56

COMPLEXITY IN EXPERIMENTAL DESIGN
I. RELATION TO EFFICIENCY

D. R. Jensen and D. E. Ramirez

Virginia Polytechnic Institute and State University
and
University of Virginia

Key Words. *Design efficiency, A and D efficiency, complexity*

AMS 1980 Subject Classifications. *62J05, 62K05*

0. ABSTRACT

Complexity based on entropy is developed as a criterion for design evaluation. Connections are drawn to such standard criteria as A- and D-efficiencies, and these are evaluated numerically in selected cases.

1. INTRODUCTION

Given a linear model $\mathbf{Y} = \mathbf{f}(\mathbf{X})\beta + \varepsilon$, a criterion $C(\cdot)$, and designs $\mathbf{X}$ and $\mathbf{Z}$ such that $C(\mathbf{X}) \leq C(\mathbf{Z})$, design $\mathbf{X}$ is said to be *more* C-*efficient* than $\mathbf{Z}$. Current usage includes A-efficiency and D-efficiency for which $C_A(\mathbf{X}) = \operatorname{tr}(\Sigma)$ and $C_D(\mathbf{X}) = |\Sigma|$, where $\Sigma = [\mathbf{f}(\mathbf{X})'\mathbf{f}(\mathbf{X})]^{-1}$ and $\operatorname{tr}(\Sigma)$ and $|\Sigma|$ are the trace and determinant of Σ, respectively. In effect, A-efficiency guarantees smaller average variances for the Gauss-Markov estimators $\hat{\beta}(\mathbf{X}) = [\mathbf{f}(\mathbf{X})'\mathbf{f}(\mathbf{X})]^{-1}\mathbf{f}(\mathbf{X})'\mathbf{Y}$, whereas D-efficiency guarantees smaller generalized variances and volumes of confidence ellipsoids for β having coefficient $1-\alpha$.

Unfortunately, even regions with minimal volume and small sections in some directions may be greatly elongated in others, reflecting highly imprecise information for one or more linear parametric functions. Moreover, these criteria often work at crossed purposes, there being no globally optimal designs apart from special cases. In particular, a less D-efficient design may give more C_A-efficient regions of greater regularity that are more readily interpreted in practice. These considerations suggest a compromise such as the ratio of $C_A(\cdot)$ to $C_D(\cdot)$ when standardized suitably. This in turn will be seen to entail the concept of *complexity* as it bears on design evaluation.

In Part I of this study we show that complexity, as defined in van Emden (1971), may be used effectively in design evaluation. Part II compares the complexities of eight small second-order designs from the literature. Complexity has been advocated as a criterion for use in model selection by Maklad and Nichols (1980) and others. Since the choice of design is an important determinant of properties in linear estimation, it is natural to consider the design itself as an essential aspect of model selection. Complexity of a design, as used here, essentially gauges the nonorthogonality of $\mathbf{X}$, as mirrored through $\Sigma = [\mathbf{f}(\mathbf{X})'\mathbf{f}(\mathbf{X})]^{-1}$, and thereby the degree of regularity of confidence ellipsoids for β. Our findings quantify fundamental relations among complexity, A- and D-efficiencies, and a gauge of discrepancy between Σ and a scalar matrix. Two complexity indices are studied in detail in Ramirez (1989). Design comparisons using any of the aforementioned criteria are not data-dependent, and thus can be done diagnostically in planning before an experiment has been carried out.

2. COMPLEXITY

The notion of the complexity of a system refers its total entropy to that of the noninteracting composition of its parts; see van Emden (1971). More precisely, the sum of the entropies of the elements of a random vector, minus their joint entropy, is commonly used to gauge their interpendency, or *complexity*. Thus if $\mathbf{Y} = [Y_1, \ldots, Y_k]'$ is random having distribution $\mathscr{L}(\mathbf{Y})$ and the dispersion matrix Λ, then the complexity of $\mathscr{L}(\mathbf{Y})$ has been defined by van Emden (1971) in terms of Λ as

$$\phi(\Lambda) = [k \ln(\mathrm{tr}(\Lambda)/k) - \ln(|\Lambda|)]/2. \tag{2.1}$$

Let $\{\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0\}$ denote the ordered eigenvalues of Λ; let $\bar{\lambda}$ be their average and $GM(\lambda)$ their geometric mean. Then $\phi(\bullet)$ has the convenient rule

$$\phi(\Lambda) = [\ln(\bar{\lambda}/\prod_{i=1}^{k} \lambda_i)]/2 = (k/2)\ln[\bar{\lambda}/GM(\lambda)]. \tag{2.2}$$

Since $\phi(\Lambda) \geq 0$ by the arithmetic-geometric mean inequality, and since $\phi(\Lambda) = 0$ only when Λ is a scalar matrix, $\phi(\bullet)$ may be used to gauge whether a design $\mathbf{X}$ yields Gauss-Markov estimators $\hat{\beta}(\mathbf{X})$ more nearly spherical in distribution than another design $\mathbf{Z}$. In fact, when evaluated at the sample dispersion matrix $\mathbf{S}$, $\phi(\mathbf{S})$ is related one-to-one through the coefficient of ellipticity to the normal-theory test that Λ is a scalar matrix; see Muirhead (1982), for example.

In order to compare designs on the basis of their relative complexity, we drop the logarithmic scale and say that design $\mathbf{X}$ is *less complex* than $\mathbf{Z}$ whenever $C_1(\mathbf{X}, \mathbf{Z}) \leq 1$, where

$$C_1(\mathbf{X}, \mathbf{Z}) = \frac{\mathrm{tr}(\Sigma)}{\mathrm{tr}(\Omega)} \left[\frac{|\Omega|}{|\Sigma|} \right]^{1/k} \tag{2.3}$$

with $\Sigma = [\mathbf{f}(\mathbf{X})'\mathbf{f}(\mathbf{X})]^{-1}$ and $\Omega = [\mathbf{f}(\mathbf{Z})'\mathbf{f}(\mathbf{Z})]^{-1}$. There is a natural connection between the standard design criteria of A- and D-efficiency and the relative complexity of two designs. In particular, A-efficiency is related directly, and D-efficiency inversely, to C_1-complexity. Details follow.

THEOREM 1. Consider designs $\mathbf{X}$ and $\mathbf{Z}$, with $\Sigma = [\mathbf{f}(\mathbf{X})'\mathbf{f}(\mathbf{X})]^{-1}$ and $\Omega = [\mathbf{f}(\mathbf{Z})'\mathbf{f}(\mathbf{Z})]^{-1}$, pertaining to the Gauss-Markov estimators $\hat{\beta}(\mathbf{X})$ and $\hat{\beta}(\mathbf{Z})$, respectively.

(i) Then $\phi(\Sigma)$, $\phi(\Omega)$, and $C_1(\mathbf{X}, \mathbf{Z})$ are related as

$$\phi(\Sigma) - \phi(\Omega) = (1/2)\ln\left[\frac{(\mathrm{tr}\,(\Sigma))^k|\Omega|}{(\mathrm{tr}\,(\Omega))^k|\Sigma|}\right] = (k/2)\ln(C_1(\mathbf{X}, \mathbf{Z})). \tag{2.4}$$

(ii) If $\mathbf{X}$ and $\mathbf{Z}$ are D-equivalent, then $\mathbf{Z}$ has greater A-efficiency than $\mathbf{X}$ if and only if $\mathbf{Z}$ is less complex than $\mathbf{X}$.

(iii) If $\mathbf{X}$ and $\mathbf{Z}$ are A-equivalent, then $\mathbf{Z}$ has greater D-efficiency than $\mathbf{X}$ if and only if $\mathbf{Z}$ is more complex than $\mathbf{X}$.

Proof. Observe that

$$\begin{aligned}\phi(\Sigma) - \phi(\Omega) &= (k/2)\ln[\mathrm{tr}\,(\Sigma)/k] - (1/2)\ln(|\Sigma|) - (k/2)\ln[\mathrm{tr}\,(\Omega)/k] + (1/2)\ln(|\Omega|) \\ &= (1/2)\ln\left(\frac{\mathrm{tr}\,(\Sigma)}{\mathrm{tr}\,(\Omega)}\right)^k - (1/2)\ln\left(\frac{|\Sigma|}{|\Omega|}\right)\end{aligned}$$

which is equivalent to assertion (i) of the theorem. Conclusions (ii) and (iii) follow directly from conclusion (i) on setting $|\Sigma| = |\Omega|$, then $\mathrm{tr}\,(\Sigma) = \mathrm{tr}\,(\Omega)$. $\square$

3. APPLICATIONS

To fix ideas, consider the second-order response model

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_{11}X_1^2 + \beta_{12}X_1X_2 + \beta_{22}X_2^2 + \varepsilon \tag{3.1}$$

together with three designs giving matrices $\mathbf{f}(\mathbf{X})$, $\mathbf{f}(\mathbf{Z})$, and and $\mathbf{f}(\mathbf{T})$, all of order (9×6). The basic designs are given as points in the (X_1, X_2)-plane. The first is the standard 3^2 factorial design consisting of the nine points $\{(0, 0), (\pm 1, 0), (0, \pm 1), (\pm 1, \pm 1)\}$. The second design derives from the first by rotating points in the (X_1, X_2)-plane clockwise through 45 degrees, giving $\{(0, 0), (\pm\sqrt{2}, 0), (0, \pm\sqrt{2}), (\pm\sqrt{0.5}, \pm\sqrt{0.5})\}$. The third design consists of eight roots of unity in the (X_1, X_2)-plane, together with the origin, giving $\{(0, 0), (\pm 1, 0), (0, \pm 1), (\pm\sqrt{0.5}, \pm\sqrt{0.5})\}$. We refer to these as designs $\mathbf{X}$, $\mathbf{Z}$ and $\mathbf{T}$, respectively.

The eigenvalues, arithmetic and geometric means, and complexity index $\phi(\cdot)$ are given in Table 2 for the matrices $\Sigma = [\mathbf{f}(\mathbf{X})'\mathbf{f}(\mathbf{X})]^{-1}$, $\Omega = [\mathbf{f}(\mathbf{Z})'\mathbf{f}(\mathbf{Z})]^{-1}$, and $\Delta = [\mathbf{f}(\mathbf{T})'\mathbf{f}(\mathbf{T})]^{-1}$ with reference to the model (3.1). The tabular results demonstrate that the

3^2 factorial design produces Gauss-Markov estimators having the least complex dispersion structure of the three designs. In fact, the relative complexities from expression (2.3) are given by $C_1(\mathbf{X}, \mathbf{Z}) = 0.3794$, $C_1(\mathbf{X}, \mathbf{T}) = 0.1852$, and $C_1(\mathbf{Z}, \mathbf{T}) = 0.4882$. It may be noted that these designs have successively larger values for the trace, namely, 2.1389, 2.5139 and 5.2500, while their determinants are 0.0001929, 0.0001929, and 0.0078129, respectively. As designs $\mathbf{X}$ and $\mathbf{Z}$ are D-equivalent, their A-efficiency and C_1-efficiency are related directly as noted in the statement of Theorem 1.

To continue, we next reconsider the allocation of design points to be placed on the perimeter and at the center of an equiradial design $\Delta(\lambda)$ consisting of n_1 points equally spaced on the unit circle, and n_0 points at the origin, where $N = n_0 + n_1$ and $\lambda = n_1/N$. We now invoke the three criteria $\{C_A(\bullet), C_D(\bullet), \phi(\bullet)\}$ to determine λ. The moment matrix $\mathbf{M}(\lambda)$ of order (6×6) satisfies

$$\mathbf{M}(\lambda)/N = (1/8)\begin{bmatrix} \mathbf{M}_{11}(\lambda) & \mathbf{M}_{12}(\lambda) & \mathbf{0} \\ \mathbf{M}_{21}(\lambda) & \mathbf{M}_{22}(\lambda) & \mathbf{0} \\ \mathbf{0}' & \mathbf{0}' & \lambda \end{bmatrix} \tag{3.2}$$

in which $\mathbf{M}_{11}(\lambda) = \mathrm{Diag}\,(8, 4\lambda, 4\lambda)$, $\mathbf{M}_{22}(\lambda)$ has diagonal elements of 3λ and off-diagonal elements λ, and $\mathbf{M}_{12}(\lambda)$ has $4\lambda[1, 1]$ as its first row and zeros elsewhere. See Wardrop and Myers (1990), for example. The eigenvalues of $\mathbf{M}(\lambda)/N$ are $\{\lambda/2, \lambda/2, \lambda/4, ((\lambda/2) + 1) \pm (1 - \lambda + (9/4)\lambda^2)^{1/2})/2\}$, and the eigenvalues of the dispersion matrix $\Xi(\lambda)$ for $\hat{\beta}(\Delta(\lambda))$ are proportional to the inverses of these values. A numerical search yields the optimal values for λ under each of the three criteria, and these values are shown in Table 2.

The dispersion matrix for $\hat{\beta}(\Delta(\lambda))$ determines the confidence ellipsoid for the parameters $\beta = [\beta_0, \beta_1, \beta_2, \beta_{11}, \beta_{12}, \beta_{22}]'$. In the class of designs $\{\Delta(\lambda); 0 \leq \lambda \leq 1\}$, the choice $\lambda = 0.7101$ minimizes the sum of squared lengths of the semiaxes of the confidence ellipsoid for β. The choice $\lambda = 0.8333$ minimizes the product of the squared lengths of the semiaxes, or equivalently, the volume of the confidence ellipsoid. Finally, the choice $\lambda = 0.5455$ minimizes the complexity of the dispersion matrix for $\hat{\beta}(\Delta(\lambda))$, yielding a confidence region that is most nearly spherical as λ is varied.

An important application of complexity is seen on choosing between possible designs, say the A-optimal and D-optimal designs pertaining to Table 2. If the researcher prefers a more regular confidence ellipsoid that is more nearly spherical, then the A-optimal design with $\lambda = 0.7101$ is to be preferred, since its complexity, with value $\phi(\Xi(0.7101)) = 1.10$, is smaller than the complexity $\phi(\Xi(0.8333)) = 1.54$ at $\lambda = 0.8333$.

REFERENCES

Emden, M. H. van (1971). "An analysis of complexity." *Mathematical Centre Tracts,* Volume 35. Mathematical Centrum, Amsterdam.

Maklad, M. S. and Nichols, S. T. (1980). "A new approach to model structure discrimination." *IEEE Transactions on System, Man, and Cybernetics* **10**: 78-84.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory.* John Wiley and Sons, New York.

Ramirez, D. E. (1989). "On two nonequivalent measures of complexity." *IEEE Transactions on System, Man, and Cybernetics* **19**: 1073-1077.

Wardrop, D. and Myers, R. H. (1990). "Some response surface designs for finding optimal conditions." *J. Stat. Planning and Inference* **25**: 7-28.

TABLE 1. Eigenvalues, arithmetic and geometric means, and complexity index for the matrices $\Sigma = [\mathbf{f}(\mathbf{X})'\mathbf{f}(\mathbf{X})]^{-1}$, $\Omega = [\mathbf{f}(\mathbf{Z})'\mathbf{f}(\mathbf{Z})]^{-1}$, and $\Delta = [\mathbf{f}(\mathbf{T})'\mathbf{f}(\mathbf{T})]^{-1}$ stemming from the three designs with reference to the model (3.1).

Matrix	Eigenvalues	$\bar{\lambda}$	$GM(\lambda)$	$\phi(\cdot)$
Σ	{ 1, 1/2, 1/4, 1/6, 1/6, 1/18 }	0.3565	0.2404	1.182
Ω	{ 1, 1, 1/6, 1/6, 1/8, 1/18 }	0.4190	0.2404	1.667
Δ	{ 3.17116, 1, 1/2, 1/4, 1/4, 0.07884 }	0.8750	0.4455	2.025

TABLE 2. Determining the optimal fraction $\lambda = n_1/N$ of points to be placed on the perimeter of an equiradial design $\Delta(\lambda)$ according to the three criteria $\{C_A(\cdot),\ C_D(\cdot),\ \phi(\cdot)\}$ with reference to the model (3.1).

Criterion	Optimal λ
$C_A(\Xi(\lambda)) = \mathrm{tr}\,(\Xi(\lambda))$	0.7101
$C_D(\Xi(\lambda)) = \lvert\Xi(\lambda)\rvert$	0.8333
$\phi(\Xi(\lambda)) = (k/2)\ln[\bar{\lambda}/GM(\lambda)]$	0.5455