Chapter 12

Factorial Designs in Incomplete Blocks

12.1 Confounding

Definition A treatment subspace estimated in stratum F is *confounded* with F.

Definition Factors F and G are *strictly orthogonal* to each other if they are orthogonal to each other in the sense of Section 10.5 and $F \wedge G = U$.

If F is strictly orthogonal to G then $V_F \cap V_0^\perp$ is orthogonal to $V_G \cap V_0^\perp$ so all contrasts in F are orthogonal to all contrasts in G.

Example 12.1 (Example 1.5 continued: Rye-grass) The treatments are all combinations of cultivar with nitrogen. These two factors are strictly orthogonal to each other. The small blocks are strips, and cultivars are confounded with strips.

Example 12.2 (Example 10.7 continued: Main-effects only design in blocks) The treatments are all combinations of levels of factors F and G. Blocks are incomplete, and we use a Latin square to ensure that F and G are both orthogonal to blocks and that F is strictly orthogonal to G.

In this chapter we develop a more general way of constructing incomplete-block designs for factorial experiments, which includes both of the above. We want to ensure that the single-replicate designs are orthogonal, in the sense of Section 10.11. We also want to identify which treatment subspace, if any, is confounded with blocks. In Figure 10.27 we used a factor Q with no real physical meaning in order to identify this confounding. This idea is generalized in this chapter.

Example 12.3 (Watering chicory) Chicory is grown in boxes in sheds. Each box is watered by a continuous flow of water from a pipe. The experimenter can vary two factors which affect this watering. One is the rate of flow of the water; this can be adjusted to one of three rates by opening the tap on the input pipe. The second factor is the depth of standing water which is permitted in the box; this can be adjusted to have one of three values by putting side bars of appropriate heights around the box.

Two sheds will be used for the experiment. Each shed contains three rooms, each of which contains three boxes of chicory. The boxes are the plots. At the end of the experiment the chociry is harvested and the saleable chicory per box is weighed.

We want each shed to be a single replicate of the treatments. How should treatments be applied?

(i) We could use a split-plot design, applying flowrate to whole rooms and depth to individual boxes. Then the main effect of flowrate is confounded with rooms, and so it is estimated less precisely than the main effect of depth or the flowrate-by-depth interaction. If rates of flow cannot be varied within a room then we must use such a design; otherwise, it is not a good choice unless we are less interested in the main effect of flowrate than in its interaction with depth.

(ii) We could confound flowrate with rooms in the first shed and confound depth with rooms in the second shed. Thus each main effect will be estimated more precisely from just *half* of the plots; in other words, each main effect has *efficiency factor* equal to $1/2$, in the sense of Section 11.6. The interaction is estimated better than the two main effects. Unless we are chiefly interested in the interaction, this is not a good design.

(iii) As in Section 9.1.3, we could use a Latin square to construct a main-effects-only design in each shed. Now part of the flowrate-by-depth interaction is confounded with rooms.

(iv) As there are four degrees of freedom for the flowrate-by-depth interaction, and only two are confounded with rooms in any one shed, it is better if we can arrange for different parts of the interaction to be confounded in the two sheds. Thus we need to know how to decompose interactions.

12.2 Decomposing interactions

Let F_1, F_2, ..., F_n be treatment factors with p values each, where p is a prime number. Suppose that treatments are all combinations of the levels of $F_1, F_2, \ldots, F_n$. Code the values of each factor by $\mathbb{Z}_p$, the integers modulo p. Then we can regard each treatment as being an n-tuple of integers modulo p.

Definition A *character* of the treatments is a function from $\mathcal{T}$ to $\mathbb{Z}_p$ which is a linear combination of the F_i with coefficients in $\mathbb{Z}_p$.

In particular, each treatment factor F_i is itself a character.

Example 12.3 revisited (Watering chicory) Here we have $p = 3$ and $n = 2$. Let A be the factor flowrate, with the three rates coded as 0, 1, 2 in any order. Similarly, let B be the factor depth, with the three depths coded as 0, 1, 2 in any order. Then treatments are ordered pairs of integers modulo 3: the first entry is the level of A and the second is the level of B. These are shown above the line in Table 12.1. The remaining characters are below the line. Each character takes a value on each treatment. For example

$$(A+2B)((2,1)) = 1\times 2+2\times 1 = 2+2 = 1.$$

(Remember that the values are all in $\mathbb{Z}_3$, so all arithmetic must be done modulo 3.) It is conventional to write I for the character whose coefficients are all zero.

characters	treatments								
A	0	0	0	1	1	1	2	2	2
B	0	1	2	0	1	2	0	1	2
$A+B$	0	1	2	1	2	0	2	0	1
$A+2B$	0	2	1	1	0	2	2	1	0
$2A+B$	0	1	2	2	0	1	1	2	0
$2A+2B$	0	2	1	2	1	0	1	0	2
$2A$	0	0	0	2	2	2	1	1	1
$2B$	0	2	1	0	2	1	0	2	1
I	0	0	0	0	0	0	0	0	0

Table 12.1: Characters in Example 12.3

Each character can be regarded as a factor. The table has some noteworthy features.

(i) There are nine treatments and nine characters.

(ii) The factor A is aliased with the factor $2A$; that is, $A \equiv 2A$.

(iii) $I \equiv U$.

(iv) The factor $A+B$ has three values, and $A+B \prec U$, so there are two degrees of freedom for contrasts between levels of $A+B$.

(v) The factor $A+B$ is strictly orthogonal to both A and B; so the contrasts between levels of $A+B$ are orthogonal to those for the main effects of both A and B. But $A \wedge B \prec A+B$ and so the two degrees of freedom for $A+B$ are part of the A-by-B interaction.

(vi) $2A+2B \equiv A+B$.

(vii) The factor $A+2B$ also accounts for two degrees of freedom from the A-by-B interaction, by similar reasoning to that used for $A+B$.

(viii) The factors $A+B$ and $A+2B$ are strictly orthogonal to each other, so the 4-dimensional subspace of treatment contrasts belonging to the A-by-B interaction is the orthogonal sum of two 2-dimensional subspaces, one corresponding to the factor $A+B$ and the other corresponding to the factor $A+2B$.

(ix) $2A+B = 2(A+2B)$ and $2A+B \equiv A+2B$.

If we use the characters $A+B$ and $A+2B$ to decompose the interaction, we obtain the Hasse diagram in Figure 12.1.

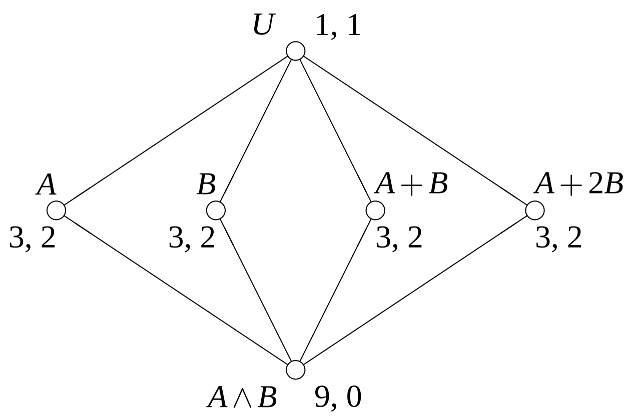

Figure 12.1: Hasse diagram for treatment structure in Example 12.3

Although we shall not prove it here, the above pattern extends to any number of factors and any prime p, as the following theorem states. In fact, we shall not prove any of the theorems in this chapter, but the reader may easily verify them on any given example.

Theorem 12.1 *Let $F_1, F_2, \ldots, F_n$ each have p values, where p is prime. Then the following hold.*

(i) There are p^n treatments and p^n characters.

(ii) If $G = \sum_i g_i F_i$ and $H = \sum_i h_i F_i$ and $G \neq I$ and $H \neq I$ then either

(a) there is some non-zero k in $\mathbb{Z}_p$ such that $g_i = kh_i$ for $i = 1, \ldots, n$, in which case $G \equiv H$; or

(b) as factors, G and H are strictly orthogonal.

Hence the $(p^n - 1)$ characters different from I split into $(p^n - 1)/(p - 1)$ sets of $(p - 1)$ characters, each set strictly orthogonal to all the others, with all the characters in any one set being equivalent as factors.

(iii) If $G = \sum g_i F_i$ *then* G *belongs to the interaction of those* F_i *for which* $g_i \neq 0$.

Example 12.4 (Two factors with five levels) If A and B both have five levels then the 24 characters apart from I fall into six sets of four as follows.

$A \equiv 2A \equiv 3A \equiv 4A$	main effect of A
$B \equiv 2B \equiv 3B \equiv 4B$	main effect of B
$A+B \equiv 2A+2B \equiv 3A+3B \equiv 4A+4B$	part of the A-by-B-interaction
$A+2B \equiv 2A+4B \equiv 3A+B \equiv 4A+3B$	part of the A-by-B-interaction
$A+3B \equiv 2A+B \equiv 3A+4B \equiv 4A+2B$	part of the A-by-B-interaction
$A+4B \equiv 2A+3B \equiv 3A+2B \equiv 4A+B$	part of the A-by-B-interaction

Thus the interaction is the orthogonal sum of four parts, each with four degrees of freedom. These parts are defined by the characters

$$A+B, \quad A+2B, \quad A+3B, \quad \text{and} \quad A+4B.$$

This representation is not unique. For example, the second part is equally well defined by the characters $2A+4B$ (which can be written as $2A-B$), $3A+B$ or $4A+3B$. It is conventional to use the character whose first non-zero coefficient is equal to 1.

Warning: many people write the characters multiplicatively. For example, they write AB^2 in place of $A+2B$. Then AB may mean $A+B$ (the character) or $A \wedge B$ (the factor) or A-by-B (the interaction).

Example 12.5 (Three factors with three levels) If there are three treatment factors A, B and C, each with three levels, then the 26 characters other than I fall into 13 sets of two as shown in Table 12.2. We have already seen how to decompose each of the two-factor interactions. There are eight degrees of freedom for the A-by-B-by-C interaction. This is the 8-dimensional space of contrasts in $A \wedge B \wedge C$ which sum to zero on each class of each of the factors $A \wedge B$, $A \wedge C$ and $B \wedge C$. By using characters we can decompose this interaction into four orthogonal subspaces, each of dimension 2. The subspaces correspond to

$$A+B+C, \quad A+B+2C, \quad A+2B+C, \quad \text{and} \quad A+2B+2C.$$

12.3 Constructing designs with specified confounding

If G is a character and we want to construct a block design in which G is confounded with blocks, the obvious method is to evaluate G on every treatment, then put into the first block all those treatments u with $G(u) = 0$, into the second block all those treatments u with $G(u) = 1$, and so on.

A	$\equiv$	$2A$	main effect of A
B	$\equiv$	$2B$	main effect of B
C	$\equiv$	$2C$	main effect of C
$A+B$	$\equiv$	$2A+2B$	part of the A-by-B interaction
$A+2B$	$\equiv$	$2A+B$	part of the A-by-B interaction
$A+C$	$\equiv$	$2A+2C$	part of the A-by-C interaction
$A+2C$	$\equiv$	$2A+C$	part of the A-by-C interaction
$B+C$	$\equiv$	$2B+2C$	part of the B-by-C interaction
$B+2C$	$\equiv$	$2B+C$	part of the B-by-C interaction
$A+B+C$	$\equiv$	$2A+2B+2C$	part of the A-by-B-by-C interaction
$A+B+2C$	$\equiv$	$2A+2B+C$	part of the A-by-B-by-C interaction
$A+2B+C$	$\equiv$	$2A+B+2C$	part of the A-by-B-by-C interaction
$A+2B+2C$	$\equiv$	$2A+B+C$	part of the A-by-B-by-C interaction

Table 12.2: Characters in Example 12.5

	First Shed									Second Shed								
	Room 1			Room 2			Room 3			Room 4			Room 5			Room 6		
	$A+B=0$			$A+B=1$			$A+B=2$			$A+2B=0$			$A+2B=1$			$A+2B=2$		
A	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
B	0	2	1	1	0	2	2	1	0	0	1	2	2	0	1	1	2	0

Table 12.3: Design confounding $A+B$ in one shed and confounding $A+2B$ in the other

Example 12.3 revisited (Watering chicory) The values of $A+B$ and $A+2B$ have already been calculated in Table 12.1. Confounding $A+B$ with rooms in the first shed and $A+2B$ with rooms in the second shed gives the design in Table 12.3.

However, there is a simpler method, that does not demand evaluation of the characters.

Each treatment u can be written as as n-tuple of integers modulo p, as $u = (u_1, u_2, \ldots, u_n)$, where $F_i(u) = u_i$. Then treatments can be added by the rule

$$u+v = (u_1+v_1, u_2+v_2, \ldots, u_n+v_n).$$

Equivalently, $F_i(u+v) = F_i(u) + F_i(v)$ for each i. Then the set $\mathcal{T}$ of treatments forms an *abelian group*. This means that:

(i) for all u and v in $\mathcal{T}$, the sum $u+v$ is also in $\mathcal{T}$;

(ii) for all u and v in $\mathcal{T}$, we have $u+v = v+u$;

(iii) $(u+v)+w = u+(v+w)$ for all u, v, w in $\mathcal{T}$;

(iv) there is a treatment $(0,0,\ldots,0)$ which satisfies

$$u+(0,0,\ldots,0)=u \quad \text{for all } u;$$

(v) for each u in $\mathcal{T}$ there is a treatment $(-u_1,-u_2,\ldots,-u_n)$, which we write as $-u$, such that

$$u+(-u)=(0,0,\ldots,0)$$

(for example, if $p=5$ and $u_1=2$ then $-u_1=3$).

Definition In a single-replicate design, the block containing the treatment $(0,0,\ldots,0)$ is the *principal* block.

Theorem 12.2 *The principal block is a subgroup of $\mathcal{T}$ (this means that if u and v are in the principal block then so is $u+v$). Every block is a coset of the principal block, in other words a subset of the form*

$$\{v+u : u \in \textit{principal block}\}$$

for some fixed v. Thus, to construct the block containing v, simply add v to every element of the principal block.

Example 12.5 revisited (Three factors with three levels) To confound $A+B+2C$ (which is part of the three-factor interaction) with blocks, the principal block consists of those treatments u for which

$$A(u)+B(u)+2C(u)=0.$$

This equation can be rewritten as

$$C(u)=A(u)+B(u),$$

because $2=-1$ modulo 3. Thus every pair of values of A and B uniquely determines a value of C so that the above equation is satisfied. The principal block therefore contains the following nine treatments:

$$\begin{array}{llllll}(0,0,0), & (0,1,1), & (0,2,2), & (1,0,1), & (1,1,2), & (1,2,0),\\ (2,0,2), & (2,1,0), & (2,2,1). & & & \end{array}$$

The treatment $(1,1,1)$ does not appear in the principal block. To obtain the block containing it we add $(1,1,1)$ to every treatment in the principal block. This gives a second block with the following nine treatments.

$$\begin{array}{llllll}(1,1,1), & (1,2,2), & (1,0,0), & (2,1,2), & (2,2,0), & (2,0,1),\\ (0,1,0), & (0,2,1), & (0,0,2). & & & \end{array}$$

The block containing $(2,2,2)$ is constructed similarly.

Since each block contains nine of the 27 treatments, it would be possible to construct the third block as simply ‘all those treatments which are not in the first two blocks’. This method gives no check on mistakes. It is better to construct every block as a coset of the principal block, and then check that every treatment appears in one and only one block.

Example 12.6 (Sugar beet) An experiment on sugar beet investigated three three-level treatment factors, whose real levels were coded by the integers modulo 3 as follows.

	0	1	2
Sowing date (D)	18 April	9 May	25 May
Spacing bewteen rows (S)	10 inches	15 inches	20 inches
Nitrogen fertilizer (N)	nil	0.3 cwt/acre	0.6 cwt/acre

There were three blocks of nine plots each. These were constructed to confound $D+S+2N$ with blocks. Thus the allocation of treatments to blocks can be deduced from Example 12.5. Table 12.4 shows the layout after randomization.

Example 12.7 (Four factors with two levels) Suppose that A, B, C and D each have two levels. To confound $A+B+C+D$ (which is the whole of the four-factor interaction) with blocks, the principal block contains those treatments u for which

$$A(u)+B(u)+C(u)+D(u)=0.$$

Solving such equations is relatively easy when $p=2$, because the only possible values are 0 and 1. Thus the principal block consists of all those treatments for which an *even* number of A, B, C, D take the value 1.

$$(0,0,0,0),\quad (0,0,1,1),\quad (0,1,0,1),\quad (0,1,1,0),\quad (1,0,0,1),\quad (1,0,1,0),$$
$$(1,1,0,0),\quad (1,1,1,1)$$

When the context is clear, we sometimes miss out the commas and the parentheses in the notation for factorial treatments. In this abbreviated notation, the principal block in Example 12.7 contains the following treatments.

$$0000,\quad 0011,\quad 0101,\quad 0110,\quad 1001,\quad 1010,\quad 1100,\quad 1111$$

Example 12.8 (Pill manufacture) Medicinal tablets are composed largely of inactive ingredients which serve to hold the active ingredients and release them into the body. In the first stage of making the tablets, the inactive ingredients are mixed. A certain volume of ingredients is put into a blender. They are blended for a length of time. Then the mixed ingredients are milled to a fine powder. Finally they are blended again. In a simple experiement each of these factors has two levels: the volume is small or large; each blending time is short or long; the milling speed is slow or fast. Only eight batches can be made up per day. The design in Example 12.7 can be used to confound the four-factor interaction with days.

12.4 Confounding more than one character

Example 12.7 revisited (Four factors with two levels) Suppose that we have four blocks of four plots (instead of two blocks of eight plots). Each character partitions the treatments into two sets of eight. Since linearly independent characters are

Block 1	$D+S+2N=1$		
Plot	Date	Spacing	Nitrogen
1	25 May	15in	0.6
2	9 May	15in	0.3
3	9 May	10in	0
4	25 May	20in	0
5	9 May	20in	0.6
6	18 April	15in	0
7	25 May	10in	0.3
8	18 April	20in	0.3
9	18 April	10in	0.6

Block 2	$D+S+2N=0$		
Plot	Date	Spacing	Nitrogen
10	18 April	15in	0.3
11	18 April	10in	0
12	25 May	10in	0.6
13	9 May	10in	0.3
14	9 May	15in	0.6
15	25 May	20in	0.3
16	25 May	15in	0
17	9 May	20in	0
18	18 April	20in	0.6

Block 3	$D+S+2N=2$		
Plot	Date	Spacing	Nitrogen
19	18 April	15in	0.6
20	25 May	15in	0.3
21	9 May	10in	0.6
22	9 May	20in	0.3
23	25 May	10in	0
24	18 April	10in	0.3
25	9 May	15in	0
26	25 May	20in	0.6
27	18 April	20in	0

Table 12.4: Layout of the sugar-beet experiment in Example 12.6

strictly orthogonal to each other, any pair of linearly independent characters partitions the treatments into four sets of four. We could confound $A+B+C+D$ and $A+B+C$ with blocks. Then

$$\begin{array}{ll} A+B+C+D & \text{is constant on each block} \\ A+B+C & \text{is constant on each block;} \end{array}$$

therefore D is also constant on each block and so the main effect of D is confounded with blocks. It might be better to confound $A+B+C$, $B+C+D$ and $A+D$.

Theorem 12.3 *If characters G and H are confounded with blocks then so is $G+H$. In fact the set of confounded characters, together with I, forms a subgroup of the group of all characters.*

Example 12.7 revisited (Four factors with two levels) If we confound $A+B+C$ and $B+C+D$ then the principal block is defined by

$$A+B+C = B+C+D = 0,$$

so

$$A = -(B+C) = B+C = D.$$

The design is shown in Table 12.5, using the abbreviated notation for the treatments.

Block	Treatments				Value of	
					$A+B+C$	$B+C+D$
Principal	0000	1011	1101	0110	0	0
Block 2	1000	0011	0101	1110	1	0
Block 3	0100	1111	1001	0010	1	1
Block 4	0001	1010	1100	0111	0	1

Table 12.5: Design for four two-level treatment factors in four blocks of four plots

The procedures of this section and the preceding one are summarized in Technique 12.4.

Technique 12.4 (Single-replicate factorial designs in incomplete blocks) Given n treatment factors, each with p levels, construct a single-replicate block design with p^s blocks of p^{n-s} plots as follows.

(i) Find a subgroup $\mathcal{G}$ of the set of all characters such that

   (a) if any main effect has to be applied to whole blocks then the corresponding character is in $\mathcal{G}$;

   (b) if a character G belongs to an effect that must be estimated precisely then $G \notin \mathcal{G}$;

(c) $|\mathcal{G}| = p^s$.

(ii) Form the principal block as those p^{n-s} treatments u satisfying

$$G(u) = 0 \quad \text{for all } G \text{ in } \mathcal{G}.$$

(iii) The remaining blocks are cosets of the principal block.

Note that Step (i) requires trial-and-error, and may be impossible. Step (ii) is straightforward, and Step (iii) is entirely automatic.

Example 12.9 (Field beans) In an experiment on field beans, there are five two-level factors, whose levels are coded by the integers modulo 2 as follows.

	0	1
Row spacing (S)	18 inches	24 inches
Dung (D)	nil	10 ton/acre
Nitrochalk (N)	nil	0.4 cwt/acre
Superphosphate (P)	nil	0.6 cwt/acre
Potash (K)	nil	1 cwt/acre

The experiment is a single replicate, in four blocks of eight plots each. Thus three characters of the form G, H and $G+H$ must be confounded. Because we are working modulo 2, if any two of these characters have an odd number of letters then the third has an even number of letters. Thus at least one of them has an even number of letters, and we should choose this number to be four, to avoid confounding any two-factor interaction. If we confound the five-letter character $S+D+N+P+K$ and *any* four-letter character then we also confound a main effect, which is probably not desirable. Thus the best we can do is to confound one four-letter character and two three-letter characters. For example, we can confound $D+N+P+K$, $S+D+P$ and $S+N+K$. This gives the design in Table 12.6.

Block	Treatments factors in order S, D, N, P, K							
Prinicpal	00000	00101	01010	01111	10011	10110	11001	11100
Block 2	10000	10101	11010	11111	00011	00110	01001	01100
Block 3	01000	01101	00010	00111	11011	11110	10001	10100
Block 4	11000	11101	10010	10111	01011	01110	00001	00100

Table 12.6: Design for the experiment on field beans in Example 12.9

12.5 Pseudofactors for mixed numbers of levels

Treatment factors with different numbers of levels can be accommodated together so long as their numbers of levels are all powers of the same prime, say p. A factor with p^m levels is represented by m pseudofactors, each with p levels.

Example 12.10 (A mixed factorial) Factors A, B and C have two levels each, and factor D has the four levels 1, 2, 3, 4. One possible correspondence between D and its pseudofactors D_1 and D_2 is

D	1	2	3	4
D_1	0	0	1	1
D_2	0	1	0	1.

Suppose that we want a single-replicate design in four blocks of eight plots each which confounds no main effect or two-factor interaction with blocks. We need to confound two different characters and their sum, and all three confounded characters must involve at least three factors. One possible choice is to confound

$A+B+D_1$	part of the A-by-B-by-D interaction
$A+C+D_2$	part of the A-by-C-by-D interaction
$B+C+D_1+D_2$	part of the B-by-C-by-D interaction.

The principal block is constructed by putting $D_1 = A+B$ and $D_2 = A+C$. The other three blocks are constructed as cosets of the principal block. Finally, the ordered pairs of values of D_1 and D_2 are translated back into values of D. The first two blocks, showing values of both pseudofactors and genuine factors, are in Table 12.7.

	Principal block								Block 2							
A	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
B	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
C	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
D_1	0	0	1	1	1	1	0	0	1	1	0	0	0	0	1	1
D_2	0	1	0	1	1	0	1	0	0	1	0	1	1	0	1	0
D	1	2	3	4	4	3	2	1	3	4	1	2	2	1	4	3

Table 12.7: Two blocks of the design on Example 12.10

We might use such a design if we can assume that the interactions A-by-B-by-D, A-by-C-by-D, B-by-C-by-D and A-by-B-by-C-by-D are all zero. The skeleton analysis-of-variance table is in Table 12.8.

Notice that the assumption that the A-by-B-by-D interaction is zero is equivalent to the assumption that there are no differences caused by the different values of the factor $A \wedge B \wedge D$ *apart from* those differences which are already accounted for by

Stratum	source	degrees of freedom
mean	mean	1
blocks		3
plots	A	1
	B	1
	C	1
	D	3
	$A \wedge B$	1
	$A \wedge C$	1
	$B \wedge C$	1
	$A \wedge D$	3
	$B \wedge D$	3
	$C \wedge D$	3
	$A \wedge B \wedge C$	1
	residual	9
	total	28
Total		32

Table 12.8: Skeleton analysis of variance in Example 12.10

treatment factors coarser than $A \wedge B \wedge D$ (that is, A, B, D, $A \wedge B$, $A \wedge D$ and $B \wedge D$). But $A \wedge B \wedge D$ is itself coarser than $A \wedge B \wedge C \wedge D$, so it would be nonsense to assume that the latter is non-zero while the former is zero. The general form of this restriction is in Principle 12.5.

Principle 12.5 (Ordering of treatment factors) If F and G are treatment factors with $F \prec G$ and the effect of F is assumed to be non-zero then the effect of G should also be assumed to be non-zero.

12.6 Analysis of single-replicate designs

There are three ways to analyse data from a single-replicate experiment in blocks. The first method was implied in Section 9.1.3. We assume that some treatment effects are zero (in accord with Principle 12.5) and then analyse in the usual way.

Example 12.6 revisited (Sugar beet) If we can assume that the three-factor interaction is zero then we obtain the skeleton analysis-of-variance table in Table 12.9. All main effects and two-factor interactions can be estimated. Their presence can be tested for, using the residual mean square in the plots stratum. That mean square can also be used to find estimates of the variances of those estimators.

On the other hand, if we cannot assume that any treatment effects are zero then we can still estimate effects and obtain the full analysis-of-variance table, but there is no residual line in any stratum.

Stratum	source	degrees of freedom
mean	mean	1
blocks	blocks	2
plots	D	2
	S	2
	N	2
	$D \wedge S$	4
	$D \wedge N$	4
	$S \wedge N$	4
	residual	6
	total	24
Total		27

Table 12.9: Skeleton analysis of variance for Example 12.6 if we can assume that the three-factor interaction is zero

Example 12.6 revisited (Sugar beet) If we cannot assume that the three-factor interaction is zero then we obtain the skeleton analysis-of-variance in Table 12.10. We can still estimate the main effects and two-factor interactions, but there are no residual degrees of freedom in the plots stratum so we cannot carry out hypothesis tests or estimate variances. We cannot realistically estimate the three-factor interaction, because we expect the blocks stratum variance to be high.

Stratum	source	degrees of freedom
mean	mean	1
blocks	$D+S+2N$ (part of D-by-S-by-N interaction)	2
plots	D	2
	S	2
	N	2
	$D \wedge S$	4
	$D \wedge N$	4
	$S \wedge N$	4
	rest of D-by-S-by-N interaction	6
	total	24
Total		27

Table 12.10: Skeleton analysis of variance for Example 12.6 if we cannot assume that the three-factor interaction is zero

In this situation, where we cannot assume in advance that any effects are zero, a graphical method is helpful. We assume the usual normal model, whether blocks are fixed or random, with plots stratum variance ξ. If all effects in the plots stratum are zero, then the sums of squares for the p-valued characters in the plots stratum are independent random variables, each of which is ξ times a χ^2 random variable on $p-1$ degrees of freedom, by Theorem 2.10(vi). There are m such characters, where $m = (p^n - p^s)/(p-1)$. Let $S_1, \ldots, S_m$ be these sums of squares, in increasing order of magnitude. For $i = 1, \ldots, m$, let Q_i be the $(i-\frac{1}{2})/m$-th quantile of the χ^2 distribution on $p-1$ degrees of freedom; that is, $\Pr[X \leq Q_i] = (i-\frac{1}{2})/m$, where X is a χ^2 random variable with $p-1$ degrees of freedom. If these sums of squares are not inflated by any non-zero treatment effects, than the graph of S_i against Q_i should be a straight line through the origin with slope ξ. Typically, such a graph shows most points on a line through the origin, with a few points at the top end lying above the line. These should be removed, m reduced accordingly, and the graph redrawn. When sufficient points have been removed for the graph to be (approximately) a straight line through the origin, the conclusion is that the effects corresponding to the removed points are non-zero, and should be investigated further. Of course, Principle 12.5 has to be observed: do not remove a two-letter character without removing the single letters in it; and do not remove part of an interaction without removing all of it.

Example 12.6 revisited (Sugar beet) Table 12.11 shows the yields of beet, in cwt/acre, for the experiment in Table 12.4. There is a column for each character not confounded with blocks, showing the total yield on each value of the character, and hence the sum of squares for the character. These are the initial values of $S_1, \ldots, S_{12}$. Since the χ^2 distribution with two degrees of freedom is just the exponential distribution with mean 2, the quantiles are readily calculated as $Q_i = -2\ln((25-2i)/24)$. Figure 12.2 shows the graph. It is clear that the main effects of D and N are non-zero, and possibly also the main effect of S.

In fact these data are only half of a two-replicate design described in Section 12.7. The full analysis of all the data does confirm that only the three main effects are needed to model the yields.

If $p > 2$ then every interaction consists of more than one character. The sums of squares for each interaction should be separated into those for the individual characters to draw the graph like the one in Figure 12.2. However, for the other two methods of analysis the sums of squares for the individual characters should be pooled into sums of squares for the various interactions. For example, in an analysis-of-variance based on Table 12.9 or Table 12.10, the sum of squares for the D-by-S interaction should be reported as 54.87 on four degrees of freedom.

12.7 Several replicates

Confounding can be used in factorial block designs with more than one replicate. Typically we use Technique 12.4 in each replicate, but confound different characters in each replicate, unless there is any factor that cannot be applied to smaller units than blocks.

Plot	yield	D	S	N	$D+S$	$D+2S$	$D+N$	$D+2N$	$S+N$	$S+2N$	$D+S+N$	$D+2S+N$	$D+2S+2N$
1	52.2	2	1	2	0	1	1	0	0	2	2	0	2
2	52.7	1	1	1	2	0	2	0	2	0	0	1	2
3	47.8	1	0	0	1	1	1	1	0	0	1	1	1
4	35.2	2	2	0	1	0	2	2	2	2	1	0	0
5	45.4	1	2	2	0	2	0	2	1	0	2	1	0
6	44.6	0	1	0	1	2	0	0	1	1	1	2	2
7	46.0	2	0	1	2	2	0	1	1	2	0	0	1
8	51.4	0	2	1	2	1	1	2	0	1	0	2	0
9	50.5	0	0	2	0	0	2	1	2	1	2	2	1
10	49.7	0	1	1	1	2	1	2	2	0	2	0	1
11	47.8	0	0	0	0	0	0	0	0	0	0	0	0
12	44.1	2	0	2	2	2	1	0	2	1	1	1	0
13	52.5	1	0	1	1	1	2	0	1	2	2	2	0
14	49.3	1	1	2	2	0	0	2	0	2	1	2	1
15	46.2	2	2	1	1	0	0	1	0	1	2	1	2
16	47.1	2	1	0	0	1	2	2	1	1	0	1	1
17	47.2	1	2	0	0	2	1	1	2	2	0	2	2
18	56.0	0	2	2	2	1	2	1	1	0	1	0	2
19	50.9	0	1	2	1	2	2	1	0	2	0	1	0
20	38.2	2	1	1	0	1	0	1	2	0	1	2	0
21	43.0	1	0	2	1	1	0	2	2	1	0	0	2
22	36.5	1	2	1	0	2	2	0	0	1	1	0	1
23	38.0	2	0	0	2	2	2	2	0	0	2	2	2
24	45.7	0	0	1	0	0	1	2	1	2	1	1	2
25	37.1	1	1	0	2	0	1	1	1	1	2	0	0
26	34.2	2	2	2	1	0	1	0	1	0	0	2	1
27	35.4	0	2	0	2	1	0	0	2	2	2	1	1
Total on 0		432.0	415.4	380.2	410.6	398.7	395.9	400.0	420.1	409.8	420.3	403.5	402.6
Total on 1		411.5	421.8	418.9	404.1	423.6	409.4	419.9	408.6	400.5	397.4	415.3	396.5
Total on 2		381.2	387.5	425.6	410.0	402.4	419.4	404.8	396.0	414.4	407.0	405.9	425.6
SS		145.15	73.92	133.47	2.87	40.12	30.91	23.96	32.29	11.14	29.29	8.64	52.33

Table 12.11: Yields of beet, and calculations of sums of squares for the characters, for the experiment in Table 12.4

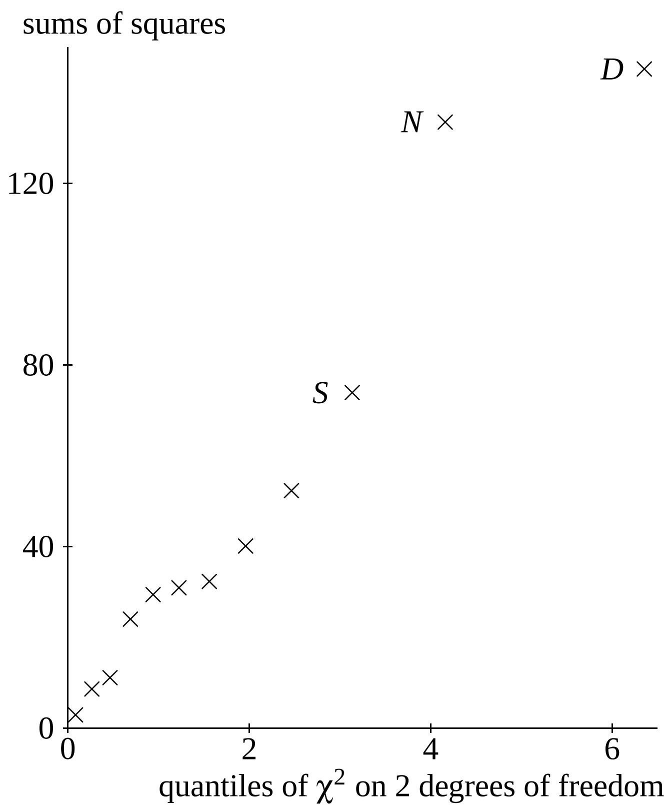

Figure 12.2: Plot of sums of squares from Table 12.11

The argument in Section 11.7 shows that if there are r replicates and a character is confounded with blocks in q of those replicates then the efficiency factor for that character is $(r-q)/r$. If we can arrange for every character to be confounded in the same number of replicates then the design is balanced.

Often we choose to confound higher-order interactions, believing them to be less important. However, sometimes we already know about main effects and want to know about the interactions: then we might choose to confound main effects more.

Example 12.3 revisited (Watering chicory) The design in Table 12.3 has two replicates. It confounds part of the interaction in one replicate and the other part of the interaction in the other replicate. Thus both main effects have full efficiency while the interaction has efficiency factor $1/2$.

Example 12.6 revisited (Sugar beet) In fact, this experiment also had two replicates. One is shown in Table 12.4, confounding $D+S+2N$. There was also a second replicate, confounding $D+S+N$.

Example 12.11 (Balanced interactions) Suppose that there are three two-level treatment factors A, B and C, and the experimental material consists of eight blocks of four plots each. We can divide these into four replicates, and confound each of $A+B+C$, $A+B$, $A+C$ and $B+C$ in one replicate. Then there is full efficiency on all main effects, and all of the interactions have efficiency factor $3/4$ because they can be estimated from three of the four replicates.

Questions for Discussion

12.1 Construct a resolved design for three three-level treatments factors in nine blocks of size nine in such a way that no main effect or two-factor interaction is confounded in any replicate and that no treatment contrast is totally confounded with blocks.

12.2 Construct a single-replicate design for six two-level treatment factors in four blocks of 16 plots each, assuming that all interactions involving four or more factors are zero. Write down the skeleton analysis-of-variance table, showing stratum, source and degrees of freedom.

12.3 There are three treatment factors, A, B and C, each with three levels. There are 54 plots, divided into 18 blocks of size three. Suppose that factor A must be confounded with blocks and that a resolved design is required.

(a) Construct one replicate (nine blocks) in such a way that no main effect apart from A is confounded with blocks.

(b) List the characters which are confounded with blocks in this replicate (ignoring characters which are multiples of each other).

(c) Explain how to construct the second replicate in such a way that no treatment effect other than A is confounded with blocks in both replicates, and no other main effect is confounded with blocks in the second replicate.

12.4 Construct a balanced incomplete-block design for eight treatments in fourteen blocks of size four, where the treatments are all combinations of three treatment factors with two levels each.