Some applications of finite group theory in the design of experiments

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

Algebra and Combinatorics seminar, St Andrews, 13 November 2019

Abstract

Group theory is used in (at least) two different ways in the design of experiments.

The first is in randomization, the process by which an initial design is turned into the actual layout for the experiment by applying a permutation of the experimental units, chosen at random from a certain group of permutations. Which group? What properties should it have?

The second is in design construction. The set of treatments is identified with a finite Abelian group, and the blocks are all translates of one or more initial blocks. The characters of this group form its dual group: they are the eigenvectors of the matrix that we need to consider to see how good the proposed design is.

Outline

First course

Randomization and permutation groups

Introductory example

I am going to do an experiment to compare 3 varieties of tomato, so see which gives me the biggest yield (in weight of fruit per plant).
My greenhouse has room for 9 tomato plants in a row.

position ("plot")	1	2	3	4	5	6	7	8	9
variety ("treatment")	A	A	A	B	B	B	C	C	C

In general,

$$\begin{aligned} \Omega &= \text{set of plots} \\ \mathcal{T} &= \text{set of treatments} \\ f\colon \Omega \to \mathcal{T} & \quad \text{tells you to put treatment } f(\omega) \text{ on plot } \omega \\ Y_\omega &= \text{yield on plot } \omega \end{aligned}$$

Simplistic assumption

$f\colon \Omega \to \mathcal{T}$	tells you to put treatment $f(\omega)$ on plot ω
$Y_\omega =$	yield on plot ω

$$Y_\omega = \tau_{f(\omega)} + \epsilon_\omega$$

τ_i is an unknown constant depending only on treatment i
(we want to estimate all of these, or, at least, their differences)
ϵ_ω is a random variable depending only on plot ω
with zero mean, same unknown variance σ^2 for all plots, and independence between different plots.

Put all the yields Y_ω into a column vector $\mathbf{Y}$. If $N = |\Omega|$ then the $N \times N$ covariance matrix $\mathrm{Cov}(\mathbf{Y})$ for $\mathbf{Y}$ has (ω,ω)-entry equal to $\mathrm{Var}(\epsilon_\omega)$ and (α,β)-entry equal to $\mathrm{Cov}(\epsilon_\alpha,\epsilon_\beta)$.
Our assumptions tell us that $\mathrm{Cov}(\mathbf{Y}) = \sigma^2 I$.
All textbooks tell you how to analyse data under this assumption, but it is unrealistically simple.

Completely randomized designs

Start with an initial layout.

Choose a permutation g at random from $\mathrm{Sym}(\Omega)$ and apply it to the initial layout before doing the experiment.

Every pair of distinct plots is equally likely to be replaced by any other pair, so this randomization "makes it reasonable" to assume that there are unknown constants κ_1 and κ_2 such that $\mathrm{Cov}(\mathbf{Y}) = \kappa_1 I + \kappa_2 (J - I)$.

The eigenspaces of this are V_0 and $V_0^\perp$,
where V_0 is the 1-dimensional subspace of constant vectors.

After the experiment, project the data vector onto $V_0^\perp$.
The covariance matrix of the projected data is effectively scalar, so it can be analysed by textbook methods.

General randomization

Sometimes we do not want to use the whole of $\mathrm{Sym}(\Omega)$ (examples coming up!).

Let G be a given transitive subgroup of $\mathrm{Sym}(\Omega)$.

Randomize by using a random permutation from G.

This lets us assume that if (α,β) and (γ,δ) are in the same orbit of G in its action on $\Omega \times \Omega$ then $\mathrm{Cov}(Y_\alpha, Y_\beta) = \mathrm{Cov}(Y_\gamma, Y_\delta)$, so that $\mathrm{Cov}(\mathbf{Y})$ is in the centralizer algebra of G.

If the permutation character of G is multiplicity-free
then we know the eigenspaces of $\mathrm{Cov}(\mathbf{Y})$
even though we do not know its entries.

So we can project the data vector onto each eigenspace and proceed as before.

An example with the direct product

The experimenter wants to compare 4 exercise regimes.
8 people will take part for 4 months, changing their regime each month. Various health indicators will be measured on each person at the end of each month.

A	B	C	D	C	B	A	D
B	A	D	C	D	C	B	A
C	D	A	B	A	D	C	B
D	C	B	A	B	A	D	C

$G = S_4 \times S_8$, the direct product of S_4 and S_8
(randomize rows; independently randomize columns).

The orbits on pairs are
$\{(\alpha,\beta) : \alpha = \beta\}$
$\{(\alpha,\beta) : \alpha \neq \beta$ but in same row$\}$
$\{(\alpha,\beta) : \alpha \neq \beta$ but in same column$\}$
$\{(\alpha,\beta) :$ other$\}$

An example with the direct product: eigenspaces

$G = S_4 \times S_8$, the direct product of S_4 and S_8

The orbits on pairs are
$\{(\alpha,\beta) : \alpha = \beta\}$
$\{(\alpha,\beta) : \alpha \neq \beta$ but in same row$\}$
$\{(\alpha,\beta) : \alpha \neq \beta$ but in same column$\}$
$\{(\alpha,\beta) :$ other$\}$

G fixes the vector subspaces V_0 (constant vectors, dimension 1),
V_R (vectors which are constant on each row, dimension 4),
V_C (vectors which are constant on each column, dimension 8),
and the whole space V (dimension 32).

The eigenspaces are

$$\begin{aligned} W_0 &= V_0 \\ W_R &= V_R \cap V_0^\perp \\ W_C &= V_C \cap V_0^\perp \\ W &= V \cap (V_0 + V_C + V_R)^\perp \end{aligned}$$

An example with the wreath product

An environmental researcher wants to compare 7 different methods of preparing the soil for a wheat crop
(such as conventional ploughing, various chemicals, etc).
7 different farmers have agreed to take part in the trial,
but each can offer only three fields.

A	B	C	D	E	F	G
B	C	D	E	F	G	A
D	E	F	G	A	B	C

$G = S_7/S_3 = S_3 \wr S_7$, the wreath product of S_3 and S_7
(randomize farms;
independently randomize fields within each farm).

The orbits on pairs are
$\{(\alpha,\beta) : \alpha = \beta\}$
$\{(\alpha,\beta) : \alpha \neq \beta$ but in same farm$\}$
$\{(\alpha,\beta) :$ other$\}$

An example with the wreath product: eigenspaces

$G = S_7/S_3 = S_3 \wr S_7$, the wreath product of S_3 and S_7
(randomize farms;
independently randomize fields within each farm).

The orbits on pairs are
$\{(\alpha,\beta) : \alpha = \beta\}$
$\{(\alpha,\beta) : \alpha \neq \beta$ but in same farm$\}$
$\{(\alpha,\beta) :$ other$\}$

G fixes the vector subspaces V_0 (constant vectors, dimension 1),
V_F (vectors which are constant on each farm, dimension 7),
and the whole space V (dimension 21).

The eigenspaces are

$$\begin{aligned} W_0 &= V_0 \\ W_F &= V_F \cap V_0^\perp \\ W &= V \cap (V_0 + V_F)^\perp \end{aligned}$$

Direct products and wreath products

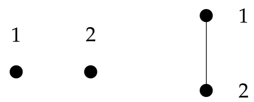
The direct product is associative and commutative.

The wreath product is associative but not commutative.

$$S_n \wr S_m \not\cong S_m \wr S_n \text{ if } n \neq m.$$

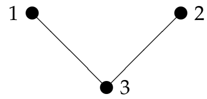
Partially ordered sets (posets)

crossing direct product $G_1 \times G_2$	nesting wreath product G_1/G_2



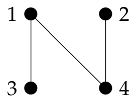
Iteration gives further posets, such as

$(G_1 \times G_2)/G_3$



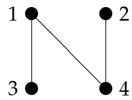
General posets

Iterated crossing and nesting gives series-parallel posets only.
It does not give posets such as



Can we start with the poset $\mathcal{P}$ with elements $1, \ldots, n$ and permutation groups $G_1, \ldots, G_n$ on $\Omega_1, \ldots, \Omega_n$ and then build the new permutation group on $\Omega_1 \times \cdots \times \Omega_n$?

That awkward poset: generalized wreath product



The elements of Ω are 4-tuples in $\Omega_1 \times \Omega_2 \times \Omega_3 \times \Omega_4$.
The permutations in the generalized wreath product of G_1, G_2, G_3 and G_4 with respect to this poset are all combinations of the following:

- permute values of the 1st coordinate by an element of G_1;
- permute values of the 2nd coordinate by an element of G_2;
- for each value of the 1st coordinate separately, permute values of the 3rd coordinate by an element of G_3;
- for each pair of values of the 1st and 2nd coordinates, permute values of the 4th coordinate by an element of G_4.

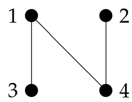
Generalized wreath product: orbits on pairs

If G is the generalized wreath product defined by a poset $\mathcal{P}$, then its orbits on pairs are as follows:

for each antichain $\mathcal{A}$ in $\mathcal{P}$
    for each i in $\mathcal{A}$
        for each non-diagonal orbit of G_i on Ω_i
            combine these with $\Omega_j \times \Omega_j$ if $j < k \in \mathcal{A}$
                $\mathrm{diag}(\Omega_j)$ otherwise.

If each G_i is 2-transitive then $\mathcal{A}$ gives a single orbit.

That example with $n = 4$



antichain	M_1	M_2	M_3	M_4
$\emptyset$	I	I	I	I
$\{1\}$	A	I	J	J
$\{2\}$	I	A	I	J
$\{3\}$	I	I	A	I
$\{4\}$	I	I	I	A
$\{1,2\}$	A	A	J	J
$\{2,3\}$	I	A	A	J
$\{3,4\}$	I	I	A	A

The orbital matrices of the generalized wreath product have the form $M_1 \otimes M_2 \otimes M_2 \otimes M_4$.

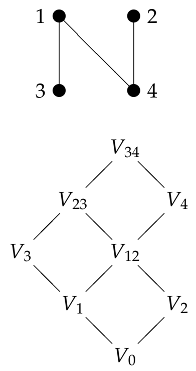
Here A denotes any non-identity orbital matrix of the relevant permutation group.

Generalized wreath product: partitions of Ω

Each antichain $\mathcal{A}$ of $\mathcal{P}$ defines a partition
(also called a block system)
of Ω preserved by the generalized wreath product G.
The parts (also known as blocks) are defined by
the values of coordinates j for which $j \geq k \in \mathcal{A}$.
This defines a corresponding vector subspace $V_{\mathcal{A}}$.

If each G_i is 2-transitive then the eigenspaces of the centralizer algebra are the subspaces $W_{\mathcal{A}}$, where, for each $\mathcal{A}$,
$W_{\mathcal{A}}$ is obtained from $V_{\mathcal{A}}$ by going orthogonal to all smaller subspaces defined by antichains.

That example with $n = 4$



$$
\begin{aligned}
W_{34} &= V_{24} \cap (V_{23} + V_4)^{\perp} \\
W_{23} &= V_{23} \cap (V_3 + V_{12})^{\perp} \\
W_4 &= V_4 \cap V_{12}^{\perp} \\
W_3 &= V_3 \cap V_1^{\perp} \\
W_{12} &= V_{12} \cap (V_1 + V_2)^{\perp} \\
W_1 &= V_1 \cap V_0^{\perp} \\
W_2 &= V_2 \cap V_0^{\perp} \\
W_0 &= V_0
\end{aligned}
$$

Some comments on history

R. A. Fisher initially advocated randomizing by choosing at random from among all layouts with a given property.

Frank Yates realised that it was sufficient for every pair of plots to have the same probability of receiving identical treatments (with appropriate modifications for blocks, rows, columns, ...).

Many people moved on to direct products and wreath products of symmetric groups.

J. A. Nelder (Proc. Roy. Soc A., 1965) formalized the iteration of these, but stated several results without proof.

O. Kempthorne, G. Zyskind, S. Addelman, T. N. Throckmorton and R. F. White (Aeronautical Research Laboratory, Technical Report, 1961) had the idea for generalized wreath products, but could not complete it because they did not know enough about permutation groups or posets.

RAB, Cheryl E. Praeger, C. A. Rowley and T. P. Speed (Proc. London Math. Soc., 1983) defined generalized wreath products, gave theory and proofs.

Back to first example: new topic

position ("plot")	1	2	3	4	5	6	7	8	9
variety ("treatment")	A	A	A	B	B	B	C	C	C

What should we do if a random permutation from S_9 gives us *AAABBBCCC*? Or *BCAAACCBB*?

Devil 1: I think that nearby plots are alike. If we use this layout and find differences between varieties, how can we know that it isn't just a difference between regions? Throw that layout away and re-randomize.

Devil 2: If you keep doing that, differences between regions will contribute more to the estimate of experimental error than they will to the estimates of differences between varieties, so you may fail to detect genuine differences between varieties.

Angel: Can we use a smaller 2-transitive subgroup G and a special initial layout that ensures that we never get a series of 3 adjacent plots with the same variety?

Restricted randomization for the first example

Let H be the elementary Abelian group of order 9.
The semi-direct product $H \rtimes \mathrm{Aut}(H)$ acts 2-transitively on H, and preserves the set of 4 partitions of H into cosets of a subgroup of order 3.

If we can arrange the elements of H in a line in such a way that none of these partitions has 3 consecutive elements in the same part, then we are done: use one partition as the initial layout, and randomize by using a random permutation from $H \rtimes \mathrm{Aut}(H)$.

00	01	20	22	10	12	02	11	21
A	A	C	C	B	B	A	B	C
D	E	D	F	D	F	F	E	E
G	H	H	G	I	H	I	G	I
J	K	L	K	K	J	L	L	J

Some comments on history

R. A. Fisher corresponded with "Student", O. Tedin and H. Jeffreys in the 1920s and 1930s about the bad consequences of simply throwing away randomized layouts with undesirable patterns. He explained this well in his 1935 book *Design of Experiments*.

This led to Frank Yates' concentration on pairs of plots.

P. M. Grundy and M. J. R. Healy realised that each symmetric group could be replaced by any 2-transitive subgroup:
then perhaps a good initial layout can be found.
They gave an example in J. Roy. Stat. Soc. B in 1950.

RAB was employed at the Agricultural Research Council Unit of Statistics because her DPhil thesis was about finite permutation groups. This led to a paper on restricted randomization in Biometrika in 1983.

Plenty of people still advocate using ordinary randomization and throwing away "bad layouts", apparently unaware of the advice from Fisher and Yates.

Second course

Abelian groups and design construction

A factorial experiment

In another experiment, treatments are all combinations of 3 varieties of tomato with 3 watering regimes. I have several greenhouses, all too small to contain all 9 combinations.

Label the varieties 0, 1, 2. Label the watering regimes 0, 1, 2. Identify $\mathcal{T}$ with the Abelian group

$$H = \langle a, b : a^3 = b^3 = 1, ab = ba \rangle;$$

here $a^i b^j$ is the combination of variety i with watering regime j.

The dual group, where $H = \langle a, b : a^3 = b^3 = 1, ab = ba \rangle$

A character of H is homomorphism from H to $(\mathbb{C}, \times)$.
The characters of H form the dual group H^*,
which is isomorphic to H.

	1	a	a^2	b	ab	a^2b	b^2	ab^2	a^2b^2
I	0	0	0	0	0	0	0	0	0
A	0	1	2	0	1	2	0	1	2
A^2	0	2	1	0	2	1	0	2	1
B	0	0	0	1	1	1	2	2	2
B^2	0	0	0	2	2	2	1	1	1
AB	0	1	2	1	2	0	2	0	1
A^2B^2	0	2	1	2	1	0	1	0	2
AB^2	0	1	2	2	1	0	1	2	0
A^2B	0	2	1	1	0	2	2	1	0

Entry s in this table should be interpreted as ω^s where $\omega = e^{2\pi i/3}$.

Main effects and interaction

	1	a	a^2	b	ab	a^2b	b^2	ab^2	a^2b^2
I	0	0	0	0	0	0	0	0	0
A	0	1	2	0	1	2	0	1	2
A^2	0	2	1	0	2	1	0	2	1
B	0	0	0	1	1	1	2	2	2
B^2	0	0	0	2	2	2	1	1	1
AB	0	1	2	1	2	0	2	0	1
A^2B^2	0	2	1	2	1	0	1	0	2
AB^2	0	1	2	2	1	0	1	2	0
A^2B	0	2	1	1	0	2	2	1	0

I is a constant vector in $\mathbb{C}^\Omega$.

A and A^2 are orthogonal to I; values depend only on the power of a (the variety): these are called the main effect of variety.

B and B^2 give the main effect of watering regime.

The other 4 characters are orthogonal to all of these; take all values equally often on each variety and each watering regime: these are the interaction between variety and watering regime.

Now get real

In practice, data are real numbers,
so we replace each pair $\chi, \bar{\chi}$ of complex vectors by
(suitable real multiples of) $\chi + \bar{\chi}$ and $i(\chi - \bar{\chi})$.

	1	a	a^2	b	ab	a^2b	b^2	ab^2	a^2b^2
I	1	1	1	1	1	1	1	1	1
A	1	ω	ω^2	1	ω	ω^2	1	ω	ω^2
A^2	1	ω^2	ω	1	ω^2	ω	1	ω^2	ω
$A + A^2$	2	-1	-1	2	-1	-1	2	-1	-1
$\frac{A-A^2}{i\sqrt{3}}$	0	1	-1	0	1	-1	0	1	-1

Applying $A + A^2$ to the data means for each of the 9 treatments enables us to estimate the difference between variety 0 and the average of the other two varieties.

But sticking with the characters makes the design process easier.

One scenario

There are 6 greenhouses, each with room for 3 plots.

1	a	a^2	1	a	a^2
a^2b	b	ab	ab	a^2b	b
ab^2	a^2b^2	b^2	a^2b^2	b^2	ab^2

$AB =$ 0 1 2 $\quad AB^2 =$ 0 1 2

we can estimate the effects of A, B and AB^2

we can estimate the effects of A, B and AB

The main effects A and B can be estimated with full efficiency.

The interaction effects AB and AB^2 have efficiency factor $1/2$, which means that the variance of their estimators is twice what it would be in an unblocked design of the same size.

Efficiency factors

Given an incomplete-block design in which all blocks have size k and all treatments occur r times, the $\mathcal{T} \times \mathcal{T}$ concurrence matrix Λ has (i,j)-entry equal to the number of blocks in which treatments i and j both occur, and the scaled information matrix is $I - (rk)^{-1}\Lambda$.

The constant vectors are in the kernel of the scaled information matrix.

The eigenvalues for the other eigenvectors are called canonical efficiency factors: the larger the better.

In the preceding example, A, A^2, B and B^2 have c.e.f. 1, while AB, A^2B^2, AB^2 and A^2B have c.e.f. $1/2$.

Helpful result

Theorem

Suppose that the set $\mathcal{T}$ of treatments for an incomplete-block design is identified with an Abelian group H.
If the (multi-)set of blocks is invariant under translation by H then

1. *the characters of H give a basis of eigenvectors of the scaled information matrix;*
2. *if χ is a character then χ and $\bar{\chi}$ have the same eigenvalue;*
3. *the eigenvalue for χ is the inner product of χ with the row of the scaled information matrix corresponding to treatment 1.*

Final scenario

There are 9 greenhouses, each with room for 4 plots.

a	a^2	1	ab	a^2b	b	ab^2	a^2b^2	b^2
a^2	1	a	a^2b	b	ab	a^2b^2	b^2	ab^2
b	ab	a^2b	b^2	ab^2	a^2b^2	1	a	a^2
b^2	ab^2	a^2b^2	1	a	a^2	b	ab	a^2b

$$\Lambda = \begin{array}{c} \\ 1 \\ a \\ \vdots \end{array} \begin{array}{ccccccccc} 1 & a & a^2 & b & ab & a^2b & b^2 & ab^2 & a^2b^2 \\ 4 & 1 & 1 & 1 & 2 & 2 & 1 & 2 & 2 \\ 1 & 4 & 1 & 2 & 1 & 2 & 2 & 1 & 2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

$A = (\ 1\ \ \omega\ \ \omega^2\ \ 1\ \ \omega\ \ \omega^2\ \ 1\ \ \omega\ \ \omega^2\ )$
eigenvalue of Λ is 1; eigenvalue of $I - \Lambda/16$ is $15/16$.

$AB = (\ 1\ \ \omega\ \ \omega^2\ \ \omega\ \ \omega^2\ \ 1\ \ \omega^2\ \ 1\ \ \omega\ )$
eigenvalue of Λ is 4; eigenvalue of $I - \Lambda/16$ is $3/4$.

$\text{c.e.f.}(B) = \text{c.e.f.}(A) = 15/16$; $\text{c.e.f.}(AB^2) = \text{c.e.f.}(AB) = 3/4$.

Some comments on history

R. A. Fisher introduced the use of elementary Abelian groups for factorial designs in Annals of Eugenics in 1942 and 1945.

D. J. Finney extended this to fractional factorial designs with many factors in 1945. Only a subgroup of H is used. If this is well chosen, and interactions among large numbers of factors can be assumed to be zero, then estimation is still possible.

R. C. Bose and K. Kishen had already used finite Eucliean geometry in Sankhyā in 1940. R. C. Bose generalized both, explaining more details, in Sankhyā in 1947. His approach using finite fields became the paradigm.

RAB showed that Fisher's method extends to arbitrary finite Abelian groups in Linear Algebra and its Applications in 1985.

Today, very few statisticians know any group theory.

Who needs theory now that we can find a design by computer search and analyse the data with standard software?

Who needs statisticians now that we all have computers?