Optimal Designs for Treatment-Control Comparisons in the Presence of Two-Way Heterogeneity

by

William I. Notz

Technical Report #84-9

Department of Statistics
Purdue University

March 1984

Research is sponsored by NSF grant no. 8301076

AMS 1980 subject classifications. Primary 62K05; secondary 62K10.

Key words and phrases. A-optimality, E-optimality, latin square design, balanced treatment block design.

Optimal Designs for Treatment-Control Comparisons in the Presence of Two-Way Heterogeneity

by

William I. Notz

Purdue University

Abstract

In this paper we shall consider the additive 3 factor linear model such as is used in Latin Square design settings. We seek optimal designs under this model for comparing p test treatments each simultaneously with a control. Conditions under which a design is A-optimal and under which a design is E-optimal are derived and examples given of when these conditions are satisfied. Since E-optimality is a criterion that involves the variances of contrasts other than those of test treatments vs. the control, we show that many E-optimal designs are also optimal under two criteria involving only the variances of test treatments vs. the control contrasts.

1. Introduction: Consider an experimental situation where it is desired to compare $p \geq 2$ test treatments to a control treatment. Let the p+1 treatments be indexed 0,1,...,p with 0 denoting the control treatment and 1,2,...,p the test treatments. Suppose the treatments can be applied in plots arranged in R rows and C columns. Assume that only one treatment can be applied in each plot. An observation Y_{ijk} is to be taken on the treatment in the plot located in row j ($1 \leq j \leq R$) and column k ($1 \leq k \leq C$) where treatment i ($0 \leq i \leq p$) is the treatment applied in this plot. If the row and column location of the plot can possibly effect the value of Y_{ijk}, we might consider modeling Y_{ijk} by the additive linear model

$$(1.1) \qquad Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}.$$

μ is the overall average response, α_i the effect of treatment i, β_j the effect of row j, γ_k the effect of column k, and ε_{ijk} the random error present in Y_{ijk}. We assume the ε_{ijk} are uncorrelated random variables with mean 0 and variance σ^2. We also make the usual assumptions that

$$\sum_{i=1}^{p} \alpha_i = \sum_{j=1}^{R} \beta_j = \sum_{k=1}^{C} \gamma_k = 0.$$

Model (1.1) is the usual model associated with Latin Square and Youden designs. It arises in many agricultural, biological, and industrial settings. Most of the design literature for model (1.1) applies to situations where one is interested in estimating all possible contrasts among the $\alpha_0, \alpha_1, \ldots, \alpha_p$. In this paper, however, we shall focus attention on the case where one is only interested in contrasts of test treatments with the control, i.e. contrasts of the form $a_0\alpha_0 - \sum_{i=1}^{p} a_i\alpha_i$ with $\sum_{i=1}^{p} a_i = a_0$, $a_0 \neq 0$. These treatment-control contrasts are to be estimated by their best linear unbiased estimates (B.L.U.E.s).

An experimental design in this setting is a particular allotment of treatments to the RC plots, one treatment per plot. Our goal is to find the best, in some sense, design for estimating all test treatment-control contrasts. To that end we introduce the following notation. For given values of p, R, and C let D(p,R,C) denote the set of all possible designs for (1.1). For $d \in D(p,R,C)$ let $r_{ij}(d)$ denote the number of times treatment i occurs in plots in row j, let $s_{ik}(d)$ denote the number of times treatment i occurs in plots in column k, let

$$r_i(d) = \sum_{j=1}^{R} r_{ij}(d) = \sum_{k=1}^{C} s_{ik}(d)$$

which is the total number of times treatment i occurs in design d, let

$$\lambda_{hi}(d) = \sum_{j=1}^{R} r_{hj}(d)\ r_{ij}(d)$$

and

$$\mu_{hi}(d) = \sum_{k=1}^{C} s_{hk}(d)\ s_{ik}(d).$$

Whenever it is clear which design d is being referred to we shall drop the (d) to simplify notation.

For estimating all possible treatment contrasts it is well known (see for example Kiefer (1958)), that the information matrix N(d), for $d \in D(p,R,C)$ is the $p+1 \times p+1$ matrix whose i,j-th entry n_{ij} is

$$(1.2) \qquad n_{ij} = \begin{cases} r_i = \lambda_{ii}/C - \mu_{ii}/R + r_i^2/RC & \text{if } i=j \\ -\lambda_{ij}/C - \mu_{ij}/C + r_i r_j/RC & \text{if } i \neq j \end{cases}$$

where $0 \leq i, j \leq p$. The row and column sums of N(d) are known to be 0. Paralleling the argument in the appendix of Bechhofer and Tamhane (1981) one can show the information matrix M(d) for estimating all test treatment-control contrasts is the $p \times p$ matrix whose i, j-th entry, m_{ij}, is

$$(1.3) \qquad m_{ij} = \begin{cases} r_i - \lambda_{ii}/C - \mu_{ii}/R + r_i^2/RC & \text{if } i=j \\ -\lambda_{ij}/C - \mu_{ij}/R + r_i r_j/RC & \text{if } i \neq j \end{cases}$$

where $1 \leq i, j \leq p$. M(d) is a nonnegative definite matrix and is nonsingular if and only if all the p contrasts $\alpha_0-\alpha_1$, $\alpha_0-\alpha_2,\ldots,\alpha_0-\alpha_p$ are estimable. If M(d) is nonsingular, $M^{-1}(d)$ is proportional to the variance-covariance matrix of the B.L.U.E. of the $p\times 1$ vector $(\alpha_0-\alpha_1, \alpha_0-\alpha_2,\ldots,\alpha_0-\alpha_p)'$, where primes on vectors indicate the transpose.

We are now in a position to state explicitly what we mean by a design $d \in D(p,R,C)$ being "best" for estimating all test treatment-control contrasts. Following the work of Kiefer (see, for example, Kiefer 1958, 1959, 1971, 1974, and 1975) a design $d \in D(p,R,C)$ is best or ϕ-optimal if it minimizes $\phi(M(d))$ over D(p,R,C) for some real valued function ϕ. Restricting ourselves to nonsingular designs, some common examples are $\phi_0(M(d)) = \det M^{-1}(d)$ (so called D-optimality), $\phi_1(M(d)) = \mathrm{tr}\, M^{-1}(d)$ (so called A-optimality), and $\phi_\infty(M(d))$ = the maximum eigenvalue of $M^{-1}(d)$ (so called E-optimality). In the present context of test treatment-control comparisons A-optimality has an appealing statistical interpretation, namely an A-optimal design minimizes $\sum_{i=1}^{p} \text{Var}(\hat{\alpha}_0 - \hat{\alpha}_i)$ over $D(p,R,C)$, where $\hat{\alpha}_i$ is the B.L.U.E. of α_i. D- and E-optimality are criteria involving the eigenvalues of $M^{-1}(d)$ which contains information about all test treatment-control contrasts as well as information about test treatment contrasts arising from contrasts in the $\alpha_0 - \alpha_i$. Thus D- and E-optimality may not be interpretable in terms of variances involving only test treatment-control contrasts. We shall see in section 3, however, that many E-optimal designs are optimal in meaningful ways involving only variances of test treatment-control contrasts. D-optimality still lacks a reasonable interpretation and so will not be considered further.

2. A-optimal designs. In this section we shall find conditions under which a design will be A-optimal over $D(p,R,C)$. The plan of attack is to find a lower bound on the value of $\text{tr}\, M^{-1}(d)$ for nonsingular $d \in D(p,R,C)$ and then to find conditions under which this bound is attained.

To begin with, suppose $d \in D(p,R,C)$ is an arbitrary design. Let $\mathcal{S}$ be the set of all permutations of test treatments $1,\ldots,p$. For $\sigma \in \mathcal{S}$, σd will denote the design resulting from d by the permutation σ of the treatments in d. For example, if σ is the simple permutation that interchanges treatments 1 and 2 leaving all other test treatments fixed, than σd is the design resulting from d by changing treatment 2 to 1 and treatment 1 to 2 in any plots containing treatments 1 and 2. All other plots are left unchanged. We define

$$\overline{M}(d) = \sum_{\sigma \in \mathcal{S}} M(\sigma d)/p! = \sum_{\pi \in \Pi} \pi' M(d)\pi/p!$$

where Π is the set of all $p \times p$ permutation matrices.

Lemma 2.1. If $d \in D(p,R,C)$ then $\overline{M}(d)$ has eigenvalues $m_1(d)$, $m_2(d) = \ldots = m_p(d)$ with

$$m_1(d) = r_0/p - \sum_{j=1}^{R} r_{0j}^2/Cp - \sum_{k=1}^{C} s_{ok}^2/Rp + r_0^2/RCp$$

$$m_2(d) = \sum_{i=1}^{p} \{r_i - \sum_{j=1}^{R} r_{ij}^2/C - \sum_{k=1}^{C} s_{ik}^2/R + r_i^2/RC\}/(p-1)$$

$$- m_1(d)/(p-1)$$

pf. The following relationships are easily verified from the definitions of r_i, the r_{ij}, the s_{ik}, the λ_{hi} and the μ_{hi}.

$$\sum_{i=1}^{p} r_{ij} = C - r_{0j}$$

$$\sum_{\substack{i=1 \\ i \neq h}}^{p} \lambda_{hi} = \sum_{\substack{i=1 \\ i \neq h}}^{p} \sum_{j=1}^{R} r_{hj} r_{ij} = \sum_{j=1}^{R} r_{hj} \sum_{\substack{i=1 \\ i \neq h}}^{p} r_{ij}$$

$$= \sum_{j=1}^{R} r_{hj} \{C - r_{0j} - r_{hj}\}$$

$$= Cr_h - \sum_{j=1}^{R} r_{hj} r_{0j} - \sum_{j=1}^{R} r_{hj}^2$$

$$\sum_{i=1}^{p} \lambda_{0i} = \sum_{i=1}^{p} \sum_{j=1}^{R} r_{0j} r_{ij} = \sum_{j=1}^{R} r_{0j}\{C - r_{0j}\}$$

$$= Cr_0 - \sum_{j=1}^{R} r_{0j}^2$$

Similar relationships hold for $\sum_{i=1}^{p} s_{ik}$, $\sum_{\substack{i=1 \\ i\neq h}}^{p} \mu_{hi}$, and $\sum_{i=1}^{p} \mu_{0i}$. Simply replace r_{ij} by s_{ij}, λ_{hi} by μ_{hi}, R by C, and C by R in the above formulas.

In addition since the j-th column sum of N(d), as given in (1.2), is 0

$$\lambda_{0j}/C + \mu_{0j}/R - r_0 r_j/RC = r_j - \sum_{\ell=1}^{R} r_{j\ell}^2/C - \sum_{\ell=1}^{C} s_{j\ell}^2/R$$

$$+ r_j^2/RC - \sum_{\substack{i=1 \\ i\neq j}}^{p} \{\lambda_{ij}/C - \mu_{ij}/R - r_i r_j/RC\}$$

Utilizing these relationships it is straightforward to calculate $\overline{M}(d)$ from M(d), as given in (1.3), and to show that $\overline{M}(d) = aI_p + bJ_{p,p}$ with I_p being the p×p identity matrix, $J_{p,p}$ the p×p matrix all of whose entries are +1, and

$$a = \sum_{i=1}^{p} \{r_i - \sum_{\ell=1}^{R} r_{i\ell}^2/C - \sum_{\ell=1}^{C} s_{i\ell}^2/R + r_i^2/RC$$

$$+ \sum_{\substack{j=1 \\ j\neq i}}^{p} \{\lambda_{ij}/C + \mu_{ij}/R - r_i r_j/RC\}/(p-1)\}/p$$

$$b = \sum_{i=1}^{p} \sum_{\substack{j=1 \\ j\neq i}}^{p} \{\lambda_{ij}/C + \mu_{ij}/R - r_i r_j/RC\}/p(p-1)$$

A matrix of the form $aI_p - bJ_{p,p}$ is known to have eigenvalues a with multiplicity p-1 and a-bp with multiplicity 1. This fact, the values for a and b above, and the relations listed at the outset of this proof can be used to verify that the eigenvalues of $\overline{M}(d)$ are as given.

Lemma 2.2. For $d \in D(p,R,C)$

$$\text{tr } \overline{M}^{-1}(d) = 1/m_1(d) + (p-1)/m_2(d)$$

where $m_1(d)$ and $m_2(d)$ are as in lemma 2.1.

pf. This follows immediately from lemma 2.1.

We shall temporarily allow the r_i, r_{ij}, and s_{ik} for $1 \le i \le p$, $1 \le j \le R$, and $1 \le k \le C$ to be arbitrary nonnegative real numbers. We continue to require r_0, the r_{0j}, and the s_{0k} to be nonnegative integers satisfying $0 \le r_{0j} \le C$ for all j and $0 \le s_{0k} \le R$ for all k. In addition we require $\sum_{i=0}^{p} r_i = RC$ and $\sum_{j=1}^{R} r_{ij} = \sum_{k=1}^{C} s_{ik} = r_i$ for $0 \le i \le p$.

Lemma 2.3. For fixed values of r_0, the r_{0j}, the s_{0k}, and the r_i, $m_2(d)$ is maximized when $r_{ij} = r_i/R$ and $s_{ik} = r_i/C$ for all $1 \le i \le p$, $1 \le j \le R$, and $1 \le k \le C$.

pf. This follows immediately from the value of $m_2(d)$ given in lemma 2.1 and the fact that $\sum_{i=1}^{p} x_i^2$ subject to $\sum_{i=1}^{p} x_i = A$ is minimized when $x_i = A/p$ for all i.

Lemma 2.4. For fixed values of r_0, the r_{0j}, and the s_{0k}, $m_2(d)$ is maximized when $r_i = (RC-r_0)/p$ for $1 \le i \le p$ and the r_{ij} and s_{ik} are as in lemma 2.3.

pf. For fixed values of r_0, the r_{0j}, the s_{0k}, and the r_i lemma 2.3 gives us that the maximum value of $m_2(d)$ is

$$m_2(d) = (RC - r_0 - \sum_{i=1}^{p} r_i^2/RC)/(p-1)$$

$$-(r_0 - \{\sum_{j=1}^{R} r_{0j}^2/C\} - \{\sum_{k=1}^{C} s_{0k}^2/R\} + r_0^2/RC)/p(p-1)$$

which is maximized over the r_i when $r_i = (RC - r_0)/p$. The lemma follows.

In what follows, $[\cdot]$ denotes the greatest integer function.

Lemma 2.5. For fixed r_0 with the r_i, the r_{ij} and the s_{ik} for $1 \leq i \leq p$, $1 \leq j \leq R$ and $1 \leq k \leq C$ as in lemmas 2.3 and 2.4, $1/m_1(d) + (p-1)/m_2(d)$ is minimized when the r_{0j}, $1 \leq j \leq R$, are either $[r_0/R]$ or $[r_0/R] + 1$ and the s_{0k}, $1 \leq k \leq C$, are either $[r_0/C]$ or $[r_0/C] + 1$.

pf. Let $Q = \sum_{j=1}^{R} r_{0j}^2/C + \sum_{k=1}^{C} s_{0k}^2/R$. For fixed r_0 with the r_i, r_{ij}, and s_{ik} as in lemmas 2.3 and 2.4 it is easy to show that

$$(2.1) \qquad 1/m_1(d) + (p-1)/m_2(d) = p/(r_0 + r_0^2/RC - Q)$$

$$+ p(p-1)^2/\{p(RC-r_0-(RC-r_0)^2/pRC) - r_0 - r_0^2/RC + Q\}$$

$$= p/(r_0+r_0^2/RC - Q)$$

$$+ p(p-1)^2/\{(p-1)(RC-r_0) - 2r_0^2/RC + Q\}$$

Differentiating this last expression with respect to Q and putting the result over a common denominator yields a ratio whose denominator is positive and whose numerator is, as a function of Q,

$$h(Q) = p\{(p-1)(RC-r_0)-2r_0^2/RC+Q\}^2 - p(p-1)^2\{r_0+r_0^2/RC-Q\}^2$$

This is a quadratic in Q with the coefficient of Q^2 negative. Since $0 \leq r_{0j} \leq C$, $0 \leq s_{0k} \leq R$, $0 \leq r_0 \leq RC$, and $\sum_{j=1}^{R} r_{0j} = \sum_{k=1}^{C} s_{0k} = r_0$ it follows that $0 \leq Q \leq 2RC$. Also $Q = 0$ only if $r_0 = 0$ and $Q = 2RC$ only if $r_0 = RC$. With these observations it is easy to check that $h(0) = p(p-1)^2R^2C^2 > 0$ and $h(2RC) = 0$. Since $h(Q)$ is a quadratic in Q with negative coefficient of Q^2 we must have $h(Q) \geq 0$ for $0 \leq Q \leq 2RC$. This implies $1/m_1(d) + (p-1)/\dot{m}_2(d)$ is increasing in Q and hence is minimized, for fixed r_0, when Q is as small as possible. For fixed r_0, Q is minimized when the r_{0j} and s_{0k}, which must be integers, are as given in the lemma.

<u>Lemma 2.6</u>. For $d \in D(p,R,C)$

$$\begin{aligned} \operatorname{tr} \overline{M}^{-1}(d) &= p/\{r_0 + r_0^2/RC - G(r_0)\} \\ &\quad + p(p-1)^2/\{(p-1)(RC-r_0) - 2r_0^2/RC + G(r_0)\} \\ &\equiv \ell(r_0) \end{aligned}$$

where

$$G(r_0) = \{r_0 + (2r_0-R)\,[r_0/R] - R[r_0/R]^2\}/C$$

$$+ \{r_0 + (2r_0-C)\,[r_0/C] - C[r_0/C]^2\}/R$$

pf. This follows from lemmas 2.2-2.5 in a straightforward manner.

Theorem 2.1. If $d \in D(p,R,C)$ is a completely symmetric design, i.e. if $M(d) = \overline{M}(d)$, satisfying

(i) $r_1 = \ldots = r_p = (RC-r_0)/p$

(ii) $r_{ij} = r_i/R$ for $1 \leq j \leq R$, $1 \leq i \leq p$

(iii) $s_{ik} = r_i/C$ for $1 \leq k \leq C$, $1 \leq i \leq p$

(iv) the r_{0j} are either $[r_0/R]$ or $[r_0/R] + 1$ for $1 \leq j \leq R$

(v) the s_{0k} are either $[r_0/C]$ or $[r_0/C] + 1$ for $1 \leq k \leq C$

(vi) r_0 is the nonnegative integer minimizing $\ell(x)$, where $\ell(x)$ is as given in lemma 2.6

then d is A-optimal over D(p,R,C).

pf. This follows easily from lemmas 2.1-2.6 and the fact that by convexity $\operatorname{tr} M^{-1}(d) \geq \operatorname{tr} \overline{M}^{-1}(d)$.

Application of this theorem involves finding the minimizing r_0 in (vi), perhaps by computer search, then computing the r_{0j}, s_{0k}, r_i, and finally the r_{ij} and s_{ik} according to (i) - (v) in theorem 2.1, and lastly verifying that there exists $d \in D(p,R,C,)$ with these design parameters. Unfortunately, more often than not, no such d exists. Some cases where such a d does exist are given in the following corollary.

<u>Corollary 2.1</u>. For any integer $m \geq 1$, suppose d* is the design obtained from a $m+m^2$ by $m+m^2$ Latin Square by changing treatment labels m^2+1, $m^2+2, \ldots,$ m^2+m to zeroes. Then d* is A-optimal over $D(m^2, m^2+m, m^2+m)$.

<u>pf</u>. In formula (2.1) for tr $\bar{M}^{-1}(d)$ if one allows the r_{0j} and s_{0k} to be possibly nonintegral and argue as in lemmas 2.5 and 2.6, one can show when $p=m^2$ and $R=C=m^2+m$

$$Q = 2r_0^2/(m^2+m)^2$$

and

$$\text{tr } \bar{M}^{-1}(d) = m^2/\{r_0 - r_0^2/(m^2+m)^2\} + m^2(m^2-1)/(\{m^2+m\}^2 - r_0)$$

$$= m^2\{(m+m^2)^2 + r_0(m^2-1)\}/r_0(\{m^2+m\}^2 - r_0)$$

Differentiating with respect to r_0, setting equal to zero, and solving, one can show tr $\bar{M}^{-1}(d)$ is minimized when $r_0 = m^3 + m^2$.

From this fact it is straightforward to verify that the conditions of theorem 2.1 require $r_0 = m^3+m^2$, all $r_{0j} = m$, all $s_{0k} = m$, all $r_{ij} = 1$, all $s_{ik} = 1$, all $r_i = m^2+m$, and that $M(d) = \bar{M}(d)$ for d to be A-optimal over $D(m^2, m^2+m, m^2+m)$. d* as stated in the corollary satisfies these conditions, and so is A-optimal.

For $p = m^2+a$ (or m^2-a), $R = C = m^2+m+a$ (or m^2+m-a), and small a, a design obtain from a R×R Latin Square by changing m of the treatment labels to zero is also likely to be A-optimal. This, however, must be checked in each case by verifying that $r_0 = m^3+m^2+ma$ (or m^3+m^2-ma) does indeed minimize $\ell(r_0)$ in lemma 2.6. For example one can check that such a design is A-optimal over D(2,3,3) using m = 1, a = 1. Likewise such a design is A-optimal over D(3,4,4) using m = 1, a = 2.

For many combinations of p, R, and C theorem 2.1 fails to yield a d in D(p,R,C) which is A-optimal. However it does suggest that if there is a d in D(p,R,C) which almost satisfies the conditions of theorem 2.1, it is likely to be nearly A-optimal. Thus theorem 2.1 may be used to suggest "good designs" under the A-optimality criterion. To find the structure of an A-optimal design in situations where no d exists satisfying theorem 2.1, a more sensitive calculation is needed that allows M(d) not to be completely symmetric. Such calculations are still being investigated.

3. E-optimal designs. Since the investigation of A-optimality in section 2 does not give A-optimal designs for all p, R, and C combinations, we turn our attention to finding E-optimal designs. Here we shall have more success and at the end of this section show that certain E-optimal designs are optimal under criteria involving only variances of treatment-control contrasts.

We begin by recalling that $d \in D(p,R,C)$ is E-optimal if it minimizes the maximum eigenvalue of $M^{-1}(d)$. Equivalently, d is E-optimal if it maximizes the minimum eigenvalue of M(d). We now state and prove a series of lemmas. Notation is as in sections 1 and 2.

Lemma 3.1. For $m_1(d)$ as in lemma 2.1, its maximum value occurs for $d \in D(p,R,C)$ such that

(1) $r_0 - R[r_0/R]$ of the r_{0j} are $[r_0/R] + 1$, the remainder being $[r_0/R]$

(ii) $r_0 - C[r_0/C]$ of the s_{0k} are $[r_0/C] + 1$, the remainder being $[r_0/C]$.

(iii) r_0 is the integer between 1 and RC which maximizes the function

$$f(x) = x/p + x^2/RCp - \{x-R[x/R] + [x/R]\,(2x-R[x/R])\}/Cp - \{x-C[x/C] + [x/C]\,(2x-C[x/C])\}/Rp$$

pf. For fixed r_0, lemma 2.1 displays $m_1(d)$ as a quadratic in the r_{0j} and s_{0k} with negative coefficients. Since the r_{0j} and s_{0k} must be nonnegative integers subject to the linear restriction $\sum_{j=1}^{R} r_{0j} = \sum_{k=1}^{C} s_{0k} = r_0$, it is well known that $m_1(d)$ will be maximized when the r_{0j} are as equal as possible and the s_{0k} are as equal as possible. This, in turn, means the r_{0j} must satisfy (i) above and the s_{0k} (ii) above. Substituting these values into the formula for $m_1(d)$ and rearranging terms yields $f(r_0)$, f as in (iii) above. The lemma follows.

We note that if one replaces $[x/R]$ by x/R, or any value between $[x/R]$ and x/R, in the formula for $f(x)$ given in (iii) of lemma 3.1, $f(x)$ increases in value. This is easily checked using calculus. Similarly if $[x/C]$ is replaced by x/C, or any value in between, in $f(x)$, $f(x)$ increases in value. This fact can be used to show $f(x) \leq x/p - x^2/RCp$. The right hand side of this inequality is maximized by $x = RC/2$. When RC is even, one can show $f(x) = x/p - x^2/RCp$ at $x = RC/2$. Thus $x = RC/2$ maximizes $f(x)$ over all integers x when RC is even. In particular, if R and C are both even the maximum value of $m_1(d)$ will then be achieved when all the $r_{0j} = C/2$ and all the $s_{0k} = R/2$.

If, say, R is even and C is odd the maximum value of $m_1(d)$ is achieved when half of the $r_{0j} = (C-1)/2$, half of the $r_{0j} = (C+1)/2$, and all the $s_{0k} = R/2$.

When RC is odd one can show f(x) achieves its maximum (over integers x) at $x = (RC-1)/2$ and $x = (RC+1)/2$. Both give the maximum.

Lemma 3.2. For $d \in D(p,R,C)$, the minimum eigenvalue of M(d) is $\leq m_1^*$, where m_1^* is the maximum value of $m_1(d)$ given by lemma 3.1.

pf. $\phi(M(d))$ = minimum eigenvalue of M(d) is a concave function over the set of p×p non-negative definite matrices and hence $\phi(M(d)) \leq \phi(\overline{M}(d) \leq m_1(d) \leq m_1^*$.

These lemmas give us the following theorem.

Theorem 3.1. If $d \in D(p,R,C)$ is such that the minimum eigenvalue of M(d) is m_1^*, as given in lemma 2.3, then d is E-optimal over D(p,R,C).

pf. This follows from lemmas 3.1 and 3.2 and the fact that the maximum of the smallest eigenvalue of M(d) equals the minimum of the largest eigenvalue of $M^{-1}(d)$.

A useful special case of this theorem is the following.

Corollary 3.1. Suppose $d \in D(p,R,C)$ is a completely symmetric design (i.e. $M(d) = \overline{M}(d)$) and the r_{0j} and s_{0k} satisfy conditions (i) - (iii) of lemma 3.1. In addition suppose

$$U = \sum_{i=1}^{p} \sum_{\substack{j=1 \\ j \neq i}}^{p} (R\lambda_{ij} + C\mu_{ij} - r_i r_j)$$

is non-negative. Then d is E-optimal.

pf. It is straightforward to use lemma 2.1 and some of the relationships contained in its proof to show $m_2(d) - m_1(d) = U/RC(p-1)$. If $U \geq 0$ then $m_1(d)$ is the minimum eigenvalue of M(d). In addition, $m_1(d) = m^*$ by lemma 3.1. By theorem 3.1 d is E-optimal over D(p,R,C).

Corollary 3.2. Let $d^* \in D(p,2p,2p)$ be the design whose i-th row, for $i \leq p$, has p+1-i zeroes followed by the integers 1,2,...,p in order, and then i-1 zeroes. For i > p, the i-th row is the integers i-p, i-p+1,...,p, followed by p zeroes, and then the integers p+1-i,...,i-p. In other words, the rows of d* are just cyclic permutations of p zeroes followed by the integers 1,2,...,p. d* is E-optimal.

pf. It is straightforward to verify that the conditions of corollary 3.1 are satisfied.

A $d \in D(p,R,C)$ will be said to be a balanced treatment block (B.T.B) design with regard to rows if $\lambda_{01} = \ldots = \lambda_{0p}$ and $\lambda_{12} = \lambda_{13} = \ldots = \lambda_{p-1,p}$. d will be said to be a B.T.B. design with regard to columns if $\mu_{01} = \ldots = \mu_{0p}$ and $\mu_{12} = \mu_{13} = \ldots = \mu_{p-1,p}$. For more information on balanced treatment block designs in an incomplete blocks design setting see Bechhofer and Tamhane (1981) and Notz and Tamhane (1983). The previous corollary can be generalized as follows.

Corollary 3.3. Suppose $d \in D(p,R,C)$ is such that the r_{0j} and s_{0k} satisfy conditions (i) - (iii) of lemma 3.1. Further suppose d is a B.T.B. design both with regard to rows and columns, that $r_1 = \ldots = r_p$, and U in corollary 3.1 is non-negative. Then d is E-optimal.

pf. If d is a B.T.B. design with respect to rows and columns and satisfies $r_1 = \ldots = r_p$ then (1.2) and the fact that the row sums of N(d) are 0 implies N(d) and hence M(d) is completely symmetric. The result now follows from corollary 3.1.

If we denote a design d by an R C matrix whose i, j-th entry is the integer representing the treatment in the plot in row i and column j, we have the following example.

Example. The following design is E-optimal over D(3,2,6)

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 0 & 0 & 0 \end{pmatrix}$$

This is easily verified since $U = 6(2\cdot2+6\cdot0-2\cdot2) = 0$, the r_i all equal 2, d is a B.T.B. design with respect to rows and columns and conditions (i) - (iii) of lemma 3.1 are satisfied.

E-optimal designs which are B.T.B. designs with regard to rows and columns, have $r_1 = \ldots = r_p$, have $U \geq 0$, and satisfy conditions (i) - (iii) of lemma 3.1 shall be called balanced treatment E-optimal (B.T.E.) designs. Such designs have some additional optimality properties.

Theorem 3.2. If $d^* \in D(p,R,C)$ is a B.T.E. design then it minimizes the variance of the B.L.U.E. of any test treatment-control contrast proportional to $\alpha_0 - \sum_{i=1}^{p} \alpha_i/p$ over D(p,R,C,).

pf. It is well known that $\phi(M) = M^{-1}$ is convex over the set of p×p positive definite matrices with the ordering $M_1 > M_2$ if and only if $M_1 - M_2$ is nonnegative definite. Since $M^{-1}(d)$, $d \in D(p,R,C)$, is the information matrix for the vector of contrasts $(\alpha_0-\alpha_1,\ldots,\alpha_0-\alpha_p)'$, it follows that the variance of the B.L.U.E.

of any contrast proportional to $\alpha_0 - \sum_{i=1}^{p} \alpha_i/p$ is proportional to $\underline{1}'M^{-1}(d)\underline{1}$, where $\underline{1}$ is the p×1 vector all of whose entries are +1. Now using the convexity of M^{-1} and the fact that $\underline{1}$ is invariant under permutations of coordinates, one gets for any $d \in D(p,R,C)$

$$\underline{1}'M^{-1}(d)\underline{1} \geq \underline{1}'\overline{M}^{-1}(d)\underline{1}$$

$$= p/m_1(d)$$

where $m_1(d)$ is as in lemma 2.1. Note that the last equality comes from the fact that the eigenvalue $1/m_1(d)$ of $\overline{M}^{-1}(d)$ has eigenvector $\underline{1}$, a consequence of the completely symmetric structure of $\overline{M}^{-1}(d)$.

Since d* is B.T.E., by Corollary 3.1 the maximum eigenvalue of $M^{-1}(d^*)$ is $1/m_1(d^*)$. Furthermore, lemma 3.1 gives $1/m_1(d) \geq 1/m_1(d^*)$ and hence

$$\underline{1}'M^{-1}(d)\underline{1} \geq p/m_1(d)$$

$$\geq p/m_1(d^*)$$

$$= \underline{1}'M^{-1}(d^*)\underline{1}$$

The theorem now follows.

The next theorem shows that a B.T.E. design is minimax with respect to the variances of the B.L.U.E.s of all test treatment-control contrasts.

Theorem 3.3. If $d^* \in D(p,R,C)$ is a B.T.E. design then it minimizes over D(p,R,C) the maximum possible variance of the B.L.U.E.s of all test treatment-control contrasts $a_0\alpha_0 - \sum_{i=1}^{p} a_i\alpha_i$ where $\sum_{i=1}^{p} a_i = a_0 \neq 0$ and $\vec{a} = (a_1,\ldots,a_p)'$ is a p×1 unit vector.

pf. Let $\mathcal{A}$ be the set of all p×1 unit vectors the sum of whose coordinates is nonzero. Clearly there is a one-to-one correspondence between $\vec{a} = (a_1,\ldots,a_p)' \in \mathcal{A}$ and test treatment-control contrasts $a_0\alpha_0 - \sum_{i=1}^{p} a_i\alpha_i$ where $\sum_{i=1}^{p} a_i = a_0 \neq 0$ and $(a_1,\ldots,a_p)'$ is a p×1 unit vector. In fact, for any $d \in D(p,R,C)$, the variance of the B.L.U.E. of $a_0\alpha_0 - \sum_{i=1}^{p} a_i\alpha_i$, where $\sum_{i=1}^{p} a_i = a_0 \neq 0$ and $\vec{a} = (a_1,\ldots,a_p)'$ is a unit vector, is proportional to $\vec{a}'M^{-1}(d)\vec{a}$. Hence the maximum possible variance of the B.L.U.E.s of such test treatment-control contrasts for any $d \in D(p,R,C)$ is proportional to

$$\sup_{\vec{a} \in \mathcal{A}} \vec{a}'M^{-1}(d)\vec{a} \geq (\underline{1}/\sqrt{p})'M^{-1}(d)(\underline{1}/\sqrt{p})$$

$$\geq \underline{1}'\overline{M}^{-1}(d)\underline{1}/p$$

$$\geq \underline{1}'M^{-1}(d^*)\underline{1}/p$$

$$= 1/m_1(d^*)$$

using theorem 3.2. Since $1/m_1(d^*)$ is the maximum eigenvalue of $M^{-1}(d^*)$ and hence $= \sup_{\vec{a} \in \mathcal{A}} \vec{a}'M^{-1}(d^*)\vec{a}$, the theorem follows.

Theorems 3.2 and 3.3 indicate that B.T.E. designs are optimal in meaningful ways with regard to test treatment-control contrasts. These theorems provide a link that makes E-optimality meaningful for treatment-control contrast design problems.

References

1. Bechhofer, R. E. and Tamhane, A. C. (1981). Incomplete block designs for comparing treatments with a control: General theory. Technometrics, 23, 45-57.

2. Kiefer, J. (1958). On the nonrandomized optimality and randomized nonoptimality of symmetrical designs. Ann. Math. Statist., 29, 675-699.

3. Kiefer, J. (1959). Optimum experimental designs. J. Roy. Statist. Soc. Ser. B, 21, 272-319.

4. Kiefer, J. (1971). The role of symmetry and approximation in exact design optimality. In Statistical Decision Theory and Related Topics (S.S. Gupta and J. Yackel, eds.) Academic, New York, 109-118.

5. Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). Ann. Statist., 2, 849-879.

6. Kiefer, J. (1975). Construction and optimality of generalized Youden designs. A Survey of Statistical Design and Linear Models (J. Srivastava ed.) North Holland, New York, 333-353.