

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2001-26

**Conditions for Optimality in
Experimental Designs**

P. DRUILHET¹

Les documents de travail ne reflètent pas la position de l'INSEE et n'engagent que leurs auteurs.

Working papers do not reflect the position of INSEE but only the views of the authors.

¹ CREST-ENSAI, Ecole Nationale de la Statistique et de l'Analyse de l'Information, Campus de Ker Lann, 35170 BRUZ.

Conditions for optimality in experimental designs

Pierre Druilhet

CREST-ENSAI, Ecole Nationale de la Statistique et de l'Analyse
de l'Information, Campus de Ker Lann, 35 170 BRUZ.

July 16, 2001

Abstract

The purpose of this paper is to establish some results on optimal criteria in experimental designs. Some relationships between optimality criteria are shown. In particular, we extend some results on the Φ_p criteria. We prove the Yeh [18] conjecture that gives a necessary and sufficient condition for a design to be universally optimal. We also give a similar result based on the eigenvalue of the information matrix.

keywords: Information matrices, Schur-convexity, majorization, universal optimality.

Résumé

L'objectif de cet article est d'établir certains résultats concernant les critères d'optimalité utilisés dans la théorie des plans d'expériences. En particulier, nous généralisons des résultats concernant les critères Φ_p largement utilisés. Nous démontrons la conjecture de Yeh [18] qui donne une condition nécessaire et suffisante d'optimalité universelle. Nous donnons également un résultat similaire basé sur les valeurs propres de la matrice d'information.

1 Introduction

The aim of the optimal design theory is to choose between a set \mathcal{D} of designs one of those that gives the “best” estimator of the parameters of interest. The optimal design depends on which sense we put in “the best”. Assume that the statistical analysis model is a linear model:

$$Y = A_d\alpha + B_d\beta + \varepsilon,$$

where α is the t-vector of interest parameters, β is the vector of nuisance parameters, A_d and B_d the design matrices and ε the vector of zero-mean constant-variance uncorrelated errors. The quality of the parameter of interest is directly related to its variance matrix V_d or equivalently to its information matrix C_d defined by

$$C_d = A_d' (I - pr_{(B_d)})A_d,$$

where $pr_{(M)} = M(M'M)^-M'$ is the matrix of the projector onto $Im(M)$. From now on, we assume that all the information matrices considered satisfy:

$$C_d\mathbb{1}_t = 0 \quad \text{and} \quad \text{Rank}(C_d) = t - 1,$$

where $\mathbb{1}_t$ is the t-vector of ones. These conditions occurs very frequently in analysis of variance models. Thus, comparing estimators is equivalent to put a preorder on the set $\mathcal{C} = \{C_d; d \in \mathcal{D}\}$ of information matrices.

A natural order on \mathcal{C} is the Loewner order that lead to the following notion of optimality:

Definition 1 *A design d^* is said to be uniformly optimal among a class \mathcal{D} of designs if for all designs d in \mathcal{D} , $C_{d^*} - C_d$ is nonnegative or equivalently if for all designs d in \mathcal{D} and all contrasts $c'\alpha$:*

$$\text{var}(c'\hat{\alpha}_{d^*}) \leq \text{var}(c'\hat{\alpha}_d),$$

where $\hat{\alpha}_d$ is an ordinary least-square estimator of α for the design d .

Strategies to obtain uniformly optimal designs can be found in Kunert [11].

Unfortunately the Loewner order is a partial order and most often uniformly optimal designs do not exist. An other way to define a preorder on \mathcal{C} is to choose a statistically meaningful criterion Φ , that is a function from \mathcal{C} to a totally ordered set, most often \mathbb{R} :

Definition 2 A design d^* is said to be Φ -optimal if:

$$\forall d \in \mathcal{D}, \quad \Phi(C_{d^*}) \leq \Phi(C_d).$$

The purpose of this paper is to present some results concerning criteria used in optimal design theory. In section 2, we recall some results on Schur convexity that is the main tool used in this paper, then we establish relationships between Φ_p criteria, with application to A-, D-, and E-optimality. In section 3, we present some results concerning universal optimality, especially, we establish a necessary and sufficient condition for a design to be universally optimal (Yeh's [18] conjecture). Then we present another necessary and sufficient conditions based on the eigenvalues of the information matrices for a modified definition of the universal optimality.

2 Optimality criteria

There exist a lot of criteria for optimal design in the literature, we refer to Hedayat [8] for a review of most of the optimality criteria and to Shah and Sinha [17] for an extended discussion on the relationships between criteria. In this section, we aim to extend some results concerning these relationships. First, we present the main tools used throughout this paper: majorization and Schur-convex functions.

2.1 Majorization and Schur convexity

The Schur convexity is an important concept, useful to derive some relationships between criteria. The best general reference on majorization and Schur convexity is Marshall & Olkin [13].

Definition 3 For x and y in \mathbb{R}^t , we denote by $x_{\downarrow i}$ the i th greatest component of x . Then x is majorized by y , denoted by $x \prec y$, if

$$\sum_{i=1}^t x_i = \sum_{i=1}^t y_i \quad \text{and} \quad \forall k = 1, \dots, t-1, \quad \sum_{i=1}^k x_{\downarrow i} \leq \sum_{i=1}^k y_{\downarrow i}$$

We also denote :

$$x \prec^w y \quad \text{if} \quad \forall k = 1, \dots, t, \quad \sum_{i=k}^t x_{\downarrow i} \geq \sum_{i=k}^t y_{\downarrow i}$$

and $x \prec_w y \quad \text{if} \quad \forall k = 1, \dots, t, \quad \sum_{i=1}^k x_{\downarrow i} \leq \sum_{i=1}^k y_{\downarrow i}$

Notation 4 We denote by P_σ the (t,t) -matrix permuting the component of a vector according to the permutation σ lying in S_t , where S_t is the symmetric group on $\{1, \dots, t\}$.

Definition 5 A real function ϕ on \mathbb{R}^t is Schur-convex if

$$x \prec y \implies \phi(x) \leq \phi(y).$$

and Schur-concave if

$$x \prec y \implies \phi(x) \geq \phi(y).$$

This definition looks more like a nondecreasing (resp. nonincreasing) condition than a convexity condition. The term ‘‘Schur-convex’’ is historical and corollary 1 establish the link between convexity and Schur-convexity.

Definition 6

A function ϕ on \mathbb{R}^t is symmetric if $\forall x \in \mathbb{R}^t$ and $\forall \sigma \in S_t$, $\phi(P_\sigma x) = \phi(x)$.

The following proposition give a characterization of majorization in term of permutation matrices. It is a combination of two theorems, one by Birkhoff [1] and the other one by Hardy, Littlewood and Pólya [7] (see Theorem 2.B.2 and 2.A.2 in Marshall & Olkin [13]).

Proposition 1 For x and y in \mathbb{R}^t , $x \prec y$ if and only if there exist nonnegative reals α_σ such that:

$$x = \sum_{\sigma \in S_t} \alpha_\sigma P_\sigma y \quad \text{with} \quad \sum_{\sigma \in S_t} \alpha_\sigma = 1.$$

Corollary 1 A convex symmetric function ϕ on \mathbb{R}^t is Schur-convex.

Note that a Schur-convex function is not necessarily convex (see Hedayat [8]). The following corollary shows how $x \prec y$ implies that the component of x are closer together around the mean than the component of y in a statistical meaning.

Corollary 2 Denote by \bar{x} (resp. \bar{y}) the arithmetic mean of the component of x (resp. of y). If $x \prec y$, then

$$\bar{x} = \bar{y} \quad \text{and} \quad \forall p \geq 1, \quad \left(\frac{1}{t} \sum_{i=1}^t |x_i - \bar{x}|^p \right)^{1/p} \leq \left(\frac{1}{t} \sum_{i=1}^t |y_i - \bar{y}|^p \right)^{1/p}.$$

Note that the converse part does not necessarily hold.

Notation 7 We denote by $\lambda(C)$ the t -vector of the eigenvalues of C in decreasing order.

Lemma 2 (Fan [6]) Let A and B be two (n, n) symmetric matrices, then

$$\lambda(A + B) \prec \lambda(A) + \lambda(B).$$

Definition 8 A criterion $C \mapsto \Phi(C)$ is Schur-convex on the eigenvalues of C if

$$\lambda(C) \prec \lambda(D) \implies \Phi(C) \leq \Phi(D).$$

Lemma 3 (Bondar [2]) If a criterion Φ is convex and satisfies $\Phi(OCO') = \Phi(C)$, then Φ is Schur-convex on the eigenvalues of C .

Lemma 4 If a criterion $C \mapsto \Phi(C)$ is Schur-convex on the eigenvalues of C , then there exists a Schur-convex function ϕ on \mathbb{R}^t such that:

$$\Phi(C) = \phi(\lambda(C)).$$

Proof : If $\Phi(C)$ is Schur-convex on the eigenvalues of C , then $\Phi(C)$ depends only on the eigenvalues of C . So $\Phi(C) = \phi(\lambda(C))$ and then $\lambda(C_1) \prec \lambda(C_2)$ implies $\phi(\lambda(C_1)) \leq \phi(\lambda(C_2))$. \square

2.2 Optimality and diagonal terms

The following result can be useful to find Φ -optimal design, when Φ is Schur convex on the eigenvalues of C and the diagonal terms are easy to calculate.

Proposition 5 Let Φ be a criterion Schur convex on the eigenvalues. By lemma 4, $\Phi(C) = \phi(\lambda(C))$. Then d^* is Φ -optimal among a class \mathcal{D} of designs if

$$\forall d \in \mathcal{D} \quad \Phi(C_{d^*}) \leq \phi(\delta(C_d))$$

where $\delta(C_d)$ is the vector of diagonal terms of C_d in decreasing order.

Proof : The proof is a direct consequence of the following lemma. \square

Lemma 6 (Schur [16]) For all symmetric real matrices,

$$\delta(C) \prec \lambda(C).$$

2.3 The Φ_p criteria

In this subsection, we present the well known Φ_p criteria introduced first by Kiefer [9]. We present some relationships among these criteria, strengthening some existing results. First we define the exponent of a non-full-rank symmetric matrix.

Notation 9 *Let M be a non-negative symmetric matrix. We denote by M^+ its Moore-Penrose inverse. Moreover, for $p > 0$, M^p is the usual matrix exponent. When $p < 0$, M^p is defined by $M^p = (M^+)^{-p}$. By continuity, $M^0 = \text{pr}_{(M)} = MM^+$.*

Definition 10 *The Φ_p criteria are defined as follow :*

$$\begin{aligned}\Phi_p(C) &= \left(\frac{1}{t-1} \sum_{i=1}^{t-1} \lambda_i^{-p}(C) \right)^{1/p} = \left(\frac{1}{t-1} \text{tr}(C^{-p}) \right)^{1/p}, \text{ for } p \in \mathbb{R} \setminus \{0\} \\ \Phi_0(C) &= \lim_{p \rightarrow 0} \Phi_p(C) = \prod_{i=1}^{t-1} \lambda_i^{-1/(t-1)}, \\ \Phi_{+\infty}(C) &= \lim_{p \rightarrow +\infty} \Phi_p(C) = \lambda_{t-1}^{-1}(C) = \max_{i=1, \dots, t-1} \lambda_i^{-1}(C), \\ \Phi_{-\infty}(C) &= \lim_{p \rightarrow -\infty} \Phi_p(C) = \lambda_1^{-1}(C) = \min_{i=1, \dots, t-1} \lambda_i^{-1}(C)\end{aligned}$$

Remark:

The Φ_0 -, Φ_1 - and Φ_∞ -optimality are equivalent respectively to the very popular D -, A - and E - optimality. The criterion $\Phi_{-1}(C) = 1/\text{tr}(C)$ will play an important role in the next section.

The following proposition is a catalog of well known results on the Φ_p criteria.

Proposition 7

- For all $p \in \bar{\mathbb{R}}$, $C \mapsto \Phi_p(C)$ is invariant by row-column permutations of C .
- For all $p \in \bar{\mathbb{R}}$, $C_{d_1} \leq C_{d_2} \implies \Phi_p(C_{d_1}) \geq \Phi_p(C_{d_2})$,
- $C \mapsto \Phi_p(C)$ is concave for $p < -1$ and convexe for $p > -1$,
- $p \mapsto \Phi_p(C)$ is nondecreasing in p .

The two following propositions generalize property 2.5 by Kiefer [10]: they show that the Φ_p criteria can be considered, in some cases, as a kind of “scale” of optimality.

Proposition 8 Let d_1 and d_2 be two designs with rank $t - 1$ information matrices. Then for $p_o \neq 0$

$$\lambda(C_{d_1}^{-p_o}) \prec \lambda(C_{d_2}^{-p_o}) \implies \begin{cases} \Phi_{p_o}(C_{d_1}) = \Phi_{p_o}(C_{d_2}) \\ \Phi_p(C_{d_1}) \leq \Phi_p(C_{d_2}) & \text{for } p > p_o \\ \Phi_p(C_{d_1}) \geq \Phi_p(C_{d_2}) & \text{for } p < p_o \end{cases}$$

Proof : For $p = 1, \dots, t - 1$,

$$\lambda_i(C^{-p}) = \lambda_i(C)^{-p} = (\lambda_i(C)^{-p_o})^{p/p_o}.$$

Then, they are several cases to be considered: $0 < p < p_o$, $0 < p_o < p$, $p < 0 < p_o$, $p_o < 0 < p$, $p < p_o < 0$ and $p_o < p < 0$. In all the cases, the scheme of the proof is the following:

a) By proposition 3.C.1.b in Marshall & Olkin [13], $x \in (\mathbb{R}_+^*)^{t-1} \mapsto \sum_{i=1}^{t-1} x_i^{p/p_o}$ is Schur-concave for $p/p_o \in]0, 1[$ and Schur-convexe otherwise. So the condition $\lambda(C_{d_1}^{-p_o}) \prec \lambda(C_{d_2}^{-p_o})$ lead to the comparison between $\sum_{i=1}^{t-1} \lambda_i(C_{d_1}^{-p})$ and $\sum_{i=1}^{t-1} \lambda_i(C_{d_2}^{-p})$.

b) Then we use the fact that $x \in (\mathbb{R}_+^*) \mapsto x^{1/p}$ is increasing for $p > 0$ and decreasing for $p < 0$ to obtain the result.

For example, when $0 < p_o < p$, $x \mapsto x^{p/p_o}$ is convex and $x \mapsto x^{1/p}$ is decreasing, so:

$$\begin{aligned} \lambda(C_{d_1}^{-p_o}) \prec \lambda(C_{d_2}^{-p_o}) &\implies \sum_{i=1}^{t-1} \lambda_i(C_{d_1}^{-p}) \leq \sum_{i=1}^{t-1} \lambda_i(C_{d_2}^{-p}) \\ &\implies \left(\frac{1}{t-1} \sum_{i=1}^{t-1} \lambda_i(C_{d_1}^{-p}) \right)^{1/p} \leq \left(\frac{1}{t-1} \sum_{i=1}^{t-1} \lambda_i(C_{d_2}^{-p}) \right)^{1/p} \end{aligned}$$

The case $p = 0$ can be obtain as the limit case when $p \rightarrow 0$. The case $p = p_o$ is obvious. □

In next proposition, we weaken both the sufficient and the necessary.

Proposition 9 Let d_1 and d_2 be two designs with rank $t - 1$ information matrices.

Then for $p_o < 0$:

$$\lambda(C_{d_1}^{-p_o}) \prec^w \lambda(C_{d_2}^{-p_o}) \implies \Phi_p(C_{d_1}) \leq \Phi_p(C_{d_2}) \text{ for } p \geq p_o$$

and for $p > 0$

$$\lambda(C_{d_1}^{-p_o}) \prec_w \lambda(C_{d_2}^{-p_o}) \implies \Phi_p(C_{d_1}) \leq \Phi_p(C_{d_2}) \text{ for } p \geq p_o$$

Proof : There are three cases to be considered: $p > p_o > 0$, $p \geq 0 > p_o$ and $0 > p > p_o$. The scheme of the proof is identical to the proof of Proposition 8 except that we use Proposition 4.B.2 or Theorem 5.A.2 by Marshall& Olkin [13]. □

Many results can be derived from the two poroposition above : the first one will be generalized in the next section.

Corollary 3 *A design d^* is Φ_p optimal among a class \mathcal{D} of designs for $p \leq -1$ if $\forall d \in \mathcal{D}$, $\lambda(C_{d^*}) \prec_w \lambda(C_d)$*

Corollary 4 *A design d^* is E -optimal among a class \mathcal{D} of designs if it is A optimal and if for all $d \in \mathcal{D}$, $\lambda(C_{d^*}^+) \prec_w \lambda(C_d^+)$*

The next application of Proposition 8 is a result of Bondar [2] concerning MV-optimality introduced by Eccleston and Hedayat [5].

Definition 11 *A design d^* among a class \mathcal{D} of designs is MV-optimal if it minimizes Φ_{-1} and if it maximizes Φ_{-2} among the subclass of design minimizing ϕ_{-1} .*

Corollary 5 *A design d^* is MS-optimal among a class \mathcal{D} if it minimizes Φ_{-1} and if $\lambda(C_{d^*}) \prec \lambda(C_d)$ for all the designs minimizing ϕ_{-1} .*

3 Conditions for universal optimality

In some cases, a design with particular information matrix can be optimal not only for a criterion but for a whole class of criteria. Following this idea, Kiefer [10] introduce the notion of universal optimality. In this section, we present different definitions of universal optimality and give necessary and sufficient conditions for a design to be universally optimal.

3.1 Kiefer's universal optimality

The following definition of universal optimality is the historical one.

Definition 12 (Kiefer [10]) *A design d^* is universally optimal among a class \mathcal{D} of designs if d^* is Φ -optimal for all criteria $\Phi(C)$ from \mathcal{C} to $]-\infty, +\infty]$ satisfying:*

- (a) Φ is invariant under each permutation of rows and (the same on) columns,
- (b) $\Phi(\alpha C)$ is nonincreasing in the scalar $\alpha > 0$,
- (c) Φ is convex.

Remark:

Many usual criteria satisfy the three conditions of definition 12: the Φ_p – criteria for $p \geq -1$, the criteria of type 1 and 2 (Cheng [3]), MV criterion,.... Proposition 1 by Kiefer [10] gives a sufficient condition for a design d^* to be universally optimal:

Proposition 10 (Kiefer [10]) *A design d^* is universally optimal among a class \mathcal{D} of designs if its information matrix is completely symmetric (i.e. invariant by row-column permutation) and maximize the trace among \mathcal{D} .*

Yeh [18] gives a more general sufficient condition and conjecture that it is also a necessary condition:

Proposition 11 (Yeh [18]) *A design d^* is universally optimal among a class \mathcal{D} of designs if it satisfies:*

- (i) $\text{tr } C_{d^*} = \max_{d \in \mathcal{D}} \text{tr } C_d$,
- (ii) $\forall d \in \mathcal{D}$, there exist scalars $a_{d\sigma} \geq 0$ satisfying:

$$C_{d^*} = \sum_{\sigma \in S_t} a_{d\sigma} P_\sigma C_d P_\sigma'.$$

Proposition 12 (Yeh's conjecture [18]) *The converse part of Proposition 11 hold.*

Proof of Prop 12.

Condition (i) is necessary because $C \mapsto -\text{tr } C$ satisfies condition (a), (b) and (c) in Definition 12. Let d^* be a universally optimal design, and assume that there exists a design d_1 for which there do not exist $\alpha_{d_1\sigma} \geq 0$ such that

$$C_{d^*} = \sum_{\sigma \in S_t} \alpha_{d_1\sigma} P_\sigma C_{d_1} P_\sigma'.$$

Let \mathcal{A} be the convex cone generated by the matrices $\{P_\sigma C_{d_1} P'_\sigma\}_{\sigma \in S_t}$, then we have (see e.g. Rockafellar [15, pp.14]):

$$\mathcal{A} = \left\{ M \mid M = \sum_{\sigma \in S_t} \alpha_{d\sigma} P_\sigma C_{d_1} P'_\sigma \text{ for some } \alpha_{d\sigma} \in \mathbb{R}^+ \right\}$$

and consider the criterion Φ defined by:

$$\Phi(C_d) = \begin{cases} 0 & \text{if } C_d \in \mathcal{A}, \\ +\infty & \text{if } C_d \notin \mathcal{A}. \end{cases}$$

For all $\sigma \in S_t$, $P_\sigma \mathcal{A} P'_\sigma = \mathcal{A}$, thus $\Phi(P_\sigma C_d P'_\sigma) = \Phi(C_d)$. The convexity of \mathcal{A} implies the convexity of Φ . Moreover, for any $\alpha > 0$, $\Phi(\alpha C_d) = \Phi(C_d)$. Hence Φ satisfies conditions (a), (b) and (c) in Definition 12. By construction of Φ , we have $\Phi(C_{d_1}) < \Phi(C_{d^*})$ that contradicts the fact that d^* is universally optimal. □

Remarks:

- One can object that the criterion Φ exhibited in the proof takes only two values: 0 and $+\infty$. However, we can replace Φ by

$$\Phi_1(C_d) = \inf_{C \in \mathcal{A}} \left\| \frac{C_d}{\text{tr } C_d} - \frac{C}{\text{tr } C} \right\|$$

where $\|\cdot\|$ is any norm on the set of the symmetric matrices invariant by row-column permutation, e.g. the Euclidean norm $\|C\| = \sqrt{\text{tr } C^2}$. The criterion $\Phi_1(C_d)$ is the distance between $\frac{C_d}{\text{tr } C_d}$ and the convex hull of $\{P_\sigma \frac{C_{d_1}}{\text{tr } C_{d_1}} P'_\sigma\}_{\sigma \in S_t}$. It can be checked that Φ_1 satisfies Conditions (a), (b) and (c) in Definition 12 and that $0 = \Phi_1(C_{d_1}) < \Phi_1(C_{d^*})$.

- A second objection is that $\forall \alpha > 0$, $\Phi_1(\alpha C_d) = \Phi_1(C_d)$. But we can define

$$\Phi_2(C) = \Phi_1(C) - \varepsilon \text{tr } C$$

with $0 < \varepsilon < (\Phi_1(C_{d^*}) - \Phi_1(C_{d_1})) / \text{tr}(C_{d^*} - C_{d_1}) \leq +\infty$. The new criterion Φ_2 satisfies condition (a), (b) and (c) in Definition 12 and the function $\alpha \mapsto \Phi_2(\alpha C_d)$ is now (strictly) decreasing. Moreover, $\Phi_2(C_{d_1}) < \Phi_2(C_{d^*})$.

3.2 Restricted universal optimality

We seek now a necessary and sufficient condition for a design to be universally optimal that only depends on the eigenvalues of the information matrix. So we are lead to restrict Kiefer's definition to criteria $\Phi(C_d)$ that depends only on the eigenvalues of C_d , thus we replace condition (a) in Definition 12 with:

(a') $\Phi(OC_dO') = \Phi(C_d)$ for any orthogonal matrix O .

Proposition 13 *A design d^* is universally optimal (with condition (a')) among a class \mathcal{D} of designs if and only if*

$$(i) \quad \text{tr } C_{d^*} = \max_{d \in \mathcal{D}} \text{tr } C_d,$$

$$(ii) \quad \lambda \left(\frac{C_{d^*}}{\text{tr } C_{d^*}} \right) \prec \lambda \left(\frac{C_d}{\text{tr } C_d} \right).$$

Proof :

Assume that conditions (i) and (ii) hold, then $\forall d \in \mathcal{D}$: condition (ii) implies that $\lambda(C_{d^*}) \prec \lambda \left(\frac{\text{tr } C_{d^*}}{\text{tr } C_d} C_d \right)$ and then by Lemme 3, $\Phi(C_{d^*}) \leq \Phi \left(\frac{\text{tr } C_{d^*}}{\text{tr } C_d} C_d \right)$. By condition (i) and (b), $\Phi(C_{d^*}) \leq \Phi(C_d)$. Thus d^* is universally optimal (with condition a').

Conversely, let d^* be universally optimal with condition (a'), then condition (i) hold. Assume that condition (ii) does not hold, then there exists a design d_1 such that

$$\lambda \left(\frac{C_{d^*}}{\text{tr } C_{d^*}} \right) \not\prec \lambda \left(\frac{C_{d_1}}{\text{tr } C_{d_1}} \right).$$

We can define a set of (t,t) nonnegative symmetric matrices:

$$\mathcal{A} = \left\{ M/M\mathbf{1}_t = 0 \text{ and } \lambda \left(\frac{M}{\text{tr } M} \right) \prec \lambda \left(\frac{C_{d_1}}{\text{tr } C_{d_1}} \right) \right\}.$$

The set \mathcal{A} is a cone and by lemma 2 it is also convex. Then, the end of the proof is identical to the proof of Proposition 12. □

Remark:

This proposition shows that the ellipsoid $x \mapsto \frac{1}{\|x\|^2} x' C_d x$ must be "as spherical as possible". The sphericity comparison is made by the majorization of the eigenvalues of C_d , that are equal to the length of the half axes of the ellipsoid, using the scale parameter $\text{tr}(C_d)$.

3.3 Schur optimality

For completeness, we mention a concept close to universal optimality: Schur optimality introduced by Magda [12] and called universal optimality by Bondar [2].

Definition 13 (Magda [12]) *A design d^* is Schur optimal among a class \mathcal{D} of designs if d^* is Φ -optimal for all criteria $\Phi(C)$ from \mathcal{C} to $]-\infty, +\infty]$ satisfying:*

$$(\alpha) : \lambda(C_1) \prec^w \lambda(C_2) \implies \Phi(C_1) \leq \Phi(C_2).$$

Remark:

By Theorem 3.A.8 in Marsall & Olkin [13] or Theorem 2.1 in Bondar [2], condition (α) is equivalent to the two conditions given historically by Magda:

(β) $\Phi(C)$ is schur-convex on the eigenvalues of C ,

$$(\gamma) \forall i, \lambda_i(C_1) \leq \lambda_i(C_2) \implies \Phi(C_1) \geq \Phi(C_2).$$

References

.

- [1] Birkhoff, G. 1946. Tres observaciones sobre el algebra lineal. Univ. Nac. Tucumán Rev. Ser. A 5:147-151.
- [2] Bondar, J. V. 1983. Universal optimality of experimental designs: definition and a criterion. The Canadian Journal of Statistics 11:325-331.
- [3] Cheng, C.-S. 1978. Optimality of certain asymmetrical experimental designs. Ann. Statist. 6:1239-1361.
- [4] Cheng, C.-S. and Wu, C. F. 1980. Balanced repeated measurements designs. Ann. Statist. 8:1272-1283.
- [5] Eccleston, J. A. and Hedayat, A. 1974. On the theory of connected designs: characterization and optimality. Ann. Statist. 2:1238-1255.
- [6] Fan, K. 1949. On a theorem of Weyl concerning eigenvalues of linear transformations. Proc. Nat. Acad. Sci. USA 37:31-35.
- [7] Hardy, G. H., Littlewood, J. E. and Pólya, G. 1929. Some inequalities satisfied by convex fonctions. Messenger Math. 58:145-152.

- [8] Hedayat, A. 1981. Study of optimality criteria in design of experiment. *Ann. Statist.* 6:619-628.
- [9] Kiefer, J. 1974. General equivalence theory for optimum designs (approximate theory). *Ann. Math. Statist.* 2:849-879.
- [10] Kiefer, J. 1975. Construction and optimality of generalized Youden designs. In: *A Survey of Statistical Design and Linear Models*. (J. N. Srivastava, Ed.), 333-353. North Holland: Amsterdam.
- [11] Kunert, J. 1983. Optimal design and refinement of the linear model with applications to repeated measurements designs. *Ann. Statist.* 11:247-257.
- [12] Magda, C. G. 1980. Circular balanced repeated measurements designs. *Commun. Statist.- Theor. Meth.* A9:1901-1918.
- [13] Marshall, A. W. and Olkin, I. 1979. *Inequalities: Theory of Majorization and Its Applications*. Academic Press.
- [14] Pukelsheim, F. 1993. *Optimal Design of Experiments*. Wiley, New York.
- [15] Rockafellar 1972. *Convex Analysis*. Princeton University Press, New Jersey.
- [16] Schur, I. 1973. *Issai Schur collected Works*. vol. II . pp. 146-427. Springer-verlag, New-York.
- [17] Shah, K. R. and Sinha, B. K. 1989. *Theory of Optimal Designs*. Springer-verlag, New-York.
- [18] Yeh, C.-M. 1986. Conditions for optimality of block designs. *Biometrika* 73:701-706.