Proceedings of the
R.C. Bose Symposium on
Probability, Statistics and Design of Experiments, Delhi, 27-30 December, 1988,
Wiley Eastern, New Delhi, 1990, 607-618

INFORMATION MATRICES IN EXPERIMENTAL DESIGN THEORY

FRIEDRICH PUKELSHEIM
Institut für Mathematik, Universitat Augsburg

Information matrices for subsystems of the mean parameters in a classical linear model are studied, providing the basic quantities for the design of experiments. The rank of information matrices reflects identifiability of the parameter system of interest. Their functional properties follow from a representation as the minimum of a set of linear functions. The information matrix mapping is upper semicontinuous on the closed cone of nonnegative definite matrices, but fails to be continuous. Most of the development is carried out for parameter systems of full rank, but is seen to generalize to the rank deficient case. The results are illustrated by the classical C-matrices for simple block designs.

1. INTRODUCTION

In the theory of block designs a central notion is that of *C-matrices*. Just where this acronym originates from is unclear. Reference to a C-matrix is made implicitly by Bose (1948), page (12), and explicitly by Chakrabarti (1963). In any case the name is appealing since the C-matrix is the

— coefficient matrix of the reduced system of normal equations for the symmetrized treatment contrasts, as well as the

— contrast information matrix,

assuming a two-way classification model with no interaction for a treatment factor and a blocking factor.

The concept of *information matrices* pertains to more general design problems rather than just block designs. Section 2 reviews the role of information matrices when the task is to estimate a parameter subsystem $K'\theta$, to test a hypothesis $K'\theta = 0$, or to find the Fisher information matrix for $K'\theta$. Here θ is a $k \times 1$ unknown parameter vector for the mean value,

Keywords and phrases: C-matrix, contrasts, singular moment matrices, Loewner ordering, Gauss-Markov theorem, identifiability, method of regularization, generalized information matrix.

AMS 1980 Subject classifications: 62K05, 62K15.

and K is a known $k \times s$ coefficient matrix of full column rank s. For the $s \times 1$ parameter subsystem $K'\theta$ a design with positive definite $k \times k$ moment matrix M has information matrix $C_K(M)$, given by the positive definite $s \times s$ matrix

$$C_K(M) = (K'M^{-1}K)^{-1}. \tag{1}$$

In this paper we study the dependence of C_K on its argument matrix M.

In design problems we need to evaluate $C_K(M)$ for varying designs, that is, for varying moment matrices M. Since an $s \times 1$ subsystem $K'\theta$ is of interest where possible $s < k$, designs with a singular moment matrix may well be optimal for $K'\theta$, and cannot be neglected. Hence, it is imperative to also cover *singular moment matrices*.

The *extension* of formula (1) from positive definite matrices M to all nonnegative definite matrices A needs careful attention. Recent work of Gaffke (1987) suggests that the following definition is appropriate:

$$C_K(A) = \min_{L \in \mathbb{R}^{s \times k}:\, LK = I_s} LAL', \tag{2}$$

whenever A is a nonnegative definite $k \times k$ matrix. The minimum in (2) is taken relative to the *Loewner ordering*, the usual ordering of symmetric matrices, given by

$$C \geqslant D \Leftrightarrow C - D \text{ is nonnegative definite.}$$

It is consistent with this notation that from now on we simply write $A \geqslant 0$ in order to indicate that A is nonnegative definite.

In Section 2 we outline that the minimum in (2) exists and is attained, and conforms with formula (1) for positive definite matrices A. In fact, this follows from a straightforward application of the *Gauß–Markov Theorem*, as anticipated by Pukelsheim and Styan (1983). This theorem also entails explicit representations for (2). Let L be an arbitrary *left inverse* of K, $LK = I_s$, and define the *residual projector*

$$R = I_k - KL.$$

The first representation is akin to the covariance adjustment formula of Rao (1967),

$$C_K(A) = LAL' - LAR'(RAR')^{-}RAL'. \tag{3}$$

There always exists at least one left inverse L_A that satisfies $L_A A R' = 0$, and then (3) simplifies to

$$C_K(A) = L_A A L'_A. \tag{4}$$

For varying matrices A the complexity of computing $C_K(A)$ hence is that of solving the linear equation $L_A(K, AR') = (I_s, 0)$ for L_A.

For the special case of the first s out of k parameters, $K' = (I_s\ 0)$, representation (3) turns into a Schur complement,

$$C_{\binom{I_s}{0}}(A) = A_{11} - A_{12}A_{22}^{-}A_{21}. \tag{5}$$

The history of Schur complement matrices and their use in statistics is reviewed by Ouellette (1981), and Styan (1987).

In Section 3 we concentrate on the *rank behaviour* of information matrices. It transpires that the information matrix $C_K(M)$ for $K'\theta$ is positive definite if and only if the parameter subsystem $K'\theta$ is identifiable under a design with moment matrix M.

Section 4 investigates the *continuity behaviour* of the information matrix mapping C_K. The central result is that C_K is matrix upper semicontinuous. As a consequence, formula (1) extends beyond positive definite matrices to all nonnegative definite matrices $A \geqslant 0$, according to

$$C_K(A) = \lim_{n\to\infty} (K'(A + \frac{1}{n} I_k)^{-1}K)^{-1}. \tag{6}$$

This result is due to Gaffke and Pukelsheim (1987), page 40. We adapt an example from Pázman (1986), page 67, to demonstrate that C_K need *not* be continuous.

In Section 5 we sketch the generalization to *degenerate* coefficient matrices K, that is, matrices K that do not have full column rank s. This leads to *generalized information matrices* which coincide with the C-matrices of the block designs from the beginning. We take this coincidence as a promising evidence for the present approach. The monograph Pukelsheim (1992) will present a comprehensive development along these lines.

2. INFORMATION MATRICES

We assume a linear model of uncorrelated homoscedastic observations. That is, the $n \times 1$ vector Y of observations is taken to have mean vector and variance-covariance matrix

$$E_{\theta;\sigma^2}[Y] = X\theta, \qquad V_{\theta;\sigma^2}[Y] = \sigma^2 I_n.$$

Here θ is a $k \times 1$ vector of unknown parameters for the mean, $\sigma^2 > 0$ is the unknown model variance, and X is the known $n \times k$ model matrix. The i^{th} row of X represents the regression vector for the i^{th} observation Y_i. The number of observations, n, and the rows of X are determined by the experimental design. The *moment matrix* M of the design then is given by

$$M = \frac{1}{n} X'X.$$

This is a nonnegative definite $k \times k$ matrix. We take M to be positive definite, for the time being; in other words, we assume the model matrix X to have full column rank k.

Let K be a known $k \times s$ coefficient matrix of full column rank s. Our interest concentrates on the $s \times 1$ parameter subsystem $K'\theta$. When the full parameter vector is of interest, $K = I_k$, then (1) plainly yields $C_{I_k}(M) = M$. This is to say that the moment matrix of the underlying design is the same as the information matrix for the full parameter vector θ. For proper subsystems K_θ the two notions differ, and the information matrices (1) start playing a role of their own. They are motivated through variance–covariance matrices of Gauß–Markov estimators, through the power of the F-test, and as Fisher information matrices, as follows.

The Gauß–Markov Theorem provides the best linear unbiased estimator for $K'\theta$. The variance–covariance matrix of this estimator is

$$\frac{\sigma^2}{n} K' M^{-1} K.$$

The smaller this matrix is in the Loewner ordering, the larger is its inverse, $(n/\sigma^2) C_K(M)$. This elucidates the role of the information matrix for $K'\theta$, when the problem is one of estimating $K'\theta$.

For the related problem of testing the hypothesis $K'\theta = 0$, we further assume that the vector Y of observations follows a normal distribution. The F-test then is based on the test statistic F which, when $n > k + 2$, has expectation

$$E_{\theta;\sigma^2}(F) = \frac{n-k}{n-2-k}\left(1 + \frac{n/\sigma^2}{s}\theta' K C_K(M) K'\theta\right). \tag{7}$$

Large values of F are significant for a deviation from the hypothesis $K'\theta = 0$. Thus, the larger the information matrix $C_K(M)$ for $K'\theta$, the larger values for F we expect, and the clearer the F-test detects a significant deviation. More precisely, the F-test then becomes uniformly more powerful. This underlines the importance of the information matrix for $K'\theta$ for the problem of testing $K'\theta = 0$.

Let us turn to general parametric modelling. It is well known that the Fisher information matrix jointly for (θ, σ^2) is

$$\frac{n}{\sigma^2}\begin{pmatrix} M & 0 \\ 0 & \dfrac{1}{2\sigma^2} \end{pmatrix}$$

It follows that the Fisher information matrix for θ alone is $(n/\sigma^2)M$. Alternatively we may apply formula (1) from which, upon setting $K = I_k$, we obtain

$$C_{I_k}(M) = M.$$

Thus the Fisher information matrix for θ coincides with the information matrix for θ as defined by (1), except for the proportionality constant n/σ^2. The same is true for subsystems $K'\theta$, as follows from the differential geometric discussion in Barndorff-Nielsen and Jupp (1988).

All this shows that the information matrix (1) for $K'\theta$ is firmly rooted in statistical inference, provided the moment matrix M of the design is positive definite.

In order to also cover singular argument matrices A of C_K, let us view formula (1) through the Gauß–Markov Theorem. To this end consider the auxiliary linear model

$$E_{\eta;\tau^2}[Z] = K\eta, \qquad V_{\eta;\tau^2}[Z] = \tau^2 M.$$

An arbitrary linear estimate LZ for η is unbiased if and only if $LK = I_s$, that is, L is a left inverse of K. The variance-covariance matrix of LZ is $\tau^2 LML'$. The Gauß–Markov Theorem states that the weighted least squares estimator solves the minimization problem of finding the unbiased estimator with smallest variance-covariance matrix,

$$(K'M^{-1}K)^{-1} = \min_{L \in \mathbb{R}^{s \times k}: LK = I_s} LML'. \tag{8}$$

The minimum is understood relative to the Loewner ordering, see for instance Witting (1985), page 303. The matrix of the left hand side of (8) is the same as in (1). However, the minimum on the right hand side of (8) exists and is attained also when M is singular, and is given by (3) or (4).

Hence, it makes sense to *define* $C_K(A)$ for nonnegative definite and possibly singular matrices A through (2). Consistency with (1) then follows from (8). When the first s out of k parameters are of interest, $K' = (I_s, 0)$, we may choose $L = K'$ in (3) in order to derive the Schur complement formula (5) Thus the functional properties of the information matrix mapping $A \rightarrow C_K(A)$ embraces as a special case the functional properties the Schur complement mapping $A \rightarrow A_{11} - A_{12}A_{22}^{-}A_{21}$.

The rank behaviour of $C_K(A)$ is related to matrix algebra and is presented in Section 3. The continuity behaviour is a topic of calculus and is discussed in Section 4.

3. RANK BEHAVIOUR

The rank of a matrix is the dimension of its range (column space). Therefore, we begin with a lemma on ranges before turning to rank.

LEMMA 1. Let the $k \times s$ coefficient matrix K have full column rank s, and let A be a nonnegative definite $k \times k$ matrix. Then the matrix $A_K = KC_K(A)K'$ is the unique matrix with the three properties

$$0 \leqslant A_K \leqslant A, \quad \text{range } A_K \subseteq \text{range } K, \quad \text{range } (A - A_K) \cap (\text{range } K) = \{0\}.$$

PROOF. The proof follows from the works of Anderson (1971), Anderson and Trapp (1975), and Mitra and Puri (1979). □

This characterization of A_K refers to the matrix K only through its range. Hence two coefficient matrices K and $\tilde{K}$ with the same range induce identical matrices $A_K = A_{\tilde{K}}$ even though $C_K(A)$ and $C_{\tilde{K}}(A)$ are, in general, distinct.

THEOREM 2. Let the $k \times s$ coefficient matrix K have full column rank s, and let A be a nonnegative definite $k \times k$ matrix. Then

$$\text{rank } C_K(A) = \dim ((\text{range } A) \cap (\text{range } K)).$$

In particular, $C_K(A)$ is positive definite if and only if the range of A includes the range of K.

PROOF. Define $A_K = KC_K(A)K'$. We know that $A = (A - A_K) + A_K$ is the sum of two nonnegative definite matrices, by Lemma 1. Hence the range of A is the algebraic sum of the ranges of $A - A_K$ and A_K, whence

$$\begin{aligned}(\text{range } A) \cap (\text{range } K) &= (\text{range } (A - A_K) \cap (\text{range } K)) \\ &+ ((\text{range } A_K) \cap (\text{range } K)) = \text{range } A_K.\end{aligned}$$

Since K has full column rank this yields

$$\text{rank } C_K(A) = \text{rank } A_K = \dim ((\text{range } A) \cap (\text{range } K)).$$

In particular, $C_K(A)$ has rank s if and only if

$$(\text{range } A) \cap (\text{range } K) = \text{range } K,$$

that is, range $A \supseteq$ range K. □

The condition that the range of A includes the range of K appears in various disguises, as identifiability condition, estimability condition. or testability condition. Theorem 2 says that the *rank* of the information matrix $C_K(M)$ for $K'\theta$ reflects the extent of identifiability (estimability, testability) of $K'\theta$ under a design with moment matrix M. The theorem suggests the following *check for identifiability*. First compute $C_K(M)$ from formula (3) or (4), then find its rank. If the rank equals s then identifiability holds, otherwise it holds not.

4. CONTINUITY BEHAVIOUR

We first list a couple of properties that are easy consequences of the defining relationship (2). Let NND(k) denote the closed convex cone of nonnegative definite $k \times k$ matrices, and let Sym(s) be the linear space of symmetric $s \times s$ matrices.

THEOREM 3. Let the $k \times s$ coefficient matrix K have full column rank s. Then the information matrix mapping

$$A \to C_K(A) = \min_{L \in \mathbb{R}^{s \times k}:\, LK = I_s} LAL'$$

from NND(k) to Sym(s) is nonnegative definite, matrix isotonic, positively homogeneous, matrix superadditive, and matrix concave:

$$\begin{aligned}
C_K(A) &\geqslant 0 && \text{for all } A \geqslant 0,\\
A \geqslant B &\Rightarrow C_K(A) \geqslant C_K(B) && \text{for all } A,\, B \geqslant 0,\\
C_K(\delta A) &= \delta C_K(A) && \text{for all } A \geqslant 0, \delta > 0,\\
C_K(A+B) &\geqslant C_K(A) + C_K(B) && \text{for all } A, B \geqslant 0\\
C_K((1-\alpha)A + \alpha B) &\geqslant (1-\alpha)C_K(A)\\
&\quad + \alpha C_K(B) && \text{for all } A,\, B \geqslant 0,\ \alpha \in (0,1).
\end{aligned}$$

PROOF. The first four properties are immediate from the definition of $C_K(A)$ as the minimum over the matrices LAL'. The last property follows since superadditivity and homogeneity imply concavity. □

Next we wish to show that the information matrix mapping is matrix upper semicontinuous. The key fact is that the functions $A \to LAL'$ are linear, whence C_K is the minimum of a family of linear functions.

THEOREM 4. Let the $k \times s$ coefficient matrix K have full colmun rank s. Then the information matrix mapping C_K is matrix upper semicontinuous, that is, for all sequences $(A_n)_{n\geqslant 1}$ in NND(k) that converge to a limit A we have

$$C_K(A_n) \geqslant C_K(A) \quad \text{for all } n \geqslant 1 \quad \Rightarrow \quad \lim_{n\to\infty} C_K(A_n) = C_K(A).$$

PROOF. Suppose the matrices $A_n \geqslant 0$ converge to A such that

$$C_K(A_n) \geqslant C_K(A).$$

With a left inverse L_A of K that satisfies (4) we obtain

$$\begin{aligned}
C_K(A) \leqslant C_K(A_n) &= \min_{L\in\mathbb{R}^{s\times k}:\, LK = I_s} LA_nL'\\
&\leqslant L_A A_n L'_A \to L_A A L'_A = C_K(A).
\end{aligned}$$

Hence the matrices $C_K(A_n)$ converge to the limit $C_K(A)$. □

A prime application consists in extending formula (1) from the open cone of positive definite matrices PD(s) to its closure, the closed cone of nonnegative definite matrices NND(s). The method is called *regularization*.

COROLLARY 5. Let the $k \times s$ coefficient matrix K have full column rank s and let B be a positive definite $k \times k$ matrix. Then

$$C_K(A) = \lim_{n\to\infty} \left(K'\left(A + \frac{1}{n}B\right)^{-1}K\right)^{-1} \text{ for all } A \geqslant 0. \tag{9}$$

PROOF. The matrices $A_n = A + \frac{1}{n}B$ evidently converge to A. They fulfill $A_n \geqslant A$, whence monotonicity of C_K entails $C_K(A_n) \geqslant C_K(A)$. Thus the matrices $C_K(A_n)$ converge to $C_K(A)$, by Theorem 4. Since A_n is positive definite, formula (1) applies and yields $C_K(A_n) = (K'A_n^{-1}K)^{-1}$.

As a particular case Corollary 5 covers formula (6) in Section 1. A general discussion of regularization methods in statistics is given by Cox (1988).

The point is that the matrices A_n converge along the ray $\{A + \delta B : \delta \geqslant 0\}$, emanating from A along the direction of B into the interior PD(k). In other words, regularization ascertains that the representation $(K' A^{-1} K)^{-1}$ permits a continuous extension from the open cone PD(s) to the closed cone NND(k), as long as the argument matrices "converge along straight lines" of positive definite matrices.

We demonstrate by example that C_K is not, in general, continuous on NND(k). Consider a model for a straight-line fit,

$$Y_{ij} = \alpha + \beta t_i + E_{ij},$$

where the experimental conditions, t_i may be chosen from the domain $[-1, 1]$. The intercept α and the slope β are unknown, and E_{ij} are uncorrelated random errors with mean 0 and variance σ^2. We study the *unsymmetric two-point designs* $\tau_{s,n}$ that assign half of the observations to the experimental conditions $1/n$ and $-s/n$, respectively, where $s \neq 1$ is an additional *support parameter*. The moment matrices $A_{s,n}$ of the designs $\tau_{s,n}$ converge to a limit A that does not depend on s,

$$A_{s,n} = \begin{pmatrix} 1 & \dfrac{1-s}{2n} \\ \dfrac{1-s}{2n} & \dfrac{1+s^2}{2n^2} \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = A. \tag{10}$$

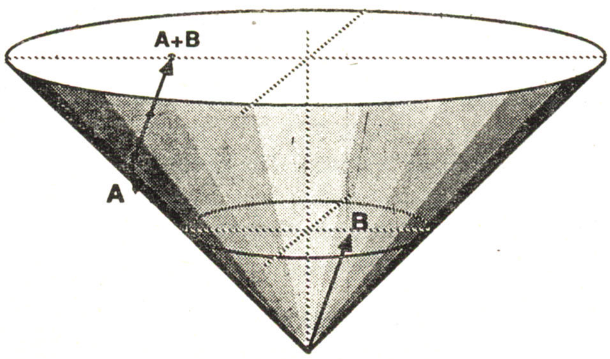

Fig. 1. Regularization of the information matrix mapping on the closed cone NND(k). Convergence to the boundary is continuous as long as it takes place along straight lines from within the open cone PD(k), as given by (9).

In fact, A is the moment matrix of the *symmetric* two-point designs $\tau_{1,n}$.

Let the intercept α be the parameter of interest. For the two-point designs $\tau_{s,n}$, the Schur complement formula (5) yields the information for α,

$$C_{\binom{1}{0}}(A_{s,n}) = 1 - \left(\frac{1-s}{2n}\right)^2 \frac{2n^2}{1+s^2} = \frac{1}{2}\frac{(1+s)^2}{1+s^2} = \phi(s), \tag{11}$$

say. The information $\phi(s)$ is constant in n, and hence equal to its limit as n tends to infinity. Therefore, with varying support parameter s, the designs $\tau_{s,n}$ exhaust all possible information values, from the minimum $\phi(-1) = 0$ to the maximum $\phi(1) = 1$, even though for n tending to infinity their moment matrices $A_{s,n}$ converge to the common limit A. See Fig. 2.

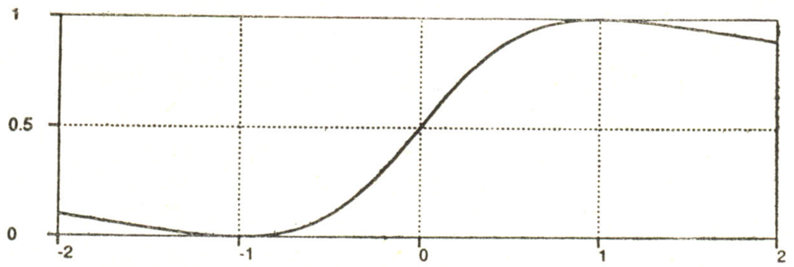

Fig. 2. Discontinuity of the information matrix mapping in a line fit model: The intercept information (11) of the unsymmetric tow-point designs $\tau_{s,n}$ attains all possible values between the minimum zero and the maximum one for varying support parameter s, even though according to (10) the moment matrices converge to a common limit not depending on s.

In summary, the unsymmetric two-point designs $\tau_{s,n}$ provide an instance of discontinuity, in that if $s \neq 1$ then

$$\lim_{n\to\infty} A_{s,n} = A$$

$$\lim_{n\to\infty} C_{\binom{1}{0}}(A_{s,n}) = \phi(s) < 1 = \phi(1) = C_{\binom{1}{0}}(A).$$

5. GENERALIZED INFORMATION MATRICES

Not all parameter systems $K'\theta$ that are of statistical interest have a coefficient matrix K that is of full column rank. It is neither helpful nor wise to remedy rank deficiency through a full rank reparametrization. In most applications the parameters have a definite meaning, and this meaning is destroyed or at least distorted by reparametrization. Instead the notion of information matrices is generalized, so as to complement the framework drawn by coefficient matrices which do have full column rank.

In view of the preceding results we define the *generalizd information* matrix for a parameter subsystem $K'\theta$, where the $K \times s$ coefficient matrix K possibly is rank deficient, to be the $k \times k$ matrix

$$AK= \min_{Q\in\mathbf{R}^{k\times k}:\ QK = K} QAQ'. \tag{12}$$

When K is of full column rank s then we have $A_K = KC_K(A)K'$; in this form the matrix A_K has made its first appearance in (7) while discussing the testing problem. It is not hard to show that it also conforms with the estimation problem. Furthermore, the mapping $A \longmapsto A_K$ enjoys the same functional properties as the mapping $A \longmapsto C_K(A)$, owing to the similarity of the two definitions (2) and (12). From the Gauβ-Markov Theorem we obtain an analogue to (3),

$$A_K = A - AR'(RAR')^- RA, \tag{13}$$

with an arbitrary generalized inverse G of K and with residual projector $R = I_k - KG$.

The major disadvantage of the definition is that the row-column number of a generalized information matrix A_K no longer exhibits the reduced dimensionality of the subsystem $K'\theta$: Both matrices are $k \times k$, and the notation A_K is meant to indicate this. It is rank, not row-column number, that provides a measure for the extent of identifiability of the subsystem $K'\theta$.

As an example we compute the generalized information matrix for the symmetrized treatment contrasts of a simple block design, in the two-way classification model with no interaction

$$Y_{ijk} = \alpha_i + \beta_j + E_{ijk}.$$

Here α_i and β_j are the fixed effects of treatments $i = 1, \ldots, \alpha$ and blocks $j = 1, \ldots, b$, respectively, while E_{ijk} are uncorrelated random errors of mean 0 and variance σ^2. Let $1_a = (1, \ldots, 1')$ be the $a \times 1$ unity vector, and define the orthodiagonal projector K_a by

$$K_a = I_a - \frac{1}{a} 1_a 1_a' .$$

Upon introducing the parameter vector $\theta = (\alpha', \beta')$ with $\alpha = (\alpha_1, \ldots, \alpha_a)'$ and $\beta = (\beta_1, \ldots, \beta_b)'$, the symmetrized treatment contrasts are the subsystem

$$\begin{pmatrix} \alpha_1 - \bar{\alpha}. \\ \alpha_a - \bar{\alpha}. \end{pmatrix} = K_a\alpha = (K_a, 0)\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = K'\theta,\ K = \begin{pmatrix} K_a \\ 0 \end{pmatrix}.$$

An experimental design for this model is given by an $a \times b$ weight matrix W whose entries w_{ij} determine the proportion of observations with treatment i in block j. A design for sample size n has weights w_{ij} which are multiples of $1/n$, $w_{ij} = n_{ij}/n$. Let $\triangle_r$ and $\triangle_s$ be the diagonal matrices formed from the row-sum vector r of W and the column-sum vector s of W, respectively. It is well known that a design with weight matrix W has moment matrix

$$M = \begin{pmatrix} \triangle_r & W \\ W' & \triangle_s \end{pmatrix}.$$

In order to compute the generalized information matrix M_K we choose

$$G = K',\ R = I_k - KK',\ (RMR)^- = \begin{pmatrix} 0 & 0 \\ 0 & \triangle_s^- \end{pmatrix}.$$

Straightforward calculation then turns (13) into

$$M_K = \begin{pmatrix} \triangle_r - W\triangle_s^- W' & 0 \\ 0 & 0 \end{pmatrix}.$$

Hence, except for vanishing subblocks, the generalized information matrix M_K for the symmetrized treatment contrasts coincides with the Schur complement of $\triangle_s$ in W,

$$C = \triangle_r - W\triangle_s^- W'.$$

This is the usual C-matrix, of the simple block design W. Thus our notion of generalized information matrices is in good agreement with the classical notion of C-matrices.

REFERENCES

W.N. Anderson, Jr (1971). Shorted operators, *SIAM Journal on Applied Mathematics*. **20**, 520-525.

W.N. Anderson, Jr and G.E. Trapp (1975). Shorted operators II, *SIAM Journal on Applied Mathematics*, **28**, 60-71.

O.E. Barndorff-Nielsen and P.E. Jupp (1988). Differential geometry, profile likelihood, L-sufficiency and composite transformation models, *The Annals of Statistics*, **16**, 1009-1043.

R.C. Bose (1948). The design of experiments, in: *Proceedings of the Thirty-Fourth Indian Science Congress, Delhi, 1947*, Indian Science Congress Association, Calcutta, (1)-(25).

M.C. Chakrabarti (1963). On the C-matrix in design of experiments, *Journal of the Indian Statistical Association*, **1**, 8-23

D.D. Cox (1988). Approximation method of regularization estimators, *The Annals of Statistics*, **16**, 694-712.

N. Gaffke (1987). Further characterizations of design optimality and admissibility for partial parameter estimation in linear regression, *The Annals of Statistics*, **15**, 942–957,

N. Gaffke and F. Pukelsheim (1987). Admissibility and optimality of experimental designs, in: *Model-Oriented Data Analysis, Proceedings, Eisenach, GDR* (V. Fedorov and H. Läuter, eds.) Springer Lecture Notes in Economics and Mathematical Systems, **297**, 37-43,

S.K. Mitra and M.L. Puri (1979). Shorted operators and generalized inverses of matrices, *Linear Algebra and Its Applications*, **25**, 45-56.

D.V. Ouellette (1981). Schur complements and statistics, *Linear Algebra and Its Applications*, **36**, 187–295.

A. Pázman (1986). *Foundations of Optimum Experimental Design*, Reidel, Dordrecht.

F. Pukelsheim (1992). *Optimality Theory for Experimental Designs in Linear Models*, Wiley, New York, forthcoming.

F. Pukelsheim and G.P.H. Styan (1983). Convexity and monotonicity properties of dispersion matrices of estimators in linear models, *Scandinavian Journal of Statistics*, **10**, 145-149.

C.R. Rao (1967). Least squares theory using an estimated dispersion matrix and its applications to measurement of signals, in: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (L. Le Cam, ed.), University of California, Berkeley CA, 1, 355-372.

G.P.H. Styan (1987). Schur complements and linear statistical models, in: *Proceedings of the First International Tampere Seminar on Linear Statistical Models and Their Applications* (T. Pukkila and S. Puntanen, eds.), University of Tampere, Tampere, 37-75.

H. Witting (1985). *Mathematische Statistik I, Parametrische Verfahren bei festem Stichprobenumfang*, Teubner, Stuttgart.