D-Optimal Designs: A Mathematical Programming Approach Using Cyclotomic Cosets

Mujahid N. SYED¹, Ilias S. KOTSIREAS², Panos M. PARDALOS¹

¹University of Florida, Department of Industrial and Systems Engineering 303 Weil Hall, P.O. Box 116595, Gainesville, FL 32611-6595

²Wilfrid Laurier University, Department of Physics and Computer Science University Avenue West Waterloo, Ontario N2L 3C5, Canada

e-mail: smujahid@ufl.edu, ikotsire@wlu.ca, pardalos@ufl.edu

Received: July 2011; accepted: September 2011

Abstract. In this paper, a novel approach involving the concepts from mathematical programming and number theory is proposed to find the D-optimal designs. In specific, we will propose a mathematical formulation for the D-optimal design. In addition to that, we will present the use of cyclotomic cosets in the mathematical formulation, in order to reduce the total number of binary variables. We will illustrate the validity of our proposed method by solving a difficult known instance (N = 126) of the D-optimal design.

Keywords: D-optimal design, cyclotomy cosets, mathematical programming.

1. Introduction

One of the widely known techniques from the theory of statistical design of experiments (Fisher, 1935) is factorial design. Specifically, two-level factorial designs find their application in numerous fields of science and technology (Mason et al., 2003). It is hard to enumerate all the applications of the two-level factorial design. On the other hand, the rationale behind two-level factorial design is simple to understand. In brief, a two-level factorial design consists of finding the relationship among different input factors and outcome. This relationship is studied by performing experiments with varying levels of the factors. For example, consider a case where two factors are studied for the relationship. In order to determine the relationship, factors are taken into two levels "High (+1)" and "Low (-1)". By observing all the possible combinations of the two factors, the relationship can be studied.

Such a design, where all possible combinations are observed, is called a full factorial design. Typically, a full factorial design can be represented by a design matrix, R. For example, a two factor two level design matrix will be:

$$R = \begin{bmatrix} +1 & +1\\ +1 & -1\\ -1 & +1\\ -1 & -1 \end{bmatrix},$$
(1)

where each column represents a factor, and each row represents an observation. In addition to that if there are N factors with two levels, then for a full factorial design there will be 2^N observations. Thus, as the total number of factors increases, a full factorial design will be expensive. To overcome the exponential growth in the required number of observations, various design methods have been proposed. The main objective of all these methods is to reduce the total number of observations, while preserving the amount of information. One such method is D-optimal design.

Let O be the number of observations required for the design. If $O = 2^N$ then the design is full factored, on the other hand, if O = N the design is called saturated design. The goal of D-optimal design is to maximize the information in the saturated design. Let \widehat{M} be the matrix defined as:

$$M = R^t R. (2)$$

Then the matrix \widehat{M} is proportional to the information matrix. In order to maximize the information, $|\widehat{M}|$ is maximized, which is equivalent to maximizing $|R|^2$ or equivalently maximizing |R|. Thus, in simple terms, D-optimal design means a saturated design matrix with the maximum determinant.

These matrices are characterized by *order* which refers to the total number of factors (or total number of observations, since O = N) in an experiment. Moreover, let ε be such that, $N \equiv \varepsilon \mod 4$. Based on different values of ε (i.e., $\varepsilon = 0, ..., 3$), different construction methods of such matrices are proposed. For example, when $\varepsilon = 0$ these matrices are similar to Hadamard matrices, and there are numerous results on their construction methods (Hedayat and Wallis, 1978; Horadam, 2007). On the other hand, when $\varepsilon \neq 0$ there is no concrete method to construct all the matrices (Chadjipantelis, 1987). Most of the methods are heuristic search based methods (Koukouvinos *et al.*, 1997; Djokovic and Kotsireas, 2011; Kotsireas *et al.*, 2010). In this paper, our focus will be to present the problem of finding D-optimal matrices as an optimization problem. Before presenting the approach, we will review some preliminary concepts.

This paper is dedicated to Professor Jonas Mockus 80th year anniversary.

2. Preliminaries

2.1. D-Optimal Design

Let $\overline{M} \in \Re^{N \times N}$ be the set of square matrices defined as, $\overline{M} = \{M: [M]_{i,j} \in \{-1, +1\}$ $\forall i, j = 1, ..., N\}$, where $[M]_{ij}$ represents the *i*th row *j*th column element of matrix M.

DEFINITION 1. A matrix $R \in \overline{M}$ is called D-optimal design of order N, if $|R| >= |M|, \forall M \in \overline{M}$.

Some D-optimal designs have a specific structure, one of the widely known structure is circulant type D-optimal design. The advantage of using circulant type design is reduction in the number of search variables in the D-optimal design (i.e., the total number of variables is reduced from N^2 to N). However, circulant type reduction can only be achieved when N is even. Ehlich has proposed the following reduction for the circulant type D-optimal designs.

DEFINITION 2. A circulant type D-optimal design is a D-optimal design (matrix) R that can be decomposed into the following block structure

$$R = \begin{bmatrix} A & B \\ -B^t & A^t \end{bmatrix},\tag{3}$$

where A, B are circulant commuting $n \times n$ matrices, with elements ± 1 , and satisfy:

$$AA^{t} + BB^{t} = (2n-2)I_{n} + 2J_{n},$$
(4)

where $n = \frac{N}{2}$, $N \equiv 2 \mod 4$, and $I_n, J_n \in \Re^{n \times n}$ are identity matrix and unit matrix (or all-ones matrix) respectively.

COROLLARY 1. Since the matrices A, B are circulant, they can be completely defined by one single row. Let a and b represent the first rows of A and B respectively. If we take each element of (4), then we have the following two equations:

(i) For the non-diagonals elements in (4), we have:

$$\sum_{i=1}^{n} a_i a_{i\oplus s} + \sum_{i=1}^{n} b_i b_{i\oplus s} = 2 \quad \forall s = 1, \dots, \left\lfloor \frac{n}{2} \right\rfloor,\tag{5}$$

where

$$i \oplus s = \begin{cases} i+s, & \text{if } i+s \leq n, \\ (i+s) \mod n, & \text{otherwise.} \end{cases}$$
(6)

The left hand side of (5) is also called periodic auto-correlation function.

(ii) In addition to (5), consider the following:

$$\left(\sum_{i=1}^{n} a_i\right)^2 + \left(\sum_{i=1}^{n} b_i\right)^2 = 2n + \sum_{s,i} (a_i a_{i\oplus s} + b_i b_{i\oplus s})$$
(7)

$$= 2n + (n-1)2 = 4n - 2,$$
(8)

where s = 1, ..., n - 1 and i = 1, ..., n.

$$\implies \left(\sum_{i=1}^{n} a_i\right)^2 + \left(\sum_{i=1}^{n} b_i\right)^2 = 4n - 2. \tag{9}$$

Equation (9) is a Diophantine equation. These two equations (5) and (9) are the basic equations, that will be used in constructing mathematical formulation.

In addition to the above definition, there is another famous definition of D-optimal design that will be presented in the following paragraph. In order to present this definition of D-optimal design, we will introduce supplementary difference sets.

2.2. Supplementary Difference Sets

Let S^1, S^2, \ldots, S^v be subsets of Z_n ¹ containing n_1, n_2, \ldots, n_v elements respectively. Let T_k represent the totality of all the differences between elements of S^k with repetition, i.e.,

$$T_k = \left\{ \left(s_i^k - s_j^k \right) \mod n \right\}_{i,j=1,\dots,n_k},\tag{10}$$

where s_i^k is the *i*th element of subset S^k . Similarly, let T be the totality of all such T_k 's, defined as, $T = \{T_k\}_{k=1,...,v}$. If each non-zero element in T is repeated a fixed number of times, say λ , then the sets will be called $v - \{n, n_1, ..., n_v; \lambda\}$ Supplementary Difference Sets (SDS).

Let S^{a-} be the set containing indices of **a**, defined as:

$$S^{a-} = \{i: a_i = -1, \, \forall i = 1, \dots, n\}$$
(11)

Similarly, let S^{b-} be the set containing indices of b (analogously defined as S^{a-}). Let n_{a-} and n_{b-} be the total number of -1's in a and b respectively.

DEFINITION 3. If S^{a-} and S^{b-} form a SDS with parameter, λ defined as:

$$\lambda = n_{a-} + n_{b-} - \frac{1}{2}(n-1), \tag{12}$$

such that

$$(n - 2n_{a-})^2 + (n - 2n_{b-})^2 = 4n - 2,$$
(13)

then the circulant matrices formed by a and b correspond to the matrices A and B of the D-optimal design (as defined in (3)).

It can be easily seen that Definitions 2 and 3 are equivalent. Moreover, SDS can be obtained using the concept of cyclotomy cosets (Storer, 1967), which will be briefly described in the following paragraph. For a detailed discussion refer (Wallis, 1973; Cusick *et al.*, 2004).

¹Or any finite Abelian group of order n.

2.3. Cyclotomy Cosets

Let F be a Galois field of order q, denoted as GF(q), let x be a primitive element of field F. Let $q = p^{\alpha}$, where p is some prime number. Now, let us write q = ef + 1, for some $e, f \in Z$.

Let $G = F \setminus \{0\}$ be an Abelian group defined as $G = \langle x \rangle$. The cyclotomic cosets (C_i) for such F are obtained as:

$$C_i = \left\{ x^{es+i} : s = 0, \dots, f-1 \right\} \quad \forall i = 0, \dots, e-1.$$
(14)

From the theory of cyclotomy, we know that the cosets $(C_i$'s) are pairwise disjoint and their union is G.

Let $[C_i]$ and \mathbf{c}_i be the incidence matrix and incidence vector of the cyclotomic coset C_i , defined as:

$$[C_i]_{a,b} = c^i_{ab},\tag{15}$$

$$\mathbf{c}_{i} = \left(c_{11}^{i}, c_{12}^{i}, \dots, c_{1(f-1)}^{i}\right)^{t},\tag{16}$$

where

$$c_{ab}^{i} = \begin{cases} 1, & \text{if } z_{b} - z_{a} \in C_{i}, \\ 0, & \text{otherwise.} \end{cases}$$
(17)

Linear combinations of $[C_i]$'s are used to obtain the D-optimal design matrices (Hunt and Wallis, 1972; Djokovic, 1991). However, this method of cyclotomic cosets using primitive root is not helpful, when n is not prime. In order to obtain D-optimal design matrices for any number n, Gysin proposed generalized cosets (Gysin and Seberry, 1998; Gysin, 1997). This relaxes the criterion of $n \in \mathbf{P}$ to $n \in \mathbb{Z}$. Moreover, the criterion of xbeing primitive root is relaxed to any number coprime to n.

3. Proposed Mathematical Formulation

From the Definitions 2 and 3, we have the two main equations that define D-optimal design. These equations are:

Constant periodic autocorrelation constraints:

$$\sum_{i=1}^{n} a_i a_{i\oplus s} + \sum_{i=1}^{n} b_i b_{i\oplus s} = 2 \quad \forall s = 1, \dots, \left\lfloor \frac{n}{2} \right\rfloor.$$

$$(18)$$

Diophantine equations:

$$\sum_{i=1}^{n} a_i = \alpha_1, \tag{19}$$

$$\sum_{i=1}^{n} b_i = \beta_1, \tag{20}$$

such that:

$$\alpha_1^2 + \beta_1^2 = 4n - 2. \tag{21}$$

Now, for the case when $n \equiv 0 \mod 3$ there are few more constraints called horizontal and vertical constraints proposed in Kotsireas *et al.* (2009), Kotsireas and Pardalos (2011):

$$\sum_{j=1}^{3} \widehat{A}_{j}^{2} + \sum_{j=1}^{3} \widehat{B}_{j}^{2} = 8m - 2,$$
(22)

$$\sum_{j=1}^{m} \widehat{C}_{j}^{2} + \sum_{j=1}^{m} \widehat{D}_{j}^{2} = 2n + 4,$$
(23)

where $m = \frac{n}{3}$, $\widehat{A}_j = \sum_{i=0}^{m-1} a_{3i+j}$, $\widehat{B}_j = \sum_{i=0}^{m-1} b_{3i+j}$, $\widehat{C}_j = \sum_{i=1}^{3} a_{j+(i-1)m}$ and $\widehat{D}_j = \sum_{i=1}^{3} b_{j+(i-1)m}$. Thus, a basic mathematical formulation that describes the D-optimal design will be:

find:

 $\{{f a},{f b}\}$

subject to:

(18) - (23)
$$a_i, b_i \in \{-1, +1\} \quad \forall i = 1, ..., n.$$
 (24)

The above formulation can be easily transformed into a 0-1 nonlinear (quadratic) formulation using the following transformation:

Let $x_i = (a_i + 1)/2$ and $y_i = (b_i + 1)/2$. Equation (18) can be transformed as:

$$\sum_{i=1}^{n} x_i x_{i \oplus s} + \sum_{i=1}^{n} y_i y_{i \oplus s} = \delta_1 \quad \forall s = 1, \dots, \left\lfloor \frac{n}{2} \right\rfloor,$$
(25)

where $\delta_1 = \frac{n+1-\alpha_1-\beta_1}{2}$. (Note that $\delta_1 = \lambda$ as defined in (12), which is the proof of equivalence of the D-optimal Definitions 2 and 3.)

Similarly, all the other equations can be transformed into binary equations. However, the formulation is non-linear (quadratic) in nature. One of the ways to solve such problems is to linearize the constraints. In the following part of this section, we will present the transformed linearized constraints:

$$\sum_{i=1}^{n} u_{i,(i\oplus s)} + \sum_{i=1}^{n} v_{i,(i\oplus s)} = \delta_1 \quad \forall s = 1, \dots, \left\lfloor \frac{n}{2} \right\rfloor,\tag{26}$$

$$\sum_{i=1}^{n} u_{i,j} \leqslant x_i, \tag{27}$$

$$\sum_{i=1}^{n} u_{i,j} \leqslant x_j, \tag{28}$$

$$\sum_{i=1}^{n} u_{i,j} \ge x_i + x_j - 1,$$
(29)

$$\sum_{i=1}^{n} v_{i,j} \leq y_i,$$

$$\sum_{i=1}^{n} v_{i,j} \leq y_j,$$
(30)
(31)

$$\sum_{j=1}^{n} v_{i,j} \leqslant y_j, \tag{31}$$

$$\sum_{i=1}^{n} v_{i,j} \ge y_i + y_j - 1, \tag{32}$$

$$\sum_{i=1}^{n} x_i = \delta_2, \tag{33}$$

$$\sum_{i=1}^{n} y_i = \delta_3, \tag{34}$$

$$\sum_{i=1}^{n} u_{i,j} = (\delta_2 - 1)x_j, \tag{35}$$

$$\sum_{i=1}^{n} v_{i,j} = (\delta_3 - 1)y_j, \tag{36}$$

$$\sum_{r=0}^{2} \sum_{\substack{i,j \\ i < j \\ i, j \in rv}} u_{(i\oplus r), (j\oplus r)} + \sum_{r=0}^{2} \sum_{\substack{i,j \\ i < j \\ i, j \in rv}} v_{(i\oplus r), (j\oplus r)} = \delta_4,$$
(37)

$$\sum_{s=0}^{m-1} \sum_{\substack{i,j\\i< j\\i,j\in rh}} u_{(i\oplus s),(j\oplus s)} + \sum_{s=0}^{m-1} \sum_{\substack{i,j\\i< j\\i,j\in rh}} v_{(i\oplus s),(j\oplus s)} = \delta_5,$$
(38)

where rh, rv are the set of indices, defined as $rh = \{i: i = 1 + rm, r = 0, ..., 2\}$ and $rv = \{i: i = 1 + 3s, s = 0, ..., (m-1)\}$. Moreover, $\delta_4 = \frac{1}{8}(8m - 2 - 6m^2 + 4(m-1))$

 $(\delta_2 + \delta_3)), \delta_5 = \frac{1}{8}(2n + 4 - 18m - 8(\delta_2 + \delta_3)), \delta_2 = \frac{n - \alpha_1}{2}$ and $\delta_3 = \frac{n - \beta_1}{2}$. Notice that, constraints (35), (36) are obtained by simple RLT technique (Sherali and Adams, 1999). Thus, the new 0-1 linear formulation will be:

find: $\{\mathbf{x}, \mathbf{y}\}\$ subject to: (26)-(38) $x_i, y_i \in \{0, 1\} \quad \forall i = 1, ..., n,$ (39)

with the above formulation, we can solve any D-optimal design problem. At this point it is a feasibility problem, and can be easily converted into an optimality problem by introducing slacks and/or by applying appropriate relaxations. However, as the value of n increases the mathematical program will be computationally inefficient in solving the D-optimal design problem. Nevertheless, in order to overcome this difficulty, we can use the help from the theory of cyclotomy.

4. Proposed Method

In this paper, we propose an extension of generalized cyclotomy cosets in deriving the D-optimal matrices. Our main hypothesis is that two different coprimes x's, can be used to construct the D-optimal design. This extension is a relaxation to the method proposed by Gysin and Seberry (1998), Gysin(1997). Moreover, the crux of our method is that we will use the knowledge of extended cyclotomic cosets in the proposed mathematical formulation. This eliminates the use of random searches in finding the solution to the D-optimal design matrices.

Let Φ be the matrix defined as:

$$\Phi = [\mathbf{c}_1, \dots, \mathbf{c}_{p_1}],\tag{40}$$

where c_i is the incidence vector of cyclotomic coset C_i with coprime w_1 . Similarly, let Ψ be the matrix defined as:

$$\Psi = [\mathbf{d}_1, \dots, \mathbf{d}_{p_2}],\tag{41}$$

where \mathbf{d}_i is the incidence vector of cyclotomic coset D_i with coprime w_2 .

Now, if the sequences **a**, **b** (or the first rows of matrices A&B) can be generated by the cyclotomic coset C_i and D_i , then the following constraints must be satisfied:

$$\mathbf{x} = \Phi \mathbf{h}_1,\tag{42}$$

$$\mathbf{y} = \Psi \mathbf{h}_2,\tag{43}$$

for some $\mathbf{h}_1 \in \{0,1\}^{p_1}$ and $\mathbf{h}_2 \in \{0,1\}^{p_2}$.

Therefore, the formulation using the generalized cyclotomic coset information will be written as:

find:

 $\{\mathbf{x}, \mathbf{y}\}$ subject to: (26)-(38) (42), (43) $x_i, y_i \ge 0 \quad \forall i = 1, \dots, n$ $\mathbf{h}_1 \in \{0, 1\}^{p_1}, \mathbf{h}_2 \in \{0, 1\}^{p_2}.$ (44)

Notice that, the resulting mathematical formulation has $p_1 + p_2$ binary variables only.

5. Results

We have used our proposed approach to find solutions of n = 63 known D-optimal designs. The results are shown in Fig. 1. It can be seen from the figure that, for $w_1 = 61$ and $w_2 = 2$, two different cosets are generated for two sequences **a**, **b**. Obviously, when $w_1 = w_2$ all the results of Gysin proposed in Gysin (1997), Gysin and Seberry (1998) can be generated. However, our goal was to show that we can relax the restriction of $w_1 = w_2$ in finding the D-optimal designs.

n	(α_1, β_1)	(w_1, w_2)	(\mathbf{a},\mathbf{b})
63	13	61	$\begin{bmatrix}++-+-+-++-+++-++++++++++++++-++++$
	9	2	[-+++++++++-+-+-+-+-+-+++++++++
63	13	61	[+-+++++++-++-++-+++-+++-+++-+++-+
	9	2	[+-++++++-+++++++++-+++++-++++

Fig. 1. D-optimal design sequences.

6. Conclusion

In the present paper, we have proposed a novel method to tackle the D-optimal design problem. This method is a part of ongoing research, where concepts from mathematical programming and number theory is being used to find the D-optimal design solutions. To the best of our knowledge, the proposed approach in this paper has not been found in the literature. By our proposed method, we have illustrated how the knowledge of optimization used in conjunction with the topics of number theory can be used in a methodological way for constructing the D-optimal designs.

References

- Chadjipantelis, T., Kounias, S., Moyssiadis, C. (1987). The maximum determinant of $21 \times 21(+1, -1)$ -matrices and D-optimal designs. *Journal of Statistical Planning and Inference*, 16, 167–178.
- Cusick, T.W., Ding, C., Renvall, A. (2004). *Stream Ciphers and Number Theory*, Vol. 66. Elsevier, Amsterdam.
- Djokovic, D.Z. (1991). On maximal (1, -1)-matrices of order 2n, n odd. *Radovi Matematicki*, 7, 371–378. Djokovic, D.Z., Kotsireas, I.S. (2011). New results on D-optimal matrices. Arxiv Preprint arXiv:1103.3626.
- Fisher, R.A. (1935). *The Design of Experiments*. Oliver & Boyd, Edinburg.
- Gysin, M. (1997). New D-optimal designs via cyclotomy and generalised cyclotomy. Australasian Journal of Combinatorics, 15, 247–256.
- Gysin, M., Seberry, J. (1998). An experimental search and new combinatorial designs via a generalisation of cyclotomy. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 27, 143–160.
- Hedayat, A., Wallis, W.D. (1978). Hadamard matrices and their applications. The Annals of Statistics, 6(6), 1184–1238.
- Horadam, K.J. (2007). Hadamard Matrices and Their Applications. Princeton University Press, Princeton.
- Hunt, D.C., Wallis, J. (1972). Cyclotomy, Hadamard arrays and supplementary difference sets. In: Proceedings

of the Second Manitoba Conference on Numerical Mathematics, pp. 351-381.

- Kotsireas, I.S., Pardalos, P.M. (2011). D-optimal matrices via quadratic integer optimization. *Journal of Heuris*tics, 1–11, DOI: 10.1007/s10732-011-9173-3.
- Kotsireas, I.S., Koukouvinos, C., Seberry, J. (2009). Weighing matrices and string sorting. Annals of Combinatorics, 13(3), 305–313.
- Kotsireas, I.S., Koukouvinos, C., Pardalos, P.M. (2010). An efficient string sorting algorithm for weighing matrices of small weight. *Optimization Letters*, 4(1), 29–36.
- Koukouvinos, C., Seberry, J., Whiteman, A.L., Xia, M. (1997). Optimal designs, supplementary difference sets and multipliers. *Journal of Statistical Planning and Inference*, 62, 1, 81–90.
- Mason, R.L., Gunst, R.F., Hess, J.L. (2003). Statistical Design and Analysis of Experiments: With Applications to Engineering and Science. Vol. 356, Wiley-Interscience, New York.

Sherali, H.D., Adams, W.P. (1999). Reformulation-linearization techniques for discrete optimization problems. In: *Handbook of Combinatorial Optimization*, Vol. 1, pp. 479–532.

Storer, T. (1967). Cyclotomy and Difference Sets, Vol. 2, Markham, Chicago.

Wallis, J.S. (1973). Some remarks on supplementary difference sets. In: Colloquia Mathematica Societatis Janos Bolyai, Vol. 10, pp. 1503–1526.

M.N. Syed is a PhD student at University of Florida, in Industrial and Systems Engineering Department. His areas of interest include mathematical programming, exact and heuristic optimization techniques, combinatorics, data mining.

I.S. Kotsirea is a professor at Wilfrid Laurier University, in Waterloo, Ontario, Canada. His research interests are in computational algebra, design theory, high-performance computing and discrete optimization. He serves on the Editorial Board of three international journals. His research is funded by grants from NSERC and SHARCNET.

P.M. Pardalos serves as distinguished professor of industrial and systems engineering at the University of Florida. He is also an affiliated faculty member of the Computer and Information Science Department, the Hellenic Studies Center, and the Biomedical Engineering Program. He is also the director of the Center for Applied Optimization. Dr. Pardalos is a world leading expert in global and combinatorial optimization. His recent research interests include network design problems, optimization in telecommunications, e-commerce, data mining, biomedical applications, and massive computing.

D-optimalus planavimas: matematinio programavimo taikymai pasinaudojant ciklotomones ko-aibes

Mujahid N. SYED, Ilias S. KOTSIREAS, Panos M. PARDALOS

Pasiūlytas naujas būdas D-optimaliems planams surasti rtemiantis matematinio programavimo ir skaičių teorijos koncepcijomis. Atskiru atveju matematiškai suformuluotas D-optimalaus plano apibrėžimas. Parodyta, kaip pasinaudojant ciklotominėmis ko-aibėmis galima sumažinti binarių kintamųjų skaičių. Pasiūlyto metodo taikymas pailiustruotas sprendžiant sunkų (N = 126) D-planavimo uždavinį.