_V u = \frac{v_1^T u}{v_1^T v_1} v_1 + \dots + \frac{v_p^T u}{v_p^T v_p} v_p = 0$$