A Method for Constructing Mixed-level Supersaturated Design Assuring Dependencies of Paired Columns

Michiyo MATSUI[1], Shu YAMADA, Takenori TAKAHASHI
Science University of Tokyo, Depertment of Industrial Management and Engineering
1-3 Kagurazaka Shinjuku, Tokyo 162-8601, Japan
E-mail:matsui@ms.kagu.sut.ac.jp, shu@ms.kagu.sut.ac.jp, takenori@ms.kagu.sut.ac.jp
and
Dennis K. J. Lin
The Pennsylvania State University, University Park, PA 16802-1913, USA
E-mail:dkl5@psu.edu

1 Introduction

Supersaturated design is a kind of fractional factorial design in which the number of columns is greater than the number of rows. It is believed the usefulness for identifying a few effective factors, because it is possible to determine the optimum condition of the effective factors in a subsequent experiment.
In this paper, we discuss a construction of supersaturated designs consisting of two-level and three-level columns. This method generates mixed-level supersaturated designs based on an initial design matrix, where the number of experimental runs in the resulting designs is a multiple of the number of experimental runs in the initial design. The advantage of this method is that the dependency among pairs of columns in the resulting design is assured in a pre-specified level. Furthermore, some resulting designs are shown with evaluation of the dependency.

2 Preliminaries

Let $\boldsymbol{c}$ and $\mathcal{C}^n$ be a n-dimensional column vector whose elements are equal numbers of 1s, 2s, and the set of $\boldsymbol{c}$, respectively. In the same manner, $\boldsymbol{d}$ and $\mathcal{D}^n$ be a n-dimensional column vector whose elements are equal numbers of 1s, 2s, 3s, and the set of $\boldsymbol{d}$, respectively. We consider construction of mixed-level supersaturated design with n runs, where n is a multiple of 2 and 3 such that $n = 2 \times 3t = 6t$ and t is a positive integer.
The $n \times p$ matrix $\boldsymbol{C}^n = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_p)$ consisting of $\boldsymbol{c} \in \mathcal{C}^n$ is called 2-level design matrix. Furthermore, $(n \times (p+q))$ matrix $(\boldsymbol{C}^n, \boldsymbol{D}^n) = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_p, \boldsymbol{d}_1, \ldots, \boldsymbol{d}_q)$ is called mixed-level design consisting of 2-level and 3-level columns, where $\boldsymbol{c}_i \in \mathcal{C}^n$ $(i = 1, \ldots, p)$ and $\boldsymbol{d}_j \in \mathcal{D}^n$ $(i = 1, \ldots, p)$. An n-dimensional design matrix $(\boldsymbol{C}^n, \boldsymbol{D}^n) = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_p, \boldsymbol{d}_1, \ldots, \boldsymbol{d}_q)$ is supersaturated when

$$v = \frac{(2-1)\,p + (3-1)\,q}{n-1} > 1 \tag{1}$$

holds (Yamada and Matsui (1998)), where v is called the degree of saturation.
Let $n_{lm}(\boldsymbol{c}, \boldsymbol{d})$ be the number of rows whose values are (l, m) in the $n \times 2$ matrix $(\boldsymbol{c}, \boldsymbol{d})$, then

$$\sum_{l=1}^{2} \sum_{m=1}^{3} n_{lm}(\boldsymbol{c}, \boldsymbol{d}) = n.$$

The χ^2 statistic defined as

$$\chi^2(\boldsymbol{c},\, \boldsymbol{d}) = \sum_{l=1}^{2}\sum_{m=1}^{3} \frac{(n_{lm}(\boldsymbol{c},\boldsymbol{d}) - n/(2\times 3))^2}{n/(2\times 3)} \tag{2}$$

is used to evaluate the dependency between two columns $\boldsymbol{c}$ and $\boldsymbol{d}$, Yamada and Lin(1998). Although above equation can be used for designs with any number of levels, such as $\chi^2(\boldsymbol{c},\boldsymbol{c})$ and $\chi^2(\boldsymbol{d},\boldsymbol{d})$.

Next, we need to consider criteria to evaluate total non-orthogonality as a whole. The criteria are based on average of χ^2 values, which are derived from an analogy from the popular design criterion $E(s^2)$ given by Booth and Cox (1962). Let us consider a mixed-level supersaturated design $(\boldsymbol{C}^n,\, \boldsymbol{D}^n) = (\boldsymbol{c}_1,\, \dots,\, \boldsymbol{c}_p,\, \boldsymbol{d}_1, \dots, \boldsymbol{d}_q)$, $\boldsymbol{c}_i \in \mathcal{C}^n$ and $\boldsymbol{d}_j \in \mathcal{D}^n$. The average χ^2 values over all pairs of columns are

$$\begin{aligned}
\mathrm{ave}\chi^2(\boldsymbol{C}^n) &= \sum_{1\le i<j\le p} \chi^2(\boldsymbol{c}_i,\boldsymbol{c}_j) \Big/ \binom{p}{2}, && (3)\\
\mathrm{ave}\chi^2(\boldsymbol{C}^n,\boldsymbol{D}^n) &= \sum_{1\le i\le p}\sum_{1\le j\le q} \chi^2(\boldsymbol{c}_i,\boldsymbol{d}_j)/\,(pq)\,, && (4)\\
\mathrm{ave}\chi^2(\boldsymbol{D}^n) &= \sum_{1\le i<j\le q} \chi^2(\boldsymbol{d}_i,\boldsymbol{d}_j) \Big/ \binom{q}{2}, && (5)
\end{aligned}$$

respectively.

Another criteria is maximum value of χ^2 values in the design matrix $(\boldsymbol{C}^n,\, \boldsymbol{D}^n)$. The criteria are defined by

$$\begin{aligned}
\max\chi^2(\boldsymbol{C}^n) &= \max\left\{\chi^2(\boldsymbol{c}_i,\boldsymbol{c}_j) \mid 1\le i<j\le p\right\} && (6)\\
\max\chi^2(\boldsymbol{C}^n,\boldsymbol{D}^n) &= \max\left\{\chi^2(\boldsymbol{c}_i,\boldsymbol{d}_j) \mid 1\le i\le p, 1\le j\le q\right\} && (7)\\
\max\chi^2(\boldsymbol{D}^n) &= \max\left\{\chi^2(\boldsymbol{d}_i,\boldsymbol{d}_j) \mid 1\le i<j\le q\right\}. && (8)
\end{aligned}$$

The problem for construction of mixed-level supersaturated design is an enumeration of the columns from the sets $\mathcal{C}^n$ and $\mathcal{D}^n$, while maintaining low dependency, *i.e.*, maintaining low level of $\mathrm{ave}\chi^2(\boldsymbol{C}^n)$, $\mathrm{ave}\chi^2(\boldsymbol{C}^n,\boldsymbol{D}^n)$, $\mathrm{ave}\chi^2(\boldsymbol{D}^n)$, $\max\chi^2(\boldsymbol{C}^n)$, $\max\chi^2(\boldsymbol{C}^n,\boldsymbol{D}^n)$ and $\max\chi^2(\boldsymbol{D}^n)$.

3 General construction method

At first, an initial design matrix with $n = 6(= 2\times 3)$ rows is constructed in considerations of the non-orthogonality, say $\left(\boldsymbol{C}^6,\, \boldsymbol{D}^6\right)$. The second, the following equation are used to generate design matrices :

$$t=2: \boldsymbol{C}^{6t} = \begin{pmatrix} 1^6 & \boldsymbol{C}^6_{(0)} & \boldsymbol{C}^6_{(0)} & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(0)} \\ 2^6 & \boldsymbol{C}^6_{(0)} & \boldsymbol{C}^6_{(1)} & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(1)} & \boldsymbol{D}^6_{(2)} \end{pmatrix}, \tag{9}$$

$$t=3: \boldsymbol{C}^{6t} = \begin{pmatrix} \boldsymbol{C}^6_{(0)} & \boldsymbol{C}^6_{(0)} & \boldsymbol{C}^6_{(1)} & 1^6 & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(1)} \\ \boldsymbol{C}^6_{(0)} & \boldsymbol{C}^6_{(1)} & \boldsymbol{C}^6_{(0)} & 2^6 & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(1)} & \boldsymbol{D}^6_{(0)} \\ \boldsymbol{C}^6_{(1)} & \boldsymbol{C}^6_{(0)} & \boldsymbol{C}^6_{(0)} & 3^6 & \boldsymbol{D}^6_{(1)} & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(0)} \end{pmatrix}, \tag{10}$$

$$t=4: \boldsymbol{C}^{6t} = \begin{pmatrix} 1^6 & 1^6 & 1^6 & \boldsymbol{C}^6_{(0)} & \boldsymbol{C}^6_{(0)} & \boldsymbol{C}^6_{(0)} & \boldsymbol{C}^6_{(0)} & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(0)} \\ 1^6 & 2^6 & 2^6 & \boldsymbol{C}^6_{(0)} & \boldsymbol{C}^6_{(0)} & \boldsymbol{C}^6_{(1)} & \boldsymbol{C}^6_{(1)} & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(1)} & \boldsymbol{D}^6_{(1)} \\ 2^6 & 1^6 & 2^6 & \boldsymbol{C}^6_{(0)} & \boldsymbol{C}^6_{(1)} & \boldsymbol{C}^6_{(0)} & \boldsymbol{C}^6_{(1)} & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(1)} & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(1)} \\ 2^6 & 2^6 & 1^6 & \boldsymbol{C}^6_{(0)} & \boldsymbol{C}^6_{(1)} & \boldsymbol{C}^6_{(1)} & \boldsymbol{C}^6_{(0)} & \boldsymbol{D}^6_{(0)} & \boldsymbol{D}^6_{(1)} & \boldsymbol{D}^6_{(1)} & \boldsymbol{D}^6_{(0)} \end{pmatrix}. \tag{11}$$

Specifically, Equation (11) generates an $6t \times pt$ design matrix $\boldsymbol{C}^{6t}$ and an $6t \times qt$ design matrix $\boldsymbol{D}^{6t}$ from an $6 \times p$ matrix $\boldsymbol{C}^6$ and $6 \times q$ matrix $\boldsymbol{D}^6$, respectively. Furthermore, $\boldsymbol{C}^{lm}_{(a)}$ and $\boldsymbol{D}^{lm}_{(a)}$ denote the matrices whose elements are replaced by $\mathrm{mod}(a+c,l)$ and $\mathrm{mod}(a+d,m)$ from the original elements c and d, respectively. The above procedure is justified by the following theorem.

Theorem 1 *Regarding the maximum χ^2 values in Equation (11), the following relations holds:*

$$\begin{aligned}
\max\{\chi^2\left(\boldsymbol{C}^{6t}\right)\} &= t \times \max\{\chi^2\left(\boldsymbol{C}^{6}\right)\}, && (12)\\
\max\{\chi^2\left(\boldsymbol{C}^{6t}, \boldsymbol{D}^{6t}\right)\} &= 0, && (13)\\
\max\{\chi^2\left(\boldsymbol{C}^{6t}\right)\} &= t \times \max\{\chi^2\left(\boldsymbol{D}^{6}\right)\}. && (14)
\end{aligned}$$

Proof 1 *Assume that there exsists an initial matrix $\boldsymbol{C}^6$ with $n = N$ rows consisting of 2-level columns.*

$$\chi^2\left(\boldsymbol{C}^6\right) = \frac{2\times 2}{N} \times \{2(n_{11} - \frac{N}{2\times 2})^2 + 2(n_{12} - \frac{N}{2\times 2})^2\} \qquad (15)$$

On the other hand,

$$\begin{aligned}
\chi^2\left(\boldsymbol{C}^{6t}\right) &= \frac{2\times 2}{tN} \times \{2(tn_{11} - \frac{tN}{2\times 2})^2 + 2(tn_{12} - \frac{tN}{2\times 2})^2\}\\
&= \frac{2\times 2}{N} \times t\{2(n_{11} - \frac{N}{2\times 2})^2 + 2(n_{12} - \frac{N}{2\times 2})^2\} = t\chi^2(C^6)
\end{aligned}$$

where tN is the number of experimental runs.

We consider a design that attains the lower dependency as above by following lemma.

Lemma 1 *Regarding the average χ^2 values (or sum of χ^2 values)in Equation (11), the following relations holds:*
in case of t=1,2,4,

$$sum\{\chi^2(C^{6t})\} = \left(\begin{array}{c} p \\ 2 \end{array}\right) \times max\{\chi^2(C^{6t})\} \times t \qquad (16)$$

on the other hand, t=3,

$$sum\{\chi^2(C^{6t})\} = \{t\left(\begin{array}{c} p \\ 2 \end{array}\right) + p\left(\begin{array}{c} t \\ 2 \end{array}\right)\} \times max\{\chi^2(C^{6t})\} + (p \times tp - \left(\begin{array}{c} t \\ 2 \end{array}\right)) \times min\{\chi^2(C^{6t})\} (17)$$

where p, q are the number of two-level columns, that of three-level ones, respectively.

4 Main results

According to constructions in the previous section, the initial design matrices with $n = 6$ runs are needed.
For any $\boldsymbol{c} \in \mathcal{C}^6$ and $\boldsymbol{d} \in \mathcal{D}^6$, the χ^2 statistics varies as follows:

$$\begin{aligned}
\chi^2\left(\boldsymbol{c}_i, \boldsymbol{c}_j\right) &= \frac{2}{3}, 6 && (18)\\
\chi^2\left(\boldsymbol{c}_i, \boldsymbol{d}_j\right) &= 0, 4 && (19)\\
\chi^2\left(\boldsymbol{d}_i, \boldsymbol{d}_j\right) &= 3, 6, 12. && (20)
\end{aligned}$$

Table 1: Evaluation of construction supersaturated design.

criteria	t	two-level $\boldsymbol{C}_a^{6t}$	$\boldsymbol{C}_b^{6t}$	three-level $\boldsymbol{D}_a^{6t}$	$\boldsymbol{D}_b^{6t}$
Number of columns	1	4	1	1	3
$max\chi^2$		0.67	—	—	3.00
$average\chi^2$		0.67	—	—	3.00
Number of columns	2	9	3	3	8
$max\chi^2$		1.33	0	6.00	6.00
$average\chi^2$		0.44	0	6.00	3.57
Number of columns	3	12	3	4	10
$max\chi^2$		2.00	2.00	0	9.00
$average\chi^2$		1.03	0.67	0	1.80
Number of columns	4	19	7	4	12
$max\chi^2$		2.67	0	12.00	12.00
$average\chi^2$		0.37	0	7.50	6.30

Since the relations $\chi^2(\boldsymbol{c}_i, \boldsymbol{c}_j) = 6$ and $\chi^2(\boldsymbol{d}_{i'}, \boldsymbol{d}_{j'}) = 12$ imply fully alias (confounded), we need to enumerate columns from the sets $\mathcal{C}^6$ and $\mathcal{D}^6$ while maintaining $\chi^2(\boldsymbol{c}_i, \boldsymbol{c}_j) = 2/3$, $\chi^2(\boldsymbol{d}_i, \boldsymbol{d}_j) = 3, 6$. The pairs of designs satisfy the followings:

$$\chi^2(\boldsymbol{c}_i, \boldsymbol{c}_j) = \frac{2}{3}, \chi^2(\boldsymbol{c}_i, \boldsymbol{d}_j) = 0, \chi^2(\boldsymbol{d}_i, \boldsymbol{d}_j) = 3. \tag{21}$$

$$(\boldsymbol{C}_a^6, \boldsymbol{D}_a^6) = \begin{pmatrix} 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 2 & 1 & 2 \\ 1 & 2 & 1 & 1 & 3 \\ 2 & 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 2 & 2 \\ 2 & 1 & 2 & 2 & 3 \end{pmatrix}, \quad (\boldsymbol{C}_b^6, \boldsymbol{D}_b^6) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 3 & 3 & 3 \\ 2 & 1 & 2 & 3 \\ 2 & 2 & 3 & 1 \\ 2 & 3 & 1 & 2 \end{pmatrix} \tag{22}$$

The general construction in the above produces the mixed-level supersaturated designs by the previous initial design matrices. For example, $\boldsymbol{C} = \boldsymbol{C}_a^6$ and $\boldsymbol{D} = \boldsymbol{D}_a^6$ are generated mixed level supersaturated design with $n = 6t = 12$ runs with $(4+1)t = 10$ columns for $t = 2$.

REFERECES

[1] Lin, D. K. J., (1993),gA New Class of Supersaturated Designs," *Technometrics*, 35, pp.28-31.

[2] Yamada, S. and Matsui, T., (1998),gOptimality of Mixed-level Supersaturated Designs," Tech. Rep. University of Tokyo, Dept. Mathmatical Engineering and Information physics, METR $98-05$.

[3] Yamada, S. and Lin, D. K. J., (1999),gThree-level supersaturated design," Statistics and Probability Letters, 45 31-39.