Journal of Statistical Planning and Inference 37 (1993) 51-68
North-Holland

Efficiency comparisons in linear inference

D.R. Jensen and D.E. Ramirez

Virginia Polytechnic Institute and State University and University of Virginia, Blacksburg, VA, USA

Received 11 February 1991; revised manuscript received 27 August 1992

Abstract: A new approach to comparative design efficiency is given in linear estimation and in tests for linear hypotheses. In contrast to others, this approach gives a complete assessment of two designs for any model on identifying precisely the subspaces of parameters for which one design is more efficient than another, or two designs are equally efficient. Local and global bounds on relative design efficiencies are given with reference to these subspaces, and connections to information functionals are noted. The relative influences of design points with regard to deletion and augmentation are examined and related to measures of influence from the literature. Our approach provides numerical diagnostics for use in design evaluation before an experiment has been run. The concepts are illustrated with reference to selected second-order designs.

AMS Subject Classification: 62K05, 62J05.

Key words: Design efficiency; Fisher, Pitman and information efficiencies; bounds; measures of influence; deletion and augmentation.

1. Introduction

Designed experiments and linear models persist at the forefront of statistical practice. Designs traditionally have been compared using criteria in estimation such as A, D, E or G efficiency (cf. Kiefer, 1959, or Fedorov, 1972, Section 1.8, for example), and more recently, the information functionals of Pukelsheim (1980). These and other criteria share the following limitations. (i) All are scalar indices, whereas the problem in intrinsically multidimensional. (ii) Efficient designs are typically model and criterion dependent, and generally there are no globally optimal designs. (iii) The criteria focus exclusively on estimation, with no thought towards efficiency in testing hypotheses. Specifically, it is not clear how design efficiency in estimation may bear on efficiency in hypothesis testing, although the two are clearly related. In response to (ii), Kiefer (1975) and others have stressed the importance of comparing designs on the basis of several different criteria and choosing one with good

Correspondence to: D.R. Jensen, Dept. of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061, USA.

overall performance. See also Galil and Kiefer (1977a, b, c, d) and Galil and Kiefer (1979).

On information-theoretic grounds consider a model $\{Y_i = f(x_i)'\boldsymbol{\beta} + \varepsilon_i;\ 1 \leqslant i \leqslant n\}$ of full rank together with linear parametric functions $A\boldsymbol{\beta}$ such that A $(r \times k)$ has rank $r \leqslant k$. For a set $\mathscr{H}$ consisting of prospective design points, consider some measurable mapping $f: \mathscr{H} \to \mathbb{R}^k$ such that the image $f(\mathscr{H})$ is compact, and let Ξ comprise the design measures on $\mathscr{H}$, i.e., the probability measures $\xi(\cdot)$ on $\mathscr{H}$ having finite support. Then under Gaussian errors $\mathscr{I}(\Xi)$ denotes the collection of all information matrices $I(\xi) = \int_{\mathscr{H}} f(x)f(x)'\,\mathrm{d}\xi$ with $\xi \in \Xi$. In particular, the information matrix for $A\boldsymbol{\beta}$ is $I(A;\xi) = [A(I(\xi))^{-1}A']^{-1}/\sigma^2$, and the design ξ_0 is said to be *uniformly optimal* for $A\boldsymbol{\beta}$ in Ξ if $I(A;\xi_0) - I(A;\xi)$ is positive semidefinite for every $\xi \in \Xi$; see Kurotschka (1978) and Pukelsheim (1980). However, Pukelsheim (1980) has shown in Theorem 7 that no design is uniformly optimal for $A\boldsymbol{\beta}$ in Ξ unless $r = 1$. In this case A is a vector, say $\boldsymbol{c}'$, and the optimal design is then called *c-optimal.* Characterizations of c-optimal designs are given in Pukelsheim (1981).

Other approaches to comparing designs have been considered. Kiefer (1975) focused primarily on scalar efficiency ratios using L_p-norms for $p = 0$, 1 or ∞. He noted, however, that the "finer comparison between two designs ξ' and ξ'' obtained by contrasting the entries of $M(\xi')$ and $M(\xi'')$, or the eigenvalues $\{\lambda_i(\xi')\}$ and $\{\lambda_i(\xi'')\}$, or by computing the eigenvalues of $M(\xi')M^{-1}(\xi'')$, does not, at least in our examples, seem worth the much greater effort required." Here $M(\xi)$ is Kiefer's notation for the information matrix $I(\xi)$.

An approach to multiparameter estimation using axiomatics and invariance is given in Jensen (1991). Suppose that $\boldsymbol{S}$ and $\boldsymbol{T}$ are unbiased for $\boldsymbol{\theta} \in \mathbb{R}^k$ having dispersion matrices $(\boldsymbol{\Sigma}, \boldsymbol{\Omega})$. If the efficiency $E(\boldsymbol{T}, \boldsymbol{S})$ of $\boldsymbol{T}$ relative to $\boldsymbol{S}$ is to depend only on second moments, and is also to be invariant to choise of basis, i.e., under $(\boldsymbol{S}, \boldsymbol{T}) \to (\boldsymbol{GS}, \boldsymbol{GT})$ for $\boldsymbol{G}$ nonsingular, then $E(\boldsymbol{T}, \boldsymbol{S})$ necessarily depends only on $(\boldsymbol{\Sigma}, \boldsymbol{\Omega})$, and then only through a maximal invariant, namely, the ordered roots of the determinantal equation $|\boldsymbol{\Sigma} - \gamma\boldsymbol{\Omega}| = 0$. This underscores that the concept of vector efficiency is intrinsically multidimensional. Moreover, by examining the eigenvalues in relation to 1.0, and subspaces spanned by the corresponding vectors, it is possible to compare the performance of $\boldsymbol{T}$ relative to $\boldsymbol{S}$ precisely as $\boldsymbol{\theta}$ ranges over its domain.

In this paper we extend these developments to include Pitman (1948) efficiency in hypothesis testing (cf. Kotz et al., 1985, pp. 731–735) as well as estimation in the context of linear models and designed experiments. This approach offers fresh insight into the detailed workings of alternative designs, and it essentially pursues the third approach mentioned above and then abandoned by Kiefer (1975) as intractable.

An outline of the paper follows. Preliminary developments comprise Section 2. Section 3 is concerned with the comparative Pitman and Fisher efficiencies of alternative designs and with connections of these to information functionals. Examples include particular second-order designs and a detailed numerical comparison of these. The basic tools are then used in Section 4 to examine the relative influence of design points and points of augmentation. These results in turn are related to measures of influence as developed in the literature, with the added feature that specific effects on linear parametric functions now can be identified precisely. Section 5 provides a brief summary.

2. Preliminaries

We establish conventions for notation and details concerning the models to be treated subsequently.

2.1. Notation

To fix notation, $\mathbb{R}^k$ and $\mathbb{R}_+^k$ denote Euclidean k-space and its positive orthant; $\boldsymbol{x}\in\mathbb{R}^k$ is of order $(k\times 1)$; $F_{n\times k}$ consists of real $(n\times k)$ matrices; and S_k, S_k^0 and S_k^+ comprise the real symmetric $(k\times k)$ matrices, their positive semidefinite and positive definite varieties, respectively. Special arrays include the unit matrix $\boldsymbol{I}_k$ and the symmetric root $\boldsymbol{A}^{1/2}\in S_k^+$ of $\boldsymbol{A}\in S_k^+$. By $\mathrm{Sp}(z_1,\dots,z_r)$ is meant the linear span of $\{z_1,\dots,z_r\}$ in $\mathbb{R}^k$, and we indicate by $\boldsymbol{B}\in\mathrm{Sp}(z_1,\dots,z_r)$ that its columns belong to $\mathrm{Sp}(z_1,\dots,z_r)$. Given a basis $\{\boldsymbol{w}_1,\dots,\boldsymbol{w}_t,\boldsymbol{w}_{t+1},\dots,\boldsymbol{w}_k\}$ of $\mathbb{R}^k$, we say that $\boldsymbol{x}$ is in the *complement* of $\mathrm{Sp}(\boldsymbol{w}_1,\dots,\boldsymbol{w}_t)$ when $\boldsymbol{x}\in\mathrm{Sp}(\boldsymbol{w}_{t+1},\dots,\boldsymbol{w}_k)$. Following Loewner (1934), we take $(S_k,\succcurlyeq)$ to be partially ordered such that $\boldsymbol{A}\succcurlyeq\boldsymbol{B}$ on S_k if and only if $\boldsymbol{A}-\boldsymbol{B}\in S_k^0$, with $\boldsymbol{A}\succ\boldsymbol{B}$ when $\boldsymbol{A}-\boldsymbol{B}$ is positive definite.

2.2. Linear models

Given a model $\{Y_i=f(\boldsymbol{x}_i)'\boldsymbol{\beta}+\varepsilon_i;\ 1\leqslant i\leqslant n\}$ of full rank with design matrix $\boldsymbol{X}=[\boldsymbol{x}_1,\dots,\boldsymbol{x}_n]'$, we write $\boldsymbol{Y}=f(\boldsymbol{X})\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ with $f(\boldsymbol{X})=[f(\boldsymbol{x}_1),\dots,f(\boldsymbol{x}_n)]'$ of order $(n\times k)$. The standard Gauss-Markov estimator $\hat{\boldsymbol{\beta}}(\boldsymbol{X})=[f(\boldsymbol{X})'f(\boldsymbol{X})]^{-1}f(\boldsymbol{X})'\boldsymbol{Y}$ is unbiased with dispersion matrix $V[\hat{\boldsymbol{\beta}}(\boldsymbol{X})]=\sigma^2[f(\boldsymbol{X})'f(\boldsymbol{X})]^{-1}$, and under Gaussian errors with information matrix $\boldsymbol{I}(\boldsymbol{X})=[f(\boldsymbol{X})'f(\boldsymbol{X})]/\sigma^2$ for some $\sigma^2>0$. Given designs $\boldsymbol{X}$ and $\boldsymbol{Z}$ and the model $f(\cdot)$, except as noted we henceforth identify $\boldsymbol{\Sigma}=[f(\boldsymbol{X})'f(\boldsymbol{X})]^{-1}$ and $\boldsymbol{\Omega}=[f(\boldsymbol{Z})'f(\boldsymbol{Z})]^{-1}$. The following analysis takes a prominent role throughout.

Let $\{\gamma_1\geqslant\cdots\geqslant\gamma_k>0\}$ be the ordered roots of $|\boldsymbol{\Sigma}-\gamma\boldsymbol{\Omega}|=0$ or, equivalently, of $\boldsymbol{\Omega}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Omega}^{-1/2}$, with $(\boldsymbol{\Sigma},\boldsymbol{\Omega})$ in S_k^+ as defined. With $\boldsymbol{K}=\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Omega}^{-1/2}$, observe that $\boldsymbol{K}'\boldsymbol{K}$ and $\boldsymbol{K}\boldsymbol{K}'$ have the same ordered eigenvalues $\{\gamma_1\geqslant\cdots\geqslant\gamma_k>0\}$, whereas their respective eigenvectors, $\{\boldsymbol{u}_1,\dots,\boldsymbol{u}_k\}$ and $\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\}$, constitute orthonormal bases by symmetry. Moreover, with $\boldsymbol{K}'\boldsymbol{K}\boldsymbol{u}=\gamma\boldsymbol{u}$, so that $\boldsymbol{K}\boldsymbol{K}'\boldsymbol{K}\boldsymbol{u}=\gamma\boldsymbol{K}\boldsymbol{u}$ and thus $\boldsymbol{K}\boldsymbol{K}'\boldsymbol{v}=\gamma\boldsymbol{v}$, we find that

$$\{\boldsymbol{v}_i=\gamma_i^{-1/2}\boldsymbol{K}\boldsymbol{u}_i,\ 1\leqslant i\leqslant k\} \tag{2.1}$$

so that their eigenvectors are related linearly. In particular, we have $\{\boldsymbol{\Sigma}^{-1/2}\boldsymbol{v}_i=\gamma_i^{-1/2}\boldsymbol{\Omega}^{-1/2}\boldsymbol{u}_i,\ 1\leqslant i\leqslant k\}$.

3. Design efficiencies

We seek a unified treatment of comparative design efficiencies in linear estimation and hypothesis testing. Basic connections are given next between the directed Fisher efficiency (3.1) and the Pitman efficiency (3.2) in linear inference. This extends work begun in Jensen (1991), but now including hypothesis tests as well as estimation.

3.1. Efficiencies

We first consider the Fisher efficiency in estimating $\theta = \boldsymbol{a}'\boldsymbol{\beta}$ in the model $Y = f(X)\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. If (S, T) are unbiased for $\theta \in \mathbb{R}^1$ having variances (σ_S^2, σ_T^2), respectively, then the Fisher efficiency of T relative to S is given by $E(T, S) = \sigma_S^2/\sigma_T^2$. For estimating $\boldsymbol{a}'\boldsymbol{\beta} \in \mathbb{R}^1$ in the direction in $\mathbb{R}^k$ whose cosines are determined by $\boldsymbol{a}' = [a_1, \ldots, a_k]$ with $\boldsymbol{a}'\boldsymbol{a} = 1$, we now take the *directed ϕ-efficiency* of design $\boldsymbol{Z}$ relative to design X to be $E_\phi(\boldsymbol{Z}, X; \boldsymbol{a}) = \phi(\boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a}/\boldsymbol{a}'\boldsymbol{\Omega}\boldsymbol{a})$ on $\mathbb{R}_+^1$, with ϕ belonging to a class $\boldsymbol{\Phi}$ of efficiency indices considered in Jensen (1991). Because $\phi(\cdot)$ is monotonic, it suffices to consider the directed Fisher efficiency $E[\boldsymbol{a}'\hat{\boldsymbol{\beta}}(Z), \boldsymbol{a}'\hat{\boldsymbol{\beta}}(X)] = \phi^{-1} \circ E_\phi(Z, X; \boldsymbol{a})$ as given in

$$E[\boldsymbol{a}'\hat{\boldsymbol{\beta}}(Z), \boldsymbol{a}'\hat{\boldsymbol{\beta}}(X)] = \frac{\boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a}}{\boldsymbol{a}'\boldsymbol{\Omega}\boldsymbol{a}} = \frac{\boldsymbol{b}'\boldsymbol{\Omega}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Omega}^{-1/2}\boldsymbol{b}}{\boldsymbol{b}'\boldsymbol{b}}, \tag{3.1}$$

with $\boldsymbol{b} = \boldsymbol{\Omega}^{1/2}\boldsymbol{a}$, where $\boldsymbol{\Omega}^{1/2}$ is the symmetric root of $\boldsymbol{\Omega}$.

Next consider the standard F-test for H: $\boldsymbol{A}\boldsymbol{\beta} = \boldsymbol{\delta}_0$ against K: $\boldsymbol{A}\boldsymbol{\beta} \neq \boldsymbol{\delta}_0$ under designs X and $\boldsymbol{Z}$. The Pitman (1948) efficiency of design $\boldsymbol{Z}$ relative to design X in testing H is given as the ratio $E(Z, X \mid A) = \lambda_A(Z)/\lambda_A(X)$ of their noncentrality parameters with $\lambda_A(X) = (\boldsymbol{A}\boldsymbol{\beta} - \boldsymbol{\delta}_0)'(\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}')^{-1}(\boldsymbol{A}\boldsymbol{\beta} - \boldsymbol{\delta}_0)/\sigma^2$, i.e.,

$$E(Z, X \mid A) = \frac{(\boldsymbol{A}\boldsymbol{\beta} - \boldsymbol{\delta}_0)'(\boldsymbol{A}\boldsymbol{\Omega}\boldsymbol{A}')^{-1}(\boldsymbol{A}\boldsymbol{\beta} - \boldsymbol{\delta}_0)}{(\boldsymbol{A}\boldsymbol{\beta} - \boldsymbol{\delta}_0)'(\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}')^{-1}(\boldsymbol{A}\boldsymbol{\beta} - \boldsymbol{\delta}_0)}. \tag{3.2}$$

With A set to $A = \boldsymbol{I}_k$ and $\boldsymbol{A}\boldsymbol{\beta} - \boldsymbol{\delta}_0 = \boldsymbol{c}$, (3.2) assumes the form

$$E(Z, X \mid \boldsymbol{I}_k) = \frac{\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{c}}{\boldsymbol{c}'\boldsymbol{\Sigma}^{-1}\boldsymbol{c}} = \frac{\boldsymbol{d}'\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Omega}^{-1}\boldsymbol{\Sigma}^{1/2}\boldsymbol{d}}{\boldsymbol{d}'\boldsymbol{d}}, \tag{3.3}$$

with $\boldsymbol{d} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{c}$. Connections between Fisher and Pitman efficiencies follow directly since both (3.1) and (3.2) are *generalized Rayleigh quotients* having standard properties as given in Bellman (1960), for example. In subsequent developments we assemble subspaces of $\mathbb{R}^k$ in natural coordinates of the parameters using vectors $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$ corresponding respectively to the roots $\{\gamma_1, \ldots, \gamma_k\}$ of $|\boldsymbol{\Sigma} - \gamma\boldsymbol{\Omega}| = 0$.

3.2. Basic results

Subsequently $E(Z, X)$ denotes the efficiency of design Z relative to design X with reference to a specified model. This applies both to the directed Fisher efficiency

$E(Z,X;\boldsymbol{a})=E[\boldsymbol{a}'\hat{\boldsymbol{\beta}}(Z),\boldsymbol{a}'\hat{\boldsymbol{\beta}}(X)]$ as in (3.1), and to the Pitman efficiency $E(Z,X\mid A)$ as in (3.2) in testing H: $A\boldsymbol{\beta}=\boldsymbol{\delta}_0$ against K: $A\boldsymbol{\beta}\neq\boldsymbol{\delta}_0$.

We next undertake a systematic assessment of these efficiencies as they depend on designs X and Z. In particular cases the ratios (3.1) and (3.2) can be evaluated numerically, but it is of further interest to examine their behavior over a range of values of the parameters. Clearly $\{\gamma_1\geqslant\cdots\geqslant\gamma_k>1\}$ if and only if $[f(Z)'f(Z)]>[f(X)'f(X)]$, in which case the ratios (3.1) and (3.2) both exceed unity. Design Z is then uniformly more efficient for estimating $\boldsymbol{\beta}$ and $\{\boldsymbol{a}'\boldsymbol{\beta};\ \boldsymbol{a}\in\mathbb{R}^k\}$, and the variance ratio test for H: $A\boldsymbol{\beta}=\boldsymbol{\delta}_0$ is uniformly more powerful for all $\{A\in F_{r\times k};\ 1\leqslant r\leqslant k\}$ at all alternatives K: $A\boldsymbol{\beta}\neq\boldsymbol{\delta}_0$. However, quantifying the precise gains in efficiency is a local phenomenon to be studied subsequently.

Generally neither design dominates globally, in which case the ratios (3.1) and (3.2) may be less than unity for some cases, possibly equal to unity for others, and greater than unity for still other cases. It is thus of interest to examine the comparative efficiencies locally. To these ends we associate with $\{\gamma_1,\ldots,\gamma_k\}$ the respective vectors $\{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_k\}$ spanning the natural parameter space, with

$$\{\boldsymbol{w}_i=\boldsymbol{\Omega}^{-1/2}\boldsymbol{u}_i=\gamma_i^{1/2}\boldsymbol{\Sigma}^{-1/2}\boldsymbol{v}_i;\ 1\leqslant i\leqslant k\} \tag{3.4}$$

as in (2.1). Now choose a subset $\{\gamma_{i_1},\ldots,\gamma_{i_t}\}$, together with the corresponding spanning vectors $\{\boldsymbol{w}_{i_1},\ldots,\boldsymbol{w}_{i_t}\}$, such that $\{\gamma_{i_1}\geqslant\cdots\geqslant\gamma_{i_t}\}$, and observe that $S(i_1,\ldots,i_t)\equiv\mathrm{Sp}(\boldsymbol{w}_{i_1},\ldots,\boldsymbol{w}_{i_t})\subset\mathbb{R}^k$ is a t-dimensional subspace of $\mathbb{R}^k$. A basic result pertaining to both local and global efficiencies in estimation is the following.

Theorem 1. *For directed Fisher efficiency, local bounds on the efficiency of design Z relative to design X for a specified model are given by*

(i) $\gamma_{i_t}\leqslant E(Z,X;\boldsymbol{a})\leqslant\gamma_{i_1}$ *uniformly for all* $\boldsymbol{a}\in S(i_1,\ldots,i_t)$, *and in particular,*

(ii) $\gamma_k\leqslant E(Z,X;\boldsymbol{a})\leqslant\gamma_1$ *uniformly for all* $\boldsymbol{a}\in\mathbb{R}^k$.

Moreover, all bounds are sharp in that equality can be attained.

Proof. To see conclusion (i), write $E(Z,X;\boldsymbol{a})=\boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a}/\boldsymbol{a}'\boldsymbol{\Omega}\boldsymbol{a}=\boldsymbol{b}'\boldsymbol{\Omega}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Omega}^{-1/2}\boldsymbol{b}$ as above with $\boldsymbol{b}'\boldsymbol{b}=1$. Then $\boldsymbol{a}\in S(i_1,\ldots,i_t)$ is equivalent to $\boldsymbol{b}\in\mathrm{Sp}(\boldsymbol{u}_{i_1},\ldots,\boldsymbol{u}_{i_t})$ from (2.1). It follows that $\boldsymbol{b}=c_1\boldsymbol{u}_{i_1}+\cdots+c_t\boldsymbol{u}_{i_t}$ for some $\boldsymbol{c}=[c_1,\ldots,c_t]'$, where $\boldsymbol{b}'\boldsymbol{b}=1$ and $\boldsymbol{c}'\boldsymbol{c}=1$ are equivalent. Moreover, the orthonormality of $\{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_k\}$, together with the spectral decomposition $\boldsymbol{\Omega}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Omega}^{-1/2}=\sum_{j=1}^k\gamma_j\boldsymbol{u}_j\boldsymbol{u}_j'$, show that

$$\begin{aligned}\boldsymbol{b}'\boldsymbol{\Omega}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Omega}^{-1/2}\boldsymbol{b}&=(c_1\boldsymbol{u}_{i_1}+\cdots+c_t\boldsymbol{u}_{i_t})\left(\sum_{j=1}^k\gamma_j\boldsymbol{u}_j\boldsymbol{u}_j'\right)(c_1\boldsymbol{u}_{i_1}+\cdots+c_t\boldsymbol{u}_{i_t})\\&=\sum_{j=1}^t c_j^2\gamma_{i_j}.\end{aligned} \tag{3.5}$$

It follows directly that

$$\gamma_{i_t}\leqslant\sum_{j=1}^t c_j^2\gamma_{i_j}\leqslant\gamma_{i_1} \tag{3.6}$$

where equality is attained on choosing c suitably, first as $c=[1\ 0\ \cdots\ 0]'$ to give equality on the right, then as $c=[0\ \cdots\ 0\ 1]'$ giving equality on the left. This establishes conclusion (i). Conclusion (ii) follows from conclusion (i) as a special case. □

We note that a version of Theorem 1 is given in Rao (1973), p. 62. We have given our proof for completeness and in preparation for the next result, which pertains to both local and global Pitman efficiencies in testing linear hypotheses.

Theorem 2. *For Pitman efficiency in testing* H: $A\beta=\delta_0$ *against* K: $A\beta\neq\delta_0$, *local bounds on the efficiency of design* Z *relative to design* X *for a specified model are given by*

(i) $\gamma_{i_t}\leqslant E(Z,X\mid A)\leqslant\gamma_{i_1}$ *uniformly for all* $A'\in S(i_1,\dots,i_t)$, *and in particular*
(ii) $\gamma_k\leqslant E(Z,X\mid A)\leqslant\gamma_1$, *uniformly for all* $\{A\in F_{r\times k};\ 1\leqslant r\leqslant k\}$.

Proof. From the spectral decomposition $\Sigma^{1/2}\Omega^{-1}\Sigma^{1/2}=\sum_{j=1}^{k}\gamma_j v_j v_j'$ we have that $\Omega=\sum_{j=1}^{k}\gamma_j^{-1}\Sigma^{1/2}v_j v_j'\Sigma^{1/2}$ and $\Sigma=\Sigma^{1/2}\sum_{j=1}^{k}v_j v_j'\Sigma^{1/2}$. Given that the rows of A belong to $S(i_1,\dots,i_t)$ by assumption, we may represent A as $A=CV\Sigma^{-1/2}$, where C $(t\times t)$ is nonsingular and $V'=[v_{i_1}\cdots v_{i_t}]$ is of order $(k\times t)$. If follows that

$$\begin{aligned}A\Omega A'&=CV\Sigma^{-1/2}\left(\sum_{j=1}^{k}\gamma_j^{-1}\Sigma^{1/2}v_j v_j'\Sigma^{1/2}\right)\Sigma^{-1/2}V'C'\\&=CV\left(\sum_{j=1}^{k}\gamma_j^{-1}v_j v_j'\right)V'C'=CD_tC'\end{aligned}\tag{3.7}$$

with $D_t=\mathrm{Diag}(\gamma_{i_1}^{-1},\dots,\gamma_{i_t}^{-1})$. The ratio $E(Z,X\mid A)$ now becomes

$$E(Z,X\mid A)=\frac{d'(A\Omega A')^{-1}d}{d'(A\Sigma A')^{-1}d}=\frac{d'(CD_tC')^{-1}d}{d'(CC')^{-1}d}=\frac{e'D_t^{-1}e}{e'e}\tag{3.8}$$

because C is invertible, and finally $E(Z,X\mid A)=\sum_{j=1}^{t}g_j^2\gamma_{i_j}$ with $g=e/\|e\|$. The proof is now completed using the final steps in the proof for Theorem 1. □

Suppose that neither design dominates. It is then of interest to be able to characterize those linear hypotheses for which design Z has efficiency greater than, or equal to, or less than that of design X. Such characterization may point to one design as being more informative than another, and thus preferable, on issues critical to a particular investigation. This is done next in Corollary 1 for Pitman efficiency based on Theorem 2. Corresponding results in estimation are given in Jensen (1991). To these ends suppose that neither $\Sigma\not\geqslant\Omega$ nor $\Omega\not\geqslant\Sigma$, in which case integers (r,s) can be found, with $r>0$ and $s\geqslant 0$, such that

$$\{\gamma_1\geqslant\cdots\geqslant\gamma_r>\gamma_{r+1}=1=\cdots=\gamma_{r+s}>\gamma_{r+s+1}\geqslant\cdots\geqslant\gamma_k\}.\tag{3.9}$$

Accordingly, identify subspaces as $L_1=\mathrm{Sp}(w_1,\dots,w_r)$, $L_2=\mathrm{Sp}(w_{r+1},\dots,w_{r+s})$, and $L_3=\mathrm{Sp}(w_{r+s+1},\dots,w_k)$. In terms of these we have the following.

Corollary 1. *Among all linear hypotheses* H: $A\boldsymbol{\beta}=\boldsymbol{\delta}_0$ *with alternatives* K: $A\boldsymbol{\beta}\neq\boldsymbol{\delta}_0$, *the Pitman efficiencies of design Z relative to X for a given model satisfy*

(i) $E(Z,X\mid A)>1$ *for all* $A'\in L_1=\mathrm{Sp}(w_1,\ldots,w_r)$,
(ii) $E(Z,X\mid A)=1$ *for all* $A'\in L_2=\mathrm{Sp}(w_{r+1},\ldots,w_{r+s})$,
(iii) $E(Z,X\mid A)<1$ *for all* $A'\in L_3=\mathrm{Sp}(w_{r+s+1},\ldots,w_k)$.

Observe that Corollary 1 holds generally without the constraints of (3.9) if at most two of L_1, L_2 and L_3 are allowed to be empty. For example, the case $s=0$ renders L_2.as an empty set. Nonetheless, Corollary 1 still holds if conclusion (ii) is removed as a vacuous statement.

Following Pukelsheim (1980), we next consider the class J of information functionals j on S_k^0 such that j is (i) nonnegative on S_k^0 and positive on S_k^+, (ii) positively homogeneous, and (iii) super-additive, i.e., $j(A+B)\geqslant j(A)+j(B)$. Let $\mathscr{I}_0$ be a compact convex subset of S_k^0, any member of which will be called an *information matrix*. Theorem 1 of Pukelsheim (1980) shows for every information matrix $\boldsymbol{I}\in\mathscr{I}_0\cap S_k^+$ that the following three statements are equivalent: (i) $\boldsymbol{I}$ has $\mathscr{I}_0$-maximal j-information for $A\boldsymbol{\beta}$ for all information functionals $j\in J$; (ii) $[A\boldsymbol{I}^{-1}A']^{-1}\succcurlyeq[A\boldsymbol{T}^{-1}A']^{-1}$ for all $\boldsymbol{T}\in\mathscr{I}_0$; and (iii) $\boldsymbol{I}$ has $\mathscr{I}_0$-maximal information for $\boldsymbol{c}'\boldsymbol{\beta}$ for all $\boldsymbol{c}$ in the range of A'.

A basic connection between Pitman efficiency and information is that $\boldsymbol{I}(A;X)=\{A[f(X)'f(X)]^{-1}A'\}^{-1}/\sigma^2$. Here $\boldsymbol{I}(A;X)$ is interpreted as the amount of information about $A\boldsymbol{\beta}$ contained in X. Corollary 1 now may be restated in terms of information functionals using Pukelsheim's (1980) result as follows.

Corollary 2. *Among all linear parametric functions* $A\boldsymbol{\beta}$ *the following normal-theory information inequalities hold*:

(i) *Design Z has greater j-information than X, for all* $A\boldsymbol{\beta}$ *with* $A'\in L_1=\mathrm{Sp}(w_1,\ldots,w_r)$, *and for all* $j\in J$,

(ii) *Designs X and Z have equivalent j-information for all* $A\boldsymbol{\beta}$ *with* $A'\in L_2=\mathrm{Sp}(w_{r+1},\ldots,w_{r+s})$, *and for all* $j\in J$,

(iii) *Design X has greater j-information than Z, for all* $A\boldsymbol{\beta}$ *with* $A'\in L_3=\mathrm{Sp}(w_{r+s+1},\ldots,w_k)$, *and for all* $j\in J$.

The foregoing developments support the detailed comparison of any pair of designs. Implicit are further comparisons between a specified design Z and members of an ensemble of designs. To these ends let $\boldsymbol{\Sigma}=\boldsymbol{\Sigma}(X)=[f(X)'f(X)]^{-1}$ to emphasize its dependence on X. Details follow.

Let $\mathscr{H}_0$ be a bounded ensemble of designs for which there are elements $\{X_m,X_M\}$ such that $\boldsymbol{\Sigma}_m\equiv\boldsymbol{\Sigma}(X_m)\preccurlyeq\boldsymbol{\Sigma}(X)\preccurlyeq\boldsymbol{\Sigma}(X_M)\equiv\boldsymbol{\Sigma}_M$ for every $X\in\mathscr{H}_0$, and let Z be a fixed design of the same order with $[f(Z)'f(Z)]^{-1}=\boldsymbol{\Omega}$ as before. Since $\boldsymbol{\Sigma}(X)\succcurlyeq\boldsymbol{\Sigma}_m$, it follows that $\boldsymbol{\Omega}^{-1/2}\boldsymbol{\Sigma}(X)\boldsymbol{\Omega}^{-1/2}\succcurlyeq\boldsymbol{\Omega}^{-1/2}\boldsymbol{\Sigma}_m\boldsymbol{\Omega}^{-1/2}$ for every $X\in\mathscr{H}_0$. If $\{\gamma_{m1}\geqslant\cdots\geqslant\gamma_{mk}\}$ are the ordered roots of $|\boldsymbol{\Sigma}_m-\gamma_{mi}\boldsymbol{\Omega}|=0$, it then follows that $\gamma_{mk}\leqslant E(Z,X)$ for both Fisher and Pitman efficiencies, uniformly for all $X\in\mathscr{H}_0$.

Table 1
Design points in the (X_1,X_2)-plane for three designs

3^2 Factorial (X)		Rotated factorial (Z)		Roots of unity (T)	
−1.0	−1.0	0.00000	−1.41421	−1.00000	0.00000
−1.0	0.0	−0.70711	−0.70711	0.00000	−1.00000
−1.0	1.0	−1.41421	0.00000	0.00000	0.00000
0.0	−1.0	0.70711	−0.70711	0.00000	1.00000
0.0	0.0	0.00000	0.00000	1.00000	0.00000
0.0	1.0	−0.70711	0.70711	−0.70711	−0.70711
1.0	−1.0	1.41421	0.00000	0.70711	−0.70711
1.0	0.0	0.70711	0.70711	−0.70711	0.70711
1.0	1.0	0.00000	1.41421	0.70711	0.70711

Similarly, if $\{\gamma_{M1}\geqslant\cdots\geqslant\gamma_{Mk}\}$ are the ordered roots of $|\boldsymbol{\Sigma}_M-\gamma_{Mi}\boldsymbol{\Omega}|=0$, then the ordering $E(Z,X)\leqslant\gamma_{M1}$ holds for both Fisher and Pitman efficiencies, uniformly for all $X\in\mathscr{H}_0$. We have proved the following.

Theorem 3. *Let $\{\gamma_{mk},\gamma_{M1}\}$ be the smallest and largest roots of $|\boldsymbol{\Sigma}_m-\gamma_{mi}\boldsymbol{\Omega}|=0$ and $|\boldsymbol{\Sigma}_M-\gamma_{Mi}\boldsymbol{\Omega}|=0$, respectively.*

(i) *Bounds on the directed Fisher efficiency, given by $\gamma_{mk}\leqslant E(Z,X;\boldsymbol{a})\leqslant\gamma_{M1}$, hold uniformly for all $\boldsymbol{a}\in\mathbb{R}^k$ and for all $X\in\mathscr{H}_0$.*

(ii) *Bounds on the Pitman efficiency, given by $\gamma_{mk}\leqslant E(Z,X\mid A)\leqslant\gamma_{M1}$, hold uniformly for all $\{A\in F_{r\times k};\ 1\leqslant r\leqslant k\}$, and for all $X\in\mathscr{H}_0$.*

3.3. Applications

To illustrate our methods consider the second-order response model

$$Y=\beta_0+\beta_1X_1+\beta_2X_2+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{12}X_1X_2+\varepsilon \tag{3.10}$$

together with three design matrices $\boldsymbol{X}$, $\boldsymbol{Z}$ and $\boldsymbol{T}$ given in Table 1 as points in the (X_1,X_2)-plane. The first is the standard 3^2 factorial design with coded values $\{-1,0,1\}\times\{-1,0,1\}$ in the (X_1,X_2)-plane; the second is obtained from the first by rotating points in the (X_1,X_2)-plane clockwise through 45 degrees; and the third consists of eight roots of unity in the (X_1,X_2)-plane plus the origin. The largest

Table 2
Eigenvalues, traces and determinants for the matrices $\boldsymbol{\Sigma}=[f(X)'f(X)]^{-1}$, $\boldsymbol{\Omega}=[f(Z)'f(Z)]^{-1}$ and $\boldsymbol{\Delta}=[f(T)'f(T)]^{-1}$ for the three designs of Table 1 with reference to the model (3.10)

Matrix	Eigenvalues	Trace	Determinant
$\boldsymbol{\Sigma}$	{1,1/2,1/4,1/6,1/6,1/18}	2.1389	0.0001929
$\boldsymbol{\Omega}$	{1,1,1/6,1/6,1/8,1/18}	2.5139	0.0001929
$\boldsymbol{\Delta}$	{3.17116,1,1/2,1/4,1/4,0.07884}	5.2500	0.0078129

Table 3
Diagonalization of $\boldsymbol{\Gamma}=\boldsymbol{\Omega}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Omega}^{-1/2}$ with $\boldsymbol{\Gamma}\boldsymbol{u}_i=\gamma_i\boldsymbol{u}_i$ and $\boldsymbol{W}=\boldsymbol{\Omega}^{-1/2}\boldsymbol{U}$, where $[\gamma_1,\ldots,\gamma_6]=[4, 1, 1, 1, 1, 1/4]'$

Columns of $\boldsymbol{W}$					
0.00000	2.92470	0.00000	0.00000	−0.66801	0.00000
0.00000	0.00000	0.00000	2.44949	0.00000	0.00000
0.00000	0.00000	2.44949	0.00000	0.00000	0.00000
−2.00000	2.17248	0.00000	0.00000	0.52955	0.00000
2.00000	2.17248	0.00000	0.00000	0.52955	0.00000
0.00000	0.00000	0.00000	0.00000	0.00000	1.00000

eigenvalue, trace and determinant are given in Table 2 for each of the matrices $\boldsymbol{\Sigma}=[f(X)'f(X)]^{-1}$, $\boldsymbol{\Omega}=[f(Z)'f(Z)]^{-1}$ and $\boldsymbol{\Delta}=[f(T)'f(T)]^{-1}$ of order (6×6) with reference to the model (3.10).

Conventional comparisons show that designs $\boldsymbol{X}$ and $\boldsymbol{Z}$ have comparable D and E efficiency, whereas $\boldsymbol{X}$ has greater A efficiency with smaller average variance of the Gauss–Markov estimators. On these grounds the 3^2 factorial design $\boldsymbol{X}$ ordinarily would be preferred to the rotated design $\boldsymbol{Z}$, a point to which we return subsequently. Further details are given in Table 3, including values for $\{\gamma_1,\ldots,\gamma_6\}$ together with the corresponding vectors $\{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_6\}$ spanning $\mathbb{R}^6$. Here we have transformed back to the natural coordinates of the original parameters using $\{\boldsymbol{w}_i=\boldsymbol{\Omega}^{-1/2}\boldsymbol{u}_i;\ 1\leqslant i\leqslant 6\}$ as in (2.1).

Applying Theorems 1 and 2 to account for both the directed Fisher efficiency in estimation and Pitman efficiency in hypothesis testing, we have the global bounds $0.25\leqslant E(Z,X)\leqslant 4$. The latter hold for all $\boldsymbol{a}\in\mathbb{R}^6$ in the case of estimation, and for all A $(r\times 6)$ in the case of hypothesis testing, with $1\leqslant r\leqslant 6$. Local bounds follow from Table 3 on interpreting the pairs $\{(\gamma_i,\boldsymbol{w}_i);\ 1\leqslant i\leqslant 6\}$. In particular, the pair $(\gamma_1,\boldsymbol{w}_1)$ implies that the rotated design $\boldsymbol{Z}$ has fourfold greater efficiency for the contrast $\beta_{22}-\beta_{11}$, whereas $(\gamma_6,\boldsymbol{w}_6)$ implies that the 3^2 factorial design $\boldsymbol{X}$ has fourfold greater efficiency for inferences regarding β_{12}.

We further infer that $1\leqslant E(Z,X)\leqslant 4$ for all inferences about the parameters *excluding* β_{12}, and that $\boldsymbol{Z}$ is at least as efficient as $\boldsymbol{X}$ for these. Similarly, $0.25\leqslant E(Z,X)\leqslant 1$ for all inferences in the space of parameters complementary to the contrast $\beta_{22}-\beta_{11}$. Finally, Corollary 1 shows the designs $\boldsymbol{X}$ and $\boldsymbol{Z}$ to be equally efficient for all linear inferences in the four-dimensional subspace complementary to the linear functions $\beta_{22}-\beta_{11}$ and β_{12}.

In terms of linear parametric functions it is now seen that our methods support a complete comparative assessment of two designs. With regard to choosing between $\boldsymbol{X}$ and $\boldsymbol{Z}$, the foregoing facts in turn point towards the rotated design whenever $\beta_{22}-\beta_{11}$ is of greater interest than β_{12}, despite the somewhat irrelevant but greater A efficiency of the factorial design.

Table 4 provides details regarding the relative efficiencies $E(T,X)$. The global bound $0.146\leqslant E(T,X)\leqslant 1.520$ applies directly. Moreover, the design $\boldsymbol{T}$ is more efficient

Table 4
Diagonalization of $\Gamma = \Delta^{-1/2}\Sigma\Delta^{-1/2}$ with $\Gamma u_i = \gamma_i u_i$ and $W = \Delta^{-1/2}U$, where $[\gamma_1, \ldots, \gamma_6] = [1.52053, 1.00000, 0.66667, 0.66667, 0.25000, 0.141615]'$

Columns of W					
−2.91580	0.00000	0.00000	0.00000	0.00000	0.70580
0.00000	0.00000	2.00000	0.00000	0.00000	0.00000
0.00000	0.00000	0.00000	−2.00000	0.00000	0.00000
−1.18500	−1.00000	0.00000	0.00000	0.00000	0.77186
−1.18500	1.00000	0.00000	0.00000	0.00000	0.77186
0.00000	0.00000	0.00000	0.00000	1.00000	0.00000

than X by a factor of 1.520 for inferences regarding $2.916\beta_0 + 1.185(\beta_{11} + \beta_{22})$; T and X are equally efficient for inferences regarding $\beta_{22} - \beta_{11}$; and T is less efficient elsewhere. It is seen that T is least efficient with regard to $0.706\beta_0 + 0.772(\beta_{11} + \beta_{22})$, where $E(T, X) = 0.146$. Comparisons between T and Z gave nearly identical results, the exception being that T and Z are equally efficient for β_{12}. For this case the matrix W containing spanning vectors is identical to the matrix W given in Table 4 except that columns 2 and 5 are interchanged. Further details are omitted.

The foregoing conclusions are both model and design-dependent. Other comparisons among these designs have been made for an altered version of the model (3.10). If the pure quadratic terms are dropped, whereby $\beta_{11} = \beta_{22} = 0$, then the 3^2 factorial design X dominates the corresponding Z and T designs uniformly, and thus is preferred by an efficiency factor ranging from 1 to 4. In particular, we find that $0.25 \leqslant E(Z, X) \leqslant 1$, where X has fourfold greater efficiency for inferences regarding β_{12}, with equal efficiency elsewhere. Design X now surpasses design T except for inferences about β_0, where they are equally efficient. Efficiencies $E(T, Z)$ are bounded globally by $0.667 \leqslant E(T, Z) \leqslant 1$; T and Z are equally efficient for inferences regarding $\{\beta_0, \beta_1\}$, whereas T is less efficient for $\{\beta_2, \beta_{12}\}$.

4. Influence and augmentation

4.1. Background

Standard diagnostics in regression examine the influence of individual points with regard to estimation or the prediction variance at those points. Design augmentation based on prediction variance is used by Wynn (1970) to obtain convergence to a D-optimal design by adding a new observation where the prediction variance is greatest. Measures of influence of subsets of observations are considered by Ghosh (1989) and Takeuchi (1991) in design evaluation, and by Mukerjee and Kageyama (1990) in the study of robustness of group divisible designs.

Here we apply results of earlier sections as tools for studying effects of deleting subsets of observations, and in choosing among candidates in design augmentation.

To illustrate we augment by replicating points of the original design, but the methods apply equally to any other points of relevance to a particular experiment. Our work transcends earlier studies in tracking precisely the effects of specific deletions and augmentations as they alter the quality of inferences regarding linear parametric functions which then may be identified explicitly.

4.2. Basic results

To fix notation let $\{i_1,\dots,i_s\}$ be a subset of $\{1,2,\dots,n\}$. Then $X(i_1,\dots,i_s)$ is the result of deleting rows $\{i_1,\dots,i_s\}$ from X; $X_2(i_1,\dots,i_s)$ consists of the deleted rows; and $X[i_1,\dots,i_s]$ is a matrix with n + s rows obtained on appending $X_2(i_1,\dots,i_s)$ to the original design as $X[i_1,\dots,i_s]=[X',(X_2(i_1,\dots,i_s))']'$. Referring to the model $Y=f(X)\boldsymbol{\beta}=\boldsymbol{\varepsilon}$, let $\Sigma(i_1,\dots,i_s)=[f(X(i_1,\dots,i_s))'f(X(i_1,\dots,i_s))]^{-1}$ whenever $X(i_1,\dots,i_s)$ is of full rank, with a similar expression for $\Sigma[i_1,\dots,i_s]$. We are concerned with relative efficiencies $E[X(i_1,\dots,i_s),X]$ and $E[X[i_1,\dots,i_s],X]$ pertaining to deletions from and augmentations to a given design X.

For short-hand notation, when the rows $\{i_1,\dots,i_s\}$ are deleted from the experimental design matrix X, we let G denote the matrix

$$G=f(X_2(i_1,\dots,i_s))\,\Sigma(i_1,\dots,i_s)f(X_2(i_1,\dots,i_s))' \tag{4.1}$$

with eigenvalues $\{\delta_1\geqslant\delta_2\geqslant\cdots\geqslant\delta_s\geqslant 0\}$, and we let H be the matrix

$$H=f(X_2(i_1,\dots,i_s))\,\Sigma f(X_2(i_1,\dots,i_s))' \tag{4.2}$$

with eigenvalues $\{\eta_1\geqslant\eta_2\geqslant\cdots\geqslant\eta_s\geqslant 0\}$.

Corresponding to the relative efficiency $E[X(i_1,\dots,i_s),X]$ are the eigenvalues of $\Sigma^{1/2}(\Sigma(i_1,\dots,i_s))^{-1}\Sigma^{1/2}$ which we have previously denoted as $\{\gamma_1\geqslant\gamma_2\geqslant\cdots\geqslant\gamma_k>0\}$. The next result relates these eigenvalues when all the matrices are of full rank.

Theorem 4. *With the foregoing notation and with all matrices of full rank, the eigenvalues of* $\Sigma^{1/2}(\Sigma(i_1,\dots,i_s))^{-1}\Sigma^{1/2}$ *are* 1.0 *with multiplicity* $k-s$, *and in addition*

$$\{\gamma_{k-r+1}=(1+\delta_r)^{-1}=1-\eta_r;\ 1\leqslant r\leqslant s\}. \tag{4.3}$$

Proof. We assume that X is partitioned as $X=[(X_2(i_1,\dots,i_s))',(X(i_1,\dots,i_s))']'$ and compute $f(X)'f(X)=f(X_2(i_1,\dots,i_s))'f(X_2(i_1,\dots,i_s))+f(X(i_1,\dots,i_s))'f(X(i_1,\dots,i_s))$. The matrix $\Sigma^{-1}\Sigma(i_1,\dots,i_s)=[f(X_2(i_1,\dots,i_s))'f(X_2(i_1,\dots,i_s))]\,\Sigma(i_1,\dots,i_s)+I_k$ has eigenvalues 1.0 with multiplicity $k-s$, since the rank of the first matrix above is s. The remaining eigenvalues are $\{(1+\delta_r);\ 1\leqslant r\leqslant s\}$. Equivalently, the eigenvalues of $\Sigma^{1/2}(\Sigma(i_1,\dots,i_s))^{-1}\Sigma^{1/2}$ are 1.0 with multiplicity $k-s$ and $\{(1+\delta_r)^{-1};\ 1\leqslant r\leqslant s\}$. Similarly, compute

$$\Sigma(\Sigma(i_1,\dots,i_s))^{-1}=[f(X)'f(X)]^{-1}[f(X)'f(X)-f(X_2)'f(X_2)]$$

$$= I_k - [f(X)'f(X)]^{-1}[f(X_2)'f(X_2)]$$
$$= I_k - \Sigma[f(X_2)'f(X_2)]$$

having eigenvalue 1.0 with multiplicity $k-s$, with the remaining eigenvalues given by $\{(1-\eta_r);\ 1 \leqslant r \leqslant s\}$. □

When $s=1$, the nonunit eigenvalue γ_k is related to the variance in the response space for prediction at x_i with x_i missing from the design. We state without proof the following corollary.

Corollary 3. *With $X(i)$ of full rank,*

(i) $$f(x_i)'\Sigma f(x_i) = f(x_i)'\Sigma(i) f(x_i)/[1+f(x_i)'\Sigma(i) f(x_i)] = 1-\gamma_k < 1,$$

(ii) $$f(x_i)'\Sigma(i) f(x_i) = f(x_i)'\Sigma f(x_i)/[1-f(x_i)'\Sigma f(x_i)] = (1-\gamma_k)/\gamma_k > 1.$$

There is an analogous theorem for designs which have been augmented with replicated rows, say $\{i_1, \ldots, i_s\}$. Let T denote the matrix

$$T = f(X_2(i_1, \ldots, i_s))\Sigma[i_1, \ldots, i_s] f(X_2(i_1, \ldots, i_s))' \tag{4.4}$$

with eigenvalues $\{\xi_1 \geqslant \xi_2 \geqslant \cdots \geqslant \xi_s \geqslant 0\}$.

Theorem 5. *With the foregoing notation and with all matrices of full rank, the eigenvalues of $\Sigma^{1/2}(\Sigma[i_1, \ldots, i_s])^{-1}\Sigma^{1/2}$ are 1.0 with multiplicity $k-s$, and*

$$\{\gamma_r = (1-\xi_r)^{-1} = 1+\eta_r;\ 1 \leqslant r \leqslant s\}. \tag{4.5}$$

Proof. This time we have $(\Sigma[i_1, \ldots, i_s])^{-1} = f(X)'f(X) + f(X_2)'f(X_2)$, with $\Sigma^{-1}\Sigma[i_1, \ldots, i_s] = I_k - f(X_2)'f(X_2)\Sigma[i_1, \ldots, i_s]$ and $\Sigma(\Sigma[i_1, \ldots, i_s])^{-1} = I_k + \Sigma f(X_2)'f(X_2)$. □

Thus there is a functional relationship between the relative efficiencies of a deleted design and the corresponding augmented design, for the rows $\{i_1, \ldots, i_s\}$. This relationship is given next where we use $\gamma(\cdot)$ to denote eigenvalues associated with the design with rows deleted, and we use $\gamma[\cdot]$ for the corresponding augmented design.

Corollary 4. *In the foregoing notation, we have*

$$\{\gamma_{k-r+1}(\cdot) + \gamma_r[\cdot] = 2;\ 1 \leqslant r \leqslant s\}. \tag{4.6}$$

We next show that the basis vectors corresponding to the two eigenvalues above are equivalent, in the sense that they differ only by a scalar factor. In other words, we have the intuitive result that the linear parametric function most harmed by deleting a row, is the same linear function most improved by replicating that row.

Theorem 6. *With the foregoing notation and with all matrices of full rank, we have*

$$\{[\gamma_{k-r+1}(\cdot)]^{-1/2} w_{k-r+1}(\cdot) = (\gamma_r[\cdot])^{-1/2} w_r[\cdot];\ \leqslant r \leqslant s\}. \tag{4.7}$$

Proof. Recall that $\Sigma(\Sigma(i_1,\dots,i_s))^{-1} = I_k - \Sigma[f(X_2)'f(X_2)]$ and $\Sigma(\Sigma[i_1,\dots,i_s])^{-1} = I_k + \Sigma[f(X_2)'f(X_2)]$ and so these two matrices have the same eigenvectors, say, $\{x_1,\dots,x_k\}$. The eigenvectors of $\Sigma^{1/2}(\Sigma[i_1,\dots,i_s])^{-1}\Sigma^{1/2}$, say $\{v_1,\dots,v_k\}$, satisfy $\{v_r = \Sigma^{-1/2}x_r;\ 1 \leqslant r \leqslant k\}$. The basis vectors from Section 3.1 are given by $w_r[i_1,\dots,i_s] = (\gamma_r[i_1,\dots,i_s])^{1/2}\Sigma^{-1/2}v_r = (\gamma_r[i_1,\dots,i_s])^{1/2}\Sigma^{-1}x_r$. Similar developments apply for the design with rows $\{i_1,\dots,i_s\}$ deleted. $\square$

4.3. Applications

To fix ideas we focus on the standard 3^2 factorial design together with the model (3.10). Three types of design points may be identified, namely, vertex, edge and interior points, as typified by rows 1, 2 and 5 of Table 1. We now proceed to evaluate the relative efficiencies $E(X(i),X)$ and $E(X(i,j),X)$ as in Section 3 for the model (3.10), and similarly for augmented designs $X[i]$ and $X[i,j]$, with $\{i,j \in \{1,2,5\}\}$. Details follow, where it suffices to report the vector $\gamma = [\gamma_1,\dots,\gamma_6]'$ of eigenvalues together with the corresponding basis vectors $\{w_1,\dots,w_6\}$ spanning subspaces of $\mathbb{R}^6$.

In particular, for $E(X(1),X)$ the eigenvalues are $\gamma(1) = [1\ 1\ 1\ 1\ 1\ 0.194]'$, indicating equal efficiency except for the linear parametric function

$$(\beta_1 + \beta_2) - (\beta_0 + \beta_{11} + \beta_{22} + \beta_{12}) \tag{4.8}$$

corresponding to $w_6(1) = 0.491[-1\ 1\ 1\ -1\ -1\ -1]'$. If we now augment the original design on duplicating the first row, we have for $E(X[1],X)$ the eigenvalues $\gamma[1] = [1.806\ 1\ 1\ 1\ 1\ 1]'$, with $w_1[1] = 1.497[-1\ 1\ 1\ -1\ -1\ -1]$. The augmented design is more efficient for $(\beta_1+\beta_2)-(\beta_0+\beta_{11}+\beta_{22}+\beta_{12})$ by a factor of 1.806, with the original and augmented designs having equal efficiencies elsewhere. Corollary 4 assures that $\gamma_6(1)+\gamma_1[1] = 2$ and Theorem 6 that $w_1[1] = (1.806/0.194)^{1/2}w_6(1)$.

Similar results for $E(X(2),X)$ are $\gamma(2) = [1\ 1\ 1\ 1\ 1\ 0.444]'$ and $w_6(2) = 0.894[-1\ 1\ 0\ -1\ 0\ 0]'$, indicating a loss of information about $\beta_1-\beta_0-\beta_{11}$, with equal efficiencies elsewhere. Augmenting X on duplicating its second row gives $\gamma[2] = [1.556\ 1\ 1\ 1\ 1\ 1]'$, with $w_1[2]$ now taking the value $w_1[2] = 1.673[-1\ 1\ 0\ -1\ 0\ 0]'$, improving by a factor of 1.556 the efficiency for inferences regarding $\beta_1-\beta_0-\beta_{11}$.

Deleting the center run gives $\gamma(5) = [1\ 1\ 1\ 1\ 1\ 0.444]'$ and $w_6(5) = 0.894[1\ 0\ 0\ 0\ 0\ 0]'$ for $E(X(5),X)$, indicating a loss of information about β_0 only. Adding an extra center run gives $\gamma[5] = [1.556\ 1\ 1\ 1\ 1\ 1]'$ for $E(X[5],X)$, the revised vector $w_1[5]$ becoming $w_1[5] = 1.673[1\ 0\ 0\ 0\ 0\ 0]'$. Thus the greatest loss of information occurs when the vertex row 1 is deleted, or equivalently from Corollary 4, the greatest information is gained when a vertex row is replicated. The loss of information is the same when an edge row or an interior row is deleted. However, the correspo linear parametric functions are very different.

Similar results were found for the design Z featured in Section 3. Somewha ferent weights appear on occasion in the spanning vectors. Additional detai omitted.

Further comparisons entail dropping pairs of points. Retaining the 3^2 fac design for generating X, the case $E(X(1,2),X)$ gives $\gamma(1,2)=[1\ 1\ 1\ 1\ 0.574\ 0.$ indicating a severe loss of information regarding the parametric function

$$0.227(\beta_2-\beta_{22}-\beta_{12})-0.359(\beta_0-\beta_1+\beta_{11})$$

corresponding to $w_6(1,2)=[-0.359\ 0.359\ 0.227\ -0.359\ -0.227\ -0.227]'$. An one-dimensional linear subspace of note is generated by the spanning vector $w_5($ $[-0.416\ 0.416\ -0.587\ -0.416\ 0.587\ 0.587]'$ corresponding to $0.416(\beta_1-\beta_0-$ $0.587(\beta_2-\beta_{22}-\beta_{12})$. Inferences about all other linear parametric functio comparable whether or not rows 1 and 2 have been deleted. If the design i augmented on duplicating rows 1 and 2, we find for $E(X[1,2],X)$ that $\gamma[$ $[1.936\ 1.426\ 1\ 1\ 1\ 1]'$, with $w_1[1,2]$ and $w_2[1,2]$ as scalar multiples of $w_6(1,2$ $w_5(1,2)$, respectively.

Similar results for $E(X(1,5),X)$ give $\gamma(1,5)=[1\ 1\ 1\ 1\ 0.487\ 0.152]'$, togethe the vectors

$$w_5(1,5)=[1.256\ -0.346\ -0.346\ 0.346\ 0.346\ 0.346]',$$

$$w_6(1,5)=[-0.245\ 0.396\ 0.396\ -0.396\ -0.396\ -0.396]',$$

corresponding to the parametric functions $1.256\beta_0-0.346(\beta_1+\beta_2-\beta_{11}-\beta_{22}$ and $0.396(\beta_1+\beta_2-\beta_{11}-\beta_{22}-\beta_{12})-0.245\beta_0$, respectively. The corresponding augmentation gives $\gamma[1,5]=[1.858\ 1.513\ 1\ 1\ 1\ 1]'$ for $E(X[1,5],X)$, with w and $w_2[1,5]$ now as scalar multiples of $w_6(1,5)$ and $w_5(1,5)$, respectively.

Conclusions for $E(X(2,5),X)$ are straightforward, giving the vector of parative efficiencies $\gamma(2,5)=[1\ 1\ 1\ 1\ 0.667\ 0.222]'$ together with the spanning v

$$w_5(2,5)=[0\ -1\ 0\ 1\ 0\ 0]'$$

and

$$w_6(2,5)=[-0.756\ 0.378\ 0\ -0.378\ 0\ 0]'.$$

On the other hand, a design augmentation yields $\gamma[2,5]=[1.778\ 1.333\ 1\ 1\ 1$ $E(X[2,5],X)$, with $w_1[2,5]$ and $w_2[2,5]$ as scalar multiples of $w_6(2,5)$ and w respectively. This augmented design has enhanced capacity for the linear co $(\beta_{11}-\beta_1)$, by a factor of 1.333, and for $0.378(\beta_1-\beta_{11})-0.755\beta_0$, by a fac 1.778. Elsewhere the 3^2 factorial design, and its augmentation via duplicatin (2,5), have comparable efficiencies. Comparing the eigenvalues $\gamma(1,2)$, $\gamma(1,$ $\gamma(2,5)$, we find that the smallest loss of information occurs when the pair o (2,5) is deleted. In the next section we discuss the interpretation of $\Sigma^{1/2}[\Sigma(\cdot)]$

To examine the robustness of a design, some authors use a scalar design cri divided by the number of observations; see, for example, Andrews and He

(1979). Our view is different. In this section on augmentation and deletion, our goal is to quantify not only the gain or loss of information, but also to identify those linear inferences most affected by replicating or deleting observations.

4.4. Measures of influence

We next relate the eigenvalues determined in the preceding section to measures of influence studied in the literature. Mukerjee and Kageyama (1990) have investigated robustness of group divisible designs to loss of sets of design points. To compare the design X to the same design with rows $\{i_1 < \cdots < i_s\}$ missing, say $X(i_1, \dots, i_s)$, they use for a measure of influence the quantity

$$\begin{aligned} e_1(\Sigma(i_1,\dots,i_s)) &= \operatorname{tr}[f(X)'f(X)]^{-1}/\operatorname{tr}[f(X(i_1,\dots,i_s))'f(X(i_1,\dots,i_s))]^{-1} \\ &= \operatorname{tr}\Sigma/\operatorname{tr}\Sigma(i_1,\dots,i_s), \end{aligned} \tag{4.14}$$

if all inverses exist, and zero if $f(X(i_1, \dots, i_s))$ is not of full rank. For example, with the model (3.10) together with the standard 3^2 factorial design having the first row (vertex) deleted, the influence is $e_1 = 2.1389/3.0952 = 0.6910$. Table 5 contains the values for e_1 corresponding to the three designs we have studied in which various rows have been deleted. This measure of influence would indicate the greatest loss of information when the edge row 2 is deleted. For a single scalar we would prefer to compute $\operatorname{tr}[\Sigma(\Sigma(i_1, \dots, i_s))^{-1}]$. We will return to this point shortly.

Ghosh (1989) has introduced a measure of influence to identify subsets of influential observations at the design stage. His measure, in the spirit of Cook's (1977) distance, is defined as $I_1(\Sigma(i_1, \dots, i_s)) = \sigma^2 \operatorname{tr}[f(X_2(i_1, \dots, i_s))\Sigma(i_1, \dots, i_s) f(X_2(i_1, \dots, i_s))']$, where $X_2(i_1, \dots, i_s)$ is the $(s \times k)$ matrix whose rows are the s rows deleted from X. In the notation of Section 4.2, this is $I_1(\Sigma(i_1, \dots, i_s)) = \sigma^2 \operatorname{tr} G = \sigma^2(\delta_1 + \cdots + \delta_s)$. In particular, with one row deleted, say row i, we have the quantity $I_1(\Sigma(i)) =$

Table 5
Measures of influence pertaining to rows 1, 2 and 5 for three designs, including e_1, $\sigma^{-2}I_1$ and the nonunit eigenvalue γ_i

Matrix	e_1	$\sigma^{-2}I_1$	γ_i
$\Sigma(1)$	0.6910	4.1428	0.1944
$\Sigma(2)$	0.8148	1.2500	0.4444
$\Sigma(5)$	0.6417	1.2500	0.4444
$\Omega(1)$	0.7596	4.1428	0.1944
$\Omega(2)$	0.7661	1.2500	0.4444
$\Omega(5)$	0.6779	1.2500	0.4444
$\Delta(1)$	0.9000	1.6667	0.3750
$\Delta(2)$	0.8514	1.6667	0.3750
$\Delta(5)$	0.0000	**	**

** Not defined.

$\sigma^2 f(x_i)'\Sigma(i)f(x_i)$. For example, given the model (3.10) and the standard 3^2 factorial design with row 2 (an edge) deleted, we compute $I_1(\Sigma(2)) = \sigma^2 f(x_2)'\Sigma(2)f(x_2) = 1.25\sigma^2$. Recalling that $\gamma(2) = [1\ 1\ 1\ 1\ 1\ 0.444]'$, we note that Theorem 4 relates the nonunit eigenvalue 0.444 to the measure of influence $I_1(\Sigma(2))$ by the rule $0.444 = 1/(1+1.25)$. Our work transcends earlier studies by Ghosh (1989) in that we are able to relate the loss of efficiency directly to the loss of information about $\beta_1 - \beta_0 - \beta_{11}$ through the vector w_6 as the last column of $W = \Sigma^{-1/2}V$, where V consists of the eigenvectors of $\Gamma(2) = \Sigma^{1/2}[\Sigma(2)]^{-1}\Sigma^{1/2}$. Values for the Ghosh (1989) index I_1, and corresponding values for the nonunit eigenvalue γ_i, are given in Table 5 for the three types of designs studied here in which various rows have been deleted.

The Ghosh (1989) measure of influence indicates the greatest loss of information when the vertex row 1 is deleted. This is consistent with the calculations in Section 4.3 since a large value of δ_1 corresponds to a small value for the efficiency γ_k.

When two or more rows have been deleted from X, the nonunit eigenvalues $\{\gamma_{i_1}, \ldots, \gamma_{i_s}\}$ of $\Gamma(i_1, \ldots, i_s) = \Sigma^{1/2}[\Sigma(i_1, \ldots, i_s)]^{-1}\Sigma^{1/2}$ will measure the loss of efficiency, and the associated eigenvectors $\{v_{i_1}, \ldots, v_{i_s}\}$ relate directly to those linear inferences $w_i'\beta$, with $\{w_i = \Sigma^{-1/2}v_i,\ 1 \leqslant i \leqslant s\}$, which will suffer a loss of information. If the researcher wishes to present a composite scalar measure of influence due to deleting rows $\{i_1, \ldots, i_s\}$, then the trace or determinant of $\Gamma(i_1, \ldots, i_s)$ can be used. As an example, with X as the 3^2 factorial design and with rows 1 and 2 (a corner and edge) deleted, we have $\gamma(1,2) = [1\ 1\ 1\ 1\ 0.574\ 0.067]'$, and so $\mathrm{tr}(\Gamma(1,2)) = 4.638$ and $|\Gamma(1,2)| = 0.0367$. The measure of influence introduced by Ghosh (1989) would be $I_1(\Sigma(1,2)) = \sigma^2 \,\mathrm{tr}\, G = 15.25\sigma^2$, where

$$G = \begin{bmatrix} f(x_1)' \\ f(x_2)' \end{bmatrix} \Sigma(1,2)\,[f(x_1) f(x_2)] = \begin{bmatrix} 11 & 6 \\ 6 & 4.35 \end{bmatrix}. \tag{4.15}$$

The matrix G has eigenvalues 14.5090 and 0.7409 which are related to 0.064 and 0.574 respectively by Theorem 4, for example, $0.064 = 1/(1+14.5090)$. The three measures of influence, $\mathrm{tr}(\Gamma)$, $|\Gamma|$, and $I_1(\Gamma)$, all indicate that the smallest loss of information occurs when the pair (2,5) is deleted. The values for $|\Gamma|$ are given in Table 6.

Table 6

Determinants of $\Gamma = \Omega^{-1/2}\Sigma\Omega^{-1/2}$ for designs obtained on deleting and augmenting rows (i) or (i,j) in a 3^2 factorial design for the model (3.10), where $\Sigma = [f(X)'f(X)]^{-1}$ and Ω is obtained on deletion or augmentation as appropriate

Row(s)	(1)	(2)	(5)	(1,2)	(1,5)	(2,5)
Deleted	0.194	0.444	0.444	0.037	0.074	0.148
Augmented	1.806	1.556	1.556	2.761	2.796	2.370

5. Summary

To help summarize these numerical results for the model (3.10), we record in Table 6 the determinant of the matrix $\boldsymbol{\Gamma} = \boldsymbol{\Omega}^{-1/2} \boldsymbol{\Sigma} \boldsymbol{\Omega}^{-1/2}$ with $\boldsymbol{\Sigma} = [f(X)'f(X)]^{-1}$ and with $\boldsymbol{\Omega}$ as the corresponding matrix for designs determined by the deletion and augmentation process. The determinants and traces are useful as composite scalar indices of efficiency. In applications the linear parametric functions are of primary interest and must be acknowledged in order to utilize effectively the fine structure contained in the eigenstructure of $\boldsymbol{\Gamma}$.

Methods presented here share common ground with those of Ghosh (1989) and Mukerjee and Kageyama (1990). Specifically, these methods depend exclusively on the designs in question and are independent of empirical observations. All such comparisons can be made numerically during planning before an experiment has been done. These methods can be contrasted with those of Cook (1977) and Takeuchi (1991), for example, which are data-dependent and thus of value in reassessing a particular experiment retrospectively.

Computations are easily programmed using standard statistical software. Extensive numerical studies to evaluate alternative designs have been undertaken by the authors, along the lines of examples reported here, using SAS and Minitab software.

6. Acknowledgment

The authors acknowledge with thanks numerous helpful comments by the referees and Professor F. Pukelsheim. Work of the second author was supported in part by the Center for Advanced Studies at the University of Virginia.

References

Andrews, D. and A. Herzberg (1979). The robustness and optimality of response surface designs. *J. Statist. Planning and Inference* **3**, 249–257.

Bellman, R. (1960). *Introduction to Matrix Analysis*. McGraw-Hill, New York.

Cook, R.D. (1977). Detection of influential observations in linear regression. *Technometrics* **22**, 459–508.

Fedorov, V.V. (1972). *Theory of Optimal Experiments*. Academic Press, New York.

Galil, Z. and J. Kiefer (1977a). Comparison of rotatable designs for regression on balls, I (Quadratic). *J. Statist. Planning and Inference* **1**, 27–40.

Galil, Z. and J. Kiefer (1977b). Comparison of design for quadratic regression on cubes. *J. Statist. Planning and Inference* **1**, 121–132.

Galil, Z. and J. Kiefer (1977c). Comparison of Box-Draper and D-optimum designs for experiments with mixtures. *Technometrics* **19**, 441–444.

Galil, Z. and J. Kiefer (1977d). Comparison of simplex designs for quadratic mixture models. *Technometrics* **19**, 445–453.

Galil, Z. and J. Kiefer (1979). Extrapolation designs and Φ_p-optimum designs for cubic regression on the q-ball. *J. Statist. Planning and Inference* **3**, 27-38.

Ghosh, S. (1989). On two methods of identifying influential sets of observations. *Statistics and Probability Letters* **7**, 241-245.

Jensen, D.R. (1991). Vector efficiency in multiparameter estimation. *Linear Algebra and Its Applications* **151**, 143-155.

Kiefer, J. (1959). Optimum experimental designs. *J.R. Statist. Soc. B* **21**, 272-319.

Kiefer, J. (1975). Optimal design: Variation in structure and performance under change of criterion. *Biometrika* **62**, 277-288.

Kotz, S., N.L. Johnson and C.B. Read (1985). *Encyclopedia of Statistical Sciences, Vol.* 6. Wiley, New York, pp. 731-735.

Kurotschka, V. (1978). Optimal design of complex experiments with qualitative factors of influence. *Commun. Statist. A-Theory Meth.* **7**, 1363-1378.

Loewner, C. (1934). Über monotone Matrix Funktionen. *Math. Z.* **38**, 177-216.

Mukerjee, R. and S. Kageyama (1990). Robustness of group divisible designs. *Commun. Statist.-Theory Meth.* **19**, 3189-3203.

Pitman, E.J.G. (1948). *Nonparametric Statistical Inference*. Lecture Notes, Institute of Statistics, University of North Carolina, Chapel Hill.

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

Pukelsheim, F. (1981). On c-optimal design measures. *Math. Operationsforsch. Statist., Ser. Statistics* **12**, 13-20.

Rao, C.R. (1973). *Linear Statistical Inference and Its Applications, Second Edition*. Wiley, New York.

Takeuchi, H. (1991). Detecting influential observations by using a new expression of Cook's distance. *Commun. Statist.-Theory Meth.* **20**, 261-274.

Wynn, H.P. (1970). The sequential generation of D-optimal designs. *Ann. Math. Statist.* **41**, 1655-1664.