SOME PROBLEMS ENCOUNTERED IN LINEAR MODEL THEORY

by

Walter T. Federer

Biometrics Unit, Cornell University, Ithaca, NY 14853

BU-1258-M* August 1994

ABSTRACT

In the course of applying linear model theory, two situations were encountered wherein it appears that gaps in the theory were discovered. The first case involved the statistical analysis for a diallel crossing treatment design with the lines falling into different maturity groups. The problem encountered was that the usual procedure for obtaining sums of squares for a set of regressions after eliminating others broke down. The second situation involved the use of Lagrangian multipliers with restrictions to bring the model to full rank. The restriction used was that the sum of the effects equals zero. Whenever a solution was possible, the value of the Lagrangian multiplier did not appear to affect the solution for effects or the variance of an estimable quantity. However, the value of measures involving the determinants of variance-covariance matrices was affected. This brings up the problem of what is the "correct" value for a Lagrangian multiplier to obtain the "correct" value for such statistical design optimality measures as D-optimality. It does not appear that this topic has been discussed in linear model texts.

* In the Technical Report Series of the Biometrics Unit, Cornell University, Ithaca, New York 14853.

1. INTRODUCTION

In the course of finding a statistical analysis for a diallel crossing experiment wherein the lines fell into maturity groups, it was found that the partitioning of sums of squares was not invariant with respect to the treatment design. The experiment was a randomized complete block design for which the treatments, the blocks and the overall mean are orthogonal to each other. It was found that standard regression procedures could not be used since the orthogonality of some treatments effects and the general mean effect changed with increasing complexity of the treatment design and consequently the response model. This would change the residual sum of squares in an orthogonal experiment design and, of course, this is unrealistic. This situation is detailed in the following section as Situation One and was found by Federer *et al.* (1993) in constructing a statistical analysis for a plant breeding experiment.

A second situation was encountered in efficiency measures used to compare various classes of repeated measures experiment designs carried on for p periods (see Federer, 1993). Such measures as D-optimality and others utilize determinants of variance-covariance matrices of effects or of variance-covariance matrices of estimable quantities of effects as a measure of efficiency for an experiment design or a class of experiment designs. To obtain a solution for the various effects, the restriction used was that the sum of the effects equals zero. The problem was the value of the various Lagrangian multipliers used in constructing the variance-covariance matrix. What values of Lagrangian multipliers are permissible when one uses such measures as D-optimality? The examples used indicate that when a solution is possible, the solution for the effects and the variances of estimable quantities of effects are invariant to the values of the Lagrangian multipliers, but this is not true for the determinants of the variance-covariance matrices. This problem is discussed in the Sections 3, 4, and 5. Some conclusions from this study are given in Section 6.

2. SITUATION ONE – DIALLEL CROSSING EXPERIMENT

A linear model commonly used for a randomized complete block design with the t treatments each appearing once in each of the b blocks is

$$Y_{ij} \;=\; \mu + \beta_i + \tau_j + \epsilon_{ij}\,, \tag{2.1}$$

where Y_{ij} is the response of treatment j, j = 1, 2, ⋯, t, in block i, i = 1, 2, ⋯, b, μ is a general mean effect, β_i is the effect of the ith block, τ_j is the effect of the jth treatment, and ϵ_{ij} is a random error effect distributed with mean zero and common variance σ^2_ϵ. The solution for μ is the arithmetic mean $\bar{y}_{..}$ when the restrictions are that the sum of the block and of the treatment effects are zero, and the effects are all orthogonal to each other.

Next, change the treatment design to be that for a diallel crossing experiment as follows:

$$Y_{ide} = \mu + \beta_i + \gamma_d + \gamma_e + \delta_{de} + \epsilon_{ide} \tag{2.2}$$

where γ_d is a general combining ability effect (gca) for line d when crossed with all other lines, δ_{de} is a specific combining ability effect (sca) of line d crossed with line e, and the other effects are as described for equation (2.1). A partitioning of the treatment sum of squares from (2.1) is a sum of squares due to general combining ability effects with v−1 degrees of freedom and a sum of squares for specific combining ability with v(v−3)/2 degrees of freedom, for t = v(v−1)/2 crosses. This partitioning is given in Table 2.1. There are no problems encountered with either (2.1) or (2.2). The total sum of squares minus the reduction due to the mean, block, and treatments regressions, $RR\left(\mu,\ \beta_i,\ \tau_i\right) = RR\left(\mu,\ \beta_i,\ \gamma_d,\ \delta_{de}\right)$, results in the sum of squares due to error.

A problem is encountered when the treatment design is changed slightly. Suppose that the t lines fall into g maturity groups, say early, medium, and late, with all possible crosses, i.e., v(v−1)/2 = t. Various response models could be formulated but the one we consider is:

$$Y_{ibcde} = \mu + \beta_i + \pi_b + \pi_c + \phi_{bc} + \gamma_{bcd} + \gamma_{bce} + \delta_{bcde} + \epsilon_{ibcde}\,, \tag{2.3}$$

where π_b is a general combining effect for group b, b = 1, 2, ⋯, g, ϕ_{bc} is a specific combining ability effect for group b with group c, and the other effects are as defined for model (2.2) for group bc. There are g(g+1)/2 groups. The gca and sma effects are considered to be peculiar to group bc.

Table 2.1. Analysis of variance for (2.1) and (2.2).

Source of variation	Degrees of freedom	Sum of squares	Mean square
Total	$bt = bv(v-1)/2$	$\sum_{i=1}^{b}\sum_{j=1}^{t} Y_{ij}^2$	
Correction for mean	1	$Y_{..}^2/bt = C$	
Blocks	$b-1$	$\sum_{i=1}^{b} Y_{i.}^2/t - C = B$	$B/(b-1)$
Treatments	$t-1$	$\sum_{j=1}^{t} Y_{.j}^2/b - C = T$	$T/(t-1)$
gma	$v-1$	$4\sum_{d=1}^{v}\left(vY_{.d.}/2 - Y_{...}\right)^2/(bv^2(v-2) = G$	$G/(v-1)$
sma	$v(v-3)/2$	$\sum_{d<}\sum_{e=2}^{v} Y_{.de}^2/b - \sum_{d=1}^{v} Y_{.d.}^2/b(v-2)$ $+2Y_{...}^2/b(v-1)(v-2) = S$	$2S/v(v-3)$
Block $\times$ treatment = error	$(b-1)(t-1)$	$\sum_{i=1}^{b}\sum_{j=1}^{t}\left(Y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}\right)^2 = E$	$E/(b-1)(t-1)$

Alternatively, they could be considered to be the same across groups as Hinkelmann (1974) does but this will not change the problem in linear model theory encountered for this treatment design and response model. An analysis of variance table for (2.3) is given in Table 2.2 for three maturity groups, early (E), medium (M), and late (L). The problem is an orthogonal partitioning of the among-groups sum of squares into that due to general combining ability effects for groups and that due to specific group combining ability effects.

Standard regression theory says that the sum of squares due to a set of regression coefficients is the sum of the products of the solutions for regression coefficients times the "right-hand sides" of the normal equations. Alternatively, it is the regression sum of squares for all regressions minus the regression sum of squares for all regressions except the ones of interest. That is, for our case, $\mathrm{RR}\left(\mu,\ \beta_i,\ \pi_b\right) - \mathrm{RR}\left(\mu,\ \beta_i\right) \neq$ regression coefficients times right-hand sides of the normal equations. The reason they are not equal is that the arithmetic mean is an estimate of μ in $\mathrm{RR}(\mu,\ \beta_i)$ but is not in the quantity $\mathrm{RR}(\mu,\ \beta_i,\ \pi_b)$. But since this is an orthogonal design, the arithmetic mean is the only reasonable solution and the resulting sum of squares can exceed the among-groups sum of squares, again an unreasonable situation. The change in the meaning of a parameter in an experiment design from that in a treatment design does not appear to be covered by standard regression procedures in linear model theory and is the problem encountered here.

SITUATION TWO – EFFICIENCY MEASURES

Consider that the goal is to compare various experiment designs (ED) using determinants of variance-covariance matrices and using the restrictions that the sum of the effects equals zero. The ED where problems occurred (see Federer, 1993) was a repeated measures (RM) ED with the following linear model:

$$Y_{ghij} = \mu + \pi_g + \delta_h + \tau_i + \rho_j + \epsilon_{ghij}\ , \tag{3.1}$$

where μ is a general mean effect, π_g is an effect associated with the gth period, g = 1, 2, $\cdots$, p periods, δ_h is an effect associated with the hth sequences, h = 1, 2, $\cdots$, s sequences, τ_i is an effect associated with the ith treatment, i = 1, 2, $\cdots$, t treatments, ρ_j is an effect associated with the jth residual (first

Table 2.2. Analysis of variance for (2.3) for $g(g+1)/2 = 6$ groups, early (E), medium (M), and late (L) maturity groups with n_e, n_m, and n_l lines, respectively.

Source of variation	Degrees of freedom	Sum of squares	Mean square
Total	bt		
Correction for mean	1		
Blocks	$b-1$	See Table 2.1.	
Treatments	$t-1$		
Among groups	$g(g+1)/2-1 = 5$	Usual one-way unequal numbers	
Group gca	$g-1 = 2$	(See Federer *et al.*, 1993)	
Group sca	$g(g-1)/2 = 3$	(See Federer *et al.*, 1993)	
Within groups	$t-6$	Usual one-way analysis	
E by E	$n_e(n_e-1)/2-1$	See Table 2.1.	
gca(E)	n_e-1		
sca(E)	$n_e(n_e-3)/2$		
M by M	$n_m(n_m-1)/2-1$	See Table 2.1.	
gca(M)	n_m-1		
sca(M)	$n_m(n_m-3)/2$		
L by L	$n_l(n_l-1)/2-1$	See Table 2.1.	
gca(L)	n_l-1		
sca(L)	$n_l(n_l-3)/2$		
E by M	$n_e n_m-1$	Usual two-way analysis	
gcaE(M)	n_e-1		
gcaM(E)	n_m-1		
sca(E by M)	$(n_e-1)(n_m-1)$		
E by L	$n_e n_l-1$	Usual two-way analysis	
gcaE(L)	n_e-1		
gcaL(E)	n_l-1		
sca(E by L)	$(n_e-1)(n_l-1)$	Usual two-way analysis	
M by L	$n_m n_l-1$		
gcaM(L)	n_m-1		
gcaL(M)	n_l-1		
sca(M by L)	$(n_m-1)(n_l-1)$		
Block × treatment $(b-1)(t-1)$ = error		See Table 2.1.	

period carryover) effect, and ϵ_{ghij} is a random error effect distributed with mean zero and variance σ_ϵ^2. For those RMEDs for which the period and mean effects are orthogonal to the other effects (see Federer, 1993), the normal equations in matrix form, $\mathbf{X'X}\beta = \mathbf{Y}^*$, are:

$$\begin{bmatrix} p\mathbf{I}_s & \mathbf{A}_{s\times t} & \mathbf{B}_{s\times t} \\ \mathbf{A}'_{t\times s} & cp\mathbf{I}_t & \mathbf{C}_{t\times t} \\ \mathbf{B}'_{t\times s} & \mathbf{C}'_{t\times t} & c(p\text{-}1)\mathbf{I}_t \end{bmatrix} \begin{bmatrix} \delta_{s\times 1} \\ \tau_{t\times 1} \\ \rho_{t\times 1} \end{bmatrix} = \begin{bmatrix} \left[Y_{\cdot h\cdot\cdot} - p\bar{y}_{\cdots\cdot}\right]_{s\times 1} \\ \left[Y_{\cdot\cdot i\cdot} - cp\bar{y}_{\cdots\cdot}\right]_{t\times 1} \\ \left[Y_{\cdots j} - c\bar{y}_{1\cdots} - c(p\text{-}2)\bar{y}_{\cdots\cdot}\right]_{t\times 1} \end{bmatrix},$$

where s = ct and $\mathbf{I}_s$ is the identity matrix. In order to obtain unique solutions for $\mathbf{X'X}\beta = \mathbf{Y}^*$, we proceed as follows:

$$\left[\begin{bmatrix} p\mathbf{I}_s & \mathbf{A}_{s\times t} & \mathbf{B}_{s\times t} \\ \mathbf{A}'_{t\times s} & cp\mathbf{I}_t & \mathbf{C}_{t\times t} \\ \mathbf{B}'_{t\times s} & \mathbf{C}'_{t\times t} & c(p\text{-}1)\mathbf{I}_t \end{bmatrix} - \begin{bmatrix} \mathbf{0}_{s\times s} & \lambda_a\mathbf{J}_{s\times t} & \lambda_b\mathbf{J}_{s\times t} \\ \lambda_d\mathbf{J}_{t\times s} & \mathbf{0}_{t\times t} & \lambda_c\mathbf{J}_{t\times t} \\ \lambda_e\mathbf{J}_{t\times s} & \lambda_f\mathbf{J}_{t\times t} & \mathbf{0}_{t\times t} \end{bmatrix}\right] \begin{bmatrix} \delta \\ \tau \\ \rho \end{bmatrix} = \mathbf{Y}^* . \tag{3.2}$$

Note that the $\mathbf{J}$ matrix, which is subtracted from $\mathbf{X'X}$, has a zero determinant for all λ_a to λ_f. Also, the restrictions $\lambda_a \mathbf{1}'_{1\times t}\tau = \lambda_b \mathbf{1}'_{1\times t}\rho = 0$, etc., are the restrictions being used to obtain solutions. Multiples, λ_a to λ_f, of zero are being added. In obtaining solutions for the vector β, it would appear that any nonzero value of λ_a, λ_b, and λ_c or any value of λ_d, λ_e, or λ_f would result in the *same* solution for the effects. This result is corroborated by the numerical example in the following section.

However, the values of λ_a to λ_f do affect the values of determinants of variance-covariance matrices as shown by the following numerical example. This raises questions about measures of optimality of designs such as D-optimality and conditions on the Lagrangian multipliers λ_a to λ_f which yield the "correct" solution, whatever "correct" means. The matrices studied are obtained from (3.2) and rewritten as

$$\begin{bmatrix} p\mathbf{I}_s & \mathbf{A}^*_{s\times t} & \mathbf{B}^*_{s\times t} \\ \mathbf{A}^{*\prime}_{s\times t} & cp\mathbf{I}_t & \mathbf{C}^*_{t\times t} \\ \mathbf{B}^{*\prime}_{s\times t} & \mathbf{C}^{*\prime}_{t\times t} & c(p\text{-}1)\mathbf{I}_t \end{bmatrix} \begin{bmatrix} \delta \\ \tau \\ \rho \end{bmatrix} = \mathbf{Y}^* . \tag{3.3}$$

Let

$$\mathbf{v} = \left[c(p\text{-}1)\mathbf{I}_t - \begin{bmatrix}\mathbf{B}^{*\prime} & \mathbf{C}^{*\prime}\end{bmatrix} \begin{bmatrix} p\mathbf{I}_s & \mathbf{A}^{*\prime} \\ \mathbf{A}^{*\prime} & cp\mathbf{I}_t \end{bmatrix}^{-1} \begin{bmatrix}\mathbf{B}^* \\ \mathbf{C}^*\end{bmatrix} \right]^{-1} ,$$

$$\mathbf{r} = \left[cp\mathbf{I}_t - \begin{bmatrix}\mathbf{B}^{*\prime} & \mathbf{C}^{*\prime}\end{bmatrix} \begin{bmatrix} p\mathbf{I}_s & \mathbf{B}^{*\prime} \\ \mathbf{B}^{*\prime} & c(p\text{-}1)\mathbf{I}_t \end{bmatrix}^{-1} \begin{bmatrix}\mathbf{B}^* \\ \mathbf{C}^*\end{bmatrix} \right]^{-1} ,$$

$$\mathbf{x} = \left[\begin{bmatrix} cp\mathbf{I}_t & \mathbf{C}^{*\prime} \\ \mathbf{C}^{*\prime} & c(p\text{-}1)\mathbf{I}_t \end{bmatrix} - \frac{1}{p} \begin{bmatrix}\mathbf{A}^{*\prime} \\ \mathbf{B}^{*\prime}\end{bmatrix} \begin{bmatrix}\mathbf{A}^* & \mathbf{B}^*\end{bmatrix} \right]^{-1} ,$$

and

$$\mathbf{K}'_{(t\text{-}1)\times t} = \text{Helmert matrix of contrasts} .$$

The determinants of $\mathbf{v}$, $\mathbf{r}$, $\mathbf{x}$, $\mathbf{K'vK} = \mathbf{K'ZK}$, $\mathbf{K'rK} = \mathbf{K'WK}$, and $\mathbf{K'xK} = \mathbf{K'UK}$ are those computed using the GAUSS program in the appendix. They refer to variance-covariance matrices for residual, direct, and residual plus direct effects. Use of the Helmert matrix $\mathbf{K}'$ considers estimable contrasts of the $\hat{\tau}_i$ and $\hat{\rho}_j$.

The following questions arise and do not appear to be answered by linear model theory:

1) What are the conditions on λ_a to λ_f making D-optimality an appropriate measure of ED efficiency?

2) Using the Kershner efficiency measure, i.e., determinants of $\mathbf{K'ZK}$, $\mathbf{K'WK}$, and $\mathbf{K'UK}$, under what conditions on λ_a to λ_f will the same results be obtained?

3) Will the solution for effects $\hat{\boldsymbol{\beta}}$ change for arbitrary values of λ_a to λ_f, or as the example indicates, is the solution invariant with respect to λ_a to λ_f?

None of the above appears to be answered by Searle (1971), Section 5.6. Using his notation, restrictions are added using

$$\mathbf{Pb} = \boldsymbol{\delta} \tag{95}$$

where $\mathbf{P}$ is the $\mathbf{J}$ matrix used above. Quoting from Searle, "**P** has full rank q. ⋯ Fitting this restricted model leads, just as in (71), to

$$\mathbf{X'X}\mathbf{b}_r^o + \mathbf{P'}\boldsymbol{\theta} = \mathbf{X'y}$$

and

$$\mathbf{P'b}_r^o = \boldsymbol{\delta} ,$$

where $2\boldsymbol{\theta}$ is a vector of Lagrangian multipliers, and the subscript r or $\mathbf{b}_r^o$ denotes that $\mathbf{b}_r^o$ is a solution to the normal equations of the restricted model."

It is not clear that the above is an answer to the queries raised herein. From the numerical example, it appears that $\lambda_a = \lambda_d$ is more important than the other lambdas in affecting values of the various determinants. Why would this be so?

4. NUMERICAL EXAMPLE OF A RMED, p = s = t = 4

For this numerical example, use the following set of effects to obtain the numerical values:

$\mu = 10$	$\pi_1 = 0$	$\delta_1 = 1$	$\tau_A = -1$	$\rho_A = -1$
	$\pi_2 = 0$	$\delta_2 = 2$	$\tau_B = 1$	$\rho_B = 1$
	$\pi_3 = 0$	$\delta_3 = 3$	$\tau_C = 2$	$\rho_C = 0$
	$\pi_4 = 0$	$\delta_4 = -6$	$\tau_D = -2$	$\rho_D = 0$

The ED and Y_{ghij} values are:

	Sequence								
Period	1		2		3		4		Totals
1	A	10	B	13	C	15	D	2	40
2	A	9	B	14	C	15	D	2	40
3	D	8	A	12	B	14	C	6	40
4	D	9	A	10	B	15	C	6	40
Totals		36		49		59		16	160

The vector of $\mathbf{Y}^*$ values after removing mean and period effects is:

$$\begin{bmatrix} Y_{\cdot 1 \cdot\cdot} - 4\bar{y}_{\cdots\cdot} & = 36 - 40 = & -4 \\ Y_{\cdot 2 \cdot\cdot} - 4\bar{y}_{\cdots\cdot} & = 49 - 40 = & 9 \\ Y_{\cdot 3 \cdot\cdot} - 4\bar{y}_{\cdots\cdot} & = 59 - 40 = & 19 \\ Y_{\cdot 4 \cdot\cdot} - 4\bar{y}_{\cdots\cdot} & = 16 - 40 = & -24 \\ Y_{\cdot\cdot A \cdot} - 4\bar{y}_{\cdots\cdot} & = 41 - 40 = & 1 \\ Y_{\cdot\cdot B \cdot} - 4\bar{y}_{\cdots\cdot} & = 56 - 40 = & 16 \\ Y_{\cdot\cdot C \cdot} - 4\bar{y}_{\cdots\cdot} & = 42 - 40 = & 2 \\ Y_{\cdot\cdot D \cdot} - 4\bar{y}_{\cdots\cdot} & = 21 - 40 = & -19 \\ Y_{\cdots A} - \bar{y}_{1\cdots} - 2\bar{y}_{\cdots\cdot} & = 27 - 30 = & -3 \\ Y_{\cdots B} - \bar{y}_{1\cdots} - 2\bar{y}_{\cdots\cdot} & = 41 - 30 = & 11 \\ Y_{\cdots C} - \bar{y}_{1\cdots} - 2\bar{y}_{\cdots\cdot} & = 35 - 30 = & 5 \\ Y_{\cdots D} - \bar{y}_{1\cdots} - 2\bar{y}_{\cdots\cdot} & = 17 - 30 = & -13 \end{bmatrix} = \mathbf{Y}^* .$$

$$
\mathbf{Z}\boldsymbol{\beta} = \left[\begin{array}{cccc|cccc|cccc}
4 & 0 & 0 & 0 & 2 & 0 & 0 & 2 & 2 & 0 & 0 & 1 \\
0 & 4 & 0 & 0 & 2 & 2 & 0 & 0 & 1 & 2 & 0 & 0 \\
0 & 0 & 4 & 0 & 0 & 2 & 2 & 0 & 0 & 1 & 2 & 0 \\
0 & 0 & 0 & 4 & 0 & 0 & 2 & 2 & 0 & 0 & 1 & 2 \\
\hline
2 & 2 & 0 & 0 & 4 & 0 & 0 & 0 & 2 & 1 & 0 & 0 \\
0 & 2 & 2 & 0 & 0 & 4 & 0 & 0 & 0 & 2 & 1 & 0 \\
0 & 0 & 2 & 2 & 0 & 0 & 4 & 0 & 0 & 0 & 2 & 1 \\
2 & 0 & 0 & 2 & 0 & 0 & 0 & 4 & 1 & 0 & 0 & 2 \\
\hline
2 & 1 & 0 & 0 & 2 & 0 & 0 & 1 & 3 & 0 & 0 & 0 \\
0 & 2 & 1 & 0 & 1 & 2 & 0 & 0 & 0 & 3 & 0 & 0 \\
0 & 0 & 2 & 1 & 0 & 1 & 2 & 0 & 0 & 0 & 3 & 0 \\
1 & 0 & 0 & 2 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 3
\end{array}\right]
\begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \tau_A \\ \tau_B \\ \tau_C \\ \tau_D \\ \rho_A \\ \rho_B \\ \rho_C \\ \rho_D \end{bmatrix}
= \begin{bmatrix} 4\mathbf{I} & \mathbf{A} & \mathbf{B} \\ \mathbf{A}' & 4\mathbf{I} & \mathbf{C} \\ \mathbf{B}' & \mathbf{C}' & 4\mathbf{I} \end{bmatrix}
\begin{bmatrix} \boldsymbol{\delta} \\ \boldsymbol{\tau} \\ \boldsymbol{\rho} \end{bmatrix} .
$$

In order to obtain unique solutions, apply the restrictions $\lambda_a \sum \hat{\tau}_i = \lambda_b \sum \hat{\rho}_j = 0$, where λ_a, λ_b, and λ_c are scalars. The matrix form of the restriction is

$$
\mathbf{J}\boldsymbol{\beta} = \begin{bmatrix} \mathbf{0}_4 & \lambda_a \mathbf{J}_4 & \lambda_b \mathbf{J}_4 \\ \lambda_d \mathbf{J}_4 & \mathbf{0}_4 & \lambda_c \mathbf{J}_4 \\ \lambda_e \mathbf{J}_4 & \lambda_f \mathbf{J}_4 & \mathbf{0}_4 \end{bmatrix}
\begin{bmatrix} \boldsymbol{\delta} \\ \boldsymbol{\tau} \\ \boldsymbol{\rho} \end{bmatrix} ,
$$

where $\mathbf{0}_4$ is a 4×4 matrix of zeros and $\mathbf{J}_4$ is a 4×4 matrix of ones. For various values of λ_a to λ_f, solutions of the effects for the following values gave the ***same*** solution for effects as was used in constructing the example:

λ_a	λ_b	λ_c	λ_d	λ_e	λ_f
1	1	1	1	1	1
1	1	1	0	0	0
2	1	1	2	1	1
2	2	1	2	2	1
2	2	2	2	2	2
3	3	3	3	3	3

Thus it would appear that the solution for effects does not depend upon the values of the λ_a to λ_f that are used.

However, in constructing measures of D-optimality the values of λ_a to λ_f do have an effect on the values of the determinant. The values (times 1,000) obtained for the various determinants given in the appended computer program are: $\lambda_a = \lambda_d$, $\lambda_b = \lambda_e$, and $\lambda_c = \lambda_f$.

λ_a	λ_b	λ_c	|v|	|r|	|x|	|K'ZK|	|K'WK|	|K'UK|
1	1	1	71	26	349	711	374	1,596
1	2	1	-51	22	- 78	711	374	1,596
1	1	2	-51	-18	-478	711	374	1,596
1	2	2	-19	8	-113	711	374	1,596
2	1	1	*	*	*	*	*	*
2	2	2	*	*	*	*	*	*
1	3	1	-10	23	64	711	374	1,596
1	3	2	- 8	17	14	711	374	1,596
1	3	3	- 5	11	-13	711	374	1,596
3	3	3	4	3	-19	711	374	1,596

* Means no solution.

5. ANOTHER EXAMPLE

For this example, we shall look at the inverses and determinants of $\mathbf{X'X-J = Z-J}$ and consider β as a single set of parameters rather than as three sets of parameters. Using the same matrix as in the numerical example, the solution vector $\hat{\beta}$, the GAUSS program used, the **J** matrix, the determinant of the **Z−J** matrix, the inverse of **Z−J**, and the determinant of the inverse of **Z−J** are given in Table 5.1.

Using all lambdas equal to one (case 1) and using $\lambda_a = \lambda_d = 2$, $\lambda_b = \lambda_c = \lambda_e = \lambda_f = 1$ (case 2), the values listed in Table 5.1 were obtained with the following results:

1. The solutions for $\hat{\beta}$ were identical for both cases, except for rounding errors.
2. The determinants of **Z−J** for cases 1 and 2 were $2.304(10^5)$ and −92160, respectively. (Note (−.4)(230400) = 92160 or −2.5(−92160) = 230400, i.e., the values of the determinants are multiples of each other.)
3. The determinants of $(\mathbf{Z-J})^{-1}$ for cases 1 and 2 were $4.340(10^{-6})$ and $-1.685(10^{-5})$, respectively.
4. All elements in the inverses differ.
5. Variances of differences, such as $\hat{\rho}_C - \hat{\rho}_D$, are the same for both inverses.

Table 5.1. Program and values for **J**, $|\mathbf{J}|$, $(\mathbf{Z}-\mathbf{J})^{-1}$, and $|(\mathbf{Z}-\mathbf{J})^{-1}|$.

```
@This is a program for studying restrictions on a linear model.@
format 3,0:
let Z[12, 12] = 4 0 0 0   2 0 0 2   2 0 0 1   0 4 0 0   2 2 0 0   1 2 0 0
                0 0 4 0   0 2 2 0   0 1 2 0   0 0 0 4   0 0 2 2   0 0 1 2
                2 2 0 0   4 0 0 0   2 1 0 0   0 2 2 0   0 4 0 0   0 2 1 0
                0 0 2 2   0 0 4 0   0 0 2 1   2 0 0 2   0 0 0 4   1 0 0 2
                2 1 0 0   2 0 0 1   3 0 0 0   0 2 1 0   1 2 0 0   0 3 0 0
                0 0 2 1   0 1 2 0   0 0 3 0   1 0 0 2   0 0 1 2   0 0 0 3;
let 04[4, 4] =  0 0 0 0   0 0 0 0   0 0 0 0   0 0 0 0;
let J4[4, 4] =  1 1 1 1   1 1 1 1   1 1 1 1   1 1 1 1;
let Y[12, 1] = -4 9 19 -24   1 16 2 -19   -3 11 5 -13;
J=(04~J4~J4) | (J4~04~J4) | (J4~J4~04); J ;
b=inv(Z - J)*Y; b';
format 3, 3;
det2=det(Z - J); det2;
D=inv(Z - J); D;
det1=det(D); det1;
```

```
  0   0   0   0   1   1   1   1   1   1   1   1
  0   0   0   0   1   1   1   1   1   1   1   1
  0   0   0   0   1   1   1   1   1   1   1   1
  0   0   0   0   1   1   1   1   1   1   1   1
  1   1   1   1   0   0   0   0   1   1   1   1
  1   1   1   1   0   0   0   0   1   1   1   1
  1   1   1   1   0   0   0   0   1   1   1   1
  1   1   1   1   0   0   0   0   1   1   1   1
  1   1   1   1   1   1   1   1   0   0   0   0
  1   1   1   1   1   1   1   1   0   0   0   0
  1   1   1   1   1   1   1   1   0   0   0   0
  1   1   1   1   1   1   1   1   0   0   0   0
  1   2   3  -6  -1   1   2  -2  -1   1 -3E-016 2E-016
2.304E+005

0.429 -5.204E-018 -0.154 -1.735E-018 -0.113 0.083 0.138 -0.083 -0.083 0.133
0.083 -0.033
8.674E-018 0.429 -1.735E-018 -0.154 -0.083 -0.113 0.083 0.138 -0.033 -0.083
0.133 0.083
-0.154 8.674E-018 0.429 -1.735E-018 0.138 -0.083 -0.113 0.083 0.083 -0.033
-0.083 0.133
-5.204E-018 -0.154 -1.735E-018 0.429 0.083 0.138 -0.083 -0.113 0.133 0.083
-0.033 -0.083
-0.113 -0.083 0.138 0.083 0.429 2.880E-019 -0.154 -2.880E-019 -0.083 -0.033
0.083 0.133
0.083 -0.113 -0.083 0.138 -2.880E-019 0.429 4.627E-018 -0.154 0.133 -0.083
-0.033 0.083
0.138 0.083 -0.113 -0.083 -0.154 8.674E-019 0.429 -8.674E-019 0.083 0.133
-0.083 -0.033
-0.083 0.138 0.083 -0.113 -8.674E-019 -0.154 -3.469E-018 0.429 -0.033 0.083
0.133 -0.083
-0.083 -0.033 0.083 0.133 -0.083 0.133 0.083 -0.033 0.533 9.252E-019 -0.133
-1.850E-018
0.133 -0.083 -0.033 0.083 -0.033 -0.083 0.133 0.083 -1.388E-018 0.533
4.626E-019 -0.133
0.083 0.133 -0.083 -0.033 0.083 -0.033 -0.083 0.133 -0.133 -4.626E-019 0.533
1.388E-018
-0.033 0.083 0.133 -0.083 0.133 0.083 -0.033 -0.083 1.850E-018 -0.133
-9.252E-019 0.533
4.340E-006
```

Table 5.1. Program and values for **J**, |**J**|, $(\mathbf{Z}-\mathbf{J})^{-1}$, and $|(\mathbf{Z}-\mathbf{J})^{-1}|$ (continued).

```
 0   0   0   0   2   2   2   2   1   1   1   1
 0   0   0   0   2   2   2   2   1   1   1   1
 0   0   0   0   2   2   2   2   1   1   1   1
 0   0   0   0   2   2   2   2   1   1   1   1
 2   2   2   2   0   0   0   0   1   1   1   1
 2   2   2   2   0   0   0   0   1   1   1   1
 2   2   2   2   0   0   0   0   1   1   1   1
 2   2   2   2   0   0   0   0   1   1   1   1
 1   1   1   1   1   1   1   1   0   0   0   0
 1   1   1   1   1   1   1   1   0   0   0   0
 1   1   1   1   1   1   1   1   0   0   0   0
 1   1   1   1   1   1   1   1   0   0   0   0
 1   2   3  -6  -1   1   2  -2  -1   1 2E-016 2E-016
-92160.000

0.189 -0.241 -0.395 -0.241 -0.322 -0.126 -0.072 -0.293 -0.233 -0.017 -0.067
-0.183
-0.241 0.189 -0.241 -0.395 -0.293 -0.322 -0.126 -0.072 -0.183 -0.233 -0.017
-0.067
-0.395 -0.241 0.189 -0.241 -0.072 -0.293 -0.322 -0.126 -0.067 -0.183 -0.233
-0.017
-0.241 -0.395 -0.241 0.189 -0.126 -0.072 -0.293 -0.322 -0.017 -0.067 -0.183
-0.233
-0.322 -0.293 -0.072 -0.126 0.189 -0.241 -0.395 -0.241 -0.233 -0.183 -0.067
-0.017
-0.126 -0.322 -0.293 -0.072 -0.241 0.189 -0.241 -0.395 -0.017 -0.233 -0.183
-0.067
-0.072 -0.126 -0.322 -0.293 -0.395 -0.241 0.189 -0.241 -0.067 -0.017 -0.233
-0.183
-0.293 -0.072 -0.126 -0.322 -0.241 -0.395 -0.241 0.189 -0.183 -0.067 -0.017
-0.233
-0.233 -0.183 -0.067 -0.017 -0.233 -0.017 -0.067 -0.183 0.433 -0.100 -0.233
-0.100
-0.017 -0.233 -0.183 -0.067 -0.183 -0.233 -0.017 -0.067 -0.100 0.433 -0.100
-0.233
-0.067 -0.017 -0.233 -0.183 -0.067 -0.183 -0.233 -0.017 -0.233 -0.100 0.433
-0.100
-0.183 -0.067 -0.017 -0.233 -0.017 -0.067 -0.183 -0.233 -0.100 -0.233 -0.100
0.433
-1.085E-005
D:\GAUSS
```

6. CONCLUSIONS AND COMMENTS

It would appear from the observations in the first situation, Section 3, that it it is not always possible to use the reduction in the sum of squares for fitting k regressions and then subtracting the sum of squares for fitting $k-r$ regressions to obtain the appropriate sum of squares for the r regressions eliminating the effect of the first $k-r$ regressions. In some cases going from one step to another may result in a different concept of a parameter. An explanation of this phenomenon would be desirable.

In using the restriction that the sum of the effects equals zero, the question arises as to how many multiples of this restriction can one use. It would appear that

1. When a solution for effects is possible, any multiple of the restriction will give the same set of solutions for the parameters.
2. The variance of an estimable contrast is invariant to the number of multiples of the restrictions used.
3. The determinant of the information matrix and of the inverse of the matrix depend upon the number of multiples of the restrictions used.
4. The question of which set of multiples of the restrictions in item 3 is correct is unresolved.
5. The results in item 4 bring into question such measures of design efficiency as D-optimality and the need for selecting a measure which is invariant to the number of multiples of the restrictions used to effect solutions.

7. LITERATURE CITED

Federer, W. T. (1993). Comparison of classes of changeover designs. BU-12??-M in the Technical Report Series of the Biometrics Unit, Cornell University, Ithaca, N. Y. 14853, November.

Federer, W. T., Ahmed, M. B., and Viands, D. R. (1993). Expected values of mean squares for a diallel crossing experiment with maturity groups. BU-1207-M in the Technical Report Series of the Biometrics Unit, Cornell University, Ithaca, N. Y. 14853, April.

Hinkelmann, K. (1974). Two-level diallel cross experiments. Silvae Genitica **33**, 18-22.

Searle, S. R. (1971). Linear Models. John Wiley & Sons, Inc., New York, Section 5.6.

APPENDIX

```
@This is Kershner efficiency for A2 designs in 4 sequences.@
q = 1;
let p = 3  4  5  6  7  8  9  10  11  12  13  14  15  16;
let p = 4;
let K[4, 3] = 1  1  1    -1  0  0    0  -1  0    0  0  -1;
M = eye(4)|eye(4);
let B4[4, 4] = 2  0  0  1   1  2  0  0   0  1  2  0   0  0  1  2;
let A4[4, 4] = 2  0  0  2   2  2  0  0   0  2  2  0   0  0  2  2;
let C4[4, 4] = 2  1  0  0   0  2  1  0   0  0  2  1   1  0  0  2;
I4 = eye(4);
I8 = eye(8);
let J0[4, 4] =  0  0  0  0   0  0  0  0   0  0  0  0   0  0  0  0;
let J1[4, 4] =  1  1  1  1   1  1  1  1   1  1  1  1   1  1  1  1;
rowp = rows(p);
do until q > 1;
    i = 1;
    do until i > rowp;
    ea = 3*J1; eb = J1;
    A = A4 - ea;
    B = B4 - ea;
```

Ins L=1 C=1 File=D:\GAUSS\FEDFA4

```
    C = C4 - ea;
    M1 = q*(p[i, 1] - 1)*I4;
    M2 = B|C;
    M3 = (p[i, 1]*I4)~A|A'~(q*p[i, 1]*I4); format /rd 10,7;
    Z = inv(M1 - M2'inv(M3)*M2); v = det(Z); v;
    det1 = det(K'Z*K);
    M1 = q*p[i, 1]*I4;
    M2 = A|C';
    M3 = (p[i, 1]*I4)~B|B'~(q*(p[i, 1] - 1)*I4));
    Z = inv(M1 - M2'inv(M3)*M2); r = det(Z); r;
    det2 = det(K'Z*K);
    M2 = A~B;
    M3 = (q*p[i, 1]*I4)~C|C'~(q*(p[i, 1] - 1)*I4));
    Z = inv(M3 - M2'(I4/p[i, 1]*M2); x = det(M'Z*M); x;
    det3 = det(K'M'Z*M*K);
    "      p        q        det(K'ZK)        det(K'WK)        det(K'UK)";
    format /rd 7,0; p[i, 1]~q;
    format /rd 10,4; det1 det2~det3;
    print; print;
    i = i + 1;
endo;
q = q + 1;
endo;
```