University of Windsor

Scholarship at UWindor

Electronic Theses and Dissertations

Theses, Dissertations, and Major Papers

1983

OPTIMALITY OF CHEMICAL BALANCE WEIGHING DESIGNS.

JOSEPH COSTANTINO. MASARO

University of Windsor

Follow this and additional works at: <https://scholar.uwindsor.ca/etd>

Recommended Citation

MASARO, JOSEPH COSTANTINO., "OPTIMALITY OF CHEMICAL BALANCE WEIGHING DESIGNS." (1983).
Electronic Theses and Dissertations. 2946.

<https://scholar.uwindsor.ca/etd/2946>

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.

CANADIAN THESES ON MICROFICHE

I.S.B.N.

THESES CANADIENNES SUR MICROFICHE

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS RECUE

OPTIMALITY OF CHEMICAL BALANCE
WEIGHING DESIGNS

by

Joseph Costantino Masaro

A Dissertation
submitted to the Faculty of Graduate Studies
and Research through the Department of Mathematics,
in Partial Fulfillment of the requirements for the
Degree of Doctor of Philosophy at
The University of Windsor

Windsor, Ontario, Canada

1983

2

Joseph Costantino Masaro 1983

All rights reserved

789737

To Ann Marie, Jessica, Ronnie,

Robert and Joey.

ABSTRACT

OPTIMALITY OF CHEMICAL BALANCE

WEIGHING DESIGNS

by

Joseph Costantino Masaro

Let N and n be positive integers with $N \geq n$ and let $\mathcal{L}(N,n)$ denote the set of all $N \times n$ matrices $X = (x_{ij})$ with $x_{ij} = -1, 0$ or 1 . Let $\mathcal{S}(N,n)$ be the set of all matrices in $\mathcal{L}(N,n)$ with entries $-1, 1$. Each such matrix in $\mathcal{S}(N,n)$ or $\mathcal{L}(N,n)$ will be called a weighing design matrix. If X_0 minimizes $\phi(X^T X)$ over $\mathcal{S}(N,n)$ for some real valued function ϕ , then X_0 is said to be ϕ -optimum over $\mathcal{S}(N,n)$. The characterization of such X_0 arises from the statistical problems of weighing designs, certain block designs and 2^n fractional factorial designs. The well-known D-, A- and E-optimality criteria are obtained by taking $\phi(X^T X) = \det(X^T X)^{-1}$, $\text{tr}(X^T X)^{-1}$ and the maximum of the eigenvalue of $(X^T X)^{-1}$ respectively. All these criteria are functions of the spectrum of $X^T X$. Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of $X^T X$. Then a more general family of criteria are the following ϕ_p -criteria:

$$\phi_p(X^T X) = \left(\frac{1}{n} \sum_{i=1}^n \lambda_i^{-p} \right)^{\frac{1}{p}}, \quad 0 < p < \infty.$$

The A-criterion is the ϕ_1 -criterion, the E-criterion is the limit of the ϕ_p -criterion as $p \rightarrow \infty$ and the D-criterion is the limit of the ϕ_p -criterion as $p \rightarrow 0^+$.

In this thesis techniques are developed for proving the optimality of designs with respect to the A- and ϕ_p -criteria, $0 \leq p \leq 1$. In particular, A-optimal designs in $\mathcal{L}(N,n)$ are classified for $n = 6$

and N arbitrary. In addition, for the cases $N \equiv 2(\text{mod } 4)$ and $N \equiv 3(\text{mod } 4)$, certain designs are shown to be ϕ_p -optimal in $\mathcal{D}(N,n)$ if N is sufficiently larger than n . In the latter case, it is also shown that A -optimality of certain designs in $\mathcal{L}(N,n)$ implies optimality with respect to the ϕ_p -criteria $0 \leq p \leq 1$. Analogous results are shown to hold for designs in $\mathcal{D}'(N,n)$. Further, in the case $N \equiv 2(\text{mod } 4)$, designs in $\mathcal{D}'(N,n)$ are shown to be optimal with respect to a large class of criteria which includes the ϕ_p -criteria, $0 \leq p \leq \infty$, for all n and N .

PREFACE

The study of optimal chemical balance weighing designs was initiated by Hotelling in 1944. Since then scattered results have appeared in this area. Ehlich (1964), Kiefer (1975), Galil and Kiefer (1980) and Cheng (1978, 1980) strengthened and unified many of these earlier results. In September of 1981 we began to study this problem. The main purpose of this thesis is to add to the existing knowledge on A-optimal weighing designs.

We would like to express our gratitude to Professor C.S. Cheng of the University of California at Berkeley for his useful comments and contributions. With his permission portions of our joint paper "Optimal Weighing Designs" are included in this dissertation (section 4.3). We would also like to thank Professor Mike Jacroux of Washington State University who independently found Theorem 3.3.1 and has generously agreed to publish this result jointly.

The author wishes to express his deepest gratitude to Professor Chi Song Wong for his advice, guidance and support and above all for his encouragement and determination, without which, this thesis would never have been written.

Thanks are also due to the committee members Professors D.S. Tracy and T. Traynor for their useful comments and suggestions, and to the external examiner Dr. A. Hedayat.

Gratitude is also extended to Mrs. Ada Rowland for typing the manuscript.

Finally the author would like to thank the Department of Mathematics at the University of Windsor for extending their patience and understanding.

TABLE OF CONTENTS

DEDICATION	iv
ABSTRACT	v
PREFACE	vii
CHAPTER	
I. INTRODUCTION	1
1.1 Methods of Comparing Results of Experiments: Optimality Criteria	1
1.2 General Optimality Criteria	5
1.3 Optimal Weighing Designs: History and Open Problems	7
1.4 Design Matrices with Entries -1 or 1	12
1.5 Contributions of this Dissertation	13
II. PRELIMINARY DEFINITIONS AND RESULTS	17
2.1 Majorization	17
2.2 Preliminary Results	18
2.3 Some Useful Lemmas	21
2.4 Hadamard Matrices	24
III. OPTIMAL $N \times n$ WEIGHING DESIGN MATRICES WITH $N \equiv 2(\text{MOD } 4)$	26
3.1 Optimal Designs in $\mathcal{D}'(N,n)$	26
3.2 A-optimal Designs in $\mathcal{D}(N,n)$	30
3.3 ϕ_p -optimal Designs in $\mathcal{D}(N,n)$	54
IV. OPTIMAL $N \times n$ WEIGHING DESIGN MATRICES WITH $N \equiv 3(\text{MOD } 4)$	59
4.1 General Results on A-optimality in $\mathcal{D}(N,n)$ and $\mathcal{D}'(N,n)$	59
4.2 A-optimal Designs in $\mathcal{D}(N,n)$	67
4.3 ϕ_p -optimality in $\mathcal{D}(N,n)$ and $\mathcal{D}'(N,n)$	85

V. PROPOSALS FOR FURTHER RESEARCH	92
REFERENCES	94
VITA AUCTORIS	96

CHAPTER I

INTRODUCTION

1.1. Methods of Comparing Results of Experiments: Optimality Criteria.

In conducting an experiment one often assumes the following linear model:

$$y = X_d \beta + e$$

where $y = (y_1, y_2, \dots, y_N)^T$ is the vector of observations,

$\beta = (\beta_1, \beta_2, \dots, \beta_n)^T$ is the vector of unknown model parameters to be estimated, X_d is an $N \times n$ matrix, $N \geq n$, called the design matrix and e is an $N \times 1$ random vector with expectation zero and dispersion matrix $\Sigma_e = \sigma^2 I_N$. If $N = n$, we say that X_d is saturated in the sense that there are no degrees of freedom left for estimating σ^2 .

The β_i 's are estimable if and only if the rank of X_d is n , i.e. $X_d^T X_d$ is nonsingular. Suppose that this is the case. The least squared estimate (or equivalently, blue) $\hat{\beta}_d$ of β is $(X_d^T X_d)^{-1} X_d^T y$ and has dispersion matrix $\sigma^2 (X_d^T X_d)^{-1}$.

Financial or other considerations may require us to take a fixed N which is not too large. Under these circumstances, it is logical for one to choose the design d (i.e. the design matrix X_d) which would produce the "most accurate" estimate $\hat{\beta}_d$ of β . The definition of "most accurate" depends, of course, on the needs of the experimenter. Let $\mathcal{L}$ be the class of all designs under consideration

and $\mathcal{M} = \{M_d\}_{d \in \mathcal{D}}$ the corresponding family of information matrices

$$M_d = X_d^T X_d . \quad \text{Three basic criteria in common use for choosing}$$

and optimal design in $\mathcal{D}$ are:

- (i) D-optimality: minimize $\det(M_d^{-1})$ over $\mathcal{M}$
- (ii) A-optimality: minimize $\text{tr}(M_d^{-1})$ over $\mathcal{M}$
- (iii) E-optimality: minimize $\lambda_1(M_d^{-1})$ over $\mathcal{M}$

where $\lambda_1(M_d^{-1})$ is the maximum eigenvalue of M_d^{-1} .

To understand the meaning of these criteria, let us return to the linear model $y = X_d \beta + e$. Under the assumption that e is normally distributed with mean zero and covariance matrix $\sigma^2 I_N$, $\hat{\beta}_d$ is normally distributed with mean β and dispersion matrix $\sigma^2 (X_d^T X_d)^{-1}$. Thus a $(1-\alpha)100\%$ confidence ellipsoid $\mathcal{E}_{d,\alpha}$ for β can be constructed for each design d in $\mathcal{D}$. In Figure 1 below the ellipsoid $\mathcal{E}_{d,\alpha}$ is presented for the two dimensional case (i.e. $n = 2$).

Figure 1.

For a given level of significance one would like the confidence region to be as "small" as possible. The fact that the confidence ellipsoid $\mathcal{E}_{d,\alpha}$ depends on d but the level of significance $1 - \alpha$ does not implies that one may choose a design d which minimizes the "size" of the confidence region while the level of significance remains constant. It is the interpretation of "size" which leads to the three optimality criteria mentioned above.

A D-optimal design minimizes the generalized variance of $\hat{\beta}_d$ and thus, under normality, minimizes the expected volume (or volume if σ^2 is known) of $\mathcal{E}_{d,\alpha}$ over $\mathcal{D}$.

The A-criterion minimizes the average variance of the parameter estimates $\hat{\beta}_{id}$ over $\mathcal{D}$ ("A" means average) or equivalently the expected mean square error $L^2(\hat{\beta}_d)$ given by $L^2(\hat{\beta}_d) = (\hat{\beta}_d - \beta)^T (\hat{\beta}_d - \beta)$.

From a geometric viewpoint an A-optimal design minimizes the expected length of the diagonal of any rectangle circumscribing the confidence ellipsoid $\mathcal{E}_{d,\alpha}$. Hence an A-optimal design provides a rectangular confidence region, with confidence level at least $1 - \alpha$, which is smallest in the sense indicated above.

An E-optimal design minimizes the expected length of the first principal axis of the confidence ellipsoid $\mathcal{E}_{d,\alpha}$ over $\mathcal{D}$. So in this case $\mathcal{E}_{d,\alpha}$ is not long and narrow. Alternatively an E-optimal design minimizes the maximum variance among all normalized linear combinations of the parameter estimates $\hat{\beta}_{id}$, $i = 1, 2, \dots, n$.

To illustrate the difference among the above criteria consider the following ellipsoidal regions in R^2 (see Figure 2.)

Figure 2.

$$\mathcal{E}_1 : \frac{x_1^2}{1^2} + \frac{x_2^2}{50^2} \leq 1 ,$$

$$\mathcal{E}_2 : \frac{x_1^2}{6^2} + \frac{x_2^2}{15^2} \leq 1 ,$$

$$\mathcal{E}_3 : \frac{x_1^2}{12^2} + \frac{x_2^2}{12^2} \leq 1 .$$

Among the three, $\mathcal{E}_1$ has minimum volume, $\mathcal{E}_2$ possesses the circumscribing rectangle with smallest diagonal and $\mathcal{E}_3$ has the smallest major axis. Thus each of the above regions is "small" in its own special way. In particular, given a choice of the above regions as confidence regions for $\beta = (\beta_1, \beta_2)^T$, one could choose $\mathcal{E}_1$ if β_1 is the parameter of interest, $\mathcal{E}_2$ if minimizing $(\hat{\beta} - \beta)^T (\hat{\beta} - \beta)$ is required and $\mathcal{E}_3$ if estimating $a_1\beta_1 + a_2\beta_2$, $a_1^2 + a_2^2 = 1$ is a primary importance.

The above discussion has focused on the notion of optimal design as a means of improving the accuracy of the parameter estimates in a given experiment. For a discussion of other aspects of optimal design theory (e.g. applications to testing hypothesis), see Kiefer [21], [22] and [23], and Hedayat [15].

1.2 General Optimality Criteria.

The intuitive appeal and computational tractability of the D-, A- and E-optimality criteria have made them the main topics of concentration. However in many cases the experimenter may choose his design to satisfy other optimality criteria. This leads to the

following definitions.

Let $\mathcal{D}$ be a class of designs for a given experiment and $\mathcal{M} = \{M_d\}_{d \in \mathcal{D}}$ the corresponding family of information matrices.

Let ϕ be a function on $\mathcal{M}$ with values in $R \cup \{+\infty\}$. A design d_1 is said to be ϕ -better than a design d_2 if $\phi(M_{d_1}) < \phi(M_{d_2})$. A design d_0 is said to be ϕ -optimal in $\mathcal{D}$ if d_0 minimizes $\phi(M_d)$ over $\mathcal{D}$. The function ϕ is called an optimality criterion.

To illustrate, Kiefer [24] introduced the following family of criteria:

$$\begin{aligned}\phi_p(M) &= \left[\frac{1}{n} \operatorname{tr} M^{-p} \right]^{\frac{1}{p}} \\ &= \left[\frac{1}{n} \sum_{i=1}^n \lambda_i^{-p} \right]^{\frac{1}{p}}, \quad 0 < p < \infty,\end{aligned}$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ are the eigenvalues of M . Clearly the A -criterion is equivalent to the ϕ_1 -criterion. Furthermore, the D -criterion is the limit of the ϕ_p -criteria as $p \rightarrow 0^+$ and the E -criterion is the limit of the ϕ_p -criteria as $p \rightarrow +\infty$. Thus the class of criteria $\{\phi_p\}_{0 \leq p \leq \infty}$ contains the three main ones.

As an example where the ϕ_p -criteria may be of interest, consider the full rank linear model $y = X_d \beta + e$ where e is normal with mean zero and variance $\sigma^2 I_N$. As mentioned in Section 1.1 an A -optimal or ϕ_1 -optimal design minimizes the expected mean squared error of $\hat{\beta}$. If controlling the size of the variance of $L^2(\hat{\beta})$ is of interest a ϕ_2 -optimal design would be appropriate since $\operatorname{var}(L^2(\hat{\beta})) = 2n \sigma^4 [\phi_2(M_d)]^2$.

We remark that the ϕ_p -criteria, for different values of p , may give rise to different designs. We refer the reader to Srivastava and Anderson [31] where examples of ϕ_p -optimal designs, $p = 0, 1, 2, \infty$, among the balanced resolution IV* designs of the 2^m series are presented and compared.

Examples of other optimality criteria are

$$(i) \quad \phi_S(M) = \text{tr}(M^2) \quad (\text{S-optimality})$$

$$(ii) \quad \phi_{1,C}(M) = \frac{1}{n} \text{tr}(CM^{-1}) \quad (\text{L-optimality}).$$

For a discussion of these and other criteria see Kiefer [23] and Hedayat [15].

A general procedure for constructing optimality criteria over $\mathcal{D}$ is:

Let $a_{\mathcal{D}} = \max_{d \in \mathcal{D}} \{\text{tr } M_d\}$. Then for any function

f on $[0, a_{\mathcal{D}}]$ with values in $R \cup \{\infty\}$, define

$$\phi_f(M) = \sum_{i=1}^n f(\lambda_i)$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ are the eigenvalues of M .

One can easily see that, by taking $f(x) = -\log x$ and $f(x) = x^{-p}$, the D and ϕ_p -criteria $0 < p < \infty$ are obtained. The E -criterion can be realized as the limit of ϕ_p -criteria as $p \rightarrow \infty$.

The fundamental problem of optimal design theory is:

(1.2.1) Given a class of functions C determine a design d in $\mathcal{D}$ which is ϕ_f -optimal for all $f \in C$, $d \in \mathcal{D}$.

1.3. Optimal Weighing Designs: History and Open Problems.

Let N and n be positive integers with $N \geq n$ and let

$\mathcal{D}(N,n)$ denote the set of all $N \times n$ matrices $X = (x_{ij})$ with $x_{ij} = -1, 0$ or 1 . Each such matrix will be called a weighing design matrix. These matrices arise, e.g., in the following statistical setting.

Suppose that we want to estimate the weights of n objects by weighing them N times on a chemical balance. Let $x_{ij} = -1, 1$ or 0 according as the j^{th} object is on the left or right scale, or is not present at the i^{th} weighing. Then the $N \times n$ matrix $X = (x_{ij})$ is called the design matrix for the underlying design of experiments. For clarity, we denote the design matrix of a weighing design d by X_d . Let $y_1, y_2, \dots, y_N$ be the readings in N weighing with the total weights on the left scale subtracted from the total weights on the right scale. Let $w_1, w_2, \dots, w_n$ be the actual weights of the n objects. Then with the usual assumption of equal variances and independence of $y_1, y_2, \dots, y_N$, we have the following linear model $y = X_d w + e$, where $y = (y_1, y_2, \dots, y_N)^T$, $w = (w_1, w_2, \dots, w_n)^T$ and e is an $N \times 1$ random vector with expectation zero and dispersion matrix $\Sigma_e = \sigma^2 I_N$.

As discussed in Section 1.1, the w_i 's are estimable if and only if $(X_d^T X_d)^{-1}$ exists and in this case the least squared estimate $\hat{w}$ of w is $(X_d^T X_d)^{-1} X_d^T y$ and has dispersion matrix $\sigma^2 (X_d^T X_d)^{-1}$.

In this context the class of all designs $\mathcal{D}$ under consideration will be identified with $\mathcal{D}(N,n)$ and the corresponding family of information matrices will be denoted by $\mathcal{C}(N,n)$. The fundamental problem then is

(1.3.1) Given a class of functions C determine a design d in $\mathcal{D}(N,n)$ which is ϕ_f -optimal for all $f \in C$, $d \in \mathcal{D}(N,n)$.

Raghavarao [27] and [28], and Bhaskararao [3] had studied this problem, but they only considered designs d such that $X_d^T X_d$ is of the form $aI_n + bJ_n$, where I_n is the $n \times n$ identity matrix and J_n is the $n \times n$ matrix with all entries equal to 1. This is a very stringent restriction. Let $\mathcal{D}^S(N,n)$ be the set of all such designs. Then there is no guarantee that the best design in $\mathcal{D}^S(N,n)$ is really optimal over $\mathcal{D}(N,n)$. Counterexamples do exist (see Cheng [5]). Furthermore, they only considered the A-, D- and E-criteria.

To summarize the more recent contributions toward solving problem (1.3.1) it is convenient to divide the value of N into four cases,

Case j : $N \equiv j \pmod{4}$, $j = 0, 1, 2, 3$.

Case 0 : $N \equiv 0 \pmod{4}$. Let X_0 be a member of $\mathcal{D}(N,n)$ such that $X_0^T X_0 = NI_n$. It has been shown (see Kiefer [24]) that X_0 is ϕ -optimal over $\mathcal{D}(N,n)$ for every nonincreasing-convex orthogonally invariant extended real-valued function ϕ defined on the nonnegative definite $n \times n$ matrices. In particular X_0 is ϕ_f -optimal for every continuous decreasing convex function f . Hotelling [17] also showed that X_0 minimizes the individual variances of the best unbiased estimators of the w_i . Thus since, for $N \leq 264$, an abundance of matrices such as X_0 exist, problem (1.3.1) is solved for the case $N \equiv 0 \pmod{4}$.

Case 1 : $N \equiv 1 \pmod{4}$. Let X_1 be a matrix in $\mathcal{D}(N,n)$ such that $X_1^T X_1 = (N-1)I_n + J_n$. Ehlich [9] showed that X_1 is

D-optimal in the saturated case (i.e. $N = n$) and Payne [26] extended this result to the unsaturated case. Cheng [5] improved this result substantially by showing that X_1 is optimal over $\mathcal{D}(N, n)$ with respect to a very general class of criteria which includes the A-, D- and E-criteria. We remark that in the case $N > n$, such an X_1 can always be obtained when the design X_0 of Case 0 in $\mathcal{D}(N-1, n)$ exists, by adjoining a row of 1's to that design. Unfortunately, in the saturated case, Raghavarao [27] showed that such an X_1 exists only if $n = (\alpha^2 + 1)/2$ for some integer α . Such designs are known for $n = 1, 5, 13, 25$. Ehlich and Zeller [11] state that for $N = n = 9$ the design matrix obtained by them, for which the above diagonal elements of $X^T X$ are all 1 except for a single 5, can be proved D-optimum. Hence, the problem of classifying optimal designs in $\mathcal{D}(N, N)$ with $N \equiv 1 \pmod{4}$ is wide open.

Case 2 : $N \equiv 2 \pmod{4}$. Let X_2, Z_2 be members of $\mathcal{D}(N, n)$ such that

$$X_2^T X_2 = \begin{bmatrix} L & 0 \\ 0 & M \end{bmatrix}$$

where $L = M = (N-2) \frac{I_n}{2} + 2J_{\frac{n}{2}}$ for n even and L and M are

$(N-2) \frac{I_{n+1}}{2} + 2J_{\frac{n+1}{2}}$ for n odd, and

$$Z_2^T Z_2 = (N-1)I_n.$$

It has been shown (see Ehlich [9], Wojtas [32] and Payne [26]) that X_2 is D-optimal in $\mathcal{D}(N, n)$. M. Jacroux recently proved that X_2 is uniquely D-optimal (see Jacroux, Wong and Masaro [18]).

Galil and Kiefer [13] remark that it is not known what other optimality properties X_2 possesses. In the aforementioned paper M. Jacroux also showed that Z_2 is uniquely E-optimal over $\mathcal{D}(N,n)$, $n \geq 3$. Cheng (private communication) noted that when $N = n$, X_2 and Z_2 have the same performance under the A-criterion and X_2 is ϕ_p -better than Z_2 for $0 \leq p \leq 1$. More generally, evidence suggests the following conjecture.

For $N \geq n$, $N \equiv 2(\text{mod } 4)$ there exists a number

$\alpha(N,n) \geq 1$ such that $\alpha(N,N) = 1$ and

X_2 is ϕ_p -optimal for $0 \leq p \leq \alpha(N,n)$ and

Z_2 is ϕ_p -optimal for $p \geq \alpha(N,n)$.

Case 3 : $N \equiv 3(\text{mod } 4)$. This is the most difficult case. Let X_3 be a matrix in $\mathcal{D}(N,n)$ such that $X_3^T X_3 = (N+1)I_n - J_n$. Then Galil and Kiefer [13] showed that X_3 is D-optimal if $N \geq 2n - 5$.

In the case $N < 2n - 5$ the situation is less pretty. For example in the case $N = n = 11$ there are at least three different forms of $X^T X$ which are D-optimal. For a discussion of this and other cases see Galil and Kiefer [13] and [14]. We remark that Galil and Kiefer do not discuss D-optimal matrices with zero entries. Concerning E-optimality, Jacroux (see Jacroux, Wong and Masaro [18]) has shown that if Z_2 in $\mathcal{D}(N-1, n)$, $n \geq 4$, is as defined in Case 2, then

any matrix of the form $\begin{bmatrix} Z \\ V^T \end{bmatrix}$, where V is any $n \times 1$ vector

with entries $-1, 0$ or 1 is E-optimal over $\mathcal{D}(N,n)$. The study of other optimality properties seems to have been neglected. In this

dissertation we make a contribution in this direction.

1.4. Design Matrices with Entries -1 or 1 .

Let $\mathcal{D}'(N,n)$ denote the set of all $N \times n$ matrices, $N \geq n$, with entries -1 or 1 and let $C'(N,n)$ be the family of all matrices of the form $X^T X$ where X ranges over $\mathcal{D}'(N,n)$.

Design matrices such as those in $\mathcal{D}'(N,n)$ occur in the setting of 2^k fractional factorial designs of resolution $2t+1$ (i.e. odd resolution). This type of design allows for the estimation of all effects up to the t -factor interactions assuming the $(t+1)$ -factor and higher interactions zero. The linear model for this type of design is

$$y_{j_1 j_2 \dots j_k} = \beta^0 + \sum \beta_i^1 x_i + \sum_{1 \leq i < j \leq k} \beta_{ij}^2 x_i x_j + \dots$$

$$+ \sum_{1 \leq i_1 < i_2 < \dots < i_t \leq k} \beta_{i_1 i_2 \dots i_t}^t x_{i_1} x_{i_2} \dots x_{i_t} + \dots + \epsilon_{j_1 j_2 \dots j_k}$$

where $x_i = 1$ or -1 according as $j_i = 1$ or 0 , $E(\epsilon_{j_1 j_2 \dots j_k}) = 0$

and $\text{var}(\epsilon_{j_1 j_2 \dots j_k}) = \sigma^2$. In matrix notation the model can be

written as $y = X_d \beta^* + e$ where X_d is an $N \times n$ matrix with entries -1 or 1,

$$\beta^{*T} = (\beta_0, \beta_1^1, \beta_2^1, \dots, \beta_k^1, \beta_{12}^2, \beta_{13}^2, \beta_{14}^2, \dots, \beta_{kk-1}^2, \dots,$$

$$\beta_{12 \dots t}^t, \dots, \beta_{k-t+1 \dots k-1k}^t)$$

and e is an $N \times 1$ random vector with mean zero and dispersion matrix $\sigma^2 I_N$. Here the number of parameters to be estimated is

$n = 1 + \binom{k}{1} + \binom{k}{2} + \dots + \binom{k}{t}$. The number of runs N is called a fraction.

The set of design matrices for a 2^k fractional factorial experiment of odd resolution with n parameters to be estimated and N runs is, in general, a proper subset of $\mathcal{D}'(N, n)$. Thus the results in Section 1.3 where the optimal designs belong to $\mathcal{D}'(N, n)$ are perfect for the problem of 2^k fractional factorial designs. In addition Cheng [5] proved that if X_3 is as defined in Case 3 of Section 1.3 then X_3 is optimal in $\mathcal{D}'(N, n)$ with respect to a large class of criteria called the generalized Type 2 criteria. Cheng also mentions a variety of results on E-optimal designs in $\mathcal{D}'(N, n)$.

Very few results have appeared concerning A-optimal (or more generally, ϕ_p -optimal) designs in $\mathcal{D}'(N, n)$ with $N \equiv 2 \pmod{4}$ and $N \equiv 3 \pmod{4}$. Some are provided in this paper. Our results, however, are only applicable to designs of resolution III.

Finally, we remark that there is a large literature on factorial designs. For a systematic development of the basic concepts of the theory and for further references, we refer the reader to Raktoe, Hedayat and Federer [29].

1.5 Contributions of this Dissertation.

It is well known that, in general, proving a design optimal with respect to a given criterion ϕ can be extremely difficult. One only has to look at the history of D-optimality (e.g. Ehlich [9] and [10], Payne [26] and Galil and Kiefer [13] and [14]) to see the effort (both mathematical and computer aided) put forth. The

difficulty, of course, is the enormous number of cases to be considered. Indeed the number of designs in $\mathcal{D}(N, n)$ is $\binom{3^n}{N}$ and in $\mathcal{D}'(N, n)$ is $\binom{2^{n-1}}{N}$. For example, $\mathcal{D}'(6, 6)$ has 906,192 members, 556,192 of which have full rank (see Raktoe, Hedayat and Federer [29]). Furthermore, the search for optimal designs in $\mathcal{D}(N, n)$ with respect to criteria, other than D-optimality, is expected to be more difficult. Indeed the D-criterion has the nice property that there always exists a matrix in $\mathcal{D}'(N, n)$ which is D-optimal over $\mathcal{D}(N, n)$. This as we shall see is not always true of the other optimality criteria.

In this dissertation A-optimal designs in $\mathcal{D}(N, n)$ are classified for $n = 6$ and N arbitrary. In addition, for the cases $N \equiv 2 \pmod{4}$ and $N \equiv 3 \pmod{4}$, ϕ_p -optimal designs $0 \leq p \leq 1$ are found when N is large relative to n . In the latter case, it is also shown that A-optimality implies optimality with respect to a large class of criteria which includes the ϕ_p -criteria $0 \leq p \leq 1$. Further, designs in $\mathcal{D}'(N, n)$, $N \geq n$, $N \equiv 2 \pmod{4}$ arbitrary, are proved optimal with respect to all relevant criteria. Also some ϕ_p -optimal designs, $0 \leq p \leq 1$, are found in $\mathcal{D}(N, n)$, $N \equiv 3 \pmod{4}$, and $n \leq 7$.

Finally, we remark that problems and techniques in optimal design theory are often related to problems and techniques in other fields. For example in graph theory there is a problem of finding a graph which has a maximum number of spanning trees among all graphs with a given number of vertices and edges (see Kelmans and Chelnokov [20] and Shier [30]). N. Gaffke [12] in his doctoral dissertation,

noted that the above problem in graph theory is related to the problem of finding a D-optimal incomplete block design. A common technique in optimal design theory as well as in other related areas is to embed the class of designs of interest in a larger family in which it is (mathematically) easier to find a best design. One then hopes that this best design belongs to the original class. Cheng [4] successfully applied this technique to optimal design theory, then, in Cheng [6], used the same technique to show that a regular complete multipartite graph has the maximum number of spanning trees among all simple graphs with the same number of vertices and edges. Techniques developed in this thesis can be used to obtain results in graph theory. As an example, we remark first that the conjecture

"A graph maximizing the total number of spanning trees over all graphs with a given number of edges and vertices is nearly balanced"

is related to the so-called Mitchell-John conjecture (see John and Mitchell [19]) in statistical design:

"If an incomplete block design is D-optimal (or A-optimal) it is a regular graph design".

Using the methods in the dissertation, it can be shown that, at least in a large system, the above conjectures are true (see Cheng, Masaro and Wong [8]).

The relations among optimal designs in graph theory, statistical theory, etc., can be further seen from the viewpoint that it is possible for two or more different sets $\mathcal{D}_i$ of designs to be embedded in the same larger set of designs $\mathcal{D}$ and a best design d

exists in $\mathcal{D}$ such that $d \in \mathcal{D}_i$ for all i . This means that a family of problems may be solved simultaneously and in a unified way.

For brevity, throughout this thesis results and examples taken from [33], [34] and [35] will not be mentioned.

CHAPTER II

PRELIMINARY DEFINITIONS AND RESULTS

2.1 Majorization.

Majorization is a useful tool in optimality theory and for this reason we include a brief discussion. For a more in depth treatment, we refer the reader to Marshall and Olkin [25].

Let $x = (x_1, x_2, \dots, x_n)^T$ be a vector in R^n . The symbol $x_{[i]}$ will denote the i^{th} largest component of x , so that $x_{[1]} \geq x_{[2]} \geq \dots \geq x_{[n]}$. If x and y are elements of R^n , we say

that x is majorized by y , and write $x \prec y$, if

$$\sum_{i=1}^k x_{[i]} \leq \sum_{i=1}^k y_{[i]}, \quad k = 1, 2, \dots, n-1 \quad \text{and} \quad \sum_{i=1}^n x_{[i]} = \sum_{i=1}^n y_{[i]}.$$

The following theorem (see Marshall and Olkin [25]) is the main reason majorization plays a fundamental role in finding optimal designs:

Theorem 2.1.1. Let $x, y \in R^n$. Then $x \prec y$ if and only if

$$\sum_{i=1}^n f(x_i) \leq \sum_{i=1}^n f(y_i) \quad \text{for all continuous convex functions } f.$$

Now let M be an $n \times n$ symmetric matrix. Let $\lambda_1(M)$ be the i^{th} largest eigenvalue of M and let $\lambda(M) = (\lambda_1(M), \lambda_2(M), \dots, \lambda_n(M))^T$.

The following theorems (see Marshall and Olkin [25]) will be used in the sequel.

Theorem 2.1.2. (Ky Fan). Let M and $\bar{M}$ be $n \times n$ symmetric matrices of the form

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}, \quad \bar{M} = \begin{bmatrix} M_{11} & 0 \\ 0 & M_{22} \end{bmatrix},$$

where M_{11} is $\ell \times \ell$ and M_{22} is $m \times m$, $\ell + m = n$.

Then $\lambda(\bar{M}) \leq \lambda(M)$.

Theorem 2.1.3. (Schur). Let M be an $n \times n$ symmetric matrix and $d = (m_{11}, m_{22}, \dots, m_{nn})^T$ the vector of diagonal elements of M . Then $d \leq \lambda(M)$.

2.2 Preliminary Results.

We begin with an application of majorization to optimal design theory. The following result is due to Kiefer [4].

Theorem 2.2.1. Let $N \equiv 0 \pmod{4}$ and let X_0 be a member of

$\mathcal{D}(N, n)$ such that $X_0^T X_0 = NI_n$. Then X_0 is ϕ_f -optimal in

$\mathcal{D}(N, n)$ for every continuous decreasing convex function f on $[0, nN]$.

Proof. Let X_0 and f be as stated above and let $X \in \mathcal{D}(N, n)$.

Let $d = (d_{11}, \dots, d_{nn})^T$ be the vector of diagonal elements of $X^T X$

and let $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$ be the vector of eigenvalues of $X^T X$.

By Theorem 2.1.3, $d \leq \lambda$; hence Theorem 2.1.1 implies that

$$\sum_{i=1}^n f(d_{ii}) \leq \sum_{i=1}^n f(\lambda_i). \quad \text{Since } N \geq d_{ii}, i = 1, 2, \dots, n \text{ and } f$$

is decreasing, we have $\sum_{i=1}^n f(N) \leq \sum_{i=1}^n f(d_{ii}) \leq \sum_{i=1}^n f(\lambda_i)$. Whence

$$\phi_f(X_0^T X_0) \leq \phi_f(X^T X). \quad \text{This completes the proof.}$$

In general, given a class of functions C , one hopes to find an

optimal design in $\mathcal{D}(N, n)$ which is ϕ_f optimal for all f in C .

For $N \equiv 0 \pmod{4}$, Theorem 2.2.1 shows that we can take C to be the family of all continuous decreasing convex functions on $[0, nN]$.

Unfortunately when N is not $\equiv 0 \pmod{4}$ the class C must be reduced in size in order to solve the problem. Such a reduction was made by Cheng [4] in the following definitions.

(a) An optimality criterion ϕ_f is of Type 1 if

- (1) f is continuous, strictly convex and strictly decreasing on $[0, nN]$. (We allow $f(0^+) = +\infty$).
- (2) f is continuously differentiable on $(0, nN)$ and f' is strictly concave on $(0, nN)$.

That is: $f' < 0$, $f'' > 0$ and $f''' < 0$ on $(0, nN)$.

(b) ϕ_f is a criterion of Type 2 if f is as in (a) except that strict concavity of f' is replaced by strict convexity, i.e. $f''' > 0$ on $(0, nN)$.

(c) ϕ is a generalized criterion of Type i ($i = 1, 2$) if it is the pointwise limit of a sequence of Type i criteria.

Note that the A -, D - and ϕ_p -criteria are of Type 1 and the E -criterion is generalized Type 1 criterion (being the limit of ϕ_p -criteria as $p \rightarrow \infty$). An example of a Type 2 criterion is ϕ_f where $f(x) = \epsilon x^3 - bx$ where $\epsilon > 0$, $b > 0$ and ϵ is small compared with b .

The following theorem, due to Cheng [4], is a useful tool for proving optimality of designs.

Theorem 2.2.2. Let $\mathcal{M} = \{M_d\}_{d \in \mathcal{D}}$ be a class of $n \times n$ symmetric

nonnegative definite matrices.

(a) Suppose that $M_{d*} \in \mathcal{M}$ is either a multiple of I_n or has two distinct nonzero eigenvalues $\lambda > \mu$ such that the multiplicity of μ is $n-1$ and

$$(2.2.1) \quad M_{d*} \text{ maximizes } \text{tr} M_d \text{ over } \mathcal{M}$$

$$(2.2.2) \quad \text{tr}(M_{d*}^2) < (\text{tr} M_{d*})^2 / (n-1)$$

$$(2.2.3) \quad M_{d*} \text{ maximizes } \text{tr} M_d - [n/(n-1)]^{1/2} \times \\ [\text{tr}(M_d^2) - (\text{tr} M_d)^2/n]^{1/2} \text{ over } \mathcal{M}.$$

Then M_{d*} is optimal over $\mathcal{M}$ with respect to any generalized criterion of Type 1.

(b) Suppose $M_{d*} \in \mathcal{M}$ is either a multiple of I_n or has two distinct nonzero eigenvalues $\lambda > \mu$ such that λ has multiplicity $n-1$ and

$$(2.2.4) \quad M_{d*} \text{ maximizes } \text{tr} M_d \text{ over } \mathcal{M}$$

$$(2.2.5) \quad M_{d*} \text{ maximizes } \text{tr} M_d - \{n(n-1)[\text{tr} M_d^2 - (\text{tr} M_d)^2/n]\}^{1/2} \\ \text{over } \mathcal{M}.$$

Then M_{d*} is optimal over $\mathcal{M}$ with respect to any generalized criterion of Type 2.

In settings where $\text{tr} M_d$ is constant for all $d \in \mathcal{D}$, (2.2.1) and (2.2.3) (or (2.2.4) and (2.2.5)) can be replaced by

$$(2.2.6) \quad M_{d*} \text{ minimizes } \text{tr} M_d^2 \text{ over } \mathcal{M}, \text{ and the condition}$$

" $f' < 0$ " in the definitions of Type 1 and Type 2 criteria can be dropped. Further, if $f(0^+) = +\infty$ then condition (2.2.2) is unnecessary.

Using Theorem 2.2.2, Cheng [5] proved the following results.

Theorem 2.2.3. Let $N \equiv 1 \pmod{4}$ and let $X_1 \in \mathcal{D}(N, n)$ be such that

$X_1^T X_1 = (N-1)I_n + J_n$. Then X_1 is optimal over $\mathcal{D}(N,n)$ with respect to any generalized criteria of Type 1.

Theorem 2.2.4. Let $N \equiv 3(\text{mod } 4)$ and let $X_3 \in \mathcal{D}'(N,n)$ be such that $X_3^T X_3 = (N+1)I_n - J_n$. Then X_3 is optimal over $\mathcal{D}'(N,n)$ with respect to any generalized Type 2 criterion.

In Chapter III we will prove a result analagous to Theorem 2.2.3 for designs in $\mathcal{D}'(N,n)$, $N \equiv 2(\text{mod } 4)$. However, for designs in $\mathcal{D}(N,n)$ and $N \equiv 2(\text{mod } 4)$ or $N \equiv 3(\text{mod } 4)$ no similar result is known at the present time.

2.3 Some Useful Lemmas.

The results of this section will be useful in reducing the number of designs among which one must search in order to find an optimum one. We will need the notion of equivalence. Two matrices X and Y in $\mathcal{D}(N,n)$ ($\mathcal{D}'(N,n)$) are said to be equivalent if Y can be obtained from X by a finite number of transformations of the form: multiplication of any row by -1 , interchanging any two rows, or transposition. It is straightforward to show that if X and Y are equivalent and $X^T X$ is nonsingular then $X^T X$ and $Y^T Y$ are similar. So, for design purposes, equivalent matrices have the same optimality properties.

The following lemma is due to Ehlich [9].

Lemma 2.3.1. Let $X \in \mathcal{D}'(N,n)$, $N \equiv 2(\text{mod } 4)$. Then X is equivalent to a matrix Y in $\mathcal{D}'(N,n)$ such that

$$Y^T Y = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

where $a_{ij} \equiv 2(\text{mod } 4)$, $c_{ij} \equiv 2(\text{mod } 4)$ and $b_{ij} \equiv 0(\text{mod } 4)$.

The following technical lemmas will be useful.

Lemma 2.3.2. (Ehlich [9]). Let x_1, x_2, x_3 be vectors in R^n , $n \equiv 2 \pmod{4}$, with entries ± 1 . If $x_1 \cdot x_2 \equiv 0 \pmod{4}$ and $x_1 \cdot x_3 \equiv 0 \pmod{4}$ then $x_2 \cdot x_3 \equiv 2 \pmod{4}$.

Lemma 2.3.3. Let $X \in \mathcal{A}(7,5)$ be such that

$$X^T X = \begin{bmatrix} 6 & \pm 1 & & & \\ & \vdots & & & \\ & \pm 1 & 6 & & \\ & & & B & \\ & & & & \pm 1 \\ & & B^T & & \\ & & & 7 & \\ & & & & 7 \end{bmatrix}.$$

Then $|b_{ij}| \geq 2$ for at least one pair (i,j) .

Proof. Suppose $b_{ij} = 0$ for all i,j . We may assume

$$X = [x_1, x_2, x_3, x_4, x_5],$$

where $x_1^T = (1, 0, 1, 1, 1, 1, 1)$ and $x_2^T = (0, 1, 1, 1, 1, -1, -1)$.

Write $x_1^T = (1, 0, u_1^T)$, $x_2^T = (0, 1, u_2^T)$ and $x_3^T = (a, b, u^T)$. Then

$u_1^T = (1, 1, 1, 1, 1)$, $u_2^T = (1, 1, 1, -1, -1)$ and the entries of x_3

are 1 or -1. By assumption, $x_1 \cdot x_3 = 0$ and $x_2 \cdot x_3 = 0$. Thus,

$$a + u_1 \cdot u = 0 \text{ and } b + u_2 \cdot u = 0.$$

Therefore

$$a + b + (u_1 + u_2) \cdot u = 0.$$

Since $u_1^T + u_2^T = (2, 2, 2, 0, 0)$, $a + b$ cannot be zero. Hence

$$a = b = \pm 1.$$

Applying similar arguments to the other columns of X , we conclude that X must have the form:

$$X = \begin{bmatrix} 1 & 0 & a_1 & a_2 & a_3 \\ 0 & 1 & a_1 & a_2 & a_3 \\ u_1 & u_2 & v_1 & v_2 & v_3 \end{bmatrix},$$

where $|a_i| = 1$ and the v_i 's are 5×1 column vectors with entries ± 1 . By multiplying columns x_3, x_4, x_5 by -1 if necessary, we may assume that $a_i = -1, i = 1, 2, 3$. Thus since $x_i \cdot x_j = 0, i = 1, 2, j = 3, 4, 5, u_1 \cdot v_j = u_2 \cdot v_j = 1, j = 1, 2, 3$. Also $|x_i \cdot x_j| = 1$ for $i \neq j, i, j = 3, 4, 5$. Thus $v_i \cdot v_j = -3$ or -1 . Since $u_1^T = (1, 1, 1, 1, 1)$ and $u_2^T = (1, 1, 1, -1, -1)$, each v_j must be of the form $v_j^T = (p^T, q^T)$ where p is the 3×1 and has exactly two entries equal to 1 and q is 2×1 and has exactly one entry equal to 1. There are only six possible choices for the v_j . They are listed as columns of the following matrix:

$$V = \begin{bmatrix} 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 & -1 & 1 \end{bmatrix}.$$

Since

$$V^T V = \begin{bmatrix} 5 & 1 & 1 & -3 & 1 & -3 \\ 1 & 5 & -3 & 1 & -3 & 1 \\ 1 & -3 & 5 & 1 & 1 & -3 \\ -3 & 1 & 1 & 5 & -3 & 1 \\ 1 & -3 & 1 & -3 & 5 & 1 \\ -3 & 1 & -3 & 1 & 1 & 5 \end{bmatrix},$$

it is clear that if the set $\{v_{i_1}, v_{i_2}, \dots, v_{j_k}\}$ satisfies

$v_{i_r} \cdot v_{j_s} = -3$ for $i_r \neq j_s$ then $k \leq 2$. But our assumption that all $b_{ij} = 0$ requires $k = 3$. This contradiction proves the lemma.

We note that Lemma 2.3.3 is the best we can do. For example,

let

$$X = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 & -1 \end{bmatrix}.$$

Then

$$X^T X = \begin{bmatrix} 6 & 1 & 0 & 0 & 2 \\ & 6 & 0 & 0 & 0 \\ & & 7 & -1 & -1 \\ & * & & 7 & -1 \\ & & & & 7 \end{bmatrix}.$$

2.4 Hadamard Matrices.

Let $N \equiv 0 \pmod{4}$. An $N \times N$ matrix H_N is said to be a Hadamard matrix of order N if each entry of H_N is -1 or 1 and $H_N^T H_N = I_N$. Theorem 2.2.1 shows that if X_0 is any $N \times n$ matrix consisting of n distinct columns of H_N , then X_0 is ϕ_f -optimal in $\mathcal{L}(N, n)$ for all continuous decreasing convex functions f . Thus Hadamard matrices play an important role in the search for optimal designs. At the moment, 268 is the smallest N for which we do not know if a Hadamard matrix of order N exists.

Galil and Kiefer [13] remark that for fixed n it is unnecessary to demand the existence of a Hadamard matrix H_N for all N . Instead

the following proposition may be used.

Proposition 2.4.1. For $n \geq 4$, let $N_n = \min\{j : j \geq n, j \equiv 0 \pmod{4}\}$.

Suppose H_j exists for all $j \equiv 0 \pmod{4}$ satisfying $N_n \leq j < 2N_n$.

Then, for all $N \geq n$ with $N \equiv 0 \pmod{4}$, there is an X_0 in $\mathcal{O}(N, n)$

with $X_0^T X_0 = NI_n$.

Thus, for example, when $n = 6$, the existence of H_8 and H_{12} implies that X_0 exists for $n = 6$ and all $N \geq n$, $N \equiv 0 \pmod{4}$.

For more on Hadamard matrices we refer the reader to Hedayat and Wallis [16]. Finally, we remark that Hadamard matrices are also useful in constructing optimal designs in $\mathcal{O}(N, n)$ even if N is not congruent to $0 \pmod{4}$.

CHAPTER III

OPTIMAL $N \times n$ WEIGHING DESIGN MATRICES

WITH $N \equiv 2(\text{MOD } 4)$

3.1 Optimal Designs in $\mathcal{D}'(N, n)$.

Throughout this chapter N will be congruent to $2(\text{MOD } 4)$ and X_2 will be a matrix in $\mathcal{D}'(N, n)$ such that

$$(3.1.1) \quad X_2^T X_2 = \begin{bmatrix} L & 0 \\ 0 & M \end{bmatrix},$$

where $L = M = (N-2) I_{\frac{n}{2}} + 2J_{\frac{n}{2}}$ if n is even and L and M are

$$(N-2) I_{\frac{n+1}{2}} + 2J_{\frac{n+1}{2}} \text{ if } n \text{ is odd} .$$

Ehlich [9] and Payne [26] proved that X_2 is D-optimal in $\mathcal{D}'(N, n)$ and recently Cheng [5] proved that X_2 is E-optimal over $\mathcal{D}'(N, n)$. In this section it is proved in a unified way that X_2 is optimal with respect to a large subclass of the Type 1 criteria. This approach gives a new proof of the result of Ehlich and Payne mentioned above.

We require the following lemma.

Lemma 3.1.1. Let $\mathcal{B}$ be the family of $n \times n$ matrices of the form:

$$B(s, t) = \begin{bmatrix} (N-2)I_s + 2J_s & 0 \\ 0 & (N-2)I_t + 2J_t \end{bmatrix},$$

where $s + t = n$. Let $s_0 = t_0 = n/2$ if n is even,

$s_0 = t_0 + 1 = (n+1)/2$ if n is odd. Then the eigenvalues of

$B(s_0, t_0)$ are majorized by the eigenvalues of all other matrices

$B(s, t)$ in $\mathcal{B}$.

Proof. Suppose N is odd, then $s_0 = (n+1)/2$ and $t_0 = (n-1)/2$.

The vector of eigenvalues of $B(s_0, t_0)$ arranged in decreasing order is $x = (N+n-1, N+n-3, N-2, \dots, N-2)^T$. For $B(s, t)$, $s \geq t \geq 0$ the corresponding vector is $y = (N+2s-2, N+2t-2, N-2, \dots, N-2)^T$. Since n is odd, $s \geq (n+1)/2$, so $2s - 2 \geq n - 1$. Hence

$$\sum_{i=1}^{n-1} x_i \leq \sum_{i=1}^{n-1} y_i \quad \text{and} \quad \sum_{i=1}^n x_i = \sum_{i=1}^n y_i. \quad \text{Thus } x \prec y. \quad \text{The proof is}$$

similar for n even.

We now prove our main result (M. Jacroux independently found this result. See Jacroux, Wong and Masaro [18]).

Theorem 3.1. Let X_2 be as in (3.1.1). Then X_2 is optimal in

$\mathcal{B}'(N, n)$ with respect to any Type 1 criterion ϕ_f such that

$\lim_{x \rightarrow 0^+} f(x) = +\infty$. Further up to similarity of $X_2^T X_2$, X_2 is unique.

Proof. Let $X \in \mathcal{B}'(N, n)$. By Lemma 2.3.1 we may assume that

$$X^T X = \begin{bmatrix} A_{s \times s} & B \\ B^T & C_{t \times t} \end{bmatrix},$$

where $|a_{ij}| \geq 2$, $|c_{ij}| \geq 2$. Thus,

$$(3.1.2) \quad \text{tr} A^2 \geq \text{tr}[(N-2)I_s + 2J_s]^2$$

and

$$(3.1.3) \quad \text{tr} B^2 \geq \text{tr}[(N-2)I_t + 2J_t]^2.$$

Hence by Theorem 2.2.2 (since $f(0^+) = +\infty$) we have

$$(3.1.4) \quad \phi_f(A) \geq \phi_f((N-2)I_s + 2J_s)$$

and

$$(3.1.5) \quad \phi_f(B) \geq \phi_f((N-2)I_t + 2J_t) .$$

So by Ky Fan's result (Theorem 2.1.2) and Lemma 3.1.1,

$$\begin{aligned} \phi_f(X^T X) &\geq \phi_f \left(\begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix} \right) \\ &= \phi_f(A) + \phi_f(C) \\ &\geq \phi_f((N-2)I_s + 2J_s) + \phi_f((N-2)I_t + 2J_t) \\ (3.1.6) \quad &= \phi_f(B(s, t)) \\ &\geq \phi_f(B(s_0, t_0)) \\ &= \phi_f(X_2^T X_2) . \end{aligned}$$

For uniqueness, suppose that $\phi_f(X^T X) = \phi_f(X_2^T X_2)$. Then all inequalities in (3.1.4), (3.1.5) and (3.1.6) become equalities. Letting $f(x) = -\log x$, we see that $\det(X^T X) = \det(A) \det(C)$. Hence

$$X^T X = \begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix} .$$

Further, since $\phi_f(A) = \phi_f((N-2)I_s + 2J_s)$, then from the proof of Theorem 2.2 of Cheng [4], A must have the same eigenvalues as $(N-2)I_s + 2J_s$. Similarly C has the same eigenvalues as $(N-2)I_t + 2J_t$. Finally, since $\phi_f(B(s, t)) = \phi_f(X_2^T X_2)$, then

$$f(N+2s-2) + f(N+2t-2) = \begin{cases} f(N+n-1) + f(N+n-2) & \text{if } n \text{ is odd} \\ 2 f(N+n-2) & \text{if } n \text{ is even} \end{cases}$$

Letting $f(x) = 1/x$ and solving for s and t we find that

$s = s_0$, $t = t_0$. Thus $B(s, t) = B(s_0, t_0) = X_2^T X_2$. Thus $X^T X$

and $X_2^T X_2$ have the same eigenvalues and therefore are similar. This

completes the proof.

We remark that in general, X_2 is not ϕ_f -optimal for every continuous convex function f . For example, let $N = n = 6$,

$$X_2 = \begin{bmatrix} 1 & 1 & -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix}$$

Then by Theorem 3.1.1, X_2 is optimal with respect to any Type 1 criteria ϕ_f such that $f(0^+) = +\infty$. However, the eigenvalues 10, 10, 4, 4, 4, 4 of $X_2^T X_2$ are not majorized by the eigenvalues of all matrices in $\mathcal{D}'(N, n)$. Indeed, if we take

$$X = \begin{bmatrix} 1 & -1 & -1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \end{bmatrix}$$

then

$$X^T X = \begin{bmatrix} 6 & -2 & -2 & -2 & 0 & 0 \\ -2 & 6 & -2 & -2 & 0 & 0 \\ -2 & -2 & 6 & 2 & 0 & 0 \\ -2 & -2 & 2 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 & -2 \\ 0 & 0 & 0 & 0 & -2 & 6 \end{bmatrix}$$

The eigenvalues of $X^T X$ are 10.4721, 8, 8, 4, 4, 1.5279 which do not majorize the eigenvalues of $X_2^T X_2$.

It should be noted that designs X_2 satisfying condition (3.1.1)

can be found for many classes $\mathcal{D}'(N, n)$. Indeed, let $n \leq N - 2$ and $X_0 \in \mathcal{D}'(N-2, n)$ such that $X_0^T X_0 = (N-2)I_n$ (from section 2.4 many such X_0 exist). Then an X_2 , as above, can be obtained by adjoining to X_0 two rows, one consisting entirely of 1's and the other consisting of $\lfloor n/2 \rfloor$ 1's and $n - \lfloor n/2 \rfloor$ -1's. If $n = N$ designs such as X_2 have been constructed, using circulant matrices, by Ehlich [9] in all cases $n \leq 38$ except $n = 22$ and 34 . Other designs in these cases were obtained by Yang [36], who in the references cited by him there also constructed X_2 for $n = 42, 46, 48, 52$. For $n = N - 1$ the appropriate matrix X_2 in $\mathcal{D}'(N, N-1)$ can be obtained by removing any column from the corresponding $\bar{X}_2$ in $\mathcal{D}'(N, N)$.

3.2. A-optimal Designs in $\mathcal{D}(N, n)$.

Let X_2 be as in (3.1.1). We conjecture that

(C1) X_2 is A-optimal in $\mathcal{D}(N, n)$.

In support of (C1), we note that (Theorem 3.3.1) for any n , X_2 above is A-optimal for sufficiently large N . The problem is left for small N which can be done with little help from computers. For example, when $N = n = 6$, $\mathcal{D}(N, n)$ has $\binom{3^6}{6}$ elements. By using Ky Fan's result on majorization (Theorem 2.1.2) we shall prove that for $n = 6$ and $N \geq 6$, X_2 is A-optimal. Our technique can be used to prove other results. It is our hope that a refined use of such methods can be applied in a more general setting.

The following well-known result gives a lower bound for $\text{tr}(X^T X)^{-1}$ over $\mathcal{D}(N, n)$. With such a lower bound, one gets a feeling in comparing one $\text{tr}(X^T X)^{-1}$ with another.

Theorem 3.2.1. Let $X \in \mathcal{B}(N, n)$. Then

$$\text{tr}(X^T X)^{-1} \geq n^2 / \text{tr}(X^T X) \geq n/N.$$

In particular $\text{tr}(X^T X)^{-1} \geq 1$ if X is saturated. Moreover

$\text{tr}(X^T X)^{-1} = n^2 / \text{tr}(X^T X)$ if and only if $X^T X$ is a scalar multiple of I_n .

Proof. With the trace inner product, we have, by the Schwarz inequality,

$$((X^T X)^{-1/2}, (X^T X)^{1/2}) \leq \|(X^T X)^{-1/2}\| \|(X^T X)^{1/2}\|,$$

i.e.

$$n \leq (\text{tr}(X^T X)^{-1})^{1/2} (\text{tr}(X^T X))^{1/2},$$

proving the result. [Another proof: Let $W = (\text{tr } X^T X / n) I_n$. Then the eigenvalues of W are majorized by the diagonal elements of $X^T X$ which are in turn majorized by the eigenvalues of $X^T X$ (see Theorem 2.1.3).

Hence $\text{tr}(X^T X)^{-1} \geq \text{tr } W^{-1} = n^2 / \text{tr } X^T X$.

The above result can be generalized to the following:

Theorem 3.2.2. Let C be a positive definite $n \times n$ matrix,

$t = \text{tr } C$, $a > 0$, and

$$b(t, a) = \frac{[(n-2)(at-n^2) + \sqrt{n^2 a^2 t^2 - 2n^2(n^2-2n+2)at + n^4(n-2)^2}]^2}{4n(n-1)a^2} + \frac{t^2}{n}.$$

Suppose that $\text{tr } C^2 \geq b(t, a)$. Then $\text{tr } C^{-1} \geq a$. Moreover, $\text{tr } C^{-1} > a$ if $\text{tr } C^2 > b(t, a)$ and $at \neq (n-2)^2$.

Proof. Let $S = \{(\lambda_1, \lambda_2, \dots, \lambda_n)' : \sum \lambda_i = t, \sum \lambda_i^2 = \text{tr } C^2,$

$\lambda_i \geq 0, i = 1, 2, \dots, n\}$ and f, g be defined by

$$f((\lambda_i)) = \sum_{i=1}^n g(\lambda_i), \text{ for } (\lambda_i) \in S,$$

$$g(x) = \frac{1}{x}, \text{ for } x > 0, \quad g(0) = +\infty.$$

Then $(\lambda_i) \in S$, where the λ_i 's are the eigenvalues of C . So S

is a nonempty compact subset of R^n and f is a continuous function of S into $[0, \infty]$. Thus there exists $(\mu_i) \in S$ such that

$f((\mu_i)) = \min f(S)$. Obviously each $\mu_i > 0$ (otherwise $f((\mu_i)) = +\infty$

which cannot be $\min f(S)$). By Lemmas A2, A3 and A6 of Cheng [4] we

may assume that $\mu_1 = \mu_2 = \dots = \mu_{n-1} \leq \mu_n$. Write $\mu = \mu_1$ and

$\lambda = \mu_n$. Since $(\mu_i) \in S$,

$$\mu = \frac{t - \sqrt{\frac{n}{n-1}} P}{n}, \quad \lambda = \frac{t + \sqrt{\frac{n}{n-1}} P}{n} \quad \text{with } P = \sqrt{\text{tr } C^2 - \frac{t^2}{n}}.$$

$\text{tr } C^{-1} \in f(S)$. So $\text{tr } C^{-1} \geq f((\mu_i))$. It then suffices to prove that

$f((\mu_i)) \geq a$, i.e.

$$\frac{n(n-1)}{t - \sqrt{\frac{n}{n-1}} P} + \frac{n}{t + \sqrt{\frac{n}{n-1}} P} \geq a$$

or

$$n(n-1)aP^2 + [n^2(n-2)\sqrt{n(n-1)} - (n-2)\sqrt{n(n-1)}]at + n^2(n-1)t - (n-1)at^2 \geq 0.$$

Let $h(P)$ be the left side of this inequality. It suffices to prove

that $P \geq P_0$, where P_0 is the positive root of $h(P) = 0$. Using

the quadratic formula, we obtain

$$P_0 = \frac{(n-2)(at - n^2) + \sqrt{n^2 a^2 t^2 - 2n^2(n^2 - 2n + 2)at + n^4(n-2)^2}}{2a\sqrt{n(n-1)}}.$$

Thus $P_0^2 = b(t, a) - \frac{t^2}{n}$. So

$$P = \sqrt{\text{tr } C^2 - \frac{t^2}{n}} \geq \sqrt{b(t, a) - \frac{t^2}{n}} = P_0.$$

Suppose now that $\text{tr } C^2 > b(t, a)$ and $at \neq (n-2)^2$. Since $b(t, \cdot)$

is continuous on $\left(\frac{(n-2)^2}{t}, \frac{n^2}{t}\right)^C$, $\text{tr } C^2 > b(t, a')$ for some

$a' > a$. By the result just proved, $\text{tr } C^{-1} \geq a' > a$.

Instead of $\text{tr } C^{-1}$, the above useful result can be stated and proved for $|C^{-1}|$, etc.

We remark that if we set $a = n^2/t$ in Theorem 3.2.2, then $b(t,a) = 0$, and Theorem 3.2.2 is reduced to Theorem 3.2.1.

We will make much use of Theorem 3.2.2. Consequently we will need to refer to Table 3.2.1 below. In this table, $a = \text{tr}(X_2^T X_2)^{-1}$, where X_2 is as in (3.1.3), $t = \text{tr}(X^T X)$ where $X \in \mathcal{O}(N,n)$, k is the number of zero entries in X and $b(t,a)$ is as in Theorem 3.2.2. We note that the absence of a value for $\{b(t,a)\} + 1$ in Table 3.2.1 indicates that $b(t,a)$ is undefined (this is easily seen to be equivalent to $a < n^2$) .

To illustrate the use of Table 3.2.1 and Theorem 3.2.2, let $X \in \mathcal{O}(6,6)$ have two zero entries and satisfy $\text{tr}(X^T X)^2 \geq 235$. Then Theorem 3.2.2 and line 3 of Table 3.2.1 show that $\text{tr}(X^T X)^{-1} > 1.2$.

Table 3.2.1

	n	k	a	$[b(\tau, a)] + 1$	$[at] + 1$
N = 6	6	1	1.2	261	43
	6	2	1.2	235	41
	6	3	1.2	209	40
	6	4	1.2	187	39
	6	5	1.2	167	38
N = 10	3	1	.30833		9
	3	2			9
	3	3			9
	4	1	.41667	387	17
	4	2			16
	4	3			16
	5	1	.52976	504	26
	5	2		471	26
	5	3			25
	6	1	.64286	625	38
	6	2		588	38
	6	3		554	37

We are now ready to prove:

Theorem 3.2.3. Let X_2 be as in (3.2.1). Then X_2 is A-optimal in $\mathcal{D}(N, n)$ for $n \leq 6$ and N arbitrary.

Proof. Part A: $N = 6, n \leq 6$.

We shall prove only the case where $n = 6$. (The case $n < 6$ is similar and easier. When $n \leq 2$, the desired result is obvious by Theorem 2.2.2.

A simple calculation gives $\text{tr}(X_2^T X_2)^{-1} = \frac{6}{5} = 1.2$. Let $X \in \mathcal{D}(6, 6)$. We must show that $\text{tr}(X^T X)^{-1} \geq 1.2$. Let $t = \text{tr}(X^T X)$, $a = 1.2$, $C = X^T X$. Suppose that some column, say $\bar{X}_1$, of X has two or more zero entries. Then by the majorization of Ky Fan (Theorem 2.1.2),

$$\text{tr} C^{-1} \geq \text{tr}(\bar{X}_1^T \bar{X}_1)^{-1} + \text{tr}(Z^T Z)^{-1}, \quad Z \in \mathcal{D}((6, 5)).$$

By this theorem with $n = 5$,

$$\begin{aligned} \text{tr}(Z^T Z)^{-1} &\geq \text{tr} \begin{bmatrix} 6 & 2 & 2 & & \\ 2 & 6 & 2 & & 0 \\ 2 & 2 & 6 & & \\ & 0 & & 6 & 2 \\ & & & 2 & 6 \end{bmatrix}^{-1} \\ &= \frac{2}{6-2} + \frac{1}{6+2(2)} + \frac{1}{6-2} + \frac{1}{6+2} = \frac{39}{40} \end{aligned}$$

whence

$$\text{tr} C^{-1} \geq \frac{1}{4} + \frac{39}{40} = \frac{49}{40} > 1.2.$$

So we may assume that no column of X has more than one zero entry.

Let $Z(X)$ denote the number of zero entries in X .

Case 1. $Z(X) = 0$. By Theorem 3.1.1, $\text{tr} C^{-1} \geq 1.2$.

Case 2. $Z(X) = 1$. We may assume that the zero entry is in the first column of X . By rearranging columns $\bar{X}_2, \bar{X}_3, \dots, \bar{X}_6$ of

$[\bar{X}_2, \bar{X}_3, \dots, \bar{X}_6]$, we may by Ehlich's result (Lemma 2.3.1) assume that

$$[\bar{X}_2, \bar{X}_3, \dots, \bar{X}_6]^T [\bar{X}_2, \bar{X}_3, \dots, \bar{X}_6] = \begin{bmatrix} A & B \\ B^T & D \end{bmatrix},$$

where $A \in M_{s \times s}$, $s \leq 2$, $D \in M_{(n-s) \times (n-s)}$, $a_{ij} \equiv 2 \pmod{4}$,

$d_{ij} \equiv 2 \pmod{4}$, $b_{ij} \equiv 0 \pmod{4}$. Since $\bar{X}_1$ has a zero entry

$|c_{1j}| \geq 1$. Thus

$$\begin{aligned} \text{tr} C^2 &\geq 10 \times 1 + 5^2 + \text{tr} \begin{bmatrix} A & B \\ B^T & D \end{bmatrix}^2 \\ &\geq 35 + 5 \times 6^2 + 4s(s-1) + 4(5-s)(4-s) \\ &= 8\left(s - \frac{5}{2}\right)^2 + 245. \end{aligned}$$

We shall use Theorem 3.2.2 and its notations. Now $n = 6$ and

* $b(t, a) = 260.8$ or more precisely $b(t, a) < 261$. By Theorem 3.2.2, we may assume that $\text{tr} C^2 \leq 260$. If $s < 2$ then $\text{tr} C^2 \geq 8\left(\frac{3}{2}\right)^2 + 245 = 263$.

So let $s = 2$. By multiplying the columns of X by -1 , we may assume that all $a_{ij} = 2 = d_{ij}$. (Recall that rearranging the columns of X or multiplying the columns of X by -1 will not affect the eigenvalues of $X^T X$ and for brevity, will be referred to as 'some basic operations' on the columns of X). Since $s = 2$,

$$(3.2.1) \quad \text{tr} C^2 \geq 8\left(2 - \frac{5}{2}\right)^2 + 245 = 247.$$

If some $b_{ij} \neq 0$, then $b_{ij}^2 = 4^2$ and

$$\text{tr} C^2 \geq 247 + 2 \cdot 4^2 = 279 > 260.$$

So we may assume that $b_{ij} = 0$ for all i, j . Thus

$$C = \begin{bmatrix} 5 & a_1 & a_2 & a_3 & a_4 & a_5 \\ a_1 & 6 & c_1 & 0 & 0 & 0 \\ -a_2 & c_1 & 6 & 0 & 0 & 0 \\ a_3 & 0 & 0 & 6 & 2 & 2 \\ a_4 & 0 & 0 & 2 & 6 & c_2 \\ a_5 & 0 & 0 & 2 & c_2 & 6 \end{bmatrix}$$

where $|c_i| = 2$, $|a_i| \geq 1$. If one of $|a_i| \neq 1$ then

$$2 a_i^2 \geq 2 \times 3^2 = 2(8+1) \text{ and therefore by (3.2.1),}$$

$$\text{tr } C^2 \geq 247 + 2 \times 8 = 263 > 260. \text{ So we may assume that all } |a_i| = 1.$$

(Note now that $\text{tr } C^2 = 247$, i.e. $\text{tr } C^2$ is minimum and this is the only case which cannot be taken care of by Theorem 3.2.2 above). By multiplying $\bar{X}_2, \bar{X}_3$ by -1 , if necessary, we may further assume that

$$a_1 = a_2 = 1. \text{ Let}$$

$$A_1 = \begin{bmatrix} 5 & 1 & 1 \\ 1 & 6 & c_1 \\ 1 & c_1 & 6 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 6 & 2 & 2 \\ 2 & 6 & c_2 \\ 2 & c_2 & 6 \end{bmatrix}, \quad B^T = \begin{bmatrix} a_3 & 0 & 0 \\ a_4 & 0 & 0 \\ a_5 & 0 & 0 \end{bmatrix}.$$

For symmetric $(a_{ij}) \in M_{3 \times 3}$ such that $(a_{ij})^{-1}$ exists,

$$(3.2.2) \quad \text{tr}(a_{ij})^{-1} = \frac{\sum_{i < j} a_{ii} a_{jj} - \sum_{i < j} a_{ij}^2}{a_{11} a_{22} a_{33} + 2a_{12} a_{13} a_{23} - a_{11} a_{23}^2 - a_{22} a_{13}^2 - a_{33} a_{12}^2}.$$

$$\text{Thus } \text{tr } A_1^{-1} = \frac{45}{74 + c_1}, \quad \text{tr } A_2^{-1} = \frac{12}{8 + c_2}. \text{ Suppose that } c_2 = -2.$$

$$\text{tr } A_1^{-1} + \text{tr } A_2^{-1} \geq \frac{45}{76} + \frac{12}{16} = \frac{51}{38} > 1.2.$$

By the (above) majorization theorem of Ky Fan, $X^T X$ majorizes

$\text{diag}(A_1, A_2)$. So

$$\text{tr } C^{-1} \geq \text{tr}(\text{diag}(A_1, A_2))^{-1} = \text{tr } A_1^{-1} + \text{tr } A_2^{-1} > 1.2.$$

Thus we may assume $c_2 = 2$. Suppose that $c_1 = -2$. Then

$$\text{tr } A_1^{-1} + \text{tr } A_2^{-1} = \frac{45}{72} + \frac{12}{20} = \frac{49}{40} > 1.2 .$$

By the majorization theorem of Ky Fan $\text{tr } C^{-1} > 1.2$. So we may

assume $c_1 = 2$, i.e.

$$C = \begin{bmatrix} 5 & 1 & 1 & a_3 & a_4 & a_5 \\ 1 & 6 & 2 & 0 & 0 & 0 \\ 1 & 2 & 6 & 0 & 0 & 0 \\ a_3 & 0 & 0 & 6 & 2 & 2 \\ a_4 & 0 & 0 & 2 & 6 & 2 \\ a_5 & 0 & 0 & 2 & 2 & 6 \end{bmatrix} , \quad a_i = \pm 1 .$$

Applying some basic operations on columns of X , we may assume that

$(a_3, a_4, a_5) = (1, 1, -1)$ or $(1, 1, 1)$. Let $f(a_1, a_2, a_3) = \text{tr } C^{-1}$.

Note that for $C = \begin{bmatrix} A_1 & B \\ B^T & A_2 \end{bmatrix}$, if A_1^{-1} and A_2^{-1} exists then

$$(3.2.3) \quad \text{tr } C^{-1} = \text{tr } A_1^{-1} + \text{tr } E D^{-1} E^T + \text{tr } D^{-1} ,$$

where $D = A_2 - B^T A_1^{-1} B$ and $E = A_1^{-1} B$. Thus (3.2.3) gives

$$f(1, 1, -1) = 1.2716 > 1.2 \quad \text{and} \quad f(1, 1, 1) = 1.213 > 1.2 .$$

Case 3. $Z(X) = 2$. Let V be the matrix obtained from X

be replacing the two zeros in X by 1. By the (above) result of

Ehlich, we may assume that

$$(3.2.4) \quad V^T V = \begin{bmatrix} P & Q \\ Q^T & R \end{bmatrix} ,$$

where $P \in M_{s \times s}$, $R \in M_{(n-s) \times (n-s)}$, $p_{ij} \equiv 2 \pmod{4}$, $r_{ij} \equiv 2 \pmod{4}$,

$q_{ij} \equiv 0 \pmod{4}$. Now $t = 34$, $b(t,a) = 234.147$.

Case 3.0. $s = 0$. Applying some basic operations on columns of X , we may assume that

$$(3.2.5) \quad C = \begin{bmatrix} 5 & * & & * \\ a & 5 & & \\ & & 6 & * \\ e_{ij} & & 6 & \\ & & & 6 \\ & & d_{ij} & 6 \end{bmatrix} ,$$

where $a \equiv 0 \pmod{4}$, $|e_{ij}| \geq 1$, $|d_{ij}| \geq 2$. Thus

$$\text{tr } C^2 \geq 2 \times 5^2 + 4 \times 6^2 + 12 \times 2^2 = 258 > 234 .$$

By Theorem 3.2.2, $\text{tr } C^{-1} \geq 1.2$.

Case 3.1. $s = 1$.

Case 3.1.1. Suppose that the two zero entries of X lie on the same row of X (and do not lie on column X_1) . Applying some basic operations on the columns of X , we may assume that

$$(3.2.6) \quad C = \begin{bmatrix} 6 & & * & & * \\ & 5 & & * & \\ b_{ij} & c_{ij} & 5 & & * \\ & & 6 & & * \\ q_{ij} & e_{ij} & & d_{ij} & 6 \\ & & & & 6 \end{bmatrix}$$

where $|b_{ij}| \geq 1$, $|c_{ij}| \geq 1$, $|e_{ij}| \geq 1$, $q_{ij} \equiv 0 \pmod{4}$, $|d_{ij}| \geq 2$.

Therefore,

$$\text{tr } C^2 \geq 4 \times 6^2 + 2 \times 5^2 + 18 \times 1 + 6 \times 2^2 = 236 > 234 .$$

By Theorem 3.2.2, $\text{tr } C^{-1} \geq 1.2$.

Case 3.1.2. Suppose that the two zero entries of X lie in different rows of X (and do not lie on column $\bar{X}_1$). Repeating the argument in Case 3.1.1, we obtain (3.2.6), except all $|c_{ij}| \geq 0$ and therefore $\text{tr } C^2 \geq 234$. Theorem 3.2.2 allows us to assume $\text{tr } C^2 = 234$, the minimum $\text{tr } C^2$ in this setting. By applying some basic operations on the columns of X ,

$$(3.2.7) \quad C = \left[\begin{array}{ccc|c} & A_1 & & * \\ \hline 0 & d_1 & d_4 & \\ 0 & d_2 & d_5 & A_2 \\ 0 & d_3 & d_6 & \end{array} \right],$$

where

$$A_1 = \begin{bmatrix} 6 & 1 & 1 \\ 1 & 5 & 0 \\ 1 & 0 & 5 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 6 & 2 & 2 \\ 2 & 6 & c \\ 2 & c & 6 \end{bmatrix}, \quad c = \pm 2, \quad |d_i| = 1.$$

By (3.2.2), $\text{tr } A_1^{-1} = \frac{83}{140}$, $\text{tr } A_2^{-1} = \frac{12}{18+c}$. If $c = -2$, then

$$\text{tr } A_1^{-1} + \text{tr } A_2^{-1} = \frac{83}{140} + \frac{3}{4} = \frac{47}{35} > 1.2$$

and therefore Ky Fan's majorization theorem $\text{tr } C^{-1} > 1.2$. So we may assume that $c = 2$. Thus

$$(3.2.8) \quad C = \left[\begin{array}{cccccc} 6 & 1 & 1 & 0 & 0 & 0 \\ 1 & 5 & 0 & d_1 & d_2 & d_3 \\ 1 & 0 & 5 & d_4 & d_5 & d_6 \\ 0 & d_1 & d_4 & 6 & 2 & 2 \\ 0 & d_2 & d_5 & 2 & 6 & 2 \\ 0 & d_3 & d_6 & 2 & 2 & 6 \end{array} \right], \quad |d_i| = 1.$$

Suppose that d_4, d_5, d_6 are not all equal. Let

$$A_1 = \begin{bmatrix} 6 & 1 \\ 1 & 5 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 5 & d_4 & d_5 & d_6 \\ d_4 & 6 & 2 & 2 \\ d_5 & 2 & 6 & 2 \\ d_6 & 2 & 2 & 6 \end{bmatrix}$$

Up to similarity (and symmetry) of C we may assume $d_4 = d_5 = 1, d_6 = -1$.

By Ky Fan's majorization theorem

$$\text{tr } C^{-1} \geq \text{tr } A_1^{-1} + \text{tr } A_2^{-1} = \frac{11}{31} + \frac{75}{86} > 1.2.$$

So we may assume that $d_4 = d_5 = d_6$. Let

$$A_1 = [6], \quad A_2 = \begin{bmatrix} 5 & 0 & d_1 & d_2 & d_3 \\ 0 & 5 & d_4 & d_5 & d_6 \\ d_1 & d_4 & 6 & 2 & 2 \\ d_2 & d_5 & 2 & 6 & 2 \\ d_3 & d_6 & 2 & 2 & 6 \end{bmatrix}$$

Up to similarity of C , we may assume

$$A_2 = \begin{bmatrix} 5 & 0 & 1 & 1 & s \\ 0 & 5 & 1 & 1 & 1 \\ 1 & 1 & 6 & 2 & 2 \\ 1 & 1 & 2 & 6 & 2 \\ s & 1 & 2 & 2 & 6 \end{bmatrix}, \quad s = \pm 1.$$

Then

$$\text{tr } C^{-1} \geq \text{tr } A_1^{-1} + \text{tr } A_2^{-1} = \begin{cases} 797/660 & \text{if } s = 1 \\ 3050/2424 & \text{if } s = -1 \end{cases} > 1.2.$$

Case 3.1.3. Suppose that one of the zero entries of X lies in $\bar{X}_1$. Applying some basic operations on the columns of X we may assume (3.2.5) and carry on arguments from there.

Case 3.2. $s = 2$.

Case 3.2.1. Suppose that both zero entries of X lie on $[\bar{X}_3, \bar{X}_4, \bar{X}_5, \bar{X}_6]$. Applying some basic operations on the columns of X , we may assume that

$$(3.2.9) \quad C = \left[\begin{array}{cc|cc|cc} 6 & 2 & & * & & * \\ 2 & 6 & & & & \\ \hline & & 5 & h & & * \\ & b_{ij} & h & 5 & & \\ \hline & & & & 6 & 2 \\ q_{ij} & e_{ij} & & & 2 & 6 \end{array} \right]$$

where $|b_{ij}| \geq 1$, $|e_{ij}| \geq 1$, $|q_{ij}| \geq 0$. Now,

$$\text{tr } C^2 \geq 2 \times 5^2 + 4 \times 6^2 + 4 \times 2^2 + 16 \times 1 = 226.$$

If any one of $|b_{ij}|$, $|e_{ij}|$ is not equal to 1, then

$$\text{tr } C^2 \geq 226 + 2(3^2 - 1) = 242 > 234$$

whence by Theorem 3.2.2, $\text{tr } C^{-1} \geq 1.2$. Similarly, if any one of the q_{ij} is not zero then

$$\text{tr } C^2 \geq 226 + 2(4^2 - 1) = 256 > 234, \text{ and}$$

Theorem 3.2.2 again implies $\text{tr } C^{-1} \geq 1.2$. So we may assume that all

$|b_{ij}| = 1$, $|e_{ij}| = 1$, $q_{ij} = 0$. Thus applying some basic

operations on the columns of X , we may write (3.2.9) as

$$(3.2.10) \quad C = \left[\begin{array}{ccc|ccc} & & & r_1 & 0 & 0 \\ & & & r_2 & 0 & 0 \\ & A_1 & & h & r_3 & r_4 \\ \hline r_1 & r_2 & h & & & \\ 0 & 0 & r_3 & A_2 & & \\ 0 & 0 & r_4 & & & \end{array} \right],$$

where

$$A_1 = \begin{bmatrix} 6 & c_1 & 1 \\ c_1 & 6 & 1 \\ 1 & 1 & 5 \end{bmatrix} \quad A_2 = \begin{bmatrix} 5 & 1 & 1 \\ 1 & 6 & c_2 \\ 1 & c_2 & 6 \end{bmatrix},$$

and $c_i = \pm 2$, $r_i = \pm 1$. If $|h| \geq 3$ then

$$\text{tr } C^2 \geq 226 + 2 \times 3^2 = 244 > 234;$$

so Theorem 3.2.2 shows $\text{tr } C^{-1} \geq 1.2$. If $|h| = 2$ then by Ky Fan's majorization theorem and (3.2.9)

$$\begin{aligned} \text{tr } C^{-1} &\geq \text{tr} \begin{bmatrix} 6 & 2 \\ 2 & 6 \end{bmatrix}^{-1} + \text{tr} \begin{bmatrix} 5 & h \\ h & 5 \end{bmatrix}^{-1} + \text{tr} \begin{bmatrix} 6 & 2 \\ 2 & 6 \end{bmatrix}^{-1} \\ &= \frac{12}{32} + \frac{10}{25-h^2} + \frac{12}{32} \\ &= \frac{103}{84} > 1.2. \end{aligned}$$

So we may assume $h = 0$ or ± 1 . Now as before, $\text{tr } A_1^{-1} = \frac{45}{74+c}$,

$\text{tr } A_2^{-1} = \frac{45}{74+c_2}$. If any one of the c_i 's is equal to -2 we have

$$\text{tr } C^{-1} \geq \text{tr } A_1^{-1} + \text{tr } A_2^{-1} \geq \frac{45}{74-2} + \frac{45}{74+2} = \frac{185}{152} > 1.2.$$

Thus we may assume $c_1 = c_2 = 2$ and

$$(3.2.11) \quad C = \begin{bmatrix} 6 & 2 & 1 & r_1 & 0 & 0 \\ 2 & 6 & 1 & r_2 & 0 & 0 \\ 1 & 1 & 5 & h & r_3 & r_4 \\ r_1 & r_2 & h & 5 & 1 & 1 \\ 0 & 0 & r_3 & 1 & 6 & 2 \\ 0 & 0 & r_4 & 1 & 2 & 6 \end{bmatrix}, \quad |r_i| = 1, \quad |h| = 0 \text{ or } 1.$$

If $r_3 = 1 = -r_4$, then

$$\operatorname{tr} C^{-1} \geq \operatorname{tr} \begin{bmatrix} 6 & 2 \\ 2 & 6 \end{bmatrix}^{-1} + \operatorname{tr} \begin{bmatrix} 5 & h & 1 & -1 \\ h & 5 & 1 & 1 \\ 1 & 1 & 6 & 2 \\ -1 & 1 & 2 & 6 \end{bmatrix}^{-1} = \begin{cases} 1.217 & \text{if } h = 0 \\ 1.239 & \text{if } h = \pm 1. \end{cases}$$

Thus $\operatorname{tr} C^{-1} > 1.2$. If $r_1 = -r_2 = 1$, a similar argument also gives

$\operatorname{tr} C^{-1} > 1.2$. Thus we may assume $r_1 = r_2 = \pm 1$ and $r_3 = r_4 = \pm 1$.

By multiplying, if necessary, the first three rows and first three columns of C by -1 , we may assume $r_1 = r_2 = 1$. Suppose that

$r_3 = r_4 = 1$. Then we have

$$(3.2.12) \quad C = \begin{bmatrix} B_1 & B_2 & B_3 \\ B_2 & B_4 & B_5 \\ B_3 & B_5 & B_6 \end{bmatrix},$$

where $B_1 = \begin{bmatrix} 6 & 2 \\ 2 & 6 \end{bmatrix}$, $B_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $B_3 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $B_4 = \begin{bmatrix} 5 & h \\ h & 5 \end{bmatrix}$,

$h = 0, \pm 1$, $B_5 = B_2$, $B_6 = B_1$. The subring S of $M_{2 \times 2}$ generated

by I_2 and the B_i 's is commutative. So, one can easily find the

characteristic polynomial ϕ of C over S :

$$(3.2.13) \quad \begin{aligned} \phi(\lambda) &= (\lambda - B_1)[(\lambda - B_4)(\lambda - B_1) - B_5^2] + B_2[-B_2(\lambda - B_6)] \\ &= (\lambda - B_1)[\lambda^2 - (B_1 + B_4)\lambda + B_1B_4 - 2B_2^2]. \end{aligned}$$

So one eigenvalue λ_1 of ϕ is B_1 . The zeros of the characteristic

polynomial of B_1 (over the real field) give two eigenvalues λ_1, λ_2

for C : $\lambda_1 = 6 - 2 = 4$, $\lambda_2 = 6 + 2 = 8$. By (3.2.13),

$$\begin{aligned}\phi(\lambda)/(\lambda - B_1) &= \lambda^2 - \begin{bmatrix} 11 & 2+h \\ 2+h & 11 \end{bmatrix} \lambda + \begin{bmatrix} 26+2h & 6+6h \\ 6+6h & 26+2h \end{bmatrix} \\ &= \begin{bmatrix} \lambda^2 - 11\lambda + 26 + 2h & -(2+h)\lambda + 6 + 6h \\ -(2+h)\lambda + 6 + 6h & \lambda^2 - 11\lambda + 26 + 2h \end{bmatrix}\end{aligned}$$

Taking the determinant of $\phi(\lambda)/(\lambda - B_1)$, we obtain the polynomial f

whose zeros $\lambda_3, \lambda_4, \lambda_5, \lambda_6$ are the remaining eigenvalues of C :

for real λ ,

$$\begin{aligned}f(\lambda) &= \det[\phi(\lambda I)/(\lambda I - B_1)] \\ &= \lambda^4 + a\lambda^3 + b\lambda^2 + c\lambda + d,\end{aligned}$$

where $c = 4(3h^2 - 2h - 137)$, $d = 32(4 + h)(5 - h)$. So

$$\frac{1}{\lambda_3} + \frac{1}{\lambda_4} + \frac{1}{\lambda_5} + \frac{1}{\lambda_6} = -\frac{c}{d} = \frac{137 + 2h - 3h^2}{8(4+h)(5-h)}.$$

Thus,

$$\begin{aligned}\text{tr } C^{-1} &= \frac{1}{4} + \frac{1}{8} + \frac{137 + 2h - 3h^2}{8(4+h)(5-h)} \\ &= \begin{cases} 197/160 & \text{if } h = 0 \\ 49/40 & \text{if } h = 1 \\ 31/24 & \text{if } h = -1 \end{cases} \\ &> 1.2.\end{aligned}$$

So we may assume that $r_3 = r_4 = -1$. By multiplying $\bar{x}_5$ by -1 , we obtain (3.2.13) with

$$\begin{aligned}B_1 &= \begin{bmatrix} 6 & 2 \\ 2 & 6 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad B_4 = \begin{bmatrix} 5 & h \\ h & 5 \end{bmatrix}, \\ h &= 0, \pm 1, \quad B_5 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad B_6 = \begin{bmatrix} 6 & -2 \\ -2 & 6 \end{bmatrix}.\end{aligned}$$

The B_i 's are

commutative with one another. So the preceding method is applicable!

We obtain $\text{tr } C^{-1} = 1.222$ if $h = 0$; $\text{tr } C^{-1} = 1.2468$ if $h = \pm 1$.

Case 3.3. $s = 3$.

Case 3.3.1. Suppose that both zero entries of X lie in $[\bar{X}_1, \bar{X}_2, \bar{X}_3]$. In terms of arguments and consequences, this is the same as Case 3.1.2.

Case 3.3.2. Suppose that both zero entries of X lie in $[\bar{X}_4, \bar{X}_5, \bar{X}_6]$. By interchanging columns $\bar{X}_1, \bar{X}_2, \bar{X}_3$ with $\bar{X}_4, \bar{X}_5, \bar{X}_6$, this is Case 3.3.1.

Case 3.3.3. Suppose that exactly one zero entry of X lies in $[\bar{X}_1, \bar{X}_2, \bar{X}_3]$. Applying some basic operations on the columns of X , we obtain (3.2.9) and carry on arguments from there.

Case 4. $Z(X) = 3$. Here $t = 33$ and $b(t, a) < 210$. We may assume that each of the first three columns of X contains one zero. Then as before we may write

$$C = \left[\begin{array}{ccc|ccc} 5 & & * & & & \\ a_1 & 5 & & & & * \\ a_2 & a_3 & 5 & & & \\ \hline b_1 & b_4 & b_7 & & & \\ b_2 & b_5 & b_8 & & & V \\ b_3 & b_6 & b_9 & & & \end{array} \right],$$

where $|b_i| \geq 1$ and V is (3.2.4) with $n = 3$. Thus there are at

least two v_{ij} 's with $i \neq j$, $|v_{ij}| = 2$. Therefore

$$\text{tr } C^2 \geq 3 \times 5^2 + 3 \times 6^2 + 2 \times 9 + 2 \times 2^2 = 209.$$

By Theorem 3.2.2, we may assume $\text{tr } C^2 = 209$. In this case we may write

$$C = \left[\begin{array}{ccc|cc|c} 5 & 0 & 0 & b_1 & b_2 & b_3 \\ 0 & 5 & 0 & b_4 & b_5 & b_6 \\ 0 & 0 & 5 & b_7 & b_8 & b_9 \\ \hline b_1 & b_4 & b_7 & 6 & 2 & 0 \\ b_2 & b_5 & b_8 & 2 & 6 & 0 \\ \hline b_3 & b_6 & b_9 & 0 & 0 & 6 \end{array} \right], \quad |b_i| = 1.$$

Let

$$P = \begin{bmatrix} A_1 & B \\ B^T & A_2 \end{bmatrix}, \quad Q = [6],$$

where

$$A_1 = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} b_1 & b_2 \\ b_4 & b_5 \\ b_7 & b_8 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 6 & 2 \\ 2 & 0 \end{bmatrix}.$$

By (3.2.3)

$$\text{tr } P^{-1} = \frac{3}{5} + \frac{2(81+a^2-10a)}{5(629+20a-a)^2} + \frac{270}{629+20a-a^2},$$

$$\text{where } a = b_1 b_2 + b_4 b_5 + b_7 b_8.$$

Since $|a| = 1$, we have

$$\begin{aligned} \text{tr } P^{-1} &= \frac{3}{5} + \frac{2(82-10a)}{5(628+20a)} + \frac{270}{628+20a} \\ &\geq \frac{3}{5} + \frac{2(72)}{5(648)} + \frac{270}{648}. \end{aligned}$$

By Ky Fan's majorization theorem,

$$\text{tr } C^{-1} \geq \text{tr } P^{-1} + \text{tr } Q^{-1} \geq \frac{3}{5} + \frac{144}{5(648)} + \frac{270}{648} + \frac{1}{6} = 1.2277 > 1.2.$$

Case 5. $Z(X) = 4$. Here $t = 32$, $b(t, a) < 188$.

We may assume that each of $\bar{X}_1, \bar{X}_2, \bar{X}_3$ and $\bar{X}_4$ contains a zero. So

$$C = \left[\begin{array}{ccc|cc} 5 & & * & & \\ & 5 & & & * \\ b_{ij} & & 5 & & \\ & & & 5 & \\ \hline & c_{ij} & & 6 & c \\ & & & c & 6 \end{array} \right],$$

where $|c_{ij}| \geq 1$, $|c| \geq 0$. Thus

$$\text{tr } C^2 \geq 4 \times 5^2 + 2 \times 6^2 + 16 \times 1 = 188 > b(t, a).$$

By Theorem 3.2.2, $\text{tr } C^{-1} \geq 1.2$.

Case 6. $Z(X) = 5$. Here $t = 31$, $b(t, a) < 168$.

We may assume that the first five columns of X each contain one zero. So

$$C = \left[\begin{array}{ccc|cc} 5 & & & & \\ & 5 & * & & \\ & & 5 & & * \\ * & & & 5 & \\ & & & & 5 \\ \hline & r_{ij} & & 6 & \end{array} \right]$$

where $|r_{ij}| \geq 1$. Thus

$$\text{tr } C^2 \geq 5 \times 5^2 + 6^2 + 10 \times 1 = 171 > 168.$$

By Theorem 3.2.2, $\text{tr } C^{-1} \geq 1.2$.

Case 7. $Z(X) = 6$. Then by Theorem 3.2.1, $\text{tr } C^{-1} \geq \frac{36}{30} = 1.2$.

This completes part A.

Part B: $N = 10$, $n \leq 6$.

Let $X \in \mathcal{D}(10, n)$. By Theorem 3.1.1, we may assume X has k zero entries $k \geq 1$. From table (3.2.1) the cases

$n = 3, k \geq 1$; $n = 4, k \geq 2$; $n = 5, k \geq 3$; $n = 6, k \geq 4$

follow from Theorem 3.2.1.

Let $n = 4$ and $k = 1$. Then $a = .41666$, $b(t, a) < 387$. By Lemma 2.3.4 of Ehlich we may assume that

$$X^T X = \begin{bmatrix} 9 & a_1 & a_2 & a_3 \\ & 10 & c & \\ & * & 10 & \\ & & & 10 \end{bmatrix}, \quad |a_i| \geq 1, \quad |c| \geq 2.$$

Thus $\text{tr}(X^T X)^2 \geq 9^2 + 3 \times 10^2 + 6 \times 1^2 + 2 \times 2^2 = 395 > b(t, a)$.

Therefore, by Theorem 3.2.2, $\text{tr}(X^T X)^{-1} \geq a$.

Let $n = 5$ and $k = 1$. Then $a = .52976$, $b(t, a) < 504$.

Again by Ehlich's result we may assume that

$$X^T X = \begin{bmatrix} 9 & a_1 & a_2 & a_3 & a_4 \\ & 10 & c_1 & & \\ & & 10 & c_2 & \\ & & & 10 & c_3 \\ & & & & 10 \end{bmatrix}, \quad |a_j| \geq 1; |c_i| \geq 2, \text{ for } \text{at least two values of } i.$$

So $\text{tr}(X^T X) \geq 9^2 + 4 \times 10^2 + 8 \times 1^2 + 4 \times 2^2 = 505 > b(t, a)$. Theorem 3.2.2 then applies.

Let $n = 5$ and $k = 2$. Then $b(t, a) < 471$. By Lemma 2.3.4, we may assume

$$X^T X = \begin{bmatrix} 9 & & & & \\ & \vdots & & & \\ & & a_{ij} & & \\ & & & \vdots & \\ & & & & 10 \\ & & & & & c \\ & & & & & & 10 \\ & & & & & & & 10 \end{bmatrix}, \quad |a_{ij}| \geq 1, \quad |c| \geq 2$$

or

$$X^T X = \begin{bmatrix} 8 & & & & & a_j \\ & 10 & & & & c_1 \\ & & 10 & & & c_2 \\ & & & 10 & & c_3 \\ & & & & 10 & \\ & & & & & 10 \end{bmatrix}, \begin{array}{l} |a_j| > 1; |c_i| > 2, \\ \text{for at least two} \\ \text{values of } i. \end{array}$$

In either case, $\text{tr}(X^T X)^2 \geq 480 > b(t, a)$. Hence by Theorem 3.2.2, $\text{tr}(X^T X)^{-1} \geq a$.

Let $n = 6$. We may assume that the zero entries of X appear in distinct columns. Indeed, suppose that the first column of X has two or more zero entries. Then we may write

$$X^T X = \begin{bmatrix} m & * \\ * & C \end{bmatrix}, \quad m \leq 8,$$

therefore, by Ky Fan's majorization theorem and the result for the case $n = 5$, we have

$$\text{tr}(X^T X)^{-1} \geq \text{tr}[m]^{-1} + \text{tr } C^{-1} \geq \frac{1}{8} + .52976 > .64286 = a.$$

Let $k = 1$. Then $b(t, a) < 625$. Thus we may assume $\text{tr}(X^T X)^2 \leq 624$. Therefore by Ehlich's Lemma we may assume

$$X^T X = \begin{bmatrix} 9 & a_1 & a_2 & a_3 & a_4 & a_5 \\ & 10 & c_1 & & & \\ & & 10 & & & \\ & & & 10 & c_2 & c_3 \\ & & & & 10 & c_4 \\ & & & & & 10 \end{bmatrix}, \quad |a_i| = 1, |c_i| = 2.$$

Up to similarity we may assume $a_1 = a_2 = 1, c_2 = c_3 = 2$. Then by

Ky Fan's theorem we have

$$\begin{aligned} \text{tr}(X^T X)^{-1} &\geq \text{tr} \begin{bmatrix} 9 & 1 & 1 \\ 1 & 10 & c_1 \\ 1 & c_1 & 10 \end{bmatrix}^{-1} + \text{tr} \begin{bmatrix} 10 & 2 & 2 \\ 2 & 10 & c_4 \\ 2 & c_4 & 10 \end{bmatrix}^{-1} \\ &= \frac{137}{422+c_1} + \frac{36}{110+c_2} \geq \frac{137}{424} + \frac{36}{112} = .64453 > a . \end{aligned}$$

Let $k = 2$. Then $b(t,a) < 588$. By Ehlich's result we assume

$$X^T X = \begin{bmatrix} 9 & & & & \\ & 9 & & & \\ & & 10 & & \\ & & & 10 & \\ & & & & 10 \end{bmatrix} \begin{matrix} a_{ij} \\ c_1 \\ c_2 \\ c_3 \\ 10 \end{matrix} , |a_{ij}| \geq 1 ; |c_i| \geq 2 ,$$

for at least two values of i .

$$\text{So } \text{tr}(X^T X)^2 \geq 2 \times 9^2 + 4 \times 10^2 + 16 \times 1^2 + 4 \times 2^2 = 594 > b(t,a) .$$

Let $k = 3$. Then $b(t,a) < 554$. As above we may assume

$$X^T X = \begin{bmatrix} 9 & & & & \\ & 9 & & & \\ & & 10 & & \\ & & & 10 & \\ & & & & 10 \end{bmatrix} \begin{matrix} a_{ij} \\ c \\ 10 \end{matrix} , |a_{ij}| \geq 1 , |c| \geq 2 .$$

$$\text{So } \text{tr}(X^T X)^2 \geq 3 \times 9^2 + 3 \times 10^2 + 18 \times 1^2 + 2 \times 2^2 = 569 > b(t,a) .$$

This completes Part B.

Part C: $N > 10$ ($N \equiv 2 \pmod{4}$) , $n \leq 6$.

For $14 \leq N \leq 42$, one can proceed as in the case $N = 10$.

The proofs are similar and easier. The cases $N \geq 46$ and $n \leq 6$

follow from the following lemma and Theorems 3.2.1 and 3.1.1.

Lemma 3.2.1. Let $N \geq 46$, $n \leq 6$, $X \in \mathcal{L}(N, n)$, $\text{tr}(X^T X) \leq nN - 1$.

$$\text{Then } \text{tr}(X_2^T X_2)^{-1} \leq \frac{n^2}{\text{tr}(X^T X)}.$$

Proof. Let $a = \text{tr}(X_2^T X_2)^{-1}$ and $t = \text{tr}(X^T X)$. Suppose first that n

is even. Then $a = \frac{n-2}{N-2} + \frac{2}{N+n-2}$. Now $at \leq n^2$ provided

$$\left[\frac{n-2}{N-2} + \frac{2}{N+n-2} \right] (nN - 1) \leq n^2. \text{ On simplifying the latter}$$

inequality we obtain that $at \leq n^2$ if $2n^2 - 5n + 4 \leq N$. Since

$\max_{0 \leq n \leq 6} (2n^2 - 5n + 4) = 46$, the result is proven for n even.

Suppose now that n is odd. Then $a = \frac{n-2}{N-2} + \frac{1}{N+n-1} + \frac{1}{N+n-3}$.

On taking $n = 1, 3, 5$ one can easily show that $at \leq n^2$ for

$N \geq 46$.

This completes the proof of Theorem 3.2.3.

The above proof suggests that the desired result for $\mathcal{L}(N, n)$ can be proved with the help of the results for $\mathcal{L}(N, n-1)$.

By the above proof of Theorem 3.2.3 (and Theorem 3.2.2), up to similarity of $X^T X$ there are two A-optimal designs in $\mathcal{L}(6, 6)$.

Let

$$X_2 = \begin{bmatrix} 1 & 1 & -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix}, \quad Z_2 = \begin{bmatrix} 0 & 1 & 1 & -1 & 1 & -1 \\ 1 & 0 & -1 & 1 & 1 & -1 \\ 1 & -1 & 0 & -1 & -1 & -1 \\ 1 & -1 & 1 & 0 & 1 & 1 \\ 1 & 1 & -1 & -1 & 0 & 1 \\ 1 & 1 & 1 & 1 & -1 & 0 \end{bmatrix}.$$

$$\text{Then } X_2^T X_2 = \begin{bmatrix} 4I_3 + 2I_3 & 0 \\ 0 & 4I_3 + 2I_3 \end{bmatrix} \text{ and } Z_2^T Z_2 = 5I_6.$$

So $\text{tr}(X_2^T X_2)^{-1} = \text{tr}(Z_2^T Z_2)^{-1} = 1.2$. Thus both X_2 and Z_2 are

A-optimal in $\mathcal{D}(6,6)$. Therefore, A-optimal designs in $\mathcal{D}(6,6)$, are not unique. Statisticians may prefer the design Z_2 over X_2 because Z_2 is, by Theorem 2.7 of Jacroux, Wong and Masaro [18], E-optimal.

With Theorem 3.2.3, we can prove the E-optimality of Z_2 in an obvious way: Let $W \in \mathcal{D}(6,6)$ and let $\{\mu_i\}$ be the eigenvalues of $W^T W$, where $\mu_1 \leq \mu_2 \leq \dots \leq \mu_6$. Then,

$$(3.2.14) \quad 1.2 \leq \text{tr}(W^T W)^{-1} = \sum_{i=1}^6 \frac{1}{\mu_i} \leq \frac{6}{\mu_1}.$$

So $\mu_1 \leq 5$. Z_2 shows that it is possible that $\mu_1 = 5$. Hence Z_2 is E-optimal. In weighing designs, using Z_2 means that every object can take a single break (i.e. not to be weighed) at different weighings. If $3 \leq n < 6$ and $Y \in \mathcal{D}(6,n)$ is formed from any n columns of Z_2 , then from Theorem 2.7 of Jacroux, Wong and Masaro [18], Y is uniquely E-optimal. Further any $X_2 \in \mathcal{D}(6,n)$ satisfying (3.1.1) is uniquely A-optimal. This shows that A-optimality does not, in general, imply E-optimality and vice-versa. Also, Theorem 3.2.3 and therefore conjecture (C1) is not an obvious statement to make. To see this, we note that Cheng [5] proved that X_2 in (3.1.1) is E-optimal over $\mathcal{D}(N,n)$ ($N \equiv 2 \pmod{4}$). But it turns out that this, X_2 is not E-optimal over $\mathcal{D}(N,n)$. So, it is not automatic that an optimal X over $\mathcal{D}'(N,n)$ will be optimal over $\mathcal{D}(N,n)$. However, D-optimal X over $\mathcal{D}'(N,n)$ is automatically D-optimal over $\mathcal{D}(N,n)$. Galil and Kiefer [13] remark that X_2 in (3.1.1) is D-optimal and ask

if it has any other optimality properties. Our conjecture says that X_2 is also A-optimal in $\mathcal{D}(N, n)$!

We will see in Section 3.3 that for each n there exists an $N_0(n)$ such that X_2 is A-optimal in $\mathcal{D}(N, n)$ if $N \geq N_0(n)$.

3.3. ϕ_p -optimal Designs in $\mathcal{D}(N, n)$.

Let X_2 and Z_2 be as in Section 1.3 (p.10) Cheng (private communication) noted that when $N = n$ X_2 and Z_2 have the same performance under A-criterion and that X_2 is ϕ_p -better than Z_2 for $0 \leq p < 1$ and Z_2 is ϕ_p -better than X_2 for $p > 1$. When $N > n$ evidence suggests that X_2 is ϕ_p -optimal for $0 \leq p \leq 1$. However, X_2 is not ϕ_p -optimal for all $p > 1$ since it is not E-optimal. In fact for $N=6, n=4$, $\phi_3(X_2^T X_2) = .206 > .2 = \phi_3(Z_2^T Z_2)$. Further Z_2 may not be ϕ_p -optimal for $p > 1$ since, for example, if $N=6, n=4$, $\phi_2(X_2^T X_2) = .197 < .2 = \phi_2(Z_2^T Z_2)$. The above discussion suggests the following conjecture

(C2) For any $N \geq n$, $N \equiv 2 \pmod{4}$ there exists a number $\alpha(N, n) \geq 1$ such that $\alpha(N, N) = 1$ and X_2 is ϕ_p -optimal for $0 \leq p \leq \alpha(N, n)$ and Z_2 is ϕ_p -optimal for $p \geq \alpha(N, n)$.

The following theorem shows that X_2 is ϕ_p -optimal for $0 \leq p \leq 1$ if N is sufficiently larger than n and hence supports conjecture (C2). The case $p = 1$ was found independently by Cheng, and Wong and Masaro; the extension to the cases $0 \leq p \leq 1$ is due to Cheng (see Cheng, Masaro and Wong [7]).

Theorem 3.3.1. For any n , there exists $N_0(n)$ such that if

$N \geq N_0(n)$, then X_2 is ϕ_p -optimal over $\mathcal{B}(N, n)$ for all $0 \leq p \leq 1$.

Proof. In Theorem 3.1.1 it was shown that X_2 is optimal over $\mathcal{B}(N, n)$ with respect to all the type 1 criteria ϕ_f , such that $f(0^+) = +\infty$;

in particular it is ϕ_p -optimal for all $p \geq 0$. So it suffices to

consider matrices C in $C(N, n) \setminus C'(N, n)$. We have to show that there exists an integer N_0 such that if $N \geq N_0$, then for any

$C \in C(N, n) \setminus C'(N, n)$, $\phi_p(C) > \phi_p(X_2^T X_2)$ for all $0 \leq p \leq 1$.

Now if $C \in C(N, n) \setminus C'(N, n)$, then $\text{tr} C \leq Nn - 1$. Since $\text{tr}\{n^{-1}(Nn - 1)I_n\} = Nn - 1$, by Proposition 1' of Kiefer [24], we have

$$(3.3.1) \quad \phi_p(C) \geq \phi_p(n^{-1}(Nn - 1)I_n) \quad \text{for all } p \geq 0.$$

By a direct comparison of $\text{tr}\{n^{-1}(Nn - 1)I_n\}^{-1} = n^2/(Nn - 1)$ with

$$\text{tr}(X_2^T X_2)^{-1} = \begin{cases} (n-2)/(N-2) + 2/(N+n-2) & \text{if } n \text{ is even,} \\ (n-2)/(N-2) + 1/(N+n-1) + 1/(N+n-3) & \text{if } n \text{ is odd,} \end{cases}$$

one can easily show that there exists an integer N_0 such that

$$(3.3.2) \quad N \geq N_0 \Rightarrow \phi_1(n^{-1}(Nn - 1)I_n) > \phi_1(X_2^T X_2).$$

Now let $\mu_1^*, \mu_2^*, \dots, \mu_n^*$ be the eigenvalues of $X_2^T X_2$. Then

$$\phi_p(X_2^T X_2) = \{n^{-1} \sum_{i=1}^n (\mu_i^*)^{-p}\}^{1/p} \quad \text{is an increasing function of } p > 0.$$

Since all the eigenvalues of $n^{-1}(Nn - 1)I_n$ are equal, we have

$$\phi_p\{n^{-1}(Nn - 1)I_n\} = n/(Nn - 1) \quad \text{for all } p > 0. \quad \text{By (3.3.2), if } N \geq N_0$$

and $0 < p \leq 1$, then

$$\phi_p(X_2^T X_2) \leq \phi_1(X_2^T X_2) < \phi_1\{n^{-1}(Nn - 1)I_n\} = \phi_p\{n^{-1}(Nn - 1)I_n\}.$$

Combining this with (3.3.1), we conclude that if $N \geq N_0$, then X_2 is ϕ_p -optimal over $\mathcal{D}(N, n)$ for all $0 < p \leq 1$. The D-optimality ($p = 0$) is obtained by passing to the limit.

We remark that if $N_0^*(n)$ is the smallest integer such that $N \geq N_0^*(n)$ implies X_2 is ϕ_p -optimal over $\mathcal{D}(N, n)$ for $0 \leq p \leq 1$, then, by the above proof, an upper bound for $N_0^*(n)$ can be obtained by comparing $\text{tr}(X_2^T X_2)^{-1}$ with $\text{tr}[n(Nn - 1)^{-1} I_n]$. This upper bound is usually quite large. For example if $n = 6$, Lemma 3.2.1 shows that $N_0^*(n) \leq 46$. So by Theorem 3.3.1, X_2 is ϕ_p -optimal over $\mathcal{D}(N, 6)$ for $0 \leq p \leq 1$ and $N \geq 46$.

Note that, in general, we cannot conclude that if X_2 is A-optimal over $\mathcal{D}(N, n)$ then it is ϕ_p -optimal, $0 \leq p \leq 1$. Indeed, the above proof only shows this to be the case if

$$\text{tr}(X_2^T X_2)^{-1} \leq \text{tr}[n(Nn - 1)^{-1} I_n].$$

Hence if $N_0(n)$ is the smallest

integer n such that $N \geq N_0(n)$ implies X_2 is A-optimal over

$\mathcal{D}(N, n)$, then $N_0(n) \leq N_0^*(n)$. It is open as to whether

$N_0(n) = N_0^*(n)$. Clearly an upper bound for $N_0^*(n)$ is also an upper

bound for $N_0(n)$. As mentioned above this bound is usually very

large. A much better upper bound for $N_0(n)$ can be obtained by the

following method due to Cheng (see Cheng, Masaro and Wong [7]).

For any $X \in \mathcal{D}(N, n)$, suppose there are k columns which contain zero entries. Without loss of generality, we may assume that $X = (Y_1 Y_2)$ where Y_1 is $N \times k$ and consists of all the columns which contain zero entries. By a result of Fan (Theorem 2.1.2) the eigen-

values of $X^T X$ majorize those of $\begin{bmatrix} Y_1^T Y_1 & 0 \\ 0 & Y_2^T Y_2 \end{bmatrix}$. Since all the diagonal

elements of $Y_1^T Y_1$ are $\leq N - 1$, by Proposition 1' of Kiefer [23], we have

$$\text{tr}(X^T X)^{-1} \geq \text{tr} \begin{bmatrix} Y_1^T Y_1 & 0 \\ 0 & Y_2^T Y_2 \end{bmatrix}^{-1} \geq \text{tr} \begin{bmatrix} (N-1)I_k & 0 \\ 0 & Y_2^T Y_2 \end{bmatrix}^{-1}$$

Now all the entries of Y_2 are ± 1 , i.e., $Y_2^T Y_2 \in C'(N, n - k)$. By an argument similar to that employed in the proof of Theorem 3.1.1, we conclude that

$$\text{tr}(X^T X)^{-1} \geq \text{tr} \begin{bmatrix} (N-1)I_k & 0 & 0 \\ 0 & (N-2)I_\ell + 2J_\ell & 0 \\ 0 & 0 & (N-2)I_{n-k-\ell} + 2J_{n-k-\ell} \end{bmatrix}$$

where $\ell = [(n-k)/2]$. Thus an upper bound of $N_0(n)$ is the smallest N such that

$$\begin{aligned} & \text{tr} \begin{bmatrix} (N-2)I_t + 2J_t & 0 \\ 0 & (N-2)I_{n-t} + 2J_{n-t} \end{bmatrix} \\ & \leq \text{tr} \begin{bmatrix} (N-1)I_k & 0 & 0 \\ 0 & (N-2)I_\ell + 2J_\ell & 0 \\ 0 & 0 & (N-2)I_{n-k-\ell} + 2J_{n-k-\ell} \end{bmatrix} \end{aligned}$$

for all k such that $1 \leq k \leq n$, where $t = [n/2]$ and $\ell = [(n-k)/2]$.

The following are some upper bounds of $N_0(n)$ obtained by the above method.

n	4	5	6	7	8	9	10	11	12
an upper bound of $N_0(n)$	10	14	14	18	18	22	22	30	30

These bounds certainly are not sharp. In fact, it has been shown in Theorem 3:2.3 that X_2 is A-optimal over $\mathcal{L}(N,n)$ for all $n \geq 6$ and all $N \geq n$. Thus $N_0(4) = N_0(5) = N_0(6) = 6$. A counter example has not been found such that X_2 is not A-optimal in $\mathcal{L}(N,n)$! But the above table is useful for eliminating a lot of cases that have to be considered in proving the A-optimality of X_2 .

CHAPTER IV

OPTIMAL $N \times n$ WEIGHING DESIGN MATRICES

WITH $N \equiv 3(\text{mod } 4)$.

4.1. General Results on A-optimality in $\mathcal{D}(N, n)$ and $\mathcal{D}'(N, n)$.

Throughout this chapter we assume $N \equiv 3(\text{mod } 4)$. Let $X_3 \in \mathcal{D}(N, n)$ be such that $X_3^T X_3 = (N+1)I_n - J_n$. In this section we will show that if N is sufficiently larger than n , X_3 is A-optimal in $\mathcal{D}(N, n)$ (and in $\mathcal{D}'(N, n)$). The following lemmas are useful in proving this result and in establishing the A-optimality of X_3 in certain cases. Recall that $C(N, n)$ ($C'(N, n)$) is the family of all matrices of the form $X^T X$ where $X \in \mathcal{D}(N, n)$ ($\mathcal{D}'(N, n)$):

Lemma 4.1.1. Let $C \in C'(N, n)$ be such that $|c_{ij}| = c$ for all $i \neq j$. Then C is similar to $(N+c)I_n - cJ_n$ or $(N-c)I_n + cJ_n$.

Proof. Let $C \in C'(N, n)$ be such that $|c_{ij}| = c$ for $i \neq j$. By

Lemma 3.1 of Ehlich [9], we may rearrange the columns of X and multiply appropriate columns by -1 to obtain a new matrix Y such that all the off-diagonal elements of $Y^T Y$ are congruent to $3(\text{mod } 4)$. Then since the absolute value of each off-diagonal element of $Y^T Y$ is also equal to c , we have $Y^T Y = (N+c)I_n - cJ_n$ or $(N-c)I_n + cJ_n$.

Lemma 4.1.2. Let $C \in C'(N, n)$. Then C is similar to a matrix $D \in C'(N, n)$ such that if $|d_{ij}| = 1, 3$, or 5 , then $d_{ij} = -1, 3$ or -5 , respectively.

Proof. Again by Lemma 3.1 of Ehlich [9] C is similar to a matrix D with all $d_{ij} \equiv 3(\text{mod } 4)$. Then D has the desired properties.

Lemma 4.1.3. Let $C \in C'(N, n)$. If $\sum_{i \neq j} c_{ij}^2 \geq n(n-1)N^2/(N-n+2)^2$,

then $\text{tr}\{(N+1)I_n - J_n\}^{-1} \leq \text{tr}C^{-1}$, where $\text{tr}C^{-1}$ is defined to be $+\infty$

if C has no inverse.

Proof. For any $C \in C'(N, n)$, we have $\text{tr}C = nN$ and

$nN^2 \leq \text{tr}C^2 \leq n^2N^2$. For any B such that $nN^2 \leq B \leq n^2N^2$, let

$\mathcal{M}(B, N, n)$ be the set of all the symmetric nonnegative definite $n \times n$

matrices M such that $\text{tr}M = nN$ and $\text{tr}M^2 = B$. Let

$Z = \{(nN - n^2N^2)/(n-1)\}^{1/2}$. Then from Lemmas A2, A3 and A6 of Cheng

[4], $\text{tr}C^{-1} \geq (n-1)/\mu + 1/\lambda$, where $\mu = (nN - Z)/n$ and

$\lambda = \{nN + (n-1)Z\}/n$. It suffices to show $\text{tr}\{(N+1)I_n - J_n\}^{-1} \leq (n-1)/\mu + 1/\lambda$,

i.e., $(n-1)/(N+1) + 1/(N-n+1) \leq (n-1)/\mu + 1/\lambda$. On substituting the expressions for μ and λ , this reduces to

$$(4.1.1) \quad (N - n + 2)Z^2 - n(n-2)Z - n^2N \geq 0.$$

Since $nN/(N - n + 2)$ is the positive root of the equation

$$(N - n + 2)x^2 - n(n-2)x - n^2N = 0, \quad (4.1.1) \text{ holds provided}$$

$$Z \geq nN/(N-n+2).$$

This is equivalent to $B - nN^2 \geq n(n-1)N^2/(N-n+2)^2$. Since for

$C \in C'(N, n)$, $\sum_{i \neq j} c_{ij}^2 = \text{tr}C^2 - nN^2$, the result follows.

Now we are ready to prove

Theorem 4.1.1. For each n , there exists a positive integer $N_0(n)$

such that for all $N \geq N_0(n)$, X_3 is A-optimal in $\mathcal{G}(N, n)$. In

particular one can take $N_0(n) = [b_n] + 1$ where

$$b_n = \max\{(n-2)\left[n^2 - n + 16 + \sqrt{n(n-1)(n^2 - n + 16)}\right]/16, n^2 - 2\} \text{ or, for}$$

$$\text{simplicity, } b_n = \max\{(n-2)(n^2 - n + 16)/8, n^2 - 2\}.$$

Proof. Let $C \in C'(N, n)$. If $\sum_{i \neq j} c_{ij}^2 = n(n-1)$, then by Lemma 4.1.1,

C is similar to $(N+1)I_n - J_n$. Thus we may assume $\sum_{i \neq j} c_{ij} \geq n(n-1)+16$.

It is straightforward to see that if

$N \geq (n-2)[n^2 - n + 16 + \{n(n-1)(n^2 - n + 16)\}^{1/2}]/16$, then

$n(n-1) + 16 > n(n-1)N^2/(N-n+2)^2$; therefore $\sum_{i \neq j} c_{ij}^2 > n(n-1)N^2/(N-n+2)^2$

and by Lemma 4.1.3, $\text{tr} C^{-1} > \text{tr}(X_3^T X_3)^{-1}$.

For $C \in C(N, n) \setminus C'(N, n)$, we have $\text{tr} C \leq nN - 1$. Thus $\text{tr} C^{-1} \geq n^2/\text{tr} C \geq n^2/(nN - 1)$. Comparing the last term with $\text{tr}(X_3^T X_3)^{-1} = (n-1)/(N+1) + 1/(N-n+1)$, we conclude that if $N \geq n^2 - 2$, then $\text{tr} C^{-1} \geq \text{tr}(X_3^T X_3)^{-1}$.

The proof is completed by taking $N_0(n) = [b_n] + 1$ where b_n is as in the statement of the theorem.

Corollary 4.1.1. X_3 is A-optimal in $\mathcal{E}'(N, n)$ if

$N \geq (n-2)[n^2 - n + 16 + \sqrt{n(n-1)(n^2 - n + 16)}]/16$ or $N \geq (n-2)(n^2 - n + 16)/8$.

To illustrate the above theorem, let

$$X = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix},$$

Then $X \in \mathcal{E}(7, 3)$ and $X^T X = \begin{bmatrix} 7 & -1 & -1 \\ -1 & 7 & -1 \\ -1 & -1 & 7 \end{bmatrix}$. Since $N = 7 = N_0(n)$, by

Theorem 4.1.1, X is A-optimal over $\mathcal{E}(7, 3)$. Galil and Kiefer [13]

show that X_3 in Theorem 4.1.1 is D-optimal over $\mathcal{L}(N,n)$ provided $N \geq 2n - 5$. There is no such nice lower bound for A-optimality. In fact, the lower bound $N_0(n)$ in Theorem 4.1.1 is sharp in the sense that X_3 ($X_3^T X_3 = 8I_4 - J_4$) is not A-optimal over $\mathcal{L}(7,4)$. Indeed let

$$Z = \begin{bmatrix} 1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & -1 & -1 & 1 \\ 1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix},$$

Then

$$Z^T Z = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 7 & -1 & -1 \\ 0 & -1 & 7 & -1 \\ 0 & -1 & -1 & 7 \end{bmatrix}$$

$$\text{tr}(X_3^T X_3)^{-1} = \frac{3}{8} + \frac{1}{4} = \frac{5}{8} = .625,$$

$$\text{tr}(Z^T Z)^{-1} = \frac{1}{6} + \frac{2}{8} + \frac{1}{5} = \frac{37}{60} < .625.$$

So Z is A-better than X_3 . We will see in Section 4.2 that Z is A-optimal over $\mathcal{L}(7,4)$.

We now prove:

Theorem 4.1.2. X_3 is A-optimal in $\mathcal{L}'(N,n)$ for $N \geq 7$, $n \leq 5$ and $N \geq 15$, $n \leq 7$.

Proof. We consider only the cases $n \leq 7$, $N = 15$ the rest being similar.

Let $N = 15$, $n = 6$. In this case $n(n-1)N^2 / (N-n+2)^2 \approx 55.8$.

It is easy to see that all matrices in $C'(15,6)$ with $\sum_{i \neq j} c_{ij}^2 \leq 55$.

must have $|c_{ij}| = 1$ for all $i \neq j$ except for at most a pair of off-diagonal elements with $|c_{ij}| = 3$. By Lemmas 4.1.2 and 4.1.3, the only competitor in $C'(15,6)$, up to equivalence, is the matrix

$$C = \begin{bmatrix} 15 & 3 & -1 & -1 & -1 & -1 \\ 3 & 15 & -1 & -1 & -1 & -1 \\ -1 & -1 & 15 & -1 & -1 & -1 \\ -1 & -1 & -1 & 15 & -1 & -1 \\ -1 & -1 & -1 & -1 & 15 & -1 \\ -1 & -1 & -1 & -1 & -1 & 15 \end{bmatrix}$$

Computer calculation gives $\text{tr}C^{-1} = 0.4151$, while $\text{tr}(16I_6 - J_6)^{-1} = 0.4125$.

Thus C_3 is A-optimal in $\mathcal{D}'(15, 6)$.

Since $N \geq (n-2)(n^2 - n + 16)/8$ for $N = 15$ and $n = 2, 3, 4, 5$, it follows from Corollary 4.1.1 that X_3 is also A-optimal in $\mathcal{D}'(15, n)$ for $n = 2, 3, 4, 5$.

In the case $n = 7$, due to our theory, the only competitors up to similarity are the following matrices:

$$C_1 = \begin{bmatrix} 15 & 3 & -1 & -1 & -1 & -1 & -1 \\ 3 & 15 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 15 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 15 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 15 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & 15 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & 15 \end{bmatrix}, \quad \text{tr}C_1^{-1} = .4876$$

$$C_2 = \begin{bmatrix} 15 & 3 & 3 & -1 & -1 & -1 & -1 \\ 3 & 15 & -1 & -1 & -1 & -1 & -1 \\ 3 & -1 & 15 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 15 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 15 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & 15 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & 15 \end{bmatrix}, \quad \text{tr}C_2^{-1} = .4926$$

$$C_3 = \begin{bmatrix} 15 & 3 & -1 & -1 & -1 & -1 & -1 \\ 3 & 15 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 15 & 3 & -1 & -1 & -1 \\ -1 & -1 & 3 & 15 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 15 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & 15 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & 15 \end{bmatrix}, \quad \text{tr}C_3^{-1} = .4896$$

$$C_4 = \begin{bmatrix} 15 & 3 & 3 & 3 & -1 & -1 & -1 \\ 3 & 15 & -1 & -1 & -1 & -1 & -1 \\ 3 & -1 & 15 & -1 & -1 & -1 & -1 \\ 3 & -1 & -1 & 15 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 15 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & 15 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & 15 \end{bmatrix}, \quad \text{tr}C_4^{-1} = .5019$$

$$C_5 = \begin{bmatrix} 15 & 3 & 3 & -1 & -1 & -1 & -1 \\ 3 & 15 & 3 & -1 & -1 & -1 & -1 \\ 3 & 3 & 15 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 15 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 15 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & 15 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & 15 \end{bmatrix}, \quad \text{tr}C_5^{-1} = .4917$$

$$C_6 = \begin{bmatrix} 15 & 3 & 3 & -1 & -1 & -1 & -1 \\ 3 & 15 & -1 & 3 & -1 & -1 & -1 \\ 3 & -1 & 15 & -1 & -1 & -1 & -1 \\ -1 & 3 & -1 & 15 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 15 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & 15 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & 15 \end{bmatrix}, \quad \text{tr}C_6^{-1} = .4974$$

$$C_7 = \begin{bmatrix} 15 & 3 & 3 & -1 & -1 & -1 & -1 \\ 3 & 15 & -1 & -1 & -1 & -1 & -1 \\ 3 & -1 & 15 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 15 & 3 & -1 & -1 \\ -1 & -1 & -1 & 3 & 15 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & 15 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & 15 \end{bmatrix}, \quad \text{tr}C_7^{-1} = .4951$$

$$C_8 = \begin{bmatrix} 15 & 3 & -1 & -1 & -1 & -1 & -1 \\ 3 & 15 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 15 & 3 & -1 & -1 & -1 \\ -1 & -1 & 3 & 15 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 15 & 3 & -1 \\ -1 & -1 & -1 & -1 & 3 & 15 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & 15 \end{bmatrix}, \quad \text{tr}C_8^{-1} = .4922$$

$$C_9 = \begin{bmatrix} 15 & -5 & -1 & -1 & -1 & -1 & -1 \\ -5 & 15 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 15 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 15 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 15 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & 15 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & 15 \end{bmatrix}, \quad \text{tr}C_9^{-1} = .5100$$

Now $\text{tr}(16I_7 - J_7)^{-1} = 0.486110$. Comparing this with $\text{tr}C_i^{-1}$, $i = 1, 2, \dots, 9$; we see that X_3 is optimal in $\mathcal{D}'(15, 7)$. Thus we have shown that X_3 is A-optimal in $\mathcal{D}'(15, n)$ for all $n \leq 7$.

To show the strength of our results, we conclude by giving an example which shows that in general $(N+1)I_n - J_n$ is not A-optimal in $C'(N, n)$. This example points out that a matrix that is A-optimal in $C'(N, n)$ need not be D-optimal in $C'(N, n)$.

Let $N = 11$, $n = 7$. From a theorem of Galil and Kiefer([13],

page 1299), the matrix $12I_7 - J_7$ is the unique D-optimal matrix in $C'(11,7)$. It should be noted that $12I_7 - J_7$ can be realized as $X^T X$ where X is an 11×7 matrix with $x_{ij} = \pm 1$. Indeed, any X such

that

$$H = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & & & \\ \vdots & & & \\ 1 & & & \end{bmatrix} \quad \text{is a Hadamard matrix of order 12 will do.}$$

Such H can be found in Hedayat and Wallis [16]. However, let

$$Z = \begin{bmatrix} 1 & 1 & -1 & -1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ -1 & -1 & 1 & -1 & -1 & -1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 & 1 & -1 & -1 \\ -1 & 1 & 1 & 1 & -1 & -1 & 1 \\ -1 & -1 & 1 & -1 & -1 & 1 & -1 \end{bmatrix};$$

then

$$Z^T Z = \begin{bmatrix} 11 & 3 & -1 & -1 & -1 & -1 & -1 \\ 3 & 11 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 11 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 11 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 11 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & 11 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & 11 \end{bmatrix}.$$

It is well known that if $M = \begin{bmatrix} P & -Q \\ Q^T & R \end{bmatrix}$ is a nonsingular matrix then $M^{-1} = \begin{bmatrix} S & T \\ T^T & U \end{bmatrix}$ where $S = P^{-1} + P^{-1}Q(R - Q^T P^{-1}Q)^{-1} Q^T P^{-1}$, $U = (R - Q^T P^{-1}Q)^{-1}$ and $T = -P^{-1}Q(R - Q^T P^{-1}Q)^{-1}$. Applying this result to $Z^T Z$ we obtain

$$(X^T X)^{-1} = \begin{bmatrix} 441/4312 & -98/4312 & & T \\ -98/4312 & 441/4312 & & \\ & & T^T & \\ & & & (1/12)I_5 + (1/66)J_5 \end{bmatrix}.$$

Then $\text{tr}(Z^T Z)^{-1} = 441/2156 + 65/132 = .69696$. But $\text{tr}[12I_7 - J_7]^{-1} = .7$.

Thus $12I_7 - J_7$ is not A-optimal in $C'(11,7)$. Also the A-optimal matrix in $C'(11,7)$ cannot be D-optimal since $12I_7 - J_7$ is the unique D-optimal matrix in $C'(11,7)$.

In conclusion we remark that the above example includes a general procedure for constructing X_3 with $X_3^T X_3 = (N+1)I_n - J_n$. Using

Proposition 2.4.1 or otherwise choose X_0 in $\mathcal{D}(N+1,n)$ such that $X_0^T X_0 = (N+1)I_n$ and the first row of X_0 consists entirely of 1's.

The appropriate X_3 is obtained by removing the first row from X_0 .

4.2. A-optimal Designs in $\mathcal{D}(N,n)$.

In this section we classify all A-optimal designs in $\mathcal{D}(N,n)$ for $n \leq 6$ and N arbitrary, $N \equiv 3 \pmod{4}$. In this case, the A-optimal matrix will be one of the following three types:

Type I: $X_3^T X_3 = (N+1)I_n - J_n$.

Type II: $Y_3^T Y_3 = \begin{bmatrix} N-1 & & 0 \\ & & \\ 0 & (N+1)I_{n-1} - J_{n-1} \end{bmatrix}$

$$\text{Type III: } Z_3^T Z_3 = \begin{bmatrix} (N-3)I_2 + 2J_2 & 0 \\ 0 & (N+1)I_{n-2} - J_{n-2} \end{bmatrix}$$

Table 4.2.1 below, is a complete list of the A-optimal design matrices in $\mathcal{D}(N,n)$ for $N \geq 7$, $N \equiv 3(\text{mod } 4)$, $n \leq 6$.

Table 4.2.1. Some A-optimal Designs in $\mathcal{D}(N,n)$, $N \equiv 3(\text{mod } 4)$.

N	n	Type of A-optimal Design
7	≤ 3	I
	4,5	II
	6	III
11	≤ 4	I
	5,6	II
15	≤ 5	I
	6	II
≥ 19	≤ 6	I

For $N \geq 35$, $n \leq 6$ the A-optimality of X_3 follows from Theorem 4.1.1. Since the cases $7 < N < 31$, $n \leq 6$ are similar and easier than the cases $N = 7$, $n \leq 6$ we only discuss the latter cases. Thus, let $N = 7$, $n \leq 6$.

Case A: $n \leq 3$. An application of Theorem 4.1.1 shows that X_3 is

A-optimal in $\mathcal{D}(7,n)$. The construction of such X_3 has already been discussed in Section 4.1.

Case B: $n = 4$. In this case the A-optimal design in $\mathcal{D}(7,4)$ is of Type II. Let

$$Y_3 = \begin{bmatrix} 0 & -1 & -1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \\ -1 & 1 & -1 & -1 \end{bmatrix},$$

then

$$Y_3^T Y_3 = \begin{bmatrix} 6 & 0 \\ 0 & 8I_3 - J_3 \end{bmatrix} \text{ and } \text{tr}(Y_3^T Y_3)^{-1} = \frac{37}{60} = .61666.$$

Let $X \in \mathcal{L}(7,4)$. We must show that $\text{tr}(X^T X)^{-1} \geq .61666$.

If $X \in \mathcal{L}(7,4)$, the result follows from Theorem 4.1.2.

Let X have one zero entry. Then,

$$X^T X = \begin{bmatrix} 6 & * \\ * & C \end{bmatrix}.$$

Thus by Ky Fan's theorem and the result for $n = 3$,

$$\text{tr}(X^T X)^{-1} \geq \text{tr}[6]^{-1} + \text{tr}C^{-1} \geq \frac{1}{6} + \text{tr}[8I_3 - J_3]^{-1} = .61666.$$

Let X have two zero entries. Then

$$X^T X = \begin{bmatrix} A & * \\ * & c \\ * & 7 \end{bmatrix}, \quad |c| \geq 1, \quad \text{tr}A = 12.$$

By Ky Fan's theorem and Theorem 3.2.1,

$$\text{tr}(X^T X)^{-1} \geq \text{tr}A^{-1} + \text{tr} \begin{bmatrix} 7 & c \\ c & 7 \end{bmatrix}^{-1} \geq \frac{4}{12} + \text{tr} \begin{bmatrix} 7 & -1 \\ -1 & 7 \end{bmatrix}^{-1} = .625 > .61666.$$

Let X have three or more zero entries. Then by Theorem 3.2.1,

$$\text{tr}(X^T X)^{-1} \geq \frac{n^2}{\text{tr}(X^T X)} \geq \frac{16}{25} = .64 > .61666..$$

Case C: $n = 5$. The A-optimal design in $\mathcal{L}(7,5)$ is of type II.

Let

$$Y_3 = \begin{bmatrix} 0 & -1 & -1 & -1 & -1 \\ -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & -1 & 1 \end{bmatrix}$$

then

$$Y_3^T Y_3 = \begin{bmatrix} 6 & 0 \\ 0 & 8I_4 - J_4 \end{bmatrix} \quad \text{and} \quad \text{tr}(Y_3^T Y_3)^{-1} = \frac{19}{24} = .79166.$$

Let $X \in \mathcal{L}(7,5)$. We must prove $\text{tr}(X^T X)^{-1} \geq .79166$.

If $X \in \mathcal{L}(7,5)$, this follows from Theorem 4.1.2.

Let X have one zero entry. Then, by using Ky Fan's theorem and Theorem 4.1.2 we can proceed in the case $n = 4$.

Let X have two zero entries. We may assume the zeros appear in distinct columns or proceed as in the case $n = 4$. Thus,

$$X^T X = \begin{bmatrix} 6 & c & & & & & \\ & 6 & & & & & \\ & & 7 & & & & \\ & & & 7 & & & \\ & & & & 7 & & \\ & & & & & 7 & \\ & & & & & & 7 \end{bmatrix}, \quad |c| \geq 0, \quad |b_{ij}| \geq 1,$$

If $|c| \geq 1$, then by Ky Fan's result and Theorem 4.1.2,

$$\text{tr}(X^T X)^{-1} \geq \text{tr} \begin{bmatrix} 6 & 1 \\ 1 & 6 \end{bmatrix}^{-1} + \text{tr}[8I_3 - J_3]^{-1} = \frac{12}{35} + \frac{2}{8} + \frac{1}{5} = .79285 > .79166.$$

Thus we may assume $c = 0$. Then the zero entries of X must lie on the same row. Thus, by Lemma 2.3.2, $|a_{ij}| \geq 2$ for at least three pairs (i,j) . Hence,

$$\text{tr}(X^T X)^2 \geq 2 \times 6^2 + 3 \times 7^2 + 6 \times 2^2 + 6 \times 1^2 = 249.$$

Now applying Theorem 3.2.2 with $t = \text{tr}(X^T X) = 33$ and $a = .79166$

we have

$$b(t, a) < 231 < \text{tr}(X^T X)^2.$$

Therefore, $\text{tr}(X^T X)^{-1} \geq a$.

Let X have three zero entries. Then

$$X^T X = \begin{bmatrix} A & * \\ 7 & c \\ & 7 \end{bmatrix}, \quad |c| \geq 1, \quad \text{tr} A = 18.$$

By Ky Fan's theorem and Theorem 3.2.1,

$$\text{tr}(X^T X)^{-1} \geq \text{tr} A^{-1} + \text{tr} \begin{bmatrix} 7 & 1 \\ 1 & 7 \end{bmatrix}^{-1} \geq \frac{9}{18} + \frac{14}{48} = .79166.$$

Let X have four or more zero entries. Then by Theorem 3.2.1,

$$\text{tr}(X^T X)^{-1} \geq \frac{25}{31} = .80645 > .79166.$$

Case D: $n = 6$. Then the A-optimal design in $\mathcal{D}(7, 6)$ is of Type III.

Let

$$Z_3 = \begin{bmatrix} 0 & 0 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 & -1 & 1 \end{bmatrix},$$

then

$$Z_3^T Z_3 = \begin{bmatrix} 4I_2 + 2J_2 & 0 \\ 0 & 8I_4 - J_4 \end{bmatrix} \quad \text{and} \quad \text{tr}(Z_3^T Z_3)^{-1} = 1.$$

Let $X \in \mathcal{D}(7, 6)$. We must show $\text{tr}(X^T X)^{-1} \geq 1$.

Let $X \in \mathcal{D}(7, 6)$. Then by Lemma 4.1.2,

$$X^T X = \begin{bmatrix} 7 & & & & & \\ & 7 & & & & \\ & & 7 & & & \\ & & & a_{ij} & & \\ & & & & 7 & \\ & & & & & 7 \end{bmatrix}, \quad a_{ij} = -1, 3 \text{ or } -5.$$

Suppose all $a_{ij} = -1$, then,

$$\text{tr}(X^T X)^{-1} = \text{tr}[8I_6 - J_6]^{-1} = \frac{9}{8} > 1.$$

Let some $a_{ij} = -5$ then by Ky Fan's theorem and Theorem 4.1.2,

$$\text{tr}(X^T X)^{-1} \geq \text{tr} \begin{bmatrix} 7 & -5 \\ -5 & 7 \end{bmatrix}^{-1} + \text{tr}[8I_4 - J_4]^{-1} = \frac{14}{24} + \frac{5}{8} = \frac{29}{24} > 1.$$

Thus we may assume $a_{ij} = -1$ or 3 . Let $t = \text{tr}(X^T X) = 42$, $a = 1$.

then $b(t, a) < 376$. Therefore, by Theorem 3.2.2, we may assume

further that $\text{tr}(X^T X)^2 \leq 375$. Hence six or less of the a_{ij} satisfy

$a_{ij} = 3$. If two of the a_{ij} equal 3 , then

$$\text{tr}(X^T X)^{-1} = \text{tr} \begin{bmatrix} 7 & 3 & & & & \\ 3 & 7 & & & & \\ & & 7 & & & \\ & & & -1 & & \\ & & & & 7 & \\ & & & & & 7 \end{bmatrix}^{-1} = 1.0625 > 1.$$

If four of the a_{ij} equal 3 , then up to similarity of $X^T X$, either

$$\text{tr}(X^T X)^{-1} = \text{tr} \begin{bmatrix} 7 & 3 & 3 & & \\ 3 & 7 & -1 & & * \\ 3 & -1 & 7 & & \\ & & & & c \\ * & & & & \end{bmatrix}^{-1} \geq \text{tr} \begin{bmatrix} 7 & 3 & 3 \\ 3 & 7 & -1 \\ 3 & -1 & 7 \end{bmatrix}^{-1} + \text{tr}[8I_3 - J_3]^{-1} = 1.11166 > 1$$

or

$$\text{tr}(X^T X)^{-1} = \text{tr} \begin{bmatrix} 7 & 3 & & & \\ 3 & 7 & & & \\ & & 7 & 3 & \\ & & 3 & 7 & \\ -1 & & & & 7 \end{bmatrix} = 1.0583 > 1.$$

If six of the a_{ij} equal 3, then the only remaining cases to be covered are

$$\text{tr}(X^T X)^{-1} = \text{tr} \begin{bmatrix} 7 & 3 & 3 & & & \\ & 7 & 3 & * & & \\ & & 7 & & & \\ & & & 7 & * & \\ & & & & 7 & \\ & & & & & 7 \end{bmatrix}^{-1} \geq \text{tr} \begin{bmatrix} 7 & 3 & 3 \\ 3 & 7 & 3 \\ 3 & 3 & 7 \end{bmatrix}^{-1} + \text{tr}[8I_3 - J_3]^{-1} = 1.026 > 1$$

or

$$\text{tr}(X^T X)^{-1} = \text{tr} \begin{bmatrix} 7 & 3 & & & & \\ 3 & 7 & & & & \\ & & 7 & 3 & & \\ & & 3 & 7 & & \\ -1 & & & & 7 & 3 \\ & & & & 3 & 7 \end{bmatrix}^{-1} \geq 3 \text{tr} \begin{bmatrix} 7 & 3 \\ 3 & 7 \end{bmatrix}^{-1} = \frac{42}{40} > 1$$

This completes the case where $X \in \mathcal{S}'(7,6)$.

Let X have one zero entry. Then one can proceed as in the case $n = 4$.

Let X have two zero entries. The usual arguments allow us to assume that the zeros appear in distinct columns. Thus,

$$X^T X = \begin{bmatrix} 6 & c & a_1 & a_2 & a_3 & a_4 \\ & 6 & a_5 & a_6 & a_7 & a_8 \\ & & 7 & b_1 & b_2 & b_3 \\ & & & 7 & b_4 & b_5 \\ & & & & 7 & b_6 \\ & & & & & 7 \end{bmatrix}, \quad |c| \geq 0, \quad |b_i| \geq 1.$$

If $|c| \geq 2$, then

$$\text{tr}(X^T X)^{-1} \geq \text{tr} \begin{bmatrix} 6 & 2 \\ 2 & 6 \end{bmatrix}^{-1} + \text{tr}[8I_4 - J_4]^{-1} = \frac{12}{32} + \frac{5}{8} = 1.$$

If $c = 0$, then the zero entries of X lie in the same row. So by

Lemma 2.3.2 $|a_i| \geq 2$ for at least four of the a_i . Thus,

$$\text{tr}(X^T X)^2 \geq 2 \times 6^2 + 4 \times 7^2 + 8 \times 2^2 + 12 \times 1^2 = 312.$$

But $b(40,1) < 314$. So by Theorem 3.2.2 and Lemma 4.1.2, we may assume that exactly four of the a_i are ± 2 , the rest being 0, and that all $b_i = -1$. Noting that Lemma 2.3.2 implies that $a_i = a_{i+4} = 0$, $i = 1, 2, 3, 4$ is not possible, we have, up to similarity of $X^T X$, the following cases:

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	$\text{tr}(X^T X)^{-1}$
2	2	2	2	0	0	0	0	1.7917
-2	2	2	2	0	0	0	0	1.2024
-2	-2	2	2	0	0	0	0	1.1042
2	-2	2	0	0	0	0	2	1.3194
-2	2	2	0	0	0	0	2	1.1144
2	2	0	0	0	0	2	2	1.225
-2	2	0	0	0	0	2	2	1.1417
-2	2	0	0	0	0	-2	2	1.075

In all cases, $\text{tr}(X^T X)^{-1} > 1$.

Now let $c = 1$ (or -1). If two of the b_i satisfy $|b_i| \geq 3$ or one b_i satisfies $|b_i| \geq 5$ then

$$\text{tr}(X^T X)^2 \geq 2 \times 6^2 + 4 \times 7^2 + 10 \times 1^2 + 4 \times 3^2 = 314 > b(40,1).$$

So, Theorem 3.2.2 applies. If say, $b_1 = 3$, $b_i = -1$, $i = 2, 3, \dots, 6$, then by Ky Fan's theorem

$$\text{tr}(X^T X)^{-1} \geq \text{tr} \begin{bmatrix} 6 & 1 \\ 1 & 6 \end{bmatrix}^{-1} + \text{tr} \begin{bmatrix} 7 & 3 & -1 & -1 \\ 3 & 7 & -1 & -1 \\ -1 & -1 & 7 & -1 \\ -1 & -1 & -1 & 7 \end{bmatrix}^{-1} = \frac{12}{35} + \frac{1184}{1792} = 1.00356 > 1.$$

Thus we may assume that all $b_i = -1$. If four or more of the a_i satisfy $|a_i| \geq 2$ or one a_i satisfies $|a_i| \geq 4$, then

$$\text{tr}(X^T X)^2 \geq 2 \times 6^2 + 4 \times 7^2 + 8 \times 2^2 + 14 \times 1^2 = 314 > b(40,1).$$

Therefore, we may assume that at most three of the a_i satisfy

$|a_i| = 2$. But by Lemma 2.3.2, at least two of the a_i have absolute value 2. Thus we are left with the following cases:

$$\text{i)} \quad a_1 = -a_5 = 2$$

$$\text{tr}(X^T X)^{-1} \geq \text{tr} \begin{bmatrix} 6 & 1 & 2 \\ 1 & 6 & -2 \\ 2 & -2 & 7 \end{bmatrix}^{-1} + \text{tr}[8I_3 - J_3]^{-1} = \frac{111}{189} + \frac{2}{8} + \frac{1}{5} = 1.037 > 1.$$

$$\text{ii)} \quad a_1 = a_5 = \pm 2$$

Lemma 2.3.2 implies that $|a_i| = 2$ for some $i \neq 1, 5$. Without loss of generality, we may assume $|a_2| = 2$. Then,

$$\begin{aligned} \text{tr}(X^T X)^{-1} &\geq \text{tr} \begin{bmatrix} 6 & 1 & a_1 & a_2 \\ 1 & 6 & a_1 & 0 \\ a_1 & a_1 & 7 & -1 \\ a_2 & 0 & -1 & 7 \end{bmatrix}^{-1} + \text{tr} \begin{bmatrix} 7 & -1 \\ -1 & 7 \end{bmatrix}^{-1} = \frac{918 - 2a_1 a_2}{1248 - 10a_1 a_2} + \frac{14}{48} \geq \frac{926}{1288} + \frac{14}{48} \\ &= 1.0106 > 1. \end{aligned}$$

$$\text{(iii)} \quad a_1 = a_2 = \pm 2.$$

We may assume $a_5 = a_6 = 0$ or else proceed as in i) and ii) above.

Thus

$$\text{tr}(X^T X)^{-1} \geq \text{tr} \begin{bmatrix} 6 & 1 & a_2 & a_2 \\ 1 & 6 & 0 & 0 \\ a_2 & 0 & 7 & -1 \\ a_2 & 0 & -1 & 7 \end{bmatrix}^{-1} + \text{tr} \begin{bmatrix} 7 & -1 \\ -1 & 7 \end{bmatrix}^{-1} = \frac{954}{1296} + \frac{14}{48} = 1.0277 > 1.$$

$$\text{iv)} \quad a_1 = -a_2 = \pm 2.$$

As above we may assume $a_5 = a_6 = 0$. If $|a_i| = 2$ for $i = 3$ or $i = 4$ then one of the above cases apply. Thus we may assume

$$X^T X = C_r \equiv \begin{bmatrix} 6 & 1 & 2 & -2 & 0 & 0 \\ & 6 & 0 & 0 & r & 0 \\ & & 7 & -1 & -1 & -1 \\ & & & 7 & -1 & -1 \\ * & & & & 7 & -1 \\ & & & & & 7 \end{bmatrix}, \quad |r| = 0 \text{ or } 2.$$

Now, C_r is similar to C_{-r} since C_{-r} can be obtained from C_r by multiplying rows one and two by -1 , columns one and two by -1 , then interchanging rows three and four and columns three and four. Therefore,

$$\begin{aligned} \text{tr}(X^T X)^{-1} &= \frac{1}{2} \text{tr} C_r^{-1} + \frac{1}{2} \text{tr} C_{-r}^{-1} \\ &\geq \text{tr} \left[\frac{1}{2} C_r + \frac{1}{2} C_{-r} \right]^{-1} \end{aligned}$$

$$= \text{tr} \begin{bmatrix} 6 & 1 & 2 & -2 & 0 & 0 \\ & 6 & 0 & 0 & 0 & 0 \\ & & 7 & -1 & -1 & -1 \\ & & & 7 & -1 & -1 \\ & & & & 7 & -1 \\ & & & & & 7 \end{bmatrix}^{-1}$$

$$= 1.0302$$

$$> 1.$$

This completes the case where X has two zero entries.

Let X have three zero entries.

If the three zero entries appear in the first column of X , then we can apply the usual arguments using Ky Fan's theorem and Theorem 4.1.2.

Suppose X has two zeros in the first column and one in the second column. Then,

$$X^T X = \left[\begin{array}{cc|cccc} 5 & c & a_1 & a_2 & a_3 & a_4 \\ & 6 & b_1 & b_2 & b_3 & b_4 \\ \hline & & 7 & c_1 & c_2 & c_3 \\ & & & 7 & c_4 & c_5 \\ & * & & & 7 & c_6 \\ & & & & & 7 \end{array} \right],$$

where $|a_i| = 1, 3, 5$, $|b_i| = 0, 2, 4, 6$, $|c_i| = 1, 3, 5$. If $|c| \geq 1$, then

$$\text{tr}(X^T X)^{-1} \geq \text{tr} \begin{bmatrix} 5 & 1 \\ 1 & 6 \end{bmatrix}^{-1} + \text{tr}[8I_4 - J_4]^{-1} = \frac{233}{232} > 1.$$

Thus assume $c = 0$. Now,

$$\begin{aligned} \text{tr}(X^T X)^2 &= 5^2 + 6^2 + 4 \times 7^2 + 2 \sum_{i=1}^4 a_i^2 + 2 \sum_{i=1}^4 b_i^2 + 2 \sum_{i=1}^6 c_i^2 \\ &= 257 + 2 \sum_{i=1}^4 a_i^2 + 2 \sum_{i=1}^4 b_i^2 + 2 \sum_{i=1}^6 c_i^2, \end{aligned}$$

and

$$b(39, 1) < 286.$$

Therefore if one $|a_i| \geq 3$ or two $|b_i| \geq 2$ or one $|c_i| \geq 3$ then

$\text{tr}(X^T X)^2 > b(39, 1)$ and Theorem 3.2.2 applies. Thus we may assume

that $|a_i| = 1$, $|b_1| = 2$ or 0 , $b_i = 0$, $i = 2, 3, 4$ and

$c_i = -1$. Hence, up to similarity

$$X^T X = C_s \equiv \left[\begin{array}{cccccc} 5 & 0 & 1 & a_2 & a_3 & a_4 \\ & 6 & s & 0 & 0 & 0 \\ & & 7 & -1 & -1 & -1 \\ & & & 7 & -1 & -1 \\ & * & & & 7 & -1 \\ & & & & & 7 \end{array} \right], \quad |s| = 0 \text{ or } 2.$$

Since C_s is similar to C_{-s} ,

$$\text{tr}(X^T X)^{-1} = \frac{1}{2} \text{tr} C_s^{-1} + \frac{1}{2} \text{tr} C_{-s}^{-1}$$

$$\geq \text{tr} \left[\frac{1}{2} C_s + \frac{1}{2} C_{-s} \right]^{-1}$$

$$= \text{tr} \begin{bmatrix} 5 & 0 & 1 & a_2 & a_3 & a_4 \\ & 6 & 0 & 0 & 0 & 0 \\ & & 7 & -1 & -1 & -1 \\ & & & 7 & -1 & -1 \\ & & & & 7 & -1 \\ & & & & & 7 \end{bmatrix}^{-1}$$

$$= \begin{cases} 1.1042 & \text{if } a_2 = a_3 = a_4 = 1 \\ 1.0452 & \text{if } a_2 = a_3 = 1, a_4 = -1 \\ 1.0278 & \text{if } a_2 = 1, a_3 = a_4 = -1. \end{cases}$$

This completes the case where two zeros occur in the first column and one zero in the second column of X .

Now assume the three zeros occur in distinct columns. Then

$$X^T X = \begin{bmatrix} 6 & a_1 & a_2 & b_1 & b_2 & b_3 \\ & 6 & a_3 & b_4 & b_5 & b_6 \\ & & 6 & b_7 & b_8 & b_9 \\ & & & 7 & c_1 & c_2 \\ & & & & 7 & c_3 \\ & & & & & 7 \end{bmatrix},$$

where $|b_i| = 0, 2, 4, 6$, $|c_i| = 1, 3, 5$.

Now, by Lemma 2.3.1, $|a_i| \geq 1$ for some i . Assume $|a_1| \geq 1$.

If $|a_i| \geq 2$ for some i , then

$$\text{tr}(X^T X)^{-1} \geq \text{tr} \begin{bmatrix} 6 & 2 \\ 2 & 6 \end{bmatrix}^{-1} + \text{tr}[8I_4 - J_4]^{-1} = \frac{12}{32} + \frac{3}{8} + \frac{1}{4} = 1.$$

Thus we may assume $a_1 = 1$. Therefore,

$$\begin{aligned} \text{tr}(X^T X)^2 &= 3 \times 6^2 + 3 \times 7^2 + 2 \times 1^2 + 2 \sum_{i=2}^3 a_i^2 + 2 \sum_{i=1}^9 b_i^2 + 2 \sum_{i=1}^3 c_i^2 \\ &= 257 + 2 \sum_{i=2}^3 a_i^2 + 2 \sum_{i=1}^9 b_i^2 + 2 \sum_{i=1}^3 c_i^2. \end{aligned}$$

If $|c_i| \geq 3$ for two or more values of i or $|c_i| \geq 5$ for one i , then

$$\text{tr}(X^T X)^2 \geq 286 \geq b(39, 1).$$

So Theorem 3.2.2 applies. If $c_1 = 3$, $c_i = -1$ for $i \neq 1$, then

$$\text{tr}(X^T X)^{-1} \geq \text{tr} \begin{bmatrix} 6 & 1 \\ 1 & 6 \end{bmatrix}^{-1} + \text{tr}[6]^{-1} + \text{tr} \begin{bmatrix} 7 & 3 & -1 \\ 3 & 7 & -1 \\ -1 & -1 & 7 \end{bmatrix}^{-1} = \frac{12}{35} + \frac{1}{6} + \frac{1}{2} = 1.0095 \geq 1.$$

Thus we may assume $c_i = -1$, $i = 1, 2, 3$. Further if $a_2 = 0$ or $a_3 = 0$ then by Lemma 2.3.2, $|b_i| \geq 2$ for at least three values of i , whence $\text{tr}(X^T X)^2 > b(39, 1)$. The same holds if $|b_i| \geq 4$ for at least one i . Thus we may assume that $a_1 = 1$, $|a_i| = 1$, $i = 2, 3$ and (by Lemma 2.3.3) that exactly two of the b_i satisfy $|b_i| = 2$, the rest being zero. Thus up to similarity of $X^T X$ we have

$$X^T X = \begin{bmatrix} 6 & 1 & a_2 & b_1 & b_2 & b_3 \\ & 6 & 1 & b_4 & b_5 & b_6 \\ & & 6 & 2 & b_8 & b_9 \\ & & & 7 & -1 & -1 \\ & & & & 7 & -1 \\ & & & & & 7 \end{bmatrix},$$

where $a_2 = \pm 1$, and exactly one $b_i = \pm 2$, the rest being zero. If

$|b_8| = 2$ or $|b_9| = 2$, then by Lemma 2.3.3 $|b_i| = 2$ for some

$i = 1, 2, \dots, 6$, whence $\text{tr}(X^T X)^2 > b(39, 1)$. Thus we may further

assume $b_8 = b_9 = 0$. The remaining cases up to similarity of $X^T X$ are computed below.

$\text{tr}(X^T X)^{-1}$		
$a_3 = 1$	$b_1 = 2$	1.0445
	$b_1 = -2$	1.0882
	$b_2 = 2$	1.0511
	$b_2 = -2$	1.0602
$a_3 = -1$	$b_1 = 2$	1.1188
	$b_1 = -2$	1.0480
	$b_2 = 2$	1.0784
	$b_2 = -2$	1.0640

This completes the case where X has three zero entries.

Now suppose X has four zero entries.

If the four zeros appear in the first two columns of X , then

$$X^T X = \begin{bmatrix} A & \vdots & \vdots & \vdots & \vdots \\ \vdots & 7 & \vdots & \vdots & \vdots \\ \vdots & \vdots & 7 & * & \vdots \\ \vdots & \vdots & \vdots & 7 & \vdots \\ \vdots & \vdots & \vdots & \vdots & 7 \end{bmatrix}, \text{tr } A = 10.$$

Thus by Ky Fan's theorem, Theorems 3.2.1 and 4.1.2

$$\text{tr}(X^T X)^{-1} \geq \text{tr} A^{-1} + \text{tr} \begin{bmatrix} 8I_4 & -J_4 \end{bmatrix}^{-1} \geq \frac{4}{10} + \frac{5}{8} = 1.025 > 1.$$

Now assume two zeros appear in the first column and one zero in the second and third columns of X . Then,

$$X^T X = \begin{bmatrix} 5 & a_1 & a_2 & b_1 & b_2 & b_3 \\ & 6 & a_3 & c_1 & c_2 & c_3 \\ & & 6 & c_4 & c_5 & c_6 \\ & & & 7 & d_1 & d_2 \\ & & & & 7 & d_3 \\ & & & & & 7 \end{bmatrix}, |b_i| \geq 1, |d_i| \geq 1, |c_i| \geq 0.$$

Therefore,

$$\begin{aligned} \text{tr}(X^T X)^2 &= 5^2 + 2 \times 6^2 + 3 \times 7^2 + 2 \sum_{i=1}^3 a_i^2 + 2 \sum_{i=1}^3 b_i^2 + 2 \sum_{i=1}^6 c_i^2 + 2 \sum_{i=1}^3 d_i^2 \\ &= 244 + 2 \sum_{i=1}^3 a_i^2 + 2 \sum_{i=1}^3 b_i^2 + 2 \sum_{i=1}^6 c_i^2 + 2 \sum_{i=1}^3 d_i^2. \end{aligned}$$

Also,

$$b(38,1) < 260.$$

If $|a_3| \geq 2$ or $|d_i| \geq 3$, for some i , then

$\text{tr}(X^T X)^2 \geq 260 \geq b(38,1)$. Thus we may assume $|a_3| \leq 1$ and $d_i = -1$

for all i . Further if $|a_3| = 1$, Lemma 2.3.3 implies $|c_i| \geq 2$ for

some i and hence $\text{tr}(X^T X)^2 > b(38,1)$. Finally, if $a_3 = 0$, then by

Lemma 2.3.2, at least three of the c_i satisfy $|c_i| \geq 2$ whence

$\text{tr}(X^T X)^2 > b(38,1)$.

Now suppose the four zeros of X lie in distinct columns. Then

$$X^T X = \begin{bmatrix} 6 & a_1 & a_2 & a_3 & b_1 & b_2 \\ & 6 & a_4 & a_5 & b_3 & b_4 \\ & & 6 & a_6 & b_5 & b_6 \\ & & & 6 & b_7 & b_8 \\ & & & & 7 & c_1 \\ & & & & & 7 \end{bmatrix}, \quad |c_1| \geq 1.$$

Thus,

$$\begin{aligned} \text{tr}(X^T X)^2 &\geq 4 \times 6^2 + 2 \times 7^2 + 2 \times 1^2 + 2 \sum_{i=1}^6 a_i^2 + 2 \sum_{i=1}^8 b_i^2 \\ &= 244 + 2 \sum_{i=1}^6 a_i^2 + 2 \sum_{i=1}^8 b_i^2. \end{aligned}$$

If two of the zeros of X occur in the same row then by Lemma 2.3.2,

$|b_i| \geq 2$ for at least two b_i . Therefore

$$\text{tr}(X^T X)^2 \geq 260 > b(38, 1).$$

Thus assume the zeros lie in distinct rows. Then $|a_i| \geq 1, i=1, 2, \dots, 6$.

Since we may assume $\text{tr}(X^T X)^2 \leq 259$, the only remaining cases up to

similarity of $X^T X$ are $a_1 = a_2 = a_3 = 1$, $|a_i| = 1, i = 4, 5, 6$,

$b_i = 0, i = 1, 2, \dots, 8$. Then,

$$\text{tr}(X^T X)^{-1} = \text{tr} \begin{bmatrix} 6 & 1 & 1 & 1 \\ & 6 & a_4 & a_5 \\ & & 6 & a_6 \\ & & & 6 \end{bmatrix}^{-1} + \text{tr} \begin{bmatrix} 7 & 1 \\ 1 & 7 \end{bmatrix}^{-1}$$

$$= \frac{792 + 2(a_4 + a_5 + a_6) + 2a_4 a_5 a_6}{1083 - 2(a_4 a_5 + a_4 a_6 + a_5 a_6) + 12(a_4 + a_5 + a_6) + 12a_4 a_5 a_6} + \frac{14}{48}$$

$$\geq \frac{800}{1125} + \frac{14}{48}$$

$$= 1.0027$$

$$> 1.$$

This completes the case where X has four zero entries.

Now let X have five zero entries.

If the five zero entries occur in the first three columns, then

$$X^T X = \left[\begin{array}{c|ccc} A & 1 & - & - & - \\ \hline - & 17 & * & - & - \\ & * & 7 & & \\ & & & & 7 \end{array} \right], \quad \text{tr} A = 16.$$

Therefore, as in previous cases,

$$\text{tr}(X^T X)^{-1} \geq \text{tr} A^{-1} + \text{tr} \left[8I_3 - J_3 \right]^{-1} \geq \frac{9}{16} + \frac{2}{8} + \frac{1}{5} = 1.0125 > 1.$$

Now suppose two zeros lie in the first column of X and one in each of the next three columns. Then

$$X^T X = \left[\begin{array}{cc|cc} 5 & & a_1 & a_2 \\ & 6 & * & \\ & & 6 & * \\ \hline & & 6 & - \\ \hline & & 7 & c \\ & & & 7 \end{array} \right], \quad |a_i| \geq 1, \quad |c| \geq 1.$$

Thus

$$\begin{aligned} \text{tr}(X^T X)^2 &\geq 5^2 + 3 \times 6^2 + 2 \times 7^2 + 6 \times 1^2 \\ &= 237 \\ &> b(37, 1). \end{aligned}$$

So Theorem 3.2.2 applies.

Finally, assume there is a zero entry in each of the first five columns of X . Then,

$$X^T X = \begin{bmatrix} 6 & a_1 & a_2 & a_3 & a_4 & b_1 \\ & 6 & a_5 & a_6 & a_7 & b_2 \\ & & 6 & a_8 & a_9 & b_3 \\ & & & 6 & a_{10} & b_4 \\ & & & & 6 & b_5 \\ & & & & & 7 \end{bmatrix}$$

Therefore,

$$\begin{aligned} \text{tr}(X^T X)^2 &\geq 5 \times 6^2 + 7^2 + 2 \sum_{i=1}^{10} a_i^2 + 2 \sum_{i=1}^5 b_i^2 \\ &= 229 + 2 \sum_{i=1}^{10} a_i^2 + 2 \sum_{i=1}^5 b_i^2 \end{aligned}$$

If $|a_i| = 0$ for some i then by Lemma 2.3.2, $|b_i| \geq 2$, for some i , hence $\text{tr}(X^T X)^2 \geq 237 > b(37,1)$. Thus assume $|a_i| \geq 1$ for all i . Then again $\text{tr}(X^T X)^2 > b(37,1)$.

Finally if X has six or more zero entries, then

$$\text{tr}(X^T X)^{-1} \geq \frac{6^2}{\text{tr}(X^T X)} = \frac{36}{36} = 1.$$

This completes the case $N = 7$, $n = 6$.

As mentioned previously the rest of Table 4.2.1 can be proved in a similar but easier manner.

It has already been discussed how Type I designs can be constructed and it is clear that such designs exist for $n \leq 6$ and all $N \geq 7$, $N \equiv 3 \pmod{4}$. To construct a Type II design for $N = 11$, $n = 6$, choose a 12×12 Hadamard matrix H such that

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 \\ & & & & & & * & & & & & \end{bmatrix}$$

Let X be the matrix obtained from H by removing the first row, the first five columns, and the last column of H . Then a Type II matrix is obtained by replacing x_{11} by 0. Similarly Type II designs can be constructed for $N = 11, n = 5$ and $N = 15, n = 6$. Thus all designs in Table 4.2.1 have been constructed.

Finally, we mention that in the saturated cases, $N = n = 7$ and $N = n = 11$ the above techniques are not very useful since they fail to reduce the number of competitors in any substantial manner.

4.3 ϕ_p -optimality in $\mathcal{D}(N, n)$ and $\mathcal{D}'(N, n)$.

In this section, we shall extend Theorem 4.1.1 to other criteria. A technique is developed to show that if a certain design is A-optimal, then it is also optimal with respect to a large class of other criteria. A result of this kind is interesting and important in its own right. The method used is a modification of the result of Cheng [4].

Lemma 4.3.1. For any positive numbers A, S and r such that $S > n^2/A$, $0 \leq r \leq n$, the system of equations $(n-r)\mu + r\lambda = A$ and $(n-r)/\mu + r/\lambda = S$, $\mu \leq \lambda$, has exactly one solution

$\lambda(r; A, S) > \mu(r; A, S) > 0$. Further, if $S = n^2/A$ there is only one solution $\lambda(r; A, S) = \mu(r; A, S) > 0$.

Proof. By solving for μ in the first equation and substituting into the second, we obtain the equation $Sr\lambda^2 + (n^2 - 2nr - SA)\lambda + rA = 0$. The discriminant of the quadratic is $h(S) = (n^2 - 2nr - SA)^2 - 4ASr^2$. The result now follows by noting that $h(n^2/A) = 0$ and $h'(S) = 2A(SA - n^2) + 4A(nr - r^2) > 0$ if $S > n^2/A$.

Lemma 4.3.2. Let $\mu(r; A, S)$ and $\lambda(r; A, S)$ be as in Lemma 4.3.1 with

$A \leq a_0$ and $S > n^2/A$. Let f be a real-valued function defined on $[0, a_0]$ such that

(i) f is continuous on $(0, a_0)$ (we allow $\lim_{x \rightarrow 0^+} f(x) = f(0) = +\infty$),

(ii) $g'' < 0$ on $(1/a_0, \infty)$, where $g(x) = f(1/x)$,

(4.3.1)

(iii) $f'' > 0$ on $(0, a_0)$,

(iv) For $a < b$ in $(0, a_0)$, $\{f(b) - f(a)\}/(b-a) < \{af'(a) + bf'(b)\}/(a+b)$,

and $F(r; A, S) \equiv (n-r)f\{\mu(r; A, S)\} + rf\{\lambda(r; A, S)\}$. Then F is a strictly decreasing function of r , strictly increasing function of S , and strictly decreasing function of A .

Proof. By differentiating the equations $(n-r)\mu + r\lambda = A$ and $(n-r)/\mu + r/\lambda = S$ with respect to r and S , then by solving the remaining equations for $\partial\mu/\partial r$, $\partial\lambda/\partial r$, $\partial\mu/\partial S$, and $\partial\lambda/\partial S$, we obtain

$$\partial\mu/\partial r = -(\lambda - \mu)\mu/(n-r)(\lambda + \mu), \quad \partial\lambda/\partial r = -(\lambda - \mu)\lambda/r(\lambda + \mu),$$

$$\partial\mu/\partial S = -\lambda^2\mu^2/(n-r)(\lambda^2 - \mu^2), \quad \partial\lambda/\partial S = \lambda^2\mu^2/r(\lambda^2 - \mu^2).$$

Thus

$$\begin{aligned} \partial F/\partial r &= -f(\mu) + (n-r)f'(\mu)\partial\mu/\partial r + f(\lambda) + rf'(\lambda)\partial\lambda/\partial r \\ &= f(\lambda) - f(\mu) - (\lambda - \mu)\{\mu f'(\mu) + \lambda f'(\lambda)\}/(\lambda + \mu) \\ &< 0, \text{ by (iv)}. \end{aligned}$$

Similarly it follows from (iii) that $\partial F/\partial S > 0$. We remark that if $f'' < 0$ on $(0, a_0)$, then F is a decreasing function of S . This fact will be used below in the proof of the decreasing monotonicity of F in A .

To show that F is a decreasing function of A , we write $F = (n-r)g(\mu') + rg(\lambda')$, where $\mu' = \mu^{-1}$, $\lambda' = \lambda^{-1}$ and $g(x) = f(x^{-1})$. Then $(n-r)\mu' + r\lambda' = S$ and $(n-r)/\mu' + r/\lambda' = A$. By

(ii) and the remark in the last paragraph, we conclude that F is a strictly decreasing function of A . This completes the proof.

Lemma 4.3.3. Let $H = \{(x_1, x_2, \dots, x_n) : x_i > 0, \sum_{i=1}^n x_i = A, \sum_{i=1}^n x_i^{-1} = S\}$,

where $A \leq a_0$ and $S \geq n^2/A$. Also let $F_f : H \rightarrow \mathbb{R}$ be defined by

$$F_f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n f(x_i), \text{ where } f \text{ is a real-valued function}$$

defined on $[0, a_0]$ which satisfies (4.3.1) and the following condition:

(4.3.2) The equation $x^2 f'(x) + \alpha x^2 - \beta = 0$ has at most two solutions in

$(0, a_0)$ for all real numbers α and β .

Then the minimum of $F_f(x_1, x_2, \dots, x_n)$ on H occurs at the point

$(\mu(n-1; A, S), \lambda(n-1; A, S), \dots, \lambda(n-1; A, S))$, where $\mu(r; A, S)$ and $\lambda(r; A, S)$

are as in Lemma 4.3.1.

Proof. First assume $S > n^2/A$. Since H is compact, the minimum of F_f on H is attained at some point $(a_1, a_2, \dots, a_n)$. Clearly not all

the a_i 's are equal (since $S > n^2/A$), so by the symmetry of F_f , we

may assume $a_1 < a_2$. Letting $g_1(x_1, x_2, \dots, x_n) = \sum_{i=1}^n x_i - A$ and

$$g_2(x_1, x_2, \dots, x_n) = \sum_{i=1}^n x_i^{-1} - S; \text{ then}$$

$\det\{D_j g_i(a_1, a_2, \dots, a_n)\}_{i,j=1,2} = a_1^{-2} - a_2^{-2} \neq 0$. So by Lagrange's theorem

(see Apostol, [1], p.381), there exist numbers α and β such that

$a_1, a_2, \dots, a_n$ satisfy the following equations:

$$(4.3.3) \quad \partial\{F_f(x_1, x_2, \dots, x_n) + \alpha g_1(x_1, x_2, \dots, x_n) + \beta g_2(x_1, x_2, \dots, x_n)\} / \partial x_i = 0,$$

$$i = 1, 2, \dots, n.$$

Now (4.3.3) simplifies to $f'(x_i) + \alpha - \beta x_i^{-2} = 0, i = 1, 2, \dots, n$. By

(4.3.2), each x_i can take on at most two possible values. So the point $(a_1, a_2, \dots, a_n)$ must be such that $a_i = a_1$ or a_2 . The result now follows from Lemma 4.3.1 and Lemma 4.3.2.

Now suppose $S = n^2/A$. Then the result follows since for $(x_1, \dots, x_n) \in H$, $(\frac{A}{n}, \frac{A}{n}, \dots, \frac{A}{n})^T = (\bar{x}, \bar{x}, \dots, \bar{x})^T \prec (x_1, \dots, x_n)^T$ and f is continuous and convex.

Now we are ready to prove the main result of this section.

Theorem 4.3.1. Let $\mathcal{M}$ be a family of $n \times n$ symmetric nonnegative definite matrices. Let $M_0 \in \mathcal{M}$ be such that

- (i) M_0 is a multiple of I_n or has two distinct eigenvalues $\lambda > \mu$ such that the multiplicity of λ is $n-1$
- (ii) M_0 maximizes $\text{tr} M$ over $\mathcal{M}$
- (iii) M_0 minimizes $\text{tr} M^{-1}$ over $\mathcal{M}$.

Then M_0 minimizes $\phi_f(M)$ over $\mathcal{M}$ for all f satisfying (4.3.1) and (4.3.2) (with $a_0 = \text{tr} M_0$).

Proof. Let $M_0 \in \mathcal{M}$ satisfy (i) - (iii) and let $S_0 = \text{tr} M_0^{-1}$, $a_0 = \text{tr} M_0$. Let $M \in \mathcal{M}$ be arbitrary and let $S = \text{tr} M^{-1}$, $a = \text{tr} M$. Then by assumption $S_0 \leq S$ and $a_0 \geq a$, so by Lemma 4.3.2 and 4.3.3 we have

$$\begin{aligned} \sum_{i=1}^n f(\mu_i) &\geq (n-1)f(\lambda(n-1; A, S)) + f(\mu(n-1; A, S)) \\ &\geq (n-1)f(\lambda(n-1; A, S_0)) + f(\mu(n-1; A, S_0)) \\ &\geq (n-1)f(\lambda(n-1; a_0, S_0)) + f(\mu(n-1; a_0, S_0)) \\ &= \sum_{i=1}^n f(\mu_i^*) \end{aligned}$$

where $\mu_1^*, \mu_2^*, \dots, \mu_n^*$ are the eigenvalues of M_0 . The last equality holds since $\text{tr} M_0 = a_0$, $\text{tr} M_0^{-1} = S_0$ and M_0 has two distinct eigenvalues with the smaller one being a simple root. This completes the proof.

Combining Theorems 4.1.1 and 4.3.1, we have

Corollary 4.3.1. For each n , there exists a positive integer $N_0(n)$ such that for all $N \geq N_0(n)$, X_3 is ϕ_f -optimal over $\mathcal{L}(N, n)$ for all f satisfying the conditions in (4.3.1) and (4.3.2).

Cheng [7] found the following simple sufficient condition for (4.3.1) and (4.3.2).

Lemma 4.3.4. If $f : [0, a_0] \rightarrow \mathbb{R}$ is such that (i), (ii) and (iii) in (4.3.1) hold, and $x^3 f''(x)$ is an increasing function on $(0, a_0)$, then all the conditions in (4.3.1) and (4.3.2) hold.

Proof. We need to verify (4.3.2) and condition (iv) in (4.3.1). Now the latter is equivalent to $\{af'(a)+bf'(b)\}(b-a)-\{f(b)-f(a)\}(a+b) > 0$. For fixed a , let $h(b) = \{af'(a)+bf'(b)\}(b-a)-\{f(b)-f(a)\}(a+b)$. Then since $h(a) = 0$, to show $h(b) > 0$ for all $b > a$, it suffices to prove $h'(b) > 0$ for all $b > a$. Now

$$h'(b) = f'(b)(b-a) + a\{f'(a) - f'(b)\} + bf''(b)(b-a) - f(b) + f(a).$$

So it is enough to prove $h''(b) > 0$ for all $b > a$. We have

$h''(b) = 3f''(b)(b-a) + bf'''(b)(b-a)$. Since $x^3 f''(x)$ is an increasing function, $d\{x^3 f''(x)\}/dx > 0$, i.e.,

$$(4.3.4) \quad xf'''(x) + 3f''(x) > 0 \quad \text{for all } x \in (0, a_0).$$

Therefore $h''(b) = \{bf'''(b) + 3f''(b)\}(b-a) > 0$ for all $b > a$. This proves condition (iv) in (4.3.1).

Now we prove (4.3.2). Suppose x and y are two distinct solutions of the equation $x^2 f'(x) + \alpha x^2 - \beta = 0$. Then $x^2 f'(x) + \alpha x^2 - \beta = y^2 f'(y) + \alpha y^2 - \beta$ and hence

$$(4.3.5) \quad \{x^2 f'(x) - y^2 f'(y)\} / (x^2 - y^2) = -\alpha.$$

Let $g(x) = x f'(\sqrt{x})$. We shall show that g is a strictly convex function; then for each fixed x , there is at most one y satisfying (4.3.5) and therefore (4.3.2) is proved. Now

$$g''(x) = 4x^{-1/2} \{f'''(\sqrt{x}) \sqrt{x} + 3f''(\sqrt{x})\} \text{ which, by (4.3.4), is positive.}$$

This completes the proof.

It is easy to see that if $0 < p < 1$, then the function $f(x) = x^{-p}$ satisfies the conditions in Lemma 4.3.4. Therefore Theorem 4.3.1 and Corollary 4.3.1 hold for all the ϕ_p -criteria with $0 < p \leq 1$. By passing to the limit or taking $f(x) = -\log x$, D-optimality also follows. We state this in the following:

Corollary 4.3.2. For each n , there exists a positive integer $N_0(n)$ such that for all $N \geq N_0(n)$, X_3 is ϕ_p -optimal over $\mathcal{D}(N, n)$ for all $0 \leq p \leq 1$; in particular, it is D-optimal.

As shown in Theorem 4.1.1, one can take $N_0(n) = \max\{(n-2)(n^2-n+16)/8, n^2-2\}$ in Corollaries 4.3.1 and 4.3.2. This is by no means the smallest bound. But since our result is much stronger than the D-optimality, the smallest $N_0(n)$ must be larger than $2n-5$, the bound found by Galil and Kiefer [13] for the D-criterion.

Also note that, unlike Theorem 3.3.1, one can choose $N_0(n)$ above to be the smallest integer $k(n)$ such that for all $N \geq k(n)$ X_3 is A-optimal in $\mathcal{D}(N, n)$. Hence we have:

Corollary 4.3.3. If X_3 is A-optimal in $\mathcal{L}(N,n)$ ($\mathcal{L}'(N,n)$) then it is ϕ_p -optimal for all $0 \leq p \leq 1$.

As an application of Corollary 4.3.3 we obtain:

Corollary 4.3.4. All Type I designs in Table 4.2.1 are ϕ_p -optimal, $0 \leq p \leq 1$, over $\mathcal{L}(N,n)$.

Proof. By definition X_3 is a Type I design. The result now follows from Corollary 4.3.3 since all designs in Table 4.2.1 are A-optimal.

Corollary 4.3.5. X_3 is ϕ_p -optimal, $0 \leq p \leq 1$, over $\mathcal{L}'(N,n)$ for $n \leq 5$, $N \geq 7$ and $n \leq 7$, $N \geq 15$.

Proof. This follows from Theorem 4.1.2 and Corollary 4.3.3.

Remark. Using a similar method, one could generalize Corollary 4.3.2 to show that for any n and $p > 0$, there exists a positive integer $N(n,p)$ such that for all $N \geq N(n,p)$, X_3 is ϕ_q -optimal over $\mathcal{L}(N,n)$ for all $0 \leq q \leq p$. However, since X_3 is not E-optimal (even when N gets large), one has $\lim_{p \rightarrow \infty} N(n,p) = \infty$. Such a result

is not very useful practically. Furthermore, for $p \neq 1$, there is no simple way to calculate a bound for $N(n,p)$ as we did in Section 4.1 for $p = 1$.

CHAPTER V

PROPOSALS FOR FURTHER RESEARCH

In this dissertation various techniques were developed for proving designs optimal in $\mathcal{D}(N, n)$ (or $\mathcal{D}'(N, n)$). We now give some proposals for future research toward improving and generalizing these methods.

Asymptotic results such as Theorem 3.3.1 and Corollary 4.3.2 enable one to show that a particular design satisfies certain optimality properties if N is sufficiently larger than some $N_0(n)$. It has been mentioned that the upper bounds given for $N_0(n)$ are by no means the smallest possible bounds. Thus further work on reducing these bounds is clearly needed.

Another technique used in determining ϕ_f -optimal designs is to minimize $\sum_{i=1}^n f(x_i)$ subject to the constraints $\sum_{i=1}^n x_i = \text{constant}$, $\sum_{i=1}^n g(x_i) = \text{constant}$, and $x_i \geq 0$. This method was successfully applied in Cheng [4] with $g(x) = x^2$ and in our Theorem 4.3.1 with $g(x) = 1/x$. It is our hope that other fruitful choices of $g(x)$ can be found.

It has been observed that A-optimal designs in $\mathcal{D}(N, n)$ may not be A-optimal in $\mathcal{D}'(N, n)$ (see for example, Table 4.2.1). Such a phenomena is in fact not uncommon. We need the following geometrical result (which we state without proof).

Theorem 5.1. Let $X \in M_{N \times n}$ such that $(X^T X)^{-1}$ exists. Let Z be obtained from X by replacing X_1 in $X = (x_{ij}) = [X_1, X_2, \dots, X_n]$ by Z_1 where $Z_1^T X_i = 0$, $i \geq 2$. Suppose further that $|X^T X| = |Z^T Z|$. Then

$$\text{tr}(Z^T Z)^{-1} \leq \text{tr}(X^T X)^{-1}$$

with equality if and only if $X_1^T X_i = 0$ for $i \geq 2$.

By Theorem 3.2.1, Theorem 5.1 and Ky Fan's majorization theorem we conclude that

$$X = \begin{bmatrix} 0 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$$

is A-optimal in $\mathcal{D}(3,3)$. Unfortunately, aside from trivial cases, Theorem 5.1 is not a very useful tool in proving designs A-optimal. A strengthening of this result would be welcome.

In general, the difficulty of proving designs optimal becomes much more difficult as N approaches n . Stronger methods are then required for these cases, especially the saturated case $N = n$. The case $N = n = 7$ is an example where our techniques are of little help in finding the A-optimal designs (although this case is statistically less important).

Finally, we hope that improved techniques will result in the proofs of conjecture (C1) of Section 3.2 and conjecture (C2) of Section 3.3.

REFERENCES

- [1] Apostol, T.M. (1974). Mathematical analysis, 2nd. ed. Addison-Wesley, Reading, Mass.
- [2] Ash, A. and Hedayat, A. (1978). An introduction to design optimality with an overview of the literature. Commun. Statist., A7, 1295-1325.
- [3] Bhaskararao, M. (1966). Weighing designs when n is odd. Ann. Math. Statist. 37, 1371-1381.
- [4] Cheng, C.S. (1978). Optimality of certain asymmetrical experimental designs. Ann. Statist. 6, 1239-1261.
- [5] _____ (1980). Optimality of some weighing and 2^n fractional factorial designs. Ann. Statist. 8, 436-446.
- [6] _____ (1981). Maximizing the total number of spanning trees in a graph: two related problems in graph theory and optimum design theory. J. Combinatorial Theory Ser. B 31, 240-248.
- [7] Cheng, C.S., Masaro, J.C. and Wong, C.S. (1983). Optimal Weighing Designs (submitted).
- [8] Cheng, C.S., Masaro, J.C. and Wong, C.S. (1983). Do nearly balanced graphs have more spanning trees? (Submitted).
- [9] Ehlich, H. (1964). Determinantenabschätzungen für binäre matrizen. Math. Zeitschr. 83, 123-132.
- [10] _____ (1964). Determinantenabschätzungen für binäre matrizen mit $n \equiv 3(\text{mod } 4)$. Math. Zeitschr. 84, 438-447.
- [11] Ehlich, H. and Zeller, K. (1962). Binäre matrizen. Z. Angew. Math. Mech. 42, T20-T21.
- [12] Gaffke, N. (1978). Optimale versuchsplanung für lineare zwei-faktor-modelle. Doctoral Dissertation, RWTH, Aachen.
- [13] Galil, Z. and Kiefer, J. (1980). D-optimum weighing designs. Ann. Statist. 8, 1293-1306.
- [14] _____ (1980). Construction methods for D-optimum weighing designs when $n \equiv 3(\text{mod } 4)$. Ann. Statist. 10, 502-510.
- [15] Hedayat, A. (1981). Study of Optimality criteria in design of experiments. In Statistics and Related Topics. (M. Csörgö, D.A. Dawson, J.N.K. Rao, A.K.Md.E. Saleh, eds.) 39-56. North Holland, Amsterdam.

- [16] Hedayat, A. and Wallis, W.D. (1978). Hadamard matrices and their applications. *Ann. Statist.* 6, 1184-1238.
- [17] Hotelling, H. (1944). Some improvements in weighing and other experimental techniques. *Ann. Math. Statist.* 15, 297-306.
- [18] Jacroux M., Wong, C.S. and Masaro, J.C. (1982). On the optimality of chemical balance weighing designs. (To appear in the *J. Statist. Plan. Inference*).
- [19] John, J.A. and Mitchell, T.J. (1977). Optimal incomplete block designs. *J.R. Statist. Soc. Ser.B*, 39, 39-43.
- [20] Kelmans, A.K. and Chelnokov, V.M. (1974). A certain polynomial of a graph and graphs with an external number of trees. *J. Combinatorial Theory, Ser.B*, 16, 197-214.
- [21] Kiefer, J. (1958). On the nonrandomized optimality and randomized nonoptimality of symmetrical designs. *Ann. Math. Statist.* 29, 675-699.
- [22] _____ (1959). Optimum experimental designs. *J. Roy. Statist. Soc. Ser.B*, 21, 272-319.
- [23] _____ (1974). General equivalence theory for optimum designs (approximate theory). *Ann. Statist.* 2, 849-879.
- [24] _____ (1975). Constructions and optimality of generalized Youden designs. In *a Survey of Statistical Designs and Linear Models*. (J.N. Srivastava, ed.) 333-353. North Holland, Amsterdam.
- [25] Marshall, A.W. and Olkin, I. (1979). *Inequalities: Theory of majorization and its applications*. Academic Press, New York.
- [26] Payne, S.E. (1974). On maximizing $\det(A^T A)$. *Discrete Math.* 10, 145-158.
- [27] Raghavarao, D. (1959). Some optimum weighing designs. *Ann. Statist.* 30, 295-303.
- [28] _____ (1960). Some aspects of weighing designs. *Ann. Math. Statist.* 31, 878-884.
- [29] Raktoe, B.L., Hedayat, A. and Federer, W.T. (1981). *Factorial designs*. John Wiley and Sons, New York.
- [30] Shier, D.R. (1974). Maximizing the number of spanning trees in a graph with n nodes and m edges. *J. Res. Nat. Bur. Standards Sect.B*, 78, 193-196.
- [31] Srivastava, J.N. and Anderson, D.A. (1974). A comparison of the determinant, maximum root and trace optimality criteria. *Commun. Statist.* 3, 933-940.

- [32] Wojtas, W. (1964). On Hadamard's inequality for the determinants of order non-divisible by 4. Colloq. Math. 12, 78-83.
- [33] Wong, C.S. (1982). Optimal design of experiments (talks given at Xian and Zhangsha, China Dec. 1982-Jan. 1983).
- [34] Wong, C.S. and Masaro, J.C. (1983). A-optimal design matrices $X = (x_{ij})_{N \times n}$ with $x_{ij} = -1, 0, 1$. (To appear in Linear and Multilinear algebra).
- [35] Wong, C.S. and Masaro, J.C. (1982). A-optimal design matrices, (submitted).
- [36] Yang, C.H. (1968). On designs of maximal $(+1, -1)$ - matrices of order $n \equiv 2 \pmod{4}$. Math. Comp. 22, 174-180.

VITA AUCTORIS

- 1949 Born, Windsor, Ontario, Canada
- 1972 B.A., Honours Mathematics, University of Windsor
- 1973 M.Sc., Mathematics, University of Windsor