INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
1560–8530/2003/05–4–571–575
http://www.ijab.org

Comparison of Different Central Composite Designs

M. AKRAM, MUNIR AKHTAR† AND M. HUSSAIN TAHIR†
Department of Statistics, B.Z. University, Multan–Pakistan
†Department of Statistics, Islamia University Bahawalpur–Pakistan

ABSTRACT

Several measures of robustness for designs, which is considered to be optimal with respect to A-, D-, E-, and G-optimality criteria. When several designs are proposed for comparison, then their optimality properties can be compared for the choice of design. To compare a class of central composite designs on the basis of optimality, a simple characterization of the optimality may be given in terms of the eigen values of $(X'X)$. Trace criterion has been used to measure the effect of missing observations on the variances of the estimates of the parameters and response. The different combinations of missing observations have different effect on the variances of the estimates of the parameters. Some combinations of these observations are more informative than others. The most informative combination of missing observations increase the variance maximum as compared to a least informative combination of missing observations.

Key Words: Central composite design; Optimality; Eigen values; Trace criterion; Informative combinations

INTRODUCTION

There are several measures of robustness for designs, which is considered to be optimal with respect to a particular criterion. Optimum experimental designs were originally developed by Kiefer (1959). For an overview of optimality criteria, see Box and Draper (1971), Box and Draper (1987), Atkinson and Donev (1992) and Pukelsheim (1993).

A k-factor central composite design (CCD) consists of n_f =factorial points 'f', n_a =2k axial points 'a' and n_c centre points 'c' at the design origin. Consider a second-order response surface model in k variables of the form: $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$, where $\underline{Y}$ is a vector of responses, X is an $n \times p$ model design matrix, $\underline{\beta}$ is a vector of unknown parameters and $\underline{\varepsilon}$ is a $p \times 1$ vector of errors. For a design matrix X, most design optimality criteria are based on optimal properties of the $X'X$ matrix and the levels or settings of the X matrix. When several designs are purposed for comparison, then their optimality properties may be compared for the choice of design.

The experimenter should be aware that a design superior to other designs by one optimality criterion might perform poorly when evaluated by another criterion. Therefore the performance of the design may be dependent on the choice of an evaluation criterion. Four commonly used criteria for design evaluation are A-, D-, E-, and G-optimality criteria. By considering the single value for the designs comparison, much information is lost regarding the design's performance.

When competing designs are compared, the E-criterion may be used over the design space. In such cases, the objective is to maximize the smallest eigen value of the information matrix $X'X$ over the design region to achieve the E-optimal design.

A measure of optimality based on the eigen values of information matrix in CCD. One of the most popular and commonly used classes of experimental designs for quadratic regression is the central composite designs. To compare a class of CCD's on the basis of optimality, a simple characterization of the optimality may be given in terms of the eigenvalues of $(X'X)$. By using the coding convention on the second-order polynomial equation, the eigen values of information matrix $(X'X)$ in CCD are

$$\lambda_1 = [e + (k-1)f + n + [(e + (k-1)f + n)^2 - 4(e \times n + (k-1)f \times n - kT^2]^{1/2}]/2$$
$$\lambda_2 = [e + (k-1)f + n - [(e + (k-1)f + n)^2 - 4(e \times n + (k-1)f \times n - kT^2]^{1/2}]/2$$

k eigen values with $\lambda_i = T$ i = 3, ..., k+2
(k-1) eigen values with $\lambda_j = e - f$ j= (k+3),..., (2k+1)
k (k-1) / 2 eigen values with $\lambda_u = f$ u= (2k+2),..., p

Where p=1+2k+ k (k-1) / 2, T= n_f +2x^2. $e = n_f + 2\alpha^4$ and $f = n_f$.

A particular member of CCD class can be identified by using the eigen values of information matrix of the model understudy. Different values of α can be used to define eigen values for various members of the class such as cuboidal, orthogonal, rotatable, BD outlier robust, spherical, minimaxloss1, minimaxloss2 and minimaxloss3 designs. The minimum of the maximum losses, of combinations of m missing observations, for the whole range of α is called minimaxloss. The CCD at this particular value of α corresponding to minimaxloss is called minimaxloss design for a combination of m (=1, 2, 3, …) missing observations.

If eigen values of the information matrix for the design understudy match with the eigen values of a compatible member of the CCD class, the design understudy could be identified as such.

Design optimality is an extremely interesting and useful tool for comparing the class of CCD particularly in terms of eigen values of the information matrix. By using these eigen values, the efficiency of different designs of

CCD class can be compared.

The eigen values λ_1, λ_2, λ_i (i = 3,..., k+2), λ_j(j=k+3, ..., 2k+1) and λ_u(u= 2k+2, ..., p) of the information matrix can be obtained as given in Table I is the case of the particular CCD used for demonstration.

Table I. Eigen values of the CCD for *k*=2, n_f =4 factorial, n_a =4 axial, and *n*=n_f+n_a+n_c design points at α=1.41421

n_c→	1	2	3	4
λ_1	24.3427	24.7047	25.0866	25.4891
λ_2	0.657281	1.2953	1.91337	2.51087
λ_3	8	8	8	8
λ_4	8	8	8	8
λ_5	8	8	8	8
λ_6	4	4	4	4

It can be seen from this table that the quantitative values of $\lambda_1$and λ_2 depend on k (variables), n_f factorial points, α the distance of the axial points from the design origin, n_c centre and n design points in the experiment under study i.e. the values of $\lambda_1$and λ_2 are changed by changing the value of any of k, n_f, n_c, n, α. The group of k from the remaining eigen values, with $\lambda_i = T = n_f + 2x^2$ (i = 3, ..., k+2), depend on n_f, the number of factorial points and the α values. The group of (k-1) from the remaining eigen values, with $\lambda_j = e - f = 2\alpha^4$ where (j=k+3, ..., 2k+1), depend only on the α values. The last group of k (k-1) / 2 eigen values, with $\lambda_u = f = n_f$ where (u= 2k+2,..., p), depends only on the n_f factorial points. For the same configurations of the central composite designs, λ_u is constant for fixed factorial points and λ_1, λ_2, λ_i and λ_j are the function of the α i.e. these eigen values are the functions of the settings of the model matrix X. λ_1 is the largest eigen value among p eigen values and increasing function of α while λ_2 is the decreasing and increasing function of α values. It can further be observed from Table I that no eigen values except λ_1 & λ_2 gets affected by the change in number of centre points.

Kiefer (1958, 1959) undertook a systematic study of the optimality of the experimental designs in a series of papers where various optimality criteria were discussed. To explain the various optimality criteria, let $M(\zeta) = \frac{(X'X)}{n}$ is the moment matrix for a design measure ζ and H denote the class of all design measures on the experimental region R. In CCD, Lucas (1974) proved that the maximum of $|X'X|$ is an increasing function of α. This is achieved only when the axial points are moved to the extremes of the experimental region R. The maximum value of $|M(\zeta)|$ is obtained, when the design measure is D-optimal. A design measure is D-optimal if it maximizes

$$|M(\zeta)| = \frac{[T]^K [f]^{{}^K C_2} [e-f]^{k-1} \left[n(e+(k-1)f) - kT^2 \right]}{n^p}$$

over H. It is equivalent to the G- optimality of minimization of the maximum variance of the estimated response, where p is the number of parameters in the model. In this case, if design points are fixed, then $|M(\zeta)|$ is the function of α values. A design measure is A-optimal, if it maximizes $\sum_{i=1}^{p} \lambda_i(\xi)$ or equivalently to minimize $tr(M(\zeta))^{-1}$ over H, where "***tr***" represent trace.

First of all Wald (1943) mentioned the E-optimality criterion for designs in testing hypothesis. The E-optimality property was proved by Ehrenfeld (1953). By Lucas (1977) approach, E-optimality can be described as minimizing the variance of the linear combination of regression coefficient having maximum variance. A design measure is E-optimal if it maximizes the smallest eigenvalue of $M(\zeta)$ or $(X'X)$ over H. The maximum of the minimum eigen value of $n \times M(\zeta) = (X'X)$ is

$$\lambda_2 = [e + (k-1)f + n - [(e + (k-1)f + n)^2 - 4(e \times n + (k-1)f \times n - kT^2]^{1/2}]/2$$

for the class of central composite designs such as orthogonal, rotatable, spherical, BD outlier robust and minimaxloss3 designs. The maximum of the minimum eigen values are given in the Table II for k=2 to 6 for the

Table II. Maximum of minimum eigen values for the range of $1 \le \alpha \le 3$

Variables (k)	Min (λ_{min})	Range of the α for minimum eigen value	Max (λ_{min}) for the range $1 \le \alpha \le 3$
2	λ_5 =2 at α =1.0	$1.0 \le \alpha \le 1.04746$	λ_2 =* λ_4 =4.0 at α =2.0
	λ_2 =2.34315 at α =1.189	$1.04746 < \alpha \le 2.0$	
	λ_4=4.0 at α=2.0	$2.0 \le \alpha$	
3	λ_5 =2 at α =1.0	$1.0 \le \alpha \le 1.16368$	λ_2 =6.68998 at α=3.0
	λ_2 =2.90033 at α =1.624	$1.16368 < \alpha$	
4	λ_5 =2 at α =1.0	$1.0 \le \alpha \le 1.2592$	λ_2 =6.89588 at α=3.0
	λ_2 =3.16702 at α =1.95	$1.2592 < \alpha$	
5	λ_5 =2 at α =1.0	$1.0 \le \alpha \le 1.3389$	λ_2 =6.39369 at α=3.0
	λ_2 =3.3216 at α =2.2125	$1.3389 < \alpha$	
6	λ_5 =2 at α =1.0	$1.0 \le \alpha \le 1.40695$	λ_2 =5.37523 at α=3.0
	λ_2 =3.42417 at α =2.4385	$1.40695 < \alpha$	

* λ_4 = is constant for specified n_f at all the values α i.e. λ_4 =4.0 for $1 \le \alpha \le 3$.

comparison of different CCD's.

Effect of *m* missing observations on different estimates of the model understudy. Missing values can occur at random by design. If any m observations are missing, which may not necessarily be the first *m* observations or even the contiguous one, we partitioned the response vector $\underline{Y}$ and model design matrix X as $\begin{bmatrix} \underline{Y}_m \\ \underline{Y}_r \end{bmatrix}$ and $\begin{bmatrix} X_m \\ X_r \end{bmatrix}$ respectively, after having shifted the missing observations along with their corresponding rows in the model design matrix at the top of both of the matrices. Thus $\underline{Y}_m$ consists of m missing observations and X_m consists of *m* respective rows. The information matrix $(X'X)$ may be written as:

$$X'X = X_m' X_m + X_r' X_r$$

Let $R_{ii} = X_i (X'X)^{-1} X_i$, where $i = r, m$

$\underline{\hat{\beta}}$ is the least square estimate of the parameter vector $\underline{\beta}$ when there are no missing observations and $\underline{\hat{\beta}}^*$ is also the least square estimate for the parameters $\underline{\beta}$ when m observations are missing. After substituting m missing values in $\underline{Y}_m$ by quantities

$\underline{F} = [F_1 \quad , \quad . \quad . \quad . \quad , \quad F_m]'$

Where (F_1, …, F_m) are the estimates of the missing observations that minimize the residual sum of square and consequently variance-covariance of estimates of the parameters. This, in turn, is equivalent to equating $\underline{F}$ to its expected value with the parameter $\underline{\beta}$ replaced by the estimate of $\underline{\beta}$ calculated for the augmented data, $\underline{\hat{\beta}}^*$.

McKee and Kshirsager (1982) showed that the least square estimates of $\underline{\beta}$ based on the remaining *(n-m)* observations $(y_{m+1}, y_{m+2} \cdots y_n)$ is given by

$$\underline{\hat{\beta}}^* = (X'X)^{-1} X' \begin{bmatrix} \underline{F} \\ \underline{Y}_r \end{bmatrix} \text{ and } \underline{\hat{Y}}^* = \begin{bmatrix} \underline{F} \\ \underline{\hat{Y}}_r \end{bmatrix} = X \underline{\hat{\beta}}^*$$

Where $\underline{Y}_r = [y_{m+1}, y_{m+2} \cdots y_n]'$

Where $\underline{\hat{Y}}^*$ is an estimate of the response vector with *m* missing values substituted by $\underline{F}$.

Akhtar and Prescott (1987), and McKee and Kshirsager (1982) proved the following results by using the estimates of the parameters due to *m* missing observations.

i) $\underline{F} = (I - R_{mm})^{-1} R_{mr} \underline{Y}_r$

ii) $Var(\underline{\hat{\beta}}^*) = \sigma^2[(X'X)^{-1} + (X'X)^{-1} X_m' (I - R_{mm})^{-1} X_m (X'X)^{-1}]$

Therefore the increase in the variance covariance matrix of $\underline{\hat{\beta}}^*$ as compared to $\underline{\hat{\beta}}$ is

iii) $V(\underline{\hat{\beta}}^*) - V(\underline{\hat{\beta}}) = \sigma^2 (X'X)^{-1} X_m' (I - R_{mm})^{-1} X_m (X'X)^{-1}$

Which can be derived and simplified as follows:

=> $\underline{F} = R_{mm}\underline{F} + R_{mr}\underline{Y}_r$

Table III. Relative increase in variance of $\underline{\hat{\beta}}^*$ and $\underline{\hat{Y}}^*$ due to a combination of three missing observations

Sr. No.	Relative increase in variances of $\underline{\hat{\beta}}^*$ due to three missing observations	Relative increase in variance of $\hat{Y}^*$ due to three missing observations	No. of Combinations of three missing observations	Combinations of three missing observations	
1	1.05882	0.500000	4	9,10,11	*ccc*
2	0.607843	0.444444	24	1,9,10	*fcc*
3	0.529412	0.444444	24	5,9,10	*acc*
4	0.764706	0.722222	8	1,4,9	*ffc*
5	0.794118	0.722222	16	1,2,9	*ffc*
6	0.567495	0.628385	32	1,5,9	*fac*
7	1.70588	1.50000	4	1,2,3	*fff*
8	0.558824	0.722222	16	5,7,9	*aac*
9	0.876021	0.967340	4	1,2,7	*ffa*
10	0.529412	0.722222	8	5,6,9	*aac*
11	1.68461	2.244629	32	1,6,9	*fac*
12	0.721467	0.967340	4	1,5,7	*faa*
13	2.17647	2.833333	8	1,4,5	*ffa*
14	1.00000	1.500000	4	5,6,7	*aaa*
15	1.94118	2.833333	8	1,5,6	*faa*
16	5.23529	6.166666	8	1,2,5	*ffa*
17	4.52941	6.166666	8	1,5,8	*faa*
18	16.1828	27.36599	4	1,2,8	*ffa*
19	18.6903	27.36599	4	1,6,8	*faa*

The sequence of relative increase in variances of $\underline{\hat{\beta}}^*$ and $\underline{\hat{Y}}^*$ is due to the combinations of three missing observations is changed by changing the values of α. The design with missing observations is the best design among designs of the same configuration which has minimum variance

$\hat{\underline{Y}}_r = R_{rm}F + R_{rr}\underline{Y}_r$

Let $R = X(X'X)^{-1}X' = \begin{bmatrix} R_{mm} & R_{mr} \\ R_{rm} & R_{rr} \end{bmatrix}$

$Var(\underline{F}) = \sigma^2 X_m (X_r' X_r)^{-1} X_m'$

$Cov(\underline{F}, \underline{Y}_r) = \sigma^2 X_m (X_r' X_r)^{-1} X_r'$

Where

$Var(\underline{Y}_r) = \sigma^2 I$

$Cov(\underline{Y}_r, \underline{F}) = \sigma^2 X_r (X_r' X_r)^{-1} X_m'$

$Var(\hat{\underline{\beta}}^*) = (X'X)^{-1} X' \; Cov\begin{bmatrix} F \\ \underline{Y}_r \end{bmatrix} X(X'X)^{-1}$

$= \sigma^2 (X'X)^{-1} \begin{bmatrix} X_m' & X_r' \end{bmatrix} \begin{bmatrix} X_m (X'X)^{-1} X_m' & X_m (X_r' X_r)^{-1} X_r' \\ X_r (X_r' X_r)^{-1} X_m' & I \end{bmatrix} \begin{bmatrix} X_m \\ X_r \end{bmatrix} (X'X)^{-1}$

$= \sigma^2 (X'X)^{-1} \begin{bmatrix} X_m' & X_r' \end{bmatrix} \begin{bmatrix} X_m (X'X)^{-1} X_m' X_m + X_m (X_r' X_r)^{-1} X_r' \\ X_r (X_r' X_r)^{-1} X_m' X_m + X_r \end{bmatrix} (X'X)^{-1}$

$= \sigma^2 (X'X)^{-1} [(X_m' X_m)(X_r' X_r)^{-1}(X_m' X_m) + (X_m' X_m)(X_r' X_r)^{-1}(X_r' X_r) + (X_r' X_r)(X_r' X_r)^{-1}(X_m' X_m) + (X_r' X_r)] (X'X)^{-1}$

$= \sigma^2 (X'X)^{-1} \left[(X_m' X_m) \{I + (X_r' X_r)^{-1} (X_m' X_m)\} + (X'X) \right] (X'X)^{-1}$

$Var(\hat{\underline{\beta}}^*) = \sigma^2 (X_r' X_r)^{-1}$

The variance of the estimate $\hat{\underline{Y}}^* = \begin{bmatrix} F \\ \hat{\underline{Y}}_r \end{bmatrix} = X\hat{\underline{\beta}}^*$ of the response vector $\underline{Y}$ due to m missing observations is

$Var(\hat{\underline{Y}}^*) = X \; Var(\hat{\underline{\beta}}^*) \; X' = \sigma^2 X (X_r' X_r)^{-1} X'$

The increase in the variance covariance matrix of $\hat{\underline{\beta}}^*$ as compared to $\hat{\underline{\beta}}$ is

$Var(\hat{\underline{\beta}}^*) - Var(\hat{\underline{\beta}}) = \sigma^2 \left[(X_r' X_r)^{-1} - (X'X)^{-1} \right]$

By using the trace criterion technique, the relative increase in variances of the estimates of the parameter $\underline{\beta}$ is

$$\frac{tr[Var(\hat{\underline{\beta}}^*)] - tr[Var(\hat{\underline{\beta}})]}{tr[Var(\hat{\underline{\beta}})]} = \frac{tr(X_r' X_r)^{-1} - tr(X'X)^{-1}}{tr(X'X)^{-1}} = \frac{tr(X_r' X_r)^{-1}}{tr(X'X)^{-1}} - 1$$

The different combinations of missing observations have different effect on the variances of the estimates of the parameters. For more detail see Akram (2002). The loss of a centre point is not as bad as the loss of a factorial point or an axial point when measuring the variances, or relative efficiency. Some combinations of these observations are more informative than the other combinations. The combination of design points may be defined as the most informative or influential points because, when it is missing, the loss efficiency is maximum and also variance is increased maximum. The most informative combination of missing observations increase the variance maximum as compared to a least informative combination of missing observations.

Similarly the change in the variance of the response estimate with missing observations may be described as.

$Var(\hat{\underline{Y}}^*) - Var(\hat{\underline{Y}}) = \sigma^2 X \left[(X_r' X_r)^{-1} - (X'X)^{-1} \right] X'$

The relative increase in the variance (RIV) of the estimate $\hat{\underline{Y}}^*$ of the response vector $\underline{Y}$ due to m missing observations may be explained as:

R.I.V.

$$= \frac{tr[Var(\hat{\underline{Y}}^*)] - tr[Var(\hat{\underline{Y}})]}{tr[Var(\hat{\underline{Y}})]} = \frac{\sigma^2 \, tr[X(X_r' X_r)^{-1} X'] - \sigma^2 \, tr[X(X'X)^{-1} X']}{\sigma^2 \, tr[X(X'X)^{-1} X']}$$

$$= \frac{tr[X(X_r' X_r)^{-1} X'] - tr(R)}{tr(R)} = \frac{tr[(X'X)(X_r' X_r)^{-1}]}{p} - 1$$

Where $tr(R) = tr[X(X'X)^{-1} X'] = \sum r_{uu} = n_f r_f + n_a r_a + n_c r_c) = p$,

r_f , r_a , r_c are the diagonal elements of R matrix and p is the number of the parameters in the model under study.

The relative increase in variances of $\hat{\underline{\beta}}^*$ and $\hat{\underline{Y}}^*$ i.e. $\frac{tr[Var(\hat{\underline{\beta}}^*)] - tr[Var(\hat{\underline{\beta}})]}{tr[Var(\hat{\underline{\beta}})]}$ and $\frac{tr[Var(\hat{\underline{Y}}^*)] - tr[Var(\hat{\underline{Y}})]}{tr[Var(\hat{\underline{Y}})]}$ due to a combination of three missing observations with k=2, n_f =4 factorial, n_a =4 axial and n_c =4 centre and n=12 design points at $\alpha = \sqrt{2}$ is shown in Table III.

The sequence of relative increase in variances of $\hat{\underline{\beta}}^*$ and $\hat{\underline{Y}}^*$ is due to the combinations of three missing observations is changed by changing the values of α . The design with missing observations is the best design among designs of the same configuration which has minimum variance.

CONCLUSION

Computations are carried out in terms of eigen values for the comparison of the class of CCD's on the basis of E-optimality.

We found the increase in the variance for CCD's when a combination of three observations is missing. Due to missing a combination of three observations, the increase in variance of the estimates is less as the number of design points is increased regardless of whether the missing combination consists of factorial, axial or centre points. The missing of an axial point may create more problem than the missing of a factorial point when measuring the variance for k $\geq$ 4. If in experiment, a most informative combination of observations is missing, the variance will then be more compared to a situation when a least informative combination of observations is missing. The more informative combinations are fff, ffa, faa, aaa that are cause to increase the variance maximum.

REFERENCES

Atkinson, A.C. and A.N. Donev, 1992. *Optimum Experimental Designs*. Oxford University Press, New York

Akhtar, M. and P. Prescott, 1987. A Review of Robust Response Surface Designs. *Pakistan J. Statist.,* 3(3) B, 11–26

Akram, M., 2002. Central Composite Designs Robust to Three Missing Observations. Unpublished *Ph.D. Thesis,* Islamia University Bahawalpur, Pakistan

Box, G.E.P. and N.R. Draper, 1971. Robust Designs. *Biometrika*, 62: 347–352

Box, G.E.P. and N.R. Draper, 1987. *Factorial Design, the* $|X'X|$ *criterion and some related maters.* John Wiley, New York
Ehranfield, S., 1953. On the Efficiency of Experimental Designs. *Ann. Math., Stat.*, 26: 247–55
Keifer, J., 1958. On the Nonrandomized Optimality and the Ranmdomized Nonoptimality of Symmetrical Designs. *Ann. Math, Statist.*, 29: 675–99
Kiefer, J., 1959. Optimum Experimental Designs (with discussions). *J. Roy. Statist., Soc.*, B21: 272–319
Lucas, J.M., 1974. Optimum Composite Designs. *Technometrics*, 16: 561–7
Lucas, J.M., 1977. Design Efficiencies for Varying Numbers of Centre Points. *Biometrka*, 64(1), 145-147.
McKee, B. and A.M. Kshirsagar, 1982. Effect of Missing Plots in Some Response Surface Designs. *ommun. Statist. – Theor. Meth.*, 11: 1525–49
Pukelsheim, F., (1993) O*ptimal Designs of Experiments.* Wiley, New York
Wald, A., 1943. On the Efficient Design of Statistical Investigations, *Ann. Math. Statist.*, 14, 34-140.

(Received 11 September 2003; Accepted 20 September 2003)