A Note on the Matrix Ordering of Special C-Matrices*

Jerzy K. Baksalary† and Friedrich Pukelsheim
Department of Mathematical and Statistical Methods
Academy of Agriculture
PL-60-637 Poznań, Poland
and
Institut für Mathematik
Universität Augsburg
D-8900 Augsburg, Federal Republic of Germany

Submitted by Ingram Olkin

ABSTRACT

The matrix inequality $(*)$ $\Delta_t - tt' \leqslant \Delta_r - rr'$ is considered, where t and r are positive stochastic vectors, and Δ_t and Δ_r are diagonal matrices with t and r on their diagonals. Necessary and sufficient conditions are established (1) for $(*)$ to hold when t and r are given, and (2) for the existence of some vector r satisfying $(*)$ when t is given. The results have applications in various parts of statistics.

1. INTRODUCTION AND RESULTS

Several problems in statistics lead to the consideration of inequalities of the form $\Delta_t - tt' \leqslant \Delta_r - rr'$, where $t = (t_1, \ldots, t_n)'$ is a positive stochastic vector in $\mathbf{R}^n$ (i.e., $t_i > 0$ and $\sum t_i = 1$), and r is a positive stochastic vector in $\mathbf{R}^n$, and Δ_t and Δ_r are diagonal matrices with t and r on their diagonals. The ordering $\leqslant$ denotes the Loewner matrix ordering; see [4, Chapter 16.E].

The matrix $\Delta_t - tt'$ appears as a special case of a C-matrix in experimental design theory, i.e. as the information matrix for the treatment contrasts of an experimental design, with treatment replication vector t, in both the two-

*The paper was essentially started at the International Statistical Conference on Linear Inference, 4–8 June 1984, Poznań, Poland. We gratefully acknowledge the opportunity to have worked together at this Conference.

†Research partly supported by the Polish Academy of Sciences Grant No. MR I.1-2/1.

 0024-3795/85/$3.30

way-classification, fixed-effects model and the interblock model associated with the two-way-classification, mixed-effects model. See [6, Theorem 4a] and [2, Lemma 2] for more details. Also, $\Delta_t - tt'$ is the dispersion matrix of a multinomial distribution with cell probability vector t.

In any case it is of interest to compare two matrices of the special form given above through the Loewner matrix ordering, and to this end we shall establish here the following results.

THEOREM 1. *Suppose t and r are positive stochastic vectors of dimension n such that $t \neq r$. Then*

$$\Delta_t - tt' \leqslant \Delta_r - rr' \qquad (*)$$

if and only if there exists some subscript i such that

$$t_i > r_i, \qquad \text{(a)}$$

$$t_j < r_j \qquad \text{for all} \quad j \neq i, \qquad \text{(b)}$$

$$\frac{r_i t_i}{t_i - r_i} \geqslant \sum_{j \neq i} \frac{r_j t_j}{r_j - t_j}. \qquad \text{(c)}$$

THEOREM 2. *Suppose t is a positive stochastic vector of dimension n. Then there exists some positive stochastic vector $r \neq t$ such that the inequality $(*)$ holds if and only if there exists some subscript i such that $t_i > \frac{1}{2}$.*

Theorem 1 says that the inequality $(*)$ is equivalent to the components of t being strictly smaller than those of r, except for one where the inequality $t_i > r_i$ goes the wrong way round as compared to $(*)$, and that altogether the distances between the components of t and r must be so as to satisfy the quantitative property (c). Theorem 2 is somewhat surprising in that the condition $t_i > \frac{1}{2}$ does not depend on the dimensionality n.

For the statistical applications Theorem 1 provides an easy means to compare the information matrices between the corresponding block designs and to compare the dispersion matrices of two multinomial distributions, by looking solely at the components of the stochastic vectors t and r. Theorem 2 offers a very simple criterion for when a design with C-matrix $\Delta_t - tt'$ is admissible, i.e. when $\Delta_t - tt'$ is maximal among all C-matrices, and when a multinomial distribution is maximally dispersed.

The necessity of the properties (a) and (b) in Theorem 1, and the direct part of Theorem 2, were established by Christof and Pukelsheim [2]. In the present note we adjoin the property (c) and prove sufficiency. We shall give a brief self-contained exposition of proofs in Section 2.

2. PROOFS

The following lemma is due to Farebrother [3, Appendix] and has been generalized by Baksalary and Kala [1, Theorem 1]. It also follows from Haynsworth's inertia formula; cf. (1.28) in [7].

LEMMA. *Suppose D is a positive definite $n \times n$ matrix, b is a nonzero vector in $\mathbf{R}^n$, and α is a positive scalar. Then*

$$D \geqslant \alpha bb' \quad \Leftrightarrow \quad 1/\alpha \geqslant b'D^{-1}b.$$

Proof. For the first part, premultiplying with $b'D^{-1}$ and postmultiplying with its transpose yields $b'D^{-1}b \geqslant \alpha(b'D^{-1}b)^2$, i.e. $1/\alpha \geqslant b'D^{-1}b$. For the converse part, the Cauchy-Schwarz inequality leads to

$$\alpha(b'x)^2 = \alpha(b'D^{-1/2}D^{1/2}x)^2 \leqslant \alpha(b'D^{-1}b)(x'Dx)$$

for every vector x in $\mathbf{R}^n$. Hence $1/\alpha \geqslant b'D^{-1}b$ implies $D \geqslant \alpha bb'$. ∎

Proof of Theorem 1. Since $t \neq r$ and $\Sigma t_i = 1 = \Sigma r_i$, there must exist some subscript i such that $t_i > r_i$. Without loss of generality we may take $i = 1$, and thus assume $t_1 > r_1$. Define the matrix $K_n = I_n - 1_n 1_n'/n$, where 1_n is the n-dimensional vector with all elements unity.

As all components of t are assumed to be positive, we have

$$(\Delta_t - tt')K_n\Delta_t^{-1}K_n = K_n.$$

Hence $K_n\Delta_t^{-1}K_n$ is seen to be the Moore-Penrose inverse of $\Delta_t - tt'$, and $\operatorname{rank}(\Delta_t - tt') = n - 1$. It now follows from Theorem 3.1 in [5] that the inequality (*) is equivalent to the converse ordering

$$K_n\Delta_t^{-1}K_n \geqslant K_n\Delta_r^{-1}K_n$$

among the Moore-Penrose inverses. Premultiplying with $(-1_{n-1}:I_{n-1})$ and postmultiplying with its transpose leads to another equivalent form of (*):

$$D \geqslant \alpha_1 1_{n-1} 1'_{n-1}, \tag{†}$$

where D is the $(n-1)\times(n-1)$ diagonal matrix with $\alpha_j = 1/t_j - 1/r_j$, for $j \geqslant 2$, on its diagonal, and $\alpha_1 = 1/r_1 - 1/t_1$. By assumption $\alpha_1 > 0$, and this forces D to be positive definite. Thus (†) entails (b); and the Lemma implies

$$\frac{1}{\alpha_1} \geqslant \sum_{j>1} \frac{1}{\alpha_j},$$

i.e. (c). Conversely (a), (b), (c) and the Lemma establish (†). ■

Proof of Theorem 2. The inequalities $t_i - t_i^2 \leqslant r_i - r_i^2$, obtained from (*), and $t_i > r_i$, given in (a), can hold simultaneously only if $t_i > \frac{1}{2}$ and $r_i \in [1 - t_i; t_i)$. This establishes the direct part. For the converse part, choose some $r_i \in [1 - t_i; t_i)$, and for $j \neq i$ define

$$r_j = \frac{1 - r_i}{1 - t_i} t_j = t_j + \frac{t_i - r_i}{1 - t_i} t_j.$$

Then $r = (r_1, \ldots, r_n)'$ is a positive stochastic vector satisfying (a) and (b). Because of

$$\sum_{j \neq i} \frac{r_j t_j}{r_j - t_j} = \frac{1 - r_i}{t_i - r_i}(1 - t_i) \leqslant \frac{r_i t_i}{t_i - r_i},$$

it also fulfills (c). The inequality (*) now follows from Theorem 1. ■

REFERENCES

1 J. K. Baksalary and R. Kala, Partial orderings between matrices one of which is of rank one, *Bull. Acad. Polon. Sci. Sér. Sci. Math.* 31:5–7 (1983).

2 K. Christof and F. Pukelsheim, Approximate design theory for a simple block design with random block effects, Tech. Report No. 34, Inst. für Mathematik, Univ. Augsburg, 1984.

3 R. W. Farebrother, Further results on the mean square error of ridge regression, *J. Roy. Statist. Soc. Ser. B* 38:248–250 (1976).

4 A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic, New York, 1979.

5 G. A. Milliken and F. Akdeniz, A theorem on the difference of the generalized inverses of two nonnegative matrices, *Comm. Statist. A — Theory Methods* 6:73–79 (1977).

6 F. Pukelsheim, On optimality properties of simple block designs in the approximate design theory, *J. Statist. Plann. Inference* 8:193–208 (1983).

7 G. P. H. Styan, 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.). Dept. of Math. Sci., Univ. Tampere, 1985, pp. 37–75.

Received 12 November 1984; revised 15 February 1985