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, France.

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 results on the Φ_p criteria. We prove the Yeh [19] conjecture that gives a necessary and sufficient condition for a design to be universally optimal. We also give a similar result based on the eigenvalues of the information matrix.

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

---

1 Introduction

The aim of optimal design theory is to choose from a set $\mathcal{D}$ of designs one that gives the "best" estimator of the parameters of interest. The optimal design depends on how "the best" is defined. For example, consider the 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 are 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

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

where $pr_{(M)} = M(M'M)^{-}M'$ is the projector onto Range(M) and I is the identity matrix. 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 occur frequently in analysis of variance. Comparing estimators is equivalent to defining a preordering on the set $\mathcal{C} = \{C_d; d \subset \mathcal{D}\}$ of information matrices. A natural ordering on $\mathcal{C}$ is the Loewner ordering that leads 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 any design d in $\mathcal{D}$, $C_{d^*} - C_d$ is non-negative or, equivalently, if for any design d in $\mathcal{D}$ and any contrast $c'\alpha$:*

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

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 [12].

Unfortunately the Loewner ordering is a partial ordering and most often uniformly optimal designs do not exist. Another way to define a preordering on $\mathcal{C}$ is to choose a statistically meaningful criterion $\Phi : \mathcal{C} \mapsto \mathcal{B}$, where $\mathcal{B}$ is a totally ordered set, such as $[0, \infty]$.

Definition 2 *A design d^* is said to be $\Phi-$optimal if:*

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

The purpose of this paper is to present some new results concerning criteria used in optimal design theory. In Section 2, we first recall some results on Schur convexity which is the main tool used in this paper. Then we establish some relationships between Φ_p criteria, with application to A-, D-, and E-optimality. In Section 3, we present some results concerning universal optimality and we establish a necessary and sufficient condition for a design to be universally optimal (Yeh's [19] conjecture). We also present another necessary and sufficient condition for an alternative definition of universal optimality

2 Optimality criteria

Many optimal design criteria are available in the literature; we refer to Hedayat [8] for a review of optimality criteria and to Shah and Sinha [18] for an extended discussion on the relationships between these 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

Schur convexity is an important concept, useful in deriving relationships between criteria. The best general reference on majorization and Schur convexity is Marshall & Olkin [14]. We recall here some basic definitions and properties.

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

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

We also denote :

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

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

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

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

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

and Schur-concave if

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

This definition looks more like a non-decreasing (resp. non-increasing) condition than a convexity condition. The term "Schur-convex" is historical and Corollary 8 establishes 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 gives a characterization of majorization in term of permutation matrices. It combines 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 [14]).

Proposition 7 *For x and y in $\mathbb{R}^t$, $x \prec y$ if and only if there exist non-negative reals α_σ such that:*

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

Corollary 8 *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 components of x are closer together around their mean than the components of y.

Corollary 9 *Denote by $\bar{x}$ (resp. $\bar{y}$) the arithmetic mean of x (resp. y). If $x \prec y$, then*

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

Note that the converse does not necessarily hold.

Notation 10 *We denote by $\lambda(C)$ the t-vector of the decreasingly ordered eigenvalues of C.*

Lemma 11 (Fan [6]) *Let A and B be two (n,n) symmetric matrices, then*

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

Definition 12 *A criterion $C \mapsto \Phi(C)$ is Schur-convex on the eigenvalues if*

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

Lemma 13 (Bondar [2]) *If a criterion Φ is convex and satisfies $\Phi(OCO') = \Phi(C)$ for any orthogonal matrix O, then Φ is Schur-convex on the eigenvalues.*

Lemma 14 *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:*

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

2.2 Optimality and diagonal terms

The following result is useful in finding $\Phi-$optimal designs when Φ is Schur convex on the eigenvalues of C and the diagonal terms are easy to calculate.

Proposition 15 *Let Φ be a Schur convex criterion on the eigenvalues. So that $\Phi(C) = \phi(\lambda(C))$. Then d^* is $\Phi-$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 16 (Schur [17]) *For all symmetric real matrices,*

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

where $\delta(C)$ is the vector of diagonal terms of C.

2.3 The Φ_p criteria

In this section, we present the well-known Φ_p criteria introduced first by Kiefer [10]. We establish some relationships between these criteria, and strengthen some existing results. First, we define the exponent of a non-full-rank symmetric matrix.

Notation 17 *Let M be a non-negative symmetric matrix. We denote by M^+ its Moore-Penrose inverse. For $p > 0$, M^p is the usual matrix exponent. When $p < 0$, M^p is defined by $M^p = (M^+)^{-p}$. By continuity, $M^0 = \mathrm{pr}_{(M)} = MM^+$.*

Definition 18 *The Φ_p criteria are defined as follow :*

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

Remark:

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

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

Proposition 19

- *For all* $p \in [-\infty, +\infty]$, $C \mapsto \Phi_p(C)$ *is invariant by row-column permutations of* C.

- *For all* $p \in [-\infty, +\infty]$, $C_{d_1} \leq C_{d_2} \Longrightarrow \Phi_p(C_{d_1}) \geq \Phi_p(C_{d_2})$,

- $C \mapsto \Phi_p(C)$ *is convex for* $p > -1$,

- $p \mapsto \Phi_p(C)$ *is non-decreasing in* p.

The following two propositions generalize property 2.5 of Kiefer [11]: they show that the Φ_p criteria can be considered, in some cases, as a kind of "scale" of optimality.

Proposition 20 *Let* d_1 *and* d_2 *be two designs with rank* $t-1$ *information matrices. Then for* $p_o \neq 0$

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

Proof : For $i = 1, ..., t-1$,

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

There 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 of Marshall & Olkin [14], $x \in (]0,\infty[)^{t-1} \mapsto \sum_{i=1}^{t-1} x_i^{p/p_o}$ is Schur-concave for $p/p_o \in]0,1[$ and Schur-convex otherwise. So the condition $\lambda(C_{d_1}^{-p_o}) \prec \lambda(C_{d_2}^{-p_o})$ leads 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) The result follows on using the fact that $x \in (]0,+\infty[) \mapsto x^{1/p}$ is increasing for $p > 0$ and decreasing for $p < 0$.

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

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

The case $p = 0$ can be obtained as the limit case when $p \to 0$. The case $p = p_o$ is obvious.

$\square$

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

Proposition 21 *Let d_1 and d_2 be two designs with rank $t-1$ information matrices.*

Then for $p_o < 0$:

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

and for $p_o > 0$

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

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 20 except that we use Proposition 4.B.2 or Theorem 5.A.2 of Marshall& Olkin [14].

$\square$

Many results can be derived from the above two propositions:

Corollary 22 *A design d^* is Φ_p-optimal among a class $\mathcal{D}$ of designs for $p \geq -1$ if $\forall\, d \in \mathcal{D}$, $\lambda(C_{d^*}) \prec^w \lambda(C_d)$*

This result will be generalized in the next section.

Corollary 23 *A design d^* is $E-$optimal among a class $\mathcal{D}$ of designs if for all $d \in \mathcal{D}$, $\lambda(C_{d^*}^+) \prec_w \lambda(C_d^+)$.*

Note that the condition "$\forall\, d \in \mathcal{D}$,$\lambda(C_{d^*}^+) \prec_w \lambda(C_d^+)$" is slightly stronger than "d is $A-$optimal".

The next application of Proposition 20 is a result by Bondar [2] concerning MS-optimality introduced by Eccleston and Hedayat [5].

Definition 24 *A design d^* of a class $\mathcal{D}$ of designs is MS-optimal if it minimizes Φ_{-1} and if it maximizes Φ_{-2} among the subclass of designs minimizing ϕ_{-1}.*

Corollary 25 *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 is optimal not only for just a single specific criterion but for a whole class of criteria. Following this idea, Kiefer [11] introduces 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 26 (Kiefer [11]) *A design d^* is universally optimal among a class $\mathcal{D}$ of designs if d^* is Φ-optimal for all the 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 non-increasing in the scalar $\alpha > 0$,

(c) Φ is convex.

Remark:

Many of the usual criteria satisfy the three conditions of Definition 26. These include $\Phi_p - criteria$ for $p \geq -1$, the criteria of type 1 and 2 (Cheng [3]) and the MS criterion.

Proposition 1 of Kiefer [11] gives a sufficient condition for a design d^* to be universally optimal.

Proposition 27 (Kiefer [11]) *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 maximizes the trace among $\mathcal{D}$.*

Yeh [19] establishes a more general sufficient condition and gives some applications. He conjectures that the condition is also necessary.

Proposition 28 (Yeh [19]) *A design d^* is universally optimal among a class $\mathcal{D}$ of designs if it satisfies:*

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

(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 29 (Yeh's conjecture [19]) *The sufficient condition in Proposition 28 is also a necessary condition.*

Proof of Proposition 29.

Condition (i) is necessary because $C \mapsto -\operatorname{tr} C$ satisfies condition (a), (b) and (c) in Definition 26 . Let d^* be a universally optimal design, and assume that there exists a design d_1 for which there are no $\alpha_{d_1\sigma} \geq 0$ 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 $\{P_\sigma \ C_{d_1} \ P'_\sigma\}_{\sigma \in S_t}$. We have (see e.g. Rockafellar [16, p.14]):

$$\mathcal{A} = \left\{ M \ / \ M = \sum_{\sigma \in S_t} \ \alpha_{d\sigma} P_\sigma \ C_{d_1} \ P'_\sigma \quad \text{for some} \quad \alpha_{d\sigma} \geq 0 \right\}.$$

Consider the criterion Φ defined by:

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

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 26. By construction of Φ, we have $\Phi(C_{d_1}) < \Phi(C_{d^*})$ that contradicts the fact that d^* is universally optimal.

$\square$

Remark:

It may be objectionable that the criterion Φ exhibited in the proof takes only two values: 0 and $+\infty$ and that $\forall \, \alpha > 0, \ \Phi(\alpha C_d) = \Phi(C_d)$. However, Φ can be replaced by

$$\Phi_1(C_d) = \inf_{C \in \mathcal{A}} \|C_d - C\|$$

where $\| \cdot \|$ is any norm on the set of symmetric matrices invariant by row-column permutation, e.g. the Euclidean norm $\|C\| = \sqrt{\operatorname{tr} C^2}$. The criterion

$\Phi_1(C_d)$ is, in fact, the distance between C_d and $\mathcal{A}$. We now check that Φ_1 satisfies conditions (a), (b) and (c) of Definition 26. Because $P_\sigma \mathcal{A} P'_\sigma = \mathcal{A}$ and because the norm is invariant by permutation, we have:

$$\Phi_1(P_\sigma C_d P'_\sigma) = \inf_{C \in \mathcal{A}} \|P_\sigma C_d P'_\sigma - C\| = \inf_{C \in \mathcal{A}} \|P_\sigma C_d P'_\sigma - P_\sigma C P'_\sigma\| = \Phi_1(C_d).$$

For $\alpha > 0$, $\alpha\,\mathcal{A} = \mathcal{A}$, thus

$$\Phi_1(\alpha\, C) = \inf_{C \in \mathcal{A}} \|\alpha\, C_d - C\| = \inf_{C \in \mathcal{A}} \|\alpha\, C_d - \alpha\, C\| = \alpha\, \Phi_1(C_d)$$

So $\Phi_1(\alpha\, C_d)$ is (strictly) increasing in $\alpha > 0$.

Moreover, $C_d \mapsto \Phi(C_d)$ is convex (see Hiriart-Urruty and Lemaréchal [9, p.153]). It remains to show that $\Phi_1(C_{d_1}) < \Phi_1(C_{d^*})$. Since the convex cone $\mathcal{A}$ is closed (see Hiriart-Urruty and Lemaréchal [9, p.102]), there exists $D \in \mathcal{A}$ such that $\Phi_1(C_{d^*}) = \|C_{d^*} - D\|$. Since $C_{d^*} \notin \mathcal{A}$, $\Phi_1(C_{d^*}) > 0$. So $0 = \Phi_1(C_{d_1}) < \Phi_1(C_{d^*})$ (QED).

3.2 Restricted universal optimality

We now seek a necessary and sufficient condition of universal optimality that only depends on the eigenvalues of the information matrix. So, we are led to restrict Kiefer's definition to criteria $\Phi(C_d)$ that depend only on the eigenvalues of C_d. Hence we replace condition (a) in Definition 26 with:

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

Proposition 30 *A design d^* is universally optimal (with condition a') among a class $\mathcal{D}$ of designs if and only if*

$$(i) \quad \operatorname{tr} C_{d^*} = \max_{d \in \mathcal{D}} \operatorname{tr} C_d,$$
$$(ii) \quad \lambda\left(\frac{C_{d^*}}{\operatorname{tr} C_{d^*}}\right) \prec \lambda\left(\frac{C_d}{\operatorname{tr} C_d}\right).$$

Proof :

Assume that conditions (i) and (ii) hold, then $\forall\, d \in \mathcal{D}$: condition (ii) implies that $\lambda(C_{d^*}) \prec \lambda\left(\frac{\operatorname{tr} C_{d^*}}{\operatorname{tr} C_d}\, C_d\right)$ and, by Lemma 13, $\Phi(C_{d^*}) \leq \Phi\left(\frac{\operatorname{tr} C_{d^*}}{\operatorname{tr} C_d} C_d\right)$. By condition (b) of Definition 26 and by condition (i), $\Phi(C_{d^*}) \leq \Phi(C_d)$. Thus d^* is universally optimal (with condition a').

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

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

We now define the following set of (t,t) non-negative symmetric matrices:

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

The set $\mathcal{A}$ is a cone and, by Lemma 11, it can be proved that $\mathcal{A}$ is convex as well. Then, the end of the proof is identical to the proof of Proposition 29.

$\square$

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 half length of the ellipsoid axes, using the scale parameter $\mathrm{tr}(C_d)$.

3.3 Schur optimality

For the sake of completeness, we mention a concept close to universal optimality: Schur optimality introduced by Magda [13] and called universal optimality by Bondar [2]. This definition show again the strong link between universal optimality and Schur convexity.

Definition 31 (Magda [13]) *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:*

$(\alpha):\ \lambda(C_1) \prec^{\mathrm{w}} \lambda(C_2) \Longrightarrow \Phi(C_1) \leq \Phi(C_2)$.

Remark:

By Theorem 3.A.8 in Marsall & Olkin [14] or Theorem 2.1 in Bondar [2], condition (α) is equivalent to the two conditions given historically by Magda[13]:

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

(γ) $\forall\, i, \lambda_i(C_1) \leq \lambda_i(C_2) \Longrightarrow \Phi(C_1) \geq \Phi(C_2)$.

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 functions. Messenger Math. 58:145-152.

[8] Hedayat, A. 1981. Study of optimality criteria in design of experiment. In: *Statistics and Related Topics*. (M. Csörgö, D.A. Dawson, J.N.K. Rao, A.K.Md.E Sahel, Eds.), 39-56. North-Holland Publishing Company.

[9] Hiriart-Urruty, J.-B. and Lemaréchal, C. 1996. *Convex Analysis and Minimization Algorithms*. vol. I . Springer-verlag, New-York.

[10] Kiefer, J. 1974. General equivalence theory for optimum designs (approximate theory). Ann. Math. Statist. 2:849-879.

[11] 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.

[12] Kunert, J. 1983. Optimal design and refinement of the linear model with applications to repeated measurements designs. Ann. Statist. 11:247-257.

[13] Magda, C. G. 1980. Circular balanced repeated measurements designs. Commun. Statist.– Theor. Meth. A9:1901-1918.

[14] Marshall, A. W. and Olkin, I. 1979. *Inequalities: Theory of Majorization and Its Applications*. Academic Press.

[15] Pukelsheim, F. 1993. *Optimal Design of Experiments*. Wiley, New York.

[16] Rockafellar 1972. *Convex Analysis*. Princeton University Press, New Jersey.

[17] Schur, I. 1973. *Issai Schur collected Works*. vol. II . pp. 146-427. Springer-verlag, New-York.

[18] Shah, K. R. and Sinha, B. K. 1989. *Theory of Optimal Designs*. Springer-verlag, New-York.

[19] Yeh, C.-M. 1986. Conditions for optimality of block designs. Biometrika 73:701-706.