Association schemes in designed experiments

R. A. Bailey
University of St Andrews
QMUL (emerita)

25th International Workshop on Matrices and Statistics
(IWMS 2016),
Funchal, June 2016

Abstract

Association schemes arise in designed experiments in many ways.

They were first used in incomplete-block designs, but they are implicit in the treatment structure of factorial designs and in many common block structures, such as row-column designs or nested blocks.

What is nice about them is the link between the matrices which show the patterns and the matrices which project onto the common eigenspaces.

In recent work, Agnieszka Łacka and I have considered designs where the treatments consist of all combinations of levels of two treatment factors and one additional control treatment. We construct nested row-column designs which have what we call control orthogonality and supplemented partial balance.

An example of a designed experiment

Example
I want to compare 6 new varieties of wheat (3 Dutch, 3 British).
I can use 30 plots, which form a 5×6 rectangle in a single field.
How do I decide which variety to plant in which plot?

In general,

- Ω is a set of experimental units,
- $\mathcal{T}$ is a set of treatments,
- a design is a function $f : \Omega \to \mathcal{T}$ such that if $\omega \in \Omega$ then $f(\omega)$ is the treatment allocated to unit ω.

Ω is given: usually it has some sort of structure.
$\mathcal{T}$ is given: usually it has some sort of structure.

My job is to choose f, possibly subject to some constraints.

Outline

I. What is an association scheme?

The structure on the set Ω of experimental units is often an association scheme.
The structure on the set $\mathcal{T}$ of treatments is often another association scheme.

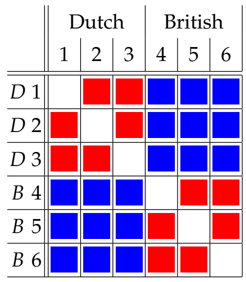
But what is an association scheme?

An association scheme on Ω with s associate classes is a partition of the unordered pairs of distinct elements of Ω into s classes such that . . .
(a technical condition that I'll tell you about later).

And similarly for $\mathcal{T}$.

First example of an association scheme ($s = 2$)

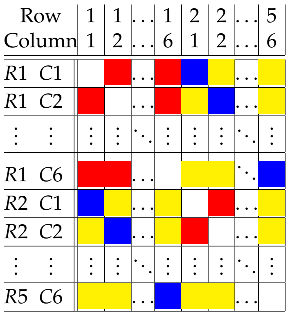
$\mathcal{T}$ consists of 6 new varieties of wheat (3 Dutch, 3 British).



Class 1, shown in red: different varieties from the same country.
Class 2, shown in blue: varieties from different countries.
Class 0, shown in white: same variety.

Second example of an association scheme ($s = 3$)

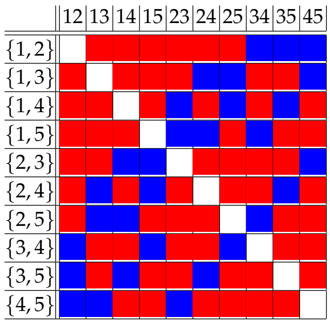
Ω is a 5×6 rectangle.



Class 1, shown in red: different plots in the same row.
Class 2, shown in blue: different plots in the same column.
Class 3, shown in yellow: plots in different rows and columns.
Class 0, shown in white: same plot.

Third example of an association scheme ($s = 2$)

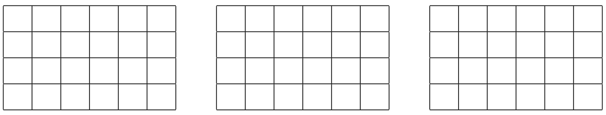
Diallel cross from 5 parental lines with no self-crosses



Class 1, in red: crosses with one parental line in common.
Class 2, in blue: crosses with no parental line in common.
Class 0, in white: same cross.

Fourth example of an association scheme ($s = 4$)

3 blocks, each consisting of 4 rows $\times$ 6 columns

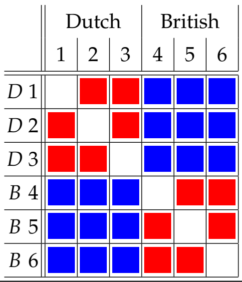


Class 0: same plot.
Class 1: different plots in the same row.
Class 2: different plots in the same column.
Class 3: plots in the same block but different rows and columns.
Class 4: plots in different blocks.

Adjacency matrices

So where do matrices come in?
Given an association scheme on a set of size N,
the adjacency matrix for class i is the $N \times N$ matrix

$$\text{whose } (\alpha, \beta) \text{ entry is equal to } \begin{cases} 1 & \text{if } (\alpha,\beta) \text{ is in class } i \\ 0 & \text{otherwise.} \end{cases}$$



$$A_0 = I$$

$$A_1 = \begin{bmatrix} 0&1&1&0&0&0 \\ 1&0&1&0&0&0 \\ 1&1&0&0&0&0 \\ 0&0&0&0&1&1 \\ 0&0&0&1&0&1 \\ 0&0&0&1&1&0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0&0&0&1&1&1 \\ 0&0&0&1&1&1 \\ 0&0&0&1&1&1 \\ 1&1&1&0&0&0 \\ 1&1&1&0&0&0 \\ 1&1&1&0&0&0 \end{bmatrix}$$

Definition of association scheme

Definition
An association scheme with s classes on a set of size N consists of $N \times N$ adjacency matrices $A_0, A_1, A_2, \ldots, A_s$ satisfying

(i) $A_0 = I$ (the identity matrix);
(ii) $A_0 + A_1 + A_2 + \cdots + A_s = J$ (the all-1 matrix);
(iii) A_i is symmetric for $i = 0, 1, 2, \ldots, s$;
(iv) For $0 \le i \le s$ and $0 \le j \le s$,
A_iA_j is a linear combination of $A_0, A_1, A_2, \ldots, A_s$.

The diagonal entries of A_i^2 are equal to the row sums of A_i,
so an easy consequence of this definition is that
there are integers $a_0 = 1, a_1, a_2, \ldots, a_s$ such that
A_i has a_i non-zero entries in every row and in every column.

First example ($N = 6$, $s = 2$)

$$A_0 = I \quad A_1 = \begin{bmatrix} 0&1&1&0&0&0 \\ 1&0&1&0&0&0 \\ 1&1&0&0&0&0 \\ 0&0&0&0&1&1 \\ 0&0&0&1&0&1 \\ 0&0&0&1&1&0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0&0&0&1&1&1 \\ 0&0&0&1&1&1 \\ 0&0&0&1&1&1 \\ 1&1&1&0&0&0 \\ 1&1&1&0&0&0 \\ 1&1&1&0&0&0 \end{bmatrix}$$

$$a_0 = 1 \qquad a_1 = 2 \qquad a_2 = 3$$

$$A_0A_i = A_iA_0 = A_i \quad \text{for } i = 0, 1, 2.$$

$$A_iJ = JA_i = a_iJ \quad \text{for } i = 0, 1, 2.$$

So it is enough to check A_1^2.

$$A_1^2 = 2I + A_1,$$

so this is an association scheme.

Bose–Mesner algebra

Given an association scheme with s classes on a set of size N, let $\mathcal{A}$ be the set of all real linear combinations of $A_0, A_1, A_2, \ldots, A_s$. Conditions (iii) and (iv) show that $\mathcal{A}$ is a commutative algebra. It is called the Bose–Mesner algebra.

Theorem

There are subspaces $W_0, W_1, \ldots, W_s$ of $\mathbb{R}^N$ such that

1. *if $j \neq k$ then W_j is orthogonal to W_k;*
2. *every vector $\mathbf{v}$ in $\mathbb{R}^N$ can be expressed uniquely as $\mathbf{v} = \sum_j \mathbf{w}_j$ with $\mathbf{w}_j \in W_j$;*
3. *W_j is contained in an eigenspace of A_i, for all i and all j;*
4. *the idempotent matrix P_j of orthogonal projection onto W_j is in $\mathcal{A}$, so it is a linear combination of $A_0, \ldots, A_s$;*
5. *if λ_{ij} is the the eigenvalue of A_i on W_j then $A_i = \sum_j \lambda_{ij} P_j$;*
6. *W_0 is the 1-dimensional space of constant vectors.*

Eigenspaces in our examples

$\mathcal{T}$ consists of 6 new varieties of wheat (3 Dutch, 3 British).
$N = 6$ and $s = 2$; $a_0 = 1$, $a_1 = 2$ and $a_2 = 3$.

subspace	description	dimension
W_0	constant vectors	1
W_1	differences between countries	1
W_2	differences within countries	4

Ω is a 5×6 rectangle.
$N = 30$ and $s = 3$; $a_0 = 1$, $a_1 = 5$, $a_2 = 4$ and $a_3 = 20$.

subspace	description	dimension
W_0	constant vectors	1
W_1	differences between rows	4
W_2	differences between columns	5
W_3	everything else	20

Eigenspaces in more of our examples

Diallel cross from 5 parental lines with no self-crosses.
$N = 10$ and $s = 2$; $a_0 = 1$, $a_1 = 6$ and $a_2 = 3$.

subspace	description	dimension
W_0	constant vectors	1
W_1	main effects of parental lines	4
W_2	everything else	5

3 blocks, each consisting of 4 rows $\times$ 6 columns.
$N = 72$ and $s = 4$; $a_0 = 1$, $a_1 = 5$, $a_2 = 3$, $a_3 = 15$ and $a_4 = 48$.

subspace	description	dimension
W_0	constant vectors	1
W_1	differences between blocks	2
W_2	differences between rows within blocks	9
W_3	differences between columns within blocks	15
W_4	everything else	45

II. Incomplete-block designs

Ω is a set of experimental units;
$\mathcal{T}$ is a set of treatments;
$f(\omega)$ is the treatment allocated to unit ω.
Suppose that Ω is partitioned into blocks of equal size, and $g(\omega)$ is the block containing unit ω.

The usual linear model for the response Y_ω is

$$E(Y_\omega) = \tau_{f(\omega)} + \beta_{g(\omega)} \qquad \text{and} \qquad \mathrm{Var}(\mathbf{Y}) = \sigma^2 I.$$

In vector form

$$\mathbf{Y} = X_1 \boldsymbol{\tau} + X_2 \boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where the ω-row of matrix X_1 picks out the treatment $f(\omega)$ and the ω-row of matrix X_2 picks out the block $g(\omega)$.

Concurrence matrix and information matrix

$$\mathbf{Y} = X_1 \boldsymbol{\tau} + X_2 \boldsymbol{\beta} + \boldsymbol{\epsilon}.$$

Put $N_{12} = X_1' X_2$, whose entries show how often each treatment occurs in each block (we assume at most once).
The entries in the concurrence matrix $N_{12} N_{12}'$ show how often each pair of treatments concur in blocks.
The information matrix is $C = X_1' X_1 - k^{-1} N_{12} N_{12}'$, where k is the block size.

Under normality, the BLUE of $\boldsymbol{\tau}$ is

$$\hat{\boldsymbol{\tau}} = C^- X_1' (I - k^{-1} X_2 X_2') \mathbf{Y}$$

and

$$\mathrm{Var}(\hat{\boldsymbol{\tau}}) = \sigma^2 C^-.$$

Statisticians needed to be able to calculate C^- easily in the pre-computer age.

Partial balance

$N_{12} = X_1' X_2$ and $C = X_1' X_1 - k^{-1} N_{12} N_{12}'$; also $\mathrm{Var}(\hat{\boldsymbol{\tau}}) = \sigma^2 C^-$.

Definition

An incomplete-block design is partially balanced if there is an association scheme on $\mathcal{T}$ such that the information matrix C is in its Bose–Mesner algebra $\mathcal{A}$.

If $C = \sum_{i=0}^{s} \mu_i P_i$, where $P_0, P_1, \ldots, P_s$ are the primitive idempotents, then $\mu_0 = 0$. If all treatment contrasts are estimable then $\mu_i > 0$ for $i = 1, \ldots, s$. So

$$C^- = \sum_{i=1}^{s} \frac{1}{\mu_i} P_i \in \mathcal{A},$$

and so $\mathrm{Var}(\hat{\tau}_l - \hat{\tau}_m)$ depends only on the associate class containing (l, m). Moreover, the variance of every normalized treatment contrast in W_j is σ^2 / μ_j.

Advantages of partial balance

1. Easy to calculate C^-.
2. If the subspaces $W_0, W_1, \dots, W_s$ of the association scheme have a practical interpretation then the equal variance of the estimators of all treatment contrasts in the same subspace is useful.
   If the association scheme comes from all combinations of two or more treatment factors (like our rectangle example), the design is said to have factorial balance (Yates, 1935).
3. For many values of the numbers of treatments, blocks, and plots per block, there is a partially balanced design which minimizes the average value of $\mathrm{Var}(\hat{\tau}_l - \hat{\tau}_m)$.

An example of a PBIBD

$\mathcal{T}$ consists of 6 new varieties of wheat (3 Dutch, 3 British).
$N = 6$ and $s = 2$; $a_0 = 1$, $a_1 = 2$ and $a_2 = 3$.
9 blocks of size 4.

D1	D2	B4	B5	D1	D2	B4	B6	D1	D2	B5	B6
D1	D3	B4	B5	D1	D3	B4	B6	D1	D3	B5	B6
D2	D3	B4	B5	D2	D3	B4	B6	D2	D3	B5	B6

$$N_{12}N_{12}' = 6I + 3A_1 + 4A_2$$

$$C = \frac{1}{4}(18I - 3A_1 - 4A_2) = 6P_1 + \frac{21}{4}P_2$$

$$C^- = \frac{1}{6}P_1 + \frac{4}{21}P_2$$

The comparison between countries has the same variance as in an unblocked design; for comparisons within each country the variance is increased by 8/7 if σ^2 is unchanged.

III. Blocking factors with random effects: an example

Ω is a 5×6 rectangle.
$N = 30$ and $s = 3$; $a_0 = 1$, $a_1 = 5$, $a_2 = 4$ and $a_3 = 20$.

subspace	description	dimension
W_0	constant vectors	1
W_1	differences between rows	4
W_2	differences between columns	5
W_3	everything else	20

If rows and columns have random effects then

$$\begin{aligned}\mathrm{Var}(\mathbf{Y}) &= \sigma^2 I + \sigma_R^2(I + A_1) + \sigma_C^2(I + A_2) \\ &= \sigma^2 P_3 + (\sigma^2 + 5\sigma_C^2)P_2 + (\sigma^2 + 6\sigma_R^2)P_1 \\ &\quad + (\sigma^2 + 5\sigma_C^2 + 6\sigma_R^2)P_0 \\ &= \xi_3 P_3 + \xi_2 P_2 + \xi_1 P_1 + \xi_0 P_0\end{aligned}$$

where ξ_3, ξ_2, ξ_1 and ξ_0 are spectral components of variance.

A mixed model

Ω is a set of experimental units; $\mathcal{T}$ is a set of treatments;
X_1 is the units-by-treatments incidence matrix.

Assume that $E(\mathbf{Y}) = X_1\boldsymbol{\tau}$
and that there is an association scheme on Ω,
with Bose–Mesner algebra $\mathcal{A}$, such that $\mathrm{Var}(\mathbf{Y}) \in \mathcal{A}$.

$$\mathrm{Var}(\mathbf{Y}) = \sum_{j=0}^{s} \xi_j P_j$$

with $\xi_j \geq 0$ for $j = 0, 1, \dots, s$. We usually ignore ξ_0, and assume no linear dependence among $\xi_1, \dots, \xi_s$.

$\mathrm{Var}(P_j\mathbf{Y}) = \xi_j P_j$, which is effectively a scalar matrix,
so the projected data $P_j\mathbf{Y}$ can be analysed in the usual way,
with information matrix C_j.

Combining information from the analyses in the strata $W_1, \dots, W_s$ is more straightforward when $C_1, \dots, C_s$ have common eigenspaces. This is called general balance (Nelder, 1965).

Nested row-column designs

b blocks, each consisting of n_1 rows $\times$ n_2 columns.
$N = bn_1n_2$ and $s = 4$; $a_0 = 1$, $a_1 = n_2 - 1$, $a_2 = n_1 - 1$,
$a_3 = (n_1 - 1)(n_2 - 1)$ and $a_4 = (b - 1)n_1n_2$.

stratum	description	dimension	info matrix
W_0	constant vectors	1	
W_1	blocks	$b - 1$	C_1
W_2	rows within blocks	$b(n_1 - 1)$	C_2
W_3	columns within blocks	$b(n_2 - 1)$	C_3
W_4	everything else	$b(n_1 - 1)(n_2 - 1)$	C_4

$$\mathrm{Var}(\mathbf{Y}) = \xi_4 P_4 + \xi_3 P_3 + \xi_2 P_2 + \xi_1 P_1 + \xi_0 P_0,$$

where we expect $0 < \xi_4 \leq \xi_3 \leq \xi_1$ and $0 < \xi_4 \leq \xi_2 \leq \xi_1$.

Corresponding information matrices C_4, C_3, C_2 and C_1.

IV. Near-factorial treatments

Treatment factor T is quantitative,
with one level 0 and t non-zero levels.
Treatment factor U is qualitative,
with u levels, which are all irrelevant when $T = 0$.

Example (Cochran and Cox, 1957, Chapter 3)

$T =$ dose, with levels 0, single and double; $t = 2$.
$U =$ type of chemical to control eelworms; $u = 4$.
If K and K' are two chemicals then
a zero dose of K is the same as a zero dose of K'.

Example (Cochran and Cox, 1957, Chapter 4)

$T =$ amount of sulphur to spread on the soil,
with levels 0, 300, 600 and 1200 lb per acre; $t = 3$.
$U =$ timing, with levels Spring and Autumn; $u = 2$.
Zero sulphur in Spring is the same as a zero sulphur in Autumn.

Control treatment plus $t \times u$ factorial

dose T ($t = 2$)	chemical U ($u = 4$) none	K_1	K_2	K_3	K_4
0	✓				
single		✓	✓	✓	✓
double		✓	✓	✓	✓

amount T ($t = 3$)	timing ($u = 2$) never	Spring	Autumn
0	✓		
300		✓	✓
600		✓	✓
1200		✓	✓

Treatment subspaces

This is not an association scheme,
but the relevant treatment subspaces are clear.

subspace	description	dimension
W_0	overall mean	1
W_{control}	control versus rest	1
W_T	main effect of T	$t-1$
W_U	main effect of U	$u-1$
W_{TU}	interaction between T and U	$(t-1)(u-1)$

V. Nested row-column designs for near-factorial treatments

stratum	description	dimension	info matrix
W_0	constant vectors	1	
W_1	blocks	$b-1$	C_1
W_2	rows within blocks	$b(n_1-1)$	C_2
W_3	columns within blocks	$b(n_2-1)$	C_3
W_4	everything else	$b(n_1-1)(n_2-1)$	C_4

subspace	description	dimension
W_0	overall mean	1
W_{control}	control versus rest	1
W_T	main effect of T	$t-1$
W_U	main effect of U	$u-1$
W_{TU}	interaction between T and U	$(t-1)(u-1)$

How can we match the treatment subspaces to the strata?

Control orthogonality

Definition
A nested row-column design for near-factorial treatments has control orthogonality if the control treatment occurs equally often in each row and occurs equally often in each column.

This implies that the contrast between the control treatment and the rest is in the null space of C_0, C_1, C_2 and C_3, so it is estimated "with full efficiency" in the bottom stratum W_4 (but its variance depends on the relative replications).

Supplemented partial balance

Definition
A nested row-column design for near-factorial treatments has supplemented partial balance if the non-control treatments are equally replicated and there is an association scheme on them such that, for $i = 1, 2, 3$ and 4,

$$C_i = \begin{bmatrix} tuc_i & -c_i & \dots & -c_i \\ -c_i & & & \\ \vdots & & L_i & \\ -c_i & & & \end{bmatrix},$$

where c_1, c_2, c_3 and c_4 are scalars and L_1, L_2, L_3 and L_4 are matrices in the Bose–Mesner algebra of the association scheme.

If the design also has control orthogonality then $c_1 = c_2 = c_3 = 0$.

Suitable association schemes

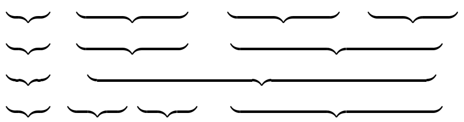
Under supplemented partial balance,
the eigenspaces of C_1, C_2, C_3 and C_4 are

W_0 with 1 df
W_{control} with 1 df
within $T \times U$, as per the association scheme.

Agnieszka Łacka and I use association schemes on the non-control treatments that are consistent with the factorial association scheme in the sense that the two sets of subspaces have a common decomposition.

Some association schemes consistent with $T \times U$

A.S.	Eigenspaces			
Factorial	mean	main effect of T	main effect of U	interaction
		W_T	W_U	W_{TU}
Rect(t,u)				
GD(t,u)				
Trivial				
EGD(p,q,u)				



"Consistent" means that
the primitive idempotents of the association scheme
commute with
the primitive idempotents of the factorial association scheme.

Equivalently, the adjacency matrices of the two association schemes commute with each other.

Some easy constructions for square blocks

1. If $n_1 = n_2 = tu + c$, start with a Latin square of order n_1.
   Replace c letters by the control treatment,
   and the remaining tu letters by the factorial treatments.
   Do this in each block.
2. If $n_1 = n_2 = tu$, start with a Latin square of order n_1 which has every letter on the main diagonal
   (this is possible because $tu \geq 4$).
   Replace all letters on the main diagonal by the control treatment.
   Do this in each block.

A	*D*	*E*	*C*	*F*	*B*
D	*B*	*F*	*E*	*C*	*A*
E	*F*	*C*	*B*	*A*	*D*
F	*C*	*A*	*D*	*B*	*E*
B	*A*	*D*	*F*	*E*	*C*
C	*E*	*B*	*A*	*D*	*F*

$\longrightarrow$

0	*D*	*E*	*C*	*F*	*B*
D	0	*F*	*E*	*C*	*A*
E	*F*	0	*B*	*A*	*D*
F	*C*	*A*	0	*B*	*E*
B	*A*	*D*	*F*	0	*C*
C	*E*	*B*	*A*	*D*	0

Some other constructions for square blocks

3. If $n_1 = n_2 = tu$ and $b = t$, start with a Latin square of order t. In block i, replace letter i by a $u \times u$ square of controls, and replace letter j by a $u \times u$ Latin square containing all factorial treatments with level j of T.

1	2
2	1

$\longrightarrow$

0	0	0	21	22	23
0	0	0	23	21	22
0	0	0	22	23	21
22	23	21	0	0	0
23	21	22	0	0	0
21	22	23	0	0	0

12	11	13	0	0	0
13	12	11	0	0	0
11	13	12	0	0	0
0	0	0	11	12	13
0	0	0	12	13	11
0	0	0	13	11	12

4. Similar, but, in each $u \times u$ subsquare of controls in block i, replace the controls on the diagonal by the factorial treatments with level i of T.

These give higher variances for the main effect of T.

Some constructions for rectangular blocks

5. If $n_1 = u \geq 3$ and $n_2 = tu$, start with a Latin square of order u which has every letter on the main diagonal. Replace all letters on the main diagonal by the control treatment. Put t copies of this square side by side. In the i-th copy, use level i of T with the non-control levels of U.

0	11	13	12	0	21	23	22	0	31	33	32
12	0	11	14	22	0	21	24	32	0	31	34
11	14	0	13	21	24	0	23	31	34	0	33
13	12	14	0	23	22	24	0	33	32	34	0

This has even larger variances for the main effect of T, but . . .

6. If $t = u$ we can confound different factorial effects (such as U, $U + T$, $U + 2T$, etc.) with columns in different blocks.

Another special construction when $t = u$

7. If $t = u$, $n_1 = t + 1$, $n_2 = t(t + 1)$, and t is a power of a prime, start with a balanced incomplete-block design for $t^2 + t + 1$ treatments in $t^2 + t + 1$ blocks of size $t + 1$.
   This can be arranged as a $(t + 1) \times (t^2 + t + 1)$ rectangle with blocks as columns and all treatments once per row.

A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*
B	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*A*
E	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*A*	*B*	*C*	*D*
G	*H*	*I*	*J*	*K*	*L*	*M*	*A*	*B*	*C*	*D*	*E*	*F*

For every letter in the last block, replace every occurrence by the control treatment. Then remove the last column.

0	*B*	*C*	0	*E*	0	*G*	*H*	*I*	*J*	*K*	*L*	0
B	*C*	0	*E*	0	*G*	*H*	*I*	*J*	*K*	*L*	0	0
E	0	*G*	*H*	*I*	*J*	*K*	*L*	0	0	*B*	*C*	0
G	*H*	*I*	*J*	*K*	*L*	0	0	*B*	*C*	0	*E*	0

References

- R. A. Bailey. *Association Schemes. Designed Experiments, Algebra and Combinatorics*. Cambridge University Press, Cambridge, 2004.
- R. A. Bailey & A. Łacka. Nested row-column designs for near-factorial experiments with two treatment factors and one control treatment. *Journal of Statistical Planning and Inference* **165** (2015), pp. 63–77.
- W. G. Cochran & G. M. Cox. *Experimental Designs*, 2nd edition. Wiley, 1957.
- J. A. Nelder. The analysis of randomized experiments with orthogonal block structure. II. Treatment structure and the general analysis of variance. *Proceedings of the Royal Society, Series A* **283** (1965), pp. 163–178.
- F. Yates. Complex experiments. *Journal of the Royal Statistical Society, Supplement* **2** (1935), pp. 181–247.