ADMISSIBILITY AND OPTIMALITY OF EXPERIMENTAL DESIGNS

Norbert Gaffke and Friedrich Pukelsheim
Institut für Mathematik der Universität Augsburg

1. INTRODUCTION

In this paper we study the relation between admissibility and optimality of experimental designs. While it is standard decision theoretic reasoning that a statistical procedure which is uniquely optimal will necessarily be admissible, we here prove a converse to the effect that an admissible design is uniquely optimal with respect to the \( E \)-criterion and a specific choice of the parameter system of interest. The general equivalence theory may then be employed to obtain necessary conditions for admissibility.

As usual we choose the experimental conditions from a compact \( k \)-dimensional experimental domain \( \mathcal{X} \subset \mathbb{R}^k \). We assume that under experimental conditions \( x \in \mathcal{X} \) the real observation \( Y(x) \) follows a linear model

\[
Y(x) = x' \theta + \sigma e(x)
\]

with uncorrelated errors \( e(x) \) of unit variance. A design \( \xi \) is a probability distribution with finite support on the experimental domain \( \mathcal{X} \), determining allocation and proportion of the experimental conditions.

The performance of a design \( \xi \) is determined through its \( k \times k \) moment matrix

\[
M(\xi) = \int_{\mathcal{X}} xx'd\xi.
\]

Let \( \mathcal{M} \) be the feasible set of moment matrices, assumed to be a convex and compact subset of nonnegative definite \( k \times k \) matrices.

We shall study admissibility of a candidate matrix \( M \) in the set \( \mathcal{M} \). It is illuminating to first discuss the case when the full parameter \( \theta \) is of interest (Section 2). Before turning to the more general case of an \( s \)-dimensional parameter system \( K' \theta \) (Section 4) we derive some intermediate results on information matrices (Section 3).

2. ADMISSIBILITY FOR THE FULL PARAMETER SET

Suppose \( M \in \mathcal{M} \) is a moment matrix whose admissibility properties we wish to investigate. We call \( M \) admissible for \( \theta \) in \( \mathcal{M} \) when no moment matrix \( A \in \mathcal{M} \) satisfies \( A \geq M \) and \( A \neq M \), relative to the Löwner ordering \( \geq \). To avoid trivialities we assume \( M \neq 0 \).
We shall show that every admissible moment matrix is $E$-optimal, i.e., it maximizes the minimum eigenvalue of an appropriate information matrix. However, the parameter system for which $E$-optimality is obtained is related to the candidate matrix $M$ in an intrinsic manner: We choose the system $H^t \theta$ from a full rank decomposition

$$M = HH'$$

where with $r = \text{rank } M$ the $k \times r$ matrix $H$ has full column rank $r$. An $E$-optimal moment matrix for $H^t \theta$ in $M$ is one which maximizes $\lambda_{\text{min}}(C_H(A))$ over $A \in M \cap \mathcal{A}(H)$, where $\mathcal{A}(H)$ is the convex cone of all nonnegative definite $k \times k$ matrices whose range contains the range of $H$, and

$$C_H(A) = (H'A^+H)^{-1} \text{ for } A \in \mathcal{A}(H).$$

We need an auxiliary lemma before turning to admissibility.

**Lemma 1.** Let $A \in M$ be a competing moment matrix. If $A$ is $E$-optimal for $H^t \theta$ in $M$ then $A \geq M$.

**Proof.** By construction the range of $M$ contains (actually coincides with) the range of $H$, and we have

$$C_H(M) = (H'M^{-1}H)^{-1} = (H'(HH')^{-1}H)^{-1} = I_r.$$

Optimality of $A$ yields $1 = \lambda_{\text{min}}(C_H(M)) \leq \lambda_{\text{min}}(C_H(A))$. Therefore $I_r \leq C_H(A)$, and pre- and postmultiplication with $H$ and $H'$ gives

$$M = HH' \leq HCH(A)H' \leq A,$$

where the last inequality may be found for instance in Pukelsheim & Styan (1983, p. 147).

We are now in a position to establish the relation between admissibility and unique $E$-optimality as announced above.

**Theorem 1.** The moment matrix $M$ is admissible for $\theta$ in $M$ if and only if $M$ is uniquely $E$-optimal for $H^t \theta$ in $M$.

**Proof.** Suppose $M$ is admissible. From Theorem 2 in Pukelsheim (1980, p. 344) we know that there exists an $E$-optimal moment matrix $A$ for $H^t \theta$ in $M$. By Lemma 1 we have $A \geq M$, and admissibility of $M$ forces $A = M$. This establishes unique $E$-optimality of $M$.

Conversely suppose $M$ is uniquely $E$-optimal. Let $A$ be a competing moment matrix satisfying $A \geq M$. Due to monotonicity $A$ will also be $E$-optimal. But then uniqueness forces $A = M$, i.e., admissibility of $M$.

Lemma 1 and Theorem 1 are closely related to Corollary 8.4 of Pukelsheim (1980, p. 359). Next we turn to the classical Theorem 7.1 on admissibility of Karlin & Studden (1966, p. 808), investigating the existence of a nonnegative definite matrix $N \neq 0$ or a positive definite matrix $N$ satisfying the system of normality inequalities

$$\text{trace } (AN) \leq \text{trace } (MN) \text{ for all } A \in M.$$

Employing customary notions of convex analysis we shall call a matrix $N$ which satisfies this system of inequalities to be normal to $M$ at $M$. 


Theorem 2. (i) If \( M \) is admissible for \( \theta \) in \( \mathcal{M} \) then there exists a nonnegative definite \( k \times k \) matrix \( N \neq 0 \) which is normal to \( \mathcal{M} \) at \( M \).

(ii) If there exists a positive definite \( k \times k \) matrix \( N \) which is normal to \( \mathcal{M} \) at \( M \) then \( M \) is admissible for \( \theta \) in \( \mathcal{M} \).

Proof. (i) From Theorem 1 we know that \( M \) is \( E \)-optimal for \( H'\theta \) in \( \mathcal{M} \). The general equivalence theory provides a necessary and sufficient condition of optimality in the following form, see Theorem 8 of Pukelsheim (1980, p. 356). Optimality holds if and only if for all \( A \in \mathcal{M} \)

\[
\text{trace}(H'GAG'H) \leq \lambda_{\text{max}}(H'M^*H) = 1/\lambda_{\text{min}}(C_H(M)),
\]

for some generalized inverse \( G \) of \( M \) and some matrix \( E \in \text{conv } S \). Here \( \text{conv } S \) denotes the convex hull of all \( r \times r \) matrices of the form \( zz' \) such that \( z \) is a normalized eigenvector of \( C_H(M) \) corresponding to \( \lambda_{\text{min}}(C_H(M)) \). However, we have seen above that \( C_H(M) = I_r \), and so \( E \) actually is an arbitrary nonnegative definite \( r \times r \) matrix with trace equal to 1.

Define the nonnegative definite matrix \( N = G'HEH'G \). Then

\[
\text{trace } AN \leq 1 = \text{trace } MN \quad \text{for all } A \in \mathcal{M}.
\]

Hence \( N \) cannot be 0, and it satisfies the normality inequalities.

(ii) Let \( A \) be a competing moment matrix satisfying \( A \geq M \). Then \( 0 \leq \text{trace } ((A - M)N) \leq 0 \). Therefore \( \text{trace } ((A - M)N) = 0 \), and positive definiteness of \( N \) forces \( A = M \). Thus admissibility is established.

Our proof provides the additional information that in Theorem 2(i) we can choose \( N \) so as to satisfy \( 1 \leq \text{rank } N \leq r = \text{rank } M \).

Note that rank 1 matrices \( M = cc' \) may well be admissible for the \( k \)-dimensional parameter \( \theta \). By Theorem 1 admissibility then holds if and only if \( M \) is uniquely optimal for \( c'\theta \) in \( \mathcal{M} \), and then Theorem 2(i) admits a rank 1 choice \( N = dd' \).

Admissibility for a subset of the full parameter system admits a similar development, with slightly more technical input concerning information matrices.

3. INFORMATION MATRICES

Consider a fixed \( s \)-dimensional parameter system \( K'\theta \) given by some \( k \times s \) matrix \( K \) of full column rank \( s \). Admissibility for \( K'\theta \) concentrates on the \( s \times s \) information matrix for \( K'\theta \) which, if \( A \in \mathcal{A}(K) \) with \( \mathcal{A}(K) \) defined as in the preceding section, is given by

\[
C_K(A) = (K'KA^{-1}K)^{-1}.
\]

Recall that for the full parameter case a rank deficient moment matrix \( M \) may be admissible. Similarly a rank deficient information matrix \( C_K(A) \) may prove admissible for \( K'\theta \), provided we exercise some care when extending the matrix function \( C_K \) from \( \mathcal{A}(K) \) to the convex cone \( \mathcal{N}_{\text{ND}}(k) \) of all nonnegative definite \( k \times k \) matrices. The appropriate definition for an arbitrary matrix \( A \in \mathcal{N}_{\text{ND}}(k) \) is

\[
C_K(A) = \lim_{\epsilon \downarrow 0}(K'(A + \epsilon I)^{-1}K)^{-1}.
\]
Then $C_K(A)$ is nonsingular if and only if $A \in \mathcal{S}(K)$ and in this case

$$C_K(A) = (K'A^{-1}K)^{-1},$$

see Lemma 2 in Müller-Funk, Pukelsheim & Witting (1985, p. 23). Another representation of the extended matrix function $C_K$ was used in Gaffke (1987), namely

$$C_K(A) = \min_{L_K} L_K AL_L'K,$$

where the minimum is taken over all left inverses $L_K$ of $K$ (i.e. $L_KK = I_s$) and is carried out relative to the Löwner matrix ordering. That the minimum exists is a consequence of the Theorem in Kraft (1983). It can also be seen using the Gauss-Markov Theorem, as follows.

Consider a linear model with expectation $K\beta$ and dispersion matrix $A$, where $\beta \in \mathbb{R}^k$ is the unknown parameter vector. The set $\{L_K\}$ of left inverses of $K$ defines the set of linear unbiased estimators for $\beta$, and the BLUE for $\beta$ corresponds to a particular member $L_K$ such that $L_K AL_K'$ is a minimum. We will call such a matrix $L_K$ a left inverse of $K$ minimizing for $A$, i.e.

$$L_KK = I_s \quad \text{and} \quad C_K(A) = L_K AL_K'.$$

Equivalently one could say that $L_K'$ is a minimum $A$-seminorm generalized inverse of $K'$, see Rao & Mitra (1971, p. 46).

Both expressions for $C_K(A)$ coincide, as shown next.

**Lemma 2.** For all nonnegative definite $k \times k$ matrices $A$ we have

$$\lim_{\epsilon \to 0} (K'(A + \epsilon I)^{-1}K)^{-1} = \min_{L_K} L_K AL_L'K.$$

**Proof.** Since for $\epsilon > 0$ the matrix $A + \epsilon I$ is positive definite, we know from the Gauss-Markov Theorem that

$$\min_{L_K} L_K(A + \epsilon I)L'_K = (K'(A + \epsilon I)^{-1}K)^{-1}.$$

Let $L_K$ be a left inverse of $K$ minimizing for $A$. Then

$$\min_{L_K} L_K AL_L'K \leq \min_{L_K} L_K(A + \epsilon I)L'_K \leq L_K(A + \epsilon I)L'_K,$$

and letting $\epsilon \to 0$ the assertion follows. \qed

With the extended definition of $C_K$ a moment matrix $M \in \mathcal{M}$ is called admissible for $K^\theta$ in $\mathcal{M}$ when no moment matrix $A \in \mathcal{M}$ satisfies $C_K(A) \geq C_K(M)$ and $C_K(A) \neq C_K(M)$.

Again we wish to study a fixed moment matrix $M \in \mathcal{M}$. However, we now choose a full rank decomposition of its information matrix (which we assume to be nonzero)

$$C_K(M) = HH',$$

where with $t = \text{rank} C_K(M)$ the $s \times t$ matrix $H$ has full column rank $t$.

We shall have to investigate the parameter system $H'K^\theta$. The information matrices relative to the representations $(KH)^\theta$ and $H'(K^\theta)$ satisfy the following decomposition rule. The matrix functions $C_{KH}$ and $C_H$ are defined as above with $KH$ and $H$ instead of $K$ and with domains $NND(k)$ and $NND(s)$, respectively.
Lemma 3. For all nonnegative definite $k \times k$ matrices $A$ we have
\[ C_{KH}(A) = C_H(C_K(A)). \]

Proof. When $A$ is positive definite then
\[ C_{KH}(A) = (H' A^{-1} K H)^{-1} = C_H((K' A^{-1} K)^{-1}) = C_H(C_K(A)). \]
Now take a nonnegative definite matrix $A$. For $\epsilon > 0$ then $C_K(A) \leq C_K(A + \epsilon I)$. Since $A + \epsilon I$ is positive definite we obtain $C_H(C_K(A)) \leq C_{KH}(A + \epsilon I)$. The right hand side becomes $C_{KH}(A)$ when $\epsilon \to 0$.

For the converse inequality let $L_H$ be a left inverse of $H$ minimizing for $C_K(A)$, and $L_K$ be a left inverse of $K$ minimizing for $A$. Obviously $L_H L_K$ is a left inverse of $KH$, and by Lemma 2
\[ C_{KH}(A) \leq L_H L_K A L_K' L_H = L_H C_K(A) L_H = C_H(C_K(A)). \]

The two inequalities force equality, and the proof is complete.

An analogous decomposition rule holds for left inverses of $KH$ minimizing for $A$.

Lemma 4. A left inverse $L_{KH}$ of $KH$ is minimizing for $A$ if and only if $L_{KH} = L_H L_K$ for some left inverse $L_K$ of $K$ minimizing for $A$ and some left inverse $L_H$ of $H$ minimizing for $C_K(A)$.

Proof. We first note that if $L_K$ is a given left inverse of $K$, then the set of all left inverses of $K$ is the linear manifold $L_K + B$ where $B$ may be any $s \times k$ matrix with $BK = 0$. From this it is easy to see that $L_K$ is minimizing for $A$ if and only if $L_K A Q_K = 0$, where $Q_K$ denotes the orthogonal projector onto the nullspace of $K'$. Similarly a left inverse $L_{KH}$ of $KH$ is minimizing for $A$ if and only if $L_{KH} A Q_{KH} = 0$, where $Q_{KH}$ is the orthogonal projector onto the nullspace of $(KH)'$.

To prove the direct part of the lemma let $L_{KH}$ be a left inverse of $KH$ minimizing for $A$. Consider the matrix equations
\[ L_{KH} K X = L_{KH}, \quad \text{and} \quad X \cdot [K, A Q_K] = [I_s, 0]. \]

Obviously each of them separately has a solution. Moreover they have a common solution for $X$, by Theorem 2.3.3 in Rao & Mitra (1971, p. 25). In order to apply this theorem we must verify $L_{KH} K [I_s, 0] = L_{KH} [K, A Q_K]$, but this holds true in view of $L_{KH} A Q_{KH} = 0$ and $Q_K = Q_{KH} Q_K$. Setting $L_K = X$ and $L_H = L_{KH} K$, we have a left inverse $L_K$ of $K$ minimizing for $A$, a left inverse $L_H$ of $H$, and $L_{KH} L_K = L_{KH}$. In fact, $L_H$ is minimizing for $C_K(A)$ since by Lemma 3
\[ L_H C_K(A) L_H' = L_H L_K A L_K' L_H = L_{KH} A L_K' L_H = C_{KH}(A) = C_K(C_K(A)). \]

The converse part is immediate: Evidently $L_{KH} L_K$ is a left inverse of $KH$, and $L_{KH} L_K A L_K' L_H = L_H C_K(A) L_H = C_K(C_K(A)) = C_{KH}(A)$.

We shall now use these intermediate results for our discussion of admissibility and optimality.

4. ADMISSIBILITY FOR PARAMETER SUBSETS

Let $M \in \mathcal{M}$ be a fixed moment matrix. We resume the discussion of $M$ being admissible for $K' \beta$ in $M$. Assume that $C_K(M) \neq 0$ and choose a full rank decomposition $C_K(M) = H H'$ as in Section 3. We first present a result similar to Lemma 1.
Lemma 5. Let $A \in \mathcal{M}$ be a competing moment matrix. If $A$ is $E$-optimal for $H'K't$ in $\mathcal{M}$ then $C_K(A) \geq C_K(M)$.

Proof. By construction the range of $C_K(M)$ contains the range of $H$. Applying Lemma 3 we obtain

$$C_{KH}(M) = (H'C_{KH}(M)^{-1}H) = (H'(HH')^{-1}H)^{-1} = I_t.$$  

Optimality of $A$ yields $1 = \lambda_{\min}(C_{KH}(M)) \leq \lambda_{\min}(C_{KH}(A))$. Therefore $I_t \leq C_{KH}(A)$, and pre- and postmultiplication with $H$ and $H'$ yields

$$C_K(M) = HH' \leq HC_H(C_K(A))H' \leq C_K(A).$$

Note that $C_H(C_K(A)) = C_{KH}(A)$ is nonsingular and hence $C_K(A) \in \mathcal{A}(H)$.

The following theorem on admissibility and $E$-optimality parallels Theorem 1.

Theorem 3. The moment matrix $M$ is admissible for $K't$ in $\mathcal{M}$ if and only if $M$ is $E$-optimal for $H'K't$ in $\mathcal{M}$ and for any other $E$-optimal moment matrix $A \in \mathcal{M}$ for $H'K't$ in $\mathcal{M}$ we have $C_K(A) = C_K(M)$.

Proof. Follow the proof of Theorem 1, with Lemma 1 replaced by Lemma 5. Use Lemma 3 for the converse part.

We are now in a position to present our main result: A proof based on $E$-optimality of Theorem 2 of Gaffke (1987).

Theorem 4. (i) If $M$ is admissible for $K't$ in $\mathcal{M}$ then there exists a nonnegative definite $s \times s$ matrix $D \neq 0$ and there exists a left inverse $L_K$ of $K$ minimizing for $M$ such that $L_KDL_K$ is normal to $\mathcal{M}$ at $M$.

(ii) If there exists a positive definite $s \times s$ matrix $D$ and a left inverse $L_K$ of $K$ minimizing for $M$ such that $L_KDL_K$ is normal to $\mathcal{M}$ at $M$ then $M$ is admissible for $K't$ in $\mathcal{M}$.

Proof. (i) By Theorem 3 the moment matrix $M$ is $E$-optimal for $H'K't$ in $\mathcal{M}$, and as shown above $C_{KH}(M) = I_t$. The general equivalence theory tells us that

$$\text{trace}(H'K'tGAG'tKH) \leq 1 \quad \text{for all } A \in \mathcal{M},$$

for some generalized inverse $G$ of $M$ and some nonnegative definite $t \times t$ matrix $E$ with trace equal to 1. Define the matrix $N = G'KHEH'K'tG$. Then

$$\text{trace}(AN) \leq 1 = \text{trace}(MN) \quad \text{for all } A \in \mathcal{M},$$

and $1 \leq \text{rank } N \leq t$. The matrix $L_{KH} = H'K'tG$ satisfies $L_{KH}KH = H'K'tGKH = (C_{KH}(M))^{-1} = I_t$ and $L_{KH}ML_{KH} = H'K'tGMC'KH = I_t = C_{KH}(M)$, and thus is a left inverse of $KH$ minimizing for $M$. Lemma 4 then ensures that $L_{KH} = L_HL_K$ where $L_K$ is a left inverse of $K$ minimizing for $M$. Setting $D = L_HEL_H$ we obtain the desired representation

$$N = L_K'REL_K = L_K'HEL_K = L_K'DL_K.$$
(ii) Let $A$ be a competing moment matrix satisfying $C_K(A) \geq C_K(M)$. Then

$$0 \leq \text{trace} \left\{ (C_K(A) - C_K(M))D \right\} \leq \text{trace} \left\{ (L_KAL_K' - L_KML_K')D \right\} = \text{trace} \left\{ (A-M)L_K'DL_K \right\} \leq 0,$$

and because of positive definiteness of $D$ therefore $C_K(A) = C_K(M)$. The proof gives the additional information that in Theorem 4(i) we can choose the $s \times s$ matrix $D$ so as to satisfy $1 \leq \text{rank} D \leq t = \text{rank} C_K(M)$.

REFERENCES


Lecture Notes in Economics and Mathematical Systems

Managing Editors: M. Beckmann and W. Krelle

297

V. Fedorov H. Läuter (Eds.)

Model-Oriented Data Analysis

Springer-Verlag 1988
Berlin Heidelberg New York London Paris Tokyo