Open Mathematics

Research Article

Czesław Stępniak*

Quotient of information matrices in comparison of linear experiments for quadratic estimation

https://doi.org/10.1515/math-2017-0135
Received April 13, 2017; accepted August 23, 2017.

Abstract: The ordering of normal linear experiments with respect to quadratic estimation, introduced by Stępniak in [Ann. Inst. Statist. Math. A 49 (1997), 569-584], is extended here to the experiments involving the nuisance parameters. Typical experiments of this kind are induced by allocations of treatments in the blocks. Our main tool, called quotient of information matrices, may be interesting itself. It is known that any orthogonal allocation of treatments in blocks is optimal with respect to linear estimation of all treatment contrasts. We show that such allocation is, however, not optimal for quadratic estimation.

Keywords: Normal linear experiment, Comparison of experiments for quadratic estimation, Nuisance parameter, Quotient of information matrices, Orthogonal block design, Nonoptimality for quadratic estimation

MSC: 62K10, 05B20, 62J05

1 Introduction

Any statistical experiment may be perceived as an information channel transforming a deterministic quantity (parameter) into a random one (observation) according to a design indicated by experimenter. The primary aim of statistician is to recover the information about the parameter from the observation. However the efficiency of this process depends not only on the statistical rule but also on the experimental design. Such design, which may be identified with the experiment, is represented by a probabilistic structure.

When observations have normal distribution the entire statistical analysis is based on their linear and quadratic forms. Thus the properties of such forms should be taken into account in any reasonable choice of statistical experiment.

Comparison of linear experiments by linear forms has been intensively studied in statistical literature. It is well known (for instance [1–6]) that almost all criteria used for comparison of two linear experiments with respect to linear estimation reduce to the Loewner order between their information matrices, say $\mathbf{M}_1$ and $\mathbf{M}_2$. However, the comparison of normal linear experiments with respect to quadratic estimation is still at the initial stage and we are looking for respective tools.

It was revealed in Stępniak [7] that the relation "to be at least as good with respect to quadratic estimation" needs some knowledge about the matrix $\mathbf{M}_1^+\mathbf{M}_2$, where symbol $^+$ means the Moore-Penrose generalized inversion. We shall refer to this matrix as quotient of $\mathbf{M}_2$ by $\mathbf{M}_1$. Properties of such quotient may be interesting themselves. It appears that the Loewner order may be expressed in terms of the quotient, but not vice versa.

***Corresponding Author: Czesław Stępniak:** Faculty of Mathematics and Natural Sciences, University of Rzeszów, Pigonia 1, PL-35-959 Rzeszów, Poland, E-mail: stepniak@umcs.lublin.pl

In this note we use the quotient of positive semidefinite matrices as the main tool in the ordering of normal linear experiments with respect to quadratic estimation. The orderings of linear experiments with respect to linear and with respect to quadratic estimation are extended here to the experiments involving nuisance parameters. Typical experiments of this kind are induced by allocations of treatments in blocks.

It is well known (see [8]) that any orthogonal allocation of treatments in blocks is optimal by means of linear estimation of all treatment contrasts. We show that this allocation is, however, not optimal for quadratic estimation.

2 Definitions and known results

In this paper the standard vector-matrix notation is used. All vectors and matrices considered here have real entries. The space of all $n \times 1$ vectors is denoted by R^n. For any matrix $\mathbf{M}$ the symbols $\mathbf{M}^T$, $R(\mathbf{M})$, $N(\mathbf{M})$ and $r(\mathbf{M})$ denote, respectively, its transpose, range (column space), kernel (null space) and rank. The symbol $\mathbf{P_M}$ stands for the orthogonal projector onto $R(\mathbf{M})$, i.e. the square matrix $\mathbf{P}$ satisfying the conditions $\mathbf{Px} = \mathbf{x}$ for $\mathbf{x} \in R(\mathbf{M})$ and zero for $\mathbf{x} \in N(\mathbf{M}^T)$. Moreover, if $\mathbf{M}$ is square then $tr(\mathbf{M})$ denotes its trace and the symbol $\mathbf{M} \geq \mathbf{0}$ means that $\mathbf{M}$ is symmetric and positive semidefinite (psd, for short).

Let $\mathbf{x}$ be a random vector with the expectation $E(\mathbf{x}) = \mathbf{A}\boldsymbol{\alpha}$ and the variance-covariance matrix $\sigma\mathbf{V}$, where $\mathbf{A}$ and $\mathbf{V}$ are known matrices while $\boldsymbol{\alpha} = (\alpha_1, ...\alpha_p)^T$ and $\sigma > 0$ are unknown parameters. In this situation we shall say that $\mathbf{x}$ is subject to the linear experiment $\mathcal{L}(\mathbf{A}\boldsymbol{\alpha}, \sigma\mathbf{V})$. If $\mathbf{V} = \mathbf{I}$, then we say that the experiment is standard. If $\mathbf{x}$ is normally distributed then instead of $\mathcal{L}(\mathbf{A}\boldsymbol{\alpha}, \sigma\mathbf{V})$ we shall use the symbol $\mathcal{N}(\mathbf{A}\boldsymbol{\alpha}, \sigma\mathbf{V})$.

Now let us consider two experiments $\mathcal{L}_1 = \mathcal{L}(\mathbf{A}_1\boldsymbol{\alpha}, \sigma\mathbf{V})$ and $\mathcal{L}_2 = \mathcal{L}(\mathbf{A}_2\alpha, \sigma\mathbf{W})$ with the same parameters and with observation vectors $\mathbf{x} \in R^m$ and $\mathbf{y} \in R^n$, respectively.

Definition 2.1 ([9]). *Experiment $\mathcal{L}_1$ is said to be at least as good as $\mathcal{L}_2$ with respect to linear estimation [notation: $\mathcal{L}_1 \trianglerighteq \mathcal{L}_2$] if for any parametric function $\psi = \mathbf{c}^T\boldsymbol{\alpha}$ and for any estimator $\mathbf{b}^T\mathbf{y}$ there exists an estimator $\mathbf{a}^T\mathbf{x}$ with uniformly not greater squared risk. If $\mathcal{L}_1 \trianglerighteq \mathcal{L}_2$ and $\mathcal{L}_2 \trianglerighteq \mathcal{L}_1$ then we say that the experiments are equivalent for linear estimation.*

The relation $\trianglerighteq$ may be expressed in terms of linear forms (see [8,9]). Namely $\mathcal{L}_1 \trianglerighteq \mathcal{L}_2$, if and only if, for any $\mathbf{b} \in R^n$ there exists $\mathbf{a} \in R^m$ such that

$$E(\mathbf{a}^T\mathbf{x}) = E(\mathbf{b}^T\mathbf{y}) \text{ and } var(\mathbf{a}^T\mathbf{x}) \leq var(\mathbf{b}^T\mathbf{y}) \tag{1}$$

for all $\boldsymbol{\alpha}$ and σ. It is worth to note that the relation $\mathcal{L}(\mathbf{A}_1\boldsymbol{\alpha}, \sigma\mathbf{V}) \trianglerighteq \mathcal{L}(\mathbf{A}_2\boldsymbol{\alpha}, \sigma\mathbf{W})$ does not depend on whether σ is known or not. Thus $\mathcal{L}_1 \trianglerighteq \mathcal{L}_2$ if and only if $\mathcal{L}(\mathbf{A}_1\boldsymbol{\alpha}, \mathbf{V}) \trianglerighteq \mathcal{L}(\mathbf{A}_2\boldsymbol{\alpha}, \mathbf{W})$.

Moreover, under the normality assumption, the condition (1) may be expressed in the form:

> For any parametric function ψ and for any $\mathbf{b} \in R^n$ there exists $\mathbf{a} \in R^m$ such that $|\mathbf{a}^T\mathbf{x} - \psi|$ is stochastically not greater than $|\mathbf{b}^T\mathbf{y} - \psi|$ for all $\boldsymbol{\alpha}$ and σ

(Sinha [10] and Stępniak [11]).

Now consider two normal linear experiments $\mathcal{N}_1 = \mathcal{N}(\mathbf{A}\boldsymbol{\alpha}, \sigma\mathbf{V})$ and $\mathcal{N}_2 = \mathcal{N}(\mathbf{B}\boldsymbol{\alpha}, \sigma\mathbf{W})$ with observation vectors $\mathbf{x} \in R^m$ and $\mathbf{y} \in R^n$. It is well known (cf. [12,13]) that such experiments are not comparable with respect to all possible statistical problems. Therefore we shall restrict our attention to quadratic estimation only.

Definition 2.2 ([7]). *Experiment $\mathcal{N}_1$ is said to be at least as good as $\mathcal{N}_2$ with respect to quadratic estimation [notation: $\mathcal{N}_1 \geq \mathcal{N}_2$] if for any quadratic form $\mathbf{y}^T\mathbf{G}\mathbf{y}$ there exists a quadratic form $\mathbf{x}^T\mathbf{H}\mathbf{x}$ such that*

$$E(\mathbf{x}^T\mathbf{H}\mathbf{x}) = E(\mathbf{y}^T\mathbf{G}\mathbf{y}) \text{ and } var(\mathbf{x}^T\mathbf{H}\mathbf{x}) \leq var(\mathbf{y}^T\mathbf{G}\mathbf{y})$$

for all $\boldsymbol{\alpha}$ and σ. If $\mathcal{N}_1 \geq \mathcal{N}_2$ and $\mathcal{N}_2 \geq \mathcal{N}_1$ then we say that the experiments are equivalent for quadratic estimation.

In the last definition the quadratic forms $\mathbf{x}^T\mathbf{H}\mathbf{x}$ and $\mathbf{y}^T\mathbf{G}\mathbf{y}$ play the role of potential unbiased estimators for parametric functions of type $\phi(\boldsymbol{\alpha},\sigma) = c\sigma + \boldsymbol{\alpha}^T\mathbf{C}\boldsymbol{\alpha}$. It is known that any mean squared error of a linearly estimable parametric function $\psi = \psi(\boldsymbol{\alpha})$ in the experiment $\mathcal{N}_1$ (or in $\mathcal{N}_2$) has such a form (Stępniak [14]). The orderings $\trianglerighteq$ and $\succeq$ possess invariance property with respect to nonsingular linear transformation both the parameter $\boldsymbol{\alpha}$ and the observation vectors $\mathbf{x}$ and $\mathbf{y}$ as well ([7], Lemmas 2.1 and 2.2).

The main tool in comparison of the standard linear experiments is the information matrix $\mathbf{M}$ defined as the Fisher information matrix $\mathbf{A}^T\mathbf{A}$ corresponding to the experiment $\mathcal{N}(\mathbf{A}\boldsymbol{\alpha}, \mathbf{I})$.

The relation $\trianglerighteq$ may be characterized by the following theorem.

Theorem 2.3 ([15], Theorem 1). *For standard linear experiments $\mathcal{L}_1 = \mathcal{L}(\mathbf{A}_1\boldsymbol{\alpha},\sigma\mathbf{I}_m)$ and $\mathcal{L}_2 = \mathcal{L}(\mathbf{A}_2\boldsymbol{\alpha}, \sigma\mathbf{I}_n)$ with information matrices $\mathbf{M}_1 = \mathbf{A}_1^T\mathbf{A}_1$ and $\mathbf{M}_2 = \mathbf{A}_2^T\mathbf{A}_2$ the followings are equivalent:*

(a) $\mathcal{L}_1 \trianglerighteq \mathcal{L}_2$,

(b) $\mathbf{M}_1 - \mathbf{M}_2$ *is psd,*

(c) $R(\mathbf{M}_2) \subseteq R(\mathbf{M}_1)$ *and the maximal eigenvalue of the matrix* $\mathbf{M}_1^+\mathbf{M}_2$ *is not greater than* 1.

A corresponding result for the relation $\succeq$ is due to Stępniak ([7], Theorem 5.1) in the form

Theorem 2.4. *For standard normal linear experiments $\mathcal{N}_1 = \mathcal{N}(\mathbf{A}_1\boldsymbol{\alpha}, \sigma\mathbf{I}_m)$ and $\mathcal{N}_2 = \mathcal{N}(\mathbf{A}_2\boldsymbol{\alpha}, \sigma\mathbf{I}_n)$ with information matrices $\mathbf{M}_1 = \mathbf{A}_1^T\mathbf{A}_1$ and $\mathbf{M}_2 = \mathbf{A}_2^T\mathbf{A}_2$ the followings are equivalent:*

(a) $\mathcal{N}_1 \succeq \mathcal{N}_2$,

(b) $\mathbf{M}_1 - \mathbf{M}_2$ *is psd and*

$$\sum_{i=1}^{q}\frac{1-\lambda_i}{1+\lambda_i} \le m - n - r(\mathbf{A}_1) + r(\mathbf{A}_2), \tag{2}$$

where λ_i, $i = 1, \dots, q$, are the positive eigenvalues of the matrix $\mathbf{M}_1^+\mathbf{M}_2$, counted with their multiplicities.

(c) $R(\mathbf{M}_2) \subseteq R(\mathbf{M}_1)$, the maximal eigenvalue of the matrix $\mathbf{M}_1^+\mathbf{M}_2$ is not greater than 1 *and the inequality* (2) *holds.*

It is interesting that the both orderings $\trianglerighteq$ and $\succeq$ may be expressed in terms of the matrix $\mathbf{M}_1^+\mathbf{M}_2$, where $\mathbf{M}_j$, $j = 1, 2$, are information matrices corresponding to the experiments $\mathcal{N}(\mathbf{A}_1\boldsymbol{\alpha}, \sigma\mathbf{I}_m)$ and $\mathcal{N}(\mathbf{A}_2\boldsymbol{\alpha}, \sigma\mathbf{I}_n)$. Matrix of this kind will be called quotient of $\mathbf{M}_2$ by $\mathbf{M}_1$.

3 Quotient of matrices in comparison of experiments

For given psd matrices $\mathbf{T}$ and $\mathbf{U}$ of the same order we shall refer to the expressions $\mathbf{Q}_1 = \mathbf{T}\mathbf{U}^+$, $\mathbf{Q}_2 = \mathbf{U}^+\mathbf{T}$, $\mathbf{Q}_3 = (\mathbf{U}^+)^{1/2}\mathbf{T}(\mathbf{U}^+)^{1/2}$ and $\mathbf{Q}_4 = \mathbf{T}^{1/2}\mathbf{U}^+\mathbf{T}^{1/2}$ as versions of the quotient of $\mathbf{T}$ by $\mathbf{U}$. We note that only $\mathbf{Q}_3$ and $\mathbf{Q}_4$ are always symmetric.

We shall start from basic properties of the quotients.

Theorem 3.1. *For arbitrary positive semidefinite matrices* $\mathbf{T}$ *and* $\mathbf{U}$ *of the same order*

(a) All versions $\mathbf{Q}_1 = \mathbf{T}\mathbf{U}^+$, $\mathbf{Q}_2 = \mathbf{U}^+\mathbf{T}$, $\mathbf{Q}_3 = (\mathbf{U}^+)^{1/2}\mathbf{T}(\mathbf{U}^+)^{1/2}$ and $\mathbf{Q}_4 = \mathbf{T}^{1/2}\mathbf{U}^+\mathbf{T}^{1/2}$ of the quotient of $\mathbf{T}$ by $\mathbf{U}$ have the same eigenvalues.

(b) All eigenvalues of arbitrary quotient are nonnegative.

(c) $\mathbf{U} - \mathbf{T}$ is psd if and only if $R(\mathbf{T}) \subseteq R(\mathbf{U})$ and all eigenvalues of arbitrary quotient $\mathbf{Q}_i$ are not greater than 1.

Proof. (a) If $\mathbf{Q}_1 w = \lambda\mathbf{w}$ then $\mathbf{U}^+\mathbf{Q}_1\mathbf{w} = \mathbf{Q}_2\mathbf{U}^+\mathbf{w} = \lambda\mathbf{U}^+\mathbf{w}$ and λ is an eigenvalue of $\mathbf{Q}_2$. Conversely, if $\mathbf{Q}_2\mathbf{w} = \lambda\mathbf{w}$ then $\mathbf{Q}_1\mathbf{T}\mathbf{w} = \lambda\mathbf{T}\mathbf{w}$. Thus $\mathbf{Q}_1$ and $\mathbf{Q}_2$ have the same eigenvalues. To prove the same for $\mathbf{Q}_3$ and $\mathbf{Q}_4$ we note that $\mathbf{Q}_3 = \mathbf{F}\mathbf{F}^T$a nd $\mathbf{Q}_4 = \mathbf{F}^T\mathbf{F}$ for $\mathbf{F} = (\mathbf{U}^+)^{1/2}\mathbf{T}^{1/2}$ and the desired correspondence follows by the implications $\mathbf{Q}_3\mathbf{w} = \lambda\mathbf{w} \Longrightarrow \mathbf{Q}_4\mathbf{F}^T\mathbf{w} = \boldsymbol{\lambda}\mathbf{F}^T\mathbf{w}$ and $\mathbf{Q}_4\mathbf{w} = \lambda\mathbf{w} \Longrightarrow \mathbf{Q}_3\mathbf{F}\mathbf{w} = \boldsymbol{\lambda}\mathbf{F}\mathbf{w}$. Thus it remains to show a similar correspondence for $\mathbf{Q}_2$ and $\mathbf{Q}_3$.

The equality $\mathbf{Q}_2\mathbf{w} = \lambda\mathbf{w}$ implies $\mathbf{Q}_4\mathbf{T}^{1/2}\mathbf{w} = \lambda\mathbf{T}^{1/2}\mathbf{w}$. Thus λ is an eigenvalue of $\mathbf{Q}_4$ and, in consequence, of $\mathbf{Q}_3$. Similarly, if $\mathbf{Q}_3\mathbf{w} = \boldsymbol{\lambda}\mathbf{w}$ then $\mathbf{Q}_2(\mathbf{U}^+)^{1/2}\mathbf{w} = \boldsymbol{\lambda}(\mathbf{U}^+)^{1/2}\mathbf{w}$. This implies the desired condition and completes the proof of the part (a).

(b) It follows immediately from (a).

(c) By (a) we only need to show the desired equivalence for $i = 3$. Implication $\mathbf{U} - \mathbf{T} \geq \mathbf{0} \implies R(\mathbf{T}) \subseteq R(\mathbf{U})$ is evident. For the remain we note that under assumption $R(\mathbf{T}) \subseteq R(\mathbf{U})$, $\lambda_{\max}(\mathbf{Q}_3) \leq 1$ if and only if $(\mathbf{U}^+)^{1/2}\mathbf{U}(\mathbf{U}^+)^{1/2} \geq (\mathbf{U}^+)^{1/2}\mathbf{T}(\mathbf{U}^+)$. This implies (c) and completes the proof of Theorem 2.4. □

Now we shall use Theorem 3.1 to comparison of normal linear experiments $\mathcal{N}_1 = \mathcal{N}(\mathbf{A}_1\boldsymbol{\alpha}, \sigma\mathbf{I}_m)$ and $\mathcal{N}_2 = \mathcal{N}(\mathbf{A}_2\boldsymbol{\alpha}, \sigma\mathbf{I}_n)$ w.r.t. quadratic estimation.

We note that $r(\mathbf{A}_1) = r(\mathbf{M}_1)$ while $n - r(\mathbf{A}_1)$ means the number of degrees of freedom in the experiment $\mathcal{N}_1$. By Theorem 2.4 we get the following result.

Lemma 3.2. *If the numbers of degrees of freedom in the experiments $\mathcal{N}_1 = \mathcal{N}(\mathbf{A}_1\boldsymbol{\alpha}, \sigma\mathbf{I}_m)$ and $\mathcal{N}_2 = \mathcal{N}(\mathbf{A}_2\boldsymbol{\alpha}, \sigma\mathbf{I}_n)$ are equal then $\mathcal{N}_1 \geq \mathcal{N}_2$ if and only if $\mathbf{M}_1 - \mathbf{M}_2$ is psd and any quotient $\mathbf{Q}_i$, $i = 3, 4$, of the information matrix $\mathbf{M}_2$ by $\mathbf{M}_1$ is idempotent, i.e. $\mathbf{Q}_i^2 = \mathbf{Q}_i$.*

Proof. Under our assumption the right side of the inequality (2) is 0 and hence each eigenvalue of any quotient $\mathbf{Q}_i$ is either 0 or 1. Since the quotients $\mathbf{Q}_i$, $i = 3, 4$, are symmetric this is equivalent their idempotency. □

The case when the numbers of observations in the both experiments are equal, i.e. $m = n$, is the most interesting. In this case by Theorem 3.1 we get

Lemma 3.3. *For standard normal experiments $\mathcal{N}_1$ and $\mathcal{N}_2$ with the same number observations the relation $\mathcal{N}_1 \geq \mathcal{N}_2$ holds if and only if $\mathbf{M}_1 = \mathbf{M}_2$, i.e. when the experiments are equivalent.*

Proof. Assume that $\mathbf{M}_1 - \mathbf{M}_2$ is psd, the inequality (2) is true and $m = n$. Then $R(\mathbf{M}_1) = R(\mathbf{M}_2)$ and, by Lemma 3.2, $\mathbf{Q}_3$ is idempotent. Thus $\mathbf{Q}_3$ is the orthogonal projector onto $R(\mathbf{M}_1) = R(\mathbf{M}_2)$, and, in consequence, $\mathbf{Q}_3 = (\mathbf{M}_1^+)^{1/2}\mathbf{M}_1(\mathbf{M}_1^+)^{1/2}$. This implies the desired result. □

Now let us consider a linear experiment where observation vector $\mathbf{x}$ depends on several parameters but only some of them are of interest. More precisely, assume that

$$E(\mathbf{x}) = \mathbf{A}\boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\beta}$$

and

$$Cov(\mathbf{x}) = \sigma\mathbf{I}$$

with unknown parameters $\boldsymbol{\alpha} \in R^p$, $\boldsymbol{\beta} \in R^k$ and $\sigma > 0$ such that $\boldsymbol{\alpha}$ (or $\boldsymbol{\alpha}$ and σ) is of interest, while $\boldsymbol{\beta}$ is treated as the nuisance one. Such experiment will be denoted by $\mathcal{L}(\mathbf{A}\boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\beta}, \sigma\mathbf{I})$ or (under the normality assumption) by $\mathcal{N}(\mathbf{A}\boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\beta}, \sigma\mathbf{I})$.

We shall say that a statistic $\mathbf{t} = t(\mathbf{x})$ is invariant (with respect to $\boldsymbol{\beta}$) if its first two moments exist and they do not depend on $\boldsymbol{\beta}$. It is evident that a linear form $\mathbf{a}^T\mathbf{x}$ is invariant in the experiment $\mathcal{L}(\mathbf{A}\boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\beta}, \sigma\mathbf{I})$ if and only if it depends on $\mathbf{x}$ only through $(\mathbf{I} - \mathbf{P_B})\mathbf{x}$. The same condition for invariance of quadratic form $\mathbf{x}^T\mathbf{H}\mathbf{x}$ follows by the well known formula

$$var(\mathbf{x}^T\mathbf{H}\mathbf{x}) = 2\sigma^2 tr\mathbf{H}^2 + 4\sigma(\boldsymbol{\alpha}^T, \boldsymbol{\beta}^T)[\mathbf{A}, \mathbf{B}]^T\mathbf{H}^2[\mathbf{A}, \mathbf{B}](\boldsymbol{\alpha}^T, \boldsymbol{\beta}^T)^T$$

for variance of quadratic forms in normal variables (cf. [16,17]).

Now let us consider two linear experiments $\mathcal{L}_1 = \mathcal{L}(\mathbf{A}_1\boldsymbol{\alpha} + \mathbf{B}_1\boldsymbol{\beta}, \sigma\mathbf{I}_m)$ and $\mathcal{L}_2 = \mathcal{L}(\mathbf{A}_2\boldsymbol{\alpha} + \mathbf{B}_2\boldsymbol{\beta}, \sigma\mathbf{I}_n)$ (or $\mathcal{N}_1 = \mathcal{N}(\mathbf{A}_1\boldsymbol{\alpha} + \mathbf{B}_1\boldsymbol{\beta}, \sigma\mathbf{I}_m)$ and $\mathcal{N}_2 = \mathcal{N}(\mathbf{A}_2\boldsymbol{\alpha} + \mathbf{B}_2\boldsymbol{\beta}, \sigma\mathbf{I}_n)$) with observation vectors $\mathbf{x} \in R^m$ and $\mathbf{y} \in R^n$.

Definition 3.4. *We shall say that $\mathcal{L}_1$ is at least as good as $\mathcal{L}_2$ w.r.t. invariant linear estimation if for any invariant statistic $\mathbf{b}^T\mathbf{y}$ there exists an invariant $\mathbf{a}^T\mathbf{x}$ such that $E(\mathbf{a}^T\mathbf{x}) = E(\mathbf{b}^T\mathbf{y})$ and $var(\mathbf{a}^T\mathbf{x}) \leq var(\mathbf{b}^T\mathbf{y})$ for all $\boldsymbol{\alpha}$ and σ. Similarly, we shall say that $\mathcal{N}_1$ is at least as good as $\mathcal{N}_2$ w.r.t. invariant quadratic estimation if for any invariant* *statistic* $\mathbf{y}^T\mathbf{G}\mathbf{y}$ *there exists an invariant* $\mathbf{x}^T\mathbf{H}\mathbf{x}$ *such that* $E(\mathbf{x}^T\mathbf{H}\mathbf{x}) = E(\mathbf{y}^T\mathbf{G}\mathbf{y})$ *and* $var(\mathbf{x}^T\mathbf{H}\mathbf{x}) \le var(\mathbf{y}^T\mathbf{G}\mathbf{y})$ *for all* $\boldsymbol{\alpha}$ *and* σ.

First we shall reduce the comparison of linear experiments with a nuisance parameter $\boldsymbol{\beta}$ to the same problem for the usual linear experiments. To this aim we need the invariance condition in a a more explicit form.

Let $\mathbf{x}$ be observation vector in a linear experiment $\mathcal{L}(\mathbf{A}\boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\beta}, \sigma\mathbf{I})$ or $\mathcal{N}(\mathbf{A}\boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\beta}, \sigma\mathbf{I})$ and let $\mathbf{b}_1, \ldots, \mathbf{b}_{n-r}$ be orthonormal basis in $N(\mathbf{B}^T)$. Then $\mathbf{I} - \mathbf{P_B}$ may be presented in the form

$$\left[\mathbf{b}_1\ , \ldots, \mathbf{b}_{n-r}\right] \begin{bmatrix} \mathbf{b}_1^T \\ \cdot \\ \cdot \\ \mathbf{b}_{n-r}^T \end{bmatrix}.$$

Define

$$\widetilde{\mathbf{A}}_i = \begin{bmatrix} \mathbf{b}_1^T \\ \cdot \\ \cdot \\ \mathbf{b}_{n-r}^T \end{bmatrix} \mathbf{A}_i. \tag{3}$$

In this way $\mathcal{L}(\mathbf{A}_1\boldsymbol{\alpha} + \mathbf{B}_1\boldsymbol{\beta}, \sigma\mathbf{I}_m)$ is at least as good as $\mathcal{L}(\mathbf{A}_2\boldsymbol{\alpha} + \mathbf{B}_2\boldsymbol{\beta}, \sigma\mathbf{I}_n)$ w.r.t. invariant linear estimation if and only if $\mathcal{L}(\widetilde{\mathbf{A}}_1\boldsymbol{\alpha}, \sigma\mathbf{I}_{m-r_1}) \ge \mathcal{L}(\widetilde{\mathbf{A}_2}\boldsymbol{\alpha}, \sigma\mathbf{I}_{n-r_2})$, where $\widetilde{\mathbf{A}}_i$ is defined by (3) and $r_i = r(\mathbf{B}_i)$. Similarly $\mathcal{N}(\mathbf{A}_1\boldsymbol{\alpha} + \mathbf{B}_1\boldsymbol{\beta}, \sigma\mathbf{I}_m)$ is at least as good as $\mathcal{N}(\mathbf{A}_2\boldsymbol{\alpha} + \mathbf{B}_2\boldsymbol{\beta}, \sigma\mathbf{I}_n)$ w.r.t. invariant quadratic estimation if and only if $\mathcal{N}(\widetilde{\mathbf{A}}_1\boldsymbol{\alpha}, \sigma\mathbf{I}_{m-r_1}) \ge \mathcal{N}(\widetilde{\mathbf{A}}_2\boldsymbol{\alpha}, \sigma\mathbf{I}_{n-r_2})$.

For convenience the matrices $\widetilde{\mathbf{A}}_i^T\widetilde{\mathbf{A}}_i$, $i = 1, 2$, will be called the reduced information matrices and will be denoted by $\widetilde{\mathbf{M}_i}$. We note that

$$\widetilde{\mathbf{M}_i} = \mathbf{A}_i^T(\mathbf{I} - \mathbf{P}_{\mathbf{B}_i})\mathbf{A}_i \tag{4}$$

As a direct consequence of Theorems 2.3 and 2.4 we get the following lemmas.

Lemma 3.5. *For arbitrary linear experiments* $\mathcal{L}_1 = \mathcal{L}(\mathbf{A}_1\boldsymbol{\alpha} + \mathbf{B}_1\boldsymbol{\beta}, \sigma\mathbf{I}_m)$ *and* $\mathcal{L}_2 = \mathcal{L}(\mathbf{A}_2\boldsymbol{\alpha} + \mathbf{B}_2\boldsymbol{\beta}, \sigma\mathbf{I}_n)$, $\mathcal{L}_1$ *is at least as good as* $\mathcal{L}_2$ *w.r.t. invariant linear estimation if and only if* $\widetilde{\mathbf{M}}_1 - \widetilde{\mathbf{M}}_2 \ge \mathbf{0}$.

Lemma 3.6. *For arbitrary normal linear experiments* $\mathcal{N}_1 = \mathcal{N}(\mathbf{A}_1\boldsymbol{\alpha} + \mathbf{B}_1\boldsymbol{\beta}, \sigma\mathbf{I}_m)$ *and* $\mathcal{N}_2 = \mathcal{N}(\mathbf{A}_2\boldsymbol{\alpha} + \mathbf{B}_2\boldsymbol{\beta}, \sigma\mathbf{I}_n)$, $\mathcal{N}_1$ *is at least as good as* $\mathcal{N}_2$ *w.r.t. invariant quadratic estimation if and only if* $\widetilde{\mathbf{M}}_1 - \widetilde{\mathbf{M}}_2 \ge \mathbf{0}$ *and*

$$\sum_{i=1}^{q} \frac{1-\lambda_i}{1+\lambda_i} \le m - r(\widetilde{\mathbf{M}}_1) - r(\mathbf{B}_1) - [n - r(\widetilde{\mathbf{M}}_2) - r(\mathbf{B}_2)],$$

where λ_i, $i = 1, \ldots, q$, *are positive eigenvalues of arbitrary version of the quotient of* $\widetilde{\mathbf{M}}_2$ *by* $\widetilde{\mathbf{M}}_1$, *counted with their multiplicities.*

In particular, if $m - r(\mathbf{B}_1) = n - r(\mathbf{B}_2)$ then by Lemma 3.2 we get

Corollary 3.7. *If* $n - r(\mathbf{B}_1) = m - r(\mathbf{B}_2)$ *then* $\mathcal{N}_1$ *is at least as good as* $\mathcal{N}_2$ *w.r.t. invariant quadratic estimation if and only if the matrix* $(\widetilde{\mathbf{M}}_1^+)^{1/2}\widetilde{\mathbf{M}}_2(\widetilde{\mathbf{M}}_1^+)^{1/2}$ *is idempotent.*

Similarly, by Lemmas 3.3 and 3.6 we get

Corollary 3.8. *If* $m = n$ *and* $\mathbf{B}_1 = \mathbf{B}_2$ *then* $\mathcal{N}_1$ *is at least as good as* $\mathcal{N}_2$ *w.r.t. invariant quadratic estimation if and only if* $\widetilde{\mathbf{M}}_1 = \widetilde{\mathbf{M}}_2$.

4 Problem of optimal allocation of treatments in blocks

Consider allocation of v treatments with replications $t_1, \ldots, t_v$ in k blocks of sizes $b_1, \ldots, b_k$, where $\sum_i t_i = \sum_j b_j = n$. Let us introduce matrices $\mathbf{B} = diag(\mathbf{1}_{b_1}, \ldots, \mathbf{1}_{b_k})$ and $\mathbf{D} = (d_{ij})$, where

$$d_{ij} = \begin{cases} 1, & \text{if the } i\text{-th observation refers to the } j\text{-th treatment,} \\ 0, & \text{otherwise.} \end{cases}$$

These matrices indicate allocation of treatments in blocks. For this reason $\mathbf{D}$ is sometimes identified with block design.

To each pair $(\mathbf{B}, \mathbf{D})$ corresponds a linear experiment $\mathcal{L} = \mathcal{L}(\mathbf{D}\boldsymbol{\alpha}+[\mathbf{1}_n, \mathbf{B}]\boldsymbol{\beta}, \sigma\mathbf{I}_n)$, where $\boldsymbol{\alpha} = (\alpha_{1,\ldots,}\alpha_v)^T$ refers to the treatment effects, while $\boldsymbol{\beta} = (\mu, \beta_1, \ldots, \beta_k)^T$ refers to the general mean and block effects. In this case the reduced information matrix (4), called also $\mathbf{C}$-matrix (see [18-20]), may be presented in the form

$$\begin{aligned} \mathbf{C} &= \mathbf{D}^T\mathbf{D} - \mathbf{D}^T diag(b_1^{-1}\mathbf{1}_{b_1}\mathbf{1}_{b_1}^T, \ldots, b_k^{-1}\mathbf{1}_{b_k}\mathbf{1}_{b_k}^T)\mathbf{D} \\ &= diag(t_1, \ldots, t_v) - \mathbf{N}diag(b_1^{-1}, \ldots, b_k^{-1})\mathbf{N}^T, \end{aligned}$$

where $\mathbf{N} = (n_{ij})$ is the incidence matrix defined as $\mathbf{N} = \mathbf{D}\mathbf{B}^T$. It is clear that $\mathbf{N}\mathbf{1}_k = \mathbf{t}$ and $\mathbf{1}_v^T\mathbf{N} = \mathbf{b}^T$, where $\mathbf{t} = (t_1, \ldots, t_v)^T$ and $\mathbf{b} = (b_1, \ldots, b_k)^T$. A design $\mathbf{D}$ is said to be orthogonal if $\mathbf{N} = \frac{1}{n}\mathbf{t}\mathbf{b}^T$.

One can verify that

$$\mathbf{N}diag(b_1^{-1}, \ldots, b_k^{-1})\mathbf{N}^T = \begin{bmatrix} \sum_j \frac{n_{1j}^2}{b_j} & \sum_j \frac{n_{1j}n_{2j}}{b_j} & \cdots & \sum_j \frac{n_{1j}n_{vj}}{b_j} \\ \sum_j \frac{n_{2j}n_{1j}}{b_j} & \sum_j \frac{n_{2j}^2}{b_j} & \cdots & \sum_j \frac{n_{2j}n_{vj}}{b_j} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_j \frac{n_{vj}n_{1j}}{b_j} & \sum_j \frac{n_{vj}n_{2j}}{b_j} & \cdots & \sum_j \frac{n_{vj}^2}{b_j} \end{bmatrix}.$$

In particular, for the orthogonal design, $\mathbf{N}diag(b_1^{-1}, \ldots, b_k^{-1})\mathbf{N}^T = \frac{1}{n}\mathbf{t}\mathbf{t}^T$.

Denote by $\mathcal{D} = \mathcal{D}(\mathbf{t}; \mathbf{b})$ the class of all possible allocations of v treatments with replications $t_1, \ldots, t_v$ in k blocks of sizes $b_1, \ldots, b_k$ for $v, k \geq 2$. Such class contains or does not contain an orthogonal design. If it does then by Stępniak [8] this design is optimal in $\mathcal{D}$ w.r.t. invariant linear estimation, i.e. it is at least as good as any other design in the class.

It is natural to ask whether the orthogonal design is also optimal w.r.t. invariant quadratic estimation. In the light of the results presented in Section 3 we are strongly convinced that the answer is negative, but for formal reasons we are ready to provide a rigorous proof of this fact. By Corollary 3.8 we only need to show that for any incidence matrix $\mathbf{N} = (n_{ij})$ corresponding to the orthogonal design there exists an incidence matrix $\mathbf{M} = (m_{ij})$ such that $\mathbf{M}diag(b_1^{-1}, \ldots, b_k^{-1})\mathbf{M}^T \neq \mathbf{N}diag(b_1^{-1}, \ldots, b_k^{-1})\mathbf{N}^T$ Define

$$m_{ij} = \begin{cases} n_{ij} + 1, & \text{if } i = 1 \text{ and } j = 1, \text{ or } i = 2 \text{ and } j = 2, \\ n_{ij} - 1, & \text{if } i = 1 \text{ and } j = 2, \text{ or } i = 2 \text{ and } j = 1, \\ n_{ij}, & \text{otherwise.} \end{cases}$$

We note that $\mathbf{M}\mathbf{1}_k = \mathbf{N}\mathbf{1}_k$ and $\mathbf{1}_v^T\mathbf{M} = \mathbf{1}_v^T\mathbf{N}$. Therefore, the designs represented by $\mathbf{M}$ and $\mathbf{N}$ belong to the same class. To show the desired inequality we only need, for instance, to compare the left upper entries, say u_{11} and u_{11}^0, of the matrices $\mathbf{M}diag(b_1^{-1}, \ldots, b_k^{-1})\mathbf{M}^T$ and $\mathbf{N}diag(b_1^{-1}, \ldots, b_k^{-1})\mathbf{N}^T$.

Since $n_{ij} = \frac{1}{n}t_ib_j$ we have

$$\begin{aligned} u_{11} - u_{11}^0 &= \frac{m_{11}^2}{b_1} + \frac{m_{12}^2}{b_2} - (\frac{n_{11}^2}{b_1} + \frac{n_{12}^2}{b_2}) \\ &= \frac{(n_{11}+1)^2}{b_1} + \frac{(n_{12}-1)^2}{b_2} - (\frac{n_{11}^2}{b_1} + \frac{n_{12}^2}{b_2}) \\ &= 2(\frac{n_{11}}{b_1} - \frac{n_{12}}{b_2}) + \frac{1}{b_1} + \frac{1}{b_2} \end{aligned}$$

$$= \frac{2}{n}(\frac{t_1 b_1}{b_1} - \frac{t_1 b_2}{b_2}) + \frac{1}{b_1} + \frac{1}{b_2} = \frac{1}{b_1} + \frac{1}{b_2} > 0.$$

This leads to the following

Conclusion 4.1. *Any orthogonal block design is not optimal w.r.t. invariant quadratic estimation. Moreover, for any* $\mathbf{t} = (t_1, ..., t_v)^T$ *and* $\mathbf{b} = (b_1, ..., b_k)^T$ *there is no optimal design in the class* $\mathcal{D} = \mathcal{D}(\mathbf{t}; \mathbf{b})$.

By the way we have demonstrated that, with reference to the orthogonal block design, the meaning of the optimality w.r.t. linear estimation may be strengthened in the sense that the words "at least as good" may be replaced by "better".

Acknowledgement: This work was partially supported by the Centre for Innovation and Transfer of Natural Sciences and Engineering Knowledge.

References

[1] Ehrenfeld S., Complete class theorem in experimental design, Proc. Third Berkeley Symp. on Math. Statist. Probab., 1955, Vol. 1, 69-75.
[2] Heyer H., Order relations for linear models: A survey on recent developments, Statist. Papers, 2006, 47, 331-372.
[3] Kiefer J., Optimum experimental designs, J. Roy. Statist. Soc. Ser. B, 1959, 21, 272-319.
[4] Goel P.K., Ginebra J., When one experiment is 'always better than' another?, Statistician, 2003, 52, 515-537.
[5] Shaked M., Suarez-Llorens A., On the comparison of reliability experiments based on convolution order, J. Amer. Statist. Assoc., 2003, 98, 693-702.
[6] Torgersen E., Comparison of Statistical Experiments. Cambridge University Press, Cambridge, England, 1991.
[7] Stępniak C., Comparison of normal linear experiments by quadratic forms, Ann. Inst. Statist. Math. A, 1997, 49, 569-584.
[8] Stępniak C., Optimal allocation of treatments in block designs, Studia Sc. Math. Hung., 1987, 22, 341-345.
[9] Stępniak C., Optimal allocation of units in experimental designs with hierarchical and cross classification, Ann. Inst. Statist. Math. A, 1983, 35, 461-473.
[10] Sinha B.K., Comparison of some experiments from sufficiency consideration, Ann. Inst. Statist. Math. A, 1973, 25, 501-520.
[11] Stępniak C., Stochastic ordering and Schur-convex functions in comparison of linear experiments, Metrika, 1989, 36, 291-298.
[12] Hansen O.H., Torgersen E., Comparison of linear experiments, Ann. Statist., 1974, 2, 265-373.
[13] Lehmann E.L., Comparing location experiments, Ann. Statist., 1988, 16, 521-533.
[14] Stępniak C., On admissibility in estimating the mean squared error of a linear estimator, Probab. Math. Statist., 1998, 18, 33-38.
[15] Stępniak C., Ordering of nonnegative matrices with application to comparison of linear models, Linear Algebra Appl., 1985, 70, 67-71.
[16] Mathai A.M., Provost S.B., Quadratic Forms in Random Variables, Marcel Dekker, New York 1992.
[17] Searle S.R., Linear Models, Wiley, New York, 1971.
[18] Chakrabarti M.C., On the C-matrix in design of experiments, J. Indian Statist. Assoc., 1963, 1, 8-23.
[19] Dey A., Theory of Block Designs, Wiley, New York, 1986.
[20] Raghavarao D., Padgett L.V., Block Designs: Analysis, Combinatorics and Applications, World Scientific Publishers, Singapore, 2005.