Journal of Statistical Planning and Inference 12 (1985) 87-91
North-Holland

ASYMPTOTICALLY OPTIMAL WEIGHING DESIGNS WITH STRING PROPERTY

Rahul MUKERJEE and Rita SAHA RAY

Division of Theoretical Statistics and Mathematics, Indian Statistical Institute, 203 Barrackpore Trunk Road, Calcutta 700035, India

Received 30 April 1984; revised manuscript received 1 August 1984
Recommended by A.S. Hedayat

Abstract: Asymptotically D- and E-optimal spring balance weighing designs with string property are obtained. The techniques applied include use of Fréchet derivatives. These asymptotic results are helpful in dealing with the more intractable design problem with a finite number of observations and some new exact optimality results follow.

AMS Subject Classification: 62K05.

Key words and phrases: D-optimality; E-optimality; Fréchet derivative.

1. Introduction

Fulkerson and Gross (1965) and Ryser (1969) considered matrices with elements 0 or 1 with exactly one run of 1's in each row. These were called by them (0,1) matrices with consecutive 1's property. Sinha and Saha (1983) indicated that spring balance weighing designs, with the design matrix having consecutive 1's property, have wide applications in a number of fields - particularly in biometry and optics - and cited several references in this regard. They called such designs spring balance designs with *string property*.

The problem of obtaining optimal spring balance designs with string property was initiated by Sinha and Saha (1983) and, in particular, the special case when n, the number of observations, equals p, the number of objects, was covered extensively. Some results for the case $n>p$ were also obtained by them and certain conjectures were made.

This paper applies techniques including the use of Fréchet derivatives to obtain asymptotically D- and E-optimal spring balance weighing designs with string property. As elaborated in the last section these asymptotic results are helpful in dealing with the more intractable design problem with a finite number of observations and for certain combinations of n and p, not covered in Sinha and Saha (1983), exactly optimal designs have been characterised.

2. Preliminaries

For $1 \le u \le v \le p$, let h_{uv} be a vector with elements 0 or 1 and having exactly one run of 1's starting at the u-th and ending at the v-th (both inclusive) positions. With p objects and under string property, each row of the design matrix must be one of these $\frac{1}{2}p(p+1)$ vectors. Let $\mathscr{X}$, the design space, be the set of the $\frac{1}{2}p(p+1)$ vectors h_{uv}. Following Silvey (1980, p. 15), let H be the class of probability distributions on the Borel sets of $\mathscr{X}$. Any $\eta \in H$ will be called a design measure. Since $\mathscr{X}$ is finite, any such η defines a discrete distribution over $\mathscr{X}$ assigning a mass, say, π_{uv} at h_{uv} $(1 \le u \le v \le p)$. For $\eta \in H$, under the standard Gauss–Markov linear model with homoscedasticity and independence of errors, define the $(p \times p)$ information matrix $\mathbf{M}(\eta) = E(x\,x')$, x being a random vector with distribution η. Denoting the (i,j)-th element of $M(\eta)$ by $m_{ij}(\eta)$, it can be checked that

$$m_{ij}(\eta) = \sum \pi_{uv}, \tag{2.1}$$

where the summation extends over $u \le i$ and $v \ge j$ $(1 \le i \le j \le p)$. Hence it follows that

$$\pi_{uv} = m_{uv}(\eta) - m_{u,v+1}(\eta) - m_{u-1,v}(\eta) + m_{u-1,v+1}(\eta), \quad 1 \le u \le v \le p, \tag{2.2}$$

$m_{0,p+1}(\eta)$, $m_{0v}(\eta)$ $(1 \le v \le p)$ and $m_{u,p+1}(\eta)$ $(1 \le u \le p)$ being interpreted as zero.

Let $\mathscr{M} = \{M(\eta) : \eta \in H\}$ and ϕ be a real-valued function defined on the class of $(p \times p)$ symmetric matrices and bounded above on $\mathscr{M}$. Then the problem is to determine η^* to maximize $\phi[M(\eta)]$ over H. Any such η^* will be called ϕ-optimal. In the sequel the following theorem (Silvey (1980, p. 19)) will be used.

Theorem 2.1. *If ϕ is concave on $\mathscr{M}$ and differentiable at $M(\eta^*)$, then η^* is ϕ-optimal if and only if $F_\phi(M(\eta^*), xx') \le 0$ for each $x \in \mathscr{X}$.*

In the above

$$F_\phi(M(\eta^*), xx') = \lim_{\varepsilon \to 0+} \varepsilon^{-1}[\phi\{(1-\varepsilon)M(\eta^*) + \varepsilon x x'\} - \phi(M(\eta^*))]$$

is the Fréchet derivative of ϕ at $M(\eta^*)$ in the direction of xx'.

3. D- and E-optimal designs

3.1. D-optimal designs

For D-optimality, $\phi[M(\eta)] = \log\det M(\eta)$ and if $M(\eta)$ is nonsingular it can be seen that (cf. Silvey (1980, pp. 21))

$$\begin{aligned} F_\phi(M(\eta), xx') &= \lim_{\varepsilon \to 0+} \varepsilon^{-1}[\log\{\det((1-\varepsilon)M(\eta) + \varepsilon x x')/\det M(\eta)\}] \\ &= x'[M(\eta)]^{-1}x - p, \end{aligned} \tag{3.1.1}$$

after some simplification.

Let S be a $(p \times p)$ matrix with elements 2 along the principal diagonal, -1 along the diagonals just above and below the principal diagonal and 0 at each other position, e.g. with $p=4$,

$$S=\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}.$$

Since for each $x \in \mathscr{X}$, $(p/2)x'Sx-p=0$, it follows by Theorem 2.1 and (3.1.1) that if η^* is such that $M(\eta^*)=(2/p)S^{-1}$, then η^* is the D-optimal design measure. As

$$S^{-1}=(p+1)^{-1}\begin{bmatrix} p & p-1 & \dots & 2 & 1 \\ p-1 & 2(p-1) & & 4 & 2 \\ \vdots & & \ddots & & \vdots \\ 2 & 4 & & 2(p-1) & p-1 \\ 1 & 2 & \dots & p-1 & p \end{bmatrix}$$

by (2.2) it becomes evident that $M(\eta^*)=(2/p)S^{-1}$ holds if and only if η^* assigns a probability mass $2/[p(p+1)]$ to each member of $\mathscr{X}$. Hence the D-optimal design measure assigns equal mass to all members of $\mathscr{X}$.

For a similar result in a different context when the cardinality of the support of the D-optimal design equals the number of parameters we refer to Silvey (1980, p. 42).

3.2. E-optimal designs

For E-optimality,

$$\phi[M(\eta)] = \operatorname*{Inf}_{y'y \neq 0} (y'M(\eta)y/y'y)$$

and let $\bar{\eta}$ be a design measure such that $M(\bar{\eta})=p^{-1}I$, where I is the identity matrix of order p. For any $y=(y_1,\dots,y_p)'$ and any $\eta \in H$, by (2.1),

$$y'M(\eta)y=\sum_{u=1}^{p}\sum_{v=u}^{p}(y_u+y_{u+1}+\dots+y_v)^2\pi_{uv}.$$

In particular, with $y=(1,-1,1,\dots,(-1)^{p-1})'=y_0$ (say), observe that the sum of no set of consecutive elements of y_0 can exceed unity in magnitude and hence the above yields

$$y_0'M(\eta)y_0 \le \sum_{u=1}^{p}\sum_{v=u}^{p}\pi_{uv}=1,$$

so that for each $\eta \in H$,

$$\phi[M(\eta)] \le y_0'M(\eta)y_0/y_0'y_0=p^{-1}.$$

Since clearly $\phi[M(\bar{\eta})]=p^{-1}$, it follows that $\bar{\eta}$ gives the E-optimal design measure. As $M(\bar{\eta})=p^{-1}I$, by (2.2), the π_{uv}'s corresponding to $\bar{\eta}$ are, say,

$$\bar{\pi}_{uu}=p^{-1} \quad (1\le u\le p), \qquad \bar{\pi}_{uv}=0 \quad (1\le u<v\le p).$$

It may be remarked that

$$\phi[M(\bar{\eta})]=\inf_{y\neq 0}(y'M(\eta)y/y'y)$$

is not differentiable at $M(\bar{\eta})$ and, therefore, the technique of Fréchet derivatives cannot be employed in obtaining the E-optimal design measure. In a recent paper, Jacroux and Notz (1983) studied extensively optimal spring balance weighing designs in general. Some of their E-optimal designs possess the string property and hence coincide with the E-optimal ones in the present context.

4. Discussion

From a practical point of view, the object of developing the above asymptotic theory is to help with the more intractable n observation design problem. With a finite number of observations n, let $\mathcal{C}_n$ denote the class of n observation spring balance designs for p objects with string property. Then the following exact optimality results follow from the findings of Subsections 3.1, 3.2.

Theorem 4.1. *If n be a multiple of $\frac{1}{2}p(p+1)$, then the design matrix in which each h_{uv} $(1\le u\le v\le p)$ occurs as a row $2n/[p(p+1)]$ times is D-optimal within $\mathcal{C}_n$.*

Theorem 4.2. *If n is a multiple of p, then the design matrix in which each h_{uu} $(1\le u\le p)$ occurs as a row np^{-1} times is E-optimal within $\mathcal{C}_n$.*

Theorem 4.1 partially settles the problem left open in Sinha and Saha (1983) regarding exact D-optimal designs for the case $n>p$. In fact, the technique of Fréchet derivatives may also be applied to obtain the A-optimal design measure, which is, however, somewhat complicated, involves trigonometric functions and does not reduce to an exact result, as in Theorems 4.1 and 4.2, even for particular combinations of n and p. The details regarding the A-optimal measure are not, therefore, presented here and for the interested reader reference may be made to Mukerjee and Saha Ray (1983).

For general n and p, starting from the asymptotic optimality results one can construct n-observation designs which are very close to optimality specially when n is not small (see e.g. Fedorov (1972, Ch. 3), Silvey (1980, p. 37)). This is of special relevance when the available resources permit a fairly large number of observations and the problem is to take these observations in an efficient manner.

In particular, regarding exact D-optimality another result is anticipated. With n

observations let n_{uv} be the number of times h_{uv} $(1 \le u \le v \le p)$ occurs as a row of the design matrix. The findings in Subsection 3.1 and Theorem 4.1 lead to the conjecture that when n is not a multiple of $\frac{1}{2}p(p+1)$, in a D-optimal design every two n_{uv}'s should differ by at most 1. A complete enumeration of the possible situations proves the conjecture for $p=2$ or 3 and numerical examples suggest that this possibly holds for general p. A rigorous proof for general p, however, appears to be difficult and is likely to involve complicated combinatorial techniques. Under this conjecture, in looking for an n-observation D-optimal design instead of $\mathscr{C}_n$ one may consider the much smaller subclass $\mathscr{C}_{n0}$ in which the n_{uv}'s differ by at most 1. With this smaller subclass even a complete enumeration does not seem infeasible and, even if the conjecture regarding completeness of $\mathscr{C}_{n0}$ in $\mathscr{C}_n$ in terms of D-optimality is not generally true, the best design in $\mathscr{C}_{n0}$ should be at least highly efficient.

Acknowledgment

The authors are thankful to Dr. Bikas K. Sinha and Dr. S. Huda, Indian Statistical Institute, Calcutta, for their kind interest in the work and helpful suggestions. Thanks are also due to the referee for his highly constructive suggestions.

References

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

Fulkerson, D.R. and O.A. Gross (1965). Incidence matrices and interval graphs. *Pacific J. Math.* 15, 835–855.

Jacroux, M. and W. Notz (1983). On the optimality of spring balance weighing designs. *Ann. Statist.* 11, 970–978.

Mukerjee, R. and R. Saha Ray (1983). Asymptotically optimal weighing designs with string property. Tech. Report No. 19/83, Stat-Math. Division, Indian Statistical Institute, Calcutta.

Ryser, H.J. (1969). Combinatorial configurations. *SIAM J. Appl. Math.* 17, 593–602.

Silvey, S.D. (1980). *Optimal Design.* Chapman and Hall, London.

Sinha, Bikas K. and R. Saha Ray (1983). Optimal weighing designs with a string property. *J. Statist. Plann. Inference* 8, 365–374.