On The Constraints of Some Combinatorial Arrays

D.V. Chopra
Wichita State University
Wichita, KS 67260-0033, U.S.A.

M. Bsharat
Quintiles, Inc
Kansas City, Missouri 64134-0708, U.S.A.

Abstract

In this paper we obtain a set of inequalities which are necessary conditions for the existence of balanced arrays of strength five, having m rows (constraints), and with two symbols. We discuss the use of these inequalities to obtain an upper bound on m, and present some illustrative examples.

1 Introduction and Preliminaries

For ease of reference we provide here some basic concepts and definitions concerning balanced arrays.

Definition 1.1 *A balanced array (B-array) T with m rows (constraints), with two symbols (say, 0 and 1) and of strength t ($\leq m$) is merely a matrix T of size $(m \times N)$, with elements 0 and 1, such that in every t rowed submatrix T^* of T every vector $\underline{\alpha}$ $(t \times 1)$ of weight i (the weight of $\underline{\alpha}$, denoted by $w(\underline{\alpha})$, means the number of ones in it, and clearly $0 \leq w(\underline{\alpha}) \leq t$) appears with the same frequency μ_i (say).*

Remark 1 *The vector $\underline{\mu}' = (\mu_0, \mu_1, \mu_2, \ldots, \mu_t)$ is called the index set of the array T, and clearly*

$$N = \sum_{i=1}^{t} \binom{t}{i} \mu_i$$

Note: The above definition can be easily extended to B-arrays with s symbols.

It is quite clear that B-arrays are generalization of orthogonal arrays (O-arrays) for which $\mu_i = \mu$ for each i (that is μ_i is independent of i). Also the incidence matrix of a balanced incomplete block design ($BIBD$) Corresponds to balanced arrays of strength two with the restriction that the weight of each column vector is the same. Balanced arrays, under certain conditions, have been extensively used to construct fractional factorial designs of varying resolutions. For example, a B-array of strength $t = 5$ would give rise to a factorial design of resolution VI (i.e. a design which allows us to estimate all the effects up to and including two-factor interactions in the presence of three-factor interactions when all other interactions are negligible). In order to obtain further information on the importance of these combinatorial arrays to combinatorics and statistics, the interested reader may consult the references (by no means it is an exhaustive list) at the end of this paper, and also further references cited therein.

From the above discussion it is quite clear that the construction and existence of such arrays become very important. It is not difficult to see that, for a given $\underline{\mu}'$, the construction of a B-array for $m(> t)$ is a non-trivial one. Furthermore the problem of obtaining the maximum possible value of m, for a given $\underline{\mu}'$, is very important in combinatorics and statistical design of experiments. These problems for B-arrays and O-arrays have been investgated, among others, by Bose and Bush [1], Chopra and/or Dios' [3,4,5,6], Rafter and Seiden[8], Rao [9,10], Seiden and Zemach [11], etc.

2 Necessary Conditions for the existence of Balanced Arrays

In the rest of the paper we restrict ourselves to B-arrays with $t = 5$. However some of the results stated here can be easily extended to a general t with some modifications.

Lemma 2.1 *The non-existence of a B-array for any $m = k$ $(k \geq 6)$ implies its non-existence for any $m > k$.*

Lemma 2.2 *A B-array with $m = t(= 5)$ always exists for any given $\underline{\mu}'$.*

Lemma 2.3 *A B-array Tof strength $t(= 5)$ is also of strength t' where 0 $\leq t' \leq t(= 5)$.*

Remark 2 *Given the index set of T is $\underline{\mu}'$, it can be easily checked that the elements of the index set of T, considered as an array of strength t', are merely a linear combination of the elements of $\underline{\mu}'$. In this case we get the following result:*

$$A(j,t') = \sum_{i=0}^{5-t'} \binom{5-t'}{i} \mu_{i+j} \tag{2.1}$$

where $A(j,t')$ is the jth element of T with its strength t'. Also $A(j,5) = \mu_j$, and $A(j,0) = N$. The next result can be easily established by counting in two ways, through columns and rows, weights of the vectors.

Lemma 2.4 *Let $x_j (j = 0, 1, .., m)$ be the number of columns of weight j in a B-array T with m rows and with $t = 5$. Then the following results hold good:*

$$B_0 = \sum_{j=0}^{m} x_j = N \tag{2.2}$$

$$B_k = \sum j^k x_j = \sum_{r=1}^{k} a_r m_r A(r,r) \qquad 1 \leq k \leq 5$$

Where m_r stands for $m(m-1)...(m-r+1)$, and a_r are positive integers which are known.

Remark 3 *The equations (2.2) clearly connect the moments of the weights of the columns of T to the polynomial functions of m, and μ_i's. Thus for a given $\underline{\mu}'$, we get merely polynomial functions in m. For computational ease, we next give various values of a_r for values of $k (1 \leq k \leq 5)$:* $[1, (1,1), (1,3,1), (1,7,6,1),$ *and* $(1,15,25,10,1)]$.

Theorem 2.5 *For a B-array T with m rows and strength $t = 5$ to exist, the following result must hold:*

$$N^4 B_5 - 2N^2 B_3 B_1^2 + B_1^5 \geq 0 \tag{2.3}$$

Proof: Consider

$$f(m) = \sum_{j=0}^{m} j(j^2 - b^2)^2 x_j \qquad where$$

$$b = \frac{\sum j x_j}{N} = \frac{B_1}{N}.$$

$$Clearly f(m) \geq 0$$

If we expand the L.H.S, and use (2.2) we will obtain the result.

Theorem 2.6 *Consider a B-array T with m rows and having strength t = 5. The following is true:*

$$N^2B_5 - 2NB_2B_3 + B_1B_2^2 \geq 0 \tag{2.4}$$

Proof: To derive (2.4), we consider the following:

$$\sum j(j^2 - a)^2 x_j \geq 0 \text{ where } a = \frac{\sum j^2 x_j}{N} = \frac{B_2}{N}$$

we obtain the result after expanding the LHS.

Theorem 2.7 *For an m-rowed B-array T with t = 5 to exist, the following must be true:*

$$N^2B_5 - 2NB_1B_4 + B_1^2B_3 \geq 0 \tag{2.5}$$

Proof: To obtain (2.5), we consider the inequality

$$\sum j^3(j - b)^2 \geq 0 \quad \text{where} \quad b = \frac{\sum jx_j}{n} = \frac{B_1}{N}$$

Expanding the LHS, and with appropriate substitution we obtain (2.5).

Theorem 2.8 *Consider a B-array T with m-rows and $\underline{\mu}' = (\mu_0, \mu_1, .., \mu_5)$. For T to exist, the following must hold:*

$$(B_5 + 3B_3 + 3B_4 + B_2)^{\frac{1}{3}} \leq (B_5)^{\frac{1}{3}} + (B_2)^{\frac{1}{3}} \tag{2.6}$$

Proof: In order to establish (2.6), we use the following result (known as Minkowski's Inequality):

$$\left[\sum_{i-1} [x_i + y_i]^p\right]^{\frac{1}{p}} \leq \left[\sum_{i=1} [x_i]^p\right]^{\frac{1}{p}} + \left[\sum_{i=1} [y_i]^p\right]^{\frac{1}{p}}$$

To derive (2.6), we set p = 3 in the Minkowski's inequality, and we take

$$x_i = j^{\frac{5}{3}} x_j^{\frac{1}{3}}, y_i = j^{\frac{2}{3}} x_j^{\frac{1}{3}}$$

, after some simplification we obtain (2.6).

Theorem 2.9 *For a B-array T with t=5 and m rows to exist, it is necessary that the following is satisfied:*

$$(B_5 + 6B_4 + 12B_3 + 8B_2)^{\frac{1}{3}} \leq 2(B_2)^{\frac{1}{3}} + (B_5)^{\frac{1}{3}} \tag{2.7}$$

Proof: Here we need extended form of Minkowski's inequality:

$$\left[\sum_{i-1}(x_i+y_i+z_i)^p\right]^{\frac{1}{p}} \le \left[\sum(x_i)^p\right]^{\frac{1}{p}} + \left[\sum(y_i)^p\right]^{\frac{1}{p}} + \left[\sum(z_i)^p\right]^{\frac{1}{p}}$$

Here also, we get (2.7) after some simplification, setting $x_i = j^{\frac{2}{3}}x_j^{\frac{1}{3}}$, $y_i = j^{\frac{2}{3}}x_j^{\frac{1}{3}}$, *and* $z_i = j^{\frac{5}{3}}x_j^{\frac{1}{3}}$ *in L.H.S. and R.H.S.*

Theorem 2.10 *Consider a B-array T of strength t=5 and with m rows. For T to exist we must have the following inequality:*

$$\sqrt[3]{B_5} \le \sqrt[3]{B_2} + \sqrt[3]{B_5 - 3B_4 + 3B_3 - B_2} \qquad (2.8)$$

Proof: In the Minkowski's inequality we make the following substitutions:

$$x_i = (j^{\frac{5}{3}} - j^{\frac{2}{3}})x_j^{\frac{1}{3}}, y_i = j^{\frac{2}{3}}x_j^{\frac{1}{3}}$$

Therefore $x_i + y_i = j^{\frac{5}{3}}x_j^{\frac{1}{3}}$ *and obtain the result after some simplification.*

Next we prepared a computer program to check if an array T exists for a given m and $\underline{\mu}'$. If anyone inequality is contradicted, then T does not exist. If we want max(m) for a given $\underline{\mu}'$, we start checking all inequalities beginning with m=6. If at least one inequality is contradicted for m=k+1 (say), then max (m) = k. However we would like to point out that if all inequalities are satisfied for a m $\ge$ 6, it does not mean that the array exists for that value of m.

Example: Consider the arrays (1, 4, 1, 1, 1, 1), (1, 1, 1, 7, 4, 3) and (2, 5, 7, 1, 1, 1). The max (m) for these arrays respectively are found to be 6, 8, and 7 by using (2.8) for the first index set, (2.3) and (2.4) for the second one, and (2.7), (2.3), (2.4) for the last one.

3 References

[1] R.C. Bose and K.A. Bush. Orthogonal arrays of strength two and three, Ann. Math. Statist. 23 (1952), 508-524.

[2] S.C. Cheng. Optimality of some weighing desings and 2^n fractional factorial designs, Ann. Statist. 8 (1980), 436-446.

[3] D.V. Chopra. On balanced arrays with two symbols, Ars Combinatioria 20A (1985), 59-64.

[4] D.V. Chopra. Balanced arrays and their applications to factorial dsigns, proceedings of the International Statistical Institute LII (1987), 71-72.

[5] D.V. Chopra and R. Dios'. On two-level balanced arrays of strength six, J. Combin. Math. Combin. Comput. 24 (1997), 249-253.

[6] R. Dios and D.V. Chopra. On the non-existence of some orthogonal arrays, J. Combin. Math. Combin. Comput. 54 (2005), 129-136.

[7] J.Q. Longyear. Arrays of strength t on two symbols, J.Statist. Plann. Inf. 10 (1984), 227-239.

[8] J.A. Rafter and E. Seiden. Comtributions to thetheory and construction of balanced arrays, Ann. Statist. 2 (1974), 1256-1273.

[9] C.R. Rao. Hyper cubes of strength d leading to confounding designs in factorial experiments, Bull. Calculta Math. Soc. 38 (1946), 67-78.

[10] C.R. Rao Factorial experiments derivable from combinatiorial arrangements of arrays, J.Roy. Statist. Soc. Suppl. 9 (1947), 128-139.

[11] E. Seiden and R. Zemach. On orthogonal arrays, Ann. Math. Statist. 27 (1966), 1355-1370.

[12] W.D. Wallis. Combinatorial designs, Marcel Dekker Inc., New York, 1988.