New Developments and Applications in Experimental Design
IMS Lecture Notes – Monograph Series (1998) Volume 34

POLYNOMIAL REPRESENTATIONS FOR RESPONSE SURFACE MODELING

BY NORMAN R. DRAPER[1] AND FRIEDRICH PUKELSHEIM[1,2]

University of Wisconsin and University of Washington

In response surface models the expected response is usually taken to be a low degree polynomial in the design variables that are coded from the factor settings. We argue that an overparameterized polynomial representation of the expected response offers great economy and transparency. As an illustration, we exhibit a constructive path of design improvement relative to the Kiefer design ordering, for polynomial regression up to degree three when the experimental domain is a ball.

1. Introduction. We overview some of the recent work on design optimality for response surface models and polynomial regression. However, our emphasis is not on scalar optimality criteria. Any such criterion singles out one—or a few—designs as being optimal, while saying little or nothing about all the other designs that are nonoptimal.

Rather, we concentrate on the Kiefer design ordering. We show that under this partial ordering there is a constructive path of design improvement. Starting with an arbitrary design, good or bad, we are lead to a small design class that turns out to be minimal complete. This is carried out for first-, second-, and third-degree polynomial response surface models when the experimental domain is a ball.

The Kiefer design ordering does not depend on how the polynomials are represented. This opens the way to write the regression function in a form that is deemed most convenient. We argue that the Kronecker product offers attractive symmetry, compact notation, and great transparency. The present paper offers a short-cut access by just verifying the results. The underlying theory for deriving these results is available in greater detail elsewhere in the literature.

A brief review of the literature is as follows. Pukelsheim (1993, p. 354) introduced the Kiefer ordering, thus extending Kiefer's (1975, p. 336) notion of universal optimality to general design settings. The Kiefer ordering combines two steps, increase in symmetry is followed by the usual enlargement of the moment matrix of a design. Kiefer (1975)

[1] Supported in part by the Alexander von Humboldt-Stiftung, Bonn, through a Max Planck-Award for cooperative research.

[2] Supported in part by the Volkswagen-Stiftung, Hannover, while on leave from the University of Augsburg.

Received July 1997; revised January 1998.

AMS 1991 subject classifications. Primary 62K15, 62J05; secondary 15A69, 15A45.

Key words and phrases. Box–Hunter representation, cubic regression, design admissibility, Kiefer design ordering, Kronecker product, minimal complete design classes, quadratic regression, rotatable designs, tensor product.

concentrated on block design settings whose natural companion is the permutation group; Cheng (1995) derives results for general permutation-invariant design regions.

For response surface models on the ball, the symmetry originates with the orthogonal group. Draper, Gaffke and Pukelsheim (1991), and Pukelsheim (1993, p. 394) use a Kronecker product regression function for the second-degree model. These sources contain many more references to work on second-degree response surface models.

A third-degree Kronecker representation was investigated by Draper and Pukelsheim (1994). That paper concludes with some remarks on higher degree rotatability and its relations with multilinear algebra. Design optimality under the standard criteria is discussed in Draper, Heiligers and Pukelsheim (1996).

Beyond polynomial regression on the ball, an application of the Kronecker algebra to the Scheffé mixture models on the simplex is proposed by Draper and Pukelsheim (1997). In either instance, the overparameterization that is inherent in the Kronecker approach creates no difficulties whatsoever.

The more widespread approach stems from the seminal paper of Box and Hunter (1957) who chose a minimal set of linearly independent monomials. Their argument that rotatability of the variance surface and rotatability of the moment matrix are equivalent is somewhat brief, and is detailed in Draper, Gaffke and Pukelsheim (1991, p. 153; 1993). For the Box–Hunter regression function, it makes a difference whether design admissibility—or optimality–is referred to the set of all designs, or to the proper subset of rotatable designs, see Karlin and Studden (1966, p. 356), or Galil and Kiefer (1979, p. 29). The reason is the following.

For the Box–Hunter regression function, the orthogonal group on the experimental domain induces a group $\widetilde{\mathcal{Q}}$ on the regression range containing matrices that are nonorthogonal. Heiligers (1991, p. 118) points out that, in order that all matrices in $\widetilde{\mathcal{Q}}$ are orthogonal, the 'biggest' group of transformations on the experimental domain is the one generated by all permutations and sign changes. Gaffke and Heiligers (1995, 1996) obtain many results for the permutation-and-sign-change group, and discuss the relation to the corresponding results for the full orthogonal group.

The present paper is organized as follows. Section 2 reviews Kronecker products of vectors and matrices. In Section 3 this is used to compactly represent polynomial regression functions, of up to degree three. In all cases the matrix group $\mathcal{Q}$ that is induced on the regression range contains orthogonal matrices only.

Section 4 discusses the first step of the Kiefer ordering, symmetrization. It transpires that rotatable moment matrices have a much simpler pattern than arbitrary moment matrices. This aids in calculating generalized inverses, and the information surfaces. Section 5 studies the second step of the Kiefer ordering, to constructively enlarge a given rotatable moment matrix relative to the usual Loewner matrix ordering.

Section 6 joins the two intermediate steps together, to obtain the Kiefer ordering. In the first-degree model, there exists a Kiefer optimal design. In the second-degree model, there is a one-parameter design family that is minimal complete. In the third-degree model, the minimal complete class has two parameters. The Kiefer ordering is invariant to a change of basis for the regression range whence our results, while conveniently derived using Kronecker algebra, continue to hold true for the Box–Hunter regression function.

2. Kronecker products. The idea underlying the use of Kronecker products is familiar from elementary statistics. For a random vector Y in $\mathbb{R}^n$, the variances and covariances of its components are redundantly assembled into an $n \times n$ dispersion matrix

$$\mathrm{D}[Y] = \begin{pmatrix} \mathrm{var}(Y_1) & \mathrm{cov}(Y_1, Y_2) & \cdots & \mathrm{cov}(Y_1, Y_n) \\ \mathrm{cov}(Y_2, Y_1) & \mathrm{var}(Y_2) & \cdots & \mathrm{cov}(Y_2, Y_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(Y_n, Y_1) & \mathrm{cov}(Y_n, Y_2) & \cdots & \mathrm{var}(Y_n) \end{pmatrix},$$

and not reduced to the distinct entries

$$\big(\mathrm{var}(Y_1), \ldots, \mathrm{var}(Y_n), \mathrm{cov}(Y_1, Y_2), \ldots, \mathrm{cov}(Y_{n-1}, Y_n)\big).$$

The benefits are that $\mathrm{D}[Y]$ is visibly exhibited as a quadratic form, and the rules for a transformation with a conformable matrix A become quite simple, $\mathrm{D}[AY] = A\big(\mathrm{D}[Y]\big)A'$.

Similarly, the Kronecker product approach bases second-degree polynomial regression in m variables $t = (t_1, \ldots, t_m)'$ on the matrix of all cross products,

$$tt' = \begin{pmatrix} t_1^2 & t_1 t_2 & \cdots & t_1 t_m \\ t_2 t_1 & t_2^2 & \cdots & t_2 t_m \\ \vdots & \vdots & \ddots & \vdots \\ t_m t_1 & t_m t_2 & \cdots & t_m^2 \end{pmatrix},$$

rather than reducing them to the Box-Hunter minimal set of monomials

$$(t_1^2, \ldots, t_m^2, t_1 t_2, \ldots, t_{m-1} t_m).$$

The benefits are that distinct terms are repeated appropriately, according to the number of times they can arise, that transformational rules with a conformable matrix R become simple, $(Rt)(Rt)' = R(tt')R'$, and that the approach extends to third-degree polynomial regression. However, the arrangement of triple products $t_i t_j t_k$ in a set of "layered" matrices appears rather awkward. This is where Kronecker products prove useful, they achieve the same goal with a more pleasing algebra.

For a $k \times m$ matrix A and an $\ell \times n$ matrix B, their Kronecker product $A \otimes B$ is defined to be the $k\ell \times mn$ block matrix

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & \ddots & \vdots \\ a_{k1}B & \cdots & a_{km}B \end{pmatrix}.$$

The Kronecker product of a vector $s \in \mathbb{R}^m$ and another vector $t \in \mathbb{R}^n$ then simply is a special case,

$$s \otimes t = \begin{pmatrix} s_1 t \\ \vdots \\ s_m t \end{pmatrix} = \underset{\text{in lexicographic order}}{(s_i t_j)_{i=1,\ldots,m;\ \ j=1,\ldots,n}} \in \mathbb{R}^{mn}.$$

A key property is the product rule $(A \otimes B)(s \otimes t) = (As) \otimes (Bt)$. This has nice implications for transposition, $(A \otimes B)' = (A') \otimes (B')$, for Moore-Penrose inversion, $(A \otimes B)^+ = (A^+) \otimes (B^+)$, and—if possible—for regular inversion, $(A \otimes B)^{-1} = (A^{-1}) \otimes (B^{-1})$. It is of specific importance to us that the Kronecker product preserves orthogonality: If A and B are individual orthogonal matrices, then their Kronecker product $A \otimes B$ is also an orthogonal matrix.

Thus while the matrix tt' assembles the cross products $t_i t_j$ in an $m \times m$ array, the Kronecker square $t \otimes t$ arranges the same numbers as a long $m^2 \times 1$ vector. The Kronecker cube $t \otimes t \otimes t$ is an even longer $m^3 \times 1$ vector, listing the triple products $t_i t_j t_k$ in lexicographic order. Yet the algebra is easy to handle. The transformation with a conformable matrix R simply amounts to $(Rt) \otimes (Rt) = (R \otimes R)(t \otimes t)$. This greatly facilitates our calculations when we now apply Kronecker products to response surface models.

3. Polynomial regression. We consider multifactor experiments, for m deterministic input factors. For $i = 1, \ldots, m$ let $t_i \in \mathbb{R}$ be the level of the factor i. Together they form the vector of experimental conditions, $t = (t_1, \ldots, t_m)' \in \mathbb{R}^m$, for which we assume the experimental domain to be the ball of radius $r > 0$ in $\mathbb{R}^m$,

$$t \in \mathcal{T} = \{t \in \mathbb{R}^m : \|t\| \leq r\}.$$

The available levels often include zeroes for the standard operating conditions, and ± 1 for a deviation of one—appropriately scaled—unit in either direction. Therefore the radius $r = \sqrt{m}$ is of particular interest, covering the full factorial design on the cube $\{\pm 1\}^m$ as well as fractions thereof. The cube $\{\pm 1\}^m$ has volume 2^m. Hence the volume for radius $\sqrt{m}$ grows exponentially and more than compensates for the shrinking volume of the unit ball (which equals $\pi^m/m!$ for even dimension $2m$, and $2^{m+1}\pi^m/\big(1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2m+1)\big)$ for odd dimension $2m+1$).

Dimension m	2	3	4	5	6	7	8	9	10
Vol. for $r = 1$	3.142	4.189	4.935	5.264	5.168	4.725	4.059	3.299	2.550
Vol. for $r = \sqrt{m}$	6.283	21.77	78.96	294.3	1116.2	4287.7	16624.5	64924.6	255016.4

Response surface models apply to scalar responses Y_t, assuming that observations under identical or distinct experimental conditions t are of equal (unknown) variance σ^2, and uncorrelated. Moreover, these models assume that the expected response $\mathrm{E}[Y_t] = \eta(t, \Theta)$ permits a fit with a low-degree polynomial in t. Making use of the Kronecker product, the first-, second-, and third-degree models then are

$$\begin{aligned}
\eta(t, \Theta) &= \theta_0 + t'\theta_{\{i\}}, \\
\eta(t, \Theta) &= \theta_0 + t'\theta_{\{i\}} + (t \otimes t)'\theta_{\{ij\}}, \\
\eta(t, \Theta) &= \theta_0 + t'\theta_{\{i\}} + (t \otimes t)'\theta_{\{ij\}} + (t \otimes t \otimes t)'\theta_{\{ijk\}}.
\end{aligned}$$

The hierarchy from lower to higher degree models is less well reflected for the Box–Hunter regression function for which the arrangement of entries varies from author to author, and occasionally for the same author from one paper to another.

The parameter vectors are, in turn,

$$\Theta = \begin{pmatrix} \theta_0 \\ \theta_{\{i\}} \end{pmatrix}, \qquad \Theta = \begin{pmatrix} \theta_0 \\ \theta_{\{i\}} \\ \theta_{\{ij\}} \end{pmatrix}, \qquad \Theta = \begin{pmatrix} \theta_0 \\ \theta_{\{i\}} \\ \theta_{\{ij\}} \\ \theta_{\{ijk\}} \end{pmatrix}.$$

The individual components have the usual interpretation, with θ_0 being the grand mean. The $m \times 1$ vector $\theta_{\{i\}} = (\theta_1, \ldots, \theta_m)'$ consists of the main effects θ_i.

The $m^2 \times 1$ vector $\theta_{\{ij\}} = (\theta_{11}, \theta_{12}, \ldots, \theta_{mm})'$ consists of the pure quadratic effects θ_{ii} and the two-way interactions θ_{ij}, with the evident second-degree restrictions $\theta_{ij} = \theta_{ji}$ for all i, j. The $m^3 \times 1$ vector $\theta_{\{ijk\}}$ comprises the pure cubic effects θ_{iii} and the two- and three-way interactions θ_{iij} and θ_{ijk}, with the evident third-degree restrictions $\theta_{ijk} = \theta_{ikj} = \theta_{jik} = \theta_{jki} = \theta_{kij} = \theta_{kji}$ for all i, j, k.

Each of these models is of the form $\eta(t, \Theta) = f(t)'\Theta$. The regression functions $t \mapsto f(t)$ conform to the parameter vectors Θ and are, in turn,

$$f(t) = \begin{pmatrix} 1 \\ t \end{pmatrix}, \qquad f(t) = \begin{pmatrix} 1 \\ t \\ t \otimes t \end{pmatrix}, \qquad f(t) = \begin{pmatrix} 1 \\ t \\ t \otimes t \\ t \otimes t \otimes t \end{pmatrix}.$$

As t varies over the experimental domain $\mathcal{T}$, the vectors $f(t)$ span spaces of respective dimensions $m + 1$, $(m + 1)(m + 2)/2$, and $(m + 1)(m + 2)(m + 3)/6$. These numbers coincide with the distinct components in the parameter vectors Θ. Thus the Kronecker models are seen to be overparameterized, from degree two onwards.

An experimental design τ on the domain $\mathcal{T}$ is a probability measure that has finite support. Suppose the support points are $t_1, \ldots, t_\ell$, and they have corresponding weights $w_1, \ldots, w_\ell$, the experimenter is then directed to draw a proportion w_j of all observations under experimental conditions t_j. For a linear model with regression function $f(t)$, the statistical properties of a design τ are captured by its moment matrix

$$M(\tau) = \sum_{j \leq \ell} w_j \, f(t_j) f(t_j)' = \int_{\mathcal{T}} f(t) f(t)' \, d\tau.$$

Because of overparameterization, any such moment matrix is rank deficient, and so is the dispersion matrix of the least squares estimator for Θ. While regular matrix inverses then do not exist, generalized inverses work just as well.

The dependence of the expected response on the experimental conditions t is described by the model response surface $t \mapsto \eta(t, \Theta)$. The parameter vector Θ is generally not known. When we replace the true parameter vector Θ by its least squares estimate $\widehat{\Theta}$, we shift interest to the estimated response surface $t \mapsto \eta(t, \widehat{\Theta}) = f(t)'\widehat{\Theta}$. When $\widehat{\Theta}$ is calculated from observations drawn according to the experimental design τ, the statistical properties of the estimated response surface are determined by the variance surface $t \mapsto v_\tau(t) = f(t)'M(\tau)^- f(t)$, or equivalently, by the information surface $t \mapsto i_\tau(t) = 1/v_\tau(t)$. These

quantities do not depend on the choice of the generalized inverse provided the vector $f(t)$ lies in the range of the matrix $M(\tau)$; otherwise a continuity argument suggests setting $v_\tau(t) = \infty$ and $i_\tau(t) = 0$, which makes good sense also statistically. The information surface $i_\tau(t)$ ranges from zero to some finite maximum, and is thus easier to show graphically than is the variance surface, which goes to infinity.

Let R be an $m \times m$ orthogonal matrix, transforming the experimental conditions t into Rt. Many response surface applications concentrate on the distance from the standard operating conditions which are usually coded to be the origin of the experimental domain. In such circumstances, it becomes desirable to choose the design τ in such a way that the information surface (and hence the variance surface) is rotatable,

$$i_\tau(t) = i_\tau(Rt) \qquad \text{for all } R \in \mathrm{Orth}(m),$$

where $\mathrm{Orth}(m)$ is the group of orthogonal $m \times m$ matrices. Such designs are characterized by an invariance property of their moment matrices.

4. Rotatable moment matrices. A linear transformation $t \mapsto Rt$ induces a linear transformation of the regression function, $f(Rt) = Q_R f(t)$. This is a consequence of the key product rule of Kronecker products. In the third-degree model, this follows from

$$f(Rt) = \begin{pmatrix} 1 \\ Rt \\ (Rt) \otimes (Rt) \\ (Rt) \otimes (Rt) \otimes (Rt) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & R & 0 & 0 \\ 0 & 0 & R \otimes R & 0 \\ 0 & 0 & 0 & R \otimes R \otimes R \end{pmatrix} \begin{pmatrix} 1 \\ t \\ t \otimes t \\ t \otimes t \otimes t \end{pmatrix}.$$

The induced transformation matrices Q_R then are, for first-, second-, and third-degree,

$$Q_R = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}, \quad Q_R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & R & 0 \\ 0 & 0 & R \otimes R \end{pmatrix}, \quad Q_R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & R & 0 & 0 \\ 0 & 0 & R \otimes R & 0 \\ 0 & 0 & 0 & R \otimes R \otimes R \end{pmatrix}.$$

Due to the Kronecker product properties, the mapping $R \mapsto Q_R$ preserves orthogonality: If R is an orthogonal matrix, so is Q_R. Hence the induced matrix groups

$$\mathcal{Q} = \{\, Q_R : R \in \mathrm{Orth}(m) \,\}$$

are proper subgroups of the orthogonal groups $\mathrm{Orth}(k)$ on the spaces $\mathbb{R}^k$, with $k = 1 + m, 1 + m + m^2, 1 + m + m^2 + m^3$, respectively. This is in sharp contrast to the Box–Hunter regression function for which the induced matrix $\widetilde{Q}_R$ need not be an orthogonal matrix although R is.

The linear transformation $f(t) \mapsto Q_R f(t)$ induces a congruence transformation among moment matrices, $M \mapsto Q_R M Q_R'$. A design has a rotatable information surface if and only if these congruence transformations leave its moment matrix invariant,

$$M = Q_R M Q_R' \qquad \text{for all } R \in \mathrm{Orth}(m).$$

Invariant moment matrices are called rotatable, and denoted by $\overline{M}$. Rotatability of moment matrices is a rather stringent requirement. It can be shown that first-, second-, and third-degree rotatable moment matrices depend on one, two, and three parameters and have the form, in turn,

$$\overline{M} = \begin{pmatrix} 1 & 0 \\ 0 & \mu_2 I_m \end{pmatrix}, \qquad \overline{M} = \begin{pmatrix} 1 & 0 & \mu_2 v_m' \\ 0 & \mu_2 I_m & 0 \\ \mu_2 v_m & 0 & \mu_{22} F_m \end{pmatrix},$$

$$\overline{M} = \begin{pmatrix} 1 & 0 & \mu_2 v_m' & 0 \\ 0 & \mu_2 I_m & 0 & \mu_{22} G_m' \\ \mu_2 v_m & 0 & \mu_{22} F_m & 0 \\ 0 & \mu_{22} G_m & 0 & \mu_{222} S_m \end{pmatrix}.$$

Here I_m is the $m \times m$ identity matrix, and v_m is its column vectorized form; F_m, G_m, and S_m are known integer matrices [Draper, Gaffke and Pukelsheim (1994, p. 145)]. The integer entries take care of the moment identities that are well-known to accompany rotatability, $\mu_4 = 3\mu_{22}$, $\mu_{42} = 3\mu_{222}$, and $\mu_6 = 15\mu_{222}$. Breaking the matrices down to their entries is space filling, and confusing rather than enlightening. It is much preferable to study them through their actions as linear mappings [see Draper and Pukelsheim (1994, p. 149)]. For three factors the matrices are listed in Figure 1.

We are now ready for the symmetrization step of the Kiefer ordering. Given an arbitrary design τ on the experimental domain $\mathcal{T}$, we define

$$\mu_2 = \frac{1}{m} \int_{\mathcal{T}} \|t\|^2 \, d\tau, \quad \mu_{22} = \frac{1}{m(m+2)} \int_{\mathcal{T}} \|t\|^4 \, d\tau, \quad \mu_{222} = \frac{1}{m(m+2)(m+4)} \int_{\mathcal{T}} \|t\|^6 \, d\tau.$$

With these values, the rotatable moment matrix $\overline{M}$ is more balanced than $M(\tau)$. It coincides with the average over the transformed moment matrices $Q_R \, M(\tau) \, Q_R'$ as R varies, $\overline{M} = \int Q_R \, M(\tau) \, Q_R' \, dR$, where the integration is relative to the Haar probability measure on the compact group $\mathrm{Orth}(m)$. Haar measure on the orthogonal group is not handled easily. Our construction circumvents evaluation of the Haar integral, and replaces it by a projection argument onto the subspace of invariant symmetric matrices [Pukelsheim (1993, pp. 348, 403); Gaffke and Heiligers (1996, p. 1158)].

In terms of the geometry of the space of symmetric $k \times k$ matrices, $\overline{M}$ is the midpoint of the convex hull of the matrices $Q_R M(\tau) Q_R'$ with $R \in \mathrm{Orth}(m)$. The Carathéodory Theorem secures the (abstract) existence of finitely many matrices $R_i \in \mathrm{Orth}(m)$, and of corresponding weights $\lambda_i > 0$ summing to one, such that $\overline{M} = \sum_i \lambda_i \, Q_{R_i} M(\tau) Q_{R_i}'$. This relation is known as matrix majorization (relative to the congruence action of the induced group $\mathcal{Q}$). It says that $\overline{M}$ is majorized by $M(\tau)$, and is denoted by

$$\overline{M} \prec M(\tau).$$

We stick to standard terminology, even though for us the emphasis is "reversed": $\overline{M}$ is superior over $M(\tau)$ since it exhibits more symmetry.

Symmetry, in the design context often called balancedness, has always been a prime attribute of good experimental designs. The other step of the Kiefer ordering concerns the usual Loewner matrix ordering. In view of the symmetrization step it suffices to study the Loewner ordering when restricted to rotatable moment matrices, a much simpler task.

$$
v_3' = \begin{array}{c} \\ 1 \end{array}
\begin{array}{c}
\begin{array}{ccccccccc} t_1t_1 & t_1t_2 & t_1t_3 & t_2t_1 & t_2t_2 & t_2t_3 & t_3t_1 & t_3t_2 & t_3t_3 \end{array} \\
\begin{pmatrix} 1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & 1 \end{pmatrix}
\end{array}
$$

$$
F_3 = \begin{array}{c}
\begin{array}{ccccccccc} t_1t_1 & t_1t_2 & t_1t_3 & t_2t_1 & t_2t_2 & t_2t_3 & t_3t_1 & t_3t_2 & t_3t_3 \end{array} \\
\begin{pmatrix}
3 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & 1 \\
\cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & 1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\
\cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\
1 & \cdot & \cdot & \cdot & 3 & \cdot & \cdot & \cdot & 1 \\
\cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & 1 & \cdot \\
\cdot & \cdot & 1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & 1 & \cdot \\
1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & 3
\end{pmatrix}
\end{array}
$$

$$
G_3' = \begin{pmatrix}
3\cdot\cdot & \cdot 1\cdot & \cdot\cdot 1 & \cdot 1\cdot & 1\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot 1 & \cdot\cdot\cdot & 1\cdot\cdot \\
\cdot 1\cdot & 1\cdot\cdot & \cdot\cdot\cdot & 1\cdot\cdot & \cdot 3\cdot & \cdot\cdot 1 & \cdot\cdot\cdot & \cdot\cdot 1 & \cdot 1\cdot \\
\cdot\cdot 1 & \cdot\cdot\cdot & 1\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot 1 & \cdot 1\cdot & 1\cdot\cdot & \cdot 1\cdot & \cdot\cdot 3
\end{pmatrix}
$$

$$
S_3 = \begin{pmatrix}
15\cdot\cdot & \cdot 3\cdot & \cdot\cdot 3 & \cdot 3\cdot & 3\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot 3 & \cdot\cdot\cdot & 3\cdot\cdot \\
\cdot 3\cdot & 3\cdot\cdot & \cdot\cdot\cdot & 3\cdot\cdot & \cdot 3\cdot & \cdot\cdot 1 & \cdot\cdot\cdot & \cdot\cdot 1 & \cdot 1\cdot \\
\cdot\cdot 3 & \cdot\cdot\cdot & 3\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot 1 & \cdot 1\cdot & 3\cdot\cdot & \cdot 1\cdot & \cdot\cdot 3 \\
\cdot 3\cdot & 3\cdot\cdot & \cdot\cdot\cdot & 3\cdot\cdot & \cdot 3\cdot & \cdot\cdot 1 & \cdot\cdot\cdot & \cdot\cdot 1 & \cdot 1\cdot \\
3\cdot\cdot & \cdot 3\cdot & \cdot\cdot 1 & \cdot 3\cdot & 3\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot 1 & \cdot\cdot\cdot & 1\cdot\cdot \\
\cdot\cdot\cdot & \cdot\cdot 1 & \cdot 1\cdot & \cdot\cdot 1 & \cdot\cdot\cdot & 1\cdot\cdot & \cdot 1\cdot & 1\cdot\cdot & \cdot\cdot\cdot \\
\cdot\cdot 3 & \cdot\cdot\cdot & 3\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot 1 & \cdot 1\cdot & 3\cdot\cdot & \cdot 1\cdot & \cdot\cdot 3 \\
\cdot\cdot\cdot & \cdot\cdot 1 & \cdot 1\cdot & \cdot\cdot 1 & \cdot\cdot\cdot & 1\cdot\cdot & \cdot 1\cdot & 1\cdot\cdot & \cdot\cdot\cdot \\
3\cdot\cdot & \cdot 1\cdot & \cdot\cdot 3 & \cdot 1\cdot & 1\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot 3 & \cdot\cdot\cdot & 3\cdot\cdot \\
\cdot 3\cdot & 3\cdot\cdot & \cdot\cdot\cdot & 3\cdot\cdot & \cdot 3\cdot & \cdot\cdot 1 & \cdot\cdot\cdot & \cdot\cdot 1 & \cdot 1\cdot \\
3\cdot\cdot & \cdot 3\cdot & \cdot\cdot 1 & \cdot 3\cdot & 3\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot 1 & \cdot\cdot\cdot & 1\cdot\cdot \\
\cdot\cdot\cdot & \cdot\cdot 1 & \cdot 1\cdot & \cdot\cdot 1 & \cdot\cdot\cdot & 1\cdot\cdot & \cdot 1\cdot & 1\cdot\cdot & \cdot\cdot\cdot \\
3\cdot\cdot & \cdot 1\cdot & \cdot\cdot 3 & \cdot 1\cdot & 1\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot 3 & \cdot\cdot\cdot & 3\cdot\cdot \\
\cdot 3\cdot & 3\cdot\cdot & \cdot\cdot\cdot & 3\cdot\cdot & \cdot 3\cdot & \cdot\cdot 1 & \cdot\cdot\cdot & \cdot\cdot 1 & \cdot 1\cdot \\
3\cdot\cdot & \cdot 3\cdot & \cdot\cdot 1 & \cdot 3\cdot & 3\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot 1 & \cdot\cdot\cdot & 1\cdot\cdot \\
\cdot\cdot\cdot & \cdot\cdot 1 & \cdot 1\cdot & \cdot\cdot 1 & \cdot\cdot\cdot & 1\cdot\cdot & \cdot 1\cdot & 1\cdot\cdot & \cdot\cdot\cdot \\
3\cdot\cdot & \cdot 1\cdot & \cdot\cdot 3 & \cdot 1\cdot & 1\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot 3 & \cdot\cdot\cdot & 3\cdot\cdot \\
\cdot 3\cdot & 3\cdot\cdot & \cdot\cdot\cdot & 3\cdot\cdot & \cdot 15\cdot & \cdot\cdot 3 & \cdot\cdot\cdot & \cdot\cdot 3 & \cdot 3\cdot \\
\cdot\cdot 1 & \cdot\cdot\cdot & 1\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot 3 & \cdot 3\cdot & 1\cdot\cdot & \cdot 3\cdot & \cdot\cdot 3 \\
\cdot\cdot\cdot & \cdot\cdot 1 & \cdot 1\cdot & \cdot\cdot 1 & \cdot\cdot\cdot & 1\cdot\cdot & \cdot 1\cdot & 1\cdot\cdot & \cdot\cdot\cdot \\
\cdot\cdot 1 & \cdot\cdot\cdot & 1\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot 3 & \cdot 3\cdot & 1\cdot\cdot & \cdot 3\cdot & \cdot\cdot 3 \\
\cdot 1\cdot & 1\cdot\cdot & \cdot\cdot\cdot & 1\cdot\cdot & \cdot 3\cdot & \cdot\cdot 3 & \cdot\cdot\cdot & \cdot\cdot 3 & \cdot 3\cdot \\
\cdot\cdot 3 & \cdot\cdot\cdot & 3\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot 3 & \cdot 3\cdot & 3\cdot\cdot & \cdot\cdot 3 & \cdot\cdot 15
\end{pmatrix}
$$

FIG. 1. *The building blocks of a third-degree rotatable moment matrix for three factors, $m = 3$. The counts reflect the moment identities $\mu_4 = 3\mu_{22}$, $\mu_{42} = 3\mu_{222}$, and $\mu_6 = 15\mu_{222}$. Dots indicate zeroes.*

5. Loewner enlargement. It is convenient to work with the uniform distribution u_ρ on the sphere $\{t \in \mathbb{R}^m : \|t\| = \rho\}$, for radius $\rho \in [0, r]$. We use the uniform distribution u_r on the boundary sphere of radius r, mixtures of u_r and center points (u_0), and mixtures of u_r and a uniform design u_ρ on an inner "nucleus" sphere of radius ρ. We call these distributions *boundary nucleus designs*, denoted by

$$\tau_{\alpha,\rho} = (1-\alpha)u_\rho + \alpha u_r.$$

Strictly speaking, these "designs" do not have a finite support and hence violate the definition of a design as given in Section 3.

However, there exist properly defined designs that have the same moments as $\tau_{\alpha,\rho}$, up to respective orders 2, 4, and 6, and they can always be taken to replace $\tau_{\alpha,\rho}$ in order to meet our definitional requirements. Moment equality is achieved by two level full factorial and fractional factorial designs and regular simplex designs for first-degree models, by central composite designs for second-degree models, and by other appropriate point sets for third-degree models (see Pukelsheim 1993, pp. 391, 402; Draper and Pukelsheim 1994, p. 156).

The following three lemmas calculate the mixing weight α and the nucleus radius ρ so that the moment matrix of $\tau_{\alpha,\rho}$ improves upon a given, general rotatable moment matrix, relative to the usual Loewner matrix ordering.

LEMMA 1. *Let $\overline{M}$ be a rotatable first-degree moment matrix. Then $M(\tau_{1,0}) \geq \overline{M}$.*

PROOF. Let μ_2 be the parameter of $\overline{M}$. From $\mu_2 = \int \|t\|^2 \, d\tau/m$ it is clear that

$$\mu_2 \in \left[0, \frac{r^2}{m}\right].$$

The uniform distribution has $\mu_2(\tau_{1,0}) = r^2/m$. Thus the difference $\delta = \mu_2(\tau_{1,0}) - \mu_2$ is nonnegative, $\delta \geq 0$, and we obtain

$$M(\tau_{1,0}) - \overline{M} = \begin{pmatrix} 0 & 0 \\ 0 & \delta I_m \end{pmatrix} \geq 0. \qquad \square$$

LEMMA 2. *Let $\overline{M}$ be a second-degree rotatable moment matrix, with parameters μ_2 and μ_{22}. Define $\alpha = m\mu_2/r^2$. Then $\alpha \in [0,1]$, and $M(\tau_{\alpha,0}) \geq \overline{M}$.*

PROOF. The range of μ_2 entails $\alpha \in [0,1]$. From $\mu_{22} = \int \|t\|^4 \, d\tau/\big(m(m+2)\big)$ we get

$$\mu_{22} \in \left[\frac{m\mu_2^2}{m+2}, \frac{r^2\mu_2}{m+2}\right],$$

where the lower bound is the Jensen inequality. We have $\mu_2(\tau_{\alpha,0}) = \alpha r^2/m = \mu_2$, and $\mu_{22}(\tau_{\alpha,0}) = \alpha r^4/\big(m(m+2)\big) = r^2\mu_2/(m+2)$. Thus the difference $\delta = \mu_{22}(\tau_{\alpha,0}) - \mu_{22}$ is nonnegative, $\delta \geq 0$, and we obtain

$$M(\tau_{\alpha,0}) - \overline{M} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \delta F_m \end{pmatrix} \geq 0. \qquad \square$$

LEMMA 3. *Let $\overline{M}$ be a third-degree rotatable moment matrix, with parameters μ_2, μ_{22}, and μ_{222}. Define $\alpha = 1$ and $\rho = 0$ when $\mu_2 = r^2/m$, and*

$$\alpha = \frac{m\mu_2 - \rho^2}{r^2 - \rho^2}, \qquad \rho^2 = \frac{r^2\mu_2 - (m+2)\mu_{22}}{r^2/m - \mu_2}$$

when $\mu_2 < r^2/m$. Then $\alpha \in [0,1]$ and $\rho \in [0,r)$, and $M(\tau_{\alpha,\rho}) \geq \overline{M}$.

PROOF. The case $\mu_2 = r^2/m$ forces $\mu_{22} = r^4/\big(m(m+2)\big)$, and $\mu_{222} = r^6/\big(m(m+2)(m+4)\big)$. These moments are attained by the uniform boundary design $\tau_{1,0}$. Otherwise the ranges of μ_2 and μ_{22} entail $\rho^2 \in [0, m\mu_2] \subset [0, r^2)$, and $\alpha \in [0,1]$. We easily verify $\mu_2(\tau_{\alpha,\rho}) = \mu_2$ and $\mu_{22}(\tau_{\alpha,\rho}) = \mu_{22}$, and we obtain

$$\mu_{222}(\tau_{\alpha,\rho}) = \frac{r^2\mu_{22}}{m+4} - \frac{\big(r^2\mu_2 - (m+2)\mu_{22}\big)^2}{(m+2)(m+4)(r^2/m - \mu_2)}.$$

To compare this with μ_{222}, we introduce $g(t) = (r^2, \|t\|^2)'$. Then the 2×2 matrix $(r^2 - \|t\|^2)\, g(t)g(t)'$ is nonnegative definite, and so is its mean under τ,

$$\begin{aligned}
\int_{\mathcal{T}} (r^2 - \|t\|^2) g(t)g(t)'\, d\tau &= \int_{\mathcal{T}} \begin{pmatrix} r^6 - r^4\|t\|^2 & r^4\|t\|^2 - r^2\|t\|^4 \\ r^4\|t\|^2 - r^2\|t\|^4 & r^2\|t\|^4 - \|t\|^6 \end{pmatrix} d\tau \\
&= \begin{pmatrix} r^6 - mr^4\mu_2 & mr^4\mu_2 - m(m+2)r^2\mu_{22} \\ mr^4\mu_2 - m(m+2)r^2\mu_{22} & m(m+2)r^2\mu_{22} - m(m+2)(m+4)\mu_{222} \end{pmatrix} \\
&= \begin{pmatrix} r^6 - a & a - b \\ a - b & b - c \end{pmatrix}, \quad \text{say.}
\end{aligned}$$

Hence the Schur complement $b - c - (a-b)^2/(r^6 - a)$ is nonnegative (Pukelsheim 1993, p. 75). This gives $c \leq b - (a-b)^2/(r^6 - a)$, that is, $\mu_{222} \leq \mu_{222}(\tau_{\alpha,\rho})$. Thus the difference $\delta = \mu_{222}(\tau_{\alpha,\rho}) - \mu_{222}$ is nonnegative, $\delta \geq 0$, and we obtain

$$M(\tau_{\alpha,\rho}) - \overline{M} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \delta S_m \end{pmatrix} \geq 0. \qquad \square$$

The argument to determine the upper bound of the moment μ_{222} is just another variant of the classical moment criterion [see Karlin and Studden (1966, p. 106)]. In the final section we now turn to the Kiefer ordering proper.

6. Kiefer design ordering. The Kiefer partial ordering is a two-stage ordering, reflecting an increase in symmetry by matrix majorization and a subsequent enlargement in the Loewner ordering. Under the Kiefer ordering, we call a moment matrix M more informative than an alternative moment matrix A when M is greater than or equal to some intermediate matrix F under the Loewner ordering, and F is majorized by A under the group action that leaves the problem invariant,

$$M \gg A \qquad \Longleftrightarrow \qquad M \geq F \prec A \quad \text{for some matrix } F.$$

We call two moment matrices M and A Kiefer equivalent when $M \gg A$ and $A \gg M$; we call M Kiefer better than A when $M \gg A$ without M and A being equivalent. We say that two designs τ and ξ are Kiefer equivalent when their moment matrices are Kiefer equivalent, and that τ is Kiefer better than ξ when $M(\tau)$ is Kiefer better than $M(\xi)$.

For response surface models on the ball, the Kiefer ordering refers to the induced group on the regression range, $\mathcal{Q} = \{ Q_R : R \in \mathrm{Orth}(m) \}$, acting by congruence, $M \mapsto Q_R M Q_R'$. Two moment matrices M and A are Kiefer equivalent if and only if $A = Q_R M Q_R'$ for some rotation $R \in \mathrm{Orth}(m)$ [Pukelsheim (1993, p. 356)]. In particular, if M and A are Kiefer equivalent and M is rotatable, then the two matrices must be equal. Our main results are the following three theorems, announcing minimal complete classes of moment matrices under the Kiefer ordering.

THEOREM 1. *In the first-degree model, the moment matrix $M(\tau_{1,0})$ of the uniform boundary design is Kiefer better than any other moment matrix.*

PROOF. Let τ be an arbitrary design on the experimental domain $\mathcal{T}$, with a moment matrix A other than $M(\tau_{1,0})$. From Section 4 and Lemma 1 we have $M(\tau_{1,0}) \geq \overline{M} \prec A$, that is, $M(\tau_{1,0}) \gg A$.

If $M(\tau_{1,0})$ and A are Kiefer equivalent then they are equal, contradicting the assumption that they are distinct. Hence M is Kiefer better than A. □

THEOREM 2. *In the second-degree model, an essentially complete class of designs for the Kiefer ordering is*

$$\mathcal{C} = \{ \tau_{\alpha,0} : \alpha \in [0,1] \},$$

and the set of moment matrices $M(\mathcal{C}) = \{ M(\tau_{\alpha,0}) : \alpha \in [0,1] \}$ constitutes a minimal complete class.

PROOF. Completeness of $M(\mathcal{C})$ means that for any moment matrix A not in $M(\mathcal{C})$ there is a member in $M(\mathcal{C})$ that is Kiefer better than A. From Section 4 and for the weight α from Lemma 2 we have $M(\tau_{\alpha,0}) \geq \overline{M} \prec A$, that is, $M(\tau_{\alpha,0}) \gg A$. Since the matrices in $M(\mathcal{C})$ are rotatable, Kiefer equivalence forces equality, contradicting the assumption that A does not lie in $M(\mathcal{C})$.

Minimal completeness of $M(\mathcal{C})$ means that no proper subset of $M(\mathcal{C})$ is complete. It suffices to show that, for a fixed weight β, the subclass $M(\mathcal{C}) \setminus \{ M(\tau_{\beta,0}) \}$ is not complete.

Indeed, the proof of Lemma 2 shows that $M(\tau_{\alpha,0}) \geq M(\tau_{\beta,0})$ implies $\alpha = \beta$. Under the Kiefer ordering, no moment matrix with $\alpha \neq \beta$ is comparable with $M(\tau_{\beta,0})$.

Essential completeness of $\mathcal{C}$ means that for every design ξ not in $\mathcal{C}$ there is a design $\mathcal{C}$ that is Kiefer better or Kiefer equivalent to ξ. This is implied by the minimal completeness of the moment matrix set $\mathcal{C}$. □

THEOREM 3. *In the third-degree model, an essentially complete class of designs for the Kiefer ordering is*

$$\mathcal{C} = \{\, \tau_{\alpha,\rho} : \alpha \in [0,1],\ \rho \in [0,r) \,\},$$

and the set of moment matrices $M(\mathcal{C}) = \{\, M(\tau_{\alpha,\rho}) : \alpha \in [0,1],\ \rho \in [0,r) \,\}$ *constitutes a minimal complete class.*

PROOF. The arguments parallel those of the foregoing proof. □

Theorem 1 can be paraphrased by saying that $M(\tau_{1,0})$ is Kiefer optimal, or that the minimal complete class degenerates to a one-point set. The fact that on the design level we "only" have essential completeness rather than minimal completeness is a great source of economy. It opens the way to conveniently choose from various designs sharing the same moment matrix. We have already made use of this option for the introduction of boundary nucleus designs $\tau_{\alpha,\rho}$, when we did not insist on the finite support requirement.

Theorem 1 solves the design problem for a first-degree model, Theorems 2 and 3 reduce it to a one- and two-parameter design class, respectively. The reduction in dimensionality does not depend on the number m of factors considered, and is enormous when compared to the dimensions of the spaces spanned by the regression vectors $f(t)$ in Section 3.

We wish to stress that these results on improving an *arbitrary* initial design τ relative to the Kiefer ordering are constructive. First the rotatability parameters $\mu_2 = \frac{1}{m}\int_{\mathcal{T}} \|t\|^2\, d\tau$ etc. are calculated, and then Lemmas 1–3 give the weight and the radius of the improving boundary nucleus design. This constructive path of improvement is not available within the optimality theory for scalar optimality criteria, such as D-, A-, and E-optimality.

Another convincing aspect is that the Kiefer ordering does not depend on the basis that is used to represent the regression function $f(t)$. While the approach using the Kronecker product has its merits, we could have used any other basis as well. In particular, Theorems 1–3 remain true for the Box–Hunter regression function $\widetilde{f}(t)$ that consists of a minimal set of monomials. We recall that information surfaces provide another important concept that does not depend on the choice of coordinates for the regression function, see Draper, Gaffke and Pukelsheim (1991, p. 158).

To prove coordinate invariance of the Kiefer ordering, let $\widetilde{f}(t) = Tf(t)$ be a change of basis for the regression function $f(t)$. Often T will be nonsingular, with a regular inverse T^{-1}. However, in the Kronecker approach the vectors $f(t)$ with $t \in \mathcal{T}$ span a proper subspace $\mathcal{L} \subset \mathbb{R}^k$ only. Therefore we make the weaker assumption that T is a square or rectangular matrix, chosen in such a way that T^+T projects onto the subspace $\mathcal{L}$.

A rotation of the experimental conditions, $t \mapsto Rt$, entails $\widetilde{f}(Rt) = Tf(Rt) = TQ_Rf(t) = TQ_RT^+Tf(t) = \widetilde{Q}_R\widetilde{f}(t)$. Thus, in the new coordinates, the group $\widetilde{\mathcal{Q}}$ of

induced transformations has members $\widetilde{Q}_R = TQ_RT^+$, and the moment matrices take the form $\widetilde{M} = \int \widetilde{f}(t)\widetilde{f}(t)'\,d\tau = TMT'$. We have $T^+\widetilde{M}T^{+\prime} = T^+TMT'T^{+\prime} = M$.

Let $\widetilde{A} = TAT'$ be another moment matrix, to be compared to $\widetilde{M}$. From

$$T^+\widetilde{Q}_R\widetilde{f}(t) = T^+TQ_RT^+Tf(t) = T^+Tf(Rt) = Q_Rf(t)$$

we get $T^+\widetilde{Q}_R\widetilde{A}\widetilde{Q}'_RT^{+\prime} = Q_RAQ'_R$. This justifies the equivalences

$$\widetilde{M} \gg \widetilde{A} \iff \widetilde{M} \geq \sum \lambda_i \widetilde{Q}_{R_i}\,\widetilde{A}\,\widetilde{Q}'_{R_i} \iff M \geq \sum \lambda_i\, Q_{R_i}\, A\, Q'_{R_i} \iff M \gg A.$$

Therefore two moment matrices are Kiefer comparable in the new coordinate system if and only if they are Kiefer comparable in the original coordinate system. The proof that the Kiefer ordering does not depend on the coordinate system is complete.

The eigenvalues of M are in general distinct from those of TMT'. When the Kiefer ordering is supplemented by a componentwise enlargement of eigenvalues, the results so obtained will again become basis dependent. For examples see Pukelsheim (1993, p. 403), Cheng (1995, p. 47), Draper, Heiligers and Pukelsheim (1996, p. 398). Even then we believe the Kronecker representation remains attractive, in that the ensuing moment matrices attain a pattern particularly suitable for the study of their eigenvalues.

Preservation of eigenvalues and orthogonality occurs when $TT' = I_d$ is the $d \times d$ identity matrix, in which case T is called a partial isometry. Then we have $T^+ = T'$, and $\widetilde{M} = TMT'$ has the same nonzero eigenvalues as $T'TM = T^+TM = M$. Furthermore the regression vectors $\widetilde{f}(t)$ span the full space $\mathbb{R}^d$, and the induced matrix $\widetilde{Q}_R = TQ_RT'$ is of full rank d. In this setting, if Q_R is an orthogonal matrix then so is $\widetilde{Q}_R$,

$$\widetilde{Q}'_R\widetilde{Q}_R\widetilde{f}(t) = TQ'_RT^+\widetilde{Q}_R\widetilde{f}(t) = TQ'_RQ_Rf(t) = \widetilde{f}(t) \qquad \Rightarrow \qquad \widetilde{Q}'_R\widetilde{Q}_R = I_d.$$

In general, orthogonality is *not* preserved. For instance, the group $\mathcal{Q}$ for the Kronecker approach contains only orthogonal matrices, while the group $\widetilde{\mathcal{Q}}$ for the Box-Hunter approach contains also nonorthogonal matrices.

The key point to watch out for is that the Kiefer ordering in the original system refers to the original group $\mathcal{Q}$ while, in a new coordinate system, it refers to the new group $\widetilde{\mathcal{Q}}$.

The dependence on the underlying group $\mathcal{Q}$ has repercussions for scalar optimality criteria. If a criterion ϕ is Loewner isotonic, concave, and invariant, then it is isotonic also under the Kiefer ordering, and an enlargement in the Kiefer ordering implies an enlargement as measured by ϕ. Loewner monotonicity and concavity are unambiguous notions. However, invariance crucially depends on the underlying group $\mathcal{Q}$.

In the present setting of response surface models, the Kronecker regression functions come with a group $\mathcal{Q}$ that is a subgroup of the group of orthogonal matrices on the space $\mathbb{R}^k$. Since all the matrix means ϕ_p are orthogonally invariant, they are Kiefer isotonic. Here, the Kiefer improvement of Theorems 1–3 implies an improvement as measured by the matrix means ϕ_p. In the Kronecker approach, the reduction by rotatability is supported by a formal, compelling argument of the theory.

For the Box–Hunter regression functions, the group $\widetilde{\mathcal{Q}}$ is not a subgroup of the orthogonal group. The matrix means ϕ_p are no longer Kiefer isotonic (except for the determinant

criterion). Theorems 1–3 remain true, but the transition to the criteria ϕ_p is no longer available. In the Box–Hunter approach, the restriction to rotatability cuts out the nonrotatable competitors, and can be justified only by an informal, persuasive decision of the investigator.

REFERENCES

BOX, G. E. P. AND HUNTER, J. S. (1957). Multi-factor experimental designs for exploring response surfaces. *Ann. Math. Statist.* **28** 195-241.

CHENG, C.-S. (1995). Complete class results for the moment matrices of designs over permutation-invariant sets. *Ann. Statist.* **23** 41–54.

DRAPER, N. R. AND PUKELSHEIM, F. (1994). On third order rotatability. *Metrika* **41** 137–161.

DRAPER, N. R. AND PUKELSHEIM, F. (1998). Mixture models based on homogeneous polynomials. *J. Statist. Plann. Inf.*, in press.

DRAPER, N. R., GAFFKE, N. AND PUKELSHEIM, F. (1991). First and second order rotatability of experimental designs, moment matrices, and information surfaces. *Metrika* **38** 129–161.

DRAPER, N. R., GAFFKE, N. AND PUKELSHEIM, F. (1993). Rotatability of variance surfaces and moment matrices. *J. Statist. Plann. Inf.* **36** 347–356.

DRAPER, N. R., HEILIGERS, B. AND PUKELSHEIM, F. (1996). On optimal third order rotatable designs. *Ann. Institute Statist. Math.* **48** 395–402.

GAFFKE, N. AND HEILIGERS, B. (1995). Computing optimal approximate invariant designs for cubic regression on multidimensional balls and cubes. *J. Statist. Plann. Inf.* **47** 347–376.

GAFFKE, N. AND HEILIGERS, B. (1996). Approximate designs for polynomial regression: Invariance, admissibility, and optimality. In *Handbook of Statistics, Volume 13* (S. Ghosh and C.R. Rao, eds.), 1149–1199. Elsevier, New York.

GALIL, Z. AND KIEFER, J. (1979). Extrapolation designs and Φ_p-optimum designs for cubic regression on the q-ball. *J. Statist. Plann. Inf.* **3** 27–38. [Also in *J. C. Kiefer Collected Papers III* (L. D. Brown, I. Olkin, J. Sacks, H. P. Wynn, eds.), 467–478. Springer, New York, 1995.]

HEILIGERS, B. (1991). Admissibility of experimental designs in linear regression with constant term. *J. Statist. Plann. Inf.* **28** 107–123.

KARLIN, S. AND STUDDEN, W. J. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics.* Interscience, New York.

KIEFER, J. C. (1975). Construction and optimality of generalized Youden designs. In *A Survey of Statistical Design and Linear Models* (J. N. Srivastava, ed.) 333–353. North-Holland, Amsterdam. [Also in *J. C. Kiefer Collected Papers III* (L. D. Brown, I. Olkin, J. Sacks, H. P. Wynn, eds.), 345–353. Springer, New York, 1995.]

PUKELSHEIM, F. (1993). *Optimal Design of Experiments.* Wiley, New York.

NORMAN R. DRAPER
DEPARTMENT OF STATISTICS
UNIVERSITY OF WISCONSIN–MADISON
MADISON WI 53706-1685, U.S.A.

FRIEDRICH PUKELSHEIM
INSTITUT FÜR MATHEMATIK
UNIVERSITÄT AUGSBURG
D-86135 AUGSBURG, GERMANY