Convexity and Monotonicity Properties of Dispersion Matrices of Estimators in Linear Models

Author(s): Friedrich Pukelsheim and George P. H. Styan
Source: *Scandinavian Journal of Statistics*, Vol. 10, No. 2 (1983), pp. 145-149
Published by: Wiley on behalf of Board of the Foundation of the Scandinavian Journal of Statistics
Stable URL: http://www.jstor.org/stable/4615914
Accessed: 31-08-2017 12:38 UTC

Scand J Statist 10: 145–149, 1983

Convexity and Monotonicity Properties of Dispersion Matrices of Estimators in Linear Models

FRIEDRICH PUKELSHEIM and GEORGE P. H. STYAN

Universität Freiburg im Breisgau and McGill University Montreal

ABSTRACT. The function $f(\mathbf{A})=(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+$ is shown to be concave and isotone and the function $g(\mathbf{A})=\mathbf{K}'\mathbf{A}^-\mathbf{K}$ is shown to be convex and antitone, when $\mathbf{A}$ varies over the set of symmetric non-negative definite matrices whose range contains the range of the matrix $\mathbf{K}$. The results are motivated and interpreted through best linear unbiased estimates and optimal experimental designs in linear model theory.

Key words: Gauss-Markov Theorem, matrix convexity, monotone matrix functions, matrix inequalities, generalized inverses

1. Introduction

The dispersion matrices discussed in this paper arise in two different fields of linear model theory. First consider a random $\mathbf{R}^n$-vector $\mathbf{Y}$ with mean vector $\mathbf{X}\boldsymbol{\beta}$ and dispersion matrix $\sigma^2\mathbf{V}$, denoted by $\mathbf{Y}\sim(\mathbf{X}\boldsymbol{\beta};\sigma^2\mathbf{V})$ for short. As usual, the $n\times k$ matrix $\mathbf{X}$ and the symmetric non-negative definite $n\times n$ matrix $\mathbf{V}$ are assumed to be given while $\boldsymbol{\beta}=(\beta_1,\dots,\beta_k)'$ and σ^2 form the unknown parameters, where the prime denotes transposition. When $\mathbf{X}$ has full column rank k and $\mathbf{V}$ is positive definite, the minimum variance linear unbiased estimator (BLUE) for $\boldsymbol{\beta}$ is $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}$, which has dispersion matrix proportional to

$$(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}. \tag{1}$$

Secondly, consider the simpler model $\mathbf{Y}\sim(\mathbf{X}\boldsymbol{\beta};\sigma^2\mathbf{I}_n)$ and suppose that the function $\mathbf{K}'\boldsymbol{\beta}$ is to be estimated, with $\mathbf{K}$ a given $k\times t$ matrix. When $\mathbf{X}$ has full column rank k, the BLUE for $\mathbf{K}'\boldsymbol{\beta}$ is $\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$, which has dispersion matrix proportional to

$$\mathbf{K}'\mathbf{A}^{-1}\mathbf{K}, \tag{2}$$

where $\mathbf{A}=\mathbf{X}'\mathbf{X}$. When $\mathbf{K}$ has full column rank t, the information matrix for $\mathbf{K}'\boldsymbol{\beta}$ is proportional to $(\mathbf{K}'\mathbf{A}^{-1}\mathbf{K})^{-1}$ which has the same form as (1). The analogy between (1) and (2) does not seem to have been noticed before, and is very helpful in that it reconciles the two objectives of making (1) small when it is viewed as a dispersion matrix, or of making it large when it is taken to be an information matrix.

We now dispose of the full rank assumptions. Define $\mathscr{A}(\mathbf{K})$ to be the set of all symmetric non-negative definite $k\times k$ matrices $\mathbf{A}$ whose range (column space) $\mathscr{R}(\mathbf{A})$ contains the range $\mathscr{R}(\mathbf{K})$:

$$\mathscr{A}(\mathbf{K})=\{\mathbf{A}\geqslant\mathbf{0}\,|\,\mathscr{R}(\mathbf{A})\supset\mathscr{R}(\mathbf{K})\}. \tag{3}$$

The set $\mathscr{A}(\mathbf{K})$ is a convex cone, i.e., $\alpha\mathbf{A}$ and $\mathbf{A}+\mathbf{B}$ lie in $\mathscr{A}(\mathbf{K})$ for all $\mathbf{A},\mathbf{B}\in\mathscr{A}(\mathbf{K})$, $\alpha>0$. Let $\mathbf{A}^-$ denote an arbitrary generalized inverse of $\mathbf{A}$, i.e., $\mathbf{A}^-$ satisfies $\mathbf{A}\mathbf{A}^-\mathbf{A}=\mathbf{A}$. For $\mathbf{A}\in\mathscr{A}(\mathbf{K})$ the matrix $\mathbf{K}'\mathbf{A}^-\mathbf{K}$ is invariant to the choice of $\mathbf{A}^-$, and $\mathbf{K}'\mathbf{A}^-\mathbf{K}$ is symmetric non-negative definite, and has the same range and rank as $\mathbf{K}'$. Let $\mathbf{A}^+$ denote the Moore-Penrose inverse of $\mathbf{A}$, i.e., $\mathbf{A}^+$ is the unique matrix satisfying $\mathbf{A}\mathbf{A}^+\mathbf{A}=\mathbf{A}$, $\mathbf{A}^+\mathbf{A}\mathbf{A}^+=\mathbf{A}^+$, $\mathbf{A}\mathbf{A}^+=(\mathbf{A}\mathbf{A}^+)'$, and $\mathbf{A}^+\mathbf{A}=(\mathbf{A}^+\mathbf{A})'$. For $\mathbf{A}\in\mathscr{A}(\mathbf{K})$ we define

10–838192

$$f(\mathbf{A})=(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+, \quad g(\mathbf{A})=\mathbf{K}'\mathbf{A}^-\mathbf{K}. \tag{4}$$

When $\mathbf{A}$ and $\mathbf{K}$ have full column rank then (4) reduces to (1) and (2). Notice that f is positively homogeneous, i.e., $f(\alpha\mathbf{A})=\alpha f(\mathbf{A})$ for all $\mathbf{A}\in\mathcal{A}(\mathbf{K})$, $\alpha>0$.

In the context of estimating the regression coefficients in the linear model $\mathbf{Y}\sim(\mathbf{X}\boldsymbol{\beta};\sigma^2\mathbf{V})$, the condition $\mathcal{R}(\mathbf{X})\subset\mathcal{R}(\mathbf{V})$ means that all of the estimation space $\mathcal{R}(\mathbf{X})$ lies in the subspace $\mathcal{R}(\mathbf{V})$ on which the distributions of $\mathbf{Y}$ are concentrated. When $\mathbf{V}=\mathbf{I}$ then the condition $\mathcal{R}(\mathbf{K})\subset\mathcal{R}(\mathbf{X}'\mathbf{X})$ ensures estimability of $\mathbf{K}'\boldsymbol{\beta}$, cf. Alalouf & Styan (1979, Th. 2). These range-inclusion conditions define a matrix pre-ordering which will be studied elsewhere. Notice that the cone $\mathcal{A}(\mathbf{K})$ is slightly bigger than the cone $PD(k)$ of positive definite $k\times k$ matrices, in that $\mathcal{A}(\mathbf{K})$ may include boundary points of $PD(k)$. Inclusion of these boundary matrices is essential for a satisfactory solution in optimal design problems, cf. Pukelsheim (1980).

For two symmetric matrices $\mathbf{A},\mathbf{B}$ of the same order the notation $\mathbf{A}\leqslant\mathbf{B}$, or equivalently, $\mathbf{B}\geqslant\mathbf{A}$, will mean that $\mathbf{B}-\mathbf{A}$ is non-negative definite. Convexity statements refer to this (Loewner-)ordering, while isotone and antitone mean order-preserving and order-reversing, respectively. We shall repeatedly use the fact that, given $\mathbf{0}\leqslant\mathbf{A}\leqslant\mathbf{B}$, then $\mathbf{B}^+\leqslant\mathbf{A}^+$ if and only if rank $\mathbf{A}=\operatorname{rank}\mathbf{B}$, cf. Milliken & Akdeniz (1977, Th. 3.1).

2. Results

We shall deduce the concavity of f from the following version of the Gauss-Markov Theorem describing which estimates $\mathbf{LY}$ are BLUE for $\mathbf{K}\boldsymbol{\beta}$ in the model $\mathbf{Y}\sim(\mathbf{K}\boldsymbol{\beta};\sigma^2\mathbf{A})$.

Lemma 1. *Suppose the symmetric $k\times k$ matrix $\mathbf{A}\in\mathcal{A}(\mathbf{K})=\{\mathbf{A}\geqslant\mathbf{0}\,|\,\mathcal{R}(\mathbf{A})\supset\mathcal{R}(\mathbf{K})\}$, where $\mathbf{K}$ is $k\times t$, and let $\mathbf{L}$ be some $k\times k$ matrix satisfying $\mathbf{LK}=\mathbf{K}$. Then $\mathbf{LAL}'\geqslant\mathbf{K}(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}'$, and equality holds if and only if $\mathbf{LA}=\mathbf{K}(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}$.*

Proof. The condition $\mathbf{A}\in\mathcal{A}(\mathbf{K})$ implies $\mathbf{K}'\mathbf{A}^-\mathbf{A}=\mathbf{K}'$; if $\mathbf{B}=\mathbf{L}-\mathbf{K}(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}'\mathbf{A}^-$ then $\mathbf{0}\leqslant\mathbf{BAB}'=\mathbf{LAL}'-\mathbf{K}(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}'$. Thus $\mathbf{LAL}'\geqslant\mathbf{K}(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}'$, with equality if and only if $\mathbf{BA}=\mathbf{0}$, or equivalently, $\mathbf{LA}=\mathbf{K}(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}'=\mathbf{LAL}'$. □

The equality condition in Lemma 1 shows that the estimator $\mathbf{LY}$ for $\mathbf{K}\boldsymbol{\beta}$ is BLUE if and only if $\mathbf{LY}$ coincides on the space of random variation $\mathbf{Y}\in\mathcal{R}(\mathbf{A})$ with the generalized least squares estimator $\mathbf{K}(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}'\mathbf{A}^-\mathbf{Y}$.

It is of interest to note that the matrix

$$\mathbf{S}=\mathbf{LAL}'-\mathbf{K}(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}'$$

is the Schur complement of $\mathbf{K}'\mathbf{A}^-\mathbf{K}$ in the partitioned matrix

$$\mathbf{M}=\begin{pmatrix}\mathbf{LAL}' & \mathbf{K}\\ \mathbf{K}' & \mathbf{K}'\mathbf{A}^-\mathbf{K}\end{pmatrix}=\begin{pmatrix}\mathbf{L}\\ \mathbf{K}'\mathbf{A}^{-\prime}\end{pmatrix}\mathbf{A}(\mathbf{L}',\mathbf{A}^-\mathbf{K}),$$

with $\mathbf{LK}=\mathbf{K}=\mathbf{AA}^-\mathbf{K}$. Then

$$\operatorname{rank}(\mathbf{M})=\operatorname{rank}(\mathbf{K})+\operatorname{rank}(\mathbf{S})=\operatorname{rank}(\mathbf{AL}',\mathbf{K}),$$

since rank is additive on the Schur complement, cf. Ouellette (1981, § 4.2). Hence $\mathbf{S}=\mathbf{0}$ if and only if $\mathcal{R}(\mathbf{AL}')\subset\mathcal{R}(\mathbf{K})$. However, $\mathbf{S}\geqslant\mathbf{0}$ implies that $\mathcal{R}(\mathbf{K})\subset\mathcal{R}(\mathbf{LA})$. Thus $\mathbf{S}=\mathbf{0}$ if and only if $\mathcal{R}(\mathbf{AL}')=\mathcal{R}(\mathbf{K})$. From Lemma 1 it follows that $\mathbf{AL}'=\mathbf{LA}$ is symmetric.

Simple least squares estimation corresponds to $\mathbf{L}=\mathbf{KK}^+$ and entails $\mathbf{K}^+\mathbf{AK}^{+\prime}\geqslant(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+$, with equality if and only if $\mathcal{R}(\mathbf{AK})=\mathcal{R}(\mathbf{K})$, cf. Zyskind (1967, Th. 2), and

Gaffke & Krafft (1977, Lemma 1). The choice $\mathbf{L}=\mathbf{I}_k$ yields $\mathbf{A}\geqslant\mathbf{K}(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}'$, with equality if and only if $\mathscr{R}(\mathbf{A})=\mathscr{R}(\mathbf{K})$, cf. Pukelsheim (1980, Cor. 8.4).

Theorem 2. (i) *Suppose* $\mathbf{A},\mathbf{B}\in\mathscr{A}(\mathbf{K})$. *Then*

$$\{\mathbf{K}'(\mathbf{A}+\mathbf{B})^-\mathbf{K}\}^+\geqslant(\mathbf{K}'\mathbf{A}^-\mathbf{K})^++(\mathbf{K}'\mathbf{B}^-\mathbf{K})^+, \tag{5}$$

and equality if and only if

$$\begin{aligned}\{\mathbf{K}'(\mathbf{A}+\mathbf{B})^-\mathbf{K}\}^+\mathbf{K}'(\mathbf{A}+\mathbf{B})^-\mathbf{A}&=(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}',\\ \{\mathbf{K}'(\mathbf{A}+\mathbf{B})^-\mathbf{K}\}^+\mathbf{K}'(\mathbf{A}+\mathbf{B})^-\mathbf{B}&=(\mathbf{K}'\mathbf{B}^-\mathbf{K})^+\mathbf{K}';\end{aligned} \tag{6}$$

a sufficient condition for equality in (5) *is*

$$(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}'\mathbf{A}^-=(\mathbf{K}'\mathbf{B}^-\mathbf{K})^+\mathbf{K}'\mathbf{B}^-, \tag{7}$$

and this is also necessary whenever $\mathbf{AA}^-=\mathbf{BB}^-$.
(ii) *Suppose* $\mathbf{A},\mathbf{B}\in\mathscr{A}(\mathbf{K})$ *and* $p,q>0$, $p+q=1$. *Then*

$$\mathbf{K}'(p\mathbf{A}+q\mathbf{B})^-\mathbf{K}\leqslant p\mathbf{K}'\mathbf{A}^-\mathbf{K}+q\mathbf{K}'\mathbf{B}^-\mathbf{K}, \tag{8}$$

and equality holds if and only if

$$\mathbf{K}'(p\mathbf{A}+q\mathbf{B})^-\mathbf{K}=\mathbf{K}'\mathbf{A}^-\mathbf{K}=\mathbf{K}'\mathbf{B}^-\mathbf{K}; \tag{9}$$

a sufficient condition for equality in (8) *is*

$$\mathbf{A}^-\mathbf{K}=\mathbf{B}^-\mathbf{K}, \tag{10}$$

and this is also necessary whenever $\mathbf{AA}^-=\mathbf{BB}^-$.

Proof. (i) Choose $\mathbf{L}=\mathbf{K}\{\mathbf{K}'(\mathbf{A}+\mathbf{B})^-\mathbf{K}\}^+\mathbf{K}'(\mathbf{A}+\mathbf{B})^-$ in Lemma 1, and so $\mathbf{LK}=\mathbf{K}$. Thus $\mathbf{LAL}'\geqslant\mathbf{K}(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}'$ and $\mathbf{LBL}'\geqslant\mathbf{K}(\mathbf{K}'\mathbf{B}^-\mathbf{K})^+\mathbf{K}$. Summation yields $\mathbf{L}(\mathbf{A}+\mathbf{B})\mathbf{L}'=\mathbf{K}\{\mathbf{K}'(\mathbf{A}+\mathbf{B})^-\mathbf{K}\}^+\mathbf{K}'\geqslant\mathbf{K}(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}'+\mathbf{K}(\mathbf{K}'\mathbf{B}^-\mathbf{K})^+\mathbf{K}'$, and pre- and postmultiplication by $\mathbf{K}^+$ and $\mathbf{K}^{+\prime}$ establish (5). Condition (6) is copied from Lemma 1, and, if $\mathbf{AA}^-=\mathbf{BB}^-$, implies (7). That (7) is always sufficient follows from the identity

$$\begin{aligned}&\{(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}'\mathbf{A}^--(\mathbf{K}'\mathbf{B}^-\mathbf{K})^+\mathbf{K}'\mathbf{B}^-\}\mathbf{B}(\mathbf{A}+\mathbf{B})^-\mathbf{K}\{\mathbf{K}'(\mathbf{A}+\mathbf{B})^-\mathbf{K}\}^+\\ &\quad=\{\mathbf{K}'(\mathbf{A}+\mathbf{B})^-\mathbf{K}\}^+-(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+-(\mathbf{K}'\mathbf{B}^-\mathbf{K})^+.\end{aligned}$$

(ii) It follows from (i) that $\mathbf{K}'(p\mathbf{A}+q\mathbf{B})^-\mathbf{K}\leqslant\{p(\mathbf{K}'\mathbf{A}^-\mathbf{K})^++q(\mathbf{K}'\mathbf{B}^-\mathbf{K})^+\}^+$. Now let s be the rank of $\mathbf{K}$, and choose some $k\times s$ matrix $\mathbf{F}$ and some $t\times s$ matrix $\mathbf{G}$ such that $\mathbf{K}=\mathbf{FG}'$ and $\mathbf{G}'\mathbf{G}=\mathbf{I}_s$. Then $(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+=\mathbf{G}(\mathbf{F}'\mathbf{A}^-\mathbf{F})^{-1}\mathbf{G}'$, and convexity of non-singular inversion (Marshall & Olkin 1979, p. 469) gives $\{p(\mathbf{K}'\mathbf{A}^-\mathbf{K})^++q(\mathbf{K}'\mathbf{B}^-\mathbf{K})\}^+=\mathbf{G}\{p(\mathbf{F}'\mathbf{A}^-\mathbf{F}^{-1}+q(\mathbf{F}'\mathbf{B}^-\mathbf{F})^{-1}\}^{-1}\mathbf{G}'\leqslant\mathbf{G}(p\mathbf{F}'\mathbf{A}^-\mathbf{F}+q\mathbf{F}'\mathbf{B}^+\mathbf{F})\mathbf{G}'=p\mathbf{K}'\mathbf{A}^-\mathbf{K}+q\mathbf{K}'\mathbf{B}^-\mathbf{K}$, with equality if and only if $\mathbf{F}'\mathbf{A}^-\mathbf{F}=\mathbf{F}'\mathbf{B}^-\mathbf{F}$, or equivalently, $\mathbf{K}'\mathbf{A}^-\mathbf{K}=\mathbf{K}'\mathbf{B}^-\mathbf{K}$. This establishes (8) and, in conjuction with (6), also (9). Since (10) is equivalent to $\mathbf{K}'\mathbf{A}^-\mathbf{K}=\mathbf{K}'\mathbf{B}^-\mathbf{K}$ and (7), the proof is complete. □

When the idempotent matrices $\mathbf{AA}^-$ and $\mathbf{BB}^-$ do not coincide then (7) is not necessary for equality to hold in (5), e.g., $\mathbf{K}=\begin{pmatrix}1\\0\end{pmatrix}$, $\mathbf{A}=\begin{pmatrix}1&0\\0&0\end{pmatrix}$, $\mathbf{B}=\begin{pmatrix}3&-1\\-1&3\end{pmatrix}$. Our proof decomposes (8) into (5) and an inversion inequality whence equality in (8) requires more than equality in (5), a simple example is $\mathbf{K}=\begin{pmatrix}1\\0\end{pmatrix}$, $\mathbf{A}=\begin{pmatrix}2&0\\0&0\end{pmatrix}$, $\mathbf{B}=\begin{pmatrix}6&-2\\-2&6\end{pmatrix}$, $p=q=\frac{1}{2}$. Nor does $\mathbf{K}'\mathbf{A}^-\mathbf{K}=\mathbf{K}'\mathbf{B}^-\mathbf{K}$ alone imply equality in (8) even when $\mathbf{AA}^-=\mathbf{BB}^-$, as demonstrated by $\mathbf{K}=\begin{pmatrix}1\\0\end{pmatrix}$, $\mathbf{A}=\begin{pmatrix}1&0\\0&1\end{pmatrix}$, $\mathbf{B}=\begin{pmatrix}2&2\\2&4\end{pmatrix}$, $p=3/4$, $q=1/4$. Equality of $\mathbf{AA}^-=\mathbf{BB}^-$ also appears in related problems: if $\mathbf{0}\leq\mathbf{A}\leq\mathbf{B}$ and the generalized inverses $\mathbf{A}^-$ and $\mathbf{B}^-$ are reflexive (i.e., $\mathbf{A}^-\mathbf{AA}^-=\mathbf{A}^-, \mathbf{AA}^-\mathbf{A}=\mathbf{A}$) and symmetric, then $\mathbf{B}^-\leq\mathbf{A}^-$ if and only if $\mathbf{AA}^-=\mathbf{BB}^-$, cf. Pukelsheim & Styan (1978, Th. 2.3), and Styan & Pukelsheim (1978). Next we turn to monotonicity properties.

Theorem 3. *Suppose* $\mathbf{A},\mathbf{B}\in\mathscr{A}(\mathbf{K})$ *and* $\mathbf{A}\leq\mathbf{B}$. *Then* $(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\leq(\mathbf{K}'\mathbf{B}^-\mathbf{K})^+$ *and* $\mathbf{K}'\mathbf{B}^-\mathbf{K}\leq\mathbf{K}'\mathbf{A}^-\mathbf{K}$, *and in each case equality holds if and only if* $\mathbf{AB}^-\mathbf{K}=\mathbf{K}$.

Proof. The two inequalities are equivalent. The second inequality and the equality characterization follow from premultiplying the block matrix $\{(\mathbf{B}-\mathbf{A})^{1/2+}\mathbf{C}, \mathbf{A}^{1/2+}\mathbf{C}\}$ by its transpose, where $\mathbf{C}=(\mathbf{B}-\mathbf{A})\mathbf{B}^-\mathbf{K}=\mathbf{K}-\mathbf{AB}^-\mathbf{K}$. □

The condition $\mathbf{AB}^-\mathbf{K}=\mathbf{K}$ is also instrumental in discussing multiplicity of optimal experimental designs, cf. Pukelsheim (1980, Cor. 5.3). Concerning the functions in (4) we may summarize as follows.

Corollary 4. *On the convex cone* $\mathscr{A}(\mathbf{K})$ *the function f is concave and isotone, and the function g is convex and antitone.* □

3. Discussion

Theorem 2(i) has an appealing statistical interpretation. Consider two linear models $\mathbf{Y}\sim(\mathbf{K}\boldsymbol{\beta};\mathbf{A})$ and $\mathbf{Z}\sim(\mathbf{K}\boldsymbol{\beta};\mathbf{B})$, with $\mathbf{A},\mathbf{B}\in\mathscr{A}(\mathbf{K})$, and assume that the observation vectors $\mathbf{Y}$ and $\mathbf{Z}$ are uncorrelated. A first estimate for $\boldsymbol{\beta}$ is obtained from averaging the individual estimates of each model. Using weights p and q this leads to $p(\mathbf{K}'\mathbf{A}^-\mathbf{K})^+\mathbf{K}'\mathbf{A}^-\mathbf{Y}+q(\mathbf{K}'\mathbf{B}^-\mathbf{K})^+\mathbf{K}'\mathbf{B}^-\mathbf{Z}$, with dispersion matrix $p^2(\mathbf{K}'\mathbf{A}^-\mathbf{K})^++q^2(\mathbf{K}'\mathbf{B}^-\mathbf{K})^+$. A second estimate, based on averaging the observations rather than the estimates, is $\{\mathbf{K}'(p^2\mathbf{A}+q^2\mathbf{B})^-\mathbf{K}\}^+\mathbf{K}'(p^2\mathbf{A}+q^2\mathbf{B})^-(p\mathbf{Y}+q\mathbf{Z})$ with dispersion matrix $\{\mathbf{K}'(p^2\mathbf{A}+q^2\mathbf{B})\mathbf{K}\}^+$. Inequality (5) shows that the first course of action is strictly preferable, unless the second estimate, restricted to the relevant spaces of random variation, $\mathscr{R}(\mathbf{A})$ or $\mathscr{R}(\mathbf{B})$, is equal to the individual estimates in the respective models.

Theorem 2(ii) may be interpreted that the dispersion matrix $\mathbf{K}'(p\mathbf{A}+q\mathbf{B})^-\mathbf{K}$, associated with the non-randomized design $p\mathbf{A}+q\mathbf{B}$, is at most $p\mathbf{K}'\mathbf{A}^-\mathbf{K}+q\mathbf{K}'\mathbf{B}^-\mathbf{K}$, obtained from randomizing between $\mathbf{A}$ and $\mathbf{B}$ with weights p and q, cf. Kiefer (1961, p. 303).

Notice that restricted to positive definite matrices the behaviour of f and g is well known (Marshall & Olkin, 1979, pp. 468–473). By continuity this behaviour extends to all of $\mathscr{A}(\mathbf{K})$, cf. Pukelsheim & Styan (1978, Sect. 3). Gaffke & Krafft (1982, Th. 4.8) deduce our inequalities (5) and (8) from more general results. Yet another proof is obtained from quasi-linear representations of f and g, given by Sibson (1974, p. 682) and by Silvey (1980, p. 69). Namely, if $\mathbf{P}=\mathbf{I}_k-\mathbf{KK}^+$ then for all $k\times t$ matrices $\mathbf{H}$

$$g(\mathbf{A})\geq\mathbf{K}'\mathbf{H}+\mathbf{H}'\mathbf{K}-\mathbf{H}'\mathbf{AH},$$

$$f(\mathbf{A})\leq(\mathbf{K}^++\mathbf{H}'\mathbf{P})\mathbf{A}(\mathbf{K}^++\mathbf{H}'\mathbf{P})',$$

with equality for $\mathbf{H}=\mathbf{A}^-\mathbf{K}$ and $\mathbf{H}=-(\mathbf{PAP})^+\mathbf{AK}^{+\prime}$, respectively. However, none of these approaches reveal that the Gauss-Markov Theorem is the common statistical denominator, nor do they exhibit that f plays a more primary role than g, nor do they lead to the equality characterizations obtained in our Theorems 2 and 3.

Acknowledgements

This research was partly carried out while the first author was a postdoctoral fellow of the Deutsche Forschungsgemeinschaft with the Department of Statistics, Stanford University; the research of the second author was supported in part by the Natural Sciences and Engineering Research Council of Canada, and by the Gouvernement du Québec, Programme de formation de chercheurs de l'action concertée. An initial version of parts of this paper appeared in Pukelsheim & Styan (1978).

References

Alalouf, I. S. & Styan, G. P. H. (1979). Characterizations of estimability in the general linear model. *Ann. Statist.* **7**, 194–200.

Gaffke, N. & Krafft, O. (1977). Optimum properties of Latin square designs and a matrix inequality. *Math. Operationsforsch. Statist. Ser. Statist.* **8**, 345–350.

Gaffke, N. & Krafft, O. (1982). Matrix inequalities in the Löwner ordering. In *Modern applied mathematics: Optimization and operations research* (ed. B. Korte), pp. 595–622. North-Holland, Amsterdam.

Kiefer, J. (1961). Optimum designs in regression problems, II. *Ann. Math. Statist.* **32**, 298–325.

Marshall, A. W. & Olkin, I. (1979). *Inequalities: Theory of majorization and its applications.* Academic Press, New York.

Milliken, G. A. & Akdeniz, F. (1977). A theorem on the difference of the generalized inverses of two nonnegative matrices. *Comm. Statist.—Theory Methods* **A6**, 73–79.

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

Pukelsheim, F. (1980). On linear regression designs which maximize information. *J. Statist. Plann. Inference* **4**, 339–364.

Pukelsheim, F. & Styan, G. P. H. (1978). Nonnegative definiteness of matrix differences and monotonicity and convexity properties of dispersion matrices. Technical Report No. 137, Dept. of Statistics, Stanford University.

Sibson, R. (1974). D_A-optimality and duality. In *Progress in statistics. European Meeting of Statisticians, Budapest (Hungary) 1972. Coll. Math. Soc. J. Bolyai 9* (ed. J. Gani, K. Sarkadi & I. Vincze), vol. II, pp. 677–692. North-Holland, Amsterdam.

Silvey, S. D. (1980). *Optimal design.* Chapman and Hall, London.

Styan, G. P. H. & Pukelsheim, F. (1978). On the nonnegative definiteness of the difference between two matrices. Mimeographed report, 7 pp. [Abstract 78t–160, *Inst. Math. Statist. Bull.* **7**, (1978), 364.]

Zyskind, G. (1967). On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models. *Ann. Math. Statist.* **38**, 1092–1109.

Received August 1981, in final form February 1983

Friedrich Pukelsheim, Lehrstuhl für Stochastik und ihre Anwendungen, Universität Augsburg, Memminger Straße 6, D-8900 Augsburg, Federal Republic of Germany