ORDERING EXPERIMENTAL DESIGNS

Friedrich Pukelsheim

Institut für Mathematik, Universität Augsburg, Federal Republic of Germany

Abstract. We present an overview of certain two-stage orderings of experimental designs which are such that they reflect an increase in information. These orderings use group majorization, in addition to the Loewner ordering of nonnegative definite matrices. The groups act through congruence on the moment matrices and information matrices of the problem, and a table of known results and open problems depending on the particular group is presented. The examples of quadratic regression on the symmetrized unit interval and of linear regression over the unit cube are discussed in some detail.

Key words: Information functionals, p-means, invariance, group majorization, information increasing orderings, universal optimality, simultaneous optimality.

AMS 1980 subject classification: 62K05, 62K10, 06F20.

1 Introduction

Experimental design orderings which reflect an increase in information are useful in that they allow to discriminate between competing designs. For a detailed technical derivation the reader is referred to Giovagnoli et al. (1986). A survey of the present state of experimental design theory will be found in Atkinson (1986) and Pukelsheim (1986) and the references given there.

2 Maximizing information

As usual in experimental design theory we consider a classical linear model

$$Y(x) = x'\beta + \sigma e$$

assuming uncorrelated observations with unit variance. The vectors $x \in \mathbb{R}^k$ represent the *experimental conditions*, and in their totality are assumed to form a compact set $\mathcal{X} \subset \mathbb{R}^k$, the *experimental domain*. A *design* ξ then is a discrete probability distribution on the experimental domain $\mathcal{X}$, determining allocations and frequencies of the observations.

2.1 Moment matrices

The essential quantity associated with the design ξ is its $k \times k$ *moment matrix*

$$M(\xi) = \int xx'd\xi = \sum_{i=1}^{l} \xi(x_i)x_i x_i'.$$

The set of all moment matrices forms a convex compact subset of nonnegative definite matrices. Since in some problems it is desirable to distinguish between feasible and non-feasible moment matrices, we simply assume to start from a set $\mathcal{M}$ of moment matrices which is convex and compact. This covers the case which is often dealt with that the set $\mathcal{M}$ consists of *all* moment matrices, as well as allowing for the possibility of $\mathcal{M}$ being a genuine subset of moment matrices.

2.2 Information matrices

We shall assume that an s-dimensional parameter system $K'\beta$ is of interest, where the $k \times s$ matrix K has rank s. Here the moment matrix may degenerate, with its rank varying between s and k, depending on whether the nuisance parameters remain identifiable or not. For the parameter system $K'\beta$ identifiability obtains if and only if the range of M contains the range of K. Thus the $s \times s$ *information matrix* for $K'\beta$ is defined to be

$$C(M) = \begin{cases} (K'M^-K)^{-1} & \text{in case of identifiability} \\ 0 & \text{otherwise.} \end{cases}$$

We shall tacitly assume that the parameter system $K'\beta$ is identifiable under at least one moment matrix in the set $\mathcal{M}$, in order to deal with a non-void problem.

As an example consider quadratic regression

$$Y(x_t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \sigma e = x_t'\beta + \sigma e.$$

We allow t to vary over the symmetrized unit interval, resulting in the experimental domain

$$\mathcal{X} = \left\{ x_t = \begin{pmatrix} 1 \\ t \\ t^2 \end{pmatrix} \mid t \in [-1, +1] \right\}.$$

Thus $\mathcal{X}$ essentially looks like a parabola in three-dimensional space. The parameter systems of interest here may be the set of all parameters, the subsets of any two parameters, or any single parameter, i. e.

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}, \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, \begin{pmatrix} \beta_0 \\ \beta_2 \end{pmatrix}, \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \beta_0, \beta_1, \beta_2.$$

2.3 Information functionals

The proper information matrix for the parameter system of interest actually is

$$\frac{n}{\sigma^2} C(M),$$

being directly proportional to sample size n and inversely proportional to the model variance σ^2. In a last step we need some real-valued functionals which appropriately preserve the properties of information matrices.

To this end we define an *information functional* ϕ to be a real-valued function on the set of nonnegative definite $s \times s$ matrices such that ϕ is

(a) nonnegative, since information ought to be a nonnegative quantity,

(b) positively homogeneous, whence we can dispose of the scalar factor n/σ^2,

(c) concave, because information cannot possibly be increased by interpolation, and

(d) increasing in the Loewner ordering, which actually is implied by (a) – (c).

This set of properties forms a minimal set of requirements for any specific application, while at the same time being sufficiently strong to build a general theory.

As an example we mention the *p-means*, with $p \in [-\infty, +1]$, i. e. the generalized means of order p of the eigenvalues of the information matrices. They are defined through

$$\begin{aligned} \phi_0(C) &= (\det C)^{1/s}, && \text{i. e. } p = 0, \\ \phi_p(C) &= (\operatorname{trace} C^p/s)^{1/p}, && \text{for } 0 \neq p \leq 1, \\ \phi_{-\infty}(C) &= \lambda_{\min}(C), && \text{i. e. } p = -\infty. \end{aligned}$$

In classical terms ϕ_0 is D-optimality, ϕ_{-1} is A-optimality, and $\phi_{-\infty}$ is E-optimality.

3 Ordering information matrices

Group majorization appears to be the right tool to model increasing symmetry or increasing balance of a design. This is closely related to the invariance properties of the underlying problem. A comprehensive treatment is presented in Giovagnoli et al. (1986), and we here outline only such details as are necessary to sketch the development.

Let $\tilde{G}$ be a subgroup of the general linear group $GL(k)$, and assume that $\tilde{G}$ acts linearly on the experimental conditions x, i. e.

$$x \to Qx, \text{ with } Q \in \tilde{G} \subset GL(k).$$

We give two simple examples. For quadratic regression over the symmetrized unit interval $[-1, +1]$ a natural candidate is the *sign-change group* which consists of the identity and of

$$x_t = \begin{pmatrix} 1 \\ t \\ t^2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} x_t = \begin{pmatrix} 1 \\ -t \\ t^2 \end{pmatrix}.$$

Here the group consists of two transformations only. As a second example consider linear regression over the unit cube $[0,1]^k$. In this case the permutation group is appropriate to catch the apparent symmetry, according to

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} \longrightarrow \pi x = \begin{pmatrix} x_{\pi(1)} \\ \vdots \\ x_{\pi(k)} \end{pmatrix}.$$

3.1 Induced group actions

Since our problem formulation heavily depends on moment and information matrices it is important to recognize that the linear group action on the experimental conditions x translates into *congruence action* on matrices:

$$M(\xi) = \int xx'd\xi \longrightarrow \int Qxx'Q'd\xi = QM(\xi)Q',$$

$$C(M) \longrightarrow (K'(QMQ')^{-}K)^{-1} = \check{Q}C(M)\check{Q}'.$$

In order for this to work out we must verify the following assumptions:

- The experimental domain $\mathcal{X}$ must be invariant.
- The set $\mathcal{M}$ of feasible moment matrices must be invariant.
- The reduction C from moment matrices to information matrices must be equivariant.
- The information functionals ϕ to be considered must be invariant.

We mention in passing that the matrices $\check{Q}$ which act on the information matrices C form a subgroup $\check{G}$ of the $s \times s$ general linear group, and that the passage from $\tilde{G}$ to $\check{G}$ is a group homomorphism.

3.2 Information increasing orderings

Assume from now on that the problem is invariant under a group $\tilde{G}$ as just outlined. We shall call a moment matrix B *more centered than* another moment matrix A when

$$B = \sum \alpha_i Q_i A Q_i' \in \text{ convex hull of the orbit of } A.$$

This is the usual concept of group majorization; our terminology of "being more centered" is tailored to the design problem.

The strongest reasonable ordering of moment matrices and of information matrices is, of course, the Loewner ordering defined by $M \geq B$ when $M - B$ is nonnegative definite.

The superposition of group majorization and Loewner ordering produces the *information increasing ordering* which has been found to be appropriate for the design problem, as follows. A moment matrix M is called *at least as informative* as another moment matrix A, denoted by $M \gg A$, when M is larger in the Loewner ordering than some matrix B which is more centered than A. Formally:

$$M \geq B \in \text{ convex hull of the orbit of } A, \quad \text{for some } B.$$

The corresponding information increasing ordering for information matrices will also be denoted by $\gg$. That these information preorderings nicely agree with the various levels of our problem is shown by the following.

Theorem. *(Giovagnoli et al. (1986))*

$$M \gg A$$
$$\Longrightarrow C(M) \gg C(A)$$
$$\Longrightarrow \phi(C(M)) \geq \phi(C(A)), \text{ for all invariant } \phi.$$

3.3 Universal optimality vs. simultaneous optimality

The preceeding theorem suggests to discriminate between the notions of *universal optimality* whenever

$$C \gg D, \text{ for all competing } D,$$

and of *simultaneous optimality* whenever

$$\phi(C) \geq \phi(D), \text{ for all competing } D \text{ and for all invariant } \phi.$$

Frequently these notions will coincide according to the following.

Theorem. *(Giovagnoli et al. (1986)) If the underlying group is compact and the information matrix C is invariant then*

$$C \text{ is universally optimal } \iff C \text{ is simultaneously optimal.}$$

When the group fails to be compact or the matrix C is not invariant it seems that the notion of simultaneous optimality is of a greater bearing. The following table gives an overview of some known results and open problems.

group	ordering	invariant functionals
$\{I_n\}$	Loewner	all ϕ
Perm(s)	?	?
Orth(s)	upper weak majorization of ordered eigenvalues	symmetric functions of ordered eigenvalues
Unim(s)	?	determinant
reflection groups	?	?
?	?	p-means

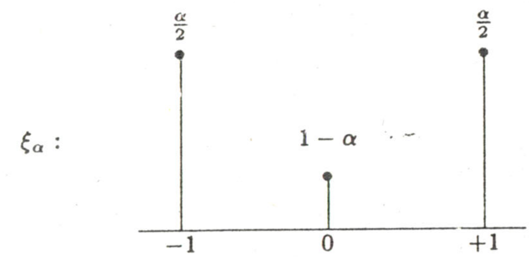
As an outstanding result we mention that this provides a further justification for the most popular criterion of D-optimality as being the sole invariant information functional under the group of unimodular linear transformations (i. e. those with determinant ± 1). On the other hand it would be of interest to study finite reflection groups as they also arise in other aspects of multivariate analysis, or to find a group such that the invariant functionals are determined by the p-means.

4 Quadratic regression; regression over the unit cube

As mentioned above the model for quadratic regression over the symmetrized unit interval $[-1,+1]$ is

$$Y(x_t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \sigma e.$$

A design ξ is invariant under the *sign-change group* if and only if ξ is symmetric about 0. This reduces the corresponding moment matrices to a two-parameter subset. If we augment this with an improvement in the Loewner ordering we obtain a reduction to the one-parameter family of *symmetric three-points designs* ξ_α given by



This approach yields the following results, cf. Preitschopf and Pukelsheim (1986): *(a) For every ξ there exists some α such that $\xi_\alpha \gg \xi$. (b) For every p there exists some $\alpha(p)$ such that $\xi_{\alpha(p)}$ is p-optimal.*

Another very instructive example is linear regression over the unit cube which has recently been resolved in a brilliant paper by Cheng (1986). With experimental conditions x varying over the k-dimensional unit cube $[0,1]^k$ the model

$$Y(x) = x'\beta + \sigma e$$

is invariant under the permutation group. Now an invariant design ξ has a moment matrix $M(\xi)$ which belongs to the two-parameter family of completely symmetric matrices (i. e. having identical on-diagonal elements and identical off-diagonal elements). At this stage the General Equivalence Theorem is invoked to obtain a further reduction to the one-parameter family of *uniform vertex designs* ξ_t, defined as follows.

A vertex x of the unit cube will be called a *c-vertex* if x has c components equal to unity and $n-c$ components equal to zero. The unique permutation invariant design which is supported by the c-vertices is the uniform distribution on the c-vertices and will be denoted by ξ_c. In addition we also need mixture designs as defined by

$$\xi_t = (1-(t-c))\xi_c + (t-c)\xi_{c+1}, \quad \text{with } t \in (c, c+1),$$

which are the permutation invariant designs supported by the c-vertices and the $(c+1)$-vertices and hence are convex combinations of ξ_c and ξ_{c+1}. The parameterization chosen evidently is continuous. The family of uniform vertex designs now is given by ξ_t with t varying continuously between 0 and k.

Cheng (1986) proves the following result: *For all p there exists some $t(p)$ such that $\xi_{t(p)}$ is p-optimal.* Moreover he derives an explicit formula for $t(p)$. A monotone behaviour emerges, in that as p increases from $-\infty$ towards 1 one has that $t(p)$ increases from the integer part of $(k+1)/2$ towards k. More precisely $t(p)$ is constantly equal to an integer value c over closed intervals of p, and strictly increasing in-between. Almost all of the variation of $t(p)$ occurs for positive values of p, a qualitative feature which is also encountered in quadratic regression.

References

Atkinson, A. C. (1986). Recent developments in the design of experiments. *Internat. Statist. Rev.*, to appear.

Cheng, C.-S. (1986). An application of the Kiefer-Wolfowitz equivalence theorem. *University of California, Berkeley*, Preprint, 14 pp.

Giovagnoli, A., Pukelsheim, F., and Wynn, H. P. (1986). Group invariant orderings and experimental designs. *J. Statist. Plann. Inference*, to appear.

Preitschopf, F., and Pukelsheim, F. (1986). Optimal designs for quadratic regression. *J. Statist. Plann. Inference*, to appear.

Pukelsheim, F. (1986). Information increasing orderings in experimental design theory. *Internat. Statist. Rev.*, to appear.

Proceedings of the Ist World Congress of the BERNOULLI SOCIETY

Tashkent, USSR
8-14 September 1986

Volume 2
Mathematical Statistics
Theory and Applications

Editors
Yu. A. Prohorov and V. V. Sazonov

VNU SCIENCE PRESS
Utrecht, The Netherlands
1987