American Journal of Theoretical and Applied Statistics
2017; 6(4): 205-208
http://www.sciencepublishinggroup.com/j/ajtas
doi: 10.11648/j.ajtas.20170604.16
ISSN: 2326-8999 (Print); ISSN: 2326-9006 (Online)

Modified Exact Single-Value Criteria for Partial Replications of the Central Composite Design

Eugene C. Ukaegbu[*], Polycarp E. Chigbu

Department of Statistics, University of Nigeria, Nsukka, Nigeria

Email address:
eugene.ukaegbu@unn.edu.ng (E. C. Ukaegbu), polycarp.chigbu@unn.edu.ng (P. E. Chigbu)
[*]Corresponding author

To cite this article:
Eugene C. Ukaegbu, Polycarp E. Chigbu. Modified Exact Single-Value Criteria for Partial Replications of the Central Composite Design. *American Journal of Theoretical and Applied Statistics*. Vol. 6, No. 4, 2017, pp. 205-208. doi: 10.11648/j.ajtas.20170604.16

Received: October 18, 2016; **Accepted**: February 3, 2017; **Published**: July 10, 2017

Abstract: Replication of the factorial (cube) and/or axial (star) portions of the central composite design (CCD in response surface exploration has gained great attention recently. Some well known metrics (called single-value functions or criteria) and graphical methods are utilized in evaluating the regression based response surface design. The single-value functions considered here are the *A*-efficiency, $A = 100p/N\left\{trace\left(M^{-1}(\xi_k)\right)\right\}$ and the *D*-efficiency, $D = \left\{\left|M(\xi_k)\right|^{1/p}\right\}/N$, where $p = (k+1)(k+2)/2$, *k* is number of factors, ξ_k is the k^{th} design measure, $M(\xi_k)$ is the design's information matrix, $M^{-1}(\xi_k)$ is its inverse and *N* is the total number of experimental runs. These two functions are very popular in parameter estimation in response surface methodology. The exact measures of these two design criteria will be developed analytically in this work to account for partial replication of the cube and/or star components of the CCD. This will alleviate the burden of manual computation of these metrics when there are partial replications and reduce over reliance on software values which, often, are approximate values and maybe inaccurate.

Keywords: Design Efficiency, Determinant, Information Matrix, Partial Replication, Response Surface Methodology, Trace

1. Introduction

In many industrial experiments, relationship between a response of interest, y, and k independent design variables, $x_1, x_2, ..., x_k$, is often adequately described by second-order response surface model

$$\mathbf{y} = \beta_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x} + \varepsilon, \tag{1}$$

where $\mathbf{y}$ is the $N\times 1$ vector of responses, N being the number of experimental runs, β_0 is a constant, $\mathbf{x}$ is a point in the design space spanned by the design, $\mathbf{b}$ is a $k\times 1$ vector of first-order regression coefficients. In addition, $\mathbf{B}$ is a $k\times k$ symmetric matrix whose main diagonal entries are the coefficients of the pure quadratic terms and the off-diagonal entries are coefficients of one-half the mixed quadratic (interaction) terms, and ε is the random error term associated with the responses (see [1]). The p number of model parameters consists of one constant, k first-order terms, k quadratic terms and $k(k-1)/2$ interaction terms. Designs for fitting second-order response surface models are called second-order response surface designs.

The central composite design (CCD) of [2] is the most popular and commonly used second-order response surface designs. The design has three major components: the factorial portion (the cube) with coordinates, $(x_1, x_2, \ldots, x_k) = (\pm 1, \pm 1, \ldots, \pm 1)$, the x_i's, i = 1,...,*k*, are the factors; the axial portion (the star) with coordinates, $(\pm\alpha, 0, 0, \ldots, 0), \ldots, (0, 0, \ldots, \pm\alpha)$, α is the distance of the star point (also called the axial distance) from the centre of the design space, and the centre point with coordinates, $(0, 0, \ldots, 0)$. The number of design runs for the CCD is $N = n_c f + 2n_s k + n_o$, where *f* is the factorial or cube run, n_c is the replication of *f*, *2k* is the star run and n_s is the replication of *2k*, n_0 is the number of centre points, *c* is the cube, *s* is the star and *o* is zero (see, for example, [3]).

The *A*-efficiency and *D*-efficiency are the most commonly used single-value functions for estimating parameters in response surface exploration using response surface designs like the CCD (see [4]). [5] have extended the use of these optimality criteria to evaluating and selecting optimal designs for functional magnetic resonance imaging (fMRI). Further extension of the use of these efficiency criteria comes in the area of blocking. [6] considered partial replications of the components of the CCD which are arranged in orthogonal blocks and the emanating designs evaluated using the *A*- and *D*-efficiencies among other criteria. Statistical packages have made the computation of these single-value functions very easy. However, and as indicated by [7], the measures of these criteria provided by some statistical softwares could be approximated values and often misleading. Hence, it is imperative to have exact forms of these criteria that could be more reliable in design evaluation.

[8] proposed exact versions of the $A-$, $D-$ and $G-$ efficiencies as well as the $IV-$criterion for the classical CCD that is based on replication of the centre point alone. The inability of these exact functions to accommodate the replication of the cube or star or both portions of the CCD is a major drawback. In [9], exact functions have been provided for the *G*-efficiency and *IV*-criterion which compensate for partial replications of the components of the CD. In this work, we propose modified versions of the exact $A-$ and $D-$ efficiencies for the CCD that accommodate the replication of the cube, star or both portions of the CCD for any given axial distance, α, for any number of centre points and in any given design region (spherical or cuboidal). The proposed modified versions will make it easy to evaluate the parameters of the regression model in response surface exploration involving the CCD where the components of the CCD are replicated to enhance the design's performance. Doubts that surround the approximate results of some statistical packages are completely eliminated by using the exact forms of the efficiency criteria.

2. Exact *A*- and *D*-Efficiency for the Partial Replication of CCD

The *D*-efficiency is derived from the determinant of the information matrix, $M(\xi_k) = X'X$, given that $M(\xi_k)$ is non-singular, such that the $D-$ efficiency, $D = \left\{100\left|X'X\right|^{1/p}\right\}\big/N$. The $A-$ efficiency is derived from the trace of, $M^{-1}(\xi_k)$, the inverse of $M(\xi_k)$, such that the $A-$ efficiency, $A = 100p\big/trace\left\{N\left(X'X\right)^{-1}\right\}$, X is the design matrix expanded to model form. According to [8] and [10], the $A-$ and $D-$ efficiency measures represent the percentage number of runs required by a particular orthogonal design to achieve the same determinant and trace.

To obtain the *D*-efficiency, we first, derive the information matrix of the partially replicated CCD, then the determinant of the information matrix. Consider the design matrix of equation (2) for the *k* factor CCD, where *k* is a fixed positive integer and *k*>1. Let x_i and $x_{i'}$, $i = 1, \ldots, k;\ i' > i$, be any two variables of the CCD. The CCD has $f = 2^{k-q}$ (*q* is an integer) support points at the vertices, *2k* support points at the axes and n_0 support points at the centre of the design space.

$$X_k = \begin{bmatrix} 1 & x_i & x_{i'} & x_i x_{i'} & x_i^2 & x_{i'}^2 \\ \left.\begin{matrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 & 1 & 1 \end{matrix}\right\} f \\ \left.\begin{matrix} 1 & \alpha & 0 & 0 & \alpha^2 & 0 \\ 1 & -\alpha & 0 & 0 & \alpha^2 & 0 \\ 1 & 0 & \alpha & 0 & 0 & \alpha^2 \\ 1 & 0 & -\alpha & 0 & 0 & \alpha^2 \end{matrix}\right\} 2k \\ \left.\begin{matrix} 1 & 0 & 0 & 0 & 0 & 0 \end{matrix}\right\} n_0 \end{bmatrix} \quad (2)$$

Let *f* have n_c multiplicity, *2k* have n_s multiplicity and with n_0 centre points, with the subscripts, *c*, *s* and *0* representing cube, star and zero, respectively. Also, let X_k be the $N \times p$ expanded design matrix from the $N \times k$ design matrix of the partially replicated *k*-factor CCD and X'_k, its transpose for all *k*>1. The information matrix, $M(\xi_k) = X'_k X_k$ is determined directly by matrix multiplication. The elements of $M(\xi_k) = X'_k X_k$ are the inner products of X'_k and X_k. Let $\omega_{j\upsilon}$, $j = 1, \ldots, p; \upsilon = 1, \ldots, N$ be the element of the j^{th} row of X'_k and $\theta_{u\upsilon}$, $u = 1, \ldots, p$, be an element of the u^{th} column of X'_k, then the inner product of X'_k and X_k is given by (see, for example, [11])

$$\omega_{j\upsilon}\theta_{u\upsilon} = \omega_{j1}\theta_{u1} + \omega_{j2}\theta_{u2} + \ldots + \omega_{jN}\theta_{uN} = \sum_{\upsilon=1}^{N} \omega_{j\upsilon}\theta_{u\upsilon}.$$

Hence, the information matrix for the *k*-factor partially replicated CCD in block form, containing sub-matrices and vectors, is given by

$$M(\xi_k) = X'_k X_k = \begin{bmatrix} N & \varphi'_1 & d.J'_k & \varphi'_3 \\ \varphi_1 & diag(d) & \varphi_2 & \varphi'_4 \\ d.J_k & \varphi_2 & \psi_k & \varphi'_5 \\ \varphi_3 & \varphi_4 & \varphi_5 & diag(F) \end{bmatrix}, \quad (3)$$

where φ_1 = $k \times 1$ zero matrix, $diag(d)$ = $k \times k$ diagonal matrix with $d = F + 2n_s\alpha^2$ as the entries, $F = n_c f$, $J_k = k \times 1$ unit column vector, $\varphi_2 = k \times k$ zero matrix, $\psi_k = k \times k$ matrix whose diagonal entries are $F + 2n_s\alpha^4$ and the off-diagonal entries are $F's$, $\varphi_3 = \binom{k}{2} \times 1$ zero matrix, $\varphi_4 = \varphi_5 = \binom{k}{2} \times k$

zero matrices and $diag(F) = \binom{k}{2} \times \binom{k}{2}$ diagonal matrix with F as the entries. For example, for k = 2, the information matrix of the CCD with partial replication of the cube and star is given by

$$M(\xi_2) = \mathbf{X}_2'\mathbf{X}_2 = \begin{bmatrix} N & 0 & 0 & F+2n_s\alpha^2 & F+2n_s\alpha^2 & 0 \\ 0 & F+2n_s\alpha^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & F+2n_s\alpha^2 & 0 & 0 & 0 \\ F+2n_s\alpha^2 & 0 & 0 & F+2n_s\alpha^4 & F & 0 \\ F+2n_s\alpha^2 & 0 & 0 & F & F+2n_s\alpha^4 & 0 \\ 0 & 0 & 0 & 0 & 0 & F \end{bmatrix} \quad (4)$$

The derivation of the determinant of the information matrix, $M(\xi_k) = X_k'X_k$ follows. Let a_{ju} , j,u = 1, 2,..., p, be the element of $M(\xi_k)$ and W_{ju} be the $(p-1)\times(p-1)$ matrix obtained by deleting the j^{th} row and u^{th} column containing a_{ju} . Then, the determinant of $M(\xi_k)$ is given by (see [11]),

$$\det M(\xi_k) = \sum_{j=1}^{p} a_{ju} (-1)^{u+j} \left|W_{ju}\right|, \quad (5)$$

for any u, where $(-1)^{u+j}\left|W_{ju}\right|$ are the cofactors of $M(\xi_k)$ obtained by expanding W_{ju} to the second-order determinant. After performing the tedious matrix algebra, the determinant of the information matrix of the partially replicated CCD is given by

$$\det M(\xi_k) = \left(2n_s\alpha^4\right)^{k-1} Q\left(F+2n_s\alpha^2\right)^k F^{k(k-1)/2}, \quad (6)$$

where $Q = 2n_s N\alpha^4 + k\rho$ and $\rho = NF - \left(F+2n_s\alpha^2\right)$. For example, for$k$ = 2, the determinant is given by $\det M(\xi_2) = \left(2n_s\alpha^4\right)\left(F+2n_s\alpha^2\right)^2 QF$, where $Q = 2n_s N\alpha^4 + 2NF - 2\left(F+2n_s\alpha^2\right)^2$.

Multiplying the p^{th} root of the determinant by 100 and dividing by N gives the percentage D- efficiency. That is,

$$D = \frac{\left\{\left(2n_s\alpha^4\right)^{k-1} Q \left(F+2n_s\alpha^2\right)^k F^{\frac{k(k-1)}{2}}\right\}^{1/p}}{N} \times 100 \quad (7)$$

To obtain the A-efficiency, the inverse of the information matrix will be derived first. To derive the inverse of $M(\xi_k)$, the cofactors of the elements of $M(\xi_k)$ are obtained as described in the case of the determinant to form the matrix of the cofactors, $M(C_k)$. This matrix is transposed to obtain the adjugate or adjoint which is multiplied by the reciprocal of the determinant to obtain the inverse, $M^{-1}(\xi_k)$. Though $M(\xi_k)$ is symmetric matrix, so, $M(C_k)$ is also symmetric and therefore, equal to the adjugate. The inverse of the information matrix for the partially replicated CCD is given by

$$M^{-1}(\xi_k) = \frac{1}{\det M(\xi_k)} M(C_k)$$

which, expressed in block form is

$$M^{-1}(\xi_k) = \left(X_k'X_k\right)^{-1} = \begin{bmatrix} \mathbf{A}_1 & \varphi_1' & \mathbf{A}_2'.J_k' & \varphi_3' \\ \varphi_1 & diag\left(d^{-1}\right) & \varphi_2 & \varphi_4' \\ \mathbf{A}_2.J_k & \varphi_2 & \mathbf{A}_3 & \varphi_5' \\ \varphi_3 & \varphi_4 & \varphi_5 & diag\left(F^{-1}\right) \end{bmatrix}, \quad (8)$$

where $\mathbf{A_1} = \left(kF+2n_s\alpha^4\right)/Q$, $\mathbf{A_2} = -\left(F+2n_s\alpha^2\right)/Q$, $\mathbf{A_3}$ is a $k\times k$ matrix with $\left(2n_sN\alpha^4+(k-1)\rho\right)/\left(2n_s\alpha^4 Q\right)$ as diagonal entries and $\rho/\left(2n_s\alpha^4 Q\right)$ as off-diagonal entries, $Q = 2n_sN\alpha^4 + k\rho$, and $\rho = NF - \left(F+2n_s\alpha^2\right)^2$.

For example, the matrix of cofactors for two-factor partially replicated CCD is given by

$$M(C_2) = \begin{bmatrix} \left(2F+2n_s\alpha^4\right)\left(2n_s\alpha^4\right)\left(F+2n_s\alpha^2\right)^2 F & 0 & 0 \\ 0 & \left(2n_s\alpha^4\right)\left(F+2n_s\alpha^2\right)^2 QF & 0 \\ 0 & 0 & \left(2n_s\alpha^4\right)\left(F+2n_s\alpha^2\right)^2 QF \\ -\left(2n_s\alpha^4\right)\left(F+2n_s\alpha^2\right)^2 F & 0 & 0 \\ -\left(2n_s\alpha^4\right)\left(F+2n_s\alpha^2\right)^2 F & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{matrix} -\left(2n_s\alpha^4\right)\left(F+2n_s\alpha^2\right)^2 F & -\left(2n_s\alpha^4\right)\left(F+2n_s\alpha^2\right)^2 F & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \left(F+2n_s\alpha^2\right)^2 QF & -\left(F+2n_s\alpha^2\right)^2 F\left\{NF-\left(F+2n_s\alpha^2\right)^2\right\} & 0 \\ -\left(F+2n_s\alpha^2\right)^2 F\left\{NF-\left(F+2n_s\alpha^2\right)^2\right\} & \left(F+2n_s\alpha^2\right)^2 QF & 0 \\ 0 & 0 & \left(2n_s\alpha^4\right)\left(F+2n_s\alpha^2\right)^2 Q \end{matrix}\Bigg]$$

Therefore, the inverse of the information matrix of the two-factor partially replicated CCD is obtained by multiplying $M\left(C_2\right)$ by the reciprocal of the determinant, $\det M\left(\xi_2\right)=\left(2n_s\alpha^4\right)\left(F+2n_s\alpha^2\right)^2 QF$, which is the matrix, $M^{-1}\left(\xi_2\right)=\left\{\det M\left(\xi_2\right)\right\}^{-1} M\left(C_2\right)$.

Let λ_{jj} be the diagonal elements of $M^{-1}\left(\xi_k\right)$: see, for example, [12]. Then the trace of $M^{-1}\left(\xi_k\right)$ is given by

$$trace\left[M^{-1}\left(\xi_k\right)\right]=\sum_{j=1}^{p}\lambda_{jj}$$

$$=\frac{kF+2n_s\alpha^4}{Q}+\frac{k}{F+2n_s\alpha^2}+\frac{k\left(2Nn_s\alpha^4+(k-1)\rho\right)}{2n_s\alpha^4 Q}+\frac{\binom{k}{2}}{F} \quad (9)$$

Hence, the percentage A-efficiency is the product of p and 100 divided by the product of the trace of $M^{-1}\left(\xi_k\right)$ and N. That is,

$$A=\frac{p}{N\left\{\mathrm{A}_1+\frac{k}{F+2n_s\alpha^2}+\frac{k\left[2n_sN\alpha^4(k-1)\rho\right]}{2n_s\alpha^4 Q}+\frac{k(k-1)}{2F}\right\}}\times 100 \quad (10)$$

These mathematical expressions for the exact A- and D-efficiencies will reduce to the [8] type of exact A- and D-efficiencies if $n_c=n_s=1$. That is when there are no partial replications of the components of the CCD.

3. Conclusion

The central composite design is a popular second-order response surface design. By replicating the cube and star portions of the CCD analytical functions have been developed for obtaining the $A-$ and $D-$ efficiencies for the partially replicated design options for any given number of centre points and for any axial distance. The higher the values of A- and D-efficiencies, the better the model's parameter estimation with minimum variance which corresponds to maximizing the information. With these results, parameter estimation in response surface exploration using the partially replicated regression-based CCD becomes less tedious and there is no risk of making inference based on sometimes misleading approximate values.

References

[1] Li, J., Liang, L., Borror, C. M., Anderson-Cook, C. M. and Montgomery, D. C. (2009). Graphical Summaries to Compare Prediction Variance Performance for Variations of the Central Composite Design for 6 to 10 Factors, *Quality Technology and Quantitative Management,* Vol. 6 (4), 433 – 449.

[2] Box, G. E. P and Wilson, K. B. (1951), On the Experimental Attainment of Optimum Conditions, *Journal of the Royal Statistical Society,* Series B, Vol. 13, 1 – 45.

[3] Chigbu, P. E. and Ukaegbu, E. C. (2017), Recent Developments on Partial Replications of Response Surface Central Composite Designs: A Review, *Journal of Statistical Application and Probability*, Vol. 6 (1), 1 – 14.

[4] Ukaegbu, E. C. and Chigbu, P. E. (2015), Graphical Evaluation of the Prediction Capabilities of Partially Replicated Orthogonal Central Composite Designs, *Quality and Reliability Engineering International*, Vol. 31, 707 – 717.

[5] Kao, J.M-H. and Stufken, J. (2015), Optimal Design for Event-Related fMRI Studies: in Handbook of Design and Analysis of Experiments (Dean, Morris, Stufkten and Bingham) as Editors, CRC Press, Taylor and Francis Group, Boca Raton, London and New York.

[6] Ukaegbu, E. C. and Chigbu, P. E. (2017). Evaluation of Orthogonally Blocked Central Composite Designs with Partial Replications, *Sankhya B,* Vol. 79(1), 112-141. DOI: 10.1007/s13571-016-0120-z.

[7] Borkowski, J. J. (2003), A Comparison of Prediction Variance Criteria for Response Surface Designs, *Journal of Quality Technology,* Vol. 35 (1), 70–77.

[8] Borkowski, J. J and Valeroso, E. S. (2001), Comparison of Design Optimality Criteria of Reduced Models for Response Surface Designs in the Hypercube, *Technometrics*, Vol. 43 (4), pp. 468–477.

[9] Ukaegbu, E. C. and Chigbu, P. E. (2015). Characterization of Prediction Variance Properties of Rotatable Central Composite Designs for 3 to 10 Factors, *International Journal of Computational and Theoretical Statistics,* Vol. 2 (2), pp. 87–97.

[10] Wong, W. K. (1994), Comparing Robust Properties of A, D, E and G-optimal Designs, *Computational Statistics and Data Analysis,* Vol. 18, pp. 441-448.

[11] Searle, S. R. (1982), *Matrix Algebra Useful for Statistics*, John Wiley and Sons, New York.

[12] Atkinson, A.C. and Donev, A.N. (1992), Optimum Experimental Designs, Oxford University Press, New York.