Optimality Criteria for Models with Random Effects

Criterios de optimalidad para los modelos con efectos aleatorios

Tisha Hooks[1,a], David Marx[2,b], Stephen Kachman[2,c], Jeffrey Pedersen[3,d]

[1]Department of Mathematics and Statistics, Winona State University, Winona, United States

[2]Department of Statistics, University of Nebraska, Lincoln, United States

[3]Department of Agronomy and Horticulture, USDA-ARS Research, University of Nebraska, Lincoln, United States

Abstract

In the context of linear models, an optimality criterion is developed for models that include random effects. Traditional information-based criteria are premised on all model effects being regarded as fixed. When treatments and/or nuisance parameters are assumed to be random effects, an appropriate optimality criterion can be developed under the same conditions. This paper introduces such a criterion, and this criterion also allows for the inclusion of fixed and/or random nuisance parameters in the model and for the presence of a general covariance structure. Also, a general formula is presented for which all previously published optimality criteria are special cases.

Key words: Optimal design, Information matrix, Nuisance parameter, Covariance structure, Mixed model.

Resumen

En el contexto de modelos lineales, los criterios de optimalidad se construyen para los modelos que incluyen efectos aleatorios. Tradicionalmente los criterios basados en la información asumen que todos los efectos en el modelo se consideran fijos. Cuando los parámetros, tratamientos o molestias son considerados efectos aleatorios, un criterio adecuado de optimalidad se puede desarrollar en las mismas condiciones. En este trabajo se introduce ese criterio, que permite la inclusión en el modelo de parámetros que representan

[a]Assistant professor. E-mail: THooks@winona.edu

[b]Professor. E-mail: DMarx1@unl.edu

[c]Professor. E-mail: SKachman1@unl.edu

[d]Geneticist and professor. E-mail: JPedersen1@unl.edu

molestias fijas o al azar, además de una estructura general de covarianza. También, se presenta una fórmula general para la cual en todos los casos publicados anteriormente, los criterios de optimalidad son casos especiales.

Palabras clave: diseño óptimo, matrix informativa, parametros molestos, estructura de covarianza, modelo mixto.

1. Introducción

There has been considerable interest among statisticians regarding the problem of design optimization. Researchers would like to choose an experimental design that maximizes the amount of information that is obtained from a fixed number of observations. To determine the optimal design among a set of candidates, it is necessary to define some criteria which allow discrimination between possible designs. Many well-known optimality criteria exist. Of these, A-optimality and D-optimality are the most commonly used (Kiefer 1974, Martin 1986). Moreover, with the advent of more complex analyses and experimental objectives in statistical research, new criteria which are relevant in these specific situations continue to be proposed in the literature (Dette & O'Brien 1999, Jacroux 2001).

The word "optimal" has several different interpretations in the context of experimental design. Usually, a researcher is concerned with finding the most accurate parameter estimates or predictions; thus, the optimal design will provide the highest quality of information concerning these parameter estimates. For linear models, information concerning the precision of the parameter estimates is contained in the variance-covariance matrix of parameter estimates. Thus, optimal designs are those experimental layouts that optimize some function of this matrix.

Many optimality criteria involve functions of the information matrix for given parameter estimates. Since primary interest often lies in the estimation of treatment effects, the most basic optimality criteria are given under the assumption that there are no effects other than those due to the intercept and treatment in the model. This criterion has been extended to allow for fixed nuisance parameters, such as block effects or covariates, in the model (Atkinson & Donev 1992), and also to permit the presence of spatial correlation (Martin 1986). These extensions will be discussed in more detail in section 2

Since various criteria have been proposed, a large segment of design literature has focused on finding specific designs that are optimal under the various criteria. Many commonly used experimental designs, *e.g.*, Latin squares, Youden squares, and even balanced incomplete block designs, have been shown to possess optimal properties (Kiefer 1958). These are intuitively appealing results since one would expect a balanced design to be more efficient if all parameters are of equal interest.

Even though these results satisfy the needs of most researchers, the existing theory does not apply to all models. These results were obtained based on criteria which operate under the assumption that all model effects are fixed. However, many well established analyses consist of fitting a linear model with at least some random effects, and in some cases, even treatment effects are regarded as random

(Sebolai et al. 2005). Schmelter (2007a) has discussed the optimality of designs for single-group designs for certain mixed models. He then extends his results for group-wise designs for linear mixed models 14 Schmelter (2007b).

Consider the case of unreplicated designs, where replicated "check" varieties are planted in the midst of unreplicated experimental varieties for comparative purposes. Such experiments are often carried out as the initial stage of a plant breeding experiment, where the objective is to select the top-performing experimental varieties for further testing. Therefore, it is important for the researcher to obtain the most accurate ranking of varieties. In the past, researchers have regarded variety as a fixed effect and based the rankings on least squares estimates of the variety effects. However, a recent simulation study which compared the efficiency of germplasm screening experiments with varying levels of checks demonstrated that the use of best linear unbiased predictors (BLUPs) was superior to the use of least squares means for selecting the highest proportion of true top-ranking genotypes (Sebolai et al. 2005). That is, Sebolai et al. (2005) demonstrated that a more accurate ranking of variety effects was obtained when treatment (variety) effects were assumed random as opposed to fixed. All previously considered optimality criteria are based on treatments being regarded as fixed effects. Given that a more accurate ranking of variety effects can be obtained when treatment effects are assumed random as opposed to fixed in an unreplicated experiment, it follows that an appropriate optimality criterion should be derived with the premise that treatments are assumed random.

The objective of this paper is to develop such a criterion. Before this criterion is given, a few well-known optimality criteria which are used in the case where only fixed treatment effects are in the model will be reviewed. Also, the extensions of these basic criteria to the case where fixed block effects are included in the model and can be viewed as nuisance parameters or to the case where observations are correlated will be discussed. Finally, the authors will extend these concepts to the case where treatment effects are assumed to be random while still allowing for nuisance parameters to be fixed and/or random and for correlated observations. Hereafter, the matrix W represents the design matrix for the parameters of interest (*i.e.*, treatment effects). Also, γ denotes the vector of fixed treatment effects, and g denotes the vector of random treatment effects.

2. Review of Some Well-Known Optimality Criteria

2.1. A-Optimality and D-Optimality

Consider the simple case where treatment effects are fixed and there are no other fixed effects, the means model. The model equation is given as

$$y = W\gamma + e \tag{1}$$

where y is the vector of observations, W is the design matrix, γ is the vector of treatment effects, and e is the vector of random errors. Assume $e \sim N\left(0, \sigma^2\right)$. Then the information matrix for the least squares estimates for treatment effects

is given by $W'W$ and so $Var(\widehat{\gamma}) = (W'W)^{-1}\sigma^2$. When treatments are of equal interest, a commonly used criterion is

$$\text{A-optimality} = \text{trace}\,(W'W)^{-1} \qquad (2)$$

Note that minimizing the trace of the inverse of the information matrix is equivalent to minimizing the average variance of the least squares estimates of the treatment effects. Another commonly used criterion based on the information matrix is

$$\text{D-optimality} = \left|(W'W)^{-1}\right| \qquad (3)$$

This is equivalent to minimizing what is known as the generalized variance of the parameter estimates (Dykstra 1971).

Other information-based criteria exist, such as G-optimality and I-optimality (which are based on the variance of prediction of the candidate points). However, A- and D-optimality are more frequently employed to evaluate classical experimental designs. An advantage of the D-optimality criterion is that optimal designs for quantitative factors do not depend on the scale of the variables, which in general is not the case for A-optimality. A solution to the problem of different scales for A-optimality has been proposed by (Dette 1995). The optimum design for estimating the maximum point of a quadratic response function with intercept is discussed in (Müller & Pazman 1998). However, for designs with all qualitative factors, such as block designs, problems of scale do not arise (Atkinson & Donev 1992) and the A-optimality criterion is frequently employed; thus, from this point forward, this paper will focus on A-optimality.

2.2. A-Optimality for Blocking Designs

The A-optimality criterion stated in equation (2) was given under the assumption that there were no effects other than fixed treatment effects in the model. Now, consider a blocking design in which both treatment and block are fixed effects. The model equation can be written as

$$y = W\gamma + X\beta + e \qquad (4)$$

where y is a vector of responses, W is the design matrix for treatment effects, X is the design matrix for block effects, γ is the vector of fixed treatment effects, β is the vector of fixed block effects, and e is the vector of random error terms. Assume $e \sim N(0, \sigma^2)$. Ideally, A-optimal designs will be those that minimize the trace of the covariance matrix of the least squares estimates of only the treatment effects; *i.e.*, block effects are viewed as nuisance parameters. Note that equation (4) can be rewritten as

$$y = P\alpha + e$$

where $P = (W\ X)$ and $\alpha' = (\gamma, \beta)'$. Specifically, the information matrix is given by $P'P = \begin{bmatrix} W'W & W'X \\ X'W & X'X \end{bmatrix}$. The inverse of this matrix can be partitioned as

$(P'P)^{-1} = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}$, where $Var(\widehat{\gamma}) = \sigma^2 D_{11}$. Recall the following result on the inverse of a partitioned matrix: Let T represent an $m \times m$ matrix, U an $m \times n$ matrix, V an $n \times m$ matrix, and S an $n \times n$ matrix. Suppose that T is nonsingular. Then $\begin{bmatrix} S & V \\ U & T \end{bmatrix}$ is nonsingular if and only if the $n \times n$ matrix $Q = S - VT^{-1}U$ is nonsingular, in which case $\begin{bmatrix} S & V \\ U & T \end{bmatrix}^{-1} = \begin{bmatrix} Q^{-1} & -Q^{-1}VT^{-1} \\ -T^{-1}UQ^{-1} & T^{-1}+T^{-1}UQ^{-1}VT^{-1} \end{bmatrix}$ (Harville 1997). It then follows that $D_{11} = (W'W - W'X(X'X)^{-1}X'W)^{-1}$. Thus, an optimal design which regards block effects as nuisance parameters (when both block and treatment effects are fixed) can be found by minimizing the trace of D_{11}. That is,

$$\text{A-optimality} = \text{trace}\left(W'W - W'X(X'X)^{-1}X'W\right)^{-1} \tag{5}$$

2.3. A-optimality in the Presence of Spatial Correlation

The previous criteria have assumed that observations are independent. Realizing that this assumption is often not met, Martin (1986) extended the general concept of A-optimality to the case where observations are correlated. Consider once again model equation (1), $y = W\gamma + e$. However, now assume $e \sim N(0, \sigma^2 R)$ where R is a positive definite matrix. The correlation structure is defined through the matrix R, and the correlation may be spatial, temporal, or of some other structure. Many known positive definite functions exist to describe spatial correlation (*e.g.*, spherical, Gaussian, and exponential functions) (Journel & Huijbregts 1978, Cressie 1993). When a specific function (*e.g.*, spherical) is chosen to construct R, it is necessary for design purposes to assume that R is known. In this case, the information matrix for the least squares estimates for treatment effects is given by $W'R^{-1}W$ and so $Var(\widehat{\gamma}) = (W'R^{-1}W)^{-1}\sigma^2$. Thus, a criterion for minimizing the average variance of the estimated treatment effects when R is a general positive definite covariance matrix of e is

$$\text{A-optimality} = \text{trace}\left(W'R^{-1}W\right)^{-1} \tag{6}$$

2.4. Criterion for Blocking Designs with Correlated Observations

Equation (6) is the criterion used in the case where treatments are fixed effects and correlation is present among the observations. It is of interest to extend this concept to allow for the inclusion of block effects. Consider model equation (4), $y = W\gamma + X\beta + e$, where we now assume $e \sim N\left(0, \sigma^2 R\right)$. Given R, a non-singular symmetric matrix $R^{\frac{1}{2}}$ exists so that $\left(R^{\frac{1}{2}}\right)' R^{\frac{1}{2}} = R^{\frac{1}{2}}R^{\frac{1}{2}} = R$ (Harville 1997). Letting $f = R^{-\frac{1}{2}}e$, it follows that $E(f) = 0$ and so $E(ff') = Var(f)$. Thus, we

have

$$Var(f) = E\left(ff'\right) = E\left(R^{-\frac{1}{2}}ee'R^{-\frac{1}{2}}\right)$$
$$= R^{-\frac{1}{2}}E\left(ee'\right)R^{-\frac{1}{2}}$$
$$= R^{-\frac{1}{2}}R^{\frac{1}{2}}R^{\frac{1}{2}}R^{-\frac{1}{2}}\sigma^2 = \sigma^2 I$$

which implies that $f \sim N(0, \sigma^2 I)$ (Draper & Smith 1998). If we premultiply equation (4) by $R^{-\frac{1}{2}}$, we have $R^{-\frac{1}{2}}y = R^{-\frac{1}{2}}W\gamma + R^{-\frac{1}{2}}X\beta + R^{-\frac{1}{2}}e$ or equivalently $Z = M\gamma + N\beta + f$. Now, new "W" and "X" matrices exist that can be substituted into equation (5) to arrive at an extension of the A-optimality criterion to the case where all effects are regarded as fixed, block effects are considered to be nuisance parameters, and observations are correlated. This criterion is given as

$$\text{A-optimality} = \text{trace}\left(M'M - M'N(N'N)^{-1}N'M\right)^{-1}$$
$$= \text{trace}\big(W'R^{-\frac{1}{2}}R^{-\frac{1}{2}}W - W'R^{-\frac{1}{2}}R^{-\frac{1}{2}}X(X'R^{-\frac{1}{2}}R^{-\frac{1}{2}}X)^{-1}$$
$$X'R^{-\frac{1}{2}}R^{-\frac{1}{2}}W\big)^{-1} \tag{7}$$
$$= \text{trace}\left(W'R^{-1}W - W'R^{-1}X(X'R^{-1}X)^{-1}X'R^{-1}W\right)^{-1}$$

If there are no nuisance parameters, then $X = 0$ and equation (7) is reduced to equation (6). Also, if there is no correlation among the observations, then $R = I\sigma^2$ and equation (7) is reduced to equation (5).

3. Extensions of the A-Optimality Criterion for Random Effects

In the previous section, the concept of A-optimality was extended to allow for both nuisance parameters and some correlation structure simultaneously; however, all treatment and block effects in this case were still assumed to be fixed. Here, a new A-optimality criterion is introduced which is premised on treatments being assumed random, and this criterion can incorporate fixed and/or random nuisance parameters.

3.1. Prediction Variance for Designs with Random Effects and Correlated Observations

The derivation of an A-optimality criterion in the case of random treatments relies on the theory of mixed models. Consider the following model,

$$y = Wg + X\beta + e \tag{8}$$

In this model, X is the design matrix for fixed effects, W is the design matrix for random effects, β is the vector of fixed effects, g is the vector of random effects, and e is the vector of random errors. A key assumption is that g and e are distributed

with $E\begin{bmatrix} g \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $Var\begin{bmatrix} g \\ e \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix}\sigma^2$. If G and R are known, $\widehat{\beta}$ is the best linear unbiased estimator (BLUE) of β and $\widehat{g}$ is the best linear unbiased predictor (BLUP) of g. It can be shown that

$$Var\begin{bmatrix} \widehat{\beta} - \beta \\ \widehat{g} - g \end{bmatrix} = \begin{bmatrix} X'R^{-1}X & X'R^{-1}W \\ W'R^{-1}X & W'R^{-1}W + G^{-1} \end{bmatrix}^{-1}\sigma^2$$

(Henderson 1975).

More specifically, using the inverse of a partitioned matrix (Harville 1997), $Var(\widehat{g} - g) = (W'R^{-1}W + G^{-1} - W'R^{-1}X(X'R^{-1}X)^{-1}X'R^{-1}W)^{-1}\sigma^2$. When treatments are considered to be random effects, an optimal design is generally found by minimizing the average prediction variance of the treatment effects.

3.2. Criterion for Blocking Designs with Random Effects and Correlated Observations

First, consider the case where treatments are random and blocks are fixed effects. Equation (8) applies to this case, where X is the design matrix for fixed blocks and W is the design matrix for the random treatment effects. Then an appropriate optimality criterion is given by

A-optimality =

$$\text{trace}\left(W'R^{-1}W + G^{-1} - W'R^{-1}X(X'R^{-1}X)^{-1}X'R^{-1}W\right)^{-1} \quad (9)$$

If there are no fixed effects in the model (which would be the case if the data were centered and hence there were no intercept in the model), then $X = 0$ and this criterion becomes

$$\text{A-optimality} = \text{trace}\left(W'R^{-1}W + G^{-1}\right)^{-1} \quad (10)$$

Next, consider the case where both treatments and blocks are random effects (again, assume the data is centered and there is no fixed intercept in the model). The corresponding model equation is given by

$$y = Wg + Zu + e \quad (11)$$

where W is the design matrix for random treatment effects, Z is the design matrix for random block effects, g is the vector of random treatment effects, and u is the vector of random block effects. Assume that g, u, and e are normally distributed with $E\begin{bmatrix} g \\ u \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ and $Var\begin{bmatrix} g \\ u \\ e \end{bmatrix} = \begin{bmatrix} G & 0 & 0 \\ 0 & D & 0 \\ 0 & 0 & R \end{bmatrix}\sigma^2$. If block effects are regarded as nuisance parameters, then this equation can be written as

$$y = Wg + \tilde{e} \quad (12)$$

where $\tilde{e} = Zu + e$ and $Var(\tilde{e}) = ZDZ' + R$. Equation (12) has no fixed effects, and so $Var(g) = G$ and $Var(\tilde{e}) = ZDZ' + R$ from equation (12) correspond to $Var(g) = G$ and $Var(e) = R$ from equation (10), respectively. Thus, from equations (10) and (12) we arrive at an A-optimality criterion for blocking designs with random effects in the presence of spatial correlation. Namely,

$$\text{A-optimality} = \text{trace}\left(W'(ZDZ' + R)^{-1}W + G^{-1}\right)^{-1}$$

Using the identity $(R + STU)^{-1} = R^{-1} - R^{-1}S(T^{-1} + UR^{-1}S)^{-1}UR^{-1}$ (Harville 1997), it is shown that this is equivalent to

$$
\begin{aligned}
\text{A-optimality} &= \text{trace}\big(W'(R^{-1} - R^{-1}Z(D^{-1} + Z'R^{-1}Z)^{-1}Z'R^{-1}) \\
&\qquad W + G^{-1}\big)^{-1} \\
&= \text{trace}\big(W'R^{-1}W + G^{-1} - W'R^{-1}Z(D^{-1} + Z'R^{-1}Z)^{-1} \\
&\qquad Z'R^{-1}W\big)^{-1}
\end{aligned}
\tag{13}
$$

If the treatment effects and the nuisance parameter (block effects) are fixed, then $D^{-1} = 0$ and $G^{-1} = 0$. Then equation (13) becomes equation (7).

3.3. Criterion for Designs with Fixed and/or Random Nuisance Parameters and Correlated Observations

In some cases, a linear model may contain both fixed nuisance parameters (*i.e.*, covariate effects) and random nuisance parameters (*i.e.*, block effects). The previous criteria can be extended to develop an appropriate optimality criterion for this situation. Let treatments be regarded as fixed effects, and consider the model

$$y = W\gamma + X\beta + Zu + e \tag{14}$$

where W is the design matrix for treatment effects, γ is the vector of fixed treatment effects, X is the design matrix for all fixed effects other than treatment, β is the vector of all fixed effects other than treatment, Z is the design matrix for random block effects, u is the vector of random block effects, and e is the vector of random error terms. Assume that $\begin{bmatrix} u \\ e \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} D & 0 \\ 0 & R \end{bmatrix}\right)$. If block effects are a nuisance parameter, then this equation can be rewritten as

$$y = W\gamma + X\beta + \tilde{e} \tag{15}$$

where $\tilde{e} = Zu + e$ and $\tilde{e} \sim N(0, ZDZ' + R)$. Both the treatment effects and nuisance parameters are fixed in the presence of correlated observations; thus, the criterion given in equation (7) applies here with $Var(e) = R$ replaced by $Var(\tilde{e}) = ZDZ' + R$ to obtain

$$
\begin{aligned}
\text{A-optimality} = \text{trace}\big[&W'(ZDZ' + R)^{-1}W - \\
&W'(ZDZ' + R)^{-1}X(X'(ZDZ' + R)^{-1}X)^{-1}X'(ZDZ' + R)^{-1}W\big]^{-1}
\end{aligned}
\tag{16}
$$

Alternatively, let treatments be regarded as random effects and consider the equation

$$y = Wg + X\beta + Zu + e \tag{17}$$

where $g \sim N(0, G)$, $u \sim N(0, D)$, $e \sim N(0, R)$, and g, u, and e are uncorrelated. Let X be the design matrix for all fixed effects and Z be the design matrix for random block effects. If blocks are random, then equation (17) can be rewritten as

$$y = Wg + X\beta + \tilde{e} \tag{18}$$

where $\tilde{e} = Zu + e$ and $\tilde{e} \sim N(0, ZDZ' + R)$. Since treatment effects are random and the nuisance parameters are fixed in the presence of correlated observations, the criterion given in equation (9) is applicable if $Var(e) = R$ is replaced by $Var(\tilde{e}) = ZDZ' + R$. Finally, we obtain

$$\text{A-optimality} = \text{trace}\{[W'(ZDZ' + R)^{-1}W + G^{-1} -$$
$$W'(ZDZ' + R)^{-1}X(X'(ZDZ' + R)^{-1}X)^{-1}X'(ZDZ' + R)^{-1}W]^{-1}\} \tag{19}$$

Note that if treatments are fixed as opposed to random, equation (19) reduces to equation (16). Also, if there are no fixed effects other than treatment, then $X = 0$ and equation (19) can be shown to reduce to equation (13). To observe how general this criterion is, note that equation (19) can be reduced back to equation (2) if treatments are regarded as fixed, there are no fixed or random nuisance parameters, and observations are uncorrelated.

4. Discussion

The criterion described by equation (19) covers the most general case, and all other previously published criteria are special cases of this formula. This equation will reduce to a simpler form based on whether treatments are fixed or random, whether fixed and/or random nuisance parameters are in the model, and whether or not observations are correlated. Table 1 presents a summary of the information matrices that would be used to derive the appropriate optimality criteria in each special case.

Equations (16) and (19) differ by only the addition of G^{-1} when treatments are random. Note that when treatments are regarded as random effects, there is an increase in the amount of information due to the assumptions which are made concerning the variability of the treatments; thus, the information matrix is greater when treatments are considered as random effects.

Also, it should be mentioned that each inverse within the information matrices listed in Table 2 can be replaced with a generalized inverse. The information matrix is invariant to the choice of generalized inverse used. However, if the entire information matrix is singular, one cannot calculate A-optimality by taking the trace of the generalized inverse of the information matrix. In this case, the A-optimality criterion is not invariant to the choice of generalized inverse used.

Finally, when observations are independent, the A-optimality criterion is said to be intuitively appealing since it both minimizes the average variance of estimates for treatment effects and the average variance of estimates for treatment differences (Martin 1986). However, when observations are correlated, A-optimality minimizes only the variance of the estimates for treatment effects. Thus, if interest lies in treatment comparisons, an appropriate optimality criterion should minimize the average variance of treatment differences. In this case, c-optimality is useful. For any estimable contrast $c'\widehat{\gamma}$ and information matrix M, we have

$$Var(c'\widehat{\gamma}) = \sigma^2 c' M^- c$$

where the choice of the g-inverse matrix M^- is arbitrary. This last expression is the c-optimality criterion (Silvey 1978, Pazman 1978). This criterion is also useful when the information matrix is singular since the c-optimality criterion is invariant to the choice of the generalized inverse.

Although our aim is not to emphasize the many search algorithms available to help construct an optimal design, the documentation of PROC OPTEX in SAS provides some enlightenment in this regard (SAS Institute Inc. 2007). A-optimal designs are harder to search for than D-optimal ones. Perhaps the easiest and fastest algorithm is the sequential search due to Dykstra (1971), which starts with no points in the design and adds successive candidate points so that the criterion is optimized after each point is added. The next fastest algorithm is the simple exchange method of Mitchell & Miller (1974). This technique tries to improve an initial design by adding a candidate point and then deleting one of the design points, stopping when the chosen criterion ceases to improve. The DETMAX algorithm of Mitchell (1974) is a widely used expansion of the simple exchange method above. Instead of requiring that each addition of a point be followed directly by a deletion, the algorithm provides for excursions in which the size of the design might vary between ND + k and ND - k. Here, ND is the specified size of the design and k is the maximum allowed size for an excursion. The three algorithms discussed so far add and delete points one at a time. The Fedorov (Fedorov 1972) algorithm is based on simultaneous switching. The modified Fedorov algorithm of Cook & Nachtsheim (1980) is another approach to the standard Federov method. Johnson & Nachtsheim (1983) introduced a generalization of both the simple exchange algorithm and the modified Fedorov search algorithm of Cook & Nachtsheim (1980). For a detailed review of the preceding search methods, see Nguyen & Miller (1998).

Even though these methods are very useful, the overall objective of this paper is to provide one equation which is appropriate for any computation of A-optimality. For example, if one were trying to decide which of four competing designs were optimal (and no others were available) when there was spatial structure, random blocking, and a fixed covariate to take into consideration, we would simply calculate the optimality criterion for each of the four scenarios and let those results help us to decide upon which design to implement. Another way to use the optimality formula would be to randomize treatment effects within a specific experimental design and compute the criterion value for each randomization choosing the most efficient. Marx & Stroup (1992) used this approach to determine the optimal 5×5 Latin Square design under spatial correlation.

5. Example

(a) Diagonal   (b) Knight's Move

FIGURE 1: Design layout for two competing Latin Squares.

Two 5×5 Latin Squares, a Diagonal design and the Knight's Move design (Figure 1), are compared. A spherical spatial structure with no nugget, a range of 5.0, and a sill of 1.0 was assumed. First, a model consisting of only fixed treatment effects is considered. Second, a model consisting of only random treatment effects (with a treatment variance of 1.0 and a treatment variance of 10.0) is considered. Finally, both fixed and random row and column nuisance parameters are added to the aforementioned models (with various values for the row and column variances).

The designs are compared on the basis of c-optimality. The reasons for using this criterion as opposed to the traditional A-optimality or D-optimality criterion are two-fold:

(i) Generally, it is more important to the researcher to minimize the average variance of treatment differences as opposed to the average variance of treatment effects.

(ii) When fixed rows and columns are included in the model, the information matrix is singular and the traditional A- or D-optimality criterion cannot be computed. However, since the choice of the g-inverse matrix M^- is arbitrary in the expression for c-optimality, this criterion can be used.

In this example, there are $\begin{bmatrix} 5 \\ 2 \end{bmatrix} = 10$ pairs of treatment differences. The variance of each pair-wise treatment difference was found using this c-optimality criterion, and the average of these variances was calculated for all models under both designs. The results are listed in Table 1. Note that for random treatment effects, as the variance component for treatments gets larger, the optimality criterion approaches that of the criterion for the model with fixed treatments. Similarly, as the variance components for row and column increase, the optimality criterion approaches that of the criterion for the model with fixed row and column effects.

6. Conclusions

Even though this paper discussed only extensions to the A-optimality criterion in detail, the ideas and results presented can be extended to any information-based

TABLE 1: Ratio of the c-optimality criteria (which minimize the average variance of treatment differences) for the Diagonal design relative to the Knight's Move design listed for various models (all assume a spherical spatial structure with nugget = 0, range = 5, and sill = 1).

	Treatments Fixed	Treatments Random $(\sigma_\tau^2 = 10)$	Treatments Random $(\sigma_\tau^2 = 1)$
No nuisance parameters	1.4077	1.4035	1.3684
Fixed row and column effects	1.4559	1.4510	1.4102
Random row and column effect $(\sigma_r^2 = 10$ and $\sigma_c^2 = 10)$	1.4533	1.4482	1.4078
Random row and column effect $(\sigma_r^2 = 1$ and $\sigma_c^2 = 10)$	1.4459	1.4410	1.4014
Random row and column effect $(\sigma_r^2 = 1$ and $\sigma_c^2 = 1)$	1.4388	1.3945	1.4399

optimality criteria without difficulty. Each criterion given was based on the trace of the information matrix, and this matrix can be used to obtain other criteria (see Table 2 for the relevant information matrices). The example in the previous section illustrates this. Another example is the D-optimality criterion. Since D-optimality is also based on the information matrix, another criterion premised on random effects which allows for both fixed and/or random nuisance parameters and spatial correlation is

$$\text{D-optimality} = \left| [W'(ZDZ' + R)^{-1}W + G^{-1} - W'(ZDZ' + R)^{-1}X(X'(ZDZ' + R)^{-1}X)^{-1}X'(ZDZ' + R)^{-1}W]^{-1} \right|$$

TABLE 2: Summary of all information matrices.

Trt	N-F	N-R	V	Information Matrix
F	N	N	I	$W'W$
F	N	N	R	$W'R^{-1}W$
F	Y	N	I	$W'W - W'X(X'X)^{-1}X'W$
F	Y	N	R	$W'R^{-1}W - W'R^{-1}X(X'R^{-1}X)^{-1}X'R^{-1}W$
R	N	N	I	$W'W + G^{-1}$
R	N	N	R	$W'R^{-1}W + G^{-1}$
R	Y	N	I	$W'W + G^{-1} - W'X(X'X)^{-1}X'W$
R	Y	N	R	$W'R^{-1}W + G^{-1} - W'R^{-1}X(X'R^{-1}X)^{-1}X'R^{-1}W$
F	N	Y	I	$W'W - W'Z(Z'Z + D^{-1})^{-1}Z'W$
F	N	Y	R	$W'R^{-1}W - W'R^{-1}Z(Z'R^{-1}Z + D^{-1})^{-1}Z'R^{-1}W$
F	Y	Y	I	$W'(ZDZ' + I)^{-1}W - W'(ZDZ' + I)^{-1}$ $X(X'(ZDZ' + I)^{-1}X)^{-1}X'(ZDZ' + I)^{-1}W$
F	Y	Y	R	$W'(ZDZ' + R)^{-1}W - W'(ZDZ' + R)^{-1}$ $X(X'(ZDZ' + R)^{-1}X)^{-1}X'(ZDZ' + R)^{-1}W$
R	N	Y	I	$W'W + G^{-1} - W'Z(Z'Z + D^{-1})^{-1}Z'W$
R	N	Y	R	$W'R^{-1}W + G^{-1} -$ $W'R^{-1}Z(Z'R^{-1}Z + D^{-1})^{-1}Z'R^{-1}W$
R	Y	Y	I	$W'(ZDZ' + I)^{-1}W + G^{-1} - W'(ZDZ' + I)^{-1}$ $X(X'(ZDZ' + I)^{-1}X)^{-1}X'(ZDZ' + I)^{-1}W$
R	Y	Y	R	$W'(ZDZ' + R)^{-1}W + G^{-1} - W'(ZDZ' + R)^{-1}$ $X(X'(ZDZ' + R)^{-1}X)^{-1}X'(ZDZ' + R)^{-1}W$

Note:

Trt = treatment (F = fixed treatment effects, R = random treatment effects)

N-F = fixed nuisance parameter (N = no, Y = yes)

N-R = random nuisance parameter (N = no, Y= yes)

$V = Var(e)$ (I = independent observations, R = correlated observations)

References

Atkinson, A. C. & Donev, A. N. (1992), *Optimum Experimental Designs*, Oxford University Press, New York, United States.

Cook, R. D. & Nachtsheim, C. J. (1980), 'A Comparison of Algorithms for Constructing Exact D-optimal Designs', *Technometrics* **22**(315-324).

Cressie, N. A. C. (1993), *Statistics for Spatial Data*, John Wiley & Sons, Inc., New York, United States.

Dette, H. (1995), 'Designing of Experiments with Respect to "Standardized" Optimality Criteria', *Journal of the Royal Statistical Society* **59**, 97–110.

Dette, H. & O'Brien, T. (1999), 'Optimality Criteria for Regression Models Based on Predicted Variance', *Biometrika* **86**, 93–106.

Draper, N. R. & Smith, H. (1998), *Applied Regression Analysis*, John Wiley & Sons, Inc., New York, United States.

Dykstra, O. J. (1971), 'The Augmentation of Experimental Data to Maximize $|X'X|$', *Technometrics* **13**, 682–688.

Fedorov, V. V. (1972), *Theory of Optimal Experiments*, Academic Press, New York, United States.

Harville, D. (1997), *Matrix Algebra from a Statistician's Perspective*, Springer-Verlag, New York, United States.

Henderson, C. R. (1975), 'Best Linear Unbiased Estimation and Prediction Under a Selection Model', *Biometrics* **31**, 423–447.

Jacroux, M. (2001), 'Determination and Construction of A-optimal Designs for Comparing two Sets of Treatments', *The Indian Journal of Statistics* **63**, 351–361.

Johnson, M. E. & Nachtsheim, C. J. (1983), 'Some Guidelines for Constructing Exact D-optimal Designs on Convex Design Spaces', *Technometrics* **25**, 271–277.

Journel, A. G. & Huijbregts, C. J. (1978), *Mining Geostatistics*, Academic Press, New York, United States.

Kiefer, J. (1958), 'On the Non-randomized Optimality and Randomized Nonoptimality of Symmetrical Designs', *The Annals of Mathematical Statistics* **29**, 675–699.

Kiefer, J. (1974), 'General Equivalence Theory for Optimum Designs (Approximate Theory)', *Annals of Statistics* **2**, 849–879.

Martin, R. J. (1986), 'On the Design of Experiments Under Spatial Correlation', *Biometrika* **73**, 247–277.

Marx, D. & Stroup, W. (1992), Designed Experiments in the Presence of Spatial Correlation, *in* 'Proceedings of the 1992 Kansas State University Conference of Applied Statistics in Agriculture', pp. 104–124.

Mitchell, T. J. (1974), 'An Algorithm for the Construction of D-optimal Experimental Designs', *Technometrics* **16**, 203–210.

Mitchell, T. J. & Miller, F. L. J. (1974), 'An Algorithm for the Construction of D-optimal Experimental Designs', *Mathematlcs Division Annual Progress Report (ORNL-4661)* pp. 130–131.

Müller, C. H. & Pazman, A. (1998), 'Applications of Necessary and Sufficient Conditions for Maximin Efficient Designs', *Metrika* **48**, 1–19.

Nguyen, N. K. & Miller, A. J. (1998), 'A Review of Exchange Algorithms for Constructing Discrete D-optimal Designs', *Metrika* **14**, 489–498.

Pazman, A. (1978), 'Computation of the Optimum Designs under Singular Information Matrices', *Annals of Statistics* **6**, 465–467.

SAS Institute Inc. (2007), *SAS OnlineDoc 9.2*, Cary, NC: SAS Institute Inc.

Schmelter, T. (2007a), 'The Optimality of Single-group Designs for Certain Mixed Models', *Metrika* **65**, 183–193.

Schmelter, T. (2007b), 'Considerations on Group-wise Identical Designs for Linear Mixed Models', *Journal of Statistical Planning and Inference* **137**, 4003–4010.

Sebolai, B., Pedersen, J. F., Marx, D. B. & Boykin, D. L. (2005), 'Effect of Grid Size, Control Plot Density, Control Plot Arrangement, and Assumption of Random or Fixed Effects on Non-replicated Experiments for Germplasm Screening', *Crop Science* **45**, 1978–1984.

Silvey, S. D. (1978), 'Optimal Design Measures with Singular Information Matrices', *Biometrika* **65**, 553–559.