A modified cyclic-coordinate exchange algorithm as illustrated by the construction of minimum-point second-order designs

Nam-Ky Nguyen[1]

Abstract

Box & Draper[1] reported five m-factor D-optimal minimum-point second order designs for $m = 2, \ldots, 5$. The m coordinates of each of the n design points lie in the interval $[-1, 1]$. This paper describes a method of constructing these designs and compares new designs with those of Box & Draper.

Keywords: computer-aided designs, D-optimality, second-order designs.

1 Introduction

A second-order response surface fitted for an experimental design with n runs and m explanatory variables may be written as:

$$y_i = \beta_0 + \sum_{j=1}^{m} \beta_j x_{ij} + \sum_{j=1}^{m} \sum_{j'=j}^{m} \beta_{jj'} x_{ij} x_{ij'} + e_i \tag{1}$$

where $i = 1, \ldots, n$; y is the response variable; and the m x's are scaled predictor variables. The minimum-point second order design problem can be formulated as being to find n design points $(x_{i1}, \ldots, x_{im})$, $i = 1, \ldots, n$ such that: (i) the coordinates of each design point lie in the interval $[-1, 1]$; (ii) n equals the number of parameters in the model, *i.e.* $(m + 1)(m + 2)/2$; and (iii) $|X'X|$ should take its maximum value where X is the $n \times n$ matrix whose ith row is $x_i' = (1, x_{i1}, \ldots, x_{im}, x_{i1}^2, x_{i1}x_{i2}, \ldots, x_{i,m-1}x_{im})$ (see Box & Draper[1], hereafter B&D.)

The following example describes a possible use of a minimum-point second-order design. In chocolate manufacture, conching converts the refined powdery ingredient mix into a suspension with the characteristic flow and flavour properties of chocolate. Over time, shear forces and heat combine to coat the suspended particles in fat and to release and develop essential flavours. Problems faced by the chocolate

[1]AV Biometrics, Department of Natural Resources & Environment, Institute for Horticultural Development, Private Bag 15, Ferntree Gully Delivery Centre, VIC 3156, Australia.
email: namky.nguyen@nre.vic.gov.au

Acknowledgment. The author would like to thank a referee for valuable comments and suggestions. This research was done while the author was a staff member of CSIRO Mathematical & Information Sciences.

manufacturer include minimizing the (costly) cocoa butter addition, reducing processing time and assessing new equipment while still maintaining product quality. A second-order design might be required to study the effects of four factors on the chocolate quality indices. These factors are (i) conching time (10-20 hours), (ii) mixing temperature (60^0-80^0C), (iii) mixing speed (100-500 RPM) and (iv) cocoa butter addition (25%-40%). While no restriction is put on the number of levels of each of the mentioned factors, the company aims to have the optimization results within one month during which a total of 15 runs could be completed.

2 Method of construction

A cyclic-coordinate exchange algorithm (CCEA) was used to construct the required designs. In a sense this algorithm is a hybrid of the adjustment algorithm (AA) of Donev & Atkinson[2] and the CCEA of Meyer & Nachtsheim[5]. The new CCEA starts with a non-singular D-optimal starting design D with points in $\{-1, 0, 1\}^m$. This design can be constructed by any of the conventional D-optimal design exchange algorithms (see Nguyen & Miller[6]). The new CCEA then follows the following steps:

1. For each row i of D, pick a coordinate $x_{ij},\ (j = 1, \ldots, m)$ at random and perturb this coordinate by an amount σ. Calculate $M = X'X + \epsilon I,\ |M|$ and M^{-1} where ϵ is a very small positive number say 1.0e-5. Note that the choice of n perturbations is empirically based and the the purpose of adding ϵI to $X'X$ is to avoid possible singularity caused by these perturbations.

2. Pick a coordinate $x_{ij},\ (i = 1, \ldots, n;\ j = 1, \ldots, m)$ at random among the nm coordinates of D.

3. Remove the ith row x_i' from X. Update $|M|$ and M^{-1} by the formula:

$$|M - x_i x_i'| = |M|(1 - x_i' M^{-1} x_i), \tag{2}$$

$$(M - x_i x_i')^{-1} = M^{-1} + u_i u_i'/(1 - x_i' u_i) \tag{3}$$

where $u_i = M^{-1} x_i$ (see Equations 3.3 and 3.4 of Nguyen & Miller[6]).

4. Reinstate the deleted ith row x_i' of X with the coordinate x_{ij} replaced by a value which circles from -1 to $+1$ in with step length σ. Let x_{ij}^* be a value which gives the biggest increase in $x_i' M^{-1} x_i$ where x_i' is now the reinstated row with x_{ij} replaced by x_{ij}^*. Update $|M|$ and M^{-1} by the formula:

$$|M + x_i x_i'| = |M|(1 + x_i' M^{-1} x_i), \tag{4}$$

$$(M + x_i x_i')^{-1} = M^{-1} - u_i u_i'/(1 + x_i' u_i) \tag{5}$$

where $u_i = M^{-1}x_i$ (see Equations 3.1 and 3.2 of Nguyen & Miller[6]).

Steps 2-4 are repeated until the number of coordinates is exhausted and no further improvement in $|M|$ is found. In these steps, σ is first set at 1 and then at 0.1 and finally at 0.01. From our experience, there is no further gain in decreasing σ further.

5. Remove the effect of ϵ in step 1 by recalculating $|X'X|$

The five steps of the above CCEA correspond to a single try. Several tries are made for each (m, n) combination and the try which corresponds to the highest $|X'X|$ will be chosen.

Remark:

1. Both the AA and the CCEA of Meyer & Nachtsheim[5] use a different starting design for each try. The new CCEA uses the same starting design for all the tries. This approach saves time as we only have to construct the starting design once. Note that there is a resemblance of this approach and the approach used in Nguyen & Williams[7] where different row-column designs are constructed from a single incomple block design by shuffling the treaments in each block of this block design.

2. With the AA, the candidate set for each coordinate x_{ij} of D is $\{x_{ij} - \sigma, x_{ij} + \sigma\}$. With the CCEA, the candidate set for each coordinate x_{ij} of D is $\{-1, -1 + \sigma, \ldots, -\sigma, 0, \sigma, \ldots, 1 - \sigma, 1\}$.

3. While the new CCEA works with all the n points, the one of Meyer & Nachtsheim[5] only works with k $(k < n)$ points which has the minimum $x_i'M^{-1}x_i$. While this practice saves some computer time, there is a substantial risk that their CCEA misses the best design.

4. While steps 3 and 4 in the new CCEA are separate steps, they are combined into single step in the one of Meyer & Nachtsheim[5]. Basically, they attempted to replace simultaneously row x_i' with another one with the coordinate x_{ij} replaced by x_{ij}^*. The update of $|M|$, unlike (2) and (4) requires more calculation (see Equations 3.5-3.9 of Nguyen & Miller[6]).

Table 1: Standardized $|X'X|$ for various designs, $2 \leq m \leq 5$

m	n	Box & Draper	Dubova & Federov	Donev & Atkinson	Nguyen
2	6	5.74e-3	5.72e-3	5.74e-3	5.74e-3
3	10	1.85e-4	1.85e-4	1.85e-4	1.85e-4
4	15	3.89e-7	3.37e-6	2.94e-6	3.45e-6
5	21	3.39e-11	-	8.46e-8	1.15e-7

3 Discussion

Table 1 presents the standardized $|X'X|$ ($= |X'X|/n^p$) for minimum point second-order designs for $m = 2, \ldots, 5$ of B&D, Dubova & Federov[3], Donev & Atkinson[2] and new ones.

The new designs for m=2 and m=3 have a pattern similar to the corresponding ones of B&D. The new design for m=4 have a pattern similar to the corresponding one of Dubova & Federov[3] (see Table 2 of B&D) except that the values of of the coordinates which are different form -1 or $+1$, i.e. $(\alpha, \beta, \gamma, \delta, \epsilon)$ are set at $(-0.22, -0.27, -0.63, 0.04, -0.04)$. Note that there was a misprint in Table 2 of B&D. The first coordinate of the first factor in this Table should be 1 instead of -1. If both δ and ϵ are set to 0, the maximum number of levels for each factor of this design (with the standardized $|X'X|$ equal to 3.44e-6) will be reduced from five to four. This is actually the design recommended for the experiment discussed in the Introduction. Table 2 gives the new design for m=5.

Table 2
Design for m=5†

x_1	x_2	x_3	x_4	x_5
-1	-1	-1	-1	-1
-1	1	1	1	1
1	-1	1	1	1
1	1	-1	1	1
1	1	1	-1	1
1	1	1	1	-1
1	1	-1	-1	-1
1	-1	1	-1	-1
1	-1	-1	-1	1
-1	1	-1	1	-1
-1	1	-1	-1	1
-1	-1	1	1	-1
-1	-1	1	-1	1
-1	-1	-1	1	1
α	-α	-1	1	-1
1	-1	-α	α	-1
-β	1	β	-1	-1
-1	β	1	-β	-1
γ	-1	1	γ	δ
-1	-γ	-γ	1	δ
-1	1	1	-1	-ϵ

†$(\alpha, \beta, \gamma, \delta, \epsilon)$=(0.31, 0.14,0.01,0.06,0.17)

Each factor of the new designs in Table 1 has at most five levels. Note that the

standardized $|X'X|$ for the corresponding D-optimal minimum-point second order designs in $\{-1, 0, 1\}^m$ for $m = 2, \ldots, 5$ are 5.49e-3, 1.33e-4 , 2.67e-6 and 7.89e-8. Hence, if there is no restriction to three level for each factor, the new designs in Table 1 are highly recommended.

As an additional note, the new CCEA improves $|X'X|$ of seven out of 17 designs (and matches $|X'X|$ of the remaining designs) in Table 1 of Donev & Atkinson[2]. The standardized $|X'X|$ of designs for (m, n)=(3,14),(3,16), (3,20), (4,15), (4,24), (5,21) and (5,26) in this table are 4.531e-4, 4.136e-4, 4.650e-4, 2.941e-6, 1.340e-5, 8.465e-8, 2.283e-7. The ones of the corresponding designs constructed by the new CCEA are 4.553e-4, 4.162e-4, 4.670e-4, 3.454e-6, 1.352e-5, 1.1474e-7 and 2.324e-7.

In this paper, minimum-point second-order designs have been constructed in order to illustrate the performance of the new CCEA. However, this CCEA should work with any practical values of m and n. The use of this CCEA in constructing large designs of the cyclic family such as cyclic incomplete block designs and Alpha designs will be discussed elsewhere (see John & Williams[4] Chapters 3-4 for the discussion of these types of designs). The byte code of the Java program implementing the CCEA discussed in this paper is available free of cost from the author.

References

[1] Box, M.J. & Draper, N.R. (1974). On minimum-point second-order designs, *Technometrics* 16, 613-616.

[2] Donev A.N. & Atkinson A.C. (1988). An adjustment algorithm for construction of the exact D-optimal experimental designs. *Technometrics* 30, 429-433.

[3] Dubova, I.S. & Federov, V.V. (1972). Tables of optimum design II (saturated optimal designs on a cube). Preprint No. 40, in Russian. Issued by Interfaculty Laboratory of Statistical methods, Moscow University.

[4] John, J.A. & Williams E.R. (1987). *Cyclic designs and computer-generated designs*. NY: Chapman & Hall.

[5] Meyer, R.K. & Nachtsheim C.J. (1995). The coordinate-exchange algorithm for constructing exact optimal experimental designs. *Technometrics* 37, 60-69.

[6] Nguyen N-K. & Miller, A.J. (1992) A review of exchange algorithms for constructing discrete D-optimal designs. *Comp. Stat. & Data Analysis* 14, 489-498.

[7] Nguyen N-K. & Williams, E.R. (1993) An algorithm for constructing optimal resolvable row-column designs. *Austral. J. of Stat.* 35, 363-370.