CDA-OPTIMUM DESIGN FOR PARAMETER ESTIMATION, MINIMIZING THE AVERAGE VARIANCE AND ESTIMATING THE AREA UNDER THE CURVE

Mohamed M.E. Abd El-Monsef & Sally M. El-Sobky
Faculty of Science, Tanta University
mmezzat@science.tanta.edu.eg
Faculty of Science, Tanta University
sally.mohamed@science.tanta.edu.eg

ABSTRACT
The aim of this paper is to introduce a new compound optimum design named CDA, by combining the C-optimality, D-optimality, and A-optimality together. The significance of the proposed compound gains from that it can be used for parameter estimation, minimizing the average variance and model estimation simultaneously.

KEYWORDS: optimum design; C-optimality; D-optimality; A-optimality; compound criteria.
1. INTRODUCTION

Cook and Wong [2] considered a compound optimality criterion that is a convex combination of the two concave criteria and so we can find the optimal design directly as if this is a single objective optimal design problem. D-optimality focuses on the variances of the estimates of the coefficients in the model, which minimizing the determinant of \( (X^T X)^{-1} \) which is equivalent to maximizing the determinant of \( X^TX \). An exact design is called D-optimal, if it minimizes the determinant \( D \) of the covariance matrix. C-optimality interest is in estimating the linear combination of the parameters \( c^T \theta \) with minimum variance, where \( c \) is a known vector of constants. In A-optimality \( tr(M^{-1}(\xi)) \), the total variance of the parameter estimates, is minimized, equivalent to minimizing the average variance. This paper is organized as follows; the C-, D-, A – Optimum Designs were introduced in Section 2. The CDA-optimality was derived in section 3 and some of its properties were discussed. The generalized CDA- Optimum Design was introduced in Section 4.

2. C-, D-, A – OPTIMUM DESIGNS

C-optimality introduced by Elfving [2] which provided a geometrical interpretation for finding c-optimal designs and developed by Silvey and Titterington [10] and Titterington[11]. Fellman [4] justified that at most \( m \) linearly independent support points are needed for a c-optimal design. Pukelsheim and Torsney [9]introduced a method for computing c-optimal weights given the support points. C-optimality minimize the variance of the best linear unbiased estimate for a given linear combination of the model parameters \( c^T \theta \), where \( c \) is \( p \times 1 \), a vector of a known constants. The c-optimality criterion to be minimized is thus

\[
\var{c^T \hat{\theta}} \propto c^T M^{-1}(\xi) c
\]

The aim of c-optimality is to obtain the best design for estimating the linear combination of the parameters

\[
c_1\theta_1 + \cdots + c_p\theta_p = c^T \theta
\]

The efficiency of any design \( \xi \) relative to C-optimum design is defined as:

\[
Eff^c(\xi) = \frac{c^T M^{-1}(\xi)}{c^T M^{-1}(\xi) c}
\]

C-optimality is defined as \( \min \var{C^T \hat{\theta}} \), which is proportional to \( C^T M^{-1}(\xi)C \). A disadvantage of c-optimum designs is that they are often singular.

D- Optimum Design

D-optimum design is one of the most commonly used design criteria for linear regression model that is also known as the Determinant criterion. This criterion introduced by Wald [12], and later was called D-optimality by Kiefer and Wolfowitz [5]. The D-Optimality is the most common criterion due to numerous applications found in the literature; see for example, Latif and Zafar Yab [6]and Poursina and Talebi [8]. D-optimality criterion is just to maximize the determinant of the Fisher information matrix, \( |X^TX| \), this means that the optimal design matrix \( X^* \) contains the \( n \) experiments which maximizes the determinant of \( X^TX \).

Maximizing the determinant of the information matrix \( X^TX \) is equivalent to minimizing the determinant of the dispersion matrix \( (X^TX)^{-1} \). Using such an idea, the D-efficiency of an arbitrary design, \( X \), is naturally defined as

\[
Eff(D) = \left\{ \frac{|M(\xi)|}{|M(\xi^*)|} \right\}^{1/2p}
\]

A-Optimum Design

A-optimality criterion introduced by Chernoff [1], who showed that the employed criterion of optimality is the one that involves the use of Fisher’s information matrix. Invariance under re-parameterization loses its appeal if the parameters of interest have a definite physical meaning. Then the average-variance criterion provides a reasonable alternative. If the coefficient matrix is partitioned into its columns, \( K = (c_1,\ldots,c_p) \), then the inverse \( 1/\phi_{11} \) can be represented as
\[
\frac{1}{s} \text{tr}\left( C_K(A) \right) = \frac{1}{s} \text{tr}\left( C_K A^{-1} \right) = \frac{1}{s} \text{tr}\left( K A^{-1} K^\top \right) = \frac{1}{s} \sum_{i=1}^{s} c_i A^{-1} c_i
\]

This is the average of the standardized variances of the optimal estimators for the scalar parameter systems \( c_i \theta, \ldots, c_i \theta \) formed from the columns of \( K \). From the point of view of computational complexity, the criterion \( \phi_{-1} \) is particularly simple to evaluate since it only requires the computation of the \( s \) diagonal entries of the dispersion matrix \( K A^{-1} K^\top \).

### 3. CDA- OPTIMUM DESIGN

To obtain parameter estimation, minimizing the average variance and model estimation of the area under the curve, a new compound criteria called CDA is introduced. CDA is constructed by combining C, D and A-optimality. By maximizing a weighted product of the efficiencies

\[
\left\{ \text{Eff}^{(c)} \right\}^k \left\{ \text{Eff}^{(D)} \right\}^{k(1-k)} \left\{ \text{Eff}^{(A)} \right\}^{(k-1)^2}
\]

Then taking the logarithm we get

\[
= k \log \left\{ \text{Eff}^{(c)} \right\} + k(1-k) \log \left\{ \text{Eff}^{(D)} \right\} + (k-1)^2 \left\{ \text{Eff}^{(A)} \right\}
\]

\[
= -k \log \left\{ c M^{-1}(\xi) c \right\} + k(1-k) \log \left\{ M(\xi) \right\}^{1/p} + (k-1)^2 \log \left\{ M^{-1}(\xi) \right\}
\]

The terms containing \( \xi_c^*, \xi_d^* \) and \( \xi_A^* \) are constants, a maximum is found over \( \xi \). Hence, the criterion that has to be maximized is given by

\[
\Phi^{(CDA)}(\xi) = -k \log \left\{ c M^{-1}(\xi) c \right\} + \frac{k(1-k)}{p} \log \left\{ M^{-1}(\xi) \right\} + (k-1)^2 \log \left\{ M^{-1}(\xi) \right\}
\]

and the derivative function for CDA-optimality is

\[
\phi^{(CDA)}(x, \xi) = -k \left\{ f^T(x) M^{-1}(\xi) c \right\}^2 + \frac{k(1-k)}{p} \left\{ f^T(x) M^{-1} f(x) \right\} + (k-1)^2 \left\{ f^T(x) M^{-1} f(x) \right\}
\]

A CDA-optimum design, \( \xi_{\text{CDA}} \), maximizes \( \Phi^{(CDA)}(\xi) \) or equivalently \( \log \Phi^{(CDA)}(\xi) \). The equivalence theorem can now be stated as follows:

**Theorem 1.**

i. A necessary and sufficient condition for a design \( \xi_{\text{CDA}}^* \) to be CDA-optimum is fulfillment of the inequality

\[
\phi^{(CDA)}(x, \xi_{\text{CDA}}^*) \leq 1, \quad x \in X.
\]

ii. The upper bound of \( \phi^{(CDA)}(x, \xi_{\text{CDA}}^*) \) is achieved at the points of the optimum design.

iii. For any non-optimum design \( \xi \) that is a design for which \( \phi^{(CDA)}(\xi) < \Phi^{(CDA)}(\xi_{\text{CDA}}^*) \) and

\[
\sup_{x \in X} \phi^{(CDA)}(x, \xi_{\text{CDA}}^*) > 1.
\]

A measure of efficiency of a design \( \xi \) relative to a CDA-optimum design is given by

\[
\text{Eff}_{\text{CDKL}}(\xi) = \frac{\phi^{(CDA)}(\xi)}{\Phi^{(CDA)}(\xi_{\text{CDA}}^*)}
\]

The proof can be made directly, since \( \Phi^{(CDA)}(\xi) \), \( 0 \leq k \leq 1 \) is a convex combination of three optimum design criteria, so the CDA-criterion is also convex and satisfying convexity conditions.
Properties of CDA-Optimality

A good design should give a small variance matrix, therefore the function $\Phi$ is related to the variance matrix, and should have following properties:

i. **Non-negativity**: $\Phi_{CDA}(M) \geq 0$.

ii. **Isotonicity**: if $(M^* - M)$ is a positive semi-definite matrix, then $\Phi[M^*] \geq \Phi[M]$.

iii. **Positive homogeneity**: $\Phi[kM] = k \Phi[M]; k > 0$.

iv. **Superadditivity**: $\Phi[M + M^*] \geq \Phi[M] + \Phi[M^*]$.

The previous properties are important to define with a proper scaling the relative efficiency of an experiment (or a design with the matrix $M$) with respect to another reference experiment with $M^*$. Pazman [7] discussed some other optimality properties for small samples.

4. THE GENERALIZED CDA-OPTIMALITY:

A generalized CDA-criterion will be introduced as:

$$
\Phi^{(GCDA)}(\xi) = -\sum_{j=1}^{m} a_j \log \left| A_j^T M^{-1} (\xi) A_j \right| + \sum_{i=1}^{n} \frac{s_i}{b_i} \log \left| A_i^T M^{-1} (\xi) A_i \right|
$$

$$
+ \sum_{s_i} c_i \log \left| A_i^T M^{-1} (\xi) A_i \right|
$$

where, $a_j, s_i, b_i$ and $c_i$ are sets of non-negative coefficients reflecting the importance of the parts of the design criteria.

REFERENCES

[1] Chernoff, H. (1953). "Locally optimal designs for estimating parameters". Annals of Mathematical Statistics, 24, 586–602.

[2] Cook, R. and Wong, W. (1994). "On the equivalence of constrained and compound optimal designs". Journal of the American Statistical Association, 89, 687–92.

[3] Elfving, G. (1952). "Optimum allocation in linear regression theory". Annals of Mathematical Statistics, 23, 255–262.

[4] Fellman, J. (1974). "On the allocation of linear observations". Commentationes Physico-Mathematicae, Helsinki, 44, 27-78.

[5] Kiefer, J. and Wolfowitz, J. (1959). "Optimum designs in regression problems". Annals of Mathematical Statistics, 30, 271–294.

[6] Latif, S. and Zafar Yab, M. (2015). "D-optimal designs for beta regression models with single predictor". Journal of Statistical Computation and Simulation, 85(9), 1709–1724.

[7] Pazman, A. (2007). "Criteria for optimal design of small-sample experiments with correlated observations". Kybernetika, 43(4), 453 – 462.

[8] Poursina, D. and Talebi, H. (2014). "Modified D-optimal design for logistic model". Journal of Statistical Computation and Simulation, 84(2), 428-437.

[9] Pukelsheim, F. and Torsney, B. (1991). "Optimal weights for experimental designs on linearly independent support points". Annals of Statistics, 19, 1614-1625.

[10] Silvey, S. and Titterington, D. (1973). "A geometric approach to optimal design theory". Biometrika, 60, 15-19.

[11] Titterington, D. (1975). "Optimal design: some geometrical aspects of D-optimality". Biometrika, 62, 313-320.

[12] Wald, A. (1943). "On the efficient design of statistical investigations". Annals of Mathematical Statistics, 14, 134–140.