Note on Agrawal's "Designs for Two-way Elimination of Heterogeneity"

W.D. Wallis and J. L. Yucas
Department of Mathematics, Southern Illinois University,
Carbondale IL, USA 62901-4408
wdwallis@math.siu.edu,jyucas@math.siu.edu

Abstract

Agrawal provided a construction for designs for two-way elimination of heterogeneity, based on a symmetric balanced incomplete block design. He could not prove the construction, although he found no counterexample. Subsequently Raghavarao and Nageswarerao published a proof of the method.

In this note we observe a flaw in the published proof.

1 Introduction

Agrawal [1] studied a class of designs for two-way elimination of heterogeneity. Such a design has three classes of constraints, called *rows*, *columns* and *letters*. Suppose there are r rows, c columns and v letters. Then the design has rc blocks of size 3, satisfying the following constraints:

- every block is a transversal of rows, columns and letters;
- every row and letter occur together in at most one block;
- every column and letter occur together in at most one block;
- every row and column occur together in exactly one block;
- there exist positive integers k, λ_{rr}, λ_{cc} and λ_{rc} such that:

TA1. each letter occurs in k blocks;

TA2. any two distinct rows contain λ_{rr} common letters;

TA3. any two distinct columns contain λ_{cc} common letters;

TA4. any row and column contain λ_{rc} common letters.

An example is

$$\begin{array}{|cccccc|} 4 & 5 & 6 & 10 & 1 & 2 \\ 1 & 8 & 3 & 4 & 7 & 6 \\ 9 & 2 & 4 & 7 & 5 & 8 \\ 8 & 3 & 9 & 5 & 6 & 10 \\ 10 & 1 & 2 & 3 & 9 & 7 \end{array}. \quad (1)$$

(This example is taken from [1, p1157].)

Such a design has a natural representation as a (binary) *row-and-column design*; if a, b and c are respectively a row, column and letter, the block abc is represented by symbol c in cell (a, b). The properties mean that the row-and-column design is *equireplicate*, with every letter appearing k times, it contains no empty cells, and if the rows and columns are treated as sets then the intersection of any two rows has size λ_{rr}, the intersection of any two columns has size λ_{cc}, and any row and column intersect in λ_{rc} elements. We shall call this array a *triple array* and denote it $TA(v, k, \lambda_{rr}, \lambda_{cc}, \lambda_{rc} : r \times c)$. The example above is a $TA(10, 3, 3, 2, 3 : 5 \times 6)$.

The design can also be represented in two *extensive* forms: as an $r \times v$ array, where block abc is represented by entry b in cell (a, c), and a $c \times v$ array with entry a in cell (b, c). We shall call these arrays the RL and CL forms of the triple array, respectively.

For any array in RL form to represent an $r \times c$ array with no repetitions in the rows or columns, it is necessary that the entries in (the occupied cells of) every column are distinct, and that the entries in (the c occupied cells of) each row should include each of $\{1, 2, \ldots, c\}$ precisely once. For the RL form of a triple array, the conditions TA1 through TA4 have the following effect:

TA1: every column has k occupied cells;

TA2: for any pair of rows, there are λ_{rr} columns in which both are occupied;

TA3: every pair of symbols must occur together in λ_{cc} columns;

TA4: columns with occupied cells in row i must contain every symbol λ_{rc} times, for all $i, 1 \leq i \leq r$.

Consider the matrix formed from this array by replacing each occupied cell by 1 and each empty cell by 0. Then TA1 and TA2 together imply that this matrix is the incidence matrix of a $(r, v, c, k, \lambda_{rr})$-BIBD. So, if the rows of a triple array are considered as treatments, and for each symbol we define a block consisting of the rows that containing that symbol, these blocks form a balanced incomplete block design with parameters $(r, v, c, k, \lambda_{rr})$, which we shall call the *row design*.

A similar analysis can be applied to the CL form, and the triple array forms a $(c, v, r, k, \lambda_{cc})$-BIBD, the *column design.* The usual parameter relations follow, and it is easy to prove that $\lambda_{rc} = k$.

In [2] we prove

Theorem 1.1 *Any triple array with $k \neq r$ and $k \neq c$ satisfies*

$$v \geq r + c - 1.$$

The extremal case $v = r + c - 1$ (in which it is easy to show that $\lambda_{cc} = r - k$) is of special interest. Agrawal [1] gave a method that started from a symmetric $(v+1, r, \lambda_{cc})$-BIBD, where $v = r + c - 1$, and constructed a $TA(v, k, \lambda_{rr}, \lambda_{cc}, \lambda_{rc} : r \times c)$. He could not prove his method, but had found it to work in every case that he tried, provided $r - \lambda_{cc} > 2$. (Not only does the method fail when $r - \lambda_{cc} = 2$, but no triple array exists in those cases.) Subsequently Raghavarao and Nageswarerao [3] claimed to prove that the method always works. However, their proof is faulty.

2 Agrawal's method

Agrawal uses two balanced incomplete block designs obtained from the symmetric $(v+1, r, \lambda_{cc})$-BIBD: the residual design, with parameters $(v+1-r, v, r, r - \lambda_{cc}, \lambda_{cc}) = (c, v, r, k, \lambda_{cc})$, obtained by deleting one block and all its members from the symmetric design, and the derived design, an $(r, v, r-1, \lambda_{cc}, \lambda_{cc} - 1)$-BIBD, obtained by deleting one block and all members of its complement.[1] Suppose the original design is based on treatment-set $T = \{1, 2, \ldots, \}$ and has blocks $B_0, B_1, \ldots, B_v$; the blocks of the residual design (with regard to B_0) are the k-sets $B_1^*, B_2^*, \ldots, B_v^*$, where $B_i^* = B_i \backslash B_0$. Let A denote the incidence matrix of the complement of the derived design with regard to the *same* original block B_0. This is a BIBD with parameters $(r, v, v-r+1, r-\lambda_{cc}, v-2r+\lambda_{cc}+1) = (r, v, c, k, \lambda_{rr})$ whose treatment-set is B_0 and whose blocks are the sets $B_0 \backslash B_j$, so A is a $(0, 1)$ matrix of size $r \times v - 1$ with k 1's per column.

Now, for each j, the entries 1 in column j of A are replaced by the elements of B_j^*. Then the elements in each column are reordered in such a way that every row contains each member of B_0. The resulting array is the RL form of a triple array $TA(v, k, \lambda_{rr}, \lambda_{cc}, \lambda_{rc} : r \times c)$.

Of course, the difficulty lies in proving that such a reordering is always possible.

[1] For some reason, both Agrawal and Raghavarao-Nageswarerao interchange the meanings of "residual" and "derived".

3 Raghavarao and Nageswarerao's analysis

Raghavarao and Nageswarerao state that Agrawal's method is equivalent to the following. Denote by $\mathcal{D}, \mathcal{D}_1$ and $\mathcal{D}_2$ the symmetric design, its residual design with regard to B_0 and the complement of its derived design with regard to B_0. Corresponding to element t_1 of B_0, select those blocks $B_{1_1}, B_{2_1}, \ldots, B_{c_1}$ of the symmetric design that do *not* contain t_1. Then find a system of distinct representatives for the corresponding blocks of D_2 (that is, the sets $B_0 \backslash B_i$ where $1 \notin B_i$). Then these elements are to be "brought to the first column" ([3, p. 198]). Clearly this should be "first row", as elements remain in their original column. However, there is a more serious problem: the elements being selected are chosen from B_0, and the elements that Agrawal requires are to be members of $T \backslash B_0$. To make sense of the construction, we assume that Raghavarao and Nageswarerao mean to ask for a system of representatives of the sets $B_i \backslash \mathcal{B}_0$. As further evidence for this assumption, we note that there are r treatments in $\mathcal{D}_2$ and c sets have been chosen, so that the scheme could only be carried out when $c < r$, which is not necessarily true.

So we assume that the plan is as follows: for each $T_j \in B_0$, consider the blocks $B_{1_j}, B_{2_j}, \ldots, B_{c_j}$ of D that do not contain t_j. Find a system of distinct representatives $b_{1_j}, b_{2_j}, \ldots, b_{c_j}$ for the corresponding sets $B_{1_j} \backslash \mathcal{B}_0, B_{2_j} \backslash \mathcal{B}_0, \ldots, B_{c_j} \backslash \mathcal{B}_0$. Then b_{i_j} is to be the (i_j, j) entry of the RL matrix.

The authors state that "an SDR exists" for the sets $B_{i,j}$. While this is true, it is not sufficient. If the algorithm is to work, then for every choice of i the elements b_{i_h} and b_{i_j} must always be different when $h \neq j$. This puts a restriction on the SDRs, and it is not obvious that it can be satisfied.

4 An example

The proof is certainly not valid as written. If it were then the construction would work when $r - \lambda_{cc} = 2$, and Agrawal noted that his method does not work in that case. But the problems go further than that. As an example, consider the $(11, 5, 2)$-design with blocks (written as columns)

1	3	2	1	4	2	3	1	1	1	2
2	4	4	3	5	5	5	4	5	2	3
3	6	7	7	6	7	8	9	6	6	6
4	7	8	8	8	9	10	10	7	8	9
5	10	11	9	9	10	11	11	11	10	11

(this is the $(11, 5, 2)$-design that was used in [1] to produce the triple array (1)).

The residual design with regard to the first block consists of the south-east corner of the array. For element 1, we require an SDR of the sets

$$\{6,7,10\}\ \{7,8,11\}\ \{6,8,9\}\ \{7,9,10\}\ \{8,10,11\}\ \{6,9,11\}$$

and one possibility is $6,7,8,9,10,11$. For element 2, the sets are

$$\{6,7,10\}\ \{7,8,9\}\ \{6,8,9\}\ \{8,10,11\}\ \{9,10,11\}\ \{6,7,11\}$$

but there are three restrictions: one cannot choose 6 from $6,7,10$, 8 from $6,8,9$ or 10 from $8,10,11$, because these choices lead to a repetition. So we really need an SDR of the sets

$$\{7,10\}\ \{7,8,9\}\ \{6,9\}\ \{8,11\}\ \{9,10,11\}\ \{6,7,11.\}$$

One choice is $7,8,9,11,10,6$. For 3, the sets are

$$\{7,8,11\}\ \{6,8,9\}\ \{7,9,10\}\ \{9,10,11\}\ \{6,7,11\}\ \{6,8,10\}$$

but after deleting the elements already used, the choice must be made from

$$\{8,11\}\ \{6\}\ \{7,10\}\ \{9,11\}\ \{7,11\}\ \{6,8,10\}$$

and we might select $8,6,7,9,11,10$. For 4, the sets are

$$\{7,8,9\}\ \{7,9,10\}\ \{8,10,11\}\ \{6,7,11\}\ \{6,8,10\}\ \{6,9,11\}$$

and after deleting the elements already used we have

$$\{7,9\}\ \{10\}\ \{8\}\ \{7\}\ \{6,8\}\ \{6,9\}$$

which have no SDR. In fact, if the first two SDRs have been chosen as above, there is no possible completion.

All three designs involved in this example — the symmetric $(11,5,2)$-design, the residual $(6,10,5,3,2)$-BIBD and the derived $(5,10,6,3,3)$-design — are uniquely determined up to isomorphism, so there is no chance that the failure is due to a poor choice of design.

5 Youden squares

The authors also state that the construction is equivalent to taking a Youden square, and deleting the first column and all its elements. (In our terminology, they assert that the resulting array is the *RL* form of a triple array.) Again, this does not work for every Youden square. If one starts with the Youden square

1	3	2	7	4	5	10	11	6	8	9
2	4	7	1	5	9	8	10	11	6	3
3	6	4	8	9	2	5	1	7	10	11
4	7	8	3	6	10	11	9	5	1	2
5	10	11	9	8	7	3	4	1	2	6

the result of deleting the first column and its elements is

–	–	7	–	–	10	11	6	8	9
–	7	–	–	9	8	10	11	6	–
6	–	8	9	–	–	–	7	10	11
7	8	–	6	10	11	9	–	–	–
10	11	9	8	7	–	–	–	–	6

If the result of this process is treated as the *RL* form of a triple array, the *RC* form is

8	3	9	10	6	7
9	2	6	5	7	8
1	8	3	4	9	10
4	1	2	7	5	6
10	5	4	3	1	2

which has neither constant row intersection size nor the property that any row and column share a constant number of symbols.

Acknowledgement

The authors would like the referee for several helpful suggestions that have substantially improved the presentation of this paper.

References

[1] H. Agrawal, Some methods of construction of designs for two-way elimination of heterogeneity. *J. Amer. Statist. Assoc.* **61**(1966), 1153–1171.

[2] J. P. McSorley, N. C. Phillips, W. D. Wallis and J. L. Yucas, Double arrays, triple arrays, and balanced grids. (submitted)

[3] D. Raghavarao and G. Nageswarerao, A note on a method of construction of designs for two-way elimination of heterogeneity. *Commun. Statist.* **3**(1974), 197–199.