APPROXIMATE DESIGN THEORY FOR A SIMPLE BLOCK DESIGN WITH RANDOM BLOCK EFFECTS

K.Christof and F.Pukelsheim
Institute of Mathematics
University of Augsburg
Federal Republic of Germany

Abstract. A simple block design with treatment effects fixed and block effects random is considered. It is shown that the sole design class which is amenable to an analysis is formed by all designs which are equi-blocksized. In this setting the block totals induce a new model, the interblock model, in which the mean vector depends linearly on the design matrix, and in which the covariance matrix is proportional to the identity matrix. The information matrix for the set of all treatment contrasts then is a convex function of the design matrix, quite distinct from the usual case where this relation is concave. The extreme points in the design space turn out to be the one-treatment-per-block designs, their information matrices depend on the treatment replication vector, only, and are called special C-matrices. We establish a disjunction which the matrix-ordering of two special C-matrices enforces on the associated replication vectors. This entails, in particular, that one-treatment-per-block designs are admissible provided no treatment gets more than one half of all observations.

1. Introduction and summary

The simple block design model with block effects random leads to a peculiar design problem since in the associated interblock model the dependence of the information matrix on the design matrix is, not concave, but convex. This is a marked difference to the usual case where this dependence is indeed concave. Here we shall discuss this

---

AMS 1980 Mathematics subject classification: Primary 62K05; 62K10.
Key words and phrases: Two-way-classification, mixed model; interblock model; one-treatment-per-block designs; convex information functions; special C-matrix; admissibility; Schur-concavity.

design problem from the approximate theory point of view, thus complementing the exact theory results obtained by Gaffke & Krafft (1980) in their 1978 Wisła Conference paper.

In Section 2 we outline the transition from the mixed effects model to the interblock model. It turns out that a manageable design problem can be formulated only among those designs which are equi-blocksized. Section 3 presents the approximate analogue of the exact results of Gaffke & Krafft (1980). As the optimization problem is one of maximizing a convex function, the one-treatment-per-block designs, i.e. the extreme points among all equi-blocksized designs, are crucial.

The information matrix for the set of all treatment contrasts of a one-treatment-per-block design depends on the treatment replication vector r, only, and takes the form $\Delta_r - rr'$ where Δ_r denotes the diagonal matrix with vector r on the diagonal. We shall call such matrices special C-matrices, they also play an important role in fixed effects models.

In Section 4 we study some implications of admissibility in the class of all special C-matrices; our results will be stated in terms of the replication vectors r which enter the definition, whereas the similar study of Giovagnoli & Wynn (1981) focused on eigenvalues. In the interblock model this yields admissibility of one-treatment-per-block designs in the class of all equi-blocksized designs, provided no treatment gets more than $1/2$ of all observations.

2. The interblock model

Consider a two-way classification $X_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$, with a fixed over-all mean $\mu \in R$, and a fixed treatment effects vector $\alpha = (\alpha_1,\dots,\alpha_v)' \in R^v$. The block effects β_j, $j = 1,\dots,b$, and the observational errors ε_{ijk} are assumed to be independent random variables with expectation 0, and variance τ^2 and σ^2, respectively. Denoting by 1_n the n-dimensional vector with all entries unity, the model has vector representation

$$X = \mu 1_n + A\alpha + B\beta + \varepsilon,$$

with known 0-1 matrices A of order $n\times v$ and B of order $n\times b$.

A design matrix N for this model is a $v \times b$ matrix with integer entries $n_{ij} \geq 0$. The treatment-block combination (i,j) does not appear in the model if $n_{ij} = 0$, and otherwise is replicated n_{ij} times, as indicated by the subscript k. The total number of observations then is $n = \Sigma\Sigma n_{ij}$. In order to discriminate between the influence of the sample size n and the distribution of observations between the various treatment-block combinations (i,j), as well as for purposes of the approximate design theory it is preferable to work with the weight matrix

$$W = N/n,$$

which is a normalized, per observation, version of the design matrix. Its marginals, i.e. row sums and column sums, equal

$$r = W1_b, \qquad s = W'1_v,$$

and are the treatment replication vector and the blocksize vector in their normalized versions, respectively.

As in Gaffke & Krafft (1980, page 135) we shall study the treatment effects through the interblock estimators based on the b block totals $Y = B'X$. Its mean vector and dispersion matrix are found to be

$$E(Y) = n\mu s + nW'\alpha = n(s:W')\binom{\mu}{\alpha},$$

$$D(Y) = n^2\tau^2\Delta_s^2 + n\sigma^2\Delta_s = \Delta_S, \text{ say.}$$

Thus Y gives rise to a linear model which we shall call the interblock model.

The components of Y are uncorrelated, but have possibly unequal variances $S_j = n^2\tau^2 s_j^2 + n\sigma^2 s_j$. The information matrix for $\binom{\mu}{\alpha}$, i.e. the inverse of the dispersion matrix of the weighted least squares estimator, turns out to be $n^2\bar{M}(W)$, with

$$\bar{M}(W) = \begin{pmatrix} s'\Delta_S^+ s & s'\Delta_S^+ W' \\ W\Delta_S^+ s & W\Delta_S^+ W' \end{pmatrix},$$

a superscript + denoting Moore-Penrose inversion. Hence any weight matrix, and also any optimal weight matrix, has information matrix depending, through S, on the sample size n, the block variance τ^2, and the error variance σ^2. This is slightly different when the designs are assumed equi-blocksized, i.e. $s = 1_b/b$. Then plainly $D(Y) = (n^2\tau^2/b^2 + n\sigma^2/b)I_b$, and simple least squares estimation leads to the information matrix $M(W)/(\tau^2/b+\sigma^2/n)$, with

$$M(W) = \begin{pmatrix} 1 & r' \\ r & bWW' \end{pmatrix}.$$

However, there seems to be no convincing reason why an optimal weight matrix should necessarily be equi-blocksized.

Namely, $\bar{M}$ is a matrix-convex function on the set of all weight matrices with block marginals given to be s, as M is convex on the set of all designs which are equi-blocksized, see Marshall & Olkin (1979, page 468). This convexity breaks down when block marginals are allowed to vary, and we shall demonstrate this failure by example. Suppose $v = 2$, and $n^2\tau^2 = 1 = n\sigma^2$. Choose $W = \begin{pmatrix} 1/4 & 1/4 \\ 1/2 & 0 \end{pmatrix}$, and $V = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$. Then

$$M(\tfrac{1}{2}W + \tfrac{1}{2}V) = \frac{1}{2145}\begin{pmatrix} 1410 & 690 & 720 \\ 690 & 362 & 328 \\ 720 & 328 & 392 \end{pmatrix},$$

$$\tfrac{1}{2}M(W) + \tfrac{1}{2}M(V) = \frac{1}{2100}\begin{pmatrix} 1360 & 710 & 650 \\ 710 & 610 & 100 \\ 650 & 100 & 550 \end{pmatrix},$$

and matrix convexity is violated because of $1410/2145 > 1360/2100$.

It then appears difficult to compare information matrices across varying block marginals, whence we shall follow Gaffke & Krafft (1980) and consider equi-blocksized designs, only.

3. Equi-blocksized designs for the interblock model

In the sequel we shall identify designs N and their weight matrices W. In the approximate theory the entries of W are not restricted to be rational numbers, i.e. we only have $w_{ij} \geq 0$, $\Sigma\Sigma w_{ij} = 1$. As we only consider equi-blocksized designs W, their information matrices for $\binom{\mu}{\alpha}$ are taken to be $M(W)$ from Section 2, thereby neglecting the common constant of proportionality $1/(\tau^2/b+\sigma^2/n)$.

Interest concentrates on the set of all treatment contrasts $(\alpha_1 - \overline{\alpha}.,\ldots,\alpha_v - \overline{\alpha}.)' = K_v\alpha$, with

$$K_v = I_v - 1_v 1_v'/v.$$

The following lemma characterizes identifiability (i.e. estimability, testability) of $K_v\alpha$, with a proof slightly different from that given by Gaffke & Krafft (1980, page 136).

Lemma 1. In the interblock model, for every equi-blocksized weight matrix W the following assertions are equivalent:

(a) $K_v\alpha$ is identifiable under W.

(b) rank $(bWW' - rr') = v-1$.

(c) rank $(M(W)) = v$.

Proof. Set $K = \binom{O}{K_v}$. (a) means that the range (column space) of K is contained in the range of $M(W)$, see Pukelsheim (1983, page 194). It is easy to see that this is equivalent to

$$\text{range } (K_v) \subseteq \text{range } (bWW' - rr'). \qquad (*)$$

But the nullspace of K_v, being spanned by 1_v, is always contained in the nullspace of $bWW' - rr'$. Thus equality holds in (*), and (a) is equivalent to (b), and also to (c), since rank is additive on the Schur complement, see Ouellette (1981, page 199).

Lemma 1(c) leaves a one-dimensional subspace of non-identifiable linear forms, a spanning vector of which obviously is $\binom{-1}{1_v}$.

The information matrix for a parameter set $K'\binom{\mu}{\alpha}$ is given by $J(M(W)) = (K'M(W)^-K)^+$. In the case of treatment contrasts K equals $\binom{O}{K_v}$, and an easy calculation yields

$$J(M(W)) = (K_v(bWW' - rr')^+K_v)^+ = bWW' - rr'.$$

Now M, as a function of W, is convex while J, as a function of M, is concave, see Marshall & Olkin (1979, page 468), Pukelsheim & Styan (1983, page 148). Hence no conclusion is possible for the composition $J \circ M$.

However, the explicit representation just obtained can be written as

$J(M(W)) = bWK_bW'$. Hence, after all, the information matrix $J \circ M$ for the treatment contrasts is a convex function of the design W. This distinct behaviour is due to the peculiar way by which in the interblock model the mean vector depends on the underlying design W.

Therefore, when the intention is to maximize information, we are led to maximize a convex functional of the design W, and the extreme designs become essential. The extreme points among all equi-blocksized designs were determined by Gaffke & Krafft (1980, page 141) to be what we shall call one-treatment-per-block designs T. By definition, such a design T has just one treatment i_j appearing in block j, for $j = 1,\dots,b$, and the constraint of uniform block marginals forces $t_{i_j j} = 1/b$. An alternative way of introducing such designs is as follows, cf. Gaffke & Krafft (1980, page 142).

Lemma 2. In the interblock model, every equi-blocksized design W with rank $(bWW' - rr') = v - 1$, where $r = W1_b$, satisfies

$$J(M(W)) = bWW' - rr' \leq \Delta_r - rr',$$

and here equality holds if and only if W is a one-treatment-per-block design.

Proof. We have $bWW' - rr' = \Delta_r - rr' - (\Delta_r - bWW')$. The matrix $\Delta_r - bWW'$ is nonnegative definite since it is an information matrix in a fixed effects model, see Pukelsheim (1983, page 201). Hence the inequality. Equality occurs if and only if $bWW' = \Delta_r$, and this is a characteristic property of one-treatment-per-block designs.

We shall call $\Delta_r - rr'$ a special C-matrix. Thus we observe the intriguing feature that the special C-matrices are uniformly optimal information matrices among all equi-blocksized designs with given treatment marginals in both, the interblock model and the fixed effects model, see Pukelsheim (1983, page 202), It follows from the fixed effects model theory that there is no uniformly optimal member in the class of all special C-matrices. However, we may then pose the question of admissibility, and to this we turn next.

4. Admissibility of special C-matrices

It is easy to verify that $\Delta_r - rr'$ has rank $v-1$ if and only if the stochastic vector r is positive, i.e. each component of r is positive. Giovagnoli & Wynn (1981, page 412) explicitly construct the eigenvalue set Λ^* of all special C-matrices, for $v = 3$ treatments. It follows from their results that $\Delta_t - tt' \leq \Delta_r - rr'$ need not imply $t = r$.

To see what can go wrong, in terms of the replication vectors, consider $t(\varepsilon) = (1-\varepsilon)e_i + \varepsilon 1_v/v$, where e_i is the i-th Euclidian unit vector. For small ε then $\Delta_{t(\varepsilon)} - t(\varepsilon)t(\varepsilon)'$ tends to vanish and thus falls below any matrix $\Delta_r - rr'$, whereas at the same time the treatment marginals $t(\varepsilon)$ converge to e_i and become extreme. In fact, this example qualitatively exhausts the possibilities, according to the following.

Theorem. Let t and r be two positive stochastic vectors in R^v. Then $\Delta_t - tt' \leq \Delta_r - rr'$ forces the alternative either (I) $t = r$, or (II) there exists some i such that $t_i > r_i$ and $t_j < r_j$, for all $j \neq i$.

Proof. Either $t_i \leq r_i$, for all i, and in view of $\Sigma t_i = \Sigma r_i$ then $t = r$. Or there exists some i such that $t_i > r_i$.
Now $\Delta_t - tt' \leq \Delta_r - rr'$ implies the converse ordering among the Moore-Penrose inverses $K_v \Delta_t^{-1} K_v$ and $K_v \Delta_r^{-1} K_v$, see Milliken & Akdeniz (1977, page 75). If $i = 1$, premultiplication by $(-1_{v-1} : I_{v-1})$ and postmultiplication by its transpose yield

$$\Delta^{-1}_{(t_2,\ldots,t_v)'} + t_1^{-1} J_{v-1} \geq \Delta^{-1}_{(r_2,\ldots,r_v)'} + r_1^{-1} J_{v-1}.$$

Its diagonal elements then satisfy $1/t_j + 1/t_1 \geq 1/r_j + 1/r_1$ and therefore $1/t_j > 1/r_j$, for all $j \geq 2$. A similar argument obviously applies when $i \neq 1$, and so the proof is complete.

Corollary 1. (a) If $\Delta_t - tt' \leq \Delta_r - rr'$ and $t_i = r_i$, for some i, then actually $t = r$.
(b) If $\Delta_t - tt' \leq \Delta_r - rr'$ and $t \neq r$ then there exists some i such that $t_i > 1/2$, $r_i \in [1-t_i ; t_i)$, and $t_j < 1/2$, $r_j \in (t_j ; 1-t_j]$, for all $j \neq i$.
(c) If $\Delta_t - tt' \leq \Delta_r - rr'$ and $t_i \leq 1/2$, for all i, then actually $t = r$.

Proof. (a) rules out alternative (II). In (b) alternative (II) yields $t_i > r_i$, while the assumption implies $t_i - t_i^2 \leq r_i - r_i^2$. On the parabolic arc $y(1-y)$ this is possible under the stated conditions, only. A similar argument applies when $t_j < r_j$. Finally (c) is a consequence of (b).

Uniform treatment marginals $t_i = 1/v$ lead to the matrix K_v/v which, according to Corollary 1(c), is always maximal among special C-matrices.

Returning to the interblock model we can now prove admissibility of one-treatment-per-block designs whenever their treatment marginals are not too "unbalanced".

Corollary 2. In the interblock model, every one-treatment-per-block design T with treatment replications $t_i \in (0;1/2]$, for all i, and every other equi-blocksized design W satisfy

$$J(M(T)) \leq J(M(W)) \Rightarrow J(M(T)) = J(M(W)).$$

Proof. Positivity of the treatment marginals yields $J(M(T)) = \Delta_t - tt'$. Hence rank $J(M(W)) \geq$ rank $(\Delta_t - tt') = v-1$, and $J(M(W)) = bWW' - rr'$, where $r = W1_b$. By Lemma 2 then

$$\Delta_t - tt' \leq bWW' - rr' \leq \Delta_r - rr'. \qquad (**)$$

Corollary 1(c) forces $t = r$, as well as equality in (**). Lemma 2 provides the additional information that W must be another one-treatment-per--block design.

Finally we remark that, as in the fixed effects models, the question of how many different one-treatment-per-block designs there are leads into the combinatorial part of the theory. Recall that permuting blocks does not change the information matrix. Hence we may assume each treatment appearing once among the first v blocks, in order to ensure positive treatment marginals. The remaining $b-v$ (indistinguishable) blocks then must be distributed between the v treatments, and this can be done in $\binom{v+(b-v)-1}{b-v} = \binom{b-1}{b-v}$ ways. A further reduction is possible when real--valued criteria $k(J)$ are considered which are permutationally invariant. In the special case of the trace criterion we can even observe a simple ordering of the values trace $(\Delta_r - rr')$. For since the latter is a

Schur-concave function in r, see Marshall & Olkin (1979, page 67), its values will grow as r becomes more uniform. This provides an alternative derivation of Lemma 2 in Gaffke & Krafft (1980, page 138).

References

[1] N.Gaffke and O.Krafft (1980). D-optimum designs for the interblock--model. Lecture Notes in Statistics 2 134-143.

[2] A.Giovagnoli and H.P.Wynn (1981). Optimum continuous block designs. Proc.Roy.Soc.London Ser.A 377 405-416.

[3] A.W.Marshall and I.Olkin (1979). Inequalities: Theory of Majorization and Its Applications. New York: Academic Press.

[4] G.A.Milliken and F.Akdeniz (1977). A theorem on the difference of the generalized inverses of two nonnegative matrices. Comm.Statist. A - Theory Methods 6 73-79.

[5] D.V.Ouellette (1981). Schur complements and statistics. Linear Algebra Appl. 36 187-295.

[6] F.Pukelsheim (1983). On optimality properties of simple block designs in the approximate design theory. J.Statist. Plann. Inference 8 193-208.

[7] F.Pukelsheim and G.P.H. Styan (1983). Convexity and monotonicity properties of dispersion matrices of estimators in linear models. Scand. J.Statist. 10 145-149.

Lecture Notes in Statistics

Edited by D. Brillinger, S. Fienberg, J. Gani, J. Hartigan, and K. Krickeberg

35

Linear Statistical Inference

Proceedings of the International Conference held at Poznań, Poland, June 4–8, 1984

N8<22196048

Edited by T. Caliński and W. Klonecki

Springer-Verlag
Berlin Heidelberg New York Tokyo