Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 63(3) 2009 219-225

A Reflection on the Choice of Covariates in the Planning of Experimental Designs*

Bikas K. Sinha**
Faculty, Bayesian and Interdisciplinary Research Unit [BIRU],
Applied Statistics Division, Indian Statistical Institute, Kolkata

SUMMARY

In the context of design of experiments, reference is drawn to well-known examples involving the use of covariates. Recent advances in the Theory of Optimal Covariates' Designs suggest possibility of significant improvements through 'optimal' choice of the covariate values. A blend of theory and applications is discussed.

Key words : Optimal designs of covariates models, A-optimality, Efficiency.

1. INTRODUCTION

Optimal Designs for Covariates Models is of relatively recent research interest and preliminary results were discussed in the papers by Lopes Troya (1982a, 1982b). After a considerable gap, renewed interest in this fascinating topic was found in the work of Liski *et al.* (2002), Das *et al.* (2002) and Rao *et al.* (2003). Since then this topic has picked up a momentum and has grown quite rapidly. See a few further references, listed at the end, on theory and applications in various design settings.

The purpose of this article is to popularize this area of research in the statistical community of teachers and researchers by narrating some of the results with examples of experiments taken from standard text books.

The settings are those of ANCOVA Models in the standard design layouts such as CRDs, RBDs, LSDs, BIBDs etc. The focus is on optimal / most efficient estimation of covariates' parameters incorporated in the model(s). The combinatorially challenging problem is to accommodate maximum number of covariates in an optimal manner in different design settings. Considerable effort and attention have been paid on these issues. We do not intend to touch upon these issues here. See the references cited above. Instead, we enter into a detailed discussion of results in some standard set-ups, following examples from standard text books such as Montgomery, Cochran and Cox, Federer, Hinkelman and Kempthorne:

Example 1. Comparison of Three Machines

Study variable: breaking strength of material in lbs;
Covariate: diameter of the material in 10^{-3} inches

Example 2. Comparison of Three Hand-trucks

Study variable: delivery time in minutes;
Covariate: volume delivered [in litres]

Example 3. Comparison of Four Glue Formulations

Study variable: strength of raw material in lbs;
Covariate: thickness of raw material in 0.01 inches

Example 4. Comparison of Three Cutting Speeds
Study variable: amount of metal removed;
Covariate: hardness of the specimen

Dr. V.G. Panse Memorial Lecture delivered at 63rd Annual Conference of the Indian Society of Agricultural Statistics held at Pusa, Samastipur, Bihar during Dec. 3-5, 2009

***E-mail address* : sinhabikas@yahoo.com

Example 5. Comparison of Three Chemical Processes
Study variable: yield or final product;
Covariate: impurity of the raw material

Example 6. Comparison of Six Varieties of Corn
Study variable: yield (lbs/plot) of ear corn;
Covariate: number of plants in each plot

2. COMPLETELY RANDOMIZED DESIGNS

Consider the simplest set-up of varietal designs viz., Completely Randomized Designs [henceforth abbreviated as CRDs] involving v treatments and n experimental units, with the treatment allocation numbers $n_1, n_2, ..., n_v$ so that $n_1 + n_2 + ... + n_v = n$. Suppose there is also available a controllable covariate X, assuming values in a finite closed interval $[a, b]$ on every experimental unit in the design layout.

Our purpose in this section is to make a comparison of the designs already used [as illustrated through the examples in text books cited above] and the alternative designs that could be suggested involving one or two covariates.

3. CRD: THEORETICAL CONSIDERATIONS

In a CRD, given n and v, various optimality considerations suggest $n_1 = n_2 = ... = n_v = n/v$ whenever n is divisible by v. We assume this divisibility condition to be satisfied and set $n_0 = n/v$. Further, by a location and scale change applied to each available covariate, it is assumed that for each one, the range of covariate values is the closed interval $[-1, 1]$. Set Z as the single covariate, which possesses values $z_{i_1}, z_{i_2}, ..., z_{i_{n0}}$ on the experimental units underlying the i-th treatment. Optimality considerations suggest $z_{ij} = +1/-1$ for each (i, j) combination and at the same time, $\Sigma_j z_{ij} = 0$ for each $i = 1, 2, ..., v$. This is possible only when there are plenty of experimental units available and one can identify such units with specified values of the covariate. We again assume this is feasible in an experimental CRD set-up. Whenever n_0 is an even number, routine split-half rule can be implemented to select the experimental units underlying each treatment, possessing exactly 50-50 per cent $+1/-1$ as the associated Z-values. This ensures most efficient CRD, as regards simultaneous inference on the treatment contrasts [of the form $(\tau_i - \tau_j)$] as also on the 'regression coefficient', say γ, involved in the model.

We now refer to Example 3 dealing with four glue formulations. The data refer to the chosen experimental units having the covariate values [thickness of the glue in suitable unit] as given below:

Formulation 1 : 12, 12, 13, 14, 14

Formulation 2 : 10, 11, 12, 12, 14

Formulation 3 : 10, 11, 11, 14, 15

Formulation 4 : 10, 11, 12, 15, 16

We notice that the minimum and maximum values of the covariate values covered in the study are respectively given by 10 and 16.

Had there been only 4 experimental units under each treatment, an optimal choice of the experimental units would correspond to those having the covariate values as 10, 10, 16 and 16. However, in this case there are 5 units under each treatment but we also have 4 treatments to be compared.

We propose to develop some comparison results in this section, based on the above features of the data.

Assume n_0 to be odd and v to be an even number. We work with the location-scale changed Z-values in the closed interval $[-1, 1]$ so that '10' corresponds to '-1' and '16' corresponds to '$+1$'. The model underlying a CRD is of the form

$$y_{ij} = \mu + \tau_i + \gamma z_{ij} + e_{ij}$$

Since n_0 is odd, we set the value '+1' for $(n_o - 1)/2$ units and the value '-1' for other $(n_0 - 1)/2$ units, leaving exactly one unit with the Z-value unspecified at z_i for treatment i.

We propose to develop results pertaining to 'optimal' choice of these $z_1, z_2, ..., z_v$ values. In order that the γ parameter is estimated 'orthogonally' [i.e., orthogonal to μ), we must have $\Sigma_i z_i = 0$ and we impose this condition to start with. Further to this, information on γ is maximized whenever the z_i's assume values $+1/-1$. Since v is even, this is also possible to achieve. On the other hand, let us check the nature of the C-matrix for the treatment parameters in the general set-up with the $z_1, z_2, ..., z_v$ values. It is shown in the Appendix that

$$C_z = n_0 I_v - n_0^2 .J/n - zz'/(n - v + T(z)_2)$$
$$= n_0(I_v - J_v/v) - zz'/(n - v + T(z)_2)$$

where $z = (z_1, z_2, ..., z_v)'$, $T(z)_2 = \Sigma z_i^2$

Therefore,

$$\text{trace } (C_z) = n_0(v - 1) - \frac{T(z)_2}{n - v + T(z)_2}$$

It turns out that the trace is maximum when $T(z)_2 = 0$ i.e., when $z_i = 0$; $i = 1, 2, ..., v$. In that case, C-matrix is also completely symmetric (c.s.) and hence this choice leads to Universally Optimal [U.O.] design for treatment parameters. But on the other hand, an expression for the information on γ is given by $n - v + T(z)_2(n_0 - 1)/n_0$ which attains its maximum when $T(z)_2 = v$ i.e., when $z_i = +1/-1$ subject to $\Sigma z_i = 0$. This is a contrasting scenario. To arrive at a compromise solution, we consider the full parameter vector of the form

$$\eta = [\gamma \;\; P\tau]$$

where $P\tau$ refers to a full set of orthonormal treatment parameter contrasts.

For average variance as also for generalized variance, we need to work with the positive eigenvalues $\lambda_1, \lambda_2, ..., \lambda_{v-1}$ of C_z-matrix which are given by n_0 with multiplicity $(v - 2)$ and $n_0 - \dfrac{T(z)_2}{n - v + T(z)_2}$.

For D-optimality, we need to maximize $\lambda_1.\lambda_2...\lambda_{v-1}.(n - v + T(z)_2(1 - n_0^{-1}))$ which is equivalent to maximizing

$$\left[n_0 - \frac{T(z)_2}{n - v + T(z)_2}\right](n - v + T(z)_2\,(1 - n_0^{-1}))$$

or, equivalently,

$$[(n - v + T(z)_2(1 - n_0^{-1}))]^2/[(n - v + T(z)_2)]$$

It turns out that this last expression attains its maximum when $T(z)_2 = v$. Hence, over-all D-optimal design corresponds to $z_i = +1/-1$-values subject to $\Sigma z_i = 0$.

For A-optimality, it trivially follows that we need the same set of solutions for the z-values.

Thus, when the treatment parameter contrasts are to be estimated without any reference to the covariate parameter γ, a characterization of UO design obtains whenever $T(z)_2 = 0$ i.e., z_i's are all 0's. This is based on the properties of (i) maximum trace and (ii) cs of the C-matrix. On the other hand, taking into consideration also the γ parameter, specific optimality criteria lead to a different characterization viz., one with z_i's equal to $+1/-1$. We would tend to recommend the latter design in practice since consideration of the covariate parameter is also of importance in such models.

Turning back to the illustrative Example 3, our recommendation [Recommendation I] would have been

Formulation 1 : 10, 10, 16, 16, 16
Formulation 2 : 10, 10, 16, 16, 16
Formulation 3 : 10, 10, 10, 16, 16
Formulation 4 : 10, 10, 10, 16, 16

and it would lead to the C-matrix given by

$$C_0 = [(3.7, -1.3, -1.2, -1.2); (3.7, -1.2, -1.2); (3.7, -1.3); (3.7)]$$

whereas the original design yields the C-matrix given by

$$C_1 = [(3.63, -1.1080, -1.1954, -1.3264); (3.5822, -1.3187, -1.1597); (3.7252, -1.2152); (3.7013)]$$

The positive eigenvalues of the two C-matrices are given by [5.00, 5.00, 4.80] and [5.00, 5.00, 4.63] respectively. Hence, trivially, a comparison between the two C-matrices results in the superiority of C_0 over the other in respect of (i) the trace, (ii) the product of the eigenvalues, (iii) the sum of reciprocals of the eigenvalues and (iv) the smallest eigenvalue.

As regards estimation of the γ-parameter, it is known that information on γ in a CRD model is proportional to $SSW(z) = \Sigma\Sigma(z_{ij} - \bar{z}_i)^2$. Computations yield :

$SSW(z) = 69.20$ for the given CRD whereas this is 172.80 for the modified design.

Further to this, from theoretical considerations, we could as well recommend [Recommendation II] the design

Formulation 1 : 10, 10, 13, 16, 16
Formulation 2 : 10, 10, 13, 16, 16
Formulation 3 : 10, 10, 13, 16, 16
Formulation 4 : 10, 10, 13, 16, 16

which would be best [UO] for inference on the treatment contrasts, without any reference to the γ parameter. The C-matrix for this design is c.s. with the scalar multiplier being 5.00 and hence it has maximum trace. Another alternative to this design [Recommendation III] would be

Formulation 1 : 10, 13, 13, 13, 16
Formulation 2 : 10, 13, 13, 13, 16
Formulation 3 : 10, 13, 13, 13, 16

Formulation 4 : 10, 13, 13, 13, 16

which would have the same form of the *C*-matrix for treatment comparisons !

As to the status of these two latter designs for inference on γ-parameter, it is evident that these two alternatives should perform in between the earlier two ! Computations yield the $SSW(z)$ values for these designs respectively as 144.00 and 72.00. Therefore, gradual improvements are seen in the order 69.20 < 72.00 < 144.00 < 172.80.

Before concluding this section, we may as well mention the context of a CRD when two covariates are involved in the experiment. This time, we want the two covariate parameters to be orthogonally estimated between them as also to be so with respect to each of the treatment parameters. This can be achieved by adopting the experimental units with the following pairs of values of the covariates, based on our earlier Recommendation I:

Formulation 1 : *Lm LM Hm HM HM*

Formulation 2 : *Lm LM Hm HM Hm*

Formulation 3 : *Lm LM Hm HM LM*

Formulation 4 : *Lm LM Hm HM Lm*

In the above, '*L*' and '*H*' denote the Lowest and Highest Values of one covariate and '*m*' and '*M*' represent those of another covariate. It turns out that both the covariate parameters are orthogonally and most efficiently estimated, each with information given by 19.20, in the standardized scale wherein the range of each covariate is taken to be [–1, 1]. For the treatment parameters, we have similar results as in Recommendation I. This time there are two covariates and so the *C*-matrix will be given by

$$C_0 = [(3.65, -1.25, -1.25, -1.15); (3.65, -1.15, -1.25); (3.65, -1.25); (3.65)]$$

whose eigenvalues would be 5.00, 4.80, 4.80.

With our Recommendation II - extended to the case of two covariates, the experimental units should be selected by adhering to the pairs of covariate values as given below :

Formulation 1 : *Lm LM Hm HM NN*

Formulation 2 : *Lm LM Hm HM NN*

Formulation 3 : *Lm LM Hm HM NN*

Formulation 4 : *Lm LM Hm HM NN*

where '*N*' refers to the mid-value of the covariate range for each covariate. This time, as before, the treatment comparisons would achieve UO status while the covariate parameters would be suboptimally estimated with SSZ = 16.00 for each one measured in the standardized scale of [–1, 1], as before.

4. RANDOMIZED BLOCK DESIGNS [RBDs]

The study of optimal use of covariates in the set-up of an RBD was initiated in Das *et al.* (2002) and continued further in subsequent work by Rao *et al.* (2003). Here we will take up an illustrative example and discuss some results.

We refer to Example 6 which deals with a Latin Square Design [LSD] of order 6 involving one covariate. Here are the covariate (*X*) values across different rows :

(18, 16, 18, 14, 15, 17)

(16, 15, 16, 19, 21, 14)

(15, 16, 16, 19, 15, 17)

(18, 18, 20, 18, 22, 17)

(15, 18, 19, 16, 16, 15)

(18, 20, 19, 17, 17, 15)

We use this example for illustrative purpose and assume that instead of a Latin Square Design, it may be treated as a Randomized Block Design [RBD] involving 6 blocks and 6 treatments with the rows representing the blocks and the columns representing the treatments. (In Section 6, we will discuss results related to the Latin Square Design. Also in Section 5, we will discuss further results, suitably modifying this example into that of a Balanced Incomplete Block Design [BIBD]).

Note that the minimum and maximum values of the covariate covered in the given design are respectively given by 14 and 22. Had there been available suffcient number of 'experimental units' with only these covariate values, we would have 'profitted' substantially by using 18 of each kind in the manner explained below.

(14, 14, 14, 22, 22, 22)

(14, 14, 14, 22, 22, 22)

(14, 14, 14, 22, 22, 22)

(22, 22, 22, 14, 14, 14)

(22, 22, 22, 14, 14, 14)

(22, 22, 22, 14, 14, 14)

In this format, in terms of the 'transformed' values [14 converted to '–1' and 22 converted to '+1'], it turns out that the 36 × 1 vector of the covariate values [written in a lexicographic order, starting with the elements of the first row, for example], is 'orthogonal' to the vectors representing the block effects as also the treatment effects. Hence the presence of the covariate values does not influence the analysis of the RBD nor is the information on the covariate parameter reduced because of the block and treatment effects. Thus we have an optimal display of the covariate values for most effcient estimation of the covariate parameter. It may be mentioned in passing that the above specification is not unique for optimal estimation of the covariate parameter. It may also be noted that most efficient estimation [in terms of orthogonality with block and treatment effects parameter-vectors] may not possible in other set-ups viz., BIBDs or LSDs. These will be studied in the next two sections.

5. BALANCED INCOMPLETE BLOCK DESIGNS [BIBDs]

This time, we slightly modify the layout to serve as a Symmetric BIBD [SBIBD] with parameters $b = v = 6$, $r = k = 5$ and $\lambda = 4$. We reproduce the design layout for covariate values, considering the rows representing the blocks and columns representing the treatments and using the incidence matrix as $\mathbf{N} = \mathbf{J} - \mathbf{I}$.

(–, 16, 18, 14, 15, 17)
(16, –, 16, 19, 21, 14)
(15, 16, –, 19, 15, 17)
(18, 18, 20, –, 22, 17)
(15, 18, 19, 16, –, 15)
(18, 20, 19, 17, 17, –)

Optimality theory has been developed for situations dealing with SBIBDs having $k = r = 0$ (mod4). Here we have $r = k = 5$ and we, therefore, propose to introduce a highly efficient design. We denote the least value of the covariate by 'L' and the maximum value of the covariate by 'M'. Here we propose use of an SBIBD with the following distribution of the covariates :

($–$, L, H, L, H, L)
(L, $–$, H, L, L, H)
(L, H, $–$, H, L, L)
(H, L, H, $–$, L, H)
(H, H, L, H, $–$, L)
(H, L, L, H, H, $–$)

As mentioned before, we will take $L = -1 < 1 = M$ as the two extreme values of the covariate. The row totals of the covariate values against the blocks are $\delta_{bl} = (-1, -1, -1, 1, 1, 1)$ and those against the column totals are $\delta_{tr} = (1, -1, 1, 1, -1, -1)$. It turns out that the incidence matrix for the treatment parameters and the covariate parameter is a $(v + 1)$ $v + 1$ matrix as outlined below. It is a partitioned matrix in the form of [P Q]; [Q' R] where

P = C-matrix for the BIBD
$Q' = \delta_{tr} - k^{(-1)} \delta'_{bl} N'$
$R = 30 - 6/k$, a scalar

Hence, the information matrix for the treatment parameters, after eliminating the effect of the covariate parameter, is given by $C - D$ where $D = QR^{(-1)}Q'$. It follows that D has rank 2 and its positive eigenvalues δ_1 and δ_2 satisfy the relations : (i) sum = (6 + 6/25)/R; (ii) product = $32/25R^2$ whence the eigenvalues are given by 0.2093 and 0.0074.

In the absence of the covariate parameter in the model, the SBIBD would have constant non-zero eigenvalue of the C-matrix given by $\lambda v/k = 4.8$. Hence the eigen-values of the C-matrix in the presence of the covariates are given by 4.8, 4.8, 4.8, 4.8 – 0.2093 = 4.5907 and 4.8 – 0.0074 = 4.7926.

Therefore, efficiency of the recommended design is computed as

Efficiency = [5/4.8] / [3/4.8 + 1/4.5907 + 1/4.7926]
= 99.07 per cent

As regards the covariate parameter, in the absence of any impact of block effects and treatment effects, the information would have been 30. However, because of non-orthogonality, it gets reduced to $R - Q'P^{(-)}Q$ which simplifies to 27.712. Hence the efficiency of the design is computed as 92.37 per cent.

6. LATIN SQUARE DESIGN

Now we turn to the Example 6 [without any modification of the layout] which describes a Latin Square Design [LSD] of order 6 involving one covariate was utilized. Here are the covariate (X) values across the different rows :

(18, 16, 18, 14, 15, 17)
(16, 15, 16, 19, 21, 14)
(15, 16, 16, 19, 15, 17)
(18, 18, 20, 18, 22, 17)

(15, 18, 19, 16, 16, 15)

(18, 20, 19, 17, 17, 15)

The original LSD is slightly modified here to demonstrate the strength of the theory for optimal choice of covariate values. Here is the ORIGINAL LSD, with treatment allocations shown row-wise :

(A, B, C, D, E, F)

(B, C, D, E, F, A)

(C, E, A, F, B, D)

(D, F, B, A, C, E)

(E, D, F, B, A, C)

(F, A, E, C, D, B)

Below we display the MODIFIED LSD :

(A, B, C, D, E, F)

(B, A, D, C, F, E)

(F, E, A, B, C, D)

(E, F, B, A, D, C)

(D, C, F, E, B, A)

(C, D, E, F, A, B)

We shall work with the modified LSD. Computations yield as follows [ignoring the variance σ^2 in the expressions below] :

(i) C-matrix for the treatment parameters, after eliminating row, column and covariate effects, is given by

(4.9929, –0.9955, –0.9839, –1.0264, –1.0110, –0.9761)
(4.9971, –1.0102, –0.9832, –0.9930, –1.0152)
(4.9633, –0.9399, –0.9751, –1.0542)
(4.9014, –1.0409, –0.9110)
(4.9832, –0.9632)
(4.9197)

The eigenvalues are 6.00020, 5.99999, 5.99995, 5.99988, 5.75759.

(ii) The information on the covariate parameter is given by $I(\gamma)$ = 487.8333.

Next note that the minimum and maximum values of the covariate covered in the given design are respectively given by 14 and 22. Had there been available sufficient number of 'experimental units' with only these covariate values, we would have 'profited' substantially by using 18 of each kind in the manner explained below.

(22, 22, 22, 14, 14, 14)

(22, 22, 14, 14, 14, 22)

(22, 14, 14, 14, 22, 22)

(14, 14, 14, 22, 22, 22)

(14, 14, 22, 22, 22, 14)

(14, 22, 22, 22, 14, 14)

For the modified LSD, it turns out that with the above recommendation on the choice of experimental units with the stated values of the covariates, the C-matrix is completely symmetric with eigenvalues each equal to 6.0000 and it has maximum trace. Hence for treatment comparisons, the LSD is UO !

Further to this, the information on the covariate parameter is given by 576.00 as against 487.8333 computed above.

So, strictly speaking, the given design is quite efficient with respect to treatment comparisons but has scope for improvement in terms of precision on the covariate parameter.

Consider again the modified LSD but this time, with deletion of the last column. Then the Row Design reduces to a BIBD ($b = v = 6$, $r = k = 5$, $\lambda = 4$) while the Column Design is still an RBD. So it is a Youden Design. In the presence of a covariate (X), we can still obtain UO Youden Design for treatment comparisons provided the covariate values are chosen as follows ['H' for highest value i.e., 22 and 'L' for lowest value i.e., 14 and '0' for mid-value i.e., 18] :

(A-H, B-L, C-H, D-L, E-0)

(B-L, A-H, D-L, C-H, F-0)

(F-H, E-L, A-L, B-H, C-0)

(E-L, F-H, B-H, A-L, D-0)

(D-H, C-L, F-L, E-H, B-0)

(C-L, D-H, E-H, F-L, A-0)

The explanation is quite straightforward. The Youden Design has all its treatment parameter vectors orthogonal to the covariate parameter vector in the model description.

7. APPENDIX

We set $\theta=(\mu, \gamma, \tau_1, \tau_2, ..., \tau_v)$ and note that for the CRD with z-values [subject to $\Sigma z_i = 0$] underlying the experimental units, as indicated, the information matrix for θ is given by

$$\mathbf{I}(\theta; z) = [(n, 0, n_0, n_0, ..., n_0); (n - v + T(z)_2, z_1, z_2, ..., z_v); (n_0, 0, 0, ..., 0); ...., (n_0)]$$

From the above, we deduce that for the vector parameter τ, the C-matrix is given by

$$\begin{aligned} C_z &= n_0.\mathbf{I} - [(n_0, z_1); (n_0, z_2); ...; (n_0, z_v)] \\ &\quad [(n, 0); (n - v + T(z)_2)]^{-1} \\ &\quad [(n_0, n_0, ..., n_0); (z_1, z_2, ..., z_v)] \\ &= n_0.\mathbf{I} - n_0^2.\mathbf{J}/n - zz'/(n - v + T(z)_2) \\ &= n_0(\mathbf{I} - \mathbf{J}/v) - zz'/(n - v + T(z)_2) \end{aligned}$$

Next, let us again set θ as before but this time, we transform the vector parameter τ to $\overline{0}\tau$ where $\overline{0}$ is an orthogonal matrix with the first row vector proportional to the vector of 1's and the submatrix of order $(v - 1) \times v$ is denoted by $\mathbf{P}$ so that $\mathbf{P}$ satisfies (i) $\mathbf{P}\ 1 = 0$, (ii) $\mathbf{P}\ \mathbf{P}' = \mathbf{I}$, (iii) $\mathbf{P}'\mathbf{P} = \mathbf{I} - \mathbf{J}/v$. Clearly, $\mathbf{P}\tau$ describes a full set of orthonormal treatment effect contrasts. We intend to derive the joint information matrix of $\mathbf{P}\tau$ and γ. Set $\Gamma = [(\mu, \gamma, \overline{0}\tau]$ so that $\Gamma = \mathbf{M}\theta$ where $\mathbf{M}$ = partitioned matrix with the first part an identity matrix of order 2 and the second part the matrix $\overline{0}$.

For this, we start with the C-matrix for θ as a whole and make a transformation from τ to $\overline{0}\tau$ and then proceed to derive the desired information matrix. We set, for brevity, T_2 for $T(z)_2$.

(i) $\mathbf{I}(\theta) = \mathbf{X}'\mathbf{X}$

(ii) $\mathbf{I}(\Gamma) = \mathbf{M}\mathbf{X}'\ \mathbf{X}\ \mathbf{M}' = [(n, T_1, n_0\sqrt{(v)}, 0, 0, ..,0); (n - v + T_2(1 - n_0^{-1}), 0, z'\mathbf{P}'), (n_0, 0, 0,...,0); (n_0, 0,0, ... 0); ...; (n_0)]$

(iii) $\mathbf{I}(\gamma, \mathbf{P}\tau) = [(n - v + T_2, z'\mathbf{P}'); (n_0, 0,0, .., 0); ..., (n_0)]$

(iv) $\mathbf{I}(\gamma) = n - v + T_2 - z'\mathbf{P}'\ \mathbf{P}z/n_0$
$= n - v + T_2 (1 - n_0^{-1})$, as expected.

(v) $\mathbf{I}(\mathbf{P}\tau) = n_0\mathbf{I} - \mathbf{P}zz'\mathbf{P}'/(n - v + T_2)$

It turns out that the eigenvalues of $\mathbf{I}(\mathbf{P}\tau)$ are n_0 with multiplicity (v-2) and $n_0 - T_2(n - v + T_2)^{-1}$.

We also note that the above expression for $\mathbf{I}(\mathbf{P}\tau)$ follows from the expression for C_z given earlier in view of

$$\mathbf{I}(\mathbf{P}\tau) = \mathbf{P}\mathbf{I}(C_z)\mathbf{P}'$$

since $\quad \mathbf{P}\mathbf{P}' = \mathbf{I} - \mathbf{J}/v$ and $\mathbf{P}1 = 0$.

ACKNOWLEDGEMENTS

I sincerely thank Professor G.M. Saha for helping me with the computations and Professors N.K. Mandal, Premadhis Das, Manisha Pal and Dr. Ganesh Dutta for giving a patient hearing while I was working on the problem.

REFERENCES

Das, K., Mandal, N.K. and Sinha, Bikas K. (2002). Optimal experimental designs with covariates. *J. Statist. Plann. Inf.*, **115**, 273-285.

Dutta, G. (2004). Optimum choice of covariates in BIBD set-up. *Cal. Stat. Assoc. Bull.*, **55**, 39-55.

Das, Premadhis, Dutta, G. and Mandal, N.K. (2007). Optimum choice of covariates for a series of SBIBDs obtained through projective geometry. *Jour. Modern Appld. Stat. Methods*, **6**, 649-656.

Das, Premadhis, Dutta, G. and Mandal, N.K. (2009). Optimum covariate designs in split-plot and strip-plot design set-ups. *Jour. Appld. Stat.*, **36**.

Das, Premadhis, Dutta, G. and Mandal, N.K. (2009). Optimum covariate designs in partially balanced incomplete block (PBIB) design set-ups. *J. Statist. Plann. Inf.*

Das, Premadhis, Dutta, G. and Mandal, N.K. (2009). Optimum covariate designs in binary proper equi-replicate block design set-up. To appear in *Discrete Mathematics*.

Das, Premadhis, Dutta, G. and Mandal, N.K. (2010). D-optimal designs for co-variate parameters in block design set-up. To appear in *Comm. Statist.–Theory Methods.*

Liski, E., Mandal, N.K., Shah, K.R. and Sinha, Bikas K. (2002). *Topics in Optimal Design*. Lecture Notes Series in Statistics, No. 163, Springer.

Lopes Troya, J. (1982a). Optimal designs for covariates models. *J. Statist. Plann. Inf.*, **6**, 373-419.

Lopes Troya, J. (1982b). Cyclic designs for a covariate model. *J. Statist. Plann. Inf.,* **7**, 49-75.

Rao, PSSNVP, Rao, S.B., Saha, G.M. and Sinha, Bikas K. (2003). Optimal designs for covariates' models and mixed orthogonal arrays. *Electronic Notes in Discrete Mathematics*, **15**, 157-160, Elsevier.