Invariance and factorial models
by
Peter McCullagh
Department of Statistics
University of Chicago

Summary

Two factors having the same set of levels are said to be homologous. This paper aims to extend the domain of factorial models to designs that include homologous factors. In doing so, it is necessary first to identify the characteristic property of those vector spaces that constitute the standard factorial models. We argue here that essentially every interesting statistical model specified by a vector space is necessarily a representation of some algebraic category. Logical consistency of the sort associated with the standard marginality conditions, is guaranteed by category representations, but not by group representations. Marginality is thus interpreted as invariance under selection of factor levels ($\mathcal{I}$-representations), and invariance under replication of levels ($\mathcal{S}$-representations). For designs in which each factor occurs once, the representations of the product category coincide with the standard factorial models. For designs that include homologous factors, the set of $\mathcal{S}$-representations is a subset of the $\mathcal{I}$-representations. It is shown that symmetry and quasi-symmetry are representations in both senses, but that not all representations include the constant functions (intercept). The beginnings of an extended algebra for constructing general $\mathcal{I}$-representations is described and illustrated by a diallel cross design.

Keywords: analysis of variance; category representation; diallel cross; group representation; homologous factors; isomorphic representations; lattice; marginality; Mendelian inheritance; nested design; quotient space; replication invariance; selection invariance.

1. Introduction

The main purpose of this paper is to extend the definition of a factorial model to designs in which two or more factors have the same set of levels. Such factors are called homologous, a property denoted by $A \cong B$. In order to achieve this goal, it is necessary first to identify the property that characterizes factorial models in the familiar setting in which each factor occurs exactly once in the design, i.e. the factors have unrelated levels. In this way, factorial models are defined, not by a constructive recipe, but by a property that can be checked for arbitrary designs. Two criteria are identified for this purpose. Selection invariance, or $\mathcal{I}$-representation, is a property connected with the behaviour of a model under selection of factor levels. Replication invariance, or $\mathcal{S}$-representation, is a property connected with the behaviour of a model under replication of factor levels. Either of these criteria implies invariance under the symmetric groups, and also marginality in the sense of Nelder (1977).

The task of listing all factorial models in the extended sense is thus reduced to the exercise of finding all subspaces that are invariant under the relevant operation. In the familiar setting in which each factor occurs once only, selection invariance and replication invariance yield the same set of invariant subspaces, the standard factorial models. These are also called hierarchical interaction models, particularly in the context of log-linear models (Haberman 1974, Lauritzen 1996). The paper lists the replication-invariant and selection-invariant subspaces of functions on the square (designs having two homologous factors). Operators are described that are sufficient to construct the invariant subspaces that occur most commonly in such designs. Examples drawn from diverse areas of application indicate that replication-invariance and selection-invariance are natural properties for statistical models.

The approach taken in this paper has similarities with work on group invariance, randomization and experimental design, notably Dawid (1977, 1988), Diaconis (1988, 1989) Tjur (1984), Speed (1987), Speed and Bailey (1987) and Bailey (1991, 1996). But there are equally important differences, not just in terms of emphasis, but in overall approach. This paper focuses primarily on logical properties of subspaces, not on symmetries of the observation vector or of its distribution. Additivity of error and systematic effects is not assumed, and the error distribution is irrelevant so far as logical properties of such subspaces are concerned. As a consequence, complications that affect the analysis, such as lack of balance, unequal replication, missing values, multiplicative effects, and non-normality, are separated from modelling issues. The invariance properties of subspaces are equally appropriate considerations for linear and generalized linear models.

2. Example 1: analysis of reciprocal crosses

The data in Table 1, taken from Yates (1947), gives the number of seeds per 100 florets pollinated in reciprocal crosses of 12 sibs of an F_1 family of *Trifolium hybridum*. The rows (factor A) identify the sib contributing pollen (male parent), and the columns (factor B) identify the female parent. These factors have the same set of levels, namely the 12 sibs. The example shows that when the design includes homologous factors, the standard factorial models are seldom relevant. The analysis of these data is complicated by the fact that, in addition to substantial self-sterility, certain groups of sibs (2, 3, 7, 11), (1, 5, 8, 9, 10), and (4, 12) are incompatible. This incompatibility is not total, so it is appropriate to analyse separately the compatible and incompatible pairs.

We use a weighted linear model, with weights zero or one chosen to select the compatible pairs (or the incompatible pairs as desired). The decomposition in Table 2, which is not complete, makes use of four subspaces, the symmetric subspace $\mathrm{sym}(A, B)$ of functions satisfying $f_{ij} = f_{ji}$, the skew-symmetric subspace $\mathrm{alt}(A, B)$ of functions satisfying $f_{ij} = -f_{ji}$, and two non-overlapping additive subspaces:

$$\mathrm{sym}(A + B) = \{f_{ij} = \alpha_i + \alpha_j\},$$
$$\mathrm{alt}(A + B) = \{f_{ij} = \alpha_i - \alpha_j\},$$

of dimension 12 and 11 respectively, so that $A + B$ is the span of these two subspaces. The entry in the row labelled $\mathrm{alt}(A + B)$ is the regression sum of squares for the projection on to the additive alternating subspace. The row labelled $\mathrm{alt}(A, B)/\mathrm{alt}(A + B)$ gives the difference between the squared lengths of the projections on to these two nested subspaces.

Table 1. *Fertility in reciprocal crosses of 12 sibs of an F_1 family of* Trifolium hybridum

						Female parent						
Sib	2	3	7	11	1	5	8	9	10	12	4	6
2	11	24	74	10	167	134	153	144	87	68	320	155
3	10	24	75	13	95	178	190	163	84	113	61	158
7	9	48	27	33	235	203	158	132	136	252	204	129
11	20	30	58	20	108	170	168	199	119	78	139	150
1	95	112	124	122	21	35	29	20	20	105	223	148
5	83	123	123	205	63	50	33	55	21	219	152	148
8	80	114	163	156	41	28	6	27	61	212	130	131
9	73	152	162	210	58	32	39	42	64	248	171	237
10	70	125	145	89	61	71	30	69	53	177	143	134
12	175	105	218	165	78	266	204	181	85	0	3	123
4	192	116	202	163	150	195	157	154	133	0	2	155
6	120	121	155	178	154	202	163	156	110	170	239	0

Table 2. *An analysis of variance for reciprocal crosses*

	Compatible pairs			Incompatible pairs		
Vector space	S.S.	d.f.	M.S.	S.S.	d.f.	M.S.
$\mathrm{sym}(A+B)/1$	67458.8	11	6132.6	11698.3	11	1063.5
$\mathrm{sym}(A,B)/\mathrm{sym}(A+B)$	92401.8	37	2497.3	3176.6	17	186.8
Total symmetric	159860.6	48		14874.9	28	
$\mathrm{alt}(A+B)$	22686.2	11	2062.4	4609.0	8	576.1
$\mathrm{alt}(A,B)/\mathrm{alt}(A+B)$	47593.3	38	1252.4	2596.1	9	288.5
Total skew	70279.5	49		7205.1	17	

The left part of this table is essentially identical to Yates's Table 9. The penultimate row furnishes an estimate of residual variance.

For the compatible pairs, in the symmetric part of the space the predominant effects are additive, the ratio of mean squares being $F = 6133/2497 = 2.46$. So there is clear evidence of fertility differences among the sibs. While the projection on to the non-additive symmetric space is significantly larger than would be expected from random variation alone ($F = 2497/1252$), the differences among sibs have a certain degree of consistency among compatible matings. In the additive space, $(A+B)/1$, the dominant component is symmetric, the ratio of mean squares being $6133/2062 = 3.0$. There is, however, a suggestion of a differential contribution of pollen and ovum with an F-ratio of $2062/1252$ and a p-value of 12%. But the major component of variability occurs in the additive symmetric subspace, attributable to variation between parental lines, male and female contributing equally.

For the incompatible pairs, the analysis of variance table provides no evidence of non-additivity in the symmetric subspace, ($F = 187/288$). The evidence for a differential contribution of pollen and ovum is suggestive, but the F-ratio of $576/288$ gives a p-value of 16%. The dominant effect is again differences among sibs.

Assuming additivity, the easiest way to check whether there is a differential contribution

of pollen and ovum is to fit the rotated factor model

$$E(Y_{ij}) = \alpha_i \cos \theta + \alpha_j \sin \theta$$

for various values of θ in $[0, \pi)$. The model implies, in effect, that whatever advantageous genes accrue to the male of sib i also accrue to the female of sib i. The progeny obtain a combination of the parental genes in a distinctly non-Mendelian manner. If the male contribution dominates, we have $0 \leq \theta < \pi/4$. If the female contribution dominates, we have $\pi/4 < \theta \leq \pi/2$. If both parents contribute equally, we have $\theta = \pi/4$. In the genetic context, values of θ outside the range $[0, \pi/2]$, or non-convex combinations, are implausible. This is a model based on algebraic convenience rather than genetic theory such as maternal inheritance through mitochondrial DNA. Despite its obvious lack of a genetic basis, this family is useful for detecting additive asymmetries of a certain type.

Plots of the residual sums of squares for compatible and incompatible pairs separately suggest that two different mechanisms are in operation. The compatible pairs are consistent with a genetic model of equal parental contributions, but the incompatible pairs are not.

For the incompatible pairs, there is no evidence of genetic variation contributed by the male parent. As it happens, clover exhibits gametophytic incompatibility, in which, if two plants have the same genotype, age and environmental conditions of the maternal plant govern acceptance or rejection of the pollen. Gametophytic incompatibility is consistent with the observed data on the incompatible pairs in which variation among the male plants is minimal. I am grateful to Deborah Charlesworth for providing the information on this incompatibility system.
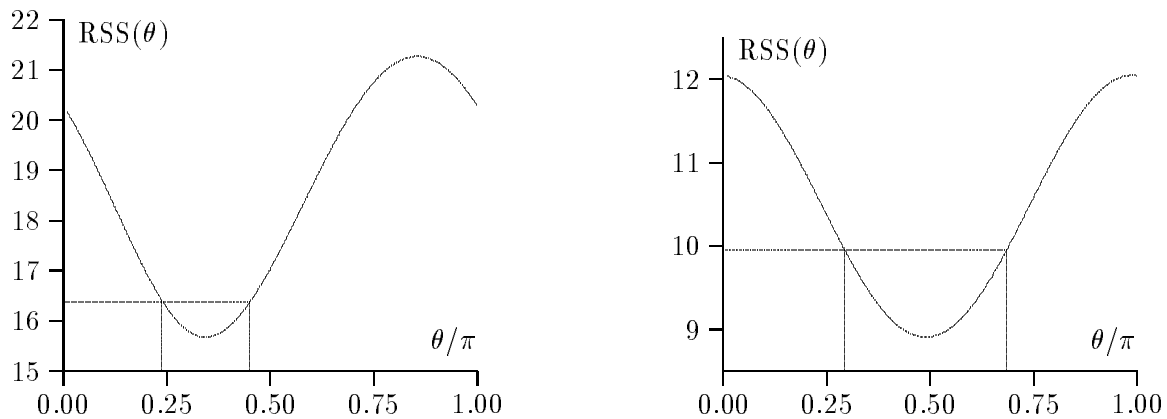

Fig. 1. Residual sum of squares for the additive model $\alpha_i \cos \theta + \alpha_j \sin \theta$ plotted against θ. The first panel is for the compatible crosses, and the second for the remaining partially incompatible crosses. Aproximate 95% confidence limits for θ are indicated. The compatible crosses are consistent with the symmetric model of equal parental contributions ($\theta = \pi/4$), but the incompatible crosses show no male parental effect.

A closer analysis of these data reveals more complicated patterns including deviations from symmetry. For example, female sib 2 appears to be partially incompatible with the members of group 2, but male sib 2 is not. Also, the conclusions suggested by Fig. 1b are entirely attributable to the anomalous behaviour of the female sib 7.

The partitioning of the sibs into four groups determines a new factor, presumably sib genotype, in two homologous copies. This makes possible analyses at the group level using further subspaces described in section 9.5.

In the sections that follow, we show that the standard factorial models and the various subspaces used in this analysis arise naturally from representation theory for the relevant category of morphisms on finite sets. As models, these representations are natural in the sense that subspaces not obeying the marginality principle do not occur.

3. Factorial models

3.1 *What is a factor?*

A factor A with levels Ω is the vector space $\mathcal{R}^\Omega$ of real-valued functions on Ω. An element α in A is thus a function that assigns to level $i \in \Omega$ a real number, $\alpha(i)$ or α_i. It is assumed throughout this paper that each factor has a finite number of levels. By this definition, the dimension of A is equal to $|\Omega|$, the number of elements in Ω.

The relation between the set of levels and the vector space is an intimate one, so much so that it is best to write $A = (\Omega, \mathcal{R}^\Omega)$ or $A = \{\Omega \mapsto \mathcal{R}^\Omega\}$. In words, *factor* is a function associating with each finite set Ω, the vector space $\mathcal{R}^\Omega$. In this paper, such a list indexed by finite sets is called a *sequence*.

From the viewpoint of applied work, this definition may seem somewhat divorced from reality because the set of statistical units is not mentioned, nor is there any mention of the mapping between the units and the factor levels. However, since statistical design is not a consideration here, this aspect of the definition turns out to be a point of strength. For comparisons with other definitions, see section 3.3.

3.2 *What is a factorial model?*

Operationally, a standard factorial model is a subspace of a tensor product space, specified by factors alone, together with operators for vector sum and tensor product (Wilkinson and Rogers, 1973). This is a descriptive recipe, not a characterization by properties. The recipe is complete and includes the subspaces that are most useful as statistical models, at least for designs in which the factors, individually and jointly, have no additional structure. But it is patently unsatisfactory for designs having homologous factors because the most useful and interesting subspaces are excluded. Nevertheless, we pursue this definition for the moment, re-casting it in a form more appropriate for our purposes.

Let $A_1, A_2, \ldots$ be factors with levels $\Omega_1, \Omega_2, \ldots$. A linear model $\mathcal{X}$ is a rule that assigns to each product set $\Omega_1 \times \Omega_2 \times \cdots$ a subspace $\mathcal{X}_{\Omega_1 \Omega_2 \ldots}$ in the tensor product space $\mathcal{R}^{\Omega_1} \otimes \mathcal{R}^{\Omega_2} \otimes \cdots$:

$$\mathcal{X} : (\Omega_1 \times \Omega_2 \times \cdots) \mapsto \mathcal{X}_{\Omega_1 \Omega_2 \ldots} \subset \mathcal{R}^{\Omega_1 \times \Omega_2 \times \cdots} \cong \mathcal{R}^{\Omega_1} \otimes \mathcal{R}^{\Omega_2} \otimes \cdots.$$

In other words, $\mathcal{X}$ is a *sequence* of vector subspaces, one such subspace for each product set.

While it may seem odd and unnecessarily cumbersome to define a model as a sequence of vector subspaces rather than a single subspace, one purpose is readily explained in statistical terms. Let Ω'_r be a subset of Ω_r, and let $A'_r = \mathcal{R}^{\Omega'_r}$ be the factor on the reduced set of levels. The rule $\mathcal{X}$ determines a sequence of subspaces, two of which are $\mathcal{X}(\Omega) \subset A_1 \otimes A_2 \cdots$, and

$\mathcal{X}(\Omega') \subset A'_1 \otimes A'_2 \cdots$. For consistency of interpretation in statistical work, it is necessary that $\mathcal{X}(\Omega')$ should be equal to the restriction of functions in $\mathcal{X}(\Omega)$ to the levels $\Omega'_1 \times \Omega'_2 \times \cdots$. This property, called selection invariance, is intrinsically a property of the sequence, not a property of any individual subspace. The condition is not satisfied by arbitrary rules: on the contrary, it is sufficiently strong to characterize standard factorial models.

The following is an example of a rather trivial, but important, sequence of subspaces that is selection-invariant. Let 1 denote the one-dimensional vector space of functions that are constant on Ω. The restriction of 1 to Ω' is equal to the vector space of functions that are constant on Ω'. In other words, the *sequence* 1 is closed under selection. Likewise, the sequence $A = \mathcal{R}^\Omega$ is also closed under selection.

For a non-trivial example in which the condition is satisfied, the model 'no interaction between treatment and blocks' assigns to the product set $\Omega_1 \times \Omega_2$ the additive subspace denoted by $A \otimes 1 + 1 \otimes B$ or, more conventionally, by $A + B$. As usual in statistical work, $A = \mathcal{R}^{\Omega_1}$ is regarded as a subspace of

$$A \otimes B = \mathcal{R}^{\Omega_1 \times \Omega_2} \cong \mathcal{R}^{\Omega_1} \otimes \mathcal{R}^{\Omega_2}$$

by the identification $A \cong A \otimes 1$. A vector v in $A + B$ is a function that assigns to the ordered pair (i, j) a real number $\alpha_i + \beta_j$, for some $\alpha \in A$ and $\beta \in B$. The restriction of v to the subset $\Omega'_1 \times \Omega'_2$ is equal to $\alpha_i + \beta_j$, in which each index is restricted to the relevant subset. In other words, the restriction of v is a vector in $A' + B'$, which is the rule *no interaction* applied to the product subset. Further, the restriction of $A + B$ coincides with $A' + B'$.

For an example in which the condition is not satisfied, it is sufficient to consider a single factor with levels Ω, and $\mathcal{X}_\Omega$ equal to the subspace $1_\Omega^\perp$ of functions that add to zero on Ω. It should be noted that this subspace is invariant under the action of the symmetric group on Ω. Nevertheless, the restriction of $1_\Omega^\perp$ to a proper subset is not equal to $1_{\Omega'}^\perp$, but rather to $\mathcal{R}^{\Omega'}$. The zero-sum constraint is not satisfied on the restricted set.

One candidate definition at this stage is as follows. *The factorial models are the selection-invariant sequences.* Whatever our ultimate definition, we shall certainly require this property. In the standard setting in which all factors are crossed and non-isomorphic, the selection-invariant sequences coincide with the conventional factorial models.

3.3 *Tjur's definitions*

Tjur's (1984) definition of a factor and a factorial model, defined in the language of set partitions, is different from the one used here. But there are important points of similarity, particularly in the emphasis on mappings between finite sets. The definitions are largely complementary, with Tjur's being more appropriate for design considerations.

Let $\mathcal{U}$ be the set of experimental or observational units, and let φ be a mapping from $\mathcal{U}$ into Ω, the set of factor levels. The inverse image assigns to each $i \in \Omega$ the set $\varphi^{-1}(i)$ of units having the factor at level i. In other words, φ^{-1} is a function from Ω into non-overlapping subsets of $\mathcal{U}$ whose union is $\mathcal{U}$. Some of these subsets may be empty. If the empty sets are discarded and the labels are ignored, φ^{-1} determines a partition of the units. It is this partition, or the vector subspace of $\mathcal{R}^{\mathcal{U}}$ of functions that are constant on the blocks of the partition, that Tjur takes as the definition of a factor. The dimension of the vector space is equal to the number of blocks of the partition, which may well fall short of the number of elements in Ω.

Given two partitions A, B of $\mathcal{U}$, if A is a sub-partition of B we write $A \leq B$, which is a partial order on the partitions. So far as the factors are concerned, the terminology is reversed: B is a subspace, or sub-factor, of A. The infimum or greatest lower bound $A \wedge B$ is the partition whose blocks are the non-empty intersections of the blocks of A and B. The vector subspace associated with $A \wedge B$ is conventionally denoted by $A.B$. Since every element in the product set $\Omega_A \times \Omega_B$ need not occur in the design, the dimension of $A.B$ may be less than the dimension of $A \otimes B$. In extreme cases, if A is a sub-partition of B, $A.B = A$. Given factors, or partitions, A, B, C, we can construct all infima together with the associated vector spaces. A factorial model is then defined to be the vector span of some subset of these vector spaces together with the constant function.

The preceding construction can be extended to Tjur block structures by including all upper bounds and the associated vector spaces. The least upper bound $A \vee B$ of A and B is the least partition such that $A \leq A \vee B$ and $B \leq A \vee B$. Standard factorial models do not ordinarily include these vector spaces.

The difference between the two definitions is principally a difference of emphasis. Tjur's definition is concerned with design aspects, i.e. the factor combinations that occur among the observed units. Our definition is concerned not so much with the observed combinations, but with the combinations that might occur in principle. The effects that are potentially of interest need not be estimable with the particular design at hand.

The relationship between the two definitions is readily explained in terms of maps between finite sets. Let A, B and C be factors corresponding to the mappings $\varphi_1 : \mathcal{U} \to \Omega_1$, $\varphi_2 : \mathcal{U} \to \Omega_2$, and $\varphi_3 : \mathcal{U} \to \Omega_3$. Then $\varphi = (\varphi_1, \varphi_2, \varphi_3)$ is a mapping from $\mathcal{U}$ into the product set $\Omega_1 \times \Omega_2 \times \Omega_3$. For any f in $A \otimes B \otimes C$, define $\varphi^* f$ by composition:

$$(\varphi^* f)(u) = f(\varphi_1(u), \varphi_2(u), \varphi_3(u))$$

for $u \in \mathcal{U}$. Thus φ^* takes f in $A \otimes B \otimes C$ into $\varphi^* f$ in $\mathcal{R}^{\mathcal{U}}$ by restriction to the observed levels and replication where necessary. In fact φ^* is a linear transformation from $A \otimes B \otimes C$ onto the subspace $A.B.C \subset \mathcal{R}^{\mathcal{U}}$, Likewise, φ^* takes $A \otimes B \otimes 1$ onto $A.B$, and so on for any factorial model.

In statistical applications, the assumptions made in connection with random variation typically have the effect of making $\mathcal{R}^{\mathcal{U}}$ an inner product space, thereby permitting an orthogonal decomposition. By contrast, $A \otimes B \otimes C$ has no inner product, so orthogonal complements do not exist in this space.

4. Categories and representation theory

4.1 *Finite sets*

Quite frequently, though certainly not always, the set of potential levels of a factor is very large. The set of levels occurring in an experiment such as a variety trial is typically a small subset of the totality of varieties. The fundamental reason for defining a model as a sequence indexed by finite sets, is to establish a logical relation between functions on one set of factor levels and functions on any related set of levels. The notion of a factor is thus intimately connected with mappings between finite sets, the so-called *category of finite sets*.

The category of finite sets $\mathcal{F}$ is the class of all finite sets together with all transformations between pairs of sets. This is a closed system, a sort of semi-group under limited

composition. For each $\varphi\colon \Omega \to \Omega'$ and $\psi\colon \Omega' \to \Omega''$, the composite function $\psi\varphi$ is a mapping from Ω to Ω'', and thus in $\mathcal{F}$. For further details, see MacLane (1971) or MacLane and Birkhoff (1988).

Let h be a real-valued function on Ω''. Then $g = h \circ \psi$ is a real-valued function on Ω' and $f = g \circ \varphi = h \circ \psi \circ \varphi$ is a real-valued function on Ω. The vector g, called the pullback of h by ψ, is obtained by linear transformation ψ^* from $\mathcal{R}^{\Omega''}$ to $\mathcal{R}^{\Omega'}$. Likewise f is obtained by the linear transformation φ^* from $\mathcal{R}^{\Omega'}$ to $\mathcal{R}^{\Omega}$. That is to say, $\psi^*h = h \circ \psi$ and $\varphi^*g = g \circ \varphi$ are linear transformations between vector spaces. The pullback map corresponding to $\psi\varphi$ is the product $\varphi^*\psi^*$ in reverse order as shown below:

$$
\begin{array}{ccccc}
\Omega & \xrightarrow{\ \varphi\ } & \Omega' & \xrightarrow{\ \psi\ } & \Omega'' \\
\downarrow & & \downarrow & & \downarrow \\
\mathcal{R}^{\Omega} & \xleftarrow{\ \varphi^*\ } & \mathcal{R}^{\Omega'} & \xleftarrow{\ \psi^*\ } & \mathcal{R}^{\Omega''}
\end{array}
$$

The top line of this diagram is a schematic diagram of $\mathcal{F}$ with arrows representing maps between sets. The lower line is a similar diagram in which each set in $\mathcal{F}$ is mapped to a vector space, and each arrow in $\mathcal{F}$ is mapped to a linear transformation between vector spaces. The mapping from top line to the lower line depicts a functor, or representation, the standard representation of $\mathcal{F}$. This functor is said to be contravariant because the arrow function $*$ that takes φ to φ^* reverses the direction of the map.

4.2 $\mathcal{F}$-representation

A representation of the category $\mathcal{F}$ is a contravariant functor from $\mathcal{F}$ into a new category whose objects are certain vector spaces and whose morphisms are certain linear transformations between these vector spaces. In other words, to each finite set Ω the functor associates a vector space $\mathcal{X}_\Omega$, and to each morphism $\varphi\colon \Omega \to \Omega'$ the functor associates a linear transformation $\varphi^*\colon \mathcal{X}_{\Omega'} \to \mathcal{X}_\Omega$. Further, the identity morphism on Ω is mapped to the identity on $\mathcal{X}_\Omega$, and the composite morphism $\psi\varphi$ is mapped to the composite linear transformation $(\psi\varphi)^* = \varphi^*\psi^*$ in reverse order.

For each fixed positive integer k, the sequence of vector spaces $(\mathcal{R}^\Omega)^{\otimes k}$ is an $\mathcal{F}$-representation. It is understood here that to each $\varphi\colon \Omega \to \Omega'$ in $\mathcal{F}$ we associate the pullback map $\varphi^*\colon (\mathcal{R}^{\Omega'})^{\otimes k} \to (\mathcal{R}^\Omega)^{\otimes k}$ by functional composition

$$
(\varphi^*f)(i_1, i_2, \ldots, i_k) = f\big(\varphi(i_1), \varphi(i_2), \ldots, \varphi(i_k)\big).
$$

A sub-representation $\mathcal{X}$ is a sequence of subspaces $\mathcal{X}_\Omega \subset (\mathcal{R}^\Omega)^{\otimes k}$ such that $\varphi^*\mathcal{X}_{\Omega'} \subset \mathcal{X}_\Omega$.

For $k = 1$ this definition does not go far. The only $\mathcal{F}$-representations that occur as subspaces of $A = \mathcal{R}^\Omega$ are zero, 1 and A. However, the definition applies directly to homologous factors in which $\mathcal{X}_\Omega$ is a subspace of $A^{\otimes 2} = \mathcal{R}^{\Omega \times \Omega}$, the set of functions on the square $\Omega \times \Omega$. The sequence $\{\mathcal{X}\}$ is an $\mathcal{F}$-representation if, for each $g \in \mathcal{X}_{\Omega'}$, the pullback φ^*g is a vector in $\mathcal{X}_\Omega$. It is clear, for example, that if g is symmetric on $\Omega' \times \Omega'$ then φ^*g is symmetric on $\Omega \times \Omega$. The sequence of symmetric subspaces $\mathrm{sym}^2(A)$ is thus an example of an $\mathcal{F}$-representation in $A^{\otimes 2}$. The following diagram gives two complete direct-sum decomposition of $A^{\otimes 2}$ by $\mathcal{F}$-representations in which sym^* is the symmetric subspace with zero diagonal:

$$
\begin{aligned}
A^{\otimes 2} &= \mathrm{sym}^* \ \oplus\ \{1 \subset \mathrm{sym}^+\} \ \oplus\ \{\mathrm{alt}^+ \subset \mathrm{alt}\} \\
&= \mathrm{sym}^* \ \oplus\ \{1 \subset A_\theta\} \ \oplus\ \{\mathrm{alt}^+ \subset \mathrm{alt}\}.
\end{aligned}
$$

The subspaces sym^+ and alt^+ are the symmetric and alternating additive subspaces consisting of functions of the form $\alpha_i + \alpha_j$ and $\alpha_i - \alpha_j$ respectively, alt is an abbreviation for $\mathrm{alt}^2(A)$, and A_θ for $\theta \neq 3\pi/4$ is the rotated factor subspace with components $\alpha_i \cos\theta + \alpha_j \sin\theta$.

Each $\mathcal{F}$-representation in $A^{\otimes 2}$ is the span of three sub-representations $U + V + W$ in which $U \in \{0, \mathrm{sym}^*\}$, $V \in \{0, 1, A_\theta\}$ for some $\theta \neq 3\pi/4$, and $W \in \{0, \mathrm{alt}^+, \mathrm{alt}\}$. The structure of this decomposition implies various 'marginality conditions'. For example, every model that includes the additive symmetric subspace necessarily includes the constant functions, or intercept. However, a model that contains the additive alternating subspace need not include the intercept.

4.3 $\mathcal{I}$-representation

The category of all morphisms between finite sets has a number of important sub-categories, one of which is the category $\mathcal{I}$ whose objects are all finite sets, and whose morphisms are all injective, or 1–1, transformations. An $\mathcal{I}$-representation is a contravariant functor from $\mathcal{I}$ into a category whose objects are certain vector spaces, and whose morphisms are surjective linear mappings between these vector spaces.

The sequence $(\mathcal{R}^\Omega)^{\otimes k}$ together with the pullback maps defined above, constitutes an $\mathcal{I}$-representation. A sequence of vector subspaces $\mathcal{X}_\Omega \subset (\mathcal{R}^\Omega)^{\otimes k}$ is an $\mathcal{I}$-representation if, for each 1–1 mapping $\varphi\colon \Omega \to \Omega'$, the restriction map φ^* that takes $(\mathcal{R}^{\Omega'})^{\otimes k}$ to $(\mathcal{R}^\Omega)^{\otimes k}$ satisfies $\varphi^* \mathcal{X}_{\Omega'} = \mathcal{X}_\Omega$. A model that is an $\mathcal{I}$-representation is said to be selection-invariant in the sense of section 3.2.

A complete direct-sum decomposition of $A^{\otimes 2}$ by $\mathcal{I}$-representations is given below in which an asterisk denotes a subspace with zero diagonal:

$$A^{\otimes 2} \;=\; \{1^* \subset \mathrm{sym}^{*+} \subset \mathrm{sym}^*\} \;\oplus\; \{1_d \subset \mathrm{diag}\} \;\oplus\; \{\mathrm{alt}^+ \subset \mathrm{alt}\}.$$

The diagonal subspace is an example of an $\mathcal{I}$-representation that is not an $\mathcal{F}$-representation. Note that, as an $\mathcal{F}$-representation, sym^* is irreducible: it contains no $\mathcal{F}$-subrepresentations. As an $\mathcal{I}$-representation, however, sym^* is not irreducible: it contains the sub-representations sym^{*+} and 1^*.

On account of multiplicities or isomorphisms, $\mathrm{alt}^+ \cong \mathrm{sym}^{*+}/1^* \cong \mathrm{diag}/1_d$ and $1^* \cong 1_d$, this decomposition is not unique, so it is not easy to characterize all of the subspace representations in $A^{\otimes 2}$. In principle, however, when isomorphic representations are identified, all $\mathcal{I}$-representations, including the rotated factor models, can be extracted from any complete decomposition. Details are given in section 9.

4.4 Representations of related categories

Category theory asks of every type of mathematical object: 'What are the morphisms?'; it suggests that these morphisms should be described at the same time as the objects. This quote, taken from MacLane (1971), is strongly appealing for statistical work. It implies that the models used should be tailored to the nature of the response scale and the levels of the explanatory factors. To say the same thing in another way, each vector space used as a statistical model should ideally be a representation of the *relevant* category of morphisms.

Most of the categories that occur in statistical work are product categories in which the objects are product sets and the morphisms are direct products. Depending on the context,

the component categories may be $\mathcal{F}$ or $\mathcal{I}$ acting on finite sets, the affine group acting on the real line, or one of the categories listed below.

(i) The category $\mathcal{S}$ of surjections: objects are all finite sets; morphisms are all surjective maps.

(ii) The category $\mathcal{F}_0$: objects all finite sets that include zero; morphisms all maps taking zero to zero and non-zero to non-zero.

(iii) The category whose objects are all finite sets that include a residual element called 'other' or 'none of the above'; the morphisms are surjections taking 'other' to 'other'.

(iv) The category $\mathcal{S}_m$ of monotone surjections: objects all finite partially ordered sets; morphisms all weakly monotone surjections.

(v) Objects all finite graphs; morphisms all neighbour-preserving maps $\varphi\colon \Omega \to \Omega'$. (If i, j are neighbouring vertices in Ω, then $\varphi(i), \varphi(j)$ are neighbours in Ω'.)

(vi) In a nested design, each object is a set of units together with the equivalence relation $i \sim j$ if units i and j belong to the same block (section 9.5). The morphisms are all equivalence-preserving injective maps $i \sim j \iff \varphi(i) \sim \varphi(j)$.

The list of variations is virtually endless, but it must be remembered that apparently different descriptions may be equivalent as abstract categories, either in the sense of isomorphism or in the sense of opposites.

The category of surjections has two interpretations that are relevant for statistical purposes. For response factors, a surjection is intimately connected with the operation of amalgamating response levels; the sub-category of monotone surjections is relevant for ordinal response scales.

For explanatory factors, it may happen that some of the levels in Ω, though labelled distinctly, are in fact equivalent. That is to say that a reduced factor with fewer levels may suffice. The surjection $\varphi\colon \Omega \to \Omega'$ maps equivalent levels in Ω to the same image in Ω', and thereby generates a partition of Ω labelled by Ω'. For each $f \in \mathcal{X}_{\Omega'}$, the pullback $\varphi^* f$ is constant on the blocks of the partition. A sequence $\mathcal{X}_\Omega \subset (\mathcal{R}^\Omega)^{\otimes k}$ is an $\mathcal{S}$-representation if two conditions are satisfied. First, $\varphi^* \mathcal{X}_{\Omega'}$ is a subspace of $\mathcal{X}_\Omega$. Second, if g lies in $\varphi^*(\mathcal{R}^{\Omega'})^{\otimes k}$ and also in $\mathcal{X}_\Omega$ then $g \in \varphi^* \mathcal{X}_{\Omega'}$. These conditions are written in the form $\varphi^{*\perp 1} \mathcal{X}_\Omega = \mathcal{X}_{\Omega'}$.

The category of monotone injections is relevant for explanatory factors having ordered levels that are non-quantitative. The categories $\mathcal{F}_0$ or $\mathcal{I}_0$ may be relevant if one of the factor levels has a special status. One example is dose, the level zero being special because zero dose of one drug is indistinguishable from zero dose of any other. Another example mentioned briefly in section 6 is an industrial process, one of whose levels is 'off' in which no physical processing occurs. The representation theory for some of these categories is discussed by McCullagh (1998).

4.5 *Monoids*

A monoid is any semi-group that contains the identity, a category with only one object. So far as this paper is concerned, the term monoid signifies the set of transformations Ω^Ω from the finite set Ω into itself. The monoid contains the symmetric group, which is the set of *invertible* transformations on Ω.

Definition: *A subspace $\mathcal{X}$ in $\mathcal{R}^\Omega$ or $\mathcal{R}^{\Omega \times \Omega}$, or in some higher-order tensor product space, is said to be monoid invariant or an M-representation, if, for every $\varphi \in \Omega^\Omega$, the pullback map satisfies $\varphi^* \mathcal{X} \subset \mathcal{X}$.*

To see what this means in terms of components, let f be a vector in $\mathcal{X} \subset A^{\otimes 2}$. Then f has components f_{ij} and $\varphi^* f$ has components $f_{\varphi(i)\varphi(j)}$ for $(i,j) \in \Omega \times \Omega$. Consider, for example, the subspace $\mathcal{X} = \text{sym}^*$ defined by

$$\mathcal{X} = \{f : f_{ij} = f_{ji};\ f_{ii} = 0\}.$$

It is readily checked that $\varphi^* f$ also belongs to $\mathcal{X}$, so $\mathcal{X}$ is M-invariant. By contrast, the diagonal subspace of functions that are zero for $i \neq j$ is not M-invariant. Let $f = \text{diag}\{3, 2, 4\}$, $\varphi: (1, 2, 3) \mapsto (1, 2, 1)$ so that

$$\varphi^* f = \begin{pmatrix} 3 & 0 & 3 \\ 0 & 2 & 0 \\ 3 & 0 & 3 \end{pmatrix},$$

which is not diagonal.

Since the monoid includes the symmetric group, it is evident that every M-invariant subspace is also group-invariant, or a representation of the symmetric group. But the great majority of group-invariant subspaces, for example the diagonal subspace, are not M-invariant.

Note that monoid- and group-invariance are properties of a single subspace, not properties of a sequence. The definition can be extended to a sequence by defining an M-representation to be any sequence of M-invariant subspaces. However, this is not a useful definition because it is satisfied by artificial sequences such as

$$\mathcal{X}_\Omega = \begin{cases} 1 & \text{if } |\Omega| \text{ is prime;} \\ \mathcal{R}^\Omega & \text{otherwise,} \end{cases}$$

which are of no interest in applied work. The natural extension of an M-invariant subspace turns out to be an $\mathcal{F}$-representation (McCullagh, 1998). There appears to be no equally natural extension for group-invariant subspaces.

4.6 *Relationships among the representations*

Let $\mathcal{X}_1$ and $\mathcal{X}_2$ be two sub-representations in $\mathcal{X}$. That is to say that $\mathcal{X}_1$, $\mathcal{X}_2$ and $\mathcal{X}$ are three representations of the same category, and $\mathcal{X}_1(\Omega)$ and $\mathcal{X}_2(\Omega)$ are both subspaces of $\mathcal{X}(\Omega)$. Then the intersection $\mathcal{X}_1 \cap \mathcal{X}_2$ and the vector span $\mathcal{X}_1 + \mathcal{X}_2$ are representations in the same sense, the linear mappings φ^* being those inherited from $\mathcal{X}$. The set of representations thus forms a lattice in which the least upper bound is the vector span, and the greatest lower bound is the intersection.

Justification for the statements listed below can be found in McCullagh (1998).
1. Every $\mathcal{F}$-representation is also an $\mathcal{I}$-representation and an $\mathcal{S}$-representation.
2. Any sequence that is both an $\mathcal{I}$-representation and an $\mathcal{S}$-representation is also an $\mathcal{F}$-representation.
3. Each subspace in an $\mathcal{F}$-representation is M-invariant.
4. To each M-invariant subspace $(\Omega_0, \mathcal{X}_0)$ there corresponds a minimal sequence $\{\mathcal{X}\}$ that is an $\mathcal{F}$-representation.
5. To each M-invariant subspace $(\Omega_0, \mathcal{X}_0)$ there corresponds a maximal sequence $\{\mathcal{X}\}$ that is an $\mathcal{F}$-representation.

6. For crossed designs in which each factor occurs once, the product $\mathcal{F}$-, $\mathcal{I}$-, M- and $\mathcal{S}$-representations in the tensor product space $\mathcal{R}^{\Omega_1 \times \Omega_2 \times \cdots}$ coincide with the standard factorial models.

7. It is conjectured that every $\mathcal{S}$-representation is also an $\mathcal{F}$-representation, i.e. that these two classes coincide.

By the first statement we mean that the restriction of any $\mathcal{F}$-functor to the sub-category $\mathcal{I}$ is a functor with the same object sequence, but with surjective linear maps. Likewise, the restriction to $\mathcal{S}$ has injective linear maps. Point 2 is a consequence of the fact that any transformation between finite sets can be represented as the composition of a surjection and an injection. Point 3 is an immediate consequence of the definition. By point 4 above is meant first that $\mathcal{X}_{\Omega_0} = \mathcal{X}_0$. Second, for each $\mathcal{F}$-representation $\{\mathcal{X}'\}$ that includes $\mathcal{X}_0$, $\mathcal{X}_\Omega \subset \mathcal{X}'_\Omega$ for every Ω. Point 5 has a similar meaning in which $\mathcal{X}'_\Omega \subset \mathcal{X}_\Omega$. A proof of 6 is given below for M-representations.

A consequence of points 3–5 is that monoid invariance and $\mathcal{F}$-representations are essentially equivalent. Conjecture 7 is based largely on absence of evidence to the contrary.

5. Product groups and product monoids

5.1 *Group representations*

The word 'group' refers here to the symmetric group acting on the factor levels. If the design contains several non-isomorphic factors, it is understood that the relevant group is the product symmetric group acting independently on each set of levels.

A group representation is a subspace that is invariant under the action of the group. The G-representations that occur in $A = \mathcal{R}^\Omega$ are 0, 1, $1^\perp$ and A. A representation is said to be irreducible if it contains no sub-representation other than zero and itself. A key simplification in representation theory for finite groups is that every G-representation is a direct sum of irreducibles. This property does not extend to $\mathcal{I}$-representations or $\mathcal{F}$-representations. The decomposition of A by G-irreducibles is $A = 1 \oplus 1^\perp$.

So far as the tensor product space $A \otimes B \otimes C$ is concerned, theorem 10 of Serre (1977) states that each irreducible representation of the product group is isomorphic with a tensor product of irreducibles:

$$A \otimes B \otimes C = \{1_A \oplus 1_A^\perp\} \otimes \{1_B \oplus 1_B^\perp\} \otimes \{1_C \oplus 1_C^\perp\}.$$

These eight irreducibles are non-isomorphic, and therefore unique. Including the zero subspace, there are $2^8 = 256$ distinct G-representations in $A \otimes B \otimes C$, one for each subset of the irreducibles. Only 20 of these, the factorial models, are also monoid-invariant.

For k factors, there are 2^k G-irreducibles in the tensor product space, and 2^{2^k} distinct G-representations. The number of factorial models also increases rapidly with k, but at a much slower rate.

5.2 *Product monoid*

We now show that the representations of the product monoid $M = M_A \times M_B \times M_C$ that occur as subspaces in $A \otimes B \otimes C$ are the standard factorial models, in 1–1 correspondence with the free distributive lattice on three, or more generally k, generators. The first point to note is that the product monoid contains the product group, so each M-representation is also G-invariant. The key step in the proof is to examine the effect of the monoid on each G-irreducible, and to find the smallest M-invariant subspace that contains the given G-irreducible. We first note that the range of the action of M_A on the G-irreducible $1_A^\perp$ is

$$\text{span}\{1_A^\perp \circ \varphi\} = \{f \circ \varphi : f \in 1_A^\perp, \varphi \in M_A\} = A,$$

provided that $|\Omega_A| \geq 2$. Likewise, the action of the product monoid on $1_A^\perp \otimes 1_B^\perp \otimes 1_C$ is

$$\text{span}\{1_A^\perp \otimes 1_B^\perp \otimes 1_C \circ \varphi\} = \{f \circ \varphi : f \in 1_A^\perp \otimes 1_B^\perp \otimes 1_C, \varphi \in M\} = A \otimes B \otimes 1_C.$$

In other words, $A.B = A \otimes B \otimes 1$ is the smallest M-invariant subspace that contains the G-irreducible $1_A^\perp \otimes 1_B^\perp \otimes 1_C$.

This argument, due to B. Totaro (personal communication), can readily be extended to k factors provided that each factor has at least two levels. Thus we arrive at the following theorem.

Theorem: *The subspaces of the tensor product space $A_1 \otimes A_2 \otimes \cdots \otimes A_k$ that are invariant under the product monoid are the standard factorial models in 1–1 correspondence with the free distributive lattice on k generators.*

6. Practical considerations

6.1 *Interaction*

In discussions concerning interaction, it is helpful to distinguish a number of different types of factors that arise in typical experimental designs or observational studies. Although factors of different types might be treated on a similar basis in a model, distinctions are usually essential and helpful for purposes of interpretation. The distinction between a treatment factor and a classification factor is fundamental. In an experiment, the levels of the treatment factor are assigned, ideally at random, by the experimenter to the observational units, or subjects. In an observational study, the levels of a treatment factor could in principle be assigned randomly, but, for ethical or practical reasons, are not. A classification factor, by contrast, is an intrinsic property of the experimental units, sex, age and medical history being examples in a medical context, soil type being an example from an agricultural context.

In a factorial design, the treatment effect can be measured at each level of the classification factor. If the treatment effect is constant, we say that there is no interaction, and if this constant is zero, we say that there is no treatment effect. At the other extreme, if the treatment effect is positive for some levels of the classification factor and negative for other levels, we say that there is qualitative interaction. For example, an acidic fertilizer that is beneficial to azaleas may be harmful to other shrubs. Qualitative interactions of this sort are probably the exception in practice. A more common phenomenon is the case in which the treatment effect is essentially zero unless certain conditions are met. For example,

the addition of trace elements to the diet of grazing animals has a dramatic effect in areas where these elements are deficient: in areas where the trace elements are naturally plentiful, supplements have no measurable effect. Certain chemical reactions require the presence of a catalyst. Some plants are sensitive to soil composition, in particular acidity and drainage: a neutral fertilizer may have little effect while the pH is outside the preferred range, but substantial effects otherwise.

Two points that are relevant in the construction of models for interaction are the following. First, if the treatment effect varies substantially with the level of the classification factor, the equally-weighted average of the treatment effects, i.e. the average effect of treatment, is rarely of great interest. Substantial interactions suggest that qualitatively different mechanisms are in operation at the levels of the classification factor. In these circumstances, it is often effective to analyse the data separately for each level of the classification factor (Cox and Snell, 1981), and to present the results separately for each level. Second, if qualitative interaction is present, it is mathematically possible that the equally-weighted average of the treatment effects is zero. However, it is extremely rare that such combinations correspond to scientifically interesting hypotheses. Even if the levels of the classification factor occur with equal frequencies in the experiment, they need not occur with the same frequencies in any relevant population or sub-population.

The marginality principle as formulated by Nelder (1977) forbids models having an interaction without the associated main effects. Although several discussants of Nelder's paper, including Lindley, Cox and Tukey, object to a universal prohibition, convincing examples of the need for such models are rare indeed. Substantial interaction in the presence of a null main effect can occur as a result of experimental design. One such example is described by Cox. A second line of argument against marginality is based on the observation that factors can be transformed so that interaction in one system becomes a main effect in the transformed system. The most convincing examples of this phenomenon involve homologous binary factors (Cox, 1972; Bloomfield, 1974). In the Cox-Bloomfield example, the two classification factors are political affiliation of the wife A, and political affiliation of the husband, B. In any analysis of data indexed in this way, it may be helpful to transform the factors to A', taking the value 'same' or 'different', and $B' = B$. Because the factors are homologous, the relevant symmetry group or set of morphisms is not the product set so the appropriate definition of marginality is less obvious. The Cox-Bloomfield example is entirely in accord with the extended definition of a factorial model as a category representation. We take A' to be either the $\mathcal{F}$-representation $1 + \mathrm{sym}^*$ or the $\mathcal{I}$-representation $1^* + 1_d$, which coincide for binary factors.

A different type of example in which transformation of factors may lead to a simpler analysis is as follows. Consider an experiment to study how chemical composition of a nutritive medium affects bacterial growth. The factors under study are chemical salts added to an aqueous solution as follows:

$$A : \mathrm{MgSO_4}$$
$$B : \mathrm{NaCl}$$
$$C : \mathrm{MgCl_2}$$
$$D : \mathrm{NaHCO_3}$$

In each case the low level denotes absence, and the high level denotes a concentration of 0.1 moles per litre. While a factorial model using these four factors might prove useful,

an analysis using the implied ionic concentrations, including H^+ and $(OH)^\perp$, seems more promising. In solution, ions of the same species are chemically indistinguishable. This line of thinking leads naturally to a factorial model in which the six ionic species

$$H^+,\ OH^\perp,\ Mg^{++},\ Na^+,\ SO_4^{\perp\perp},\ Cl^\perp.$$

are the factors, with log concentration as numerical value. Since $[H^+]$ and $[OH^\perp]$ are linearly dependent, there are in effect only five factors.

To take a second example from an industrial context, consider a production line for metal castings in which each item goes through a sequence of manufacturing steps in a definite order from rough casting, stamping, grinding, milling and tempering followed by polishing and peening. The parameters at each step are adjustable, for example grit size and speed at the grinding step, temperature and oil composition at the tempering step, and so on. Further, certain steps are mandatory, but other steps such as polishing and peening may be primarily cosmetic, and thus optional. Optional factors differ from mandatory factors in one important way, namely that they have a level 'off' in which no processing occurs. Because of the sequential nature of the manufacturing process, it may well happen, and indeed it is desirable, that errors and misalignments introduced at earlier stages are corrected, at least partially, by subsequent manufacturing steps. Only if the subsequent steps are set at 'off' do the effects of earlier steps persist in the final product. For an instance of this phenomenon, see example 1 in Hamada and Wu (1992) where, among the seven factors studied, only the final two have an effect that is detectable in the product, and the penultimate factor has an effect only if the final factor is off.

One possibility for modelling designs with optional factors is to use representation theory for the category $\mathcal{F}_0$ (McCullagh, 1998). If all factors are optional with two levels, all one-dimensional subspaces are isomorphic $\mathcal{F}_0$-representations. Thus, the decomposition by $\mathcal{F}_0$-representations is not useful unless the number of levels is at least three. A different option for modelling data from an industrial production-line experiment with optional factors A, B, C in reverse order is as follows. Suppose that each factor has a level 'off' coded as zero, and that any effect of non-trivial processing at one stage is reduced by a multiplicative factor λ by subsequent processing. Then a plausible model for the response might be

$$E(Y_{ijk}) = \mu_0 + \alpha_i + \lambda_i \beta_j + \lambda'_{ij} \gamma_k$$

in which $\lambda_0 = \lambda'_{00} = 1$, and $\alpha_0 = \beta_0 = \gamma_0 = 0$. In other words, if the final two factors are set at 'off' the contribution is γ_k from factor C. If factor A is off, the contribution is $\beta_j + \lambda'_{0j} \gamma_k$: if B is off, the contribution is $\alpha_i + \lambda'_{i0} \gamma_k$.

The main lesson to be learned from these examples is that, in applications, recorded factors frequently have additional properties or further structure that can, and should, be exploited by statistical models. The essence of this structure can frequently be described by a suitable category of morphisms on certain finite sets. The mere fact that the observations are indexed by a product set is not, in itself, justification for restricting attention to factorial models.

6.2 *Category representation versus group representation*

The chief aim of statistical models is to gain understanding by simplification. Simplification can be achieved in many ways, for example by non-linear transformation of the response to remove interactions (Box and Cox, 1964), by re-definition of factors (Cox, 1972; Bloomfield, 1974), or by splitting the data on the levels of a classification factor if appropriate. Simplification is also achieved by combining factor levels into subsets that are equivalent in their effect. In a variety trial, for example, certain varieties that are labelled differently might be genetically identical, or otherwise similar in terms of yield, response to fertilizer and method of cultivation. In molecular genetics, distinct base triplets coding for the same amino acid are biochemically equivalent. A sequence $\mathcal{X}$ is an $\mathcal{S}$-representation if, whenever f in $\mathcal{X}_\Omega$ is constant on certain blocks, i.e. f lies in the range of φ^* in $\mathcal{X}_\Omega$, then $f \in \varphi^* \mathcal{X}_{\Omega'}$, so that the reduced factor with levels Ω' suffices.

The set of levels occurring in an experiment or observational study need not be exhaustive, and might be slightly different if the study were repeated. Any model for the design actually performed should be compatible with the model for the notional design in which all levels are represented. This argument, in itself, does not lead to selection invariance because the factor levels might have a quantitative structure that can be exploited to ensure compatibility. For example, dose in a pharmaceutical study is usually restricted to a small number of arbitrary quantitative levels. A model that is linear in the dose, or log dose, satisfies the compatibility criterion. However, if permutation invariance is granted, compatibility under selection leads to models that are $\mathcal{I}$-representations.

Either of these arguments leads to the standard set of factorial models. Since both criteria are plausible in most applications, either one can be taken as the *definition* of marginality and applied to homologous factors where the issues are unsettled (Nelder, 1994, Lindsey, 1995). Monoid-invariant models, or $\mathcal{F}$-representations, are those that satisfy both conditions.

6.3 *Response factors*

All of the arguments given above address factors in the role of explanatory variable, and may be inappropriate for response factors such as occur in multinomial-response models. A model for a response factor A is set of probability distributions, or measures, on the response levels Ω. While amalgamation of response levels is frequently a relevant operation, selection is less compelling. The usual device for modelling response factors is to associate with each probability distribution on Ω, a point in some vector space. For log-linear models, the vector of log probabilities determines a point in $\mathcal{R}^\Omega/1$, an irreducible $\mathcal{S}$-representation (McCullagh, 1999). For ordinal response models, the cumulative distribution function and its logistic transformation determine a point in $\mathcal{R}^{\Omega \setminus 0}$, an $\mathcal{I}_0$-representation that includes sub-representations. The relevant representation theory for these categories and their duals is given by McCullagh (1998).

7. Statistical tests and irreducible quotient spaces

Let $\mathcal{V}_0$ be a subspace of a vector space $\mathcal{V}$. For any $x \in \mathcal{V}$, the set of points $x + \mathcal{V}_0$, the translate of $\mathcal{V}_0$ by x, is called a coset of $\mathcal{V}_0$ in $\mathcal{V}$. The set of cosets of $\mathcal{V}_0$, the quotient space $\mathcal{V}/\mathcal{V}_0$, is a vector space of dimension $\dim(\mathcal{V}) - \dim(\mathcal{V}_0)$, and the mapping $x \mapsto x + \mathcal{V}_0$ from $\mathcal{V}$ to $\mathcal{V}/\mathcal{V}_0$ is the quotient projection. If $\mathcal{V}$ is an inner product space, the natural inner product in the quotient space is the inner product in $\mathcal{V}_0^{\perp}$, i.e. $\langle x + \mathcal{V}_0, \, y + \mathcal{V}_0 \rangle = \langle Qx, \, Qy \rangle$, in which Q is the orthogonal projection from $\mathcal{V}$ onto $\mathcal{V}_0^{\perp}$.

In statistics, quotient spaces, rather than subspaces, are intimately connected with statistical tests. In a randomized blocks design, for example, the usual test for equality of treatment effects is based on the difference between the residual sum of squares for the nested models

$$blocks \quad \text{and} \quad blocks + treatment.$$

We interpret this difference as the squared length of the projection of y on to the quotient space

$$(blocks + treatment)/blocks,$$

i.e. treatment eliminating additive block effects. In order to compute these orthogonal projections, it is necessary that $\mathcal{V}, \mathcal{V}_0$ be identified with subspaces of the inner product space $\mathcal{R}^{\mathcal{U}}$ of functions on the statistical units. In the case of generalized linear models, deviances are used in place of residual sums of squares.

In those rare cases in which the residual variance is known, projection on to the quotient space $\mathcal{V}/\mathcal{V}_0$ furnishes a direct test of the model $\mathcal{V}_0$ with $\mathcal{V}$ as alternative. In most cases, however, the residual variance is estimated on the basis of the residuals after projection on to $\mathcal{V}$ using a statistic of the form $\mathrm{RSS}(\mathcal{V})/(|\mathcal{U}| - \dim(\mathcal{V}))$. An F-test is then used in place of the direct chi-squared test.

An M-invariant space is said to be M-irreducible if it contains no non-trivial M-invariant subspace. In the context of conventional factorial models, this is not a useful concept because all M-invariant subspaces contain 1, which is the only M-irreducible subspace. However, the monoid also acts on quotient spaces, so it is natural to investigate the properties of quotient spaces in respect of this action. The M-invariant quotient spaces are ordered pairs of M-invariant subspaces $\mathcal{V}/\mathcal{V}_0$ in which $\mathcal{V}_0$ is a proper subspace of $\mathcal{V}$. For any M-invariant quotient space $\mathcal{V}/\mathcal{V}_0$, there may exist an M-invariant subspace $\mathcal{V}_1$, not equal to $\mathcal{V}_0$ or $\mathcal{V}$, such that $\mathcal{V}_0 \subset \mathcal{V}_1 \subset \mathcal{V}$. In that case $\mathcal{V}/\mathcal{V}_0$ contains the M-invariant subspace $\mathcal{V}_1/\mathcal{V}_0$. Otherwise, if no such strictly intermediate M-invariant subspace $\mathcal{V}_1$ exists, $\mathcal{V}/\mathcal{V}_0$ is an irreducible representation of the monoid. From a statistical perspective, the most interesting irreducible M-representations occur as quotient spaces, not as subspaces.

Reverting to factorial model formulae, although $A/1$ and $(A + B)/B$ are isomorphic M-irreducible representations, they represent different quotient spaces. For linear models, the squared projections on to the two spaces are

$$\mathrm{RSS}(1) - \mathrm{RSS}(A) \quad \text{and} \quad \mathrm{RSS}(B) - \mathrm{RSS}(A + B)$$

respectively, which are ordinarily different. Each M-irreducible quotient space has, in principle at least, a natural statistical interpretation as a mean square in an analysis of variance, or an analysis of deviance, table. The M-irreducibles are obtained automatically from the lattice diagram of spaces and subspaces.

In connection with the discussion at the end of section 3.3, it should be pointed out that the linear mapping φ^* that maps $A \otimes B \otimes C$ onto $A.B.C \subset \mathcal{R}^\mathcal{U}$, also maps quotient representations onto quotients of $A.B.C$. For example, the irreducible representation $(A \otimes B + C)/(A + B + C)$ is mapped onto the sub-quotient $(A.B + C)/(A + B + C)$ in $\mathcal{R}^\mathcal{U}$. By contrast, the orthogonal complement of $A + B + C$ in $(A.B + C) \subset \mathcal{R}^\mathcal{U}$ is not the image under φ^* of any M-representation in $A \otimes B \otimes C$. For this reason, each line in an analysis-of-variance table is logically associated with a quotient or sub-quotient representation, and only indirectly with a subspace of $\mathcal{R}^\mathcal{U}$, (McCullagh, 1999).

The preceding discussion applies equally to $\mathcal{F}$-representations and $\mathcal{I}$-representations.

8. Homologous factors

8.1 *Statistical examples*

Suppose that the observations are indexed by the same factor twice. To say the same thing in a different way, suppose that A and B are homologous factors ($A \cong B$) having the same set of levels, Ω, and the observations are indexed by the square $\Omega \times \Omega$. In a sense emphasized by the mathematics, A and B represent the same vector space. As subspaces of the space of functions on the square, however, they represent different vector subspaces having only the constant subspace in common. From the viewpoint of abstract algebra, one symbol is sufficient. For statistical and computational purposes, two symbols are essential in order to identify the distinct subspaces. For these reasons, the symbols $A^{\otimes 2}$ and $A \otimes B$ are used synonymously in this section.

Statistical designs of this sort may seem a little unusual at first glance, but examples are surprisingly common as the following list shows.

Citation analysis. A common practice in citation analyses is to begin by tabulating the number of citations y_{ij} in journal i to journal j (Stigler, 1994). In the present framework, A is the factor 'citing journal' and B is the homologous factor 'cited journal.' Within a particular subject area, two phenomena are usually apparent, a strong tendency for self-citations, and an academic hierarchy reminiscent of social classes. The applied journals speak only to the methodological journals, and the methodological journals speak only to the theoretical journals, and the theoretical journals speak only to god. A similar phenomenon occurs in geographical studies of migration, in which y_{ij} is the number of individuals, families, or other units, moving from region i to region j in a specified period. Short migrations within regions are much more numerous than longer migrations between regions. At the same time, certain regions become more popular, or economically more active, so that the relative migration rate from region i to region j depends on the difference in the growth rates of the two regions. In both cases the null model is that the expected value of Y_{ij} is the same as the expected value of Y_{ji}. The non-null model, usually of the Bradley-Terry form,

$$\log \mu_{ij} = \phi_{ij} + \alpha_i - \alpha_j,$$

where $\phi_{ij} = \phi_{ji}$, is designed to yield a linear ranking of the journals or regions. A matrix μ satisfying this condition is said to be quasi-symmetric (Caussinus, 1965; McCullagh, 1982). In citation studies, quasi-symmetry is known as the 'export scores model' (Stigler, 1994); in migration studies, the term 'gravity model' is used. (Stewart 1948, Upton 1985).

Plant breeding. In plant breeding studies, males of several varieties are crossed with females of the same varieties. The response in such an experiment is a function on the varieties crossed with themselves, i.e. an element in the tensor product space under study. The Mendelian inheritance model implies that the offspring of male variety i crossed with female variety j is, in a statistical sense, genetically identical to male variety j crossed with female variety i. In other words, the expected response $E(Y_{ij})$ lies in the symmetric subspace $\mathrm{sym}(A, B)$. A stronger form of the model implies that the genetic contribution of a male is independent of the variety of the female. In practice, the stronger condition is interpreted as

$$g(\mu_{ij}) = \alpha_i + \alpha_j$$

for some monotone transformation g. It should be emphasized that Mendel's laws do not imply additivity, and that favourable or unfavourable conjunctions of genes guarantee non-additivity in most cases. Nevertheless, this symmetric additive model often accounts for a large fraction of the systematic variation provided that incompatible pairs are excluded.

Tournament analysis. Consider a tournament in which teams or players compete with each other in pairs. The results of such a tournament can be compiled in a table indexed by the set of teams crossed with itself, with entries

$$y_{ij} = \begin{cases} 1 & \text{player } i \text{ beat player } j; \\ -1 & \text{player } j \text{ beat player } i; \\ 0 & \text{draw.} \end{cases}$$

By construction, y is skew-symmetric, so any model for y must focus on the invariant skew-symmetric subspaces of functions on the square. Two slightly different models for the probabilities $\pi_{ij} = \mathrm{pr}(Y_{ij} = 1)$ are as follows:

$$\log(\pi_{ij}/\pi_{ji}) = \alpha_i - \alpha_j$$
$$\mathrm{logit}\,\pi_{ij} = \alpha'_0 + \alpha'_i - \alpha'_j$$

If draws do not occur, $\alpha'_0 = 0$ and the models coincide. Both models involve the additive skew-symmetric subspace: the second subspace also includes the constant function and is denoted by $1 + \mathrm{alt}(A + B)$.

Matched pairs design. Consider a matched pairs design in which the response for each subject is a value in the finite set Ω. In a study of identical twins reared apart, for example, the response might be some form of preference, for say, clothing, colour, music, political party or sexual orientation. In a medical context, the response might be a measure of pain relief or improvement in medical condition. In each of these cases, the response for each pair is a point in $\Omega \times \Omega$, and y_{rs}, the number of pairs responding (r, s), are the components of a vector in $A^{\otimes 2}$.

Absence of a treatment effect implies symmetry, i.e. $\mu \in \mathrm{sym}^2(A)$. Various non-null models have been derived or proposed including quasi-symmetry and marginal homogeneity. Marginal homogeneity states that the expected number of treated members responding in category r is the same as the expected number of controls in category r, $\mu_{r\cdot} = \mu_{\cdot r}$. Clearly, marginal homogeneity and quasi-symmetry are vector subspaces of $A^{\otimes 2}$ invariant under permutation of the labels in Ω. However, as shown in the next section, quasi-symmetry is a monoid-invariant subspace but marginal homogeneity is not.

Diallel cross design. A slightly more complicated version of the plant-breeding example occurs when the males and females are both indexed by genotype. Then Ω is the set of alleles, and A is the vector space of functions on Ω. A function f on genotypes is a function on $\Omega \times \Omega$ such that $f(i, j) = f(j, i)$. In other words, $\text{sym}^2(A)$ may be identified with the vector space of real-valued functions on genotypes. Each observation is associated with an ordered pair (female, male) of genotypes

$$((r, s), (i, j)) \mapsto y_{r,s;i,j}$$

such that $y_{rs;ij} = y_{sr;ij} = y_{rs;ji}$. In other words, y is an element of the tensor product space $A^{\otimes 4}$ in the subspace $(\text{sym}^2(A))^{\otimes 2}$. An example of a model that might be contemplated for a design of this type is the additive symmetric model

$$g(\mu_{r,s;i,j}) = \alpha_r + \alpha_s + \alpha_i + \alpha_j$$

in which all four alleles contribute equally and additively without interaction. The single-locus Mendelian model is

$$\alpha_{ri} + \alpha_{sj} + \alpha_{rj} + \alpha_{si},$$

with $\alpha_{ij} = \alpha_{ji}$, an equally-weighted combination of the four genotypes. This is called the additive model with allelic interaction (Mather and Jinks, 1982). If it is thought that the sexes contribute unequally, an extended vector space containing functions of the form

$$\alpha_r + \alpha_s + \beta_i + \beta_j$$

or $\alpha_{rs} + \beta_{ij}$ might be tried. In practice, if the plants are self-sterile, any model may be restricted to the compatible pairs.

Other examples. The following list indicates that homologous factors arise in a large number of contexts.
1. The study of finger-print similarity in related individuals (Galton, 1892, p. 175);
2. Matrices of transition counts in Markov chains or more general processes with a discrete state space;
3. Allele-sharing of affected sib-pairs or triplets in the study of genetic diseases;
4. Social or occupational mobility tables;
5. Import-export analyses of international trade;
6. Studies in which the observational unit is the nuclear family indexed by the political or religious affiliation of both parents.

All of the models listed above determine vector subspaces not of the factorial type. Essential use is made of the fact that the factors are homologous. We now investigate the invariance properties of these vector spaces.

9. Decomposition of $A^{\otimes 2}$

9.1 *Group decomposition*

The principal difference between the tensor products $A \otimes B$ and $A^{\otimes 2}$ is the implied set of transformations or symmetries. For non-homologous factors, $A \otimes B$ is a representation of the product group in a space of dimension $|\Omega_a \times \Omega_b|$. By contrast, $A^{\otimes 2}$ is a representation of the symmetric group in a space of dimension $a^2 = |\Omega_a|^2$.

The standard decomposition of $A^{\otimes 2}$ by G-irreducibles proceeds in two steps. In the first step, the space is split into three non-overlapping subspaces, the diagonal subspace, the symmetric off-diagonal subspace, and the skew-symmetric subspace.

$$A^{\otimes 2} = \operatorname{diag}(A^{\otimes 2}) \oplus \operatorname{sym}^*(A) \oplus \operatorname{alt}^2(A).$$

Each of these invariant subspaces splits into irreducibles labelled by partitions of the number a (James and Liebeck, 1993). Thus

$$\operatorname{diag}(A^{\otimes 2}) \cong (a) \oplus (a-1,1) \cong A$$
$$\operatorname{sym}^*(A) \cong (a) \oplus (a-1,1) \oplus (a-2,2)$$
$$\operatorname{alt}^2(A) \cong (a-1,1) \oplus (a-2,1,1)$$
$$A^{\otimes 2} = (a)^{\oplus 2} \oplus (a-1,1)^{\oplus 3} \oplus (a-2,1) \oplus (a-2,1,1)$$

The two irreducibles labelled (a) are one-dimensional subspaces spanned by δ_{ij} and $1 - \delta_{ij}$ respectively, where δ_{ij} is Kronecker's delta. The three irreducibles labelled $(a-1,1)$ are isomorphic representations of dimension $a - 1$. In the decomposition given here, they represent the following subspaces:

$$(i,j) \mapsto \alpha_i \delta_{ij}, \quad (i,j) \mapsto (\alpha_i + \alpha_j)(1 - \delta_{ij}), \quad \text{and} \quad (i,j) \mapsto \alpha_i - \alpha_j,$$

in which $\sum \alpha_i = 0$.

A canonical set of unique G-invariant subspaces can be obtained by combining isomorphic representations into a single non-irreducible invariant subspace (Diaconis 1989, Bailey 1991). However, there are compelling reasons in statistical applications to avoid the canonical subspaces for purposes of modelling. First, the additive symmetric space, and the additive skew-symmetric space occur frequently, not in combination, in numerous statistical applications. Second, the canonical subspaces are ordinarily not closed under other important operations such as selection or monoid action. As models, they fail the test of marginality. In other words, the set of canonical group-invariant subspaces includes subspaces that are uninteresting as statistical models, and excludes spaces such as $\alpha_i - \alpha_j$ that occur frequently in a great variety of applications.

9.2 *$\mathcal{I}$-decomposition*

To obtain the $\mathcal{I}$-representations in $A^{\otimes 2}$, we first note that each of the subspaces $\operatorname{diag}(A^{\otimes 2})$, $\operatorname{sym}^*(A)$, and $\operatorname{alt}^2(A)$ is an $\mathcal{I}$-representation. The diagonal subspace is isomorphic with A, and contains the representation 1_d of constant diagonal functions. In the subspace sym^*, the only sub-representations are

$$1^* \cong (a), \quad \text{and} \quad \operatorname{sym}^{*+} \cong (a) \oplus (a-1,1).$$

In the skew-symmetric subspace, the only sub-representation is alt^+ of additive functions, isomorphic with $(a-1,1)$. Thus we obtain the decomposition shown in section 4.3. On account of multiplicities, however, this is not the complete list of $\mathcal{I}$-representations in $A^{\otimes 2}$.

9.3 *Homomorphisms and isomorphisms*

Let $\mathcal{W}$ and $\mathcal{X}$ be two representations of the category $\mathcal{C}$. A homomorphism $g\colon\mathcal{W}\to\mathcal{X}$ is a covariant functor such that each object map $g_\Omega\colon\mathcal{W}_\Omega\to\mathcal{X}_\Omega$ is a linear transformation. The arrow map taking $\psi^*\colon\mathcal{W}_{\Omega'}\to\mathcal{W}_\Omega$ in $\mathcal{W}$ to $\varphi^*\colon\mathcal{X}_{\Omega'}\to\mathcal{X}_\Omega$ in $\mathcal{X}$ satisfies $\varphi^*g_{\Omega'}=g_\Omega\psi^*$, so that the arrow diagram

$$
\begin{array}{ccc}
\mathcal{W}_{\Omega'} & \overset{\psi^*}{\longrightarrow} & \mathcal{W}_\Omega \\
g_{\Omega'}\downarrow & & g_\Omega\downarrow \\
\mathcal{X}_{\Omega'} & \overset{\varphi^*}{\longrightarrow} & \mathcal{X}_{\Omega'}
\end{array}
$$

is commutative. Note that there may be arrows in $\mathcal{X}$ that are not in the range of g.

If $\mathcal{W}$ is a sub-representation in $\mathcal{X}$, the insertion map from $\mathcal{W}$ into $\mathcal{X}$ is a homomorphism and g is an injection. For all of the representations considered here, if $\mathcal{X}$ is a sub-representation in $\mathcal{W}$, the quotient map $\mathcal{W}\to\mathcal{W}/\mathcal{X}$ is a surjective homomorphism. Apart from the zero map, there need not exist a homomorphism of $\mathcal{W}$ into a sub-representation $\mathcal{X}$. Consider, for example, the sub-representations of $\mathcal{C}=\mathcal{I}\times\mathcal{I}$ in $\mathcal{W}=A+B\subset A\otimes B$. Although $\mathcal{W}$ contains three proper sub-representations, it does not contain a complementary pair, so the only homomorphisms of $\mathcal{W}$ into itself are scalar multiples.

An isomorphism is a homomorphism such that each arrow in $\mathcal{X}$ is the image of one arrow in $\mathcal{W}$ and each object map g_Ω is an invertible linear transformation. For all of the representations considered in this paper, the kernel of the homomorphism $g\colon\mathcal{W}\to\mathcal{X}$ is a sub-representation in $\mathcal{W}$, and $g\mathcal{W}\subset\mathcal{X}$ is a representation isomorphic with $\mathcal{W}/\ker(g)$.

Let $g\colon\mathcal{W}\to\mathcal{X}$ and $h\colon\mathcal{W}\to\mathcal{X}$ be two homomorphisms of $\mathcal{W}$ into $\mathcal{X}$. Then every linear combination $\alpha g+\beta h$ is also a homomorphism of $\mathcal{W}$ into $\mathcal{X}$. Thus $(\alpha g+\beta h)\mathcal{W}$ is a sub-representation in $\mathcal{X}$, derived from the representations $g\mathcal{W}$ and $h\mathcal{W}$. This derived sub-representation must not be confused with the span $g\mathcal{W}+h\mathcal{W}$.

By way of example, let $\mathcal{C}=\mathcal{I}$, $\mathcal{W}=A=\mathcal{R}^\Omega$ and $\mathcal{X}=A^{\otimes 2}$. A vector w in $\mathcal{W}$ is a function assigning to each $i\in\Omega$ a real number w_i. A vector x in $\mathcal{X}$ is a function assigning to each ordered pair (i,j) in $\Omega\times\Omega$ a value x_{ij}. Let gw be the vector in $A^{\otimes 2}$ whose (i,j)-component is w_i, and let hw be the vector whose (i,j)-component is w_j. In other words, $g\mathcal{W}$ is the row factor, and $h\mathcal{W}$ is the column factor. Since $\ker(g)=\ker(h)=0$, both $g\mathcal{W}$ and $h\mathcal{W}$ are isomorphic with $\mathcal{W}$. Also, $(g+h)\mathcal{W}=\mathrm{sym}^+$, the symmetric additive subspace, $(g\cos\theta+h\sin\theta)\mathcal{W}=A_\theta$, and $(g-h)\mathcal{W}=\mathrm{alt}^+$. The kernel of $g-h$ is the subspace of constant functions, so $\mathrm{alt}^+\cong A/1$. Note that, although the dimensions are equal, the representation $1+\mathrm{alt}^+$ is equal to the span of two proper sub-representations, and thus cannot be $\mathcal{I}$-isomorphic with $\mathrm{sym}^+\cong A$.

9.4 *$\mathcal{F}$-decomposition*

The $\mathcal{F}$-representations are a subset of the $\mathcal{I}$-representations. In the decomposition given in section 4.2, there are five irreducible $\mathcal{F}$-representations, three occurring as subspaces and two as quotient spaces. Two of these are isomorphic: $\mathrm{alt}^+\cong\mathrm{sym}^+/1$ for $|\Omega|>2$. For each $\theta\in[0,\pi)$, the mapping from $\mathcal{R}^\Omega$ into $(\mathcal{R}^\Omega)^{\otimes 2}$ given by

$$
A_\theta=\{f_{rs}=\alpha_r\cos\theta+\alpha_s\sin\theta:\alpha\in\mathcal{R}^\Omega\}
$$

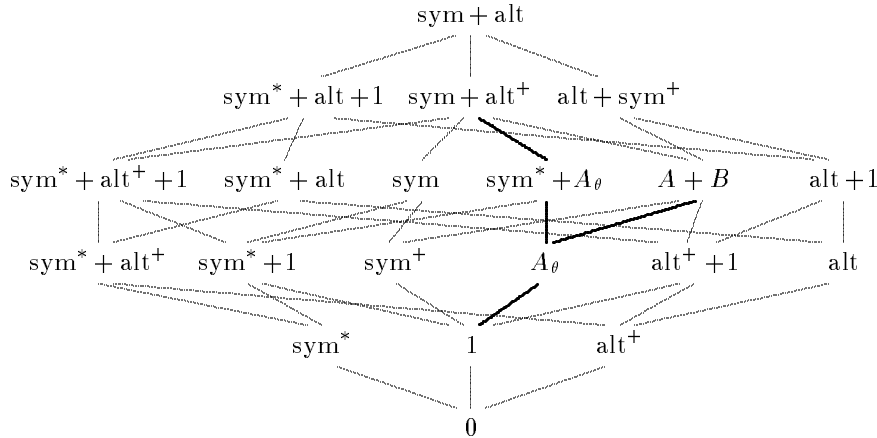is an $\mathcal{F}$-representation, so there is an infinite number of additive representations. Four of

Figure 2: *Hasse diagram for the lattice of $\mathcal{F}$-representations in $A^{\otimes 2}$. sym^* is the symmetric subspace with zero diagonal; sym is the full symmetric subspace; alt is the full skew-symmetric subspace; sym^+ and alt^+ are the additive symmetric and alternating subspaces; the remaining subspaces isomorphic with A have been telescoped into one, denoted by A_θ.*

these are of sufficient importance in statistical work to require special notation.

θ	Components	Notation
0	$(r,s) \mapsto \alpha_r$	A
$\pi/2$	$(r,s) \mapsto \alpha_s$	B
$\pi/4$	$(r,s) \mapsto \alpha_r + \alpha_s$	$\mathrm{sym}(A+B)$
$3\pi/4$	$(r,s) \mapsto \alpha_r - \alpha_s$	$\mathrm{alt}(A+B)$

The symmetric additive subspace occurs in certain plant-breeding models in which male and female parents contribute equally and additively to characteristics of the offspring (Yates, 1947; Griffing, 1956). The alternating additive subspace occurs in competition models, citation analyses (Stigler, 1994), migration studies, matched pairs designs (McCullagh, 1984) reversible Markov chains, and a number of other areas.

The complete set of $\mathcal{F}$-representations in $A^{\otimes 2}$ forms the lattice of subspaces illustrated in Fig. 2.

It is important to note that A and B are factors, by which we mean that there is a partition of the units into $|\Omega|$ blocks such that each function in A is constant on the blocks. Although $\mathrm{sym}(A+B) \cong A$, it is not possible to represent the additive symmetric subspace as a partition. The coarsest partition for which $\mathrm{sym}(A+B)$ is constant on blocks is the symmetry partition $(i,j) \sim (j,i)$ corresponding to the subspace $\mathrm{sym}(A,B)$. In addition, as a vector space of functions on the units, a factor has the property

$$A.A = \{fg : f,g \in A\} = A.$$

By contrast, the additive subspaces satisfy $\mathrm{sym}^+ . \mathrm{sym}^+ = \mathrm{sym}$, which includes sym^+, and $\mathrm{alt}^+ . \mathrm{alt}^+ = \mathrm{sym}^*$, which does not include alt^+. Finally, even though sym^* is not the vector subspace corresponding to any partition, $\mathrm{sym}^* . \mathrm{sym}^* = \mathrm{sym}^*$, showing that the property of closure under functional products is not unique to factors.

9.5 *Nested designs*

The structure of the units in a nested design is best described by the category $\mathcal{ND}$ in which each object is a finite set of units together with the equivalence relation $i \sim j$ if units i and j belong to the same block. The maps $\varphi : \Omega \to \Omega'$ in $\mathcal{ND}$ are injective for both blocks and units, i.e. genotypes and sibs in the example in section 2. In other words, $i \sim j$ in Ω if and only if $\varphi(i) \sim \varphi(j)$ in Ω'. The set of maps from Ω to itself in $\mathcal{ND}$ constitutes a group. When all blocks are of equal size, this group is known as the wreath product of symmetric groups. The sub-representations of $\mathcal{ND}$ in $A = \mathcal{R}^\Omega$ are $1 \subset B \subset A$, where B is the subspace of functions that are constant on each block.

Denote by (i, g) a sib with genotype $g(i)$, so that the elements of Ω^2 are ordered pairs $((i, g), (i', g'))$. Each of the subsets of Ω^2

$$\mathcal{O}_1 = \{i = i'\}; \quad \mathcal{O}_2 = \{i \neq i', g = g'\}, \quad \text{and} \quad \mathcal{O}_3 = \{g \neq g'\},$$

is an orbit conserved by $\mathcal{ND}$, so that $(\mathcal{R}^\Omega)^{\otimes 2}$ is the direct sum of the $\mathcal{ND}$-representations $\mathcal{R}^{\mathcal{O}_1} \oplus \mathcal{R}^{\mathcal{O}_2} \oplus \mathcal{R}^{\mathcal{O}_3}$. Now, $\mathcal{R}^{\mathcal{O}_1}$ is isomorphic with $\mathcal{R}^\Omega$, and the remaining two spaces split further into symmetric and skew-symmetric subspaces. The structure of a complete decomposition is as follows:

$$\mathcal{R}^{\mathcal{O}_1} : \quad A \supset B \supset 1$$
$$\mathcal{R}^{\mathcal{O}_2} : \quad \begin{cases} \mathrm{sym}^2(A) \supset \mathrm{sym}^+(A) \supset B \supset 1 \\ \mathrm{alt}^2(A) \supset \mathrm{alt}^+(A) \quad (\cong A/B). \end{cases}$$

The symmetric subspace of $\mathcal{R}^{\mathcal{O}_3}$ contains six proper sub-representations, four of which are 1, $\mathrm{sym}^+(B)$, $\mathrm{sym}^2(B)$ and $\mathrm{sym}^+(A)$. In addition, $\mathrm{sym}^+(A) + \mathrm{sym}^2(B)$ is a sub-representation in $\mathrm{sym}^+(A).\mathrm{sym}^2(B)$. The alternating subspace of $\mathcal{R}^{\mathcal{O}_3}$ contains five sub-representations, three of which are $\mathrm{alt}^+(B)$, $\mathrm{alt}^2(B)$ and $\mathrm{alt}^+(A)$. In addition, $\mathrm{alt}^+(A) + \mathrm{alt}^2(B)$ is a sub-representation in $\mathrm{alt}^+(A).\mathrm{sym}^2(B)$. Various multiplicities such as $\mathrm{alt}^+(B) \cong \mathrm{sym}^+(B)/1$ and $\mathrm{alt}^+(A) \cong \mathrm{sym}^+(A)/1$, give rise to an infinite number of rotated representations. Each model for the compatible pairs in section 2 is a sub-representation in $\mathcal{R}^{\mathcal{O}_3}$; each model for the incompatible pairs is a sub-representation in $\mathcal{R}^{\mathcal{O}_1} \oplus \mathcal{R}^{\mathcal{O}_2}$.

9.6 *Marginal homogeneity and dual representations*

To each category $\mathcal{C}$ there corresponds a dual, or opposite, category, $\mathcal{C}^\dagger$. The objects on which $\mathcal{C}^\dagger$ acts are the same as the objects on which $\mathcal{C}$ acts, but all arrows are reversed. Representations for $\mathcal{C}^\dagger$ follow immediately from the representations of $\mathcal{C}$. To each $\mathcal{C}$-representation $\mathcal{X}$ in $A^{\otimes k}$ there corresponds a dual $\mathcal{C}^\dagger$-representation $\mathcal{X}^0$ of linear functionals taking the value zero on $\mathcal{X}$. The dimension of $\mathcal{X}^0$ is equal to the co-dimension of $\mathcal{X}$.

Marginal homogeneity (MH) is a dual vector subspace of measures on $A^{\otimes 2}$, of dimension $a^2 - a + 1$ consisting of those measures P satisfying $P(\omega, \Omega) = P(\Omega, \omega)$ for each $\omega \subset \Omega$. In other words, the two marginal measures are equal. Marginal homogeneity is the dual subspace $(\mathrm{alt}^+)^0$, and is thus a representation of the dual category $\mathcal{F}^\dagger$.

To illustrate the meaning of the categories $\mathcal{F}$ and $\mathcal{F}^\dagger$, let Y be a random variable with values in Ω and probability measure P defined on the subsets of Ω. For any $\varphi : \Omega \to \Omega'$ $Y' = \varphi Y$ is a random variable with values in Ω'. The distribution of Y' is a measure $P \circ \varphi^\dagger$ on Ω', the restriction of P to the sigma algebra of sets $\varphi^{-1}(\omega)$ for $\omega \subset \Omega'$.

In the statistical literature on contingency tables, tests for marginal homogeneity, also called marginal symmetry, occur in the work of Stuart (1953, 1955) and in subsequent work by Bhapkar (1966), Kullback (1971), Fryer (1971), Mantel and Fleiss (1975), Plackett (1981), and Agresti (1983). Firth (1989) and Lipsitz, Laird and Harrington (1990) discuss ways to construct a basis in the space. For a review, see White, Landis and Cooper (1982). For the most part, however, marginal homogeneity is put forward as a criterion to be tested, not as a model to be entertained seriously in a scientific sense. Unlike symmetry or quasi-symmetry, no mathematical derivation has been proposed in the context of matched pairs designs where the test is most commonly used. One minor exception (Lindsey, 1995) occurs when the two margins refer to two time points of a process: *MH* is then equivalent to stationarity. A number of authors have criticized the use of *MH* as a model, for example the first edition of McCullagh and Nelder (1983, p. 125) and Nelder (1994). The reasoning given by McCullagh and Nelder in the context of matched pairs is that absence of a treatment effect implies symmetry in addition to marginal homogeneity. 'Marginal homogeneity corresponds to a hypothesis about a null main effect in the presence of interactions involving that effect. Hypotheses that violate the marginality principle ... are rarely of scientific interest ...'. The disappearance of this paragraph in the second edition was due to error, not change of heart.

Despite these reservations, marginal homogeneity cannot be ruled out as an uninteresting model in all circumstances. In an import-export analysis of international trade statistics, marginal homogeneity has the easily-understood meaning of nationally balanced trade figures. Exports and imports are best envisaged as measures than as functions, so the dual vector space is arguably relevant.

10. The subspace algebra of sym and alt

For purposes of constructing statistical models, we would like to have available a small number of operators, capable of being used in combination with existing operators, ('+' and '.'), that are sufficient to generate all $\mathcal{I}$-representations and product representations. This goal is not feasible at present even for fairly simple spaces such as the representations of $\mathcal{I}$ in $A^{\otimes 3}$, or the representations of $\mathcal{F} \times \mathcal{F}$ in $A^{\otimes 2} \otimes B^{\otimes 2}$. As a compromise, we describe a small number of operators capable of generating only the most useful $\mathcal{I}$-representations.

We suppose for the sake of discussion that we are dealing with representations of $\mathcal{I} \times \mathcal{I} \times \cdots$ in the tensor product space $A^{\otimes d_1} \otimes B^{\otimes d_2} \otimes \cdots$. In other words, factors $A, B, \ldots$ are non-isomorphic; factor A occurs in d_1 homologous copies, factor B in d_2 homologous copies, and so on. The complete design is thus the product set $\mathcal{U}_0 = \Omega_1^{d_1} \times \Omega_2^{d_2} \times \cdots$, although, in practice, some of these elements may be equivalent. The mapping φ from the observed units $\mathcal{U}$ into $\mathcal{U}_0$ determines the design. Certain elements in $\mathcal{U}_0$ may be missing from $\mathcal{U}$; others may be replicated. To any representation $\mathcal{X}$ in $\mathcal{R}^{\mathcal{U}_0}$, the pullback map determines a subspace $\mathcal{V} = \varphi^* \mathcal{X} \subset \mathcal{R}^{\mathcal{U}}$. If $\mathcal{X}_1, \mathcal{X}_2$ are isomorphic representations, we say that the subspaces $\mathcal{V}_j = \varphi^* \mathcal{X}_j$ in $\mathcal{R}^{\mathcal{U}}$ are isomorphic. Because of the dependence on the design, however, there is no guarantee that the subspaces $\mathcal{V}_j$ have the same dimension. In a citation study in which the citations are ascertained by source, the set of citing journals is usually a subset of the cited journals.

Let $\mathcal{X}_1, \mathcal{X}_2, \ldots$ be representations in $\mathcal{R}^{\mathcal{U}_0}$, all homomorphic with $\mathcal{X}$. The homorphisms are $g_r \colon \mathcal{X} \to \mathcal{X}_r$; in practice, these maps are implied by the context. The operator plus with

syntax
$$\text{plus}(\mathcal{X}_1, \ldots, \mathcal{X}_k, (c_1, \ldots, c_k))$$

signifies the subspace $(c_1 g_1 + \cdots + c_k g_k)\mathcal{X}$. In practice, $\mathcal{X}$ is determined by a model matrix X, and $\mathcal{X}_r$ by the model matrix $X_r = g_r X$. The space $(c_1 g_1 + \cdots + c_k g_k)\mathcal{X}$ is spanned by the columns of $c_1 X_1 + \cdots + c_k X_k$. As a subspace of $\mathcal{R}^{\mathcal{U}}$, $\text{plus}(\mathcal{V}_1, \ldots, \mathcal{V}_k, (c_1, \ldots, c_k))$ signifies the subspace $\varphi^*(c_1 g_1 + \cdots + c_k g_k)\mathcal{X}$ by restriction to the observed units. The model matrix is $c_1 V_1 + \cdots + c_k V_k$ in which $V_r = \varphi^* X_r$ by restriction to the observed units.

Two special cases of plus are sufficiently important to be given special names:

$$\text{symadd}(\mathcal{V}_1, \ldots, \mathcal{V}_k) = \text{plus}(\mathcal{V}_1, \ldots, \mathcal{V}_k, (1, \ldots, 1)),$$
$$\text{altadd}(\mathcal{V}_1, \mathcal{V}_2) = \text{plus}(\mathcal{V}_1, \mathcal{V}_2, (1, -1)).$$

In particular, if A, B are homologous factors, the implicit isomorphism gives $\text{symadd}(A, B) = \text{sym}(A + B)$ and $\text{altadd}(A, B) = \text{alt}(A + B)$. The vector spaces appearing as arguments must be isomorphic, but they need not be factors. For example, if A, B, C, D are homologous factors,

$$\text{symadd}\big(\text{altadd}(A, B), \text{altadd}(C, D)\big)$$

is the vector space of additive functions with components $\alpha_r - \alpha_s + \alpha_t - \alpha_u$. Because of the implied isomorphism, this is not the same subspace as

$$\text{symadd}\big(\text{altadd}(A, B), \text{altadd}(D, C)\big).$$

Although $\text{altadd}(A, B) = \text{altadd}(B, A)$ the implied isomorphism is not the identity but the negative of the identity.

By the symbol $\mathcal{V}_1.\mathcal{V}_2$ we understand the subspace tensor product, the span of all products fg with $f \in \mathcal{V}_1$ and $g \in \mathcal{V}_2$. Isomorphism is not implied.

By expressions such as $\text{sym}(\mathcal{V}_1, \mathcal{V}_2)$, $\text{diag}(\mathcal{V}_1, \mathcal{V}_2)$ and $\text{alt}(\mathcal{V}_1, \mathcal{V}_2)$ we understand the following.

(i) The subspaces $\mathcal{V}_j$ are isomorphic as representations.

(ii) The set of units $\mathcal{U}$ is is a subset of $\Omega \times \Omega$, i.e. each unit is an ordered pair $u = (r, s)$ with r and s in some set Ω. To each unit $u = (r, s)$ we associate the reverse unit $u' = (s, r)$.

(iii) Functions in $\mathcal{V}_1$ depend only on the first component of u; functions in $\mathcal{V}_2$ depend only on the second component.

The definitions are as follows.

$$\text{sym}(\mathcal{V}_1, \mathcal{V}_2) = \text{span}\{u \mapsto f(u)g(u) + f(u')g(u') : f \in \mathcal{V}_1, g \in \mathcal{V}_2\}$$
$$\text{alt}(\mathcal{V}_1, \mathcal{V}_2) = \text{span}\{u \mapsto f(u)g(u) - f(u')g(u') : f \in \mathcal{V}_1, g \in \mathcal{V}_2\}$$
$$\text{diag}(\mathcal{V}_1, \mathcal{V}_2) = \{f \in \mathcal{V}_1.\mathcal{V}_2 : u \neq u' \Rightarrow f(u) = 0\}$$
$$\text{sym}^*(\mathcal{V}_1, \mathcal{V}_2) = \{f \in \text{sym}(\mathcal{V}_1, \mathcal{V}_2) : u = u' \Rightarrow f(u) = 0\}.$$

Each of these can be extended to an arbitrary number of arguments. In particular, $\text{sym}^k(\mathcal{V}_1, \ldots, \mathcal{V}_k)$ is the span of the symmetrized products

$$u \mapsto \sum_{\pi \in S_k} f_1(\pi u) f_2(\pi u) \cdots f_k(\pi u)$$

with $f_j \in \mathcal{V}_j$ and summation over the permutations of $u = (r_1, \ldots, r_k)$.

To give a few examples, let A, B, C, D be homologous factors. Certain expressions such as $\mathrm{sym}\big(\mathrm{alt}(A,B), \mathrm{alt}(B,C)\big)$ are not permitted because of assumptions (ii) and (iii) above. The expression $\mathrm{sym}\big(\mathrm{alt}(A,B), \mathrm{alt}(C,D)\big)$ represents the vector space of functions f satisfying

$$f_{rstu} = f_{turs} = -f_{srtu}.$$

Finally, if $A \cong C$ and $B \cong D$, we have

$$\mathrm{sym}(A+B, C+D) = \mathrm{sym}(A,C) + \mathrm{sym}(A.D + C.B) + \mathrm{sym}(B,D),$$
$$\mathrm{alt}(A+B, C+D) = \mathrm{alt}(A,C) + \mathrm{alt}(A.D + C.B) + \mathrm{alt}(B,D).$$

A partial implementation in 'S' is described by Glonek and McCullagh (1997).

11. Example 2: a diallel cross design

In a study of the genetics of seed dormancy, Lane and Lawrence (1995) observed the germination rates of *Papaver rhoeas* seeds for various crosses of the wild type (W) with the 'Shirley' cultivar, (S). The observed germination rates, given for all 16 second-generation families in Table 3, constitute the data for our analyses. This design is unusually complete in that, the two types of F_1 crosses, female Shirley crossed with wild male (S, W) and wild female crossed with male Shirley (W, S), are distinguished.

Table 3. *Germination rates for second-generation Papaver rhoeas crosses*

Female parent	Male parent			
	(S,S)	(S,W)	(W,S)	(W,W)
(S,S)	96.0 ± 2.5	97.0 ± 0.9	95.3 ± 1.4	92.3 ± 2.6
(S,W)	87.2 ± 5.0	53.0 ± 4.4	50.5 ± 4.5	38.9 ± 5.1
(W,S)	90.0 ± 2.4	55.9 ± 4.8	60.0 ± 3.3	42.6 ± 4.4
(W,W)	98.3 ± 2.0	41.3 ± 5.1	42.2 ± 5.3	15.8 ± 6.0

Each of the four factors A, B, C, D has two levels, 'Shirley' and 'wild type'. Factor A is the cultivar of the mother's mother, B is the cultivar of the mother's father, C is the father's mother, and D is the father's father. The rows of the table are indexed by (A, B) and the columns by (C, D).

In order to construct a model, it is helpful to assume, at least tentatively, that dormancy is controlled by a single gene, that Shirley is homozygous of type SS, and the wild type is homozygous of type ss. Then the F_1 crosses (S, W) and (W, S) have genotypes Ss and sS, which under normal circumstances, are genetically indistinguishable. Under Mendelian inheritance, when genotype (i, j) is crossed with (r, s), the four mixed types (i, r), (j, s), (i, s) and (j, r), and only these types, occur in equal proportion. If duplicate genotypes occur in this list the proportions are added. The single-locus model thus leads to consideration of the subspace of functions of the form

$$(i, j, r, s) \mapsto \alpha_{ir} + \alpha_{js} + \alpha_{is} + \alpha_{jr}$$

in which α is symmetric. In terms of the factors A, B, C, D, we denote this space by

$$L_1 = \mathrm{symadd}\big(\mathrm{sym}(A,C),\ \mathrm{sym}(B,D),\ \mathrm{sym}(A,D),\ \mathrm{sym}(B,C)\big),$$

which has dimension $a(a+1)/2 = 3$. Mather and Jinks (1982) refer to this as the additive model with dominance. In the absence of dominance, if the response of the F_1 heterozygotes is equal to the average of the parents, we obtain the additive subspace

$$L_0 = \mathrm{symadd}(A,\ B,\ C,\ D)$$

of dimension equal to the number of levels.

The preceding derivations are based on a single-locus assumption and equal viability of genotypes. If either of these assumptions is not satisfied, in particular if the character under study is governed by genes at two or more loci, linked or unlinked, the standard genetic model leads to the subspace

$$L_2 = \mathrm{sym}(\mathrm{sym}(A,B),\ \mathrm{sym}(C,D))$$

of dimension $a(a+1)(a^2+a+2)/8 = 6$. In other words, regardless of the number of loci, the mean response depends only on the unordered pair of unordered pairs of parental genotypes.

A significant projection on to the space L_2/L_1 is interpreted as evidence either that dormancy is controlled by more than one locus, or that there are symmetric parental contributions of a non-Mendelian type, the latter being less plausible. A significant projection on to any space orthogonal to L_2 is an asymmetric, and therefore non-Mendelian, contribution. One such form of inheritance, known as imprinting, occurs when the response of genotype (i, r) differs from that of (r, i). Alleles are inherited in the standard manner, but their effect depends on whether the allele is maternally inherited or paternally inherited. Imprinting is characterized by a non-zero projection on to the space

$$I_1 = \mathrm{symadd}(\mathrm{alt}(A,C),\ \mathrm{alt}(B,D),\ \mathrm{alt}(A,D),\ \mathrm{alt}(B,C))$$

of dimension $a(a-1)/2$. The additive subspace of dimension $a-1$,

$$\begin{aligned}
I_1' &= \mathrm{symadd}(\mathrm{alt}(A+C),\ \mathrm{alt}(B+D),\ \mathrm{alt}(A+D),\ \mathrm{alt}(B+C)) \\
&= \mathrm{sym}(\mathrm{alt}(A+C)+\mathrm{alt}(B+D)) \\
&= \mathrm{alt}(\mathrm{sym}(A+B)+\mathrm{sym}(C+D)),
\end{aligned}$$

spanned by functions of the form $\alpha_i + \alpha_j - \alpha_k - \alpha_l$, coincides with I_1 when $a = 2$. Note that $L_2 \cap I_1 = 0$ and that

$$L_1 + I_1 = \mathrm{sym}^4(A.B + B.D + A.D + B.C).$$

Another type of non-Mendelian effect that can be detected occurs when parental types (i, r) and (r, i) contribute unequally to the character of the offspring. Such an effect is characterized by a non-zero projection on to the space

$$I_2 = \mathrm{alt}(A,B) \otimes \mathrm{alt}(C,D)$$

of dimension $a^2(a-1)^2/4$. The subspaces

$$\mathrm{sym}(\mathrm{alt}(A+B), \mathrm{alt}(C+D)) \subset \mathrm{sym}(\mathrm{alt}(A,B), \mathrm{alt}(C,D)),$$

of dimensions $a-1$ and $a(a-1)(a^2-a+2)/8$, coincide with I_2 when $a=2$.

A third type of non-Mendelian effect occurs if the two parents contribute relevant genetic material unequally to the offspring. Such a contribution may be of the type $A.B + C.D$ or the rotated factor $\alpha_{ij}\cos\theta + \alpha_{kl}\sin\theta$ for some θ. In the context of a multi-locus genetic contribution, only the skew-symmetric projection, or quotient modulo L_2, of either space is relevant. Maternal, or paternal effects are characterized by a non-zero projection on to the additive skew-symmetric space

$$P = \mathrm{alt}(A.B + C.D)$$

of dimension $a^2 - 1$. This space can be decomposed into two non-overlapping subspaces

$$P = \mathrm{alt}(\mathrm{sym}(A,B) + \mathrm{sym}(C,D)) + \mathrm{alt}(\mathrm{alt}(A,B) + \mathrm{alt}(C,D))$$

of dimensions $(a^2 + a - 2)/2$ and $a(a-1)/2$ respectively. The first component contains the subspace $I_1' = \mathrm{alt}(\mathrm{sym}(A+B) + \mathrm{sym}(C+D))$ of dimension $a-1$, associated with imprinting. The second component contains the additive doubly-skew subspace $\mathrm{alt}(\mathrm{alt}(A+B) + \mathrm{alt}(C+D))$, also of dimension $a-1$, and coincides with this space if $a=2$. The latter space consisting of functions of the form $(\alpha_i - \alpha_j) - (\alpha_k - \alpha_l)$, coincides with $\mathrm{alt}(\mathrm{sym}(A+D), \mathrm{sym}(B+C))$.

Table 4 gives a summary of the fits obtained from a weighted untransformed analysis by projecting on to various monoid-invariant subspaces. The first five subspaces correspond to the single-locus model and the various extensions discussed above. It is immediately apparent that the single-locus model does not fit, and that the extended model including the subspace $\mathrm{sym}(A,B) + \mathrm{sym}(C,D)$, with $RSS = 24.12$ on 10 d.f, results in a considerably improved fit. Most of this improvement in fit comes from the two degrees of freedom in the additive skew-symmetric subspace, so the evidence for a maternal effect seems to be strong. The F-ratio is $15.6/2.4$, or 6.5 on 2,10 degrees of freedom.

Table 4. *Weighted residual sums of squares for various models*

Model	RSS	d.f.
$L_1 = \mathrm{sym}^4(\mathrm{sym}(A,C) + \mathrm{sym}(A,D) + \mathrm{sym}(B,C) + \mathrm{sym}(B,D))$	69.72	13
$L_1 + I_1$	58.94	12
$L_1 + \mathrm{sym}(\mathrm{sym}(A,B) + \mathrm{sym}(C,D))$	55.27	12
$L_1 + \mathrm{sym}(\mathrm{sym}(A,B) + \mathrm{sym}(C,D)) + \mathrm{alt}(\mathrm{sym}(A+B) + \mathrm{sym}(C+D))$	37.28	11
$L_1 + \mathrm{sym}(\mathrm{sym}(A,B) + \mathrm{sym}(C,D)) + \mathrm{alt}(\mathrm{sym}(A,B) + \mathrm{sym}(C,D))$	24.12	10
$L_2 = \mathrm{sym}(\mathrm{sym}(A,B), \mathrm{sym}(C,D))$	15.31	10
$L_2 + I_1$	9.53	9
$L_2 + I_2$	14.19	9
$L_2 + \mathrm{alt}(\mathrm{sym}(A,B) + \mathrm{sym}(C,D))$	5.31	8
$L_2 + P$	3.28	7
$\mathrm{sym}(A.B, C.D)$	8.99	6

Abandoning the single-locus assumption, the F-test for imprinting using the spaces L_2 and $L_2 + I_1$, yields an F-ratio of 5.78/1.06 on 1,9 degrees of freedom, and a p-value of 4.4%. In the absence of an alternative explanation, these data appear to provide moderate evidence for imprinting. By contrast, the projection on to I_2 is negligible, so there is no evidence that the offspring of parental types (i, r) and (r, i) differ in the character under study. The overall test for additive non-Mendelian parental effects yields an F-ratio of 4.01/0.47 on (3,7) degrees of freedom, or a p-value of 1%.

The fitted values under the multi-locus Mendel model L_2, and under the extended model, $L_2 + P$, with additive parental effects, are as follows.

$$
\begin{matrix}
96.00 & 95.74 & 95.74 & 90.42 \\
95.74 & 55.73 & 55.73 & 41.34 \\
95.74 & 55.73 & 55.73 & 41.34 \\
90.42 & 41.34 & 41.34 & 15.80
\end{matrix}
\quad \text{and} \quad
\begin{matrix}
96.00 & 97.04 & 95.07 & 92.87 \\
88.27 & 55.78 & 53.82 & 38.96 \\
90.24 & 57.74 & 55.78 & 40.92 \\
88.96 & 43.82 & 41.85 & 15.80
\end{matrix}.
$$

Because of the unequal weights, the difference between these fitted values does not lie in the alternating subspace P.

Acknowledgements

References

Agresti, A. (1983) Testing marginal homogeneity for ordinal categorical variables. *Biometrics* **39**, 505–510.

Bailey, R.A, (1991) Strata for randomized experiments (with discussion). *J. R. Statist. Soc.* B **53**, 27–78.

Bailey, R.A. (1996) Orthogonal partitions in designed experiments. *Design codes and cryptography* **8**, 45–77.

Bhapkar, V.P. (1966) A note on the equivalence of two test criteria for hypotheses in categorical data. *J. Am. Statist. Assoc.* **61**, 228–35.

Bloomfield, P. (1974) Linear transformations for multivariate binary data. *Biometrics* **30**, 609–617.

Box, G.E.P. and Cox, D.R. (1964) An analysis of transformations (with discussion). *J. R. Statist. Soc.* B **26**, 211–252.

Bradley, R.A. and Terry, M.A. (1952) Rank analysis of incomplete block designs. *Biometrika* **39**, 324–345.

Cox, D.R. (1972) The analysis of multivariate binary data. *Appl. Statist.* **21**, 113–120.

Cox, D.R. and Snell, E.J. (1981) *Applied Statistics*. Chapman and Hall, London.

Caussinus, H. (1965) Contribution à l'analyse statistique des tableaux de correlation. *Ann. Fac. Sci. Univ. Toulouse* **29**, 77–182.

Dawid, A.P. (1977) Invariant distributions and analysis of variance models. *Biometrika* **64**, 291–297.

Dawid, A.P. (1988) Symmetry models and hypotheses for structured data layouts (with discussion). *J. R. Statist. Soc.* B **50**, 1–34.

Diaconis, P. (1988) *Group Representations in Probability and Statistics.* IMS Lecture Notes **11**, Hayward, CA.

Diaconis, P. (1989) A generalization of spectral analysis with application to ranked data. *Ann. Statist.* **17**, 949–79.

Firth, D. (1989) Marginal homogeneity and the superposition of latin squares. *Biometrika* **76**, 179–182.

Fryer, J.G. (1971) On the homogeneity of the marginal distributions of a multidimensional contingency table. *J. R. Statist. Soc.* A **134**, 368–371.

Galton, F. (1892) *Finger Prints.* London, Macmillan.

Glonek, G. and McCullagh, P. (1997) An extended model syntax for designs involving homologous factors. Technical Report 448, Department of Statistics, University of Chicago.

Griffing, B. (1956) Concept of general and specific combining ability in relation to diallel crossing systems. *Austral. J. Biol. Sci.* **9**, 463–93.

Hamada, M. and Wu, C.F.J (1992) Analysis of designed experiments with complex aliasing. *J. Quality Technology* **24**, 130–137.

James, G. and Liebeck, M. (1993) *Representations and Characters of Groups.* Cambridge University Press.

Kullback, S. (1971) Marginal homogeneity of multidimensional contingency tables. *Ann. Math. Statist.* **42**, 594–606.

Lane, M.D. and Lawrence, M.J. (1995) Genetics of seed dormancy in *Papaver rhoeas.* *Heredity* **75**, 84–91.

Lauritzen, S.L. (1996) *Graphical Models.* Oxford University Press.

Lindsey, J.K. (1995) The uses and limits of linear models. *Statistics and Computing* **5**, 87–98.

Lipsitz, S.R., Laird, N. and Harrington, D.P. (1990) Finding the design matrix for the marginal homogeneity model. *Biometrika* **77**, 353–358.

MacLane, S. (1971) *Categories for the Working Mathematician.* Springer, New York.

MacLane, S. and Birkhoff, G. (1988) *Algebra.* New York, Chelsea.

Mantel, N. and Fleiss, J.L. (1975) The equivalence of the generalized McNemar tests for marginal homogeneity in 2^3 and 3^2 tables. *Biometrics* vol31 727–729.

Mather, K. and Jinks, J.L. (1982) *Biometrical Genetics.* Chapman and Hall, London.

McCullagh, P. (1980) Regression models for ordinal data (with discussion). *J. R. Statist. Soc.* B **42**, 109–142.

McCullagh, P. (1982) Some applications of quasisymmetry. *Biometrika* **69**, 303–8.

McCullagh, P. and Nelder, J.A. (1983) *Generalized Linear Models.* Chapman and Hall, London.

McCullagh, P. (1998) Category representations and factorial models. Technical Report 475, Department of Statistics, University of Chicago.

McCullagh, P. (1999) Quotient spaces and statistical models. *Can. J. Statist.* (to appear).

Nelder, J.A. (1977) A reformulation of linear models, (with discussion). *J. R. Statist. Soc.* A **140**, 48–77.

Nelder, J.A. (1994) The statistics of linear models: back to basics. *Statistics and Computing* **4**, 221–234.

Plackett, R.L. (1981) *The Analysis of Categorical Data.* Griffin, London.

Serre, J.-P. (1977) *Linear Representations of Finite Groups.* Springer-Verlag, New York.

Speed, T.P. (1987) What is an analysis of variance? (with discussion). *Ann. Statist.* **15**, 885–941.

Speed, T.P. and Bailey, R.A. (1987) Factorial dispersion models. *Int. Statist. Rev.* **55**, 261–277.

Stewart, J.Q. (1948) Demographic gravitation: Evidence and applications. *Sociometry* **1**, 31–58.

Stigler, S. (1994) Citation patterns in the journals of statistics and probability. *Statistical Science* **9**, 94–108.

Stuart, A. (1953) The estimation and comparison of strengths of association in contingency tables. *Biometrika* **40**, 105–10.

Stuart, A. (1955) A test for homogeneity of the marginal distributions in a two-way classification. *Biometrika* **42**, 412–16.

Tjur, T. (1984) Analysis of variance models in orthogonal designs (with discussion). *Int. Statist. Rev.* **52**, 33–81.

Upton, G.J.G. (1985) Modelling cross-tabulated regional data. In *Measuring the Unmeasurable* Ed. P. Nijkamp. NATO Scientific Publications, 1985, 197–218.

White, A.A., Landis, J.R., and Cooper, M.M. (1982) A note on the equivalence of several marginal homogeneity test criteria for categorical data. *Int. Statist. Rev.* **50**, 27–34.

Wilkinson, G.N. and Rogers, C.E. (1973) Symbolic description of factorial models for analysis of variance. *Appl. Statist.* **22**, 392–9.

Yates, F. (1947) Analysis of data from all possible reciprocal crosses between a set of parental lines. *Heredity* **1**, 287–301.