A Short Overview of Orthogonal Arrays

John Stufken

Department of Statistics
University of Georgia

Isaac Newton Institute
September 5, 2011

Starting points

Web libraries of Orthogonal Arrays:

Neil Sloane's library:
http://www2.research.att.com/ njas/oadir/index.html

Warren Kuhfeld's library:
http://support.sas.com/techsup/technote/ts723.html

A book on Orthogonal Arrays:

Hedayat, Sloane and Stufken; Orthogonal Arrays: Theory and Applications. Springer, 1999.

Too many references to list here

Outline

Definition

An Orthogonal Array of strength t with N rows, k columns ($k \geq t$) and based on s symbols is an $N \times k$ array with entries $0, 1, \ldots, s-1$, say, so that every $N \times t$ subarray contains each of the s^t possible t-tuples equally often as a row (say λ times)

N must be a multiple of s^t, and $\lambda = N/s^t$ is the index of the array

Notation: OA($N,\ k,\ s,\ t$) or sometimes OA($N,\ s^k, t$)

Introduced by Rao (1946, 1947, 1949) for use in fractional factorial experiments

Terminology

N: Number of rows, level combinations or runs

k: Number of columns, constraints or factors

s: Number of symbols or levels

t: Strength

Use in factorial experiments

The strength t of an OA is related to estimability of parameters under certain models when using the rows of the OA as runs in the experiment

What is 'orthogonal' about OAs? Will return to this shortly

Example: An OA(8, 4, 2, 3)

0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	0
1	1	1	1

Example: An OA(8, 4, 2, 3)

0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	0
1	1	1	1

Simple observations

An OA of strength t is also of strength t', $t' \leq t$

Invariant to permutations of rows, columns and symbols within a column; isomorphic arrays

Existence of an OA$(N, \ k, \ s, \ t)$ implies existence of OA$(N/s, \ k-1, \ s, \ t-1)$

$$\left[\begin{array}{c|c} 0 & \\ \vdots & A_0 \\ 0 & \\ \hline \vdots & \vdots \\ \hline s-1 & \dots \\ \vdots & \vdots \\ s-1 & \dots \end{array}\right]$$

There are many methods to construct larger OAs from smaller ones.

Simple observations

An OA of strength t is also of strength t', $t' \leq t$

Invariant to permutations of rows, columns and symbols within a column; isomorphic arrays

Existence of an OA$(N,\ k,\ s,\ t)$ implies existence of OA$(N/s,\ k-1,\ s,\ t-1)$

$$\left[\begin{array}{c|c} 0 & \\ \vdots & A_0 \\ 0 & \\ \hline \vdots & \vdots \\ \hline s-1 & \dots \\ \vdots & \vdots \\ s-1 & \dots \end{array}\right]$$

There are many methods to construct larger OAs from smaller ones.

Some basic questions

For which values of $t \geq 2$, $k \geq t$, $s \geq 2$ and $N \equiv 0 \pmod{s^t}$ does an OA(N, k, s, t) exist?

Reformulation 1: What is the maximum k for which an OA(N, k, s, t) with $t \geq 2$, $s \geq 2$ and $N \equiv 0 \pmod{s^t}$ exists? Call this $f(N, s, t)$

Reformulation 2: What is the minimum N for which an OA(N, k, s, t) with $t \geq 2$, $s \geq 2$, and given k exists? Call this $F(k, s, t)$

Relationships: $F(k, s, t) = \min\{N : f(N, s, t) \geq k\}$ and $f(N, s, t) \leq \max\{k : F(k, s, t) \leq N\}$

Establishing values for $f(N, s, t)$ and $F(k, s, t)$ is typically through a combination of obtaining a bound and constructing an OA that attains that bound; but many values are still unknown

General bounds: Rao's inequalities and the LP bound; many methods of construction, especially when s is a prime or prime power

Orthogonality and bounds

What is "orthogonal" about an OA(N, k, s, t)?

Let $\phi_i(u)$, $i = 1, \ldots, s-1$, be a set of $s-1$ orthogonal polynomials on $0, \ldots, s-1$, degree$(\phi_i) = i$

If $t = 2$, then replace $u \in \{0, \ldots, s-1\}$ in the OA everywhere by $(\phi_1(u), \ldots, \phi_{s-1}(u))$, and add a constant column

All columns of the resulting matrix are orthogonal: $N \geq 1 + k(s-1)$; main-effects model

If $t = 3$, then $N \geq 1 + k(s-1) + (k-1)(s-1)^2$

If $t = 4$, can add contrasts for two-factor interactions and maintain orthogonality: $N \geq 1 + k(s-1) + \binom{k}{2}(s-1)^2$

General form of Rao's inequalities

$t = 2u$: $N \geq 1 + k(s-1) + \binom{k}{2}(s-1)^2 + \ldots + \binom{k}{u}(s-1)^u$

$t = 2u + 1$:
$N \geq 1 + k(s-1) + \binom{k}{2}(s-1)^2 + \ldots + \binom{k}{u}(s-1)^u + \binom{k-1}{u}(s-1)^{u+1}$

Only explicit bounds that apply for any OA

Lower bound for N or $F(k, s, t)$; but also upper bound for k or $f(N, s, t)$

Bounds can be attained for some parameters, but improvements are possible for others

Constructions and bounds

Example 1. Zero-sum array: Form an $s^t \times t$ array from all t-tuples based on $0, \ldots, s-1$, and add one column so that the entries in each row add up to 0 modulo s. This is an OA(s^t, $t+1$, s, t)

Surprisingly, $f(s^t, s, t) = t + 1$ if $s \leq t$

Example 2. If $s \geq t$ is a prime power, Bush (1952) obtained an OA(s^t, $s+1$, s, t) using polynomials of degree $\leq t$ over a Galois field GF(s); one extra column can be added if $t = 3$ and s is a power of 2

Conjecture: If s is a prime power, then

$$f(s^t, s, t) = \begin{cases} s+1\,, & \text{if } 2 \leq t \leq s \\ t+1, & \text{if } t \geq s \end{cases}$$

except that $f(s^3, s, 3) = s + 2$, if $s = 2^m$, and
$f(s^{s-1}, s, s-1) = s + 2$, if $s = 2^m$

Rao-Hamming arrays

For a prime power s, $n \geq 2$, Rao-Hamming arrays are arrays $\text{OA}(s^n, (s^n - 1)/(s - 1), s, 2)$ (strength 2)

Give equality for Rao's bounds, so that $f(s^n, s, 2) = (s^n - 1)/(s - 1)$ for prime power s

One possible construction: Form an $s^n \times n$ matrix consisting of all n-tuples over GF(s). If $C_1, \ldots, C_n$ denote the columns, form the columns $\alpha_1 C_1 + \ldots + \alpha_n C_n$ where not all α_i's are zero and the first nonzero α_i is 1

Can alternatively also start with an $n \times (s^n - 1)/(s - 1)$ array, and take all linear combinations of the rows

Outline

Coding theory perspective

Example: A code of length $k = 4$ with an alphabet of size $s = 2$ (binary) with $N = 8$ codewords (size) and minimal distance $d = 2$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Notation: A $(k,\ N,\ d)_s$ code; can correct up to $\lfloor (d-1)/2 \rfloor$ errors

For given $k,\ d$ and s, what is the maximum value for N?

A relationship between codes and OAs

In a linear OA or code over GF(s), the rows form a vector space

Regular fractions in statistics are linear OAs or translates of linear OAs

A linear code C must have $N = s^n$ for some n: $(k,\ s^n,\ d)_s$ code

The vectors in the null space of C also form a vector space over GF(s), and form therefore also a linear code, say $C^\perp$

$C^\perp$ is a $(k,\ s^{k-n},\ d^\perp)_s$ code, where $d^\perp$ is called the dual distance

If C is a linear code $(k,\ N,\ d)_s$ over $GF(s)$ with dual distance $d^\perp$, then its codewords form an OA$(N,\ k,\ d^\perp - 1, s)$ (Kempthorne, 1947) (Clare College graduate); converse statement (Bose, 1961)

Thus the dual distance $d^\perp$, and not the minimal distance d, provides information about the performance of C as an OA

A relationship between codes and OAs

In statistics, for a linear OA, or for one of its translates, $d^\perp$ is called the resolution of the OA

Statisticians also study the length of the words in $C^\perp$ through the so called word length pattern: B_i, $1 \le i \le k$, is the number of words in $C^\perp$ with i nonzero elements, i.e. with Hamming distance i to the all-zero row

The minimum aberration criterion selects designs that sequentially minimize $B_1, B_2, \ldots, B_k$ (e.g. Cheng and Tang, 2005)

As we will see next, the B_i's are also of interest in coding theory

Linear codes can lead to interesting OAs

But what connection is there between codes and OAs if C is not linear?

A relationship between codes and OAs

For C a code $(k,\ N,\ d)_s$, and a word $u \in C$, define the weight distribution w.r.t. u as $(A_0(u), A_1(u), \ldots, A_k(u))$, where $A_i(u)$ denotes the number of words in C with a Hamming distance of i to u

Define the weight distribution of C as $(A_0, A_1, \ldots, A_k)$ where $A_i = \frac{1}{N}\sum_{u \in C} A_i(u)$

For a linear code, $A_i(u)$ is the same for all u, and is equal to the number of words with i nonzero elements

For any $(k,\ N,\ d)_s$ code, $\sum_i A_i = N$, $A_0 \geq 1$ (with equality iff the code is simple), $A_1 = \ldots = A_{d-1} = 0$, $A_i \geq 0$ for all i

Define the weight enumerator of C as the homogeneous polynomial:

$$W_C(x, y) = \sum_{i=0}^{k} A_i x^{k-i} y^i$$

A relationship between codes and OAs

The MacWilliams identity for linear codes:

$$W_{C^\perp}(x,y) = \frac{1}{N} W_C(x + (s-1)y, x - y)$$

(Jessie MacWilliams was a Cambridge graduate)

Using this identity, the weight distribution $(A_0^\perp, A_1^\perp, \ldots, A_k^\perp)$ of $C^\perp$, can be expressed in terms of the weight distribution of C:

$$A_i^\perp = \frac{1}{N} \sum_{j=0}^{k} A_j P_i(j), \;\; 0 \le i \le k,$$

where the $P_i(j)$ are the Krawtchouk polynomials.

$\sum_i A_i^\perp = s^k/N, \; A_0^\perp = 1, \; A_1^\perp = \ldots = A_{d^\perp - 1}^\perp = 0, \; A_i^\perp \ge 0$ for all i

If C is not linear ...

Linear programming bound

Punch line: If C is an arbitrary code, there is no $C^{\perp}$, but the $A_i^{\perp}$'s and $d^{\perp}$ can still be defined as in the MacWilliams identity, and will still meet the previous conditions (Delsarte, 1973)

LP bound for OAs (Delsarte, 1973): For given k, s and t, in an OA$(N,\ k,\ s,\ t)$, it holds that $N \geq N_{LP}$, where N_{LP} is the solution of:

$$\min \sum_i A_i \text{ subject to}$$

$$A_0 \geq 1,\ A_i \geq 0,\ 1 \leq i \leq k$$

$$A_0^{\perp} = 1,\ A_1^{\perp} = \ldots = A_t^{\perp} = 0,\ A_i^{\perp} \geq 0,\ t+1 \leq i \leq k,$$

where $A_i^{\perp}$'s are defined by MacWilliams identity (and $d^{\perp} = t+1$)

LP bound is always at least as good as Rao's bound

$A_i^{\perp}$'s are also at the basis of Generalized Minimum Aberration (Xu and Wu, 2001)

Outline

Hadamard matrices

A Hadamard matrix H_n of order n is an $n \times n$ matrix with entries 1 and -1 so that $H_n H_n^T = nI_n$

Can only exist for $n = 1, 2$, and $n = 4u$ for some u ...

Can always take it of the form $H_n = \begin{bmatrix} 1 & A \end{bmatrix}$; then A is an OA$(n,\ n-1,\ 2,\ 2)$ that meets Rao's bound

Further, $\begin{bmatrix} H_n \\ -H_n \end{bmatrix}$ is an OA$(2n,\ n,\ 2,\ 3)$ that meets Rao's bound (foldover)

Since arrays with $s = 2$ are useful in statistical applications, these relationships have been explored extensively

Difference schemes

A difference scheme (of strength 2) $D(r,\ c,\ s)$ is an $r \times c$ array based on the s elements of a group G so that for any two columns the element-wise "differences" contain every element of G equally often

Clearly $r = \lambda s$ for some λ, the index

A Hadamard matrix H_n is a difference scheme $D(n,\ n,\ 2)$

Another example, $D(6, 6, 3)$:

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 2 & 0 \\ 0 & 2 & 1 & 1 & 0 & 2 \\ 0 & 2 & 2 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 & 2 & 1 \\ 0 & 1 & 0 & 2 & 1 & 2 \end{bmatrix}$$

Difference schemes and OAs

Always, $c \leq r$; a Generalized Hadamard Matrix if $r = c$

If $D = D(r,\ c,\ s)$, then $\begin{bmatrix} D+0 \\ \vdots \\ D+(s-1) \end{bmatrix}$ is an OA$(rs,\ c,\ s,\ 2)$

("foldover") to which at least one more column can be added

Difference schemes of strength t (Hedayat, Stufken, Su, 1996); but ...

Best (?) known values for c:

$\lambda \backslash s$	2	3	4	5	6
1	2	3	4	5	2
2	4	6	8	10	6
3	2	9	12	7	2
4	8	12	16	20	6
5	2	9	8	25	2

Blue = Generalized Hadamard matrix
Black = Known to be best
Red = Best known ?

Outline

Mixed OAs

Some columns with s_1 levels, others with s_2, etc.

Important in statistics, but no coding theory equivalent

Definition of OA extends trivially; so do Rao's bounds

Construction is, typically, mathematically more challenging

There is also an extension of the LP bound (Sloane and Stufken, 1996); for columns of two types, need to define a two-dimensional weight distribution A_{i_1,i_2}, a weight enumerator using four variables, and a MacWilliams identity that defines $A^{\perp}_{i_1,i_2}$

$A^{\perp}_{i_1,i_2}$'s can be used to define Generalized Minimum Aberration for mixed OAs (Xu and Wu, 2001); but less knowledge about construction, existence and enumeration of arrays

Mixed OAs: Example 1

Question: For which $N, \ m, \ k \geq 2$ can we find an OA$(N, \ m^1 \ 2^k, \ 2)$?

Rao: $k \leq N - m$; also N must be a multiple of 4 and $2m$

Best results for $m = 3$:

N	12	24	36	48	60	72	84	96
k	4	16	27	40	30	63	28	88

Blue = Known to be the best possible

k is generally quite far from $N - m$ (and never more than $N - 8$)

Mixed OA: Illustration 2

Largest known arrays $OA(N, (N/4)^1\, 2^k, 2)$, $N \leq 100$,

N	k	N	k	N	k	N	k
12	4	16	12	20	8	24	14
28	12	32	24	36	16	40	22
44	15	48	36	52	16	56	30
60	17	64	48	68	18	72	38
76	19	80	60	84	20	88	46
92	21	96	72	100	22		

Rao: $k \leq N - m = 3N/4$; don't get close, except for $N \equiv 0 \pmod{16}$

Reach $N/2 + 2$ for $N \equiv 8 \pmod{16}$

Results for other cases mostly from computer search by Kuhfeld and Suen (2005)

Outline

The lattice of OAs of strength 2 with N columns

For a given N, what are the parameters for which an OA(N, $s_1^{k_1} s_2^{k_2} \ldots,$ 2) exists?

An N-row OA B is dominated by an N-row OA A if B can be obtained from A by one or more steps based on the expansive replacement method

In the expansive replacement method, the symbols in a column of A are replaced by rows of a smaller OA with a number of rows that is equal to the number of symbols in that column, yielding an new OA B

Can lead to more columns in B, but also to fewer if a column in A becomes constant

Examples of expansive replacement

OA$(8,\ 2^4\ 4^1,\ 2)$ dominates OA$(8,\ 2^7,\ 2)$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 2 & 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 0 & 1 \\ 3 & 0 & 1 & 0 & 1 \\ 3 & 1 & 0 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Converse is not as simple

Examples of expansive replacement

OA$(8,\ 2^4\ 4^1,\ 2)$ dominates OA$(8,\ 2^3\ 4^1,\ 2)$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 2 & 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 0 & 1 \\ 3 & 0 & 1 & 0 & 1 \\ 3 & 1 & 0 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 1 & 0 \\ 2 & 0 & 0 & 0 & 1 \\ 3 & 0 & 1 & 0 & 1 \\ 3 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Lattice

Domination provides for each N a partial ordering of the possible parameter sets $(N,\ s_1^{k_1} s_2^{k_2} \ldots)$

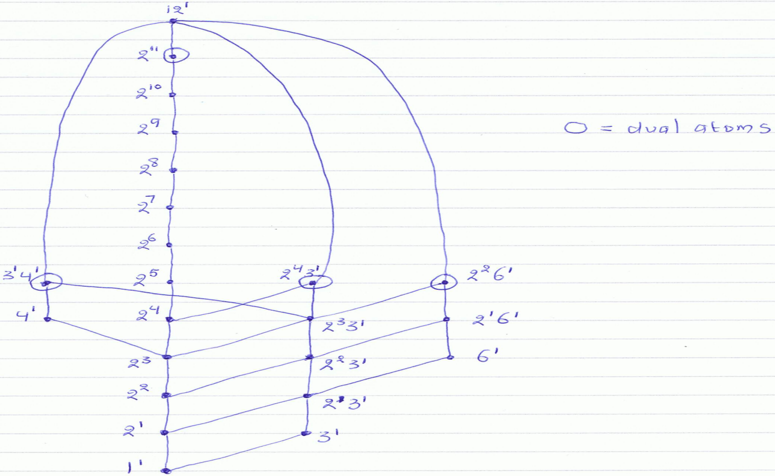
With the sets as nodes, this partial ordering is actually a lattice (with a meet and join)

Minimal element $(N,\ 1^1)$; maximal element $(N,\ N^1)$

Atoms are (N, p^1) for primes p that divide N; what is really of interest though are the dual atoms, which are the nodes just below the maximal element

The dual atoms (and thus the lattice) are completely known for some N, but for most they are not

Lattice of 12-row OAs



Height = 12 ; number of nodes = 23

number of dual atoms = 4

Dual atoms for selected N

$N = 8$: $(8,\ 2^4\ 4^1)$ is the single dual atom

$N = 12$: The 4 dual atoms are $(12,\ 2^{11})$, $(12,\ 2^4\ 3^1)$, $(12,\ 2^2\ 6^1)$, $(12,\ 3^1\ 4^1)$

$N = 24$: The 4 dual atoms are $(24,\ 2^{20}\ 4^1)$, $(24,\ 2^{13}\ 3^1\ 4^1)$, $(24,\ 2^{12}\ 12^1)$, $(24,\ 2^{11}\ 4^1\ 6^1)$ (Schoen, Eendebak, Nguyen, 2010)

$N = 28$: The 4 dual atoms are $(28,\ 2^{27})$, $(28,\ 2^{12}\ 7^1)$, $(28,\ 2^2\ 14^1)$, $(28,\ 4^1\ 7^1)$ (Schoen, Eendebak, Nguyen, 2010)

$N = 32$: The 2 dual atoms are $(32,\ 2^{16}\ 16^1)$, $(32,\ 4^8\ 8^1)$

$N = 64$: There are 7 dual atoms, not listed here, but all known (Rains, Sloane, Stufken, 2002)

Smallest unresolved case appears to be $N = 36$

Non-isomorphic OAs

Could take this a step further: For a particular parameter set, how many non-isomorphic OAs exist?

Theoretical results are very limited in scope (e.g. Seiden and Zemach, 1966; Hedayat, Seiden and Stufken, 1997; Stufken and Tang, 2007)

Schoen, Eendebak and Nguyen (2010) use an algorithm to find an almost complete enumeration of all OAs with (i) $t = 2,\ N \leq 28$; (ii) $t = 3,\ N \leq 64$; and (iii) $t = 4,\ N \leq 168$

Lattice of N-row OAs for strength $t \geq 3$?

Outline

Orthogonal Latin hypercube designs

A Latin hypercube is an $N \times k$ array in which each column is a permutation of $1, \ldots, N$ (LH(N, k))

Achieves uniformity in 1-dimensional projections

Tang (1993) showed how to obtain a LH(N, k) from an OA(N, k, s, t): Essentially, for each column in the OA, replace 0 by $1, \ldots, N/s$, replace 1 by $N/s + 1, \ldots, 2N/s$, and so on (orthogonal Latin hypercubes)

The strength t of the OA assures uniformity properties for the LH(N, k) in t-dimensional projections

Many variations, some using OAs, especially of strength 2, aiming at small values of N for given k in the resulting LH

For example, Lin, Mukerjee and Tang (2009), use an OA$(25, 6, 5, 2)$ to construct a LH$(25, 12)$ with uniformity properties in 2-dimensional projections

Nested Orthogonal Arrays

Nested orthogonal arrays were introduced to construct nested space filling designs for situations that combine low accuracy and high accuracy experiments

Looking for an OA($N_1,\ k,\ s_1,\ t$), say A, that contains an OA($N_2,\ k,\ s_2,\ t$), say B, as a subarray

Interesting combinatorial problem, with special cases that have been studied before

If A and B have the same index, then deleting the rows of B results in an incomplete orthogonal array (Maurin, 1985; Hedayat and Stufken, 1992)

There are bounds and methods of construction ...

Bounds for nested orthogonal arrays

Rao-type bounds from Mukerjee, Qian and Wu (2008) ($r = s_1/s_2$):

$t = 2u$: $N_1 \geq N_2 \left(1 + k(r-1) + \ldots + \binom{k}{u}(r-1)^u\right)$
$t = 2u + 1$:
$N_1 \geq N_2 \left(1 + k(r-1) + \ldots + \binom{k}{u}(r-1)^u + \binom{k-1}{u}(r-1)^{u+1}\right)$

In addition:

$$k \leq \frac{N_1 - N_2 r^{t-2}}{N_2 r^{t-2}(r-1)} + t - 2$$

Bounds only depend on N_1/N_2 and s_1/s_2

While the first bound is very similar to Rao's bounds for OAs, there is no similar proof (yet); equality is possible

Dey (2010) gives methods of construction (and extends the concept to nesting of mixed OAs)

Sliced OAs

Qian and Wu (2009) introduced the concept of sliced OAs and provided a number of constructions, including through the use of difference schemes

An OA($N_1,\ k,\ s_1,\ t$), say A, is called a sliced OA if its rows can be partitioned into N_1/N_2 arrays B_i, each with N_2 rows, such that after mapping the s_1 symbols to s_2 symbols, each B_i turns into an OA(N_2, k, s_2, t).

Motivation from computer experiments with quantitative and qualitative factors: When A is used to construct a LH, it can be done assuring good uniformity properties at each of the level combinations for the qualitative factors

For which parameters do these designs exist?

An example of a sliced OA

$$
\left[\begin{array}{ccc}
0 & 0 & 0 \\
0 & 1 & 1 \\
1 & 0 & 1 \\
1 & 1 & 0 \\
\hline
0 & 3 & 3 \\
0 & 2 & 2 \\
1 & 3 & 2 \\
1 & 2 & 3 \\
\hline
3 & 3 & 0 \\
3 & 2 & 1 \\
2 & 3 & 1 \\
2 & 2 & 0 \\
\hline
3 & 0 & 3 \\
3 & 1 & 2 \\
2 & 0 & 2 \\
2 & 1 & 3
\end{array}\right]
\qquad
\begin{array}{l}
0,\ 3 \rightarrow 0 \\
1,\ 2 \rightarrow 1
\end{array}
\qquad
\left[\begin{array}{ccc}
0 & 0 & 0 \\
0 & 1 & 1 \\
1 & 0 & 1 \\
1 & 1 & 0 \\
\hline
0 & 0 & 0 \\
0 & 1 & 1 \\
1 & 0 & 1 \\
1 & 1 & 0 \\
\hline
0 & 0 & 0 \\
0 & 1 & 1 \\
1 & 0 & 1 \\
1 & 1 & 0 \\
\hline
0 & 0 & 0 \\
0 & 1 & 1 \\
1 & 0 & 1 \\
1 & 1 & 0
\end{array}\right]
$$

Outline

OA of Type II or semi-balanced array

A semi-balanced array of strength t with N rows, k columns ($k \geq t$) and s symbols is an $N \times k$ array with s symbols so that no symbol is repeated in a row, and the rows of every $N \times t$ subarray contain each of the $\binom{s}{t}$ possible t-subsets λ times

Also known as an OA of Type II

For $t = 2$, semi-balanced arrays often have good properties for ordered units within blocks (correlation or interference effects in blocks) (Majumdar and Martin, 2004; Majumdar and Stufken, 2008)

Must have $k \leq s$; for $t = 2$, N must be a multiple of $s(s-1)/2$, and (if $k \geq 3$) an even multiple if s is odd

Existence and construction well established for s a prime or prime power (suffices to focus on smallest λ)

Related structures: OA of Type I; perpendicular arrays; perpendicular difference arrays; pairwise orthogonal Latin squares

Outline

Ordered orthogonal arrays

An Ordered Orthogonal Array, OOA(N, (k, ℓ), s, t) of strength t, is an $N \times k\ell$ array based on s symbols, in which the $k\ell$ columns are partitioned into k groups of ℓ ordered columns each, so that for every k-tuple $(d_1, \ldots, d_k)$ of integers with $0 \leq d_i \leq \ell$ and $\sum_i d_i = t$, the rows of the $N \times t$ subarray consisting of the first d_1 columns of the first group, the first d_2 columns of the second group, etc., contain every possible t-tuple equally often

Example: An OOA$(4, (2, 2), 2, 2)$:

$$\left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right]$$

Ordered orthogonal arrays

An Ordered Orthogonal Array, OOA$(N,\ (k,\ \ell),\ s,\ t)$ of strength t, is an $N \times k\ell$ array based on s symbols, in which the $k\ell$ columns are partitioned into k groups of ℓ ordered columns each, so that for every k-tuple $(d_1, \ldots, d_k)$ of integers with $0 \leq d_i \leq \ell$ and $\sum_i d_i = t$, the rows of the $N \times t$ subarray consisting of the first d_1 columns of the first group, the first d_2 columns of the second group, etc., contain every possible t-tuple equally often

Example: An OOA$(4, (2, 2), 2, 2)$:

$$\left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right]$$

Ordered orthogonal arrays

An Ordered Orthogonal Array, OOA(N, (k, ℓ), s, t) of strength t, is an $N \times k\ell$ array based on s symbols, in which the $k\ell$ columns are partitioned into k groups of ℓ ordered columns each, so that for every k-tuple $(d_1, \ldots, d_k)$ of integers with $0 \leq d_i \leq \ell$ and $\sum_i d_i = t$, the rows of the $N \times t$ subarray consisting of the first d_1 columns of the first group, the first d_2 columns of the second group, etc., contain every possible t-tuple equally often

Example: An OOA$(4, (2, 2), 2, 2)$:

$$\left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right]$$

Ordered orthogonal arrays

An Ordered Orthogonal Array, OOA$(N,\ (k,\ \ell),\ s,\ t)$ of strength t, is an $N \times k\ell$ array based on s symbols, in which the $k\ell$ columns are partitioned into k groups of ℓ ordered columns each, so that for every k-tuple $(d_1, \ldots, d_k)$ of integers with $0 \leq d_i \leq \ell$ and $\sum_i d_i = t$, the rows of the $N \times t$ subarray consisting of the first d_1 columns of the first group, the first d_2 columns of the second group, etc., contain every possible t-tuple equally often

Example: An OOA$(4, (2, 2), 2, 2)$:

$$\left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right]$$

Ordered orthogonal arrays

For $\ell = 1$, an OOA is just an OA($N, \ k, \ s, \ t$)

Existence of OOA implies existence of certain OAs, which can be used to come up with bounds

But both a Rao-type bound and an LP bound have now been developed specially for OOAs (Martin & Stinson, 1999)

Motivation for OOAs?

(T, M, S)-nets and (T, S)-sequences in base b

Used in numerical integration; occasionally in computer experiments

Let $I = [0, 1)^S$ be the S-dimensional unit hypercube, and define the elementary intervals in base b ($b \geq 2$) as

$$E = \prod_{i=1}^{S} [a_i b^{-d_i}, (a_i + 1) b^{-d_i}),$$

for non-negative integers a_i, d_i, $0 \leq a_i < b^{d_i}$. The volume is $b^{-\sum_i d_i}$

Points $\{x_1, \ldots, x_{b^M}\}$ in I form a (T, M, S)-net in base b if every interval E of volume b^{T-M} contains exactly b^T of the points

An infinite sequence $x_1, x_2, \ldots$ in I is a (T, S)-sequence in base b if, for any integers $k \geq 0$, $M > T$, the points $\{x_n : kb^M < n \leq (k+1)b^M\}$ form a (T, M, S)-net in base b (Sobol sequence if $b = 2$)

Relationship to ordered orthogonal arrays

A (T, M, S)-net in base b exists if and only if an OOA$(b^M, (S, M-T), b, M-T)$ exists

Uses only the subclass of OOAs with group size and strength equal

Previously, bounds for nets were based on bounds for OAs:
$S \leq \min_{T+2 \leq m \leq M} f(b^m, b, m-T)$ (Lawrence, 1995)

The Rao-type bound and an LP bound by Martin & Stinson, 1999, give generally better bounds for the existence of nets

Other applications for OOAs?

Outline

Definition of BOMA

A proper Balanced Orthogonal Multi-Array (BOMA) of strength 2 is an $N \times k$ array in which the entry in position (i, j), α_{ij}, is a subset of size n_j of $S_j = \{0, 1, \ldots, s_j - 1\}$ so that:

1. In each column, the symbols that can appear must appear equally often, and the sets form the blocks of a BIBD (2-design) if $n_j \geq 2$
2. For every two columns j_1 and j_2, and for $\ell_1 \in S_{j_1}$ and $\ell_2 \in S_{j_2}$, the cardinality of $\{i : 1 \leq i \leq N,\ \ell_1 \in \alpha_{ij_1},\ \ell_2 \in \alpha_{ij_2}\}$ does not depend on the choice of ℓ_1 and ℓ_2

Notation: BOMA$(N,\ s_1 s_2 \ldots s_k, n_1 n_2 \ldots n_k)$

If $n_j = 1$ for all j, this is just an OA of strength 2

Introduced and studied by Sitter (1993), Mukerjee (1998)

Example: A BOMA(12, $4^3\,3^1$, $2^3\,1^1$)

$$
\begin{bmatrix}
01 & 01 & 01 & 0 \\
01 & 23 & 23 & 0 \\
02 & 02 & 02 & 1 \\
02 & 13 & 13 & 1 \\
03 & 03 & 03 & 2 \\
03 & 12 & 12 & 2 \\
12 & 03 & 12 & 2 \\
12 & 12 & 03 & 2 \\
13 & 02 & 13 & 1 \\
13 & 13 & 02 & 1 \\
23 & 01 & 23 & 0 \\
23 & 23 & 01 & 0
\end{bmatrix}
$$

Bound for BOMAs

For a BOMA(N, $s_1 s_2 \ldots s_k, n_1 n_2 \ldots n_k$), it must hold that

$$N \geq \sum_{j=1}^{k} (s_j - 1)$$

Equality holds for previous example; but often equality cannot be achieved

If $n_j = 1$ for all j, this is precisely Rao's bound for strength 2 OAs

Note that bound does not depend on n_j's

There is a general construction method using α-resolvable BIBDs and OAs

Outline

Miscellaneous

Pairwise orthogonal Latin squares and OAs of strength 2 and index unity: An OA($s^2, k+2, s, 2$) exists if and only if there are k pairwise orthogonal Latin squares of order s

Pairwise orthogonal F-squares and OAs of strength 2: Pairwise orthogonal F-squares can easily be converted to OAs of strength 2 (including mixed OAs), but there is no simple converse of this

Simple Orthogonal Multi-Arrays SOMA(k, n)

Covering Arrays

Outline

Conclusions

OAs have been studied extensively ...

... but remain an active area of research

Includes

- Existence and bounds (especially for mixed OAs)
- Methods of construction
- Enumeration
- Arrays with additional properties
- Several structures closely related to OAs