Statistica Sinica **9**(1999), 605-610

# A RESOLUTION RANK CRITERION FOR SUPERSATURATED DESIGNS

Lih-Yuan Deng, Dennis K. J. Lin and Jiannong Wang

*University of Memphis, Pennsylvania State University and Covance Clinical and Periapproval Service, Inc.*

*Abstract:* Hadamard matrices are found to be useful in constructing supersaturated designs. In this paper, we study a special form of supersaturated designs using Hadamard matrices. Properties of such a supersaturated design are discussed. It is shown that the popular $E(s^2)$ criterion is in general inadequate to measure the goodness of a supersaturated design. A new criterion based upon the projection property, called resolution rank (r-rank), is proposed. Furthermore, an upper bound for r-rank is given for practical use.

*Key words and phrases:* Hadamard matrix, projective designs.

## 1. Introduction

When the number of factors is large and a small number of runs is desired, a supersaturated design can save considerable cost. A two-level supersaturated design is a fraction of a factorial design with $n$ observations in which the number of factors $k$ is larger than $n-1$. The usefulness of such a supersaturated design relies upon the realism of effect sparsity, namely, that the number of dominant active factors is small. The goal is to identify these active factors with so-called screening experimentation. (A brief review of early work on supersaturated designs is available from Lin (1991).)

Apart from some ad hoc procedures and computer-generated designs, the construction problem has not been addressed until very recently (see Lin (1993a) and (1995), Wu (1993), and Tang and Wu (1997)). Most of these supersaturated designs were constructed based on Hadamard matrices. In this paper, a special form of supersaturated designs using Hadamard matrices is studied. Furthermore, a criterion based upon the projection property called resolution rank is proposed to further differentiate among designs.

## 2. The Construction Method

Consider a supersaturated design of the form

$$\mathbf{X_c} = [\mathbf{H}_1, \mathbf{H}_2\mathbf{C}], \tag{2.1}$$

where $\mathbf{H}_1, \mathbf{H}_2$ are two normalized Hadamard matrices of dimension $n \times n$, and $\mathbf{C}$ is an $n \times (n-c)$ matrix representing the operation of column selection so that all fully aliased columns are removed. For example, if $c$ columns of $\mathbf{H}_1$ need to be removed, the corresponding $\mathbf{C}$ can be obtained by deleting $c$ columns from $\mathbf{I}_n$, the identity matrix of order $n$. Clearly, the column of $1's$ (or $-1's$) in $\mathbf{H}_1$ will not be assigned to any factor and the total number of factors (columns) in (2.1) that can be considered is $k = (n-1) + (n-c) = 2n - c - 1$. To make a clearer and simpler presentation of our results, we keep the column of $1's$ in (2.1). It can be seen that this method is equivalent to a method proposed by Tang and Wu (1997).

The supersaturated design $\mathbf{X}_c$ proposed here includes several interesting special cases. First, when $\mathbf{H}_2 = \mathbf{D}(\mathbf{h}_i)\mathbf{H}_1$, where $\mathbf{D}(\mathbf{h}_i)$ is the diagonal matrix with diagonal elements equal to the elements of $\mathbf{h}_i$, the $i$th column of $\mathbf{H}_1, 1 < i \le n$, and $\mathbf{C}$ is the matrix corresponding to the deletion of the first and $i$th columns, $\mathbf{X}_c$ is the same as the design obtained by the product method proposed by Wu (1993). When $\mathbf{P}$ is a matrix corresponding to a permutation of the row vectors in $\mathbf{H}_1$, $\mathbf{H}_2 = \mathbf{P}\mathbf{D}(\pm\mathbf{h}_i)\mathbf{H}_1$ corresponds to the operation of permuting rows, following the product method. A more general form of supersaturated designs is given in Cheng (1997). We will focus on the design of (2.1) here, however. Note also that (2.1) can be easily extended to the form of $(\mathbf{H}\mathbf{C}_1, \mathbf{H}_2\mathbf{C}_2, \ldots, \mathbf{H}_K\mathbf{C}_K)$, if so desired.

## 3. Main Results

Denote the $i$th column of $\mathbf{H}_1$ as $\mathbf{h}_i$, and the $j$th column of $\mathbf{H}_2$ as $\mathbf{k}_j$. For any supersaturated design $\mathbf{X}_c$ in (2.1), it can be shown that

$$\mathbf{X}_c'\mathbf{X}_c = \begin{pmatrix} n\mathbf{I}_n & \mathbf{H}_1'\mathbf{H}_2\mathbf{C} \\ \mathbf{C}'\mathbf{H}_2'\mathbf{H}_1 & n\mathbf{I}_{n-c} \end{pmatrix} = \begin{pmatrix} n\mathbf{I}_n & \mathbf{W}\mathbf{C} \\ \mathbf{C}'\mathbf{W}' & n\mathbf{I}_{n-c} \end{pmatrix},$$

where

$$\mathbf{W} = \mathbf{H}_1'\mathbf{H}_2 = (w_{ij}) = (\mathbf{h}_i'\mathbf{k}_j).$$

Let $s_{ij}$ denote the $(i,j)$ entry of $\mathbf{X}_c'\mathbf{X}_c$. Then $E(s^2)$ of $\mathbf{X}_c$, proposed by Booth and Cox (1962), can be defined as

$$E(s^2) = \sum_{i<j} s_{ij}^2 / \binom{k}{2},$$

where $k = 2n - c - 1$ is the number of columns (excluding the column of 1's) in $\mathbf{X}_c$. Clearly, $E(s^2)$ is equal to a constant times $\sum_{j\in\mathbf{C}}\sum_{i=1}^{n} w_{ij}^2$ (here, $\sum_{j\in\mathbf{C}}$ denotes a summation over the $(n-c)$ columns selected by the matrix $\mathbf{C}$), which stays constant as shown in Theorem 1.

**Theorem 1.** *Let* $\mathbf{W} = \mathbf{H}_1'\mathbf{H}_2$*, where* $\mathbf{H}_1, \mathbf{H}_2$ *are Hadamard matrices of dimension* $n \times n$*. Then we have*

(1) $\frac{1}{n}\mathbf{W}$ *is an* $n \times n$ *orthogonal matrix;*

(2) $n^2 = \sum_{i=1}^n w_{ij}^2 = \sum_{j=1}^n w_{ij}^2$;

(3) $w_{ij}$ *is always a multiple of* 4.

**Corollary 1.** *For any* $\mathbf{C}, \mathbf{H}_1$ *and* $\mathbf{H}_2$, *all* $\mathbf{X}_c$ *in* (2.1) *have* $E(s^2) = \frac{2n^2(n-c)}{(2n-c-1)(2n-c-2)}$.

Corollary 1 implies that the popular criterion $E(s^2)$ used in supersaturated design theory is invariant for many choices of $\mathbf{C}$. Therefore, it is not effective in comparing supersaturated designs. Wu (1993) and Deng, Lin and Wang (1996a) extend the classical design optimalities and propose to compute the average $D_f$ ($D$-optimal) and $A_f$ ($A$-optimal) criteria over the projected submatrices of $f$ columns to select a better supersaturated design. A large average $D_f$, however, does not ensure that every projective design is nonsingular (namely, the identifiability in terms of the projective design). One important feature of the goodness of a supersaturated design is such a projective property (see Lin (1993b) and Cheng (1995)). We thus consider the $r$-rank property as defined below.

**Definition.** Let $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_k)$ be an $n \times k$ matrix. The *resolution-rank* of $\mathbf{X}$ ($r$-rank, for short) is defined as $r = \max\{c:$ for any $(\mathbf{x}_{i1}, \ldots, \mathbf{x}_{ic})$ of $\mathbf{X}, \mathbf{x}_{i1}, \ldots, \mathbf{x}_{ic}$ are linearly independent$\}$.

Clearly, if a supersaturated design $\mathbf{X}$ has an r-rank of $f$, then when $\mathbf{X}$ is projected to any submatrix of $f$ (or fewer) factors, the main effects of the projected design are all estimable. Moreover, in many situations where two supersaturated designs have nearly identical $D_f$ and $A_f$ values (differences are less than 1%), their $r$-ranks can be very different. As an illustrative example, let $\mathbf{H}_1$ be the Hadamard matrix of "V. 3/8 Group" given in Hall (1961). Define $\mathbf{X}_a = [\mathbf{H}_1, \mathbf{H}a\mathbf{C}]$ and $\mathbf{X}_b = [\mathbf{H}_1, \mathbf{H}_b\mathbf{C}]$, where $\mathbf{H}_a$ and $\mathbf{H}_b$ are permutation matrices resulting from the row orders $(12, 15, 11, 7, 6, 3, 2, 14, 16, 5, 1, 4, 10, 13, 8, 9)$ and $(10, 15, 14, 12, 3, 9, 16, 7, 4, 6, 13, 5, 1, 2, 8, 11)$ of $\mathbf{H}_1$, respectively, and $\mathbf{C}$ is a matrix deleting the first column of the identity matrix of order $n$. It can be shown that both $\mathbf{X}_a$ and $\mathbf{X}_b$ have $E(s^2) = 8.83$ while their r-ranks are 4 and 7, respectively. For other examples, see Deng, Lin and Wang ((1994), Table 1).

**Remark.** It has come to our attention that the r-rank criterion is closely related to the $P_t$ property given by Srivastava (1975) in the context of search design. The present paper is a condensed version of Deng, Lin and Wang's 1994 IBM Technical Reprot where the idea of $r$-rank was first proposed. There has been some subsequent work since then. For example, Deng, Lin and Wang (1996b) used the $r$-rank criterion as an optimality criterion for adding columns to any orthogonal array.

One way to check the r-rank of a supersaturated design is to check all $\binom{k}{f}$ sub-matrices for $f = 2, 3 \ldots$ etc. This is very time consuming, even for moderate

$k$ and $f$, although it is faster than evaluating $D_f$ and $A_f$. We shall derive a simple upper bound on the $r$-rank.

**Lemma 1.** *Let* $\mathbf{W} = \mathbf{H}_1'\mathbf{H}_2 = (w_{ij})$. *Then* $n\mathbf{k}_j = \sum_{i=1}^n w_{ij}\mathbf{h}_i$, *where* $\mathbf{h}_i$ *is the* $i$*th column of* $\mathbf{H}_1$ *and* $\mathbf{k}_j$ *is the* $j$*th column of* $\mathbf{H}_2$.

As we can see from Lemma 1, if many $w_{ij}$ are zero, then $\mathbf{k}_j$ will be a linear combination of only a few $\mathbf{h}_i$. Let $r_1 = \min\{|U_j|, j \in \mathbf{C}\}$, where $U_j = \{i|w_{ij} \neq 0, i = 1, 2, \ldots, n\}$, $|S|$ is the number of elements in a set $S$, and $j \in \mathbf{C}$ means the $j$th column of $\mathbf{H}_2$ selected by $\mathbf{C}$. Similarly, for fixed $j_1 \neq j_2 \in \mathbf{C}$, let $r_2 = \min\{|U^+_{j1,j2}|, |U^1_{j1,j2}|, j_1 \neq j_2 \in \mathbf{C}\} + 1$, where $U^+_{j1j2} = \{i|w_{ij_1} + w_{ij_2} \neq 0, i = 1, 2, \ldots, n\}$ and $U^-_{j1j2} = \{i|w_{ij_1} - w_{ij_2} \neq 0, i = 1, 2, \ldots, n\}$. Then we have the following theorem.

**Theorem 2.** $r$*-rank* $\leq \min\{r_1, r_2\} = r_e$.

This upper bound for the $r$-rank has some obvious advantages. First, it is very easy to compute. According to our study, it is at least 10,000 times faster than calculating the actual $r$-rank (and $D_f, A_f$ criteria). Second, $r_e$ can easily screen out many undesirable (e.g., low $r$-rank) supersaturated designs $\mathbf{X}_c$ of the form in (2.1). Third, empirical studies show that $r_e$ is a good estimate of the $r$-rank. Consider the two supersaturated designs $\mathbf{X}_a$ and $\mathbf{X}_b$ based on a Hadamard matrix of order 16 as previously given, for example. For $\mathbf{X}_a$, $r_1 = 4$ and $r_2 = 6$, and thus $r_e = \min(r_1, r_2) = 4$. For $\mathbf{X}_b$, $r_1 = 7$ and $r_2 = 8$, and thus $r_e = \min(r_1, r_2) = 7$. A straightforward evaluation of all possible column combinations confirms that the $r$-ranks of $\mathbf{X}_a$ and $\mathbf{X}_b$ are, indeed, 4 and 7, respectively.

The special case $\mathbf{H}_2 = \mathbf{D}(\mathbf{h}_l)\mathbf{H}_1$ results in some nice properties of the $\mathbf{W} = \mathbf{H}_1'\mathbf{D}(\mathbf{h}_l)\mathbf{H}_1$ matrix as stated in Theorem 3.

**Theorem 3.** *Let* $\mathbf{W} = \mathbf{H}_1'\mathbf{D}(\mathbf{h}_l)\mathbf{H}_1$, *where* $\mathbf{H}_1$ *is a Hadamard matrix of size* $n = 4t$ *and* $\mathbf{D}(\mathbf{h}_l)$ *is the diagonal matrix associated with* $\mathbf{h}_l$, *the* $l$*th column vector of* $\mathbf{H}_1$, $\mathbf{h}_l \neq \pm 1$, *for* $2 \leq l \leq n$. *For a column vector of* $\mathbf{W}$ *such that none of the entries has the value of* $\pm n$, *we have*

(1) *If* $t$ *is odd, then there are exactly three* 0 *in each column of* $\mathbf{W}$. *The rest of* $w_{ij}$ *in* $\mathbf{W}$ *can only be of the form* $\pm 8w + 4$, *for some non-negative integer* $w$;

(2) *If* $t$ *is even, then every entry* $w_{ij}$ *in* $\mathbf{W}$ *can only be of the form* $\pm 8w$, *for some non-negative integer* $w$.

There are some implications for the product method from the above theorem:

(1) When $t$ is even, let $w_{ij} = 8u_{ij}$, where $u_{ij}$ is an integer. From Theorem 2, the $u_{ij}$'s satisfy $\sum_{i=1}^n u_{ij}^2 = \sum_{j=1}^n u_{ij}^2 = t^2/4$. Hence, there are at most $t^2/4$ of $w_{ij}$ that are non-zero. For example, for any Hadamard matrix with $n = 16 (t = 4)$, there are at most $4^2/4 = 4$ non-zero elements in any column of $\mathbf{W}$. Furthermore, every non-zero element must be $\pm 8$ or $\pm 16$. The case of $w_{ij} = \pm 16$ corresponds

to having the $j$-column of $\mathbf{D}(\mathbf{h}_l)\mathbf{H}_1$ fully aliased with the $i$th column of $\mathbf{H}_1$. Therefore, there are exactly 4 elements in each column vector of $\mathbf{WC}$ with values $\pm 8$, and thus the $r$-rank is at most 4. In general, the product method produces designs with a smaller $r_1$ than that from the permutation method. Hence, the product method is not recommended when $t$ is even.

(2) When $t$ is odd, there are $(n-3)$ non-zero $w_{ij}$ which may make $r_1$ large. For example, for a Hadamard matrix with $n = 12 (t = 3)$, there are exactly $n - 3 = 9$ non-zero elements in any column in $\mathbf{WC}$. Furthermore, every non-zero entry of $w_{ij}$ must be $\pm 4$. In this case, we can easily see that the value of $r_1$ by the product method can be the maximum value.

## Appendix. Proofs

**Proof of Theorem 1.** Parts (1) and (2) are true because $\mathbf{W}'\mathbf{W} = \mathbf{H}_2'\mathbf{H}_1\mathbf{H}_1'\mathbf{H}_2 = n\mathbf{H}_2'\mathbf{H}_2 = n^2\mathbf{I}$. Part (3) is a well known fact in the weighing design literature, since both $\mathbf{h}_i$ and $\mathbf{k}_j$ have even number of $\pm 1$'s.

**Proof of Lemma 1.** Pre-multiplying $\mathbf{H}_1$ in both sides of the equation $\mathbf{W} = \mathbf{H}_1'\mathbf{H}_2$ and using the fact that $\mathbf{H}_1\mathbf{H}_1' = n\mathbf{I}$, we have

$$n\mathbf{H}_2 = \mathbf{H}_1\mathbf{W} = \mathbf{H}_1(\mathbf{w}_1, \ldots, \mathbf{w}_n) = (\mathbf{H}_1\mathbf{w}_1, \ldots, \mathbf{H}_1\mathbf{w}_n).$$

Therefore the $j$th column of $n\mathbf{H}_2$ is

$$n\mathbf{k}_j = (\mathbf{h}_1, \ldots, \mathbf{h}_n)(w_{1j}, \ldots, w_{nj})' = \sum_{i=1}^{n} w_{ij}\mathbf{h}_i.$$

**Proof of Theorem 2.** From Lemma 1 we have $n\mathbf{k}_j = \sum_{i\in U_j} w_{ij}\mathbf{h}_i$. Therefore, the $r$-rank $\le |U_j|$ for each $j \in \mathbf{C}$, and is $\le r_1$. Similarly, by applying Lemma 1 for fixed $j_1 \ne j_2$, we have $r$-rank $\le |U^+_{j1j2}| + 1, r$-rank $\le |U^-_{j1j2}| + 1$. Theorem 2 now follows easily.

**Proof of Theorem 3.** Without loss of generality, we assume $\mathbf{h}_1 = \pm 1$ is the first column of $\mathbf{H}_1$. For a fixed column say, column $j$ $(j \ge 2)$ of $\mathbf{W}$, it is easy to see that when $i = 1$, or $i = j$ or $i = l$, $w_{ij} = \mathbf{h}'_i\mathbf{D}(\mathbf{h}_l)\mathbf{h}_j = 0$. Suppose now that $(i, j, l)$ are all different and all are not equal to 1. Let $d_{++} = \{m | h_{im} = h_{lm} = +1\}$, $d_{+-} = \{m | h_{im} = +1, h_{lm} = -1\}$, $d_{-+} = \{m | h_{im} = -1, h_{lm} = +1\}$, $d_{--} = \{m | h_{im} = h_{lm} = -1\}$. Then $|d_{++}| = |d_{+-}| = |d_{-+}| = |d_{--}| = n/4 = t$. Let $q_{++} = \{m \in d_{++} | h_{jm} = +1\}$, $q_{+-} = \{m \in d_{+-} | h_{jm} = +1\}$, $q_{-+} = \{m \in d_{-+} | h_{jm} = +1\}$, $q_{--} = \{m \in d_{--} | h_{jm} = +1\}$. Using the conditions $\mathbf{h}_j'\mathbf{h}_i = 0, \mathbf{h}_j'\mathbf{h}_l = 0, \mathbf{h}_j'1 = 0$, and letting $Q$ be the number of elements in $q_{++}$ , we can see that $|q_{++}| = |q_{--}| = Q, \quad |q_{+-}| = |q_{-+}| = t - Q$. Then

$$\begin{aligned} w_{ij} &= \sum_{m=1}^{n} h_{im}h_{lm}h_{jm} = \sum_{m\in d_{++}} h_{jm} - \sum_{m\in d_{+-}} h_{jm} - \sum_{m\in d_{-+}} h_{jm} + \sum_{m\in d_{--}} h_{jm} \\ &= [Q-(t-Q)] + [Q-(t-Q)] + [Q-(t-Q)] + [Q-(t-Q)] = 4(2Q-t). \end{aligned}$$

If $t$ is odd, then $w_{ij}$ is of the form $\pm 8w + 4$ and if $t$ is even, then $w_{ij}$ is of the form $8w$.

## References

Booth, K. H. V. and Cox, D. R. (1962). Some systematic supersaturated designs. *Technometrics* **4**, 489-495.

Cheng, C. S. (1995). Some projection properties of orthogonal arrays. *Ann. Statist.* **23**, 1223-1233.

Cheng, C. S. (1997). $E(s^2)$-optimal supersaturated designs. *Statist. Sinica* **7**, 929-939.

Deng, L. Y., Lin, D. K. J. and Wang, J. (1994). Supersaturated design using Hadamard matrix. IBM Research Report, RC 19470, IBM Watson Research Center.

Deng, L. Y., Lin, D. K. J. and Wang, J. (1996a). A measurement of multi-factor orthogonality. *Statist. Probab. Lett.* **28**, 203-209.

Deng, L. Y., Lin, D. K. J. and Wang, J. (1996b). Marginally oversaturated designs. *Commun. Statist.-Theory Meth.* **25**, 2557-2573.

Hall, M. J. (1961). Hadamard matrix of order 16. *Jet Propulsion Lab Summary* **1**, 21-26.

Lin, D. K. J. (1991). Systematic supersaturated designs. Working Paper No.264, Department of Statistics, University of Tennessee.

Lin, D. K. J. (1993a). A new class of supersaturated designs. *Technometrics* **35**, 28-31.

Lin, D. K. J. (1993b). Another look at first-order saturated designs: The $p$-efficient designs. *Technometrics* **35**, 284-292.

Lin, D. K. J. (1995). Generating systematic supersaturated designs. *Technometrics* **37**, 213-225.

Srivastava, J. N. (1975). Design for searching non-negligible effects. In *A Survey of Statistical Design and Linear Models* (Edited by J. N. Srivastava), 507-519. North-Holland Amsterdam.

Tang, B. and Wu, C. F. J. (1997). A method for constructing Supersaturated Designs and its $E(s^2)$ optimality. *Canad. J. Statist.* **25**, 191-201.

Wu, C. F. J. (1993). Construction of supersaturated designs through partially aliased interactions. *Biometrika* **80**, 661-669.

Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152, U.S.A.
E-mail: DENGL@MSCI.MEMPHIS.EDU

Department of Management Science and Information Systems, The Pennsylvania State University, University Park, PA 16802-1913, U.S.A.
E-mail: DKL5@PSU.EDU

Department of Statistics and Clinical Programming, Covance Clinical and Periapproval Services, Inc., Princeton, NJ 08540, U.S.A.
E-mail: jim.wang@Covance.Com

(Received December 1996; accepted September 1998)