Hindawi Publishing Corporation
Journal of Probability and Statistics
Volume 2011, Article ID 182049, 15 pages
doi:10.1155/2011/182049

Research Article

Determinant Efficiencies in Ill-Conditioned Models

D. R. Jensen

Department of Statistics, Virginia Polytechnic Institute, Blacksburg, VA 24061, USA

Correspondence should be addressed to D. R. Jensen, djensen@vt.edu

Received 18 May 2011; Accepted 1 August 2011

Academic Editor: Michael Lavine

The *canonical correlations* between subsets of *OLS* estimators are identified with *design linkage parameters* between their regressors. Known *collinearity indices* are extended to encompass angles between each regressor vector and remaining vectors. One such angle quantifies the collinearity of regressors with the intercept, of concern in the corruption of all estimates due to ill-conditioning. Matrix identities factorize a determinant in terms of principal subdeterminants and the canonical *Vector Alienation Coefficients* between subset estimators—by duality, the *Alienation Coefficients* between subsets of regressors. These identities figure in the study of D and D_s as determinant efficiencies for estimators and their subsets, specifically, D_s-efficiencies for the constant, linear, pure quadratic, and interactive coefficients in eight known small second-order designs. Studies on D- and D_s-efficiencies confirm that designs are seldom efficient for both. Determinant identities demonstrate the propensity for D_s-inefficient subsets to be masked through near collinearities in overall D-efficient designs.

1. Introduction

Given $\{\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}\}$ of full rank with homogeneous, uncorrelated errors, the *OLS* estimators $\widehat{\boldsymbol{\beta}}$ are unbiased with second-moment matrix $V(\widehat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$. Such moment matrices pervade experimental design, to include determinants as gauges of D- and D_s-efficiencies for estimators and their subsets. Early references trace to [1–4], and more recently to [5–10] and others. Finding D_s-efficient designs for polynomial models is considered in [11–20], for example. Studies examining the D_s-efficiencies of D-efficient designs confirm that designs are seldom efficient for both; see [13, 21–23]. From those beginnings, the study of D- and D_s-efficiencies continues apace. To wit, a recent key-word search in the *Current Index to Statistics* shows in excess of 60 listings from 2006 to 2010, and more than 100 from 2001 to 2010. Moreover, these ideas bear fruit in a widening diversity of applications as evidenced in the following.

To fix ideas, let D correspond to a polynomial P^c of degree c, namely, $g(\mu) = \sum_{i=0}^{c} \beta_i t^i$. In toxicology studies, a two-stage experiment is proffered in [24], seeking D-efficiency in estimating $k = c+1$ overall parameters at the first stage, then D_1-efficiency at the second stage in estimating a critical "threshold parameter," using quasilikelihood in nonlinear models. Coupled with this is the D_{k-1}-efficiency for the remaining $k-1$ parameters at the second stage. In related work [25], experiments with c chemicals in combination are to be examined along fixed-ratio rays. When restricted to a specified ray, the fundamental hypothesis of noninteracting factors can be rejected when higher-order polynomial terms are required in the total dose-response model $g(\mu) = \beta_0 + \beta_1 t + \sum_{i=2}^{c} \beta_i t^i$ in the linear predictor t. Here D_2 refers to $[\beta_0, \beta_1]$ and D_{c-1} to D_s-efficiency in the critical estimation of $[\beta_2, \ldots, \beta_c]$, which vanish under the conjectured additivity. Moreover, in [11, 12, 21] D refers to P^{k+1}, D_k to P^k, and D_1 to β_{k+1}, the highest-order term in P^{k+1}, for example. In short, users often are properly concerned with both D- and D_s-efficiencies, and connections between these basic criteria deserve further study, to be undertaken here.

Ill-conditioning, as near-collinearity among the columns of $\mathbf{X}$, "causes crucial elements of $\mathbf{X}'\mathbf{X}$ to be large and unstable," "creating inflated variances," and estimates that are "very sensitive to small changes in $\mathbf{X}$," having "degraded numerical accuracy;" see [26–28], for example. Diagnostics include the *condition number* $c_1(\mathbf{X}'\mathbf{X})$, the ratio of largest to smallest eigenvalues; and the *Variance Inflation Factors* $\{\mathrm{VIF}(\widehat{\beta}_j) = v_{jj} w_{jj};\ 1 \le j \le p\}$ with $\mathbf{W} = \mathbf{X}'\mathbf{X}$ and $\mathbf{V} = (\mathbf{X}'\mathbf{X})^{-1}$, that is, ratios of actual (v_{jj}) to "ideal" $(1/w_{jj})$ variances had the columns of $\mathbf{X}$ been orthogonal. In models with intercept, "collinearity with the intercept can quite generally corrupt the estimates of *all parameters* in the model whether or not the intercept is itself of interest and whether or not the data have been (mean) centered," as noted in [29].

To the foregoing list of ills from ill-conditioning, we add that not only are designs seldom efficient for both, but D_s-inefficient estimators may be masked in overall D-efficient designs, and conversely. This masking may be quantified in terms of structural dependencies, specifically, through determinant identities linking D- and D_s-efficiencies to various gages of nonorthogonality of the data. The latter include nonvanishing inner products between columns of regressors, Hotelling's [30] canonical correlations among *OLS* solutions, and VIFs. An outline follows.

Section 2 contains supporting material. Details surrounding collinearity diagnostics are topics in Section 3, to include duality of angles between subspaces of the design and parameter spaces, and their connections to VIFs. Section 4 develops basic determinant identities and inequalities of independent interest. Section 5 revisits eight small second-order designs with regard to D_s-efficiencies in estimating the constant, linear, pure quadratic, and interactive coefficients, to include the masking of inefficient estimators. Though in wide usage, with no apparent accounting for collinearity, these designs are seen to exhibit varying degrees of collinearity of regressors with the constant. Since computations proceed from the design matrix itself, an advantage is that prospective designs can be evaluated beforehand in regard to issues studied here, before committing to an actual experiment. Section 6 concludes with a brief summary.

2. Preliminaries

2.1. Notation

Spaces of note include $\mathbb{R}^k$ as Euclidean k-space; $\mathbb{R}_+^k$ as its positive orthant; $\mathbb{F}_{n\times k}$ as the real matrices of order $(n \times k)$; $\mathbb{S}_k$ as the $(k \times k)$ real symmetric matrices; and $\mathbb{S}_k^+$ as their positive definite varieties. The transpose, inverse, trace, and determinant of $\mathbf{A} \in \mathbb{S}_k^+$ are $\mathbf{A}', \mathbf{A}^{-1}$, $\operatorname{tr}(\mathbf{A})$, and $|\mathbf{A}|$; and $\mathbf{A}^{1/2}$ is its spectral square root. Special arrays include the unit vector $\mathbf{1}_n' = [1, 1, \ldots, 1] \in \mathbb{R}^n$, the identity $\mathbf{I}_n$ of order $(n \times n)$, the block-diagonal matrix $\operatorname{Diag}(\mathbf{A}_1, \mathbf{A}_2) \in \mathbb{S}_k$, the idempotent form $\mathbf{B}_n = (\mathbf{I}_n - n^{-1}\mathbf{1}_n\mathbf{1}_n')$, and $\mathcal{O}(k)$ as the real orthogonal group of $(k \times k)$ matrices. For $\mathbf{X}(n \times p)$ of rank $p \leq n$, designate a pseudoinverse as $\mathbf{X}^\dagger$, its ordered singular values as $\sigma(\mathbf{X}) = \{\xi_1 \geq \xi_2 \geq \cdots \geq \xi_p > 0\}$, and by $\mathcal{S}_p(\mathbf{X}) \subset \mathbb{R}^n$, the linear span of columns of $\mathbf{X}$. Its condition number is $c_2(\mathbf{X}) = \xi_1/\xi_p$, specifically, $c_2(\mathbf{X}) = [c_1(\mathbf{X}'\mathbf{X})]^{1/2}$.

The mean, dispersion matrix, and generalized variance for a random $\mathbf{U} \in \mathbb{R}^k$ are designated as $E(\mathbf{U}) = \boldsymbol{\mu} \in \mathbb{R}^k, V(\mathbf{U}) = \boldsymbol{\Sigma} \in \mathbb{S}_k^+$, and $GV(\mathbf{U}) = |\boldsymbol{\Sigma}|$, respectively. To account for dimension, consider $G(\mathbf{U}) = [GV(\mathbf{U})]^{1/k} = |\boldsymbol{\Sigma}|^{1/k}$ as a function homogeneous of unit degree. The class $\mathrm{M}_0 : \{\mathbf{Y} = \beta_0\mathbf{1}_n + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}\}$, comprising models with intercept and dispersion $V(\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}_n$, is our principal focus. Unless stated otherwise, we take $\sigma^2 = 1.0$, since variance ratios are scale-invariant. A distinction is drawn between *centered* and *uncentered* VIFs, namely, VIF_cs and VIF_us, the former from columns of $\mathbf{X}$ centered to their means. The latter, designated as $\{\mathrm{VIF}_u(\widehat{\beta}_j);\ j = 0, 1, \ldots, k\}$, are diagonal elements of $(\mathbf{X}_0'\mathbf{X}_0)^{-1}$ divided by reciprocals of diagonals of $\mathbf{X}_0'\mathbf{X}_0$ itself. These are of subsequent interest. Special distributions on $\mathbb{R}_+^1$ include the Snedecor-Fisher distribution $F(\cdot; \nu_1, \nu_2, \lambda)$ having (ν_1, ν_2) degrees of freedom and noncentrality λ.

3. Collinearity Diagnostics

Ill-conditioned models $\{\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}\}$, burdened with difficulties as cited, trace to nonorthogonality among columns of $\mathbf{X}$. To examine aspects of near collinearity, we first establish duality between *design linkage parameters* among columns of $\mathbf{X}$, and collinearity among the *OLS* solutions as quantified by Hotelling's [30] *canonical correlations*.

3.1. Duality Results

Partition a generic $\mathbf{X} \in \mathbb{F}_{n \times p}$ as $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2]$ with $\{\mathbf{X}, \mathbf{X}_1, \mathbf{X}_2\}$ of orders $\{(n \times p), (n \times r), (n \times s)\}$, respectively, having ranks $\{p, r, s\}$ such that $r \leq s$ and $r + s = p < n$. Accordingly, write $\{\mathbf{Y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\epsilon}\}$, taking $\boldsymbol{\beta}' = [\boldsymbol{\beta}_1', \boldsymbol{\beta}_2']$, and denoting by $\mathcal{S}_p(\mathbf{X}_1)$ and $\mathcal{S}_p(\mathbf{X}_2)$, the subspaces of $\mathbb{R}^n$ spanned by columns of $\mathbf{X}_1$ and $\mathbf{X}_2$. We seek a canonical form preserving these subspaces and linkage between $(\mathbf{X}_1, \mathbf{X}_2)$, a geometric concept independent of bases for representing $\mathcal{S}_p(\mathbf{X}_1)$ and $\mathcal{S}_p(\mathbf{X}_2)$. Accordingly, let $\mathbf{G}_1 = (\mathbf{X}_1'\mathbf{X}_1)^{-1/2}\mathbf{P}$ and $\mathbf{G}_2 = (\mathbf{X}_2'\mathbf{X}_2)^{-1/2}\mathbf{Q}$, with $\mathbf{P} \in \mathcal{O}(r)$ and $\mathbf{Q} \in \mathcal{O}(s)$ to be stipulated. The original model becomes $\{\mathbf{Y} = \mathbf{Z}_1\boldsymbol{\alpha}_1 + \mathbf{Z}_2\boldsymbol{\alpha}_2 + \boldsymbol{\epsilon}\}$ with $\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2]$ and $\boldsymbol{\alpha}' = [\boldsymbol{\alpha}_1', \boldsymbol{\alpha}_2']$, such that $\mathbf{Z}_1 = \mathbf{X}_1\mathbf{G}_1$, $\mathbf{Z}_2 = \mathbf{X}_2\mathbf{G}_2$, $\boldsymbol{\alpha}_1 = \mathbf{G}_1^{-1}\boldsymbol{\beta}_1$, and $\boldsymbol{\alpha}_2 = \mathbf{G}_2^{-1}\boldsymbol{\beta}_2$.

Following [31], cosines of angles between $\mathcal{S}_p(\mathbf{X}_1)$ and $\mathcal{S}_p(\mathbf{X}_2)$ are found as singular values generated by $(\mathbf{X}_1, \mathbf{X}_2)$, to be designated as *design linkage parameters* $\{\delta_1, \ldots, \delta_r\}$. To these ends, observe that $\mathbf{X}'\mathbf{X}$ in partitioned form transitions into $\mathbf{Z}'\mathbf{Z}$ through

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} \mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{X}_1 & \mathbf{X}_2'\mathbf{X}_2 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{G}_1'\mathbf{X}_1'\mathbf{X}_1\mathbf{G}_1 & \mathbf{G}_1'\mathbf{X}_1'\mathbf{X}_2\mathbf{G}_2 \\ \mathbf{G}_2'\mathbf{X}_2'\mathbf{X}_1\mathbf{G}_1 & \mathbf{G}_2'\mathbf{X}_2'\mathbf{X}_2\mathbf{G}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_r & \mathbf{P}'\mathbf{R}\mathbf{Q} \\ \mathbf{Q}'\mathbf{R}'\mathbf{P} & \mathbf{I}_s \end{bmatrix} = \begin{bmatrix} \mathbf{I}_r & \mathbf{D} \\ \mathbf{D}' & \mathbf{I}_s \end{bmatrix} = \mathbf{Z}'\mathbf{Z}. \tag{3.1}$$

Here $\mathbf{R} = (\mathbf{X}_1'\mathbf{X}_1)^{-1/2}\mathbf{X}_1'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-1/2}$; its singular decomposition is $\mathbf{R} = \mathbf{P}\mathbf{D}\mathbf{Q}'$, where $\mathbf{D} = [\mathbf{D}_\delta, \mathbf{0}]$; and elements of $\mathbf{D}_\delta = \operatorname{Diag}(\delta_1, \ldots, \delta_r)$ comprise the singular values of $\mathbf{R}$. In particular,

$\{\phi_j = \arccos(\delta_j);\ 1 \le j \le r\}$ defines the *design linkage angles* between $\mathcal{S}_p(\mathbf{X}_1)$ and $\mathcal{S}_p(\mathbf{X}_2)$ as subspaces of $\mathbb{R}^n$.

To continue, partition $V(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\Sigma} = [\boldsymbol{\Sigma}_{ij}]$ conformably with $\boldsymbol{\beta}' = [\boldsymbol{\beta}_1', \boldsymbol{\beta}_2']$, $\boldsymbol{\beta}_1 \in \mathbb{R}^r$, $\boldsymbol{\beta}_2 \in \mathbb{R}^s$; designate their inner product space as $(\mathbb{R}^r \oplus \mathbb{R}^s, (\cdot,\cdot)_{\boldsymbol{\Sigma}})$, where $\mathbb{R}^r \oplus \mathbb{R}^s$ is the direct sum and $(\cdot,\cdot)_{\boldsymbol{\Sigma}}$ their inner product, as in Eaton [32, page 409]. Denote by $\{\rho_1, \ldots, \rho_r\}$ Hotelling's [30] *canonical correlations*. Then by Proposition 10.2 of [32], $\{\rho_1, \ldots, \rho_r\}$ are cosines of angles between $(\mathbb{R}^r, \mathbb{R}^s)$ as subspaces of $(\mathbb{R}^r \oplus \mathbb{R}^s, (\cdot,\cdot)_{\boldsymbol{\Sigma}})$. In keeping with earlier usage, identify $\{\mathcal{S}_p(\boldsymbol{\beta}_1), \mathcal{S}_p(\boldsymbol{\beta}_2)\}$ with $\{\mathbb{R}^r, \mathbb{R}^s\}$. As Hotelling's canonical correlations are invariant under affine transformations $\{\boldsymbol{\beta}_1 \to \mathbf{A}\boldsymbol{\beta}_1 + \mathbf{c}_1,\ \boldsymbol{\beta}_2 \to \mathbf{B}\boldsymbol{\beta}_2 + \mathbf{c}_2\}$, parameters may be redefined linearly, preserving subspaces, thus leaving the canonical correlations invariant. Retracing steps leading to the canonical design model embodied in (3.1), but now to preserve $\{\mathcal{S}_p(\boldsymbol{\beta}_1), \mathcal{S}_p(\boldsymbol{\beta}_2)\}$, it thus suffices to begin with the canonical model $\{\mathbf{Y} = \mathbf{Z}_1\boldsymbol{\alpha}_1 + \mathbf{Z}_2\boldsymbol{\alpha}_2 + \boldsymbol{\epsilon}\}$, where $V(\widehat{\boldsymbol{\alpha}}) = \sigma^2(\mathbf{Z}'\mathbf{Z})^{-1}$ with $\mathbf{Z}'\mathbf{Z}$ as the rightmost matrix of (3.1).

We next establish connections between the design linkage parameters $\mathbf{D}_\delta$ from (3.1), and the corresponding canonical correlations $\mathbf{D}_\rho = \operatorname{Diag}(\rho_1, \ldots, \rho_r)$, as derived eventually from $\boldsymbol{\Sigma} = (\mathbf{X}'\mathbf{X})^{-1}$. A critical duality result is encoded in the following.

Theorem 3.1. *Consider the* design linkage parameters $\mathbf{D}_\delta$ *between* $\{\mathcal{S}_p(\mathbf{X}_1), \mathcal{S}_p(\mathbf{X}_2)\}$ *as subspaces of* $\mathbb{R}^n$ *and Hotelling's [30]* canonical correlations $\mathbf{D}_\rho$ *between* $\{\mathcal{S}_p(\boldsymbol{\beta}_1), \mathcal{S}_p(\boldsymbol{\beta}_2)\}$ *as subspaces of* $(\mathbb{R}^r \oplus \mathbb{R}^s, (\cdot,\cdot)_{\boldsymbol{\Sigma}})$. *Then* $\mathbf{D}_\delta$ *and* $\mathbf{D}_\rho$ *coincide.*

Proof. In view of invariance of $\{\rho_1, \ldots, \rho_r\}$ under nonsingular linear transformations of $\boldsymbol{\beta}_1 \in \mathbb{R}^r$ and of $\boldsymbol{\beta}_2 \in \mathbb{R}^s$, canonical correlations between $(\widehat{\boldsymbol{\beta}}_1, \widehat{\boldsymbol{\beta}}_2)$ proceed as in expression (3.1), but beginning instead on the left with $V(\widehat{\boldsymbol{\alpha}}) = (\mathbf{Z}'\mathbf{Z})^{-1}$ *in lieu of* $\mathbf{X}'\mathbf{X}$. Specifically, with $\mathbf{D} = [\mathbf{D}_\delta, \mathbf{0}]$, and using rules for block-partitioned inverses, we have

$$(\mathbf{Z}'\mathbf{Z})^{-1} = \begin{bmatrix} \mathbf{I}_r & [\mathbf{D}_\delta, \mathbf{0}] \\ [\mathbf{D}_\delta, \mathbf{0}]' & \mathbf{I}_s \end{bmatrix}^{-1} = \begin{bmatrix} \left(\mathbf{I}_r - \mathbf{D}_\delta^2\right)^{-1} & -\mathbf{D}(\mathbf{I}_s - \mathbf{D}_0)^{-1} \\ -\mathbf{D}'\left(\mathbf{I}_r - \mathbf{D}_\delta^2\right)^{-1} & (\mathbf{I}_s - \mathbf{D}_0)^{-1} \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{I}_r & [\mathbf{D}_\delta, \mathbf{0}] \\ [\mathbf{D}_\delta, \mathbf{0}]' & \mathbf{I}_s \end{bmatrix}, \tag{3.2}$$

where equality at the first step follows using $\mathbf{D}\mathbf{D}' = \mathbf{D}_\delta^2$ and $\mathbf{D}'\mathbf{D} = \operatorname{Diag}(\mathbf{D}_\delta^2, \mathbf{0}) = \mathbf{D}_0$. The succeeding step utilizes the factors $(\mathbf{I}_r - \mathbf{D}_\delta^2)^{1/2}$ and $(\mathbf{I}_s - \mathbf{D}_0)^{1/2}$, taking the principal diagonal blocks of $(\mathbf{Z}'\mathbf{Z})^{-1}$ into $(\mathbf{I}_r, \mathbf{I}_s)$ as in the rightmost matrix of (3.2), and its off-diagonal block from

$$\left(\mathbf{I}_r - \mathbf{D}_\delta^2\right)^{1/2} [\mathbf{D}_\delta, \mathbf{0}] (\mathbf{I}_s - \mathbf{D}_0)^{-1} (\mathbf{I}_s - \mathbf{D}_0)^{1/2} = [\mathbf{D}_\delta, \mathbf{0}], \tag{3.3}$$

since diagonal matrices commute. But the off-diagonal block is precisely $[\mathbf{D}_\rho, \mathbf{0}]$, the canonical correlations between $(\widehat{\boldsymbol{\beta}}_1, \widehat{\boldsymbol{\beta}}_2)$, to complete our proof. □

For subsequent reference, designate $\boldsymbol{\delta}(\mathbf{X}_1 : \mathbf{X}_2) = (\delta_1, \ldots, \delta_r)$ and $\boldsymbol{\rho}(\widehat{\boldsymbol{\beta}}_1 : \widehat{\boldsymbol{\beta}}_2) = (\rho_1, \ldots, \rho_r)$. Moreover, the foregoing analysis applies for models $\mathbf{X}_0 = [\mathbf{1}_n, \mathbf{X}]$ in M_0, where $r = 1$ and $s = k$ as partitioned. In short, we have the following equivalences.

Corollary 3.2. *(i) Consider the design linkage parameters* $\{\cos(\phi_j) = \delta_j;\ 1 \le j \le r\}$*, gaging collinearity between* $\{\mathcal{S}_p(\mathbf{X}_1), \mathcal{S}_p(\mathbf{X}_2)\}$ *as subspaces of* $\mathbb{R}^n$ *and the canonical correlations* $\{\cos(\phi_j) = \rho_j;\ 1 \le j \le r\}$*, between* $\{\mathcal{S}_p(\boldsymbol{\beta}_1), \mathcal{S}_p(\boldsymbol{\beta}_2)\}$ *as subspaces of* $(\mathbb{R}^r \oplus \mathbb{R}^s, (\cdot,\cdot)_{\boldsymbol{\Sigma}})$*. Then angles between these pairs of subspaces correspond one-to-one, that is,* $\{\phi_j = \arccos(\text{ffi}_\text{j}) = \arccos(\text{æ}_\text{j});\ 1 \le \text{j} \le \text{r}\}$.

(ii) For models $\mathbf{X}_0 = [\mathbf{1}_n, \mathbf{X}]$ *in* M_0*, the element* $\boldsymbol{\delta}(\mathbf{1}_n : \mathbf{X}) = \delta_1$ *generates the angle* $\cos(\phi_1) = \delta_1$ *between the regressor vectors and the constant vector. Equivalently, this is given by* $\cos(\phi_1) = \rho_1 = \boldsymbol{\rho}(\widehat{\beta}_0 : \widehat{\boldsymbol{\beta}})$ *from duality.*

3.2. Collinearity Indices

Stewart [33] reexamined numerical aspects of ill-conditioning, to the following effects for $\mathbf{X}_0 = [\mathbf{1}_n, \mathbf{X}]$. Taking $\mathbf{X}_0^{\dagger} = (\mathbf{X}_0'\mathbf{X}_0)^{-1}\mathbf{X}_0'$ as the pseudoinverse of note, and letting $\mathbf{x}_j^{\dagger}$ be its jth row, each *collinearity index* in the collection

$$\left\{\kappa_j = \|\mathbf{x}_j\| \cdot \left\|\mathbf{x}_j^{\dagger}\right\|;\ j = 0, 1, \ldots, k\right\} \tag{3.4}$$

is constructed to be scale-invariant. Clearly $\|\mathbf{x}_j^{\dagger}\|^2$ is found along the principal diagonal of $[(\mathbf{X}_0^{\dagger})(\mathbf{X}_0^{\dagger})'] = (\mathbf{X}_0'\mathbf{X}_0)^{-1}$. In addition, the conventional VIF_us are squares of the collinearity indices, that is, $\{\text{VIF}_u(\widehat{\beta}_j) = \kappa_j^2;\ \ j = 0, 1, \ldots, k\}$. In particular, since $\mathbf{x}_0 = \mathbf{1}_n$ in $\mathbf{X}_0$, we have $\kappa_0^2 = n\|\mathbf{x}_0^{\dagger}\|^2$.

Transcending Stewart's analysis, we connect his collinearity indices to angles between subspaces as follows. Choose a typical $\mathbf{x}_j$ in $\mathbf{X}_0$; rearrange $\mathbf{X}_0$ as $[\mathbf{x}_j, \mathbf{X}_{[j]}]$ and similarly $\boldsymbol{\beta}'$ as $[\beta_j, \boldsymbol{\beta}_{[j]}']$; and seek elements of

$$\left\{\mathbf{Q}_j'(\mathbf{X}_0'\mathbf{X}_0)^{-1}\mathbf{Q}_j = \begin{bmatrix} \mathbf{x}_j'\mathbf{x}_j & \mathbf{x}_j'\mathbf{X}_{[j]} \\ \mathbf{X}_{[j]}'\mathbf{x}_j & \mathbf{X}_{[j]}'\mathbf{X}_{[j]} \end{bmatrix}^{-1};\ j = 0, 1, \ldots, k\right\} \tag{3.5}$$

as reordered by each permutation matrix $\mathbf{Q}_j$. From the clockwise rule, the $(1,1)$ element of each inverse is

$$\left\{\left[\mathbf{x}_j'(\mathbf{I}_n - \mathbf{P}_j)\mathbf{x}_j\right]^{-1} = \left[\mathbf{x}_j'\mathbf{x}_j - \mathbf{x}_j'\mathbf{P}_j\mathbf{x}_j\right]^{-1} = \left\|\mathbf{x}_j^{\dagger}\right\|^2;\ j = 0, 1, \ldots, k\right\}, \tag{3.6}$$

where $\mathbf{P}_j = \mathbf{X}_{[j]}[\mathbf{X}_{[j]}'\mathbf{X}_{[j]}]^{-1}\mathbf{X}_{[j]}'$ is the projection operator onto the subspace $\mathcal{S}_p(\mathbf{X}_{[j]}) \subset \mathbb{R}^n$. These relationships in turn enable us to connect $\{\kappa_j^2;\ j = 0, 1, \ldots, k\}$ to the geometry of ill-conditioning as follows.

Theorem 3.3. *For models in* M_0*, let* $\{\text{VIF}_u(\widehat{\beta}_j) = \kappa_j^2;\ j = 0, 1, \ldots, k\}$ *be conventional* VIF_u*s in terms of Stewart's collinearity indices. These in turn quantify collinearities between subspaces through angles (in* deg*) as follows.*

(i) *Angles between* $[\mathbf{x}_j, \mathbf{X}_{[j]}]$ *are given by* $\phi_j = \arccos[(1 - 1/\kappa_j^2)^{1/2}]$*, in succession for* $\{j = 0, 1, \ldots, k\}$.

(ii) *Equivalently,* $\{\kappa_j^2 = 1/[1-\boldsymbol{\delta}^2(\mathbf{x}_j : \mathbf{X}_{[j]})] = 1/[1-\boldsymbol{\rho}^2(\widehat{\beta}_j : \widehat{\boldsymbol{\beta}}_{[j]})];\ j = 0,1,\dots,k\}$.

(iii) *In particular,* $\phi_0 = \arccos[(1-1/\kappa_0^2)^{1/2}]$ *quantifies the degree of collinearity between the regressor vectors and the constant vector.*

Proof. From the geometry of the right triangle formed by $(\mathbf{x}_j, \mathbf{P}_j\mathbf{x}_j)$, the squared lengths satisfy $\|\mathbf{x}_j\|^2 = \|\mathbf{P}_j\mathbf{x}_j\|^2 + RS_j$, where $RS_j = \|\mathbf{x}_j - \mathbf{P}_j\mathbf{x}_j\|^2$ is the residual sum of squares from the projection. Accordingly, the principal angle between $(\mathbf{x}_j, \mathbf{P}_j\mathbf{x}_j)$ is given by

$$\cos(\phi_j) = \frac{\mathbf{x}_j'\mathbf{P}_j\mathbf{x}_j}{\|\mathbf{x}_j\|\cdot\|\mathbf{P}_j\mathbf{x}_j\|} = \frac{\|\mathbf{P}_j\mathbf{x}_j\|}{\|\mathbf{x}_j\|} = \left(1-\frac{RS_j}{\|\mathbf{x}_j\|^2}\right)^{1/2} = \left(1-\frac{1}{\kappa_j^2}\right)^{1/2} \tag{3.7}$$

for $\{j = 0,1,\dots,k\}$, to give conclusion (i) and conclusion (ii) by duality. Conclusion (iii) follows on specializing $(\mathbf{x}_0, \mathbf{P}_0\mathbf{x}_0)$ with $\mathbf{x}_0 = \mathbf{1}_n$ and $\mathbf{P}_0 = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, to complete our proof. □

Remark 3.4. The foregoing developments specialize from Section 3.1 in that the partition $[\mathbf{x}_j, \mathbf{X}_{[j]}]$ always has $r = 1$ and $s = k$, giving a single angle ϕ_j. Rules-of-thumb in common use for problematic VIFs include those exceeding 10, as in [34], or even 4 as in [35], for example. In angular measure, these correspond respectively to $\phi_j < 18.435$ deg and $\phi_j < 30.0$ deg.

3.3. Case Study 1

Consider the model $\mathrm{M}_0 : \{Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon_i\}$, the design $\mathbf{X}_0 = [\mathbf{1}_5, \mathbf{X}_1, \mathbf{X}_2]$ of order (5×3), and $\mathbf{X}_0'\mathbf{X}_0$ and its inverse as in

$$\mathbf{X}_0' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0.5 & 0.5 & 1 & 0 \\ -1 & 1 & 1 & 0 & 0 \end{bmatrix}, \qquad \mathbf{X}_0'\mathbf{X}_0 = \begin{bmatrix} 5 & 3 & 1 \\ 3 & 2.5 & 0 \\ 1 & 0 & 3 \end{bmatrix},$$
$$(\mathbf{X}_0'\mathbf{X}_0)^{-1} = \begin{bmatrix} 0.9375 & -1.1250 & -0.3125 \\ -1.1250 & 1.7500 & 0.3750 \\ -0.3125 & 0.3750 & 0.4375 \end{bmatrix}. \tag{3.8}$$

Note first that $\mathrm{VIF}_u(\widehat{\beta}_0) = \kappa_0^2 = 0.9375 \times 5 = 4.6875$. Next apply first principles to find

$$\boldsymbol{\delta}^2(\mathbf{1}_5 : [\mathbf{X}_1, \mathbf{X}_2]) = (5)^{-1}\ [3\ \ 1]\begin{bmatrix} 2.5 & 0 \\ 0 & 3 \end{bmatrix}^{-1}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = 0.786666,$$
$$\boldsymbol{\rho}^2\left(\widehat{\beta}_0 : \left[\widehat{\beta}_1, \widehat{\beta}_2\right]\right) = (0.9375)^{-1}\ [-1.1250\ \ -0.3125]\begin{bmatrix} 1.7500 & 0.3750 \\ 0.3750 & 0.4375 \end{bmatrix}^{-1}\begin{bmatrix} -1.1250 \\ -0.3125 \end{bmatrix} = 0.786666, \tag{3.9}$$

both equal to $1-(1/\kappa_0^2)$ as in Theorem 3.3(ii). The remaining VIF_us are found directly as $\text{VIF}_u(\hat{\beta}_1)=1.7500\times 2.5=4.3750$ and $\text{VIF}_u(\hat{\beta}_2)=0.4375\times 3=1.3125$. Using duality and earlier findings, we further compute

$$\begin{aligned}
\boldsymbol{\delta}^2(\mathbf{X}_1:[\mathbf{1}_5,\mathbf{X}_2]) &= \boldsymbol{\rho}^2\left(\hat{\beta}_1:\left[\hat{\beta}_0,\hat{\beta}_2\right]\right)=1-\left(\frac{1}{\kappa_1^2}\right)=1-\frac{1}{4.3750}=0.771429,\\
\boldsymbol{\delta}^2(\mathbf{X}_2:[\mathbf{1}_5,\mathbf{X}_1]) &= \boldsymbol{\rho}^2\left(\hat{\beta}_2:\left[\hat{\beta}_0,\hat{\beta}_1\right]\right)=1-\left(\frac{1}{\kappa_2^2}\right)=1-\frac{1}{1.3125}=0.238095,
\end{aligned}\tag{3.10}$$

thereby preempting the need to undertake singular decompositions as required heretofore.

4. Determinant Identities

4.1. Background

The generalized variance, as a design criterion for $\{\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\epsilon}\}$, rests in part on the geometry of ellipsoids of the type

$$R(\boldsymbol{\beta})=\left\{\boldsymbol{\beta}\in\mathbb{R}^k:\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\mathbf{X}'\mathbf{X}\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\le c^2\right\}.\tag{4.1}$$

Choices for c^2 in common usage give first (i) a confidence region for $\boldsymbol{\beta}$, whose normal-theory confidence coefficient is $1-\alpha$ on taking $c^2=S^2c_\alpha^2$, with S^2 as the residual mean square and c_α^2 the $100(1-\alpha)$ percentage point of $F(\cdot;k,n-k)$; and otherwise admitting a lower Chebychev bound as in [36, page 92]. The alternative choice $c^2=k+2$ gives (ii) Cramér's [37] *ellipsoid of concentration* for $\hat{\boldsymbol{\beta}}$, that is, the measure uniform over $R(\boldsymbol{\beta})$ having the same mean and dispersion matrix as $\hat{\boldsymbol{\beta}}$. The generalized variance $GV(\hat{\boldsymbol{\beta}})=|V(\hat{\boldsymbol{\beta}})|$ is proportional to the squared volumes of these ellipsoids, smaller volumes reflecting tighter concentrations.

4.2. Factorizations

To continue, let some $\mathbf{T}(\mathbf{Y})=\hat{\boldsymbol{\theta}}\in\mathbb{R}^k$ be random having $E(\hat{\boldsymbol{\theta}})=\boldsymbol{\theta}$ and $V(\hat{\boldsymbol{\theta}})=\boldsymbol{\Sigma}\in\mathbb{S}_k^+$; partition $\boldsymbol{\theta}'=[\boldsymbol{\theta}_1',\boldsymbol{\theta}_2']$ and $\boldsymbol{\Sigma}=[\boldsymbol{\Sigma}_{ij}]$ conformably, with $\boldsymbol{\theta}_1\in\mathbb{R}^r$ and $\boldsymbol{\theta}_2\in\mathbb{R}^s$ such that $r\le s$ and $r+s=k$; and let $G(\hat{\boldsymbol{\theta}})=|\boldsymbol{\Sigma}|^{1/k}$. The canonical correlations [30], as singular values of $\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1/2}$, are now to be designated as $\boldsymbol{\rho}(1:2)=[\rho_1,\ldots,\rho_r]$, *in lieu of* $\boldsymbol{\rho}(\hat{\boldsymbol{\beta}}_1:\hat{\boldsymbol{\beta}}_2)$, and to be ordered as $\{\rho_1\ge\rho_2\ge\cdots\ge\rho_r\ge 0\}$. Moreover, the quantity $\boldsymbol{\gamma}(1:2)=\Pi_{i=1}^r(1-\rho_i^2)$ is the *Vector Alienation Coefficient* of Hotelling [30]. The factorization $|\boldsymbol{\Sigma}|=|\boldsymbol{\Sigma}_{11}||\boldsymbol{\Sigma}_{22}|$ for $\boldsymbol{\Sigma}=\text{Diag}(\boldsymbol{\Sigma}_{11},\boldsymbol{\Sigma}_{22})$ extends directly as an upper bound for any $\boldsymbol{\Sigma}=[\boldsymbol{\Sigma}_{ij}]$, with further ramifications as follows.

Theorem 4.1. *Consider* $\hat{\boldsymbol{\theta}}'=[\hat{\boldsymbol{\theta}}_1',\hat{\boldsymbol{\theta}}_2']\in\mathbb{R}^k$ *having* $E(\hat{\boldsymbol{\theta}}')=[\boldsymbol{\theta}_1',\boldsymbol{\theta}_2']$ *and* $V(\hat{\boldsymbol{\theta}})=[\boldsymbol{\Sigma}_{ij}]$, *such that* $\boldsymbol{\theta}_1\in\mathbb{R}^r$ *and* $\boldsymbol{\theta}_2\in\mathbb{R}^s$ *with* $r\le s$ *and* $r+s=k$.

(i) *The determinant of* $\boldsymbol{\Sigma}=[\boldsymbol{\Sigma}_{ij}]$ *admits the factorization*

$$|\boldsymbol{\Sigma}|=|\boldsymbol{\Sigma}_{11}||\boldsymbol{\Sigma}_{22}|\boldsymbol{\gamma}(1:2)\tag{4.2}$$

such that $|\boldsymbol{\Sigma}|\le|\boldsymbol{\Sigma}_{11}||\boldsymbol{\Sigma}_{22}|$ *and* $\boldsymbol{\gamma}(1:2)=\Pi_{i=1}^r(1-\rho_i^2)\le 1$.

(ii) *If* $\boldsymbol{\Sigma} = \mathrm{Diag}(\boldsymbol{\Sigma}_{11}, \boldsymbol{\Sigma}_{22})$*, then* $G(\widehat{\boldsymbol{\theta}})$ *is the geometric mean* $G(\widehat{\boldsymbol{\theta}}) = [G(\widehat{\boldsymbol{\theta}}_1)]^{r/k}[G(\widehat{\boldsymbol{\theta}}_2)]^{s/k}$ *of the quantities* $G(\widehat{\boldsymbol{\theta}}_1)$ *and* $G(\widehat{\boldsymbol{\theta}}_2)$*.*

(iii) *Generally, for any* $\boldsymbol{\Sigma}$*, the quantity* $G(\widehat{\boldsymbol{\theta}})$ *becomes*

$$G\left(\widehat{\boldsymbol{\theta}}\right) = \left[G\left(\widehat{\boldsymbol{\theta}}_1\right)\right]^{r/k}\left[G\left(\widehat{\boldsymbol{\theta}}_2\right)\right]^{s/k}[\boldsymbol{\gamma}(1:2)]^{1/k} \tag{4.3}$$

in terms of $\{G(\widehat{\boldsymbol{\theta}}_1), G(\widehat{\boldsymbol{\theta}}_2), \boldsymbol{\gamma}(1:2)\}$*.*

(iv) *If* $\widehat{\boldsymbol{\theta}} = [\widehat{\boldsymbol{\theta}}_1', \widehat{\boldsymbol{\theta}}_2', \widehat{\boldsymbol{\theta}}_3]'$ *and* $\boldsymbol{\Sigma} = [\boldsymbol{\Sigma}_{ij};\ 1 \le i, j \le 3]$ *are partitioned conformably, with* $\widehat{\boldsymbol{\theta}}_1 \in \mathbb{R}^r$*,* $\widehat{\boldsymbol{\theta}}_2 \in \mathbb{R}^s$*, and* $\widehat{\boldsymbol{\theta}}_3 \in \mathbb{R}^t$*, such that* $r + s + t = k$*, then* $G(\widehat{\boldsymbol{\theta}})$ *admits the factorization*

$$G\left(\widehat{\boldsymbol{\theta}}\right) = \left[G\left(\widehat{\boldsymbol{\theta}}_1\right)\right]^{r/k}\left[G\left(\widehat{\boldsymbol{\theta}}_2\right)\right]^{s/k}\left[G\left(\widehat{\boldsymbol{\theta}}_3\right)\right]^{t/k}[\boldsymbol{\gamma}(1:23)\boldsymbol{\gamma}(2:3)]^{1/k}, \tag{4.4}$$

with $\boldsymbol{\gamma}(1:23)$ *and* $\boldsymbol{\gamma}(2:3)$ *as the Vector Alienation Coefficients between* $\{\widehat{\boldsymbol{\theta}}_1, [\widehat{\boldsymbol{\theta}}_2, \widehat{\boldsymbol{\theta}}_3]\}$ *and between* $\{\widehat{\boldsymbol{\theta}}_2, \widehat{\boldsymbol{\theta}}_3\}$*, respectively.*

Proof. As in Section 3.1 with $\mathbf{R} = \boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1/2} = \mathbf{PDQ}'$ and $\mathbf{D} = [\mathbf{D}_\rho, \mathbf{0}]$, we have

$$\begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{W}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_2 \end{bmatrix}\begin{bmatrix} \mathbf{I}_r & \mathbf{D} \\ \mathbf{D}' & \mathbf{I}_s \end{bmatrix}\begin{bmatrix} \mathbf{W}_1' & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_2' \end{bmatrix} \tag{4.5}$$

with $\mathbf{W}_1 = \boldsymbol{\Sigma}_{11}^{1/2}\mathbf{P}$ and $\mathbf{W}_2 = \boldsymbol{\Sigma}_{22}^{1/2}\mathbf{Q}$. The middle factor on the right has determinant $|\mathbf{I}_r - \mathbf{DD}'| = \Pi_{i=1}^r(1-\rho_i^2)$ from the clockwise rule, so that $|\boldsymbol{\Sigma}| = |\boldsymbol{\Sigma}_{11}||\boldsymbol{\Sigma}_{22}|\Pi_{i=1}^r(1-\rho_i^2)$ to give conclusion (i). Conclusion (ii) follows directly from $G(\widehat{\boldsymbol{\theta}}) = [GV(\widehat{\boldsymbol{\theta}})]^{1/k}$, and conclusion (iii) on combining (i) and (ii). Conclusion (iv) now follows on applying (iii) twice, first on partitioning $\widehat{\boldsymbol{\theta}}$ into $\{\widehat{\boldsymbol{\theta}}_1, [\widehat{\boldsymbol{\theta}}_2, \widehat{\boldsymbol{\theta}}_3]\}$, whose canonical correlations are $\boldsymbol{\rho}(1:23)$, then $[\widehat{\boldsymbol{\theta}}_2, \widehat{\boldsymbol{\theta}}_3]$ into $\{\widehat{\boldsymbol{\theta}}_2, \widehat{\boldsymbol{\theta}}_3\}$ having canonical correlations $\boldsymbol{\rho}(2:3)$, to complete our proof. □

Remark 4.2. In short, Theorem 4.1 links determinants and principal subdeterminants precisely through angles between subspaces. Moreover, arguments leading to conclusion (iv) may be iterated recursively to achieve a hierarchical decomposition for four or more factors, as in the following with $k = r + s + t + v$, namely,

$$G\left(\widehat{\boldsymbol{\theta}}\right) = \left[G\left(\widehat{\boldsymbol{\theta}}_1\right)\right]^{r/k}\left[G\left(\widehat{\boldsymbol{\theta}}_2\right)\right]^{s/k}\left[G\left(\widehat{\boldsymbol{\theta}}_3\right)\right]^{t/k}\left[G\left(\widehat{\boldsymbol{\theta}}_4\right)\right]^{v/k}[\boldsymbol{\gamma}(1:234)\boldsymbol{\gamma}(2:34)\boldsymbol{\gamma}(3:4)]^{1/k}. \tag{4.6}$$

Remark 4.3. Hotelling's [30] Vector Alienation Coefficient $\boldsymbol{\gamma}(1:2) = \Pi_{i=1}^r(1-\rho_i^2)$ is a composite index of linkage between $\{\mathcal{S}_p(\boldsymbol{\beta}_1), \mathcal{S}_p(\boldsymbol{\beta}_2)\}$ as subspaces of $(\mathbb{R}^r \oplus \mathbb{R}^s, (\cdot,\cdot)_{\boldsymbol{\Sigma}})$, decreasing in each $\{\rho_i^2;\ 1 \le i \le r\}$. Equivalently, duality asserts that $\boldsymbol{\gamma}(1:2) = \Pi_{i=1}^r(1-\delta_i^2)$ is the identical composite index of linkage between $\{\mathcal{S}_p(\mathbf{X}_1), \mathcal{S}_p(\mathbf{X}_2)\}$ as subspaces of $\mathbb{R}^n$.

Theorem 4.1 anticipates that D_s-inefficient subset estimators may be masked in a design exhibiting good overall D-efficiency. Conversely, a D_s-inefficient subset may contraindicate, incorrectly, the overall D-efficiency of a design. Details are provided in case studies to follow.

5. Case Studies

5.1. The Setting

Our tools are informative in input-output studies. In particular, specify $\{\mathbf{Y} = \mathbf{X}_0\boldsymbol{\beta} + \boldsymbol{\epsilon}\}$ as a second-order model $Y(x_1, x_2, x_3)$ in three regressors and $p = 10$ parameters, namely,

$$\Big\{Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_{11} x_{i1}^2 + \boldsymbol{\beta}_{22} x_{i2}^2 + \beta_{33} x_{i3}^2 \\ +\beta_{12} x_{i1} x_{i2} + \beta_{13} x_{i1} x_{i3} + \beta_{23} x_{i2} x_{i3} + \epsilon_i;\ i = 1, 2, \ldots, n\Big\}. \tag{5.1}$$

Next partition $\boldsymbol{\beta}' = [\beta_0, \boldsymbol{\beta}'_L, \boldsymbol{\beta}'_Q, \boldsymbol{\beta}'_I]$ with $\boldsymbol{\beta}'_L = [\beta_1, \beta_2, \beta_3]$ as slopes; $\boldsymbol{\beta}'_Q = [\beta_{11}, \beta_{22}, \beta_{33}]$ as pure quadratic terms reflecting diminishing (−) or increasing (+) returns to inputs; and $\boldsymbol{\beta}'_I = [\beta_{12}, \beta_{13}, \beta_{23}]$ as interactive terms reflecting synergistic (+) or antagonistic (−) effects for pairs of regressors in combination. Further let $\boldsymbol{\beta}'_M = [\boldsymbol{\beta}'_L, \boldsymbol{\beta}'_Q, \boldsymbol{\beta}'_I]$ exclusive of β_0, the latter a base line for $Y(0, 0, 0)$. We proceed under conventional homogeneous and uncorrelated errors, the minimizing solution $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'_0\mathbf{X}_0)^{-1}\mathbf{X}'_0\mathbf{Y}$ being unbiased with $V(\widehat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'_0\mathbf{X}_0)^{-1}$. We take σ^2, although unknown, to reflect natural variability in experimental materials and protocol, and thus applicable in a given setting independently of the choice of design. Accordingly, for present purposes we may standardize to $\sigma^2 = 1.0$ for reasons cited earlier.

5.2. The Designs

Early polynomial response designs made use of factorial experiments, setting levels as needed to meet the required degree. For example, the second-order model (5.1) in three regressors would require $3^3 = 27$ runs. However, in the early 1950s such designs were seen to be excessive, in carrying redundant interactions beyond the pairs required in the model (5.1). In industrial and other settings where parsimony is desired, several small second-order designs have evolved, often on appending a few additional runs to two-level factorials or fractions thereof.

Eight such small designs of note here are the hybrids (H310, H311B) of [38], the small composite SCD [39], the BBD [40], the central composite rotatable design CCD [41], and designs ND [42], HD [43], and BDD [44]. The designs [H310, H311B, SCD, BBD, CCD, ND, HD, BDD] have numbers of runs as [11, 11, 11, 13, 15, 11, 11, 11], respectively. These follow on adding a center run to all but design ND, rendering all as unsaturated having at least one degree of freedom for error. Specifically, the design ND of [42] already has 11 runs and is unsaturated. All designs have been scaled to span the same range for each regressor; and none strictly dominates another under the positive definite dispersion ordering. All determinants as listed derive from the respective $V(\widehat{\boldsymbol{\beta}}) = (\mathbf{X}'_0\mathbf{X}_0)^{-1}$ and its submatrices. Subset efficiencies for $\{\boldsymbol{\beta}_L, \boldsymbol{\beta}_Q, \boldsymbol{\beta}_I\}$ were examined in [45] for selected designs using criteria other than D- and D_s-efficiencies. Our usage here, as elsewhere in the literature, considers $GV(\widehat{\boldsymbol{\beta}})$ and $G(\widehat{\boldsymbol{\beta}})$ to be *efficiency indices* for $\widehat{\boldsymbol{\beta}}$ specific to a particular design, to include subsets $\{\widehat{\boldsymbol{\beta}}_i;\ i \in \mathsf{I}\}$, and smaller values reflect greater efficiencies through smaller volumes of concentration ellipsoids. On the other hand, the *comparative efficiencies* of two designs for estimating $\boldsymbol{\beta}$ or $\{\boldsymbol{\beta}_i;\ i \in \mathsf{I}\}$ are found as ratios of these quantities.

5.3. Numerical Studies

Details for these designs are listed in the accompanying tables. Table 1 gives values $G(\cdot) = [GV(\cdot)]^{1/\dim}$ for $\widehat{\boldsymbol{\beta}}$ and selected subsets, with dim as the order of the determinant. Also listed are angles $\phi(\mathbf{1}_n : \mathbf{X})$ deg between regressors and the constant, to be noted subsequently. Table 2 displays the squared canonical correlations $\boldsymbol{\rho}^2(\widehat{\boldsymbol{\beta}}_i : \widehat{\boldsymbol{\beta}}_j)$ between designated subsets, and Table 3 the corresponding Vector Alienation Coefficient $\boldsymbol{\gamma}(\widehat{\boldsymbol{\beta}}_i : \widehat{\boldsymbol{\beta}}_j) = \Pi_{i=1}^{r}(1-\rho_i^2)$, for specified pairs. Here $\{0\text{L}, \text{QI}\}$ refers to the pair $\{[\widehat{\beta}_0, \widehat{\boldsymbol{\beta}}_L], [\widehat{\boldsymbol{\beta}}_Q, \widehat{\boldsymbol{\beta}}_I]\}$, for example. Moreover, values of the composite indices $\boldsymbol{\gamma}(\widehat{\boldsymbol{\beta}}_i : \widehat{\boldsymbol{\beta}}_j) = \boldsymbol{\gamma}(\mathbf{X}_i : \mathbf{X}_j)$, if much less than unity, serve to alert the user as to potential problems with ill-conditioning.

5.3.1. An Overview

To fix ideas, observe for the CCD that $G(\widehat{\boldsymbol{\beta}}_Q, \widehat{\boldsymbol{\beta}}_I) = 1.13633$, $G(\widehat{\boldsymbol{\beta}}_Q) = 1.03300$, and $G(\widehat{\boldsymbol{\beta}}_I) = 1.25000$ from Table 1. These not only are comparable in magnitude, but are commensurate, in having been adjusted for dimensions and thus homogeneous of unit degree, as are all entries in Table 1. Moreover, since $(\widehat{\boldsymbol{\beta}}_Q, \widehat{\boldsymbol{\beta}}_I)$ are uncorrelated and $\gamma(\widehat{\boldsymbol{\beta}}_Q, \widehat{\boldsymbol{\beta}}_I) = 1.0$ for the CCD from Table 3, $G(\widehat{\boldsymbol{\beta}}_Q, \widehat{\boldsymbol{\beta}}_I)$ is the geometric mean $1.13633 = (1.03300)^{3/6}(1.25000)^{3/6}$ from Theorem 4.1(ii). A further rough spot check of Table 1 may be summarized as follows.

Summary Properties

(P1) Compared with $G(\widehat{\boldsymbol{\beta}})$, values for $G(\widehat{\beta}_0)$ appear excessive throughout.

(P2) Values for $G(\widehat{\boldsymbol{\beta}}_L)$ are roughly comparable across designs.

(P3) The eight designs sort essentially into two groups.

(P4) Designs $\{\text{H310}, \text{H311B}, \text{SCD}, \text{BBD}, \text{CCD}\}$ overall are comparatively D- and D_s-efficient, with the noted exception being $G(\widehat{\boldsymbol{\beta}}_I) = 4.16667$ for the SCD.

(P5) The designs $\{\text{ND}, \text{HD}, \text{BDD}\}$ are considerably less D-efficient, with their generalized variances $GV(\widehat{\boldsymbol{\beta}})$ being $\{1192.09, 4768.37, 2886.03\}$, respectively, in comparison with $\{57.342, 11.852, 74.422, 2.722, 0.523\}$ for the remaining designs; and each of the former is burdened by unequivocal D_s-inefficiency for $\boldsymbol{\beta}_Q$, to be treated subsequently.

5.3.2. Further Details

We next examine Hartley's [39] SCD in some detail, first in terms of generalized variances. Values for $GV(\widehat{\boldsymbol{\beta}})$, $GV(\widehat{\beta}_0)$, and $GV(\widehat{\boldsymbol{\beta}}_M)$ appear in the first row of Table 4, along with $\boldsymbol{\gamma}(\widehat{\beta}_0 : \widehat{\boldsymbol{\beta}}_M) = (1.0 - 0.909090) = 0.090909$ using $\boldsymbol{\rho}^2(0 : M) = 0.909090$ from Table 2. Theorem 4.1 (i) now asserts that $GV(\widehat{\boldsymbol{\beta}}) = GV(\widehat{\beta}_0)GV(\widehat{\boldsymbol{\beta}}_M)\boldsymbol{\gamma}(\widehat{\beta}_0 : \widehat{\boldsymbol{\beta}}_M)$, as verified numerically through $74.4216 = 10.00(81.8638)(0.090909)$. In a similar manner, $\widehat{\boldsymbol{\beta}}_M$ partitions into $\{\widehat{\boldsymbol{\beta}}_L, [\widehat{\boldsymbol{\beta}}_Q, \widehat{\boldsymbol{\beta}}_I]\}$, where $GV(\widehat{\boldsymbol{\beta}}_L) = 4.6296$ and $GV([\widehat{\boldsymbol{\beta}}_Q, \widehat{\boldsymbol{\beta}}_I]) = 81.8638$ from Table 4. The squared canonical correlations between $\{\widehat{\boldsymbol{\beta}}_L, [\widehat{\boldsymbol{\beta}}_Q, \widehat{\boldsymbol{\beta}}_I]\}$ are $\boldsymbol{\rho}^2(L : QI) = [0.4000, 0.4000, 0.4000]'$ from Table 2, so that $\boldsymbol{\gamma}(L : QI) = 0.21600$ as in Table 3. Theorem 4.1 (i) again recovers $GV(\widehat{\boldsymbol{\beta}}_M)$ as $81.8638 = 4.6926(81.8638)(0.21600)$ since $G(\widehat{\boldsymbol{\beta}}_L)$ and $\boldsymbol{\gamma}(L : QI)$ are reciprocals in this instance. Moreover,

Table 1: Roots $G(\widehat{\boldsymbol{\beta}})$ and $G(\widehat{\boldsymbol{\beta}}_i)$ of generalized variances for $\widehat{\boldsymbol{\beta}}$ and subset estimators $\{\widehat{\boldsymbol{\beta}}_i \in \mathbb{R}^t\}$, and angles $\phi(1_n : \mathbf{X})$deg between regressors and the constant, against reference values $\phi^{**} < 18.44$ and $\phi^{*} < 30.00$, in eight small second-order designs.

Design	$G(\widehat{\boldsymbol{\beta}})$	$G(\widehat{\beta}_0)$	$G(\widehat{\boldsymbol{\beta}}_L)$	$G(\widehat{\boldsymbol{\beta}}_Q)$	$G(\widehat{\boldsymbol{\beta}}_I)$	$G(\widehat{\boldsymbol{\beta}}_Q, \widehat{\boldsymbol{\beta}}_I)$	$G(\widehat{\boldsymbol{\beta}}_M)$	$\boldsymbol{\phi}(\mathbf{1}_n : \mathbf{X})$
H310	1.49916	5.0477	1.18821	2.08026	1.61045	1.83034	1.58484	25.11^{*}
H311B	1.28050	10.0000	1.0000	1.17680	1.99997	1.53413	1.33017	17.55^{**}
SCD	1.53876	10.0000	1.66667	1.04210	4.16667	2.08376	1.63142	17.55^{**}
BBD	1.10533	10.0000	0.83333	1.64583	1.11111	1.35229	1.15077	16.10^{**}
CCD	0.93725	10.0000	0.71429	1.03300	1.25000	1.13633	0.97341	14.96^{**}
ND	2.03063	8.7500	1.25000	7.80031	1.25000	3.12256	2.22030	18.80^{*}
HD	2.33258	5.0000	1.91189	6.61313	1.95275	3.57444	2.59004	25.24^{*}
BDD	2.21835	4.6477	1.58809	4.96625	2.25752	3.22673	2.44950	26.25^{*}

Table 2: Squared canonical correlations between designated subsets $(\widehat{\boldsymbol{\beta}}_i, \widehat{\boldsymbol{\beta}}_j)$ of estimators for eight small second-order designs.

Subsets	H310	H311B	SCD	BBD	CCD	ND	HD	BDD
{0L, QI}	0.81990	0.90909	0.90909	0.92308	0.93333	0.89333	0.79476	0.81304
			0.40000			0.11111	0.54261	0.11721
			0.40000			0.11111	0.11111	0.11721
			0.40000				0.11111	0.03524
{L, Q}						0.11111	0.11111	0.06461
						0.11111	0.11111	0.01413
						0.08333	0.00631	0.01413
{L, QI}			0.40000			0.11111	0.54456	0.11721
			0.40000			0.11111	0.11111	0.11721
			0.40000			0.08333	0.11111	0.09890
{Q, I}							0.03153	0.10418
								0.10418
								0.00196
{0, L}						0.10714	0.11790	0.02281
{0, Q}	0.81990	0.90909	0.90909	0.92308	0.93333	0.89286	0.64103	0.79815
{0, I}							0.04918	0.00220
{0, LQI}	0.81990	0.90909	0.90909	0.92308	0.93333	0.89610	0.81818	0.80440

Table 3: Vector Alienation Coefficients $\gamma(\widehat{\boldsymbol{\beta}}_i : \widehat{\boldsymbol{\beta}}_j)$ between subsets $(\widehat{\boldsymbol{\beta}}_i, \widehat{\boldsymbol{\beta}}_j)$ of estimators, and the factor $abc = [\gamma(0 : LQI)\gamma(L : QI)\gamma(Q : I)]^{1/10}$, for each of eight small second-order designs.

Subsets	H310	H311B	SCD	BBD	CCD	ND	HD	BDD
{0L, QI}	0.18010	0.09091	0.01964	0.07692	0.06667	0.08428	0.07417	0.14057
{L, Q}	1.00000	1.00000	1.00000	1.00000	1.00000	0.72428	0.78514	0.90914
{L, QI}	1.00000	1.00000	0.21600	1.00000	1.00000	0.72428	0.35986	0.70225
{Q, I}	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	0.96847	0.80092
{0, L}	1.00000	1.00000	1.00000	1.00000	1.00000	0.89286	0.88210	0.97719
{0, Q}	0.18010	0.09091	0.09091	0.07692	0.06667	0.10714	0.35897	0.20185
{0, I}	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	0.95082	0.99780
{0, LQI}	0.18010	0.09091	0.09091	0.07692	0.06667	0.10390	0.18182	0.19566
abc	0.84246	0.78679	0.67500	0.77376	0.76277	0.77206	0.75890	0.80197

Table 4: Generalized variances and Vector Alienation Coefficients between designated subsets for Hartley's [39] SCD in $k = 3$ regressors.

$GV(\hat{\boldsymbol{\beta}})$	$GV(\hat{\beta}_0)$	$GV(\hat{\boldsymbol{\beta}}_M)$	$\boldsymbol{\gamma}(0 : LQI)) = (1-\rho_1^2)$
74.4216	10.0000	81.8638	0.090909
$GV(\hat{\boldsymbol{\beta}}_M)$	$GV(\hat{\boldsymbol{\beta}}_L)$	$GV(\hat{\boldsymbol{\beta}}_Q, \hat{\boldsymbol{\beta}}_I)$	$\gamma(L : QI) = \Pi_{i=1}^{r}(1-\rho_i^2)$
81.8638	4.6296	81.8638	0.216000
$GV(\hat{\boldsymbol{\beta}}_Q, \hat{\boldsymbol{\beta}}_I)$	$GV(\hat{\boldsymbol{\beta}}_Q)$	$GV(\hat{\boldsymbol{\beta}}_I)$	$\boldsymbol{\gamma}(Q : I) = \Pi_{i=1}^{r}(1-\rho_i^2)$
81.8638	1.1317	72.3379	1.000000

$GV(\hat{\boldsymbol{\beta}}_Q, \hat{\boldsymbol{\beta}}_I) = GV(\hat{\boldsymbol{\beta}}_Q)GV(\hat{\boldsymbol{\beta}}_I)\boldsymbol{\gamma}(Q : I)$ translates into $81.8638 = 1.1317(72.3379)(1.0)$, where $\boldsymbol{\gamma}(Q : I) = 1.0$ since elements of $\{\hat{\boldsymbol{\beta}}_Q, \hat{\boldsymbol{\beta}}_I\}$ are mutually uncorrelated from Table 2. In summary, the value $G(\hat{\boldsymbol{\beta}})$ for the SCD admits the factorization of (4.6), on identifying $\{\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2, \hat{\boldsymbol{\theta}}_3, \hat{\boldsymbol{\theta}}_4\}$ with $\{\hat{\beta}_0, \hat{\boldsymbol{\beta}}_L, \hat{\boldsymbol{\beta}}_Q, \hat{\boldsymbol{\beta}}_I\}$, respectively, given numerically from Tables 1 and 3 as

$$1.53876 = (10.00)^{1/10}(1.66667)^{3/10}(1.04210)^{3/10}(4.16667)^{3/10}[(0.090909)(0.21600)(1.0)]^{1/10}. \tag{5.2}$$

Corresponding factorizations proceed similarly for other designs. Details are left to the reader, but values for $[\boldsymbol{\gamma}(0 : LQI)\boldsymbol{\gamma}(L : QI)\boldsymbol{\gamma}(Q : I)]^{1/10}$, the rightmost factor of (5.2), are supplied for each design as the final row of Table 3. Although the tables, together with Theorem 4.1, support other factorizations, the one featured here seems most natural in terms of the parameters $\{\beta_0, \boldsymbol{\beta}_L, \boldsymbol{\beta}_Q, \boldsymbol{\beta}_I\}$, together with their central roles in identifying noteworthy treatment effects in second-order models.

5.4. Masking

The D-efficiency index of the SCD, at $GV(\hat{\boldsymbol{\beta}}) = 74.4216$, is larger but roughly comparable to that of H310 at $GV(\hat{\boldsymbol{\beta}}) = 57.3418$. What cannot be anticipated from these facts alone, however, is that the (3×3) determinant $GV(\hat{\boldsymbol{\beta}}_I) = 72.3379$ for the SCD is comparable to its (10×10) determinant $GV(\hat{\boldsymbol{\beta}}) = 74.4216$, despite their disparate dimensions. Adjusting for dimensions gives $G(\hat{\boldsymbol{\beta}}) = (74.4216)^{1/10} = 1.53876$ and $G(\hat{\boldsymbol{\beta}}_I) = (72.3379)^{1/3} = 4.16667$ for the SCD. This illustrates the masking of a remarkably inefficient estimator for $\boldsymbol{\beta}_I$, despite the value $G(\hat{\boldsymbol{\beta}}) = 1.53876$ in estimating all parameters. This masking stems from the nonorthogonality of subset estimators as reflected in their canonical correlations and Vector Alienation Coefficients. In contrast are the corresponding commensurate values for the H310 design, namely, $G(\hat{\boldsymbol{\beta}}) = (57.3418)^{1/10} = 1.49916$ and $G(\hat{\boldsymbol{\beta}}_I) = (4.1768)^{1/3} = 1.61045$. It may be noted that the condition number $c_1(\mathbf{X}_0'\mathbf{X}_0)$ is 21.59 for H310, with the somewhat larger value 54.01 for the SCD.

We next examine the D_s-inefficiencies of ND and HD for $\boldsymbol{\beta}_Q$ as noted earlier, with $G(\hat{\boldsymbol{\beta}}_Q)$ taking values 7.80031 and 6.61313, respectively. Our reference for masking is $G(\hat{\boldsymbol{\beta}}_L, \hat{\boldsymbol{\beta}}_Q)$. These values are not listed in Table 1, but may be recovered from Tables 2 and 3 as follows. Specifically, for ND we have

$$\begin{gathered}\left[GV\left(\hat{\boldsymbol{\beta}}_L\right)GV\left(\hat{\boldsymbol{\beta}}_Q\right)\gamma(L : \mathbf{Q})\right]^{1/6} = \left[GV\left(\hat{\boldsymbol{\beta}}_L, \hat{\boldsymbol{\beta}}_Q\right)\right]^{1/6} = G\left(\hat{\boldsymbol{\beta}}_L, \hat{\boldsymbol{\beta}}_Q\right), \\ (1.25000)^{3/6}(7.80031)^{3/6}(0.72428)^{1/6} = 2.95912,\end{gathered} \tag{5.3}$$

where neither $G(\widehat{\boldsymbol{\beta}}_L) = 1.25000$ nor $G(\widehat{\boldsymbol{\beta}}_L, \widehat{\boldsymbol{\beta}}_Q) = 2.95912$ appears excessive. In consequence, that $G(\widehat{\boldsymbol{\beta}}_Q) = 7.80031$ is excessive would be masked on examining $G(\widehat{\boldsymbol{\beta}}_L)$ and $G(\widehat{\boldsymbol{\beta}}_L, \widehat{\boldsymbol{\beta}}_Q)$ only. Parallel steps for HD give the factorization $(1.91189)^{3/6}(6.61313)^{3/6}(0.78514)^{1/6} = 3.41528$, with similar conclusions in regard to masking.

5.5. Collinearity with the Constant

Advocates for these and other small designs have focused on D, D_s, and other efficiency criteria, as well as the parsimony of small designs and their advantage in industrial experiments. To the knowledge of this observer, none has considered prospects for ill-conditioning and its consequences, despite the fact that columns of $\mathbf{X}$ are necessarily inter-linked as a consequence of second-order from first-order effects. Nonetheless, from Section 3.1 and Corollary 3.2, we may compute angles between the constant vector and the span of the regressors using duality together with the information at hand. This may prove to be critical in view of the admonition [29] that "collinearity with the intercept can quite generally corrupt the estimates of *all parameters* in the model." As noted in Remark 3.4, rules-of-thumb for problematic VIFs include those exceeding 10 or 4 or, in angular measure, $\phi^{**} < 18.435\,\text{deg}$ and $\phi^{*} < 30.00\,\text{deg}$. From the last row of Table 2, the angles $\phi(\mathbf{1}_n : \mathbf{X})$ have been computed for each of the eight designs, as listed in the final column of Table 1. For example, $\arccos(\sqrt{0.81990}) = 25.1116\,\text{deg}$ for H310. It is seen that all designs are flagged as potentially problematic using rules-of-thumb as cited. This adds yet another layer of concerns, heretofore unrecognized, in seeking further to implement these designs already in wide usage.

6. Conclusions

Duality of (i) Hotelling's [30] *canonical correlations* $\{\rho_1, \ldots, \rho_r\}$ between the *OLS* estimators $\{\widehat{\boldsymbol{\beta}}_1, \widehat{\boldsymbol{\beta}}_2\}$ and (ii) the *design linkage parameters* $\{\delta_1, \ldots, \delta_r\}$ between $\{\mathbf{X}_1, \mathbf{X}_2\}$ is established at the outset. Stewart's [33] *collinearity indices* are then extended to encompass angles $\{\phi_0, \phi_1, \ldots, \phi_k\}$ between each column of $\mathbf{X}_0 = [\mathbf{1}_n, \mathbf{X}_1, \ldots, \mathbf{X}_k]$ and remaining columns. In particular, ϕ_0 quantifies numerically the collinearity of regressors with the intercept, of concern in the prospective corruption of all estimates due to ill-conditioning.

Matrix identities factorize a determinant in terms of principal subdeterminants and the *Vector Alienation Coefficients* of [30] between $\{\widehat{\boldsymbol{\beta}}_1, \widehat{\boldsymbol{\beta}}_2\}$. By duality, the latter also are *Alienation Coefficients* between $\{\mathbf{X}_1, \mathbf{X}_2\}$. These identities in turn are applied in the study of D_s-efficiencies for the parameters $\{\beta_0, \boldsymbol{\beta}_L, \boldsymbol{\beta}_Q, \boldsymbol{\beta}_I\}$ in eight small second-order designs from the literature. Studies on D_s- and D-efficiencies, as cited in our opening paragraph, confirm that designs are seldom efficient for both. Our determinant identities support a rational explanation. In particular, these identities unmask the propensity for D_s-inefficient subset estimators to be masked through near collinearities in overall D-efficient designs.

Finally, the evidence suggests that all eight designs are vulnerable, to varying degrees, to the corruption of all estimates due to ill-conditioning. In short, we have exposed quantitatively the structural origins of masking through Hotelling's [30] canonical correlations, and their equivalent design linkage parameters. This analysis in turn proceeds from the design matrix itself rather than empirical estimates, so that any design can be evaluated beforehand with regard to masking and possible subset inefficiencies, rather than retrospectively after having committed to a given design in a particular experiment.

References

[1] J. C. Kiefer, "Optimum experimental designs," *Journal of the Royal Statistical Society. Series B*, vol. 21, pp. 272–319, 1959.

[2] J. C. Kiefer, "Optimum designs in regression problems. II," *Annals of Mathematical Statistics*, vol. 32, pp. 298–325, 1961.

[3] J. C. Kiefer and J. Wolfowitz, "Optimum designs in regression problems," *Annals of Mathematical Statistics*, vol. 30, pp. 271–294, 1959.

[4] A. Wald, "On the efficient design of statistical investigations," *Annals of Mathematical Statistics*, vol. 14, pp. 134–140, 1943.

[5] M. J. Box and N. R. Draper, "Factorial designs, the $|X'X|$ criterion, and some related matters," *Technometrics*, vol. 13, pp. 731–742, 1971.

[6] J. M. Lucas, "Which response surface design is best," *Technometrics*, vol. 18, pp. 411–417, 1976.

[7] R. H. Myers, A. I. Khuri, and W. H. Carter, Jr., "Response surface methodology: 1966–1988," *Technometrics*, vol. 31, no. 2, pp. 137–157, 1989.

[8] R. Sibson, "D_A-optimality and duality," in *Progress in Statistics (European Meeting Statisticians, Budapest, 1972), Vol. II*, J. Gani, K. Sarkadi, and I. Vincze, Eds., North-Holland, Amsterdam, The Netherlands, 1974.

[9] S. D. Silvey, *Optimal Design*, Chapman & Hall, London, UK, 1980.

[10] R. C. St. John and N. R. Draper, "D-optimality for regression designs: a review," *Technometrics*, vol. 17, pp. 15–23, 1975.

[11] H. Dette, "A generalization of D- and D_1-optimal designs in polynomial regression," *The Annals of Statistics*, vol. 18, no. 4, pp. 1784–1804, 1990.

[12] H. Dette and T. Franke, "Robust designs for polynomial regression by maximizing a minimum of D- and D_1-efficiencies," *The Annals of Statistics*, vol. 29, no. 4, pp. 1024–1049, 2001.

[13] S. Huda, "On some D_s-optimal designs in spherical regions," *Communications in Statistics*, vol. 20, no. 9, pp. 2965–2985, 1991.

[14] S. Huda and A. A. Al-Shiha, "On D-optimal designs for estimating slope," *Sankhyā. Series A*, vol. 61, no. 3, pp. 488–495, 1999.

[15] C. M.-S. Lee, "D-optimal designs for polynomial regression when lower degree parameters are more important," *Utilitas Mathematica*, vol. 34, pp. 53–63, 1988.

[16] Y. B. Lim and W. J. Studden, "Efficient D_s-optimal designs for multivariate polynomial regression on the q-cube," *The Annals of Statistics*, vol. 16, no. 3, pp. 1225–1240, 1988.

[17] W. J. Studden, "D_s-optimal designs for polynomial regression using continued fractions," *The Annals of Statistics*, vol. 8, no. 5, pp. 1132–1141, 1980.

[18] D. M. Wardrop and R. H. Myers, "Some response surface designs for finding optimal conditions," *Journal of Statistical Planning and Inference*, vol. 25, no. 1, pp. 7–28, 1990.

[19] H. P. Wynn, "The sequential generation of D-optimum experimental designs," *Annals of Mathematical Statistics*, vol. 41, pp. 1655–1664, 1970.

[20] H. P. Wynn, "Results in the theory and construction of D-optimum experimental designs," *Journal of the Royal Statistical Society. Series B. Methodological*, vol. 34, pp. 133–147, 1972.

[21] H. Dette, "On a mixture of the D- and D_1-optimality criterion in polynomial regression," *Journal of Statistical Planning and Inference*, vol. 35, no. 2, pp. 233–249, 1993.

[22] S. Huda, "On D_s-efficiency of some D-optimal designs," *Calcutta Statistical Association Bulletin*, vol. 34, no. 133-134, pp. 119–122, 1985.

[23] S. Huda and I. H. Khan, "On efficiency of some D-optimal designs in subset estimation," *Journal of the Indian Society of Agricultural Statistics*, vol. 47, no. 2, pp. 151–164, 1995.

[24] P. F. Schwartz, C. Gennings, L. K. Teuschler, and M. W. Farris, "Optimizing the precision of toxicity threshold estimation using a two-stage experimental design," *Journal of Agricultural, Biological, and Environmental Statistics*, vol. 6, pp. 409–428, 2001.

[25] S. L. Meadows-Shropshire, C. Gennings, W. H. Carter, and J. E. Simmons, "Analysis of mixtures of drugs/chemicals along a fixed-ratio ray without single-chemical data to support an additivity model," *Journal of Agricultural, Biological, and Environmental Statistics*, vol. 9, pp. 500–514, 2004.

[26] D. A. Belsley, "Centering, the constant, first-differencing, and assessing conditioning," in *Model Reliability*, D. A. Belsley and E. Kuh, Eds., pp. 117–153, MIT Press, Cambridge, Mass, USA, 1986.

[27] E. R. Mansfield and B. P. Helms, "Detecting multicollinearity," *American Statistician*, vol. 36, pp. 158–160, 1982.

[28] S. D. Simon and J. P. Lesage, "The impact of collinearity involving the intercept term on the numerical accuracy of regression," *Computer Science in Economics and Management*, vol. 1, no. 2, pp. 137–152, 1988.
[29] D. A. Belsley, "Comment: Well-conditioned collinearity indices," *Statistical Science*, vol. 2, no. 1, pp. 86–91, 1987.
[30] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 321–377, 1936.
[31] Å. Björck and G. H. Golub, "Numerical methods for computing angles between linear subspaces," *Mathematics of Computation*, vol. 27, pp. 579–594, 1973.
[32] M. L. Eaton, *Multivariate Statistics: A Vector Space Approach*, John Wiley & Sons, New York, NY, USA, 1983.
[33] G. W. Stewart, "Collinearity and least squares regression," *Statistical Science*, vol. 2, no. 1, pp. 68–100, 1987.
[34] R. H. Myers, *Classical and Modern Regression with Applications*, PWS–KENT, Boston, Mass, USA, 2nd edition, 1990.
[35] R. M. O'Brien, "A caution regarding rules of thumb for variance inflation factors," *Quality and Quantity*, vol. 41, no. 5, pp. 673–690, 2007.
[36] S. S. Wilks, *Mathematical Statistics*, John Wiley & Sons, New York, NY, USA, 1962.
[37] H. Cramér, *Mathematical Methods of Statistics*, Princeton Mathematical Series, vol. 9, Princeton University Press, Princeton, NJ, USA, 1946.
[38] K. G. Roquemore, "Hybrid designs for quadratic response surfaces," *Technometrics*, vol. 18, pp. 419–423, 1979.
[39] H. O. Hartley, "Smallest composite designs for quadratic response surfaces," *Biometrics. Journal of the Biometric Society*, vol. 15, pp. 611–624, 1959.
[40] G. E. P. Box and D. W. Behnken, "Some new three level designs for the study of quantitative variables," *Technometrics*, vol. 2, pp. 455–475, 1960.
[41] G. E. P. Box and K. B. Wilson, "On the experimental attainment of optimum conditions," *Journal of the Royal Statistical Society. Series B*, vol. 13, pp. 1–45, 1951.
[42] W. Notz, "Minimal point second order designs," *Journal of Statistical Planning Inference*, vol. 6, no. 1, pp. 47–58, 1982.
[43] A. T. Hoke, "Economical second-order designs based on irregular fractions of the 3^n factorial," *Technometrics*, vol. 16, pp. 375–384, 1974.
[44] M. J. Box and N. R. Draper, "On minimum-point second-order designs," *Technometrics*, vol. 16, pp. 613–616, 1974.
[45] D. R. Jensen, "Efficiencies of some small second-order designs," *Statistics & Probability Letters*, vol. 21, no. 4, pp. 255–261, 1994.