DS-optimal designs in one way ANOVA

Rita SahaRay and Subir Kumar Bhandari

Stat-Math Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata-700 108, India
(e-mail: rita@isical.ac.in, subir@isical.ac.in)

Abstract. Characterization and construction of optimal designs using the familiar optimality criteria, for example A-, D- and E-optimality are well studied in the literature. However the study of the Distance Optimality (DS-) criterion introduced by Sinha (1970) has very recently drawn attention of researchers. In the present article, we consider the singularly estimable full rank problem of estimating the full set of elementary treatment contrasts using the DS optimality criterion in the set up of a one way ANOVA model. Using a limit argument it turns out that a CRD in which difference between any two allocation numbers is at the most unity is uniquely DS-optimal.

AMS Subject Classification (1970): 62K05

Key words: DS-optimality criterion, majorization, weakly supermajorization, one-way ANOVA model, Okamoto Lemma

1 Introduction

The problem of characterization and construction of optimal designs under both discrete and continuous set up using the well known A-, D- and E-optimality criteria has been extensively studied in the literature. See for example, Shah and Sinha (1989), Pukelsheim (1993). However, the study of the distance optimality criterion put forward by Sinha (1970) has received relatively less attention. Recently there has been a growing interest in this direction (cf. Liski, Luoma, Mandal and Sinha (1998); Liski, Luoma and Zaigraev (1999); Mandal, Shah and Sinha (2000)).

We start with a classical linear model

$$\underline{Y} \sim N(X\underline{\beta}, \sigma^2 I_N),$$

where the $N \times 1$ response vector $\underline{Y} = (Y_1, \ldots, Y_N)'$ follows a multivariate normal distribution, $X = (\underline{X}_1, \ldots, \underline{X}_N)'$ is the $N \times k$ design matrix, and $\underline{\beta} = (\beta_1, \ldots, \beta_k)'$ is the $k \times 1$ parameter vector. $E(\underline{Y}) = X\underline{\beta}$, and $D(\underline{Y}) = \sigma^2 I_N$ are respectively the expectation vector and the dispersion matrix of $\underline{Y}$. Since σ^2 is irrelevant to the results derived in this paper, for the sake of simplicity, in the remaining part of the paper we assume that $\sigma^2 = 1$.

Let $\eta_{l\times 1} = L_{l\times k}\beta_{k\times 1}$ be the vector of the linear parametric functions of interest to us. We confine only to the class $\mathscr{C}$ of the designs d (i.e. the so called design matrix X_d) under which all the components of $\underline{\eta}$ are estimable. Let the Best Linear Unbiased Estimator (BLUE) of $\underline{\eta}$ using the design d be denoted by $\hat{\underline{\eta}}_d$,

$$\hat{\underline{\eta}}_d = L\hat{\underline{\beta}}_d,$$

where $\hat{\underline{\beta}}_d$ is a Least Squares Estimator (LSE) of $\underline{\beta}$ using the design d. We are interested in characterizing an experimental design d_* which maximises the probability

$$P_\varepsilon = Pr[\|\hat{\underline{\eta}}_d - \underline{\eta}\| < \varepsilon] \quad \forall \varepsilon > 0 \tag{1.1}$$

over the class $\mathscr{C}$ of all competing designs d where $\|\hat{\underline{\eta}}_d - \underline{\eta}\| = [(\hat{\underline{\eta}}_d - \underline{\eta})'(\hat{\underline{\eta}}_d - \underline{\eta})]^{1/2}$, the Euclidian norm of $\hat{\underline{\eta}}_d - \underline{\eta}$. As this criterion aims at minimising the distance between the true parameter value and its estimate in a stochastic sense, it is abbreviated as the DS-optimality criterion in the literature. A design d_* is said to be $DS(\varepsilon)$ optimal for the LSE of $\underline{\eta}$ if for a given $\varepsilon > 0$, it maximizes the probability $Pr[\|\hat{\underline{\eta}}_d - \underline{\eta}\| < \varepsilon]$. When d_* is $DS(\varepsilon)$ optimal for all $\varepsilon > 0$, we say that d_* is DS-optimal. Sinha (1970) introduced this criterion for optimal allocation of observations with a given total in a CRD model:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1, \ldots, v;\ j = 1, \ldots, n_i, \tag{1.2}$$

where $\underline{\tau} = (\tau_1, \ldots, \tau_v)'$ is the vector of treatment effects, the ith treatment being allocated n_i times, $n_i \geq 1$, $1 \leq i \leq v$, $\sum_{i=1}^{v} n_i = n$. In that paper, the parametric vector of interest is the mean vector $\underline{\eta} = (\mu + \tau_1, \ldots, \mu + \tau_v)'$ and the 'symmetrical allocation' with $n_i = n/v$, $\forall i$ is shown to be uniquely DS-optimal when n is divisible by v. In this context, the general case when n is not divisible by v is implicitly resolved in a recent work (Liski et al. (1998)). By repeated application of Lemma 2 in that paper it can be shown that whenever allocation numbers for a pair of treatments differ by more than 1, successively reducing thier difference by 2, but keeping the total fixed, a better design can be obtained and finally, a most symmetrical allocation with $|n_i - n_j| \leq 1$, $\forall i \neq j$, turns out to be DS-optimal as is expected. Mandal, Shah and Sinha (2000) considers the problem of comparison of one test treatment with a set of v control treatments using this optimality criterion in both the CRD and the Block design settings. All the problems considered so far are nonsingularly estimable full rank problem. Our purpose in this paper is to take up the singularly estimable full rank problem of estimating the full set of elementary treatment contrasts which is of practical interest in a one-way ANOVA model and characterize a DS-optimal design. We organise our paper as follows.

In Section 2 we elaborate the distance optimality criterion in the context of the problem just mentioned and quote relevant results on majorization and matrix theory which are useful in deriving the main result. In Section 3 DS-optimal designs are characterized for all v.

2 Preliminaries

We assume the one way ANOVA model (1.2). The set of all elementary contrasts of the form $\tau_i - \tau_j$, $i < j$ *viz*.

$$\underline{\eta}' = (\tau_1 - \tau_2, \ldots, \tau_1 - \tau_v, \ldots . \tau_{v-1} - \tau_v)$$

is of interest to us. For a design $d \in \mathscr{C}$ let n_{di} denote the allocation number of the *ith* treatment, $i = 1, \ldots, v$ with $\sum_{i=1}^{v} n_{di} = n$. However, in the sequel, we sometimes skip the suffix d to avoid the notational complexity. Let $\bar{y}_{i.} = \sum_{j=1}^{n_{di}} y_{ij}/n_{di}$ stand for the mean of all observations receiving the *ith* treatment, $i = 1, \ldots, v$ under the design d. Then $\hat{\underline{\eta}}'_d = (\bar{y}_{1.} - \bar{y}_{2.}, \ldots, \bar{y}_{1.} - \bar{y}_{v.}, \ldots, \bar{y}_{(v-1).} - \bar{y}_{v.})$ is the BLUE of $\underline{\eta}'$ using the design d under the model (1.2).

Let P be a $v - 1 \times v$ submatrix of an orthogonal $v \times v$ matrix such that

$$P_{v-1\times v} P'_{v\times v-1} = I_{v-1}, \quad P'P = (I - J/v), \tag{2.1}$$

$$D = Diag(1/n_{d1}, \ldots, 1/n_{dv}),$$
$$\text{and} \quad D^{1/2} = Diag(1/\sqrt{n_{d1}}, \ldots, 1/\sqrt{n_{dv}}). \tag{2.2}$$

Writing $\underline{\eta} = L\underline{\tau}$, so that $\hat{\underline{\eta}}_d = L\hat{\underline{\tau}}_d$, where $\hat{\underline{\tau}}_d = (\bar{y}_{1.}, \ldots, \bar{y}_{v.})$, we have from (1.1)

$$\begin{aligned} P_\varepsilon &= Pr[(\hat{\underline{\eta}}_d - \underline{\eta})'(\hat{\underline{\eta}}_d - \underline{\eta}) \le \varepsilon^2] \\ &= Pr[(\hat{\underline{\tau}}_d - \underline{\tau})'L'L(\hat{\underline{\tau}}_d - \underline{\tau}) \le \varepsilon^2] \\ &= Pr[(\hat{\underline{\tau}}_d - \underline{\tau})'(I - J/v)(\hat{\underline{\tau}}_d - \underline{\tau}) \le \varepsilon^2/v] \\ &= Pr[(\hat{\underline{\tau}}_d - \underline{\tau})'P'P(\hat{\underline{\tau}}_d - \underline{\tau}) \le \varepsilon^2/v] \\ &= Pr[(\underline{\xi}'\underline{\xi}) \le \varepsilon^2/v], \end{aligned} \tag{2.3}$$

where $\underline{\xi} = P(\hat{\underline{\tau}}_d - \underline{\tau}) \sim N_{v-1}(0, \Sigma)$ with $\Sigma = PDP'$.

It is not hard to verify that for any $d \in \mathscr{C}$, $\Sigma = PDP'$ is nonsingular. Let $\underline{\lambda}_d = (\lambda_{d1}, \ldots, \lambda_{d,(v-1)})'$ denote the vector of ordered eigenvalues of PDP' where $\lambda_{d1} \le \lambda_{d2} \le \cdots \le \lambda_{d,(v-1)}$. Let $T\Lambda T' = \Sigma$ be the spectral decomposition of Σ, the dispersion matrix of $\underline{\xi}$, where T is an orthogonal $(v-1) \times (v-1)$ matrix and $\Lambda = Diag(\lambda_{d1}, \ldots, \lambda_{d,(v-1)})$ is the diagonal matrix of the eigenvalues of Σ, in other words of PDP'. Define

$$\underline{Z} = \Lambda^{-1/2} T' \underline{\xi},$$

so that

$$\underline{Z} \sim N_{v-1}(0, I_{v-1}).$$

Then from (2.3),

$$Pr[\|\hat{\underline{\eta}}_d - \underline{\eta}\| < \varepsilon] = Pr[\underline{Z}'\Lambda\underline{Z} \leq \varepsilon^2/v]$$
$$= Pr\Big[\sum \lambda_{di} Z_i^2 \leq \delta^2\Big] \qquad (2.4)$$

for $\delta^2 = \varepsilon^2/v$.

Thus $\forall\delta^2 > 0$, the DS(ε) optimality criterion $Pr[\|\widehat{\underline{\eta_d}} - \underline{\eta}\| < \varepsilon]$ depend on the design d with allocation numbers n_{di}, $i = 1, \ldots, v$ only through the eigenvalues $\lambda_{d1}, \ldots, \lambda_{d,(v-1)}$ of the matrix (PDP') where D is as given in (2.2).

Remark 1. Instead of $\underline{\eta}'$ if we had considered the set of all contrasts of the form $\tau_i - \tau_j$, $i \neq j$, the problem would have remained the same except that δ^2 in (2.4) would change to a scalar multiple of it, viz $\delta^2/2$.

We define the $DS(\varepsilon)$ criterion function Ψ_ε or equivalently $\Psi_\delta(\underline{\lambda}_d)$ as

$$\Psi_\delta(\underline{\lambda}_d) = Pr\left[\sum_{i=1}^{v-1} \lambda_{di} Z_i^2 \leq \delta^2\right]. \qquad (2.5)$$

A design $d_* \in \mathscr{C}$ is said to be DS-optimal if

$$\Psi_\delta(\underline{\lambda}_{d_*}) \geq \Psi_\delta(\underline{\lambda}_d), \quad \forall\delta > 0 \text{ and } \forall d \in \mathscr{C}. \qquad (2.6)$$

In order to characterize a DS-optimal design in the case when n is a multiple of v we make use of the following Lemma due to Okamoto (1970), proof of which can be found in Marshall & Olkin (1979, p. 303).

Lemma 2.1. *Let $Z_1^2, \ldots, Z_p^2$ be independent random variables and suppose that Z_i^2 has a χ^2 distribution with n_i degrees freedom, $i = 1, \ldots, p$. If $\lambda_i > 0$, $i = 1, \ldots, p$ then*

$$Pr\Big(\sum \lambda_i Z_i^2 \leq \delta^2\Big) \leq Pr(\lambda Z \leq \delta^2),$$

where $\lambda = (\prod_{i=1}^p \lambda_i^{n_i})^{1/n}$, $n = \sum_{i=1}^p n_i$ and Z has a χ^2 distribution with n degrees of freedom.

In the case when n is not divisible by v, the notion of majorization proves very useful in the study of the function $\Psi_\delta(\underline{\lambda}_d)$. Majorization concerns the diversity of the components of a vector (cf. Marshall & Olkin (1979, p. 5)). For a ready reference we quote below the relevant definitions and results on majorization used in this paper.

Let $\underline{a} = (a_1, \ldots, a_p)'$, and $\underline{b} = (b_1, \ldots, b_p)'$ be two $p \times 1$ vectors and $a_{(1)} \leq \cdots \leq a_{(p)}$, $b_{(1)} \leq \cdots \leq b_{(p)}$ be the ordered components.

Definition 2.2: *For $\underline{a}, \underline{b} \in \mathbb{R}^p$, $\underline{a}$ is said to majorize $\underline{b}$, written $\underline{a} \succ \underline{b}$ if*

$$\left.\begin{aligned} \sum_{i=1}^{k} a_{(i)} &\leq \sum_{i=1}^{k} b_{(i)} \quad k = 1, \ldots, p-1, \\ \sum_{i=1}^{p} a_{(i)} &= \sum_{i=1}^{p} b_{(i)}. \end{aligned}\right\} \tag{2.7}$$

Definition 2.3: *For $\underline{a}, \underline{b} \in \mathbb{R}^p$, $\underline{a}$ is said to weakly supermajorize $\underline{b}$, written $\underline{a} \,{}^{w}\!\!\succ \underline{b}$ if*

$$\sum_{i=1}^{k} a_{(i)} \leq \sum_{i=1}^{k} b_{(i)} \quad k = 1, \ldots, p. \tag{2.8}$$

Definition 2.4: *A function $f(\underline{x}) : \mathbb{R}^p \to \mathbb{R}$ is said to be a Schur Concave function if for $\underline{x}, \underline{y} \in \mathbb{R}^p$ the relation $\underline{x} \succ \underline{y}$ implies $f(\underline{x}) \leq f(\underline{y})$. Thus the value of $f(\underline{x})$ becomes greater when the components of $\underline{x}$ become less diverse.*

Let $\underline{\lambda}_d^{-1}$ denote the vector of reciprocals of the elements of $\underline{\lambda}_d = (\lambda_{d1}, \ldots, \lambda_{d,(v-1)})'$. For any two designs d_1 and $d_2 \in \mathscr{C}$ let $\underline{a} = \underline{\lambda}_{d_1}^{-1}$ and $\underline{b} = \underline{\lambda}_{d_2}^{-1}$ denote respectively the vectors of eigenvalues of $(PD_1P')^{-1}$ and $(PD_2P')^{-1}$. Using Proposition 7.4.2 of Tong (1990), the following theorem regarding Schur concavity of the DS-optimality criterion is proved in Liski, Luoma & Zaigraev (1999).

Theorem 2.5. *If $\underline{a} \succ \underline{b}$ and $\underline{Z} \sim N_{v-1}(0, I_{v-1})$, then*

$$Pr\left(\sum \frac{Z_i^2}{a_i} \leq \delta^2\right) \leq Pr\left(\sum \frac{Z_i^2}{b_i} \leq \delta^2\right) \quad \forall \delta > 0.$$

The following result (Marshall & Olkin p. 11) plays a crucial role in the determination of main results of this paper.

Theorem 2.6. *If $\underline{a} \,{}^{w}\!\!\succ \underline{b}$, then there exists a vector $\underline{a_0}$ such that $\underline{a_0} \geq \underline{a}$ and $\underline{a_0} \succ \underline{b}$.*

The next theorem is an immediate consequence of the above two theorems.

Theorem 2.7. *If $\underline{a} \,{}^{w}\!\!\succ \underline{b}$ and $\underline{Z} \sim N_{v-1}(0, I_{v-1})$*

$$Pr\left(\sum \frac{Z_i^2}{a_i} \leq \delta^2\right) \leq Pr\left(\sum \frac{Z_i^2}{b_i} \leq \delta^2\right) \quad \forall \delta > 0.$$

Proof:

$$Pr\left(\sum \frac{Z_i^2}{a_i} \leq \delta^2\right) \leq Pr\left(\sum \frac{Z_i^2}{a_{0i}} \leq \delta^2\right) \leq Pr\left(\sum \frac{Z_i^2}{b_i} \leq \delta^2\right) \quad \forall \delta > 0.$$

The first inequality follows from the fact that the event $E_1 : \left\{ \underline{Z} : \sum \frac{Z_i^2}{a_i} \leq \delta^2 \right\}$ implies the event $E_2 : \left\{ \underline{Z} : \sum \frac{Z_i^2}{a_{di}} \leq \delta^2 \right\}$ as $a_{di} \geq a_i$, $\forall i = 1, \ldots, v-1$. The last inequality now follows from Theorem 2.5.

The following result involving the eigenvalues of a matrix will be used in the course of derivation of main results.

Result 2.1. If A is an $m \times n$ matrix and B is an $n \times m$ matrix where $m \leq n$, then the n eigenvalues of BA are the m eigenvalues of AB together with $n - m$ zeros.

The proof of the following theorem on the comparisons of eigenvalues can be found in Marshall Olkin (1979, p. 245).

Theorem 2.8. *If A and B are $m \times m$ complex matrices such that $aA + bB$ has real eigenvalues for all $a, b \in \mathbb{R}$ then*

$$(\lambda_1(A+B), \ldots, \lambda_m(A+B)) \prec (\lambda_1(A) + \lambda_1(B), \ldots, \lambda_m(A) + \lambda_m(B)).$$

In particular, for any real number $0 < \alpha < 1$,

$$\begin{aligned}&(\lambda_1(\alpha A + (1-\alpha)B), \ldots, \lambda_m(\alpha A + (1-\alpha)B)) \\ &\qquad \prec (\alpha\lambda_1(A) + (1-\alpha)\lambda_1(B), \ldots, \alpha\lambda_m(A) + (1-\alpha)\lambda_m(B)).\end{aligned}$$

Using the fact that the eigenvalues depend continuously on the entries of a matrix, the following theorem is immediate.

Theorem 2.9. *For any two $m \times m$ positive semidefinite matrices A and B*

$$\lim_{\theta \to 0} \lambda_i(A + \theta B) = \lambda_i(A) \quad \forall i = 1, \ldots, m.$$

3 Main results

In this section we present main results concerning the characterization of a DS-optimal design for estimation of η in a one way ANOVA model.

It is interesting to note that for any design $d \in \mathscr{C}$, the positive eigenvalues of (PDP'), denoted by λ_{di}, $i = 1, \ldots, v-1$ do not depend on the choice of the P matrix where $P'P = I - J/v$ and $PP' = I_{v-1}$. Define $A = PD^{1/2}$ and $B = A'$. Noting that $AB = PDP'$ is nonsingular but BA is singular, the following theorem is immediate as an application of Result 2.1.

Theorem 3.1. *The +ve eigenvalues of PDP' and $D^{1/2}(I - J/v)D^{1/2}$ are equal.*

Furthermore, using Lemma 2.2 of Bischoff (1995) it directly follows that $|PDP'| = |D|.|\frac{1}{v}\underline{1}'D^{-1}\underline{1}|$, and hence the next theorem.

Theorem 3.2. $|PDP'| = \prod_{i=1}^{v-1} \lambda_{di} = \frac{n}{v\prod_{i=1}^{v} n_{di}}$.

In the sequel, we first deal with the case when n is divisible by v.

Theorem 3.3. *When n is divisible by v, the CRD $d_* \in \mathscr{C}$ with symmetrical allocation viz. $n_{d_*i} = n/v, \forall i = 1, \ldots v$ is DS-optimal for estimating $\underline{\eta}$, the vector of elementary contrasts of treatment effects.*

Proof: For any design $d \in \mathscr{C}$, from (2.4) using Lemma 2.1 and Theorem 3.2 we get

$$Pr\left(\sum_{i=1}^{v-1} \lambda_{di} Z_i^2 \leq \delta^2\right) \leq Pr\left[\left(\prod_{i=1}^{v-1} \lambda_{di}\right)^{1/(v-1)} \chi^2_{v-1} \leq \delta^2\right]$$

$$= Pr\left[\left(n\Big/\left(v\prod_{i=1}^{v} n_{di}\right)\right)^{1/v-1} \chi^2_{v-1} \leq \delta^2\right].$$

Let $d_* \in \mathscr{C}$ correspond to the design with $n_{d_*i} = n/v$, $\forall i = 1, \ldots v$. Then the eigenvalues of $PD_*P' = \frac{v}{n}I$ are all equal. Now using the condition that $\sum_{i=1}^{v} n_{di} = n$ and the well known inequality between the Arithmetic Mean (A.M) and the Geometric Mean (G.M) of a set of positive quantities, viz $A.M \geq G.M$ we get

$$\prod n_{di} \leq \left(\sum n_{di}/v\right)^v = (n/v)^v = \left(\prod n_{d_*i}\right).$$

So

$$Pr\left(\sum_{i=1}^{v-1} \lambda_{di} Z_i^2 \leq \delta^2\right) \leq Pr\left[\left(n\Big/\left(v\prod_{i=1}^{v} n_{di}\right)\right)^{1/v-1} \chi^2_{v-1} \leq \delta^2\right]$$

$$\leq Pr\left[\left(n\Big/\left(v\prod_{i=1}^{v} n_{d_*i}\right)\right)^{1/v-1} \chi^2_{v-1} \leq \delta^2\right]$$

(follows by implication of events)

$$= Pr\left[\left(\prod_{i=1}^{v-1} \lambda_{d_*i}\right)^{1/(v-1)} \chi^2_{v-1} \leq \delta^2\right]$$

$$= Pr\left(\sum_{i=1}^{v-1} \lambda_{d_*i} Z_i^2 \leq \delta^2\right) \quad \text{(as } \lambda_{d_*i}\text{'s are all equal).}$$

Thus whenever n is divisible by v, the design d_* with symmetrical allocation turns out to be DS-optimal.

The case when n is not divisible by v is dealt below. From now onwards we assume that $n = vk + t$, $t \geq 1$, where $k = [n/v]$ denotes the greatest integer less than or equal to n/v. In this case, it needs a close argument to reveal that a most symmetrical allocation viz. the design $d_* \in \mathscr{C}$ with $\underline{n}_{d_*} = (k, \ldots, k, k+1, \ldots, k+1)'$, k occuring $v - t$ times is DS-optimal.

Let

$$D_* = Diag(\underbrace{1/k, \ldots, 1/k}_{v-t \ times}, \underbrace{1/(k+1), \ldots, 1/(k+1)}_{t \ times}),$$

$$\text{and} \quad A = P'P = (I - J/v).$$

We start with the comparison of the vector of eigenvalues of $D^{-1}(A + \varepsilon^* I)$ and $D_*^{-1}(A + \varepsilon^* I)$ where $\varepsilon^* \in \mathbb{R}$, $\varepsilon^* \neq 0$ and -1.

Theorem 3.4. $\underline{\lambda}(D^{-1}(A + \varepsilon^* I)) \succ \underline{\lambda}(D_*^{-1}(A + \varepsilon^* I))$, $\forall \varepsilon^* \in \mathbb{R}, \varepsilon^* \neq 0$ *and* -1.

Proof: We first note that, for any $\varepsilon \in \mathbb{R}$, $\varepsilon \neq 0$ and -1, $A + \varepsilon I$ is a nonsingular matrix. Writing $D^{-1/2} = Diag(\sqrt{n_{d1}}, \ldots, \sqrt{n_{dv}})$, as an application of Result 2.1, we observe that the eigenvalues of $D^{-1/2}(A + \varepsilon I)D^{-1/2}$ and $D^{-1}(A + \varepsilon I)$ are identical. It is clear that for any design d $(\neq d_*) \in \mathscr{C}$, there exists at least one pair of treatment symbols i' and j' such that $(n_{di'} - n_{dj'}) \geq 2$ and $\sum n_{di} = n$. We permute i' and j' treatment symbols, keeping others fixed and obtain $\tilde{d} \in \mathscr{C}$ as

$$n_{\tilde{d}i} = n_{di} \quad \forall i \neq i', j'$$

$$n_{\tilde{d}i'} = n_{dj'}$$

$$n_{\tilde{d}j'} = n_{di'}$$

and hence $\tilde{D}^{-1} = QD^{-1}Q'$, where Q represents the corresponding permutation matrix. In view of the relation $Q'Q = QQ' = I$, and $Q(A + \varepsilon^* I)Q' = A + \varepsilon^* I$,

$$\underline{\lambda}(\tilde{D}^{-1}(A + \varepsilon^* I)) = \underline{\lambda}(D^{-1}(A + \varepsilon^* I))$$

as

$$\begin{aligned} \tilde{D}^{-1}(A + \varepsilon^* I) &= QD^{-1}Q'(A + \varepsilon^* I) = QD^{-1}Q'Q(A + \varepsilon^* I)Q' \\ &= QD^{-1}(A + \varepsilon^* I)Q'. \end{aligned} \tag{3.1}$$

Now it is easy to see that for some $0 < \alpha < 1$, $(n_{di'} - 1, n_{dj'} + 1)$ can be represented as a convex combination of $(n_{di'}, n_{dj'})$ and $(n_{dj'}, n_{di'})$. Choosing this α, it follows that

$$\begin{aligned}\underline{\lambda}(D^{-1}(A+\varepsilon^* I)) &= \alpha\underline{\lambda}(D^{-1}(A+\varepsilon^* I)) + (1-\alpha)\underline{\lambda}(QD^{-1}(A+\varepsilon^* I)Q') \\ &= \alpha\underline{\lambda}(D^{-1}(A+\varepsilon^* I)) + (1-\alpha)\underline{\lambda}(\tilde{\boldsymbol{D}}^{-1}(A+\varepsilon^* I)) \\ &\succ \underline{\lambda}(\alpha D^{-1}(A+\varepsilon^* I) + (1-\alpha)\tilde{\boldsymbol{D}}^{-1}(A+\varepsilon^* I)) \\ &\quad \text{(using Theorem 2.8)} \\ &= \underline{\lambda}((\alpha D^{-1} + (1-\alpha)\tilde{\boldsymbol{D}}^{-1})(A+\varepsilon^* I)) \\ &= \underline{\lambda}(D_0^{-1}(A+\varepsilon^* I)), \end{aligned} \tag{3.2}$$

where

$$D_0^{-1} = \alpha D^{-1} + (1-\alpha)\tilde{\boldsymbol{D}}^{-1}$$

$$= Diag(n_{d1}, \ldots, n_{d(i'-1)}, n_{di'} - 1, n_{d(i'+1)}, \ldots, n_{d(j'-1)}, n_{dj'} + 1, n_{d(j'+1)} \ldots, n_{dv}).$$

Thus when the pair of allocation numbers $(n_{di'}, n_{dj'})$ is transferred to a pair $(n_{di'} - 1, n_{dj'} + 1)$, reducing their mutual difference by two, but keeping the total fixed, we get the above result (3.2) on majorization of the eigenvalues of $D^{-1}(A+\varepsilon^* I)$. Note that starting from D^{-1} successive averaging by taking convex combination of any two co-ordinates of $\underline{n} = (n_{d1}, \ldots n_{dv})'$ in the above sense, while keeping the rest of the co-ordinates fixed, we will eventually get D_*^{-1} or $Q_* D_*^{-1} Q_*'$ where Q_* is a permutation matrix and similar successive steps of majorization will yield

$$\underline{\lambda}(D^{-1}(A+\varepsilon^* I)) \succ \underline{\lambda}(D_0^{-1}(A+\varepsilon^* I)) \succ \cdots \succ \underline{\lambda}(D_*^{-1}(A+\varepsilon^* I)). \tag{3.3}$$

Remark 3.1. The convex coefficient α may change at different steps.

Remark 3.2. The condition of Theorem 2.8 that for any two real numbers a and b $(aD^{-1}(A+\varepsilon^* I) + b\tilde{\boldsymbol{D}}^{-1}(A+\varepsilon^* I))$ has real eigenvalues is trivially satisfied as the required matrix has the same set of eigenvalues as that of the real symmetric matrix $((\sqrt{(1+\varepsilon^*)}A + \sqrt{\varepsilon^*}J/v)(aD^{-1} + b\tilde{\boldsymbol{D}}^{-1})((\sqrt{(1+\varepsilon^*)}A + \sqrt{\varepsilon^*}J/v)$.

In the remainig portion of the paper, using Theorem 3.4 we establish the required result on weak supermajorization and characterize the DS-optimal design whenever n is not a multiple of v.

Theorem 3.5. *Let $n = vk + t$ and $d_* \in \mathscr{C}$ be a CRD with $\underline{n}_{d_*} = (k, \ldots, k, k+1, \ldots, k+1)'$, k occuring $v - t$ times. Then d_* is DS-optimal.*

Proof: In order to establish d_* to be DS-optimal, in view of (2.4) and Theorem 2.7, it suffices to show that

$$\underline{\lambda}_d^{-1} \overset{w}{\succ} \underline{\lambda}_{d_*}^{-1}$$

where $\underline{\lambda}_d^{-1}$ and $\underline{\lambda}_{d_*}^{-1}$ denote respectively the vectors of eigenvalues of $(PDP')^{-1}$ and $(PD_*P')^{-1}$. Not to obscure the essential steps, we first note that (3.3) yields

$$\underline{\lambda}^{-1}(D(A+\varepsilon^* I)^{-1}) \succ \underline{\lambda}^{-1}(D_*(A+\varepsilon^* I)^{-1}). \tag{3.4}$$

It is easy to check that as $A = I - J/v$,

$$(A + \varepsilon^* I)^{-1} = \frac{1}{1+\varepsilon^*}\left(A + \frac{1+\varepsilon^*}{\varepsilon^*} J/v\right), \quad \varepsilon^* \in \mathbb{R}, \varepsilon^* \neq 0 \text{ and } -1. \tag{3.5}$$

Call $\frac{1+\varepsilon^*}{\varepsilon^*} = \theta$. Thus for any $\theta > 0$, (3.4) can be rewritten as

$$\underline{\lambda}^{-1}(D(A + \theta J/v)) \succ \underline{\lambda}^{-1}(D_*(A + \theta J/v)). \tag{3.6}$$

Applying Theorem 2.9

$$\lim_{\theta \to 0} \lambda_1(D(A + \theta J/v)) = 0,$$

$$\text{and} \quad \lim_{\theta \to 0} \lambda_i(D(A + \theta J/v)) = \lambda_i(DA), \quad i \neq 1.$$

Recalling $A = P'P$ and Result 2.3, we note that $\lambda_i(DA) = \lambda_{i-1}(PDP')$, $\forall i = 2, \ldots, v$. Thus from (3.6) using the first $(v-1)$ inequalilties discussed in (2.7), and taking limit we conclude that

$$\underline{\lambda}_d^{-1} = (1/\lambda_{d1}, \ldots, 1/\lambda_{d(v-1)}) \,{}^{w}\!\!\succ (1/\lambda_{d_*1}, \ldots, 1/\lambda_{d_*(v-1)}) = \underline{\lambda}_{d_*}^{-1} \tag{3.7}$$

and hence the theorem.

Remark 3.3. Using similar arguments as given in (3.1) $\underline{\lambda}(D_*^{-1}(A + \varepsilon^* I)) = \underline{\lambda}(Q_* D_*^{-1} Q_*'(A + \varepsilon^* I))$ and hence d_* as well as any permutation of d_* is DS-optimal.

Remark 3.4. The design d_* upto permutation, is uniquely DS-optimal since if any other DS-optimal design d_{0*} exists

$$E\left(\sum Z_i^2 \lambda_{d_*i}\right) = E\left(\sum Z_i^2 \lambda_{d_{0*}i}\right).$$

But clearly

$$E\left(\sum Z_i^2 \lambda_{d_*i}\right) = \sum \lambda_{d_*i} = \left(1 - \frac{1}{v}\right) \sum \frac{1}{n_{d_*i}}$$
$$< \left(1 - \frac{1}{v}\right) \sum \frac{1}{n_{di}} = E\left(\sum Z_i^2 \lambda_{di}\right)$$

for any other design $d \in \mathscr{C}$.

Acknowledgement. We are thankful to Prof. B. K. Sinha for suggesting the problem to us and many helpful discussions with him. We are also thankful to the referees for drawing our attention to the reference of Bischoff (1995) and many helpful comments.

References

1. Bischoff W (1995) Determinant formulas with applications to designing when the observations are correlated. Ann Inst Statist Math 47:385–399
2. Liski EP, Luoma A, Mandal NK, Sinha BK (1998) Pitman nearness, distance criterion and optimal regression designs. Calcutta Statistical Association Bulletin 48(191–192):179–194
3. Liski EP, Luoma A, Zaigraev A (1999) Distance optimality design criterion in linear models. Metrika 49:193–211
4. Mandal NK, Shah KR, Sinha BK (2000) Comparison of test vs. control treatments using distance optimality criterion. Metrika 52:147–162
5. Marshall AW, Olkin I (1979) Inequalities: Theory of majorization and its applications. Academic press, New York
6. Okamoto (1960) An inequality for the weighted sum of χ^2 variates. Bulletin of Mathematical Statistics 9:69–70
7. Pukelshiem F (1993) Optimal design of experiments. John Wiley & Sons, Inc., New York
8. Shah KR, Sinha BK (1989) Theory of optimal designs. Springer-Verlag Lecture Notes in Statistics Series, No. 54
9. Sinha BK (1970) On the optimality of some designs. Calcutta Statistical Association Bulletin 20:1–20
10. Tong YL (1990) The multivariate normal distribution. Springer-Verlag, New York