On a Class of Arrangements

By C. Radhakrishna Rao

In this paper, I introduce a class of arrangements called *arrays of strength d* and discuss methods of constructing them with the help of finite geometrical configurations and algebraic groups involving elements of a Galois field. The definitions of arrays of strength d and other configurations that are used are given below.

Arrays of strength d. Let there be n factors A_1, $A_2 \ldots$, A_n each of which can assume s values, those corresponding to A_i being represented by $i_1, i_2, \ldots, i_s$. An ordered set $(1_a 2_b \ldots n_k)$, denoted simply by $(a\ b \ldots k)$, may be called a combination or an *assembly*. There are altogether s^n assemblies of which a subset of N assemblies is called an *array* and represented by (N, n, s). This array is said to be *of strength d* if all s^d assemblies corresponding to any d factors chosen out of n occur an equal number of times. The array of strength d is represented by (N, n, s, d). This array is called a *hypercube of strength d* when N is of the form s^m. These hypercubes when s is a prime or a prime power have been used by the author (Rao: 1946)[1] in arriving at certain arrangements in the theory of experimentation in statistics.

Web geometry. A system of points and lines is defined as a web and denoted by $W(n, s)$ if the following incidence relations hold:

  (i) two lines can be incident with one and only one point,

  (ii) not more than one line can pass through two points,

  (iii) all lines form n classes or pencils each containing s lines with s points on each such that no two lines of the same pencil intersect but two lines of two different pencils intersect.

Orthogonal Latin squares. An arrangement of s^2 elements consisting of s values each taken s times, in the cells of a square formed by s rows and columns such that each value is repeated once and

---

[1] See the list of references at the end of the paper.

only once in every row and column, is called a *Latin square*. Two squares are said to be *orthogonal* if all s^2 combinations of s values occur once when they are superimposed.

We now prove the following theorems.

THEOREM 1. *The following relationships hold good among the parameters N, n, s and d of the array (N, n, s, d).*

(a) $N = \lambda s^d$ where λ is an integer.

(b) $N \geqq {}^nC_0 + {}^nC_1 (s - 1) + \ldots + {}^nC_{d/2} (s - 1)^{d/2}$
$$\text{when } d \text{ is even.}$$

(c) $N \geqq {}^nC_0 + {}^nC_1 (s - 1) + \ldots + {}^nC_{(d - 1)/2} (s - 1)^{(d - 1)/2}$
$$+ {}^{n-1}C_{(d-1)/2} (s - 1)^{(d + 1)/2}$$
$$\text{when } d \text{ is odd.}$$

Proof. The result (a) follows from definition. To prove (b) we define the symbolic sum involving the levels $i_1, i_2, \ldots, i_s$ of the i-th factor

$$i^a = \sum_j m_{ja}\, i_j$$

such that $\sum_j m_{ja} = 0, \qquad a = 1, 2, \ldots, s - 1$

$$\sum_j m_{ja}\, m_{jb} = 0, \qquad a \neq b,$$

$$m_{ja} = 1 \text{ for all } j \text{ when } a = s.$$

The symbolic product $1^a\, 2^b \ldots n^k$ when expanded gives a linear function of the assemblies which may be represented by $[a\, b \ldots k]$ to differentiate it from the assembly $(a\, b \ldots k)$. If in the linear function $[a\, b \ldots k]$ we retain only those assemblies present in the array (N, n, s, d) we get another function which may be represented by $\{a\, b \ldots k\}$.

Consider the case $d = 4$. Since all combinations of every four factors are equally repeated it follows that all combinations of every f factors, where $f < 4$, are equally repeated. Hence we see that the functions

$\{a\, b\, s \ldots s\}$ and $\{s\, s\, c\, d\, s \ldots s\}$, except when $a = b = c = d = s$,

$\{a\, b\, s \ldots s\}$ and $\{s\, b'\, c'\, s \ldots s\}$, except when $a = b = a' = b' = c' = s$,

considered as linear functions of assemblies, are mutually orthogonal and hence independent. Since the total number of assemblies in the array is N, the number of independent linear functions of the

assemblies is not greater than N. The numbers of functions of the type $\{a\ b \ldots k\}$ with all s's, $(n-1)$ s's and $(n-2)$ s's are 1, $^nC_1\ (s-1)$ and $^nC_2\ (s-1)^2$ respectively, and these are all independent, being orthogonal. Hence

$$N \geqq 1 + {}^nC_1\ (s-1) + {}^nC_2\ (s-1)^2.$$

The extension to general d is obvious, and the result (b) is proved.

To prove (c) let us consider the case $d = 5$. Besides the orthogonal functions enumerated above for $d = 4$, we have $^{n-1}C_2\ (s-1)^3$ more functions obtained by fixing any value in the first position and keeping any two other values different from s. Hence we get, when $d = 5$,

$$N \geqq 1 + {}^nC_1\ (s-1) + {}^nC_2\ (s-1)^2 + {}^{n-1}C_2\ (s-1)^3.$$

This can be extended to general d, and the result (c) is proved.

THEOREM 2. *The configuration of the web $W\ (n,\ s)$, $(n-2)$ mutually orthogonal squares of side s, and the array $(s^2,\ n,\ s,\ 2)$ of strength 2 exist together.*

Proof. Given the configuration of $W\ (n, s)$ we have to identify the n pencils with n factors, s values of a factor with the s lines of a pencil. There are s^2 points through each of which pass n lines belonging to n different pencils. Hence a point can be identified by the nature of lines passing through it or an assembly. The s^2 assemblies arising out of the s^2 points, supply the array $(s^2,\ n,\ s,\ 2)$ in virtue of the properties (i), (ii) and (iii) of a Web.

If we identify the s values of a factor with the rows and s values of another with the columns of a square and fill in the cells represented by combinations of two factors with the corresponding values of the third factor, we get a Latin square arrangement. The values corresponding to the fourth factor give an orthogonal square and those of the fifth factor give a square orthogonal to the first two. The $(n-2)$ orthogonal squares are thus derivable from $(s^2,\ n,\ s,\ 2)$. It is easy to see that by retracing the steps the array can be derived from $(n-2)$ mutually orthogonal squares. This shows that any one of $W(n,\ s)$, $(s^2,\ n,\ s,\ 2)$ and $(n-2)$ mutually orthogonal squares of side s, can be derived from the others, thus establishing Theorem 2.

COROLLARY (2.1). *The maximum value of n in $(s^2, n, s, 2)$ is $(s+1)$ when s is a prime or prime power.*

When s is a prime or a prime power it is known that the web $W\ (n, s)$ for $n = s + 1$ exists, for it is equivalent to a finite Euclidean

geometry with $(s+1)$ points on a line. That this is the maximum possible value of N is given by the inequality (b) of Theorem 1 when $d = 2$.

COROLLARY (2.2). *The maximum value of n in $(6^2, n, 6, 2)$ is* 3.

Euler conjectured that when s is of the form $4t + 2$ no two orthogonal squares exist. If this is correct then the maximum value of n in $((4t + 2)^2, n, 4t + 2, 2)$ is 3. In particular it has been verified by Fisher and Yates (1934) that when $s = 6$ Euler's conjecture is correct; hence the result.

Let s be a prime or a prime power, in which case there exists a Galois field $GF(s)$ with s elements. The s values of a factor may be represented by the s elements of $GF(s)$. In the following theorems only these elements will be used.

If we represent the factors by $\lambda_1 F_1 + \ldots + \lambda_r F_r$, where $\lambda_1, \ldots, \lambda_r$ are elements in $GF(s)$ with the convention that not all λ's are zero and the first non-zero coefficient is unity, and determine an assembly by giving the value $\lambda_1 a_1 + \ldots + \lambda_r a_r$, corresponding to the factor $\lambda_1 F_1 + \ldots + \lambda_r F_r$, where $(a_1 \ldots a_r)$ is a combination of r elements in $GF(s)$, then we get an array (s^r, n, s) with s^r assemblies by allowing each a in $(a_1 \ldots a_r)$ to assume all possible values in $GF(s)$, n being the number of factors used.

THEOREM 3. *The array with s^r assemblies constructed above is of strength 2, and this is possible with the maximum number of factors $n = (s^r - 1)/(s - 1)$.*

To prove that the array (s^r, n, s) is of strength 2 we must show that in the s^r combinations

$$(\lambda_1 a_1 + \ldots + \lambda_r a_r), (\mu_1 a_1 + \ldots + \mu_r a_r),$$

where $a_1 \ldots a_r$ assume all possible values in $GF(s)$, every ordered pair of values in $GF(s)$ occurs s^{r-2} times if $\lambda_i \neq \sigma \mu_i$ for all i, where σ is an element in $GF(s)$. If β_1, β_2 is an ordered pair of values in $GF(s)$ then we require the number of solutions of the equations

$$\lambda_1 a_1 + \ldots + \lambda_r a_r = \beta_1$$
$$\mu_1 a_1 + \ldots + \mu_r a_r = \beta_2,$$

which is evidently s^{r-2} if the above condition regarding λ's and μ's is satisfied.

Since a factor is represented by $\lambda_1 F_1 + \ldots + \lambda_r F_r$ with the provision that the first non-zero λ is the unit element, we find that the number of different factors is $(s^r - 1)/(s - 1)$.

If in the inequality (b) of Theorem 1 corresponding to $d = 2$,

$$N \geqq 1 + {}^nC_1 (s - 1),$$

we substitute $N = s^r$, we find $n \leqq (s^r - 1) / (s - 1)$, so that the number of factors found above is the maximum possible. This completes the proof of the above theorem and gives the method of constructing arrays (s^r, n, s, d) of strength 2 for the maximum possible number of factors.

As an example we may construct the array $(2^3, 7, 2, 2)$. The elements of the field $GF(2)$ are 0 and 1 with the law of addition $1 + 1 = 0, 1 + 0 = 1$.

The array $(8, 7, 2, 2)$.

Factors

	F_1	F_2	F_3	$F_1 + F_2$	$F_1 + F_3$	$F_2 + F_3$	$F_1 + F_2 + F_3$
	0	0	0	0	0	0	0
	0	0	1	0	1	1	1
	0	1	0	1	0	1	1
Assemblies	0	1	1	1	1	0	0
	1	0	0	1	1	0	1
	1	0	1	1	0	1	0
	1	1	0	0	1	1	0
	1	1	1	0	0	0	1

If we want the array with s^r assemblies to be of strength 3, we have to choose the factors so that the number of solutions corresponding to any three factors

$$\lambda_1 a_1 + \ldots + \lambda_r a_r = \beta_1$$
$$\mu_1 a_1 + \ldots + \mu_r a_r = \beta_2$$
$$\nu_1 a_1 + \ldots + \nu_r a_r = \beta_3$$

is s^{r-3}. This means that the vector $(\nu_1 \ldots \nu_r)$ does not depend on the vectors $(\lambda_1 \ldots \lambda_r)$ and $(\mu_1 \ldots \mu_r)$. The maximum number of factors for which an array of strength 3 can be constructed by the above method is equal to the number of vectors in a space of r dimensions such that no vector is a linear combination of any two vectors. This can be generalised to the construction of arrays of strength d. Hence we get the following theorem.

THEOREM 4. *A set containing a maximum number of vectors in an r-space such that no vector in the set is a linear combination of $(d-1)$ vectors in the set, gives rise to the arrangement of an array with s^r*

assemblies involving the maximum number of factors if the above method of construction is followed.

The method of building up this set is to choose the elementary vectors corresponding to the factors $F_1, F_2 \ldots, F_r$ and go on adding vectors which satisfy the above condition. An application of this leads to the following theorem in the case $s = 2$.

THEOREM 5.   *The array* $(2^r, n, 2, d)$ *can be constructed for the maximum number of factors* $n = 2^{r-1}$ *when* $d = 3$.

It is easy to verify that if we choose factors involving an odd number of F's, then the condition of Theorem 4 is satisfied when $d=3$. This is due to the fact that all linear combinations of two factors involving an odd number of F's consist of two original factors and another factor involving an even number of F's. Thus in the case of $r = 3$ we find that the choice of $F_1, F_2, F_3, F_1 + F_2 + F_3$ leads to an array of strength 3. The number of factors involving an odd number of F's in the general case is $^rc_1 + {}^rc_3 + \ldots = 2^{r-1}$. To prove that this is the maximum possible, we have to substitute $N = 2^r$ in the relation (c) of Theorem 1 for $d = 3$:

$$2^r \geqq 2n \quad \text{or} \quad n \leqq 2^{r-1}.$$

This shows that the maximum possible n is 2^{r-1}, and the theorem is proved.

We now consider the problem of constructing arrays of maximum strength for a given number of assemblies of the form s^r and a given number of factors n.

THEOREM 6.   *The group of solutions of the* $(n - r)$ *independent equations*

$$\sum_j \lambda_{ij} x_j = 0, \qquad i = 1, 2, \ldots, n - r$$

gives an array of strength d *if no equation in the group of equations derived from the above set contains fewer than* $(d + 1)$ *non-zero coefficients.*

The number of solutions is s^r. Since every solution satisfies all the equations in the group of equations, it follows that if any equation has $c, < (d + 1)$, non-zero coefficients, all the combinations of c factors corresponding to the non-zero coefficients are not found in the group of solutions. Otherwise the condition of strength is satisfied in the array formed by the group of solutions. An important theorem giving a relationship among s, n, r and d is as follows.

THEOREM 7.

$$d \leqq \frac{n \left(s^{n-r} - s^{n-r-1} \right)}{s^{n-r} - 1} - 1.$$

The number of distinct equations in the group of equations is $(s^{n-r} - 1) / (s - 1)$. The number of zero coefficients corresponding to x_j in all the equations is equal to the number of solutions of

$$b_1 \lambda_{1j} + b_2 \lambda_{2j} + \ldots + b_{n-r} \lambda_{n-r,j} = 0.$$

This is $(s^{n-r-1} - 1) / (s - 1)$. The mean number of zero coefficients in all the equations is

$$n \frac{s^{n-r-1} - 1}{s - 1} \div \frac{s^{n-r} - 1}{s - 1} = \frac{n(s^{n-r-1} - 1)}{s^{n-r} - 1}$$

If $d + 1$ is the least number of non-zero coefficients for equations of the above group, then it follows that the maximum number of zeros, $n - d - 1$, must not be less than the mean number of zeros. Hence we get

$$n - d - 1 \geqq n (s^{n-r-1} - 1) / (s^{n-r} - 1)$$

or
$$d \leqq n \frac{s^{n-r} - s^{n-r-1}}{s^{n-r} - 1} - 1.$$

An analogous result has been derived by Bhattacharya (1942) in a certain minimal problem in the design of experiments.

The method of constructing these arrays when the number of assemblies is not a power of the levels of a factor requires investigation. An important advance in this direction is made by Plackett and Burman (1946), who have supplied an almost complete solution to the problem of arrays of strength 2 when the number of levels of a factor is 2.

REFERENCES.

A. Bhattacharya, "A note on Ramamurti's problem of maximal sets," *Sankhya*, 6 (1942), 189.

R. A. Fisher and F. Yates, "The 6 × 6 Latin squares," *Proc. Camb. Phil. Soc.*, 146 (1934), 1.

R. L. Plackett and J. P. Burman, "The design of optimum multifactorial experiments," *Biometrika*, 33 (1946), 305.

C. R. Rao, "On hypercubes of strength d and a system of confounding in factorial experiments," *Bull. Calcutta Math. Soc.*, 38 (1946), 1.

KING'S COLLEGE,
CAMBRIDGE.