STATISTIQUE ET ANALYSE DES DONNÉES

ZIMING ZHANG
CHANGYU WANG

On entropy optimality of orthogonal design

Statistique et analyse des données, tome 8, nº 2 (1983), p. 64-80
<http://www.numdam.org/item?id=SAD_1983__8_2_64_0>

Statistiques et Analyse de données .64.
1983 - Vol. 8 n° 2 - pp. 64-80

ON ENTROPY OPTIMALITY OF ORTHOGONAL DESIGN

ZHANG ZIMING, WANG CHANGYU

Mathematical Department
Qu Fu Teacher's College
QUFU, SHANDONG, 273164
P.R. CHINA

Résumé : *Cet article présente une étude de l'optimalité des plans orthogonaux, et compare les notions de D, A et E optimalité et d'entropie.*

Abstract : *This paper presents a study on optimality of orthogonal arrays and compares the notions of D, A and E optimality and of entropy.*

Mots clés : *Plans d'expérience orthogonaux*

Note de l'éditeur :

Nous avons choisi de publier cet article en témoignage des activités de recherche en statistique en République Populaire de Chine, malgré un certain nombre d'insuffisances. En particulier nous avons complété la bibliographie par quelques références plus accessibles en Europe préparées par l'un des rapporteurs qui ont lu le manuscrit.

I - INTRODUCTION AND MATHEMATICAL MODEL

Orthogonal Design is a method of arranging several factor tests by a set of orthogonal tables. By using this method better technological conditions may be selected with less tests. This method of design is widely used in China since 1974.

The purpose of this paper is to analyse the orthogonal design from the theoretical point of view. At first we shall introduce the mathematical model of tests design.

Consider the problem of the index μ affected by m factors F^1 (S_1 levels in whole), F^2 (S_2 levels in whole),...F^m (S_m levels in whole). The theoretical value (or real value) of the index μ is $\mu(\lambda_1,\lambda_2,\ldots,\lambda_m)$ when F^j takes the level λ_j ($j = 1,2,\ldots,m$). Note that there are altogether $\prod_{j=1}^{n} S_j$ theoretical values $\mu(\lambda_1,\lambda_2,\ldots,\lambda_m)$ $(1 \leq \lambda_j \leq S_j,\ j = 1,2,\ldots,m)$ which exist but can not be observed. We can only observe the test value

$$\mu(\lambda_1,\lambda_2,\ldots,\lambda_m) = \mu(\lambda_1,\lambda_2,\ldots,\lambda_m) + \varepsilon(\lambda_1,\lambda_2,\ldots,\lambda_m)$$

where $\varepsilon(\lambda_1,\lambda_2,\ldots,\lambda_m)$ is called test error which is a random variable.

The purpose of test design is to select parts of combinatorics as little as possible from $\prod_{j=1}^{m} S_j$ level combinatorics. Then statistics deduction is carried by making use of the test values obtained.

<u>Definition 1</u>

Assume $S_1, S_2,\ldots,S_m$ are positive integers not smaller than 2. Matrix $\Lambda(\lambda_{ij})_{n \times m}$ is called a design of D_n $(S_1 \times S_2 \times\ldots\times S_m)$ type if

$$\lambda_{ij} \in \{1,2,\ldots,S_j\},\ (1 \leq i \leq n,\ 1 \leq j \leq m)$$

Let

$$\theta_o = \frac{\sum_{\lambda_1\ldots\lambda_m} \mu(\lambda_1,\lambda_2,\ldots,\lambda_m)}{m}$$

where λ_1 runs from 1 to $S_1,\ldots,\lambda_m$ runs from 1 to S_m. Obviously θ_0 is an average value of all the theoretical value.

Let

$$\theta_{j\lambda_j} = \frac{\sum\limits_{\lambda_1,\ldots\lambda_{j-1}\lambda_{j+1}\ldots\lambda_m} \mu(\lambda_1,\ldots,\lambda_{j-1},\lambda_j,\lambda_{j+1},\ldots,\lambda_m)}{\prod\limits_{\substack{i=1\\i\neq j}}^{m} S_i} - \theta_0$$

$\theta_{j\lambda_j}$ is called the main effect of factor F^j under the level (effect for brief).

<u>Definition 2</u>

The theoretical value $\mu(\lambda_1,\lambda_2,\ldots,\lambda_m)$ of factors F^1, $F^2,\ldots,F^m$ is called accordance with addible model if

$$\mu(\lambda_1,\lambda_2,\ldots,\lambda_m) = \theta_0 + \sum_{j=1}^{m} \theta_{j\lambda_j} \tag{1.1}$$

holds for all $1 \le \lambda_1 \le S_1$,..., $1 \le \lambda_m \le S_m$.

From the definition of the main effect we have

$$\sum_{\lambda_j=1}^{S_j} \theta_{j\lambda_j} = 0 \tag{1.2}$$

holds for all j.

Obviously it is sufficient to estimate the parameters θ_0, $\theta_{j\lambda_j}$ $(1 \le \lambda_j \le S_j,\ j = 1,2,\ldots,m)$ in order to study the property of $\mu(\lambda_1^j,\lambda_2,\ldots,\lambda_m)$ of the addible model. Since the test values have error, it's better to have more accurate estimate of these parameters. Accuracy of estimate is closely related to the number of tests and test design. Faculty of Division of Probability Theory, Mathematics Department, Beijing University, China, have proved that accuracy of parameter estimate satisfies to so called A-optimality, E-optimality, D-optimality etc. If we apply the orthogonal design. On the basis of their work, we prove here that accuracy of parameter estimate satisfies entropy optimality if orthogonal design is applied. We have known that Shannon entropy is numerical index reflecting uncertainty of random variables. The entropy optimality of orthogonal design shows that the estimation of effect of various levels of each factor has the smallest uncertainty. More significantly, A-optimality, E-optimality, D-optimality etc. of the orthogonal design can be looked as some special cases of entropy optimality In addition, D-optimality is further strengthened in this paper.

This paper discusses the addible model only.

Suppose that we get n test values from (1.1) under the design $\Lambda = (\lambda_{ij})_{n \times m}$ of $D_n(S_1 \times S_2 \times \dots \times S_m)$ type.

$$\begin{cases} y_1 = \theta_o + \sum_{j=1}^{m} \lambda_{j\lambda_{1j}} + \varepsilon_1 \\ y_2 = \theta_o \quad \sum_{j=1}^{m} \theta_{j\lambda_{2j}} + \varepsilon_2 \\ \vdots \\ y_i = \theta_o + \sum_{j=1}^{m} \theta_{j\lambda_{ij}} + \varepsilon_i \\ y_n = \theta_o + \sum_{j=1}^{m} \theta_{j\lambda_{nj}} + \varepsilon_n \end{cases} \tag{1.3}$$

where $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ are errors of tests. (1.3) will be denoted by matrices in the following. Let

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \qquad \theta_j = \begin{bmatrix} \theta_{j1} \\ \theta_{j2} \\ \vdots \\ \theta_{jS_j} \end{bmatrix} \quad (j = 1,2,\dots,m),$$

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}, \qquad \mathbb{1}_S = \left.\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}\right\} S \text{ row}, \quad \theta = \begin{bmatrix} \theta_o \\ \theta_1 \\ \vdots \\ \theta_m \end{bmatrix}$$

$$x_{ik}^{(j)} = \delta(\lambda_{ij}, k), \qquad X_j = (x_{ik}^{(j)})_{n \times S_j}$$

$$X = (\mathbb{1}_n, X_1, \dots, X_m)$$

where

$$\delta(u,v) = \begin{cases} 1, & \text{when } u = v \\ 0, & \text{when } u \neq v \end{cases}$$

Then (1.3) can be written as

$$y = X\theta + \varepsilon \tag{1.4}$$

where X is determined uniquely by the design Λ , called the subordinate matrix of Λ and denoted as $X(\Lambda)$ (or X for brief).
We further assume that

$$E\varepsilon = 0, \qquad D(\varepsilon) = \sigma I_n \tag{1.5}$$

where E is mathematical expectation, D is covariance and I_n is $n \times n$ unit matrix.

We can eliminate useless parameters through transformation since parameters $\theta_{j\lambda_j}$ satisfy constrained condition (1.2).

Let Γ_j be S_j order orthogonal matrix, having the structure

$$\Gamma_j = (\frac{1}{S_j} \mathbb{1}_{S_j}, T_j) \qquad (j = 1,2,\ldots,m) \tag{1.6}$$

Let $\eta_j = T_j'\theta_j$ $(j = 1,2,\ldots,m)$, $\eta_0 = \theta_0$.

Then

$$T_j\eta_j = T_jT_j'\theta_j = (I_{S_j} - \frac{\mathbb{1}_{S_j}\mathbb{1}_{S_j}'}{S_j})\,\theta_j$$

$$= \theta_j - \frac{1}{S_j}\mathbb{1}_{S_j}(\mathbb{1}_{S_j}'\theta_j) = \theta_j,$$

and

and $$X\theta = (\mathbb{1}_n, X_1, \ldots, X_m) \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_m \end{bmatrix}$$

$$= (\mathbb{1}_n, X_1, \ldots, X_m) \begin{bmatrix} \eta_o \\ T_1 \eta_1 \\ \vdots \\ T_m \eta_m \end{bmatrix}$$

$$= (\mathbb{1}_n, X_1 T_1, \ldots, X_m T_m) \begin{bmatrix} \eta_o \\ \eta_1 \\ \vdots \\ \eta_m \end{bmatrix}$$

Denote

$$\tilde{X} = (\mathbb{1}_n, X_1 T_1, \ldots, X_m T_n), \quad \eta = \begin{bmatrix} \eta_o \\ \eta_I \\ \vdots \\ \eta_m \end{bmatrix}$$

Then (1.4) can be written as

$$y = \tilde{X}\eta + \varepsilon \qquad (1.7)$$

The component of η is unconstrained. The design Λ is restricted as follows in order to guarantee that the design can be estimated.

Definition 3

The design $\Lambda = (\lambda_{ij})$ of $D_n(S_1 \times S_2 \times \ldots \times S_n)$ type is feasible if the rank of subordinate matrix $X(\Lambda)$ of Λ is equal to $\sum_{j=1}^{m} S_j - m + 1$ i.e.

$$R(X(\Lambda)) = \sum_{j=1}^{m} S_j - m + 1$$

It is easy to prove that $R(\tilde{X}) = R(X(\Lambda))$ from (1.6). According to the theory of linear model, the components of η can be estimated if and only if the columns of $\tilde{X}$ are linear independent i.e. $R(\tilde{X}) = 1 + \sum_{j+1}^{m} (S_j - 1)$. In other words, the components of Λ can be estimated if and only if is a feasible design. From now on we shall discuss feasible design only. The optimal estimation of η is

$$\hat{\eta} = (\tilde{X}' \tilde{X})^{-1} \tilde{X}' y \qquad (1.8)$$

for feasible design Λ .

The optimal estimation of θ is

$$\hat{\theta} = \begin{pmatrix} 1 & & & \\ & T_1 & & \\ & & \ddots & \\ & & & T_m \end{pmatrix} \hat{\eta} \qquad (1.9)$$

The optimal estimation of $\mu(\lambda_1,\lambda_2,\ldots,\lambda_m)$ is

$$\hat{\mu}(\lambda_1,\lambda_2,\ldots,\lambda_m) = \hat{\theta}_0 + \sum_{j=1}^{m} \hat{\theta}_{j\lambda_j} \qquad (1.10)$$

where $\hat{\eta}$ and $\hat{\theta}$ are related to the design Λ . The definition of orthogonal design is as follows.

Definition 4

$\Lambda = (\lambda_{ij})_{n \times m}$ is orthogonal design of $D_n(S_1 \times S_2 \times \ldots \times S_m)$ type if it satisfies :

1°) $\lambda_{ij} \in \{1,2,\ldots,S_j\}$ for any j ($j = 1,2,\ldots,m$) and the number of elements of the set $\{i : \lambda_{ij} = K\}$ is $\frac{n}{S_j}$ for any $K \in \{1,2,\ldots,S_j\}$.

2°) The number of element of the set

$$\{i : \lambda_{ij} = u, \quad \lambda_{i1} = v\}$$

is $\frac{n}{S_j S_1}$ for any $j \neq 1$, $u,v \in \{1,2,\ldots,S_j\}$

(Note : the orthogonal design of $D_n(S_1 \times \ldots \times S_m)$ type may not exists for any given positive integers n, $s_1, s_2,\ldots,s_m$).

Definition 5 (A-optimality)

The feasible design $\hat{\Lambda}$ of $D_n(S_1 \times S_2 \times \ldots \times S_m)$ type is A-optimal if

$$D[\hat{e}_o(\Lambda)] \geq D[\hat{\theta}_o(\Lambda)]$$

$$\frac{\sum_{k=1}^{S_j} D[\hat{e}_{jk}(\Lambda)]}{S_j} \geq \frac{\sum_{k=1}^{S_j} D[\hat{\partial}_{jk}(\tilde{\Lambda})]}{S_j} \quad (1 \leq j \leq m) \qquad (1.11)$$

holds for any feasible design Λ of $D_n(S_1 \times S_2 \times \ldots \times S_m)$ type. (Note (1.11) is equivalent to $tr[D\ \hat{\theta}_j(\Lambda)] \geq tr\ D[\hat{\theta}_j(\tilde{\Lambda})]$ where tr denote the trace of matrix).

Definition 6 (E-optimality)

The feasible design $\tilde{\Lambda}$ of the $D_n(S_1 \times S_2 \times \ldots \times S_m)$ type is called E optimality if

$$D[\hat{\theta}_o(\Lambda)] \geq D[\hat{\theta}_o(\tilde{\Lambda})]$$

$$\max_{\alpha_j:\alpha_j' 1 S_j=0} \frac{D[\alpha_j' \hat{\theta}_j(\Lambda)]}{\alpha_j' \alpha_j} \geq \max_{\alpha_j:\alpha_j' 1 S_j=0} \frac{D[\alpha_j' \hat{\theta}_j(\tilde{\Lambda})]}{\alpha_j' \alpha_j}$$

$$(1 \leq j \leq m)$$

hold for any feasible design of the $D_n(S_1 \times \ldots \times S_m)$ type, where the components of

$$\alpha_j = \begin{bmatrix} \alpha_{j1} \\ \vdots \\ \alpha_{jS_j} \end{bmatrix}$$

are not all zero.

Definition 7 (D-optimality)

The feasible design $\tilde{\Lambda}$ of the $D_n(S_1 \times \ldots \times S_m)$ type is called D optimality if

$$D[\hat{\theta}_o(\Lambda)] \geq D[\hat{\theta}_o(\tilde{\Lambda})]$$

$$|D[\hat{n}_j(\Lambda)]| \geq |D[\hat{n}_j(\tilde{\Lambda})]|$$

$$(1 \leq j \leq m)$$

where $\hat{\eta}_j(\Lambda) = T_j'\hat{\theta}_j(\Lambda)$, T_j' is any matrix satisfying (1.6), "$|\ \ |$" denotes determination.

2 - A-, E-, D-OPTIMALITY OF ORTHOGONAL DESIGN

We present briefly the results obtained by the Division of Probability, Department of Mathematics, Beijing University in this section [2].

LEMMA :

Any orthogonal design is feasible design of $D_n(S_1 \times S_2 \times \ldots \times S_m)$ type.

THEOREM 2.1

Let Λ be a feasible design of the $D_n(S_1 \times \ldots \times S_m)$ type then

$$D[\hat{\theta}_o(\Lambda)] \geq \frac{\sigma^2}{n}$$

$$\mathrm{tr}[D\ \hat{\theta}_j(\Lambda)] \geq \sigma^2\ \frac{S_j(S_j-1)}{n}$$

The necessary and sufficient condition for all the equality to hold is that Λ is an orthogonal design.

THEOREM 2.2

For any feasible design Λ of $D_n(S_1 \times \ldots \times S_m)$ type, all the following hold :

(1) $$D[\hat{\theta}_o(\Lambda)] \geq \frac{\sigma^2}{n}$$

(2) $$\max_{\alpha_j\ \Omega_j} \frac{D[\alpha_j'\hat{\theta}_j(\Lambda)]}{\alpha_j'\ \alpha_j} \geq \frac{\sigma^2 S_j}{n} \qquad (1 \leq j \leq m)$$

where

$$\Omega_j = \{(\alpha_{j1},\ldots,\alpha_{jS_j})' : \sum_{i=1}^{S_j} \alpha_{ji} = 0, \quad \text{but}$$

$$\alpha_{ji}\ (1 \leq i \leq S_j) \text{ are not all zero}\}.$$

The necessary and sufficient condition for all the equality in (1) and (2) to hold is that Λ is an orthogonal design.

THEOREM 2.3

For any feasible design Λ of $D_n(S_1 \times \ldots \times S_m)$ type, all the following hold :

(1) $$D[\hat{\theta}_0(\Lambda)] \geq \frac{\sigma^2}{n}$$

(2) $$|D[T_j^{\prime}\hat{\theta}_j(\Lambda)]| \geq \left(\frac{\sigma^2 S_j}{n}\right)^{S_j-1} \qquad (1 \leq j \leq n)$$

The necessary and sufficient condition for the equality in (1) and (2) to hold is that Λ is an orthogonal design.

The following theorem further strengthens Theorem 2.3.

THEOREM 2.4

For any feasible design Λ of $D_n(S_1 \times S_2 \times \ldots \times S_m)$ type,

$$|D[\hat{\eta}(\Lambda)]| \geq \frac{\sigma^2}{n} \prod_{j=1}^{m} \left(\frac{\sigma^2 S_j}{n}\right)^{S_j-1}$$

holds and the necessary and sufficient condition for the equality to hold is that Λ is an orthogonal design, where $\hat{\eta}(\Lambda)$ is determined by (1.9).

THEOREM 2.5

For any feasible design Λ of $D_n(S_1 \times \ldots \times S_m)$ type

$$\frac{1}{N} \operatorname{tr}[D\, \hat{\mu}(\Lambda)] \geq \frac{P}{n} \sigma^2$$

holds and the necessary and sufficient condition for the equality to hold is that Λ is an orthogonal design, where

$$N = \prod_{j=1}^{m} S_j, \quad P = 1 + \sum_{j=1}^{m} (S_j - 1),$$

$\hat{\mu}(\Lambda)$ is estimation of engineering mean value (See Definition 8 and (1.10)).

3 - ENTROPY OPTIMALITY OF ORTHOGONAL DESIGN

We shall generalize and strengthen D-optimality of orthogonal design by using the property of normal distribution and unifying A-optimality, E-optimality and D-optimality by establishing entropy optimality.

Theorem 2.3 and 2.4 are only the special case of the following Theorem 3.1, which shows that orthogonal design has more strong D-optimality.

THEOREM 3.1

$$|D[\hat{\eta}_{j1}(\Lambda)', \hat{\eta}_{j2}(\Lambda)', \dots, \hat{\eta}_{jr}(\Lambda)']| \geq \frac{\sigma^2}{n} \prod_{p=1}^{r} \left(\frac{\sigma^2 S_{j_p}}{n}\right)^{S_{j_p}^{-1}}$$

holds for any feasible design and any subset $(j_1, j_2, \dots, j_r) \subset (1, 2, \dots, m)$. The necessary and sufficient condition for the equality to hold is that Λ is an orthogonal design.

The following Lemma is first proved in order to prove this theorem.

LEMMA 3.1

Assume that k-dimension random vector $(\xi_1, \xi_2, \dots, \xi_k)$ subordinate normal distribution $N(a, B)$, where $a = (a_1, a_2, \dots, a_k)$, B is k order positive definite symmetrical matrix, then any sub-vector $(\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_t})$, $(t \leq k)$ of $(\xi_1, \xi_2, \dots, \xi_k)$ subordinates normal distribution $N(\tilde{a}, \tilde{B})$ too, where $\tilde{a} = (a_{i_1}, a_{i_2}, \dots, a_{i_t})$, $\tilde{B}$ is at order submatrix of B containing $i_1, i_2, \dots, i_t$ columns and rows.

The proof is in [1]. In addition, we need two well-known inequalities in matrix theory.

LEMMA 3.2

Let A be a positive definite matrix and

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1t} \\ A_{21} & A_{22} & \cdots & A_{2t} \\ \cdots & \cdots & \cdots & \cdots \\ \Delta & \Delta & & \Delta \end{pmatrix}$$

where A_{ii} is square matrix (i = 1, 2,...t), let

$$A^{-1} = \begin{pmatrix} A^{11} & A^{12} & \dots & A^{1t} \\ A^{21} & A^{22} & \dots & A^{2t} \\ \hdashline A^{t1} & A^{t2} & \dots & A^{tt} \end{pmatrix}$$

then $$A^{ii} \geq A_{ii}^{-1} \qquad (i = 1, 2, \dots, t)$$

and the necessary and sufficient condition for the equality to hold is that $A_{ij} = 0 \quad (i \neq j,\ i,j = 1,2,\dots,t)$

LEMMA 3.3.

Let

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1t} \\ A_{21} & A_{22} & \dots & A_{2t} \\ \hdashline A_{t1} & A_{t2} & \dots & A_{tt} \end{pmatrix}$$

be positive definite matrix, where A_{ii} is square matrix (i = 1,2,...,t) then

$$|A| \leq \prod_{i=1}^{t} |A_{ii}|$$

LEMMA 3.4

If A is p order non-negative definite matrix, then $|A| \leq (\frac{1}{p}\,\mathrm{tr}\,A)^p$, and the necessary and sufficient condition for the equality to hold is that $A = \lambda I_p$ (λ is a non-negative number).

The proof of Theorem 3.1 : From (1.8) we can get

$$D[\hat{\eta}(\Lambda)] = \sigma^2(\tilde{X}',\tilde{X})^{-1}.$$

where $\tilde{X}'X = (\mathbb{1}_n, X_1T_1, \ldots, X_mT_m)'(\mathbb{1}_n, X_1T_1, \ldots, X_mT_m)$

$$= \begin{pmatrix} n & \mathbb{1}_n'X_1T_1 & \cdots & \mathbb{1}_n'X_jT_j & \cdots & \mathbb{1}_n'X_mT_m \\ T_1'X_1'\mathbb{1}_n & T_1'X_1'X_1T_1 & \cdots & T_1'X_1'X_jT_j & \cdots & T_1'X_1'X_mT_m \\ \hline T_m'X_m'\mathbb{1}_n & T_m'X_m'X_1T_1 & \cdots & T_m'X_m'X_jT_j & \cdots & T_m'X_m'X_mT_m \end{pmatrix}$$

Divide $D[\hat{n}(\Lambda)]$ into submatrices according to submatrices of :

$$D[\hat{n}(\Lambda)] = \begin{pmatrix} D_{00} & D_{01} & \cdots & D_{0j} & \cdots & D_{0m} \\ D_{10} & D_{11} & \cdots & D_{1j} & \cdots & D_{1m} \\ \hline D_{m0} & D_{m1} & \cdots & D_{mj} & \cdots & D_{mm} \end{pmatrix}$$

Rearrange $D[\hat{n}(\Lambda)]$ by moving $i_1, i_2, \ldots, i_t$ row blocks and $i_1, i_2, \ldots, i_t$ column blocks into 1, 2,...,t row blocks and 1,2,...,t column blocks. $D[\hat{n}(\Lambda)]$ is changed into

$$\tilde{D}[\hat{n}(\Lambda)] = \begin{pmatrix} D_\Omega & X \\ X & X \end{pmatrix} \tag{3.1}$$

the inverse of $\tilde{D}\,\hat{n}(\Lambda)$ will be

$$\tilde{D}[\hat{n}(\Lambda)]^{-1} = \begin{pmatrix} R & X \\ X & X \end{pmatrix} \tag{3.2}$$

where

$$\frac{1}{\sigma^2} R = \begin{pmatrix} T_{i_1}' X_{i_1}' X_{i_1} T_{i_1} & T_{i_1}' X_{i_1}' X_{i_2} T_{i_2} & \cdots & T_{i_1}' X_{i_1}' X_{i_t} T_{i_t} \\ T_{i_2}' X_{i_2}' X_{i_1} T_{i_1} & T_{i_2}' X_{i_2}' X_{i_2} T_{i_2} & \cdots & T_{i_2}' X_{i_2}' X_{i_t} T_{i_t} \\ \hline T_{i_t}' X_{i_t}' X_{i_1} T_{i_1} & T_{i_t}' X_{i_t}' X_{i_2} T_{i_2} & \cdots & T_{i_t}' X_{i_t}' X_{i_t} T_{i_t} \end{pmatrix}$$

(3.3)

(Let $X_{i_r} T_{i_r} = \mathbb{1}_n$ when $i_r = 0$).

From Lemma 3.1, we have

$$D[\hat{\eta}_{i_1}(\Lambda)', \hat{\eta}_{i_2}(\Lambda)', \ldots, \eta_{i_t}(\Lambda)'] = D_\Omega. \tag{3.4}$$

If Lemma 3.2 is applied to (3.1) and (3.2),

$$\sigma^2 D[\hat{\eta}_{i_1}', \hat{\eta}_{i_2}', \ldots, \hat{\eta}_{i_t}']^{-1} \le R \tag{3.5}$$

holds. If Lemma 3.3 is applied to (3.3)

$$\left| \begin{pmatrix} T_{i_1}' X_{i_1}' X_{i_1} T_{i_1} & \cdots & T_{i_1}' X_{i_1}' X_{i_t} T_{i_t} \\ \hline T_{i_t}' X_{i_t}' X_{i_1} T_{i_1} & \cdots & T_{i_t}' X_{i_t}' X_{i_t} T_{i_t} \end{pmatrix} \right| \le \prod_{r=1}^{t} |T_{i_r}' X_{i_r}' X_{i_r} T_{i_r}| \tag{3.6}$$

holds. From (3.3),(3.5) and (3.6)

$$D[\hat{\eta}_{i_1}, \hat{\eta}_{i_2}, \ldots, \hat{\eta}_{i_t}]^{-1} \le (S_{i_r}-1) \prod_{r=1}^{t} |T_{i_r}' X_{i_r}' X_{i_r} T_{i_r}| \tag{3.7}$$

holds for any subset $\{i_1, i_2, \ldots, i_m\} \subset \{0, 1, \ldots, m\}$.
From the definition of feasible design, (1.6) and Lemma 3.4 we have

$$|T_{i_r}' X_{i_r}' T_{i_r} X_{i_r}| \le \left(\frac{n}{S_{i_r}}\right)^{S_{i_r}-1}$$

From (3.7) we have

$$D[\hat{\eta}_{i_1}, \hat{\eta}_{i_2}, \ldots, \hat{\eta}_{i_t}] \ge \prod_{r=1}^{t} \left(\frac{\sigma^2 S_{i_r}}{n}\right)^{S_{i_r}-1}$$

From Theorem 2.3 we know that the necessary and sufficient condition for the equality above to hold is that Λ is an orthogonal design.

Definition 3.1

Let $C(\Lambda)$ be a non-zero non-negative definite matrix and Λ be a feasible design of $D_n(S_1 \times S_2 \times \ldots \times S_m)$ type. The feasible design is called entropy optimality under transformation C, if the inequality

$$H[C(\Lambda)\hat{\eta}(\Lambda)] \geq H[C(\tilde{\Lambda})\hat{\eta}(\tilde{\Lambda})]$$

holds for any feasible design Λ .
Where $H(\xi)$ denotes Shonnon entropy of random vector ξ.

THEOREM 3.2

Let $C(\Lambda)$ be a non-negative definite matrix satisfying

$$|C(\Lambda)| = D[\hat{\theta}_0(\Lambda)] \prod_{j=1}^{m} \text{tr } D[\hat{\theta}_j(\Lambda)]$$

Then the necessary and sufficient condition for $\tilde{\Lambda}$ to be an orthogonal design is that $\tilde{\Lambda}$ is entropy optimal under transformation C i.e.

$$H[C(\Lambda)\hat{\eta}(\Lambda)] \geq H[C(\tilde{\Lambda})\hat{\eta}(\tilde{\Lambda})]$$

Proof : Let $P(X_1,X_2,\ldots,X_p)$ be distribution density of $\hat{\eta}(\Lambda)$. When $P(X_1,X_2,\ldots,X_p)$ is normal distribution it is easy to compute

$$H[C(\Lambda)\hat{\eta}(\Lambda)] = H[\hat{\eta}(\Lambda)]+$$

$$+\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} P(x_1,x_2,\ldots,x_p) \text{ Log}|C(\Lambda)|dx_1,\ldots,dx_p$$

$$= \text{Log } \sqrt{(2\pi e)^p . |D[\hat{\eta}(\Lambda)]|} + \text{Log}|C(\Lambda)|$$

$$= \text{Log } \sqrt{(2\pi e)^p . |D[\hat{\eta}(\Lambda)]|} + \text{Log}|D[\hat{\theta}_0(\Lambda)]| + \sum_{j=1}^{m} \text{Log tr } D[\hat{\theta}_j(\Lambda)]$$

From Theorem 2.1 and Theorem 2.4, the necessary and sufficient condition for $\tilde{\Lambda}$ to be an orthogonal design is that $\tilde{\Lambda}$ is entropy optimal under transformation C.

THEOREM 3.3

Let $C(\Lambda)$ be non-negative definite matrix satisfying

$$|C(\Lambda)| = D[\hat{\theta}_0(\Lambda)] \prod_{j=1}^{m} \max_{\alpha_j \, \Omega_j} \frac{D[\alpha_j'\hat{\theta}_j(\Lambda)]}{\alpha_j'\alpha_j}$$

where

$$\Omega_j = \{(\alpha_{j_1},\ldots,\alpha_{jS_j})' : \sum_{\lambda_j=1}^{S_j} \alpha_{j\lambda_j} = 0, \; \alpha_{j\lambda_j} \; (1 \le \lambda_j \le S_j) \text{ are not all zero}\}$$

then the necessary and sufficient condition for $\tilde{\Lambda}$ to be an orthogonal design is that $\tilde{\Lambda}$ is entropy optimal under transformation C.

Proof : We can imitate the proof of Theorem 3.2 by using Theorem 2.2 and 2.4.

THEOREM 3.4

Let

$$C(\Lambda) = C_{j_1 j_2 \ldots j_t}(\Lambda) = \begin{bmatrix} J_0 & & & & O \\ & J_1 & & & \\ & & \ddots & & \\ & & & J_j & \\ O & & & & \ddots \; J_m \end{bmatrix}$$

where

$$J_j = \begin{cases} I_{S_j-1} & j \in \{j_1, j_2, \ldots, j_t\} \\ 0 & j \bar{\in} \{j_1, j_2, \ldots, j_t\} \end{cases}$$

(Let $S_0 = 1$, $I_0 = 1$). Then the necessary and sufficient condition for $\tilde{\Lambda}$ to be an orthogonal design is that $\tilde{\Lambda}$ is entropy optimal under transformation C.

Proof :

$$H[C_{j_1 \ldots j_t}(\Lambda)\hat{\eta}(\Lambda)] = H[\hat{\eta}_{j_1}(\Lambda)', \hat{\eta}_{j_2}(\Lambda)', \ldots, \hat{\eta}_{j_t}(\Lambda)']$$

$$= \text{Log} \sqrt{(2\pi e)^p \, | D[\hat{\eta}_{j_1}, \hat{\eta}_{j_2}, \ldots, \hat{\eta}_{j_t}]|}$$

From Theorem 3.1, the results is followed.

Finally we point out that Theorem 2.5 can be looked as entropy optimal under some transformation C.

BIBLIOGRAPHIE

(1) H. Cramer, Mathematical Methods of statistics.
Princeton 1946

(2) Division of Probability, Department of Mathematics, Beijing University.
"On optimality of orthogonal design (I), (II)".
Acta Mathematicae Applicatae Sinica n°1, n°2 (1977).

BIBLIOGRAPHIE COMPLEMENTAIRE

(1) Raktoe B.L., Hedayat A., Federer W.T. (1981)
Factorial Designs, Jonh Wiley & Sons, N.Y., 209 p.

(2) Rao C.R. (1973)
Some combinatorial problems of arrays and applications to design of experiments. In A Survey of Combinatorial Theory (J.N. Srivastava Ed). p. 349-359 North Holland, Amsterdam.

(3) Seber G.A.F. (1964)
Orthogonality in Analysis of Variance. Ann. Math. Statist., 35, p.705-710

(4) Mukerjee R. (1980)
Orthogonal fractional factorial plans. Calcutta Stat Ass. Bull., 29, p. 143-160.

(5) Cheng C.S. (1980)
Orthogonal arrays with variable number of symbols.
Annals. Statist., 8, p. 447-453.

(6) Mukerjee R. (1982)
Universal optimality of fractional factorial plans derivable trough orthogonal arrays. Calcutta Stat. Ass. Bull., 31, p. 63-68

(7) Stone M. (1959)
Application of a measure of information to the design and comparison of regression experiments. Ann. Math. Statist. 30, p. 55-70.

(8) Pazman A. (1980)
Some features of the optimal design theory - A survey. Math. Operations forsch. Statist., Ser Statistics, II, p. 415-446.