Steen Arne Andersson

The Lattice Structure of Orthogonal Linear Models and Orthogonal Variance Component Models

Preprint
December
1987

11

Institute of Mathematical Statistics
University of Copenhagen

ISSN 0902-8846

Steen Arne Andersson

THE LATTICE STRUCTURE OF ORTHOGONAL LINEAR MODELS AND ORTHOGONAL VARIANCE COMPONENT MODELS

Preprint 1987 No. 11

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

December 1987

THE LATTICE STRUCTURE OF ORTHOGONAL LINEAR MODELS AND ORTHOGONAL VARIANCE COMPONENT MODELS

BY

STEEN ARNE ANDERSSON

INSTITUTE OF MATHEMATICAL STATISTICS

UNIVERSITY OF COPENHAGEN

This research was supported in part by the Danish Research Council and by U.S. National Science Foundation Research Grant No. DMS 86-03489 at the Department of Statistics, University of Washington.

SUMMARY.

Tjur (1984) showed that an orthogonal (= balanced) analysis of variance (ANOVA) design may be described and analyzed in terms of an associated factor structure diagram. In this paper an extended class of orthogonal designs is defined and studied, the class of geometrically orthogonal (g.o.) designs of linear regression models, which includes all well-behaved ANOVA and regressions designs. It is shown that such designs may be characterized and analyzed most naturally in terms of the lattice structure of $\mathcal{L}$, the family of regression subspaces in the design. For example, a design is g.o. only if $\mathcal{L}$ is distributive, and the ANOVA table is determined by the contrast subspaces indexed by $J(\mathcal{L})$, the set of join-irreducible elements of $\mathcal{L}$. Furthermore, any g.o. design may be extended in a natural way to a family of canonical variance component (c.v.c.) models, called a geometrically orthogonal variance component design, whose structure and analysis are also determined by $\mathcal{L}$ and $J(\mathcal{L})$. A necessary and sufficient condition is given for a random effect modelassociated with the design to be a c.v.c. model, hence well-behaved.

AMS 1980 subject classification. Primary 62H05; secondary 62J05.

Key words and phrases. Analysis of variance, ANOVA table, linear regression model, geometrically orthogonal design, distributive lattice, partially ordered set, Möbius function, join-irreducible elements, variance component model, random effect model, normal distribution.

0. INTRODUCTION.

Linear regression models and variance component models lie near the center of statistical theory and practice. This paper presents a unified framework for the study of so-called orthogonal (or balanced) designs consisting of models of these types. All analysis of variance (ANOVA) designs with orthogonal factors and all well-behaved random effect models are included in the framework.

As mentioned in the Summary, this investigation may be regarded as an extension of the paper by Tjur (1984), who studied orthogonal ANOVA designs in terms of their factor structure diagrams. In the more general case of geometrically orthogonal (g.o.) linear model designs considered here, the role of this diagram is assumed by the partially ordered set (poset) $J(\mathcal{L})$ consisting of all join-irreducible elements of the lattice $\mathcal{L}$ of linear subspaces in the design. All information needed for the statistical analysis of the models in the design, such as the structure of its ANOVA table, is determined by $J(\mathcal{L})$. Furthermore, this statement is also true for the extended design of canonical variance component (c.v.c.) models determined by the original design.

The main mathematical concepts used are those of finite partially ordered sets (posets), finite distributive lattices, and finite-dimensional vector spaces. These concepts are very simple and the complications, if any, are of a combinatorial character. By disregarding all extraneous structure which is not essential for the definition and analysis of a statistical problem, one is able to obtain a mathematically efficient formulation of the problem, which we refer to as the "invariant" formulation. This formulation (hopefully) leads to a precise characterization of the class of models with the desired statistical properties, and to a unified and efficient mathematical analysis of the models. All pertinent definitions and results are directly suggested by the invariant formulation. For example, we shall obtain a definition of the ANOVA table, which for the case of analysis of variance with orthogonal factors, is slightly different from the definition in Tjur (1984). We hope that our approach is in the spirit of Bailey's points (i)-(iii) in her discussion of Tjur's paper (bottom of p.73)

The necessary concepts from the theory of partially ordered sets and lattices are presented in Section 1. The main reference for this section is Grätzer (1978). We have omitted all proofs of standard results. On the other hand, the one-to-one correspondence between the categories of finite partially ordered sets and finite distributive lattices is treated (Theorem 1.2, Proposition 1.2 and 1.3) because of its importance for our study of variance component models. Furthermore, we prove a useful condition for a distributive lattice to be finite (Proposition 1.1).

In Section 2, the results from Section 1 are applied to the lattice $\mathcal{L}$ of subspaces of a finite-dimensional vector space V. The main aim of this section is to describe the decomposition of a vector space into a direct sum determined by a distributive lattice $\mathcal{L}$ of subspaces (Theorem 2.1 and 2.2). It is shown that $\mathcal{L}$ is distributive if and only if there exists an inner product δ on V such that the subspaces in $\mathcal{L}$ are geometrically orthogonal (perpendicular) with respect to δ (Proposition 2.1 and 2.2). This existence of an inner product adapted to the lattice $\mathcal{L}$ is used not only in the present paper but also in a forthcoming paper on normal models given by conditional independence with respect to a distributive lattice of subspaces (Andersson and Perlman (1988)).

In Section 3 the invariant formulation and solution of an ordinary normal linear regression model is briefly reviewed. A geometrically orthogonal design of linear models is defined, together with the associated decomposition of the observation space into a direct sum of independent components and the ANOVA table. For the special case of an orthogonal ANOVA design (ie. a design generated by orthogonal factors), we compare our treatment to that of Tjur (1984). The section concludes with a series of examples.

The class of canonical variance component (c.v.c.) models extending a geometrically orthogonal design of linear models is defined and analyzed in Section 4. A necessary and sufficient condition that a random effect model be a c.v.c. model is derived in Theorem 4.1. For the special case of a completely balanced multiway ANOVA design we compare our formulation and results to those of Jensen (1979) and make similar comparisons with the work of Tjur (1984) in the case of a balanced (= orthogonal) ANOVA design. It is shown that the extension presented in this section includes some new interesting examples of variance component models. Finally, we also discuss the question of the statistical interpretability of random effect models in general, and of our canonical variance component models in particular.

Any list of references that one couldt readily compile for the statistical topics of linear models, analysis of variance, and variance component models would be far from comprehensive. Furthermore our formulation and point of view are somewhat different from those of most authors.

The reader is best advised to begin by referring to the paper by Tjur (1984) mentioned above, the discussions therein by Bailey, Speed, and Wynn, and the papers in the combined lists of references, in particular Speed and Bailey (1982). Our approach and ideas relate most closely to those of Jensen (1979) and Tjur (1984).

1. POSETS AND DISTRIBUTIVE LATTICES.

1.1. Posets.

A set $\mathscr{P}$ equipped with an ordering relation $\leq$ which is

(P1) reflexive: $\forall x \in \mathscr{P}: x \leq x$

(P2) antisymmetric: $\forall x,y \in \mathscr{P}: x \leq y$ and $y \leq x \Rightarrow x = y$

(P3) transitive: $\forall x,y,z \in \mathscr{P}: x \leq y$ and $y \leq z \Rightarrow x \leq z$

is called a partially ordered set or simply a poset. We use the notation $x < y$ if $x \leq y$ and $x \neq y$, $x,y \in \mathscr{P}$.

For a subset S of $\mathscr{P}$, $x \in \mathscr{P}$ is an upper bound (lower bound) of S if $y \leq x$ ($x \leq y$) $\forall y \in S$; x is a supremum (infimum) of S if $x \leq z$ ($z \leq x$) $\forall$ upper (lower) bound z of S. If a supremum (infimum) of S exists, it is unique and it is denoted by sup S (inf S). If it exists, the element sup $\mathscr{P}$ (inf $\mathscr{P}$) is called the unit (zero) element and is denoted by 1 (0); in this case $\mathscr{P}$ is called a poset with unit (poset with zero).

A finite sequence $x_1 < x_2 < \cdots < x_n$ of elements from $\mathscr{P}$ is called a chain of length n. If there exists $n \in \mathbb{N} = \{1,2,\cdots\}$ such that every chain in $\mathscr{P}$ has length less than n, then $\mathscr{P}$ is said to have finite length.

A mapping $\psi:\mathscr{P}_1 \to \mathscr{P}_2$ between two posets is called increasing or a (poset) homomorphism if $\forall x,y \in \mathscr{P}_1$: $x \leq y \Rightarrow \psi(x) \leq \psi(y)$. A composition of homomorphisms is a homomorphism and the identity mapping of a poset onto itself is a homomorphism. If ψ is bijective and $\psi^{-1}:\mathscr{P}_2 \to \mathscr{P}_1$ is a homomorphism then ψ is called a (poset) isomorphism and $\mathscr{P}_1$ and $\mathscr{P}_2$ are said to be isomorphic.

If $(\mathcal{P}_i | i \in I)$ is a family of posets then the product $X(\mathcal{P}_i | i \in I)$ equipped with the obvious (component-wise) ordering relation is itself a poset.

Any subset $\mathcal{P}_o \subseteq \mathcal{P}$ equipped with the restriction of the ordering $\leq$ on $\mathcal{P}$ is itself a poset, called a subposet, and the embedding $u: \mathcal{P}_o \to \mathcal{P}$ is a poset homomorphism.

An element $x \in \mathcal{P}$ is said to cover an element $y \in \mathcal{P}$ if $x > y$ and there is no $z \in \mathcal{P}$ such that $x > z > y$. This concept is needed to describe the representation of a finite poset by a (directed) graph. The points of the graph correspond to the elements of $\mathcal{P}$. A line connects two points if and only if the left-most point covers the other one (cf. Figures 3.1 - 3.15).

1.2. The Möbius function of a poset.

Let $\mathcal{P}$ be a finite poset. The Möbius function $\mu: \mathcal{P} \times \mathcal{P} \to \mathbb{Z} := \{0,1,-1,2,-2,3,-3,\cdots\}$ of $\mathcal{P}$ is defined as

$$\mu(x,y) = \begin{cases} 1 & \text{for } x = y \\ -\Sigma(\mu(x,z) | x \leq z < y) & \text{for } x < y \end{cases}$$

and 0 otherwise.

Lemma 1.1. (Möbius Inversion Formula for a finite poset). Let f and g be two functions defined on the finite poset $\mathcal{P}$ and assuming values in the same vector space. Then

$$f(x) = \Sigma(g(y) | y \leq x), \qquad x \in \mathcal{P},$$

if and only if

$$g(x) = \Sigma(\mu(y,x) f(y) | y \leq x), \quad x \in \mathcal{P}.$$

Proof: The matrix $(\mu(x,y)\,|\,(x,y)\in\mathcal{P}\times\mathcal{P})$ is the inverse of the matrix $(e(x,y)\,|\,(x,y)\in\mathcal{P}\times\mathcal{P})$, where $e(x,y) = 1$ if $y \leq x$ and 0 otherwise.□

As a space-saving convention, when we describe the Möbius function μ for a particular $\mathcal{P}$, we will only specify the values of $\mu(x,y)$ for those $x < y$, $x,y \in \mathcal{P}$, such that $\mu(x,y) \neq 0$.

Remark 1.1. If μ_i is the Möbius function for a finite poset $\mathcal{P}_i$, $i \in I$, and $|I| < \infty$, then the Möbius function μ for the product $\mathcal{P} = X(\mathcal{P}_i\,|\,i\in I)$ is given by $\mu((x_i\,|\,i\in I),(y_i\,|\,i\in I)) = \Pi(\mu(x_i,y_i)\,|\,i\in I)$.

The Möbius function for a finite chain poset $\{x_1 < x_2 < \cdots < x_n\}$ is specified by $\mu(x_i,x_{i+1}) = -1$, $i = 1,\ldots,n-1$.

If J is a finite set and $\mathcal{D}(J)$ denotes all subsets of J, then $\mathcal{D}(J)$ becomes a poset under the relation $\subseteq$. Since $\mathcal{D}(J)$ and the product $\{0,1\}^J$ are isomorphic as posets and $\{0,1\}$ is a chain, the Möbius function for $\mathcal{D}(J)$ is readily obtained as $\mu(A,B) = (-1)^{|B|-|A|}$, $A \subset B$, $A,B \in \mathcal{D}(J)$, where $|D|$ denotes the number of elements in $D \in\mathcal{D}(J)$.□

1.3. Lattices.

A set $\mathcal{L}$ equipped with two binary operations $\wedge$ and $\vee$ called meet and join, respectively, is called a lattice if $\wedge$ and $\vee$ are

(L1) idempotent: $\forall x \in \mathcal{L}$: $x \wedge x = x$ and $x \vee x = x$

(L2) commutative: $\forall x,y \in \mathcal{L}$: $x \wedge y = y \wedge x$ and $x \vee y = y \vee x$

(L3) associative: $\forall x,y,z \in \mathcal{L}$: $(x \wedge y) \wedge z = x \wedge (y \wedge z)$ and

$$(x \vee y) \vee z = x \vee (y \vee z)$$

and satisfy the

(L4) absorption identity:

$$\forall x,y \in \mathcal{L}: x \wedge (x \vee y) = x \text{ and } x \vee (x \wedge y) = x.$$

The property (L3) allows us to write expressions involving only $\wedge$'s or only $\vee$'s without using parentheses.

Every lattice $\mathcal{L}$ can be considered as a poset in the following natural way. If one defines the relation $\leq$ on $\mathcal{L}$ by $x \leq y$ iff $x \wedge y = x$ (or equivalently $x \leq y$ iff $x \vee y = y$) then $\mathcal{L}$ equipped with $\leq$ becomes a pose, denoted by $P(\mathcal{L})$ (or simply $\mathcal{L}$) with the additional property

(P4) For all non-empty finite subsets $S \subseteq \mathcal{L}$ the elements sup S and inf S exist.

Note that (P4) is equivalent to the condition that $\sup\{x,y\}$ and $\inf\{x,y\}$ exist for every two-point subset $\{x,y\} \subseteq \mathcal{L}$.

On the other hand, if $\mathcal{P}$ is a poset with the additional property (P4) then $\mathcal{P}$ equipped with the binary operations $\wedge$, $\vee$ defined by $x \wedge y = \inf\{x,y\}$ and $x \vee y = \sup\{x,y\}$, becomes a lattice denoted by $L(\mathcal{P})$ (or simply $\mathcal{P}$). It is easy to see that $L(P(\mathcal{L})) = \mathcal{L}$ and $P(L(\mathcal{P})) = \mathcal{P}$.

A finite lattice can be represented by a meet-and-join table, but since it is also a poset it can also be represented by a graph as described above.

A mapping $\varphi: \mathcal{L}_1 \rightarrow \mathcal{L}_2$ between two lattices is called a (<u>lattice</u>) <u>homomorphism</u> if $\forall x,y \in \mathcal{L}_1$: $\varphi(x \vee y) = \varphi(x) \vee \varphi(y)$ and $\varphi(x \wedge y) = \varphi(x) \wedge \varphi(y)$. The composition of homomorphisms is a homomorphism and the identity mapping of a lattice onto itself is a homomorphism. If φ is bijective then $\varphi^{-1}:\mathcal{L}_2 \to \mathcal{L}_1$ is also a homomorphism; φ is called a (lattice) <u>isomorphism</u> and $\mathcal{L}_1$ and $\mathcal{L}_2$ are said to be <u>isomorphic</u>.

If $\varphi:\mathcal{L}_1 \to \mathcal{L}_2$ is a lattice homomorphism then $\varphi:P(\mathcal{L}_1) \to P(\mathcal{L}_2)$ is a poset homomorphism. It is not true, however, that if $\psi:\mathcal{P}_1 \to \mathcal{P}_2$ is a poset homomorphism between posets $\mathcal{P}_1$, $\mathcal{P}_2$ satisfying (P4) then $\psi:L(\mathcal{P}_1) \to L(\mathcal{P}_2)$ is a lattice homomorphism.

If $(\mathcal{L}_i|i\in I)$ is a family of lattices then the <u>product</u> $X(\mathcal{L}_i|i\in I)$ equipped with the obvious component-wise binary operations is itself a lattice.

A subset $\mathcal{L}_0 \subseteq \mathcal{L}$ is called a <u>sublattice</u> if $x,y \in \mathcal{L}_0 \Rightarrow x\wedge y,\ x\vee y \in \mathcal{L}_0$; in this case the embedding $u:\mathcal{L}_0 \to \mathcal{L}$ becomes a lattice homomorphism.

A representation of an element $x \in \mathcal{L}$ as $x = \bigvee(x_i|i\in I)$, where I is a finite set and $x_i \in \mathcal{L}$, $i \in I$, is called <u>join-irredundant</u> if there does not exist a proper subset $J \subset I$ such that $x = \bigvee(x_i|i\in J)$. An element $z \in \mathcal{L}$ is called <u>join-irreducible</u> if $z = x \vee y$ implies $z = x$ or $z = y$. The set $J(\mathcal{L})$ of all join-irreducible elements of $\mathcal{L}$ plays an important role in the study of finite distributive lattices.

<u>1.4. Distibutive lattices.</u>

A lattice $\mathcal{L}$ is called <u>distributive</u> if

(DL) $$\forall x,y,z \in \mathcal{L}:\ x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z).$$

This condition is equivalent to

(DL) $\forall x,y,z \in \mathcal{L}: x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$.

An important example of a distributive lattice is the following: Let M be a non-empty set and $\mathfrak{R}$ a set of subsets of M. $\mathfrak{R}$ is called a ring if $R_1 \cap R_2 \in \mathfrak{R}$ and $R_1 \cup R_2 \in \mathfrak{R}$ for every $R_1,R_2 \in \mathfrak{R}$. It is obvious that $\mathfrak{R}$ equipped with $\cap$ and $\cup$ as $\wedge$ and $\vee$ becomes a distributive lattice. The following is a major result in the theory of lattices.

Theorem 1.1. (Birkhoff (1933), Stone (1936)). A lattice is distributive if and only if it is isomorphic to a ring of subsets of a set.

Proof: See Grätzer, (1978) Theorem 19, p.64.□

Since we shall mainly be interested in finite distributive lattices the following result is useful:

Proposition 1.1. Let $\mathcal{L}$ be a distributive lattice of finite length. Then $\mathcal{L}$ is finite.

Proof: Zorn's lemma and the finite length assumption implies that $\mathcal{L}$ has a maximal element $x_1 \in \mathcal{L}$. If $x \in \mathcal{L}$ then $x \vee x_1 = x_1$, hence $x \leq x_1$ for every $x \in \mathcal{L}$. Thus $x_1 = 1 := \sup \mathcal{L}$; similary, one shows that $0 := \inf \mathcal{L}$ exists.

Next we show that for any $x \in \mathcal{L}$ such that $x \neq 0$, there exists a finite chain $x = z_0 > z_1 > \cdots > z_m = 0$ such that z_i covers z_{i+1}, $i = 0,\cdots,m-1$. If x does not cover 0 then there exists y_1 such that $x > y_1 > 0$. If x does not cover y_1 then there exists y_2 such that $x > y_2 > y_1 > 0$. This process can continue for at most n steps, where n is the maximal length of any chain in $\mathcal{L}$. Thus there exists an z_1 such that $x > z_1 > 0$

and x covers z_1. If z_1 does not cover 0 we repeat the argument with x replaced by z_1 and obtain z_2 such that $x > z_1 > z_2 > 0$ and z_1 covers z_2. Again this process must terminate within at most n steps, hence the assertion is established.

Now define $\mathcal{L}_0 = \{0\}$, $\mathcal{L}_1 = \{x\in\mathcal{L} | x \text{ covers } 0\}$, $\mathcal{L}_2 = \cup(\{x\in\mathcal{L} | x \text{ covers } z\} | z\in\mathcal{L}_1), \cdots, \mathcal{L}_n = \cup(\{x\in\mathcal{L} | x \text{ covers } z\} | z\in\mathcal{L}_{n-1})$. Then the above assertion implies that $\mathcal{L} = \cup(\mathcal{L}_i | i=0,1,\cdots,n)$.

If $x_1 \neq x_2$ both cover $y \in \mathcal{L}$ then neither $x_1 < x_2$ nor $x_2 < x_1$ can occur, hence $x_1 > x_1 \wedge x_2 \geq y$. Since x_1 covers y, this implies that $x_1 \wedge x_2 = y$.

Next we claim that the set $\{x\in\mathcal{L} | x \text{ covers } y\}$ is finite for every $y \in \mathcal{L}$. If not, let $x_1, x_2, \cdots$ be an infinite sequence of distinct elements in $\mathcal{L}$ such that x_i covers y, $i = 1,2,\cdots$. Then $x_1 < (x_1 \vee x_2) < \cdots < (x_1 \vee \cdots \vee x_r) < \cdots$ is an infinite chain, since $x_1 \vee \cdots \vee x_r = x_1 \vee \cdots \vee x_r \vee x_{r+1} \Rightarrow x_{r+1} \leq x_1 \vee \cdots \vee x_r \Rightarrow x_{r+1} = (x_1 \vee \cdots \vee x_r) \wedge x_{r+1} = (x_1 \wedge x_{r+1}) \vee \cdots \vee (x_r \wedge x_{r+1}) = y \vee \cdots \vee y = y$ contradicting that x_{r+1} covers y. Since every chain in $\mathcal{L}$ is finite, the claim is established. It then follows that $\mathcal{L} = \cup(\mathcal{L}_i | i=0,1,\cdots,n)$ is finite.□

<u>Remark 1.2</u>. It can be shown that Proposition 1.1 remains valid under the weaker condition that every chain in $\mathcal{L}$ has finite length.□

<u>Remark 1.3</u>. For a finite distributive lattice $\mathcal{L}$, the Möbius function μ for the poset $\mathcal{L}$ is specified by: $\mu(x,y) = (-1)^k$ if y is the join of k distinct elements covering x (Grätzer (1978), Exercise 36, p.191).□

1.5. Finite distributive lattices and posets with zero.

For the remainder of this section, $\mathcal{L}$ denotes a finite distributive lattice. Then $\mathcal{L}$ is a lattice with zero and with unit and the set

$$J(\mathcal{L}) = \{ x \in \mathcal{L} \mid \sup\{y\in\mathcal{L} | y<x\} < x \} \cup \{0\}$$

is the set of all join-irreducible elements of $\mathcal{L}$. Since $0 \in J(\mathcal{L})$, the subposet $J(\mathcal{L})$ of $\mathcal{L}$ is a poset with zero - see the right hand graphs in figures 3.1-3.15. (Grätzer (1978) excludes 0 from $J(\mathcal{L})$ but our definition is more convenient for our purposes.) Conversely, let $\mathcal{P}$ be a finite poset with zero. A non-empty subset S of $\mathcal{P}$ is called <u>hereditary</u> if $x \in S$ and $y \leq x$ implies that $y \in S$. Let $H(\mathcal{P})$ denote the set of all hereditary subsets of $\mathcal{P}$. Then $H(\mathcal{P})$ becomes a finite distributive lattice under the binary operations $\wedge := \cap$ and $\vee := \cup$. The mappings $\mathcal{L} \to J(\mathcal{L})$ and $\mathcal{P} \to H(\mathcal{P})$ determine a fundamental correspondence between the class of all finite distributive lattices and the class of all finite posets with zero (cf. Grätzer (1978), Theorem 9 and Corollary 10, pp. 61-62).

<u>Theorem 1.2</u>. (i): Let $\mathcal{L}$ be a finite distributive lattice. Then the mapping

$$\begin{aligned} &\mathcal{L} \to H(J(\mathcal{L})) \\ &x \to r(x) := \{y\in J(\mathcal{L}) | y\leq x\} \end{aligned} \tag{1.1}$$

is a lattice isomorphism i.e., $H(J(\mathcal{L})) \cong \mathcal{L}$.

(ii): Conversely let $\mathcal{P}$ be a poset with zero. Then the mapping

$$(1.2) \qquad \begin{aligned} \mathcal{P} &\to J(H(\mathcal{P})) \\ x &\to s(x) := \{y\in\mathcal{P} \mid y\leq x\} \end{aligned}$$

is a poset isomorphism i.e., $J(H(\mathcal{P})) \cong \mathcal{P}$.

The proof of (i) requires the following lemma.

<u>Lemma 1.2</u>. Let $x \in J(\mathcal{L})$ and $x_1,\cdots,x_k \in \mathcal{L}$. If $x \leq x_1 \vee\cdots\vee x_k$ then there exists $i \in \{1,\cdots,k\}$ such that $x \leq x_i$.

<u>Proof</u>: $x \leq x_1 \vee\cdots\vee x_k \Rightarrow x = x \wedge(x_1 \vee\cdots\vee x_k) = (x \wedge x_1)\vee\cdots\vee(x \wedge x_k) \Rightarrow \exists\, i: x \wedge x_i = x \Rightarrow \exists\, i: x \leq x_i$.□

<u>Proof of Theorem 1.2</u>: (i) That $r(x) \in H(J(\mathcal{L}))$ and $r(x\wedge y) = r(x)\wedge r(y)$ follow trivially, while Lemma 1.2 implies that $r(x\vee y) = r(x)\vee r(y)$. Since $\mathcal{L}$ is finite, $x = \vee r(x)$ for every $x \in \mathcal{L}$, hence the mapping (1.1) is one to one. To show that (1.1) is onto, choose $R \in H(J(\mathcal{L}))$ and define $x = \vee R$. Clearly, $R \subseteq r(x)$. If $y \in r(x)$, then $y = y\wedge x = y\wedge(\vee R) = \vee(y\wedge z \mid z\in R)$; Since $y \in J(\mathcal{L})$, this implies that $y = y\wedge z$ for some $z \in R$, i.e., $y \leq z$. As R is hereditary, $y \in R$, hence $r(x) = R$.

(ii): To see that $s(x) \in J(H(\mathcal{P}))$ suppose that $s(x) = S_1\cup S_2$ for $S_1,S_2 \in H(\mathcal{P})$. Since $x \in s(x)$, without loss of generality assume that $x \in S_1$. Then $s(x) \subseteq S_1$, hence $S_2 \subseteq S_1$, so $s(x) \in J(H(\mathcal{P}))$. Trivially, $x \leq y \Rightarrow s(x) \subseteq s(y)$, $x,y \in P$. Since $x = \sup s(x)$, the mapping (1.2) is one to one. To see that (1.2) is onto, choose $S \in J(H(\mathcal{P}))$ and let $x_1,\cdots,x_k$

denote the maximal elements in S. Since S is a hereditary subset of $\mathcal{P}$, S $= \cup(s(x_i)|i=1,\cdots,k)$, while, since S is join-irreducible, $\exists$ i such that S $= s(x_i)$.□

Corollary 1.1. Every element in a finite distributive lattice has a unique join-irredundant representation as a join of join-irreducible elements.

Proof: See Grätzer (1978), Corollary 13, p. 62.□

The next proposition establihes the natural correspondence between homomorphisms of finite distributive lattices and homomorphisms of finite posets with zero.

Proposition 1.2. (i): Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be two posets with zero and let $\psi:\mathcal{P}_1 \to \mathcal{P}_2$ be a poset homomorphism such that $\psi(0) = 0$. The mapping

$$(1.3) \qquad \begin{aligned} H(\psi):\ H(\mathcal{P}_2) &\to H(\mathcal{P}_1) \\ h &\to \psi^{-1}(h) \end{aligned}$$

is a lattice homomorphism with the property $H(\psi)(1) = 1$.

(ii): Conversely, let $\mathcal{L}_1$ and $\mathcal{L}_2$ be two finite distributive lattices and φ: $\mathcal{L}_1 \to \mathcal{L}_2$ a lattice homomorphism such that $\varphi(1) = 1$. The mapping

$$(1.4) \qquad \begin{aligned} J(\varphi):\ J(\mathcal{L}_2) &\to J(\mathcal{L}_1) \\ x &\to \inf\{x' \in \mathcal{L}_1 | \varphi(x') \geq x\} \end{aligned}$$

is a poset homomorphism such that $J(\varphi)(0) = 0$.

Proof: (i): To see that $\psi^{-1}(S) \in H(\mathcal{P}_1)$, suppose that $x \in \psi^{-1}(S)$ and $y \leq x$. Then $\psi(y) \leq \psi(x) \in S$, hence $y \in \psi^{-1}(S)$. It is trivial to verify that $H(\psi)(S_1 \cup S_2) = H(\psi)(S_1) \cup H(\psi)(S_2)$ and the analogous property for $\cap$. Also, $H(\psi)(1) = H(\psi)(\mathcal{P}_2) = \psi^{-1}(\mathcal{P}_2) = 1$.

(ii): To show that $J(\varphi)(x) \in J(\mathcal{L}_1)$, note first that $\{x' \in \mathcal{L}_1 | \varphi(x') \geq x\} \neq \emptyset$ for every $x \in \mathcal{L}_2$ because $\varphi(1) = 1$. Let $J(\varphi)(x) = y_1 \vee \cdots \vee y_k$ be the unique join-irredundant representation of $J(\varphi)(x)$ as the join of join-irreducible elements (cf. Corollary 1.1). To see that $k = 1$, note that $\varphi(J(\varphi)(x)) = \varphi(y_1) \vee \cdots \vee \varphi(y_k) \geq x \in J(\mathcal{L}_2)$. By Lemma 1.2, this implies that there exists i such that $\varphi(y_i) \geq x$. Thus we have $y_i \geq J(\varphi)(x) = y_1 \vee \cdots \vee y_k$ and therefore $k = 1$. Next for $x,y \in J(\mathcal{L}_2)$ with $x \leq y$ we have $J(\varphi)(x) = \inf\{x' \in \mathcal{L}_1 | \varphi(x') \geq x\} \leq \inf\{x' \in \mathcal{L}_1 | \varphi(x') \geq y\} = J(\varphi)(y)$. Finally, $J(\varphi)(0) = \inf\{x \in \mathcal{L}_1 | \varphi(x) \geq 0\} = \inf\{x \in \mathcal{L}_1\} = 0$.□

Remark 1.4. If the lattice $\mathcal{L}_i$ is identified with $H(J(\mathcal{L}_i))$ through (1.1) and the poset $\mathcal{P}_i$ is identified with $J(H(\mathcal{P}_i))$ through (1.2), i=1,2, then $J(H(\psi)) = \psi$ and $H(J(\varphi)) = \varphi$.□

Remark 1.5. For the identity mappings $id_{\mathcal{P}}$ and $id_{\mathcal{L}}$ we have $H(id_{\mathcal{P}}) = id_{H(\mathcal{P})}$ and $J(id_{\mathcal{L}}) = id_{J(\mathcal{L})}$. It can also be seen that if $\psi_1 : \mathcal{P}_1 \to \mathcal{P}_2$ and $\psi_2 : \mathcal{P}_2 \to \mathcal{P}_3$ are homomorphisms of finite posets with zero such that $\psi_1(0) = 0$ and $\psi_2(0) = 0$, then $H(\psi_2 \circ \psi_1) = H(\psi_1) \circ H(\psi_2)$. Likewise if $\varphi_1 : \mathcal{L}_1 \to \mathcal{L}_2$ and $\varphi_2 : \mathcal{L}_2 \to \mathcal{L}_3$ are homomorphisms of finite distributive lattices such that $\varphi_1(1) = 1$ and $\varphi_2(1) = 1$, then $J(\varphi_2 \circ \varphi_1) = J(\varphi_1) \circ J(\varphi_2)$.□

Proposition 1.3. (i): Let $u:\mathcal{L}_0 \to \mathcal{L}$, where $\mathcal{L}_o$ and $\mathcal{L}$ are finite distributive lattices, be an injective lattice homomorphism such that $u(1) = 1$. Then $J(u):J(\mathcal{L}) \to J(\mathcal{L}_0)$ is surjective.

(ii): Conversely, let $p:\mathcal{P} \to \mathcal{P}_0$, where $\mathcal{P}$ and $\mathcal{P}_0$ are posets with zero, be a surjective poset homomorphism such that $p(0) = 0$. Then $H(p):H(\mathcal{P}_0) \to H(\mathcal{P})$ is injective.

Proof: (i): Let $x_0 \in J(\mathcal{L}_0)$. We shall find $x \in J(\mathcal{L})$ such that $J(u)(x) = x_0$. Since $u(x_0) \in \mathcal{L}$, there is a unique join-irredundant representation $u(x_0) = x_1 \vee\cdots\vee x_k$ of $u(x_0)$ as the join of join-irreducible elements. Since $x_i \leq u(x_0)$ we have $J(u)(x_i) \leq x_0$, $i=1,\cdots,k$, hence $J(u)(x_1) \vee\cdots\vee J(u)(x_n) \leq x_0$. On the other hand, $x_i \leq u(J(u)(x_i))$, $i=1,\cdots,k$, so $x_1 \vee\cdots\vee x_k \leq u(J(u)(x_1)) \vee\cdots\vee u(J(u)(x_k)) \leq u(x_0)$. Thus $x_0 = J(u)(x_1) \vee\cdots\vee J(u)(x_k)$. Since $x_0 \in J(\mathcal{L}_0)$, there exists an i such that $J(u)(x_i) = x_0$, as claimed. The proof of (ii) is trivial.□

2. THE LATTICE OF SUBSPACES OF A VECTOR SPACE.

2.1. Distributive lattices of subspaces.

Let V be a finite-dimensional real vector space with zero element $\underset{\sim}{0}$. A subspace of V is a pair (L,u) consisting of a vector space L and an injective linear mapping $u:L \to V$. (For example, in the definition of a linear model (cf. Section 3), L represents the parameter space and u the parametrization mapping.) Let (L,u) and (L',u') be two subspaces of V. If there exists an injective linear mapping $v:L' \to L$ such that $u' = u \circ v$, then we write $(L',u') \subseteq (L,v)$. Note that v is unique and that (L',v) is a subspace of L. The inclusion relation $\subseteq$ is not antisymmetric on the set of all subspaces of V, but is antisymmetric on the set $\mathcal{L}(V)$ of all equivalence classes of subspaces of V determined by the following equivalence relation: $(L,u) \sim (M,v)$ if $u(L) = v(M)$. Equipped with the ordering relation induced by $\subseteq$ (also denoted $\subseteq$), $\mathcal{L}(V)$ becomes a poset. The usual representation of $\mathcal{L}(V)$ is the set of all embedded subspaces in the classical sense, that is, all pairs (L,u) where L is a subset of V closed under the vector space operations in V and $u:L \to V$ is the embedding ($u(x) = x$). (Usually we omit the embedding u and simply write L for (L,u).) For this representation, the relation $\subseteq$ on $\mathcal{L}(V)$ is the usual inclusion relation for embedded subspaces. We shall study the structure of the poset $\mathcal{L}(V)$ through this representation. The usual vector space concepts such as intersection, sum, direct sum ($\oplus$) and complement may be defined in $\mathcal{L}(V)$.

Since the poset $\mathcal{L}(V)$ satisfies conditon (P4) of Section 1, $\mathcal{L}(V)$ is a lattice with $\wedge$ and $\vee$ defined by $L \wedge M := L \cap M$ and $L \vee M := L+M = \text{span}\{L,M\}$, $L,M \in \mathcal{L}(V)$. In fact, $\mathcal{L}(V)$ satisfies a stronger condition namely:

(2.1) For any subset $S \subseteq \mathscr{L}(V)$ the elements $\Sigma S := \sup(S)$ (=spanS) and $\cap S = \inf(S)$ exist. Furthermore, there exist finite subsets $S_{1,} S_0 \subseteq S$ such that $\Sigma S = \Sigma S_1$ and $\cap S = \cap S_0$.

$\mathscr{L}(V)$ is a lattice with unit and zero given by the subspaces V and $\{\underset{\sim}{0}\}$ respectively. If $\dim(V) \geq 2$ then $\mathscr{L}(V)$ is not distributive and $|\mathscr{L}(V)| = \infty$, but $\mathscr{L}(V)$ has finite length (= $\dim(V)+1$).

Lemma 2.1. Let $L, M \in \mathscr{L}(V)$ and let V_L, V_M be complements of $L \cap M$ in L and M respectively. Then $L + M = V_L \oplus (L \cap M) \oplus V_M$.

Proof: Staightforward from the definitions of direct sum and complement.□

Let $\mathscr{L} \subseteq \mathscr{L}(V)$ be a distributive sublattice. Since $\mathscr{L}(V)$ has finite length, so does $\mathscr{L}$, hence $\mathscr{L}$ is a finite by Proposition 1.1. Since $\mathscr{L}$ is finite, it has a unit $1 = 1_{\mathscr{L}}$ and a zero $0 = 0_{\mathscr{L}}$. Note that in general, $1 \subseteq V$ and $0 \supseteq \{\underset{\sim}{0}\}$..

Theorem 2.1. (Decomposition Theorem). Let $\mathscr{L} \subseteq \mathscr{L}(V)$ be a distributive sublattice with $V \in \mathscr{L}$. For $L \in J(\mathscr{L})\backslash\{0\}$, let V_L be a complement of $J(L) := \Sigma(L' \in \mathscr{L} | L' \subset L)$ in L; for $L = 0$, let $V_L = 0$. Then

$$V = \oplus\ (V_{L'}\ |L' \in J(\mathscr{L})) \tag{2.2}$$

Proof: We shall need the fact that for any $L \in \mathscr{L}$, the set $\mathscr{L}^L := \{L' \in \mathscr{L} | L' \subseteq L\}$ is a distributive lattice with L as the unit element and

(2.3) $$J(\mathcal{L}^L) = J(\mathcal{L})\cap\mathcal{L}^L$$

To prove (2.2) proceed by induction on $|J(\mathcal{L})|$. If $|J(\mathcal{L})| = 1$ or 2, (2.2) is immediate. For $n\geq 3$ suppose that (2.2) is true whenever $|J(\mathcal{L})| \leq n-1$, and assume that $|J(\mathcal{L})| = n$. Suppose first that $V\in J(\mathcal{L})$. Then $J(V) \subset V$ and $|J(\mathcal{L}^{J(V)})|=|J(\mathcal{L})\setminus\{V\}| = n-1$ by (2.3), hence

$$J(V) = \oplus(V_L, |L'\in J(\mathcal{L}), L'\subseteq J(V))$$

by the induction hypothesis and (2.3). But $V = V_V\oplus J(V)$, hence (2.2) holds.

Next, suppose that $V\notin J(\mathcal{L})$. Then $V = L + M$, where $L,M \in \mathcal{L}$ with $L \subset V$ and $M \subset V$. It follows from Lemma 1.2 that

(2.4) $$\{L'\in J(\mathcal{L})\,|L'\subseteq L\}\cup\{L'\in J(\mathcal{L})\,|L'\subseteq M\} = J(\mathcal{L}).$$

By (2.3) this implies that $|J(\mathcal{L}^L)|<n$, $|J(\mathcal{L}^M)|<n$ and $|J(\mathcal{L}^{L\cap M})|<n$. By induction and (2.3),

$$L = \oplus(V_L, |L'\in J(\mathcal{L}), L'\subseteq L\ ) =$$
$$(\oplus(V_L, |L'\in J(\mathcal{L}), L'\subseteq L, L'\not\subseteq L\cap M)) \oplus (\oplus(V_L, |L'\in J(\mathcal{L}), L'\subseteq L\cap M)) =$$
$$\oplus(V_L, |L'\in J(\mathcal{L}), L'\subseteq L, L'\not\subseteq L\cap M)\oplus(L\cap M),$$

and an analogous formula holds for M, namely

$$M = \oplus(V_L, |L'\in J(\mathcal{L}), L'\subseteq M, L'\not\subseteq L\cap M)\oplus(L\cap M).$$

By Lemma 2.1 and (2.4),

$$V = L + M =$$
$$\oplus(V_L, |L'\in J(\mathcal{L}), L'\subseteq L, L'\not\subseteq L\cap M)) \oplus (\oplus(V_L, |L'\in J(\mathcal{L}), L'\subseteq L\cap M)) \oplus$$
$$(\oplus(V_L, |L'\in J(\mathcal{L}), L'\subseteq M, L'\not\subseteq L\cap M)) = \oplus(V_L, |L'\in J(\mathcal{L})).\square$$

For any $L \in \mathcal{L}$, we may apply Theorem 2.1 with V and $\mathcal{L}$ replaced by L and $\mathcal{L}^L$ and invoke (2.3) to obtain

$$(2.5) \qquad L = \oplus(V_{L'} \,|L'\in J(\mathcal{L}), L'\subseteq L), \quad L\in\mathcal{L}.$$

We call (2.5) a <u>decomposition of</u> $L\in\mathcal{L}$ <u>with respect to</u> $\mathcal{L}$. It is not unique, since for every $L\in J(\mathcal{L})$ the complement V_L can be choosen in many ways. (Compare with Theorem 2.2.)

<u>Remark 2.1</u>. In fact any (abstract) finite distributive lattice $\mathcal{M}$ can be represented as a distributive sublattice $\mathcal{L} \subseteq \mathcal{L}(V)$ for some finite -dimensional vector space V with $V \in \mathcal{L}$. One such representation is constructed as follows. For each $x \in J(\mathcal{M})$ let V_x be an arbitrary finite -dimensional vectorspace of dimension ≥ 1. Define

$$V = \oplus(V_{x'} \,|x'\in J(\mathcal{M}))$$

and

$$L_x = \oplus(V_{x'} \,|x'\in J(\mathcal{M}), x'\leq x), \quad x \in \mathcal{M}.$$

Then $\mathcal{L} := \{L_x | x\in\mathcal{M}\} \subseteq \mathcal{L}(V)$ is a distributive sublattice and the mapping $x \to L_x$ of $\mathcal{M}$ onte $\mathcal{L}$ is a lattice isomorphism.

2.2. Geometrically orthogonal lattices of subspaces.

Linear statistical models and variance component models are defined on a real vector spaces V equipped with an inner product δ (cf. Section 3 and 4). In the remainder of Section 2 we study the interplay between the distributive property for a sublattice $\mathcal{L} \subseteq \mathcal{L}(V)$ and a geometric orthogonality property (with respect to δ) of the subspaces in $\mathcal{L}$.

Let δ be an inner product on V, i.e., $\delta: V\times V \to \mathbb{R}$ is a positive definite form. If (L,u) is a subspace of V the mapping $\delta\circ(u\times u)$ becomes an inner product on L. The adjoint linear mapping $u^*: V \to L$ with respect to (wrt) δ and $\delta\circ(u\times u)$ defined by the equation $\delta(x,u(z)) = \delta\circ(u\times u)(u^*(x),z)$

$(\equiv\delta(u(u^*(x)),u(z))$, $\forall x\in V$, $\forall z\in L$ is called the orthogonal projection onto (L,u) (or just onto L) wrt δ. This should not be confused with the linear mapping $q_L := u\circ u^*: V \to V$ which is usually called the orthogonal projection on L. The mapping q_L depends on (L,u) only through its equivalence class. The complement $q_L^{-1}(0) = (\in \mathscr{L}(V))$ of L in V is called the orthogonal complement of L in V wrt δ, and is denoted by $L^{\perp}$.

Definition 2.1. Two subspaces $L,M \in \mathscr{L}(V)$ are called geometrically orthogonal (g.o.) wrt δ if the orthogonal projections q_L and q_M commute, i.e., $q_Lq_M = q_Mq_L$. A subset $\mathscr{M} \subseteq \mathscr{L}(V)$ is called geometrically orthogonal wrt δ if every pair $L,M \in \mathscr{M}$ is g.o.□

If L and M are g.o. wrt δ then

(2.6) $$q_{L\cap M} = q_Lq_M,$$

(2.7) $$q_{L+M} = q_L+q_M-q_Lq_M$$

Furthermore, one can easily see that L and M are g.o. wrt δ if and only if the two subspaces $L\cap(L\cap M)^{\perp}$ and $M\cap(L\cap M)^{\perp}$ are orthogonal wrt δ.

Remark 2.2. Let $\mathscr{M} \subseteq \mathscr{L}(V)$ be a g.o. subset and let $\mathscr{L}$ be the smallest lattice containing $\mathscr{M}$. Any element in $\mathscr{L}$ is obtained from $\mathscr{M}$ by means of finitely many binary operations using $\cap$ and $+$ (cf. Grätzer (1978) p.27 Lemma 3). It follows from (2.6) and (2.7) that $\mathscr{L}$ is also g.o.□

Proposition 2.1. Let the sublattice $\mathscr{L} \subseteq \mathscr{L}(V)$ be geometrically orthogonal wrt δ. Then $\mathscr{L}$ is distributive, hence finite.

Proof: Let $L,M,N \in \mathscr{L}$ and let q_L, q_M and q_N be the corresponding orthogonal projections. Since they all commute, it follows from (2.6) and (2.7) that the orthogonal projections on $L\cap(M+N)$ and $(L\cap M)+(L\cap N)$ are $q_L(q_M+q_N-q_Mq_N)$ and $q_Lq_M+q_Lq_N-(q_Lq_M)(q_Lq_N)$ respectively. Since these two orthogonal projections are identical it follows that $L\cap(M+N) = (L\cap M)+(L\cap N)$, and thu that $\mathscr{L}$ is distributive. The final assertion is immediate from Proposition 1.1.□

By Proposition 2.1 and Remark 2.2, if $\mathscr{M} \subseteq \mathscr{L}(V)$ is g.o., it must be finite.

The following result is a partial converse to Proposition 2.1.

Proposition 2.2. Let $\mathscr{L} \subseteq \mathscr{L}(V)$ be a distributive sublattice. Then there exists an inner product δ on V such that $\mathscr{L}$ is geometrically orthogonal wrt δ.

Proof: Without loss of generality we may suppose that $V \in \mathscr{L}$. Consider a decomposition (2.2) of V wrt $\mathscr{L}$. Then we can choose an inner producer δ such that the direct sum in (2.2) becomes orthogonal wrt δ. It follows from (2.5) that all pairs $L,M \in \mathscr{L}$ are g.o. wrt δ. □

Theorem 2.2. (Orthogonal Decomposition Theorem). Let the sublattice $\mathscr{L} \subseteq \mathscr{L}(V)$ be geometrically orthogonal wrt δ. For $L \in J(\mathscr{L})$ choose each complement V_L in Theorem 2.1 to be the orthogonal complement of $J(L)$ in L wrt δ, i.e., $V_L = L\cap J(L)^{\perp}$. Then the direct sums (2.2) and (2.5) become orthogonal wrt δ.

Conversely, if the direct sum in (2.2) is orthogonal wrt δ, then $V_L = L \cap J(L)^{\perp}$, $L \in J(\mathcal{L})\setminus\{0\}$.

Proof: Let $L, M \in J(\mathcal{L})$ with $L \neq M$. Then the orthogonal projections on V_L and V_M become $r_L = q_L - q_{J(L)}$ and $r_M = q_M - q_{J(M)}$ respectively. Since all q_N, $N \in \mathcal{L}$, commute, r_L and r_M commute, hence V_L and V_M are g.o. wrt δ. Since $V_L \cap V_M = \{\underset{\sim}{0}\}$, they must be orthogonal wrt δ.□

Remark 2.3. When V is equipped with an inner product δ and a direct sum is orthogonal wrt δ, it will be written with the symbol $\perp$ rather than $\oplus$.□

3. GEOMETRICALLY ORTHOGONAL DESIGNS OF NORMAL LINEAR MODELS.

3.1. The linear statistical model.

Let $V \neq \{\underset{\sim}{0}\}$ be a finite-dimensional vector space and δ a fixed inner product on V. Every subspace (P,u) of V together with δ determines a normal regression model in the following well-known way: the observation space is V, the parameter space is $P\times\mathbb{R}_+$, where $\mathbb{R}_+ =]0,\infty[$, and the set of unknown probability measures on V consists of the normal distributions with mean $u(\xi)$ and precision $\sigma^{-2}\delta$, where $(\xi,\sigma^2) \in P\times\mathbb{R}_+$. The vector space P and the mapping u are the parameter space for and the parametrization of the mean value subspace ($\equiv$ regression subspace) $L = u(P)$. It is well known that the maximum likelihood (ML) estimator $(\hat{\xi},\hat{\sigma}^2)$ for $(\xi,\sigma^2) \in P\times\mathbb{R}_+$ exists with probability 1 if and only if P is a proper subspace of V. In this case the ML estimator is unique and is given by

$$(3.1) \qquad (\hat{\xi},\hat{\sigma}^2)(x) = (p(x),\delta(x-q(x))/n), \ x \in V,$$

where $p:V \to P$ is the orthogonal projection wrt δ, $q = u\circ p$, $n = \dim(V)$, and $\delta(x):=\delta(x,x)$, $x \in V$. The ML estimator is a minimal sufficient statistic and its distribution can be described in the following way: $\hat{\xi}$ and $\hat{\sigma}^2$ are independent, $\hat{\xi}$ is normally distributed on P with mean ξ and precision $\sigma^{-2}\delta\circ(u\times u)$, and $\hat{\sigma}^2$ is χ^2-distributed with n-n' degrees of freedom and scale σ^2/n, where $n' = \dim(P)$. Usually one uses the unbiased estimator $s^2 = \hat{\sigma}^2 n/(n-n')$. It is thus seen that the solution to the likelihood inference problem is reduced to the algebraic problem of representing p and calculating dim(P). In the calculation of $\hat{\sigma}^2$ or s^2 one uses the formula

$\delta(x-q(x)) = \delta(x)-\delta(q(x)),\ x\in V.$

3.2. Nested designs of linear models.

We now discuss a simple case of a geometrically orthogonal (g.o.) design of normal linear models (the general case is defined in Section 3.3). Let

$$(3.2)\qquad (P_0,u_0) \subseteq (P_1,u_1) \subseteq\cdots\subseteq (P_{k-1},u_{k-1}) \subseteq (P_k,u_k) \equiv (V,id_V),$$

where id_V is the identity mapping, be a nested family of subspaces in V. These subspaces are trivially pairwise g.o. and together with δ they determine k+1 linear models. The likelihood ratio (LR) statistic Q_{ij} for testing model i against model j, $0\leq i<j<k$, is given by

$$(3.3)\qquad Q_{ij}(x)^{2/n} = \frac{\delta(x-q_j(x))}{\delta(x-q_i(x))},\quad x \in V,$$

where $q_i:V \rightarrow V$ is the orthogonal projection onto $u_i(P_i)$, $i=0,1,\cdots,k$. Under model i, $Q_{ij}^{2/n}$ has the beta distribution with n_k-n_j and n_j-n_i degrees of freedom, where $n_i = \dim(P_i)$, $(n \equiv n_k)$ which does not depend on the unknown parameters. Furthermore, $Q_{k-2,k-1},\cdots,Q_{i,i+1}$ and the ML estimator $(\hat{\xi},\hat{\sigma}^2)$ in model i are independent under model i, $i = 0,\cdots,k-2$. The calculation of these statistics do not require any new quantities. Therefore, the testing problems also reduce to the algebraic problem described above.

The ANOVA table associated with the design (3.2) is as follows:

$$(3.4) \qquad (SSD_0, f_0),\ (SSD_1, f_1), \cdots, (SSD_k, f_k),$$

where $SSD_i(x) = \delta(q_i(x)-q_{i-1}(x))$, $i=1,\cdots,k$, and $SSD_0(x) = \delta(q_0(x))$, $x \in V$, $f_i = n_i-n_{i-1}$, $i=1,\cdots,k$, and $f_0 = n_0$. All quantities needed to calculate the LR test statistics Q_{ij}, $0\leq i<j<k$, and the ML estimators of σ^2 under the models $0,1,\cdots,k-1$ are readily obtained from this table. The ANOVA table (3.4) arises from the orthogonal decomposition of V determined by the subspaces (3.2) and δ:

$$(3.5) \qquad V = L_0 \perp (L_1 \cap L_0^{\perp}) \perp \cdots \perp (L_{k-1} \cap L_{k-2}^{\perp}) \perp (L_k \cap L_{k-1}^{\perp})$$

where $L_i = u_i(P_i)$, $i=0,1,\cdots,k$.

3.3. Geometrically orthogonal designs.

The reader is undoubtedly familiar with the classical examples of balanced ANOVA designs, where such orthogonal decompositions of V and ANOVA tables may also be defined and possess analogous properties to (3.4) and (3.5), e.g., balanced multi-way ANOVA and split plot designs. Furthermore, it is wellknown that in unbalanced designs, no such decompositions and tables exist (cf. Examples 3.4 and 3.13). It is the main aim of this section to characterize those designs (= families of linear regression models) where suitable decompositions and ANOVA tables exist. In this generality, this question has not yet been adequately answered in the literature. It will be shown that these are exactly the geometrically orthogonal (g.o.) designs.

We shall show that a design is g.o. only if its lattice $\mathcal{L}$ of regression subspaces is distributive. Furthermore, in this case the structure and analysis of the design (eg., the contrast estimates and ANOVA table) are uniquely and unambiguously determined by the contrast subspaces indexed by $J(\mathscr{L})$, the poset of join-irreducible elements of $\mathscr{L}$.

Let $\mathscr{M} \subseteq \mathscr{L}(V)$ be a family of embedded subspaces of V. Let (P_L, u_L) be a subspace of V such that $u_L(P_L) = L$, $L \in \mathscr{M}$. We refer to the family of linear models determined by the subspaces (P_L, u_L), $L \in \mathscr{M}$, and by the fixed inner product δ, as a design of linear models. If $\mathscr{M}$ is g.o. wrt δ it follows from Remark 2.2 and Proposition 2.1 that the smallest lattice $\mathscr{L}$ containing $\mathscr{M}$ is g.o., hence distributive, and thus finite. The extension from $\mathscr{M}$ to the lattice $\mathscr{L}$ is statistically natural, since interest in the models corresponding to the subspaces in $\mathscr{M}$ implies interest in the models corresponding the subspaces obtained from $\mathscr{M}$ by means of the operations + and $\cap$. The g.o. condition on $\mathscr{M}$ insures $\mathscr{L}$ is finite. (Without the orthogonality condition $\mathscr{L}$ may be infinite even if $\mathscr{M}$ is finite.)

Definition 3.1. For a fixed inner product δ on V, let $\mathscr{L}$ be a geometrically orthogonal (hence distributive and finite) lattice of embedded subspaces of V. Let (P_L, u_L) be a subspace of V such that $u_L(P_L) = L$, $L \in \mathscr{L}$. The design determined by $((P_L, u_L) | L \in \mathscr{L})$ and δ is called a geometrically orthogonal design of linear models. When there is no danger of confusion we simply refer to the g.o. design determined by $\mathscr{L}$.□

Remark 3.1. If $\mathscr{L}$ is a g.o. lattice of embedded subspaces of V, then the lattice $\mathscr{L} \cup \{V\}$ is also a g.o. lattice. For notational convenience we hereafter assume without loss of generality in Definition 3.1 that $V \in \mathscr{L}$. The nested chain of linear models given by (3.2) is trivially a g.o. design.□

3.4. Contrast vector estimates and the ANOVA table.

We now show that each g.o. design of linear models determines a decomposition of V together with an associated ANOVA table from which every LR test statistic Q and variance estimator s^2 within the design may be obtained, together with their distributions.

Consider a g.o. design of linear models determined by the family $((P_L, u_L)|L\in\mathcal{L})$ of subspaces in V and the inner product δ. From Theorem 2.2 and Remark 2.3 we obtain the orthogonal decompositions

$$(3.6) \qquad M = \perp(V_L|L\in J(\mathcal{L}), L\subseteq M),\ M \in \mathcal{L},$$

$$(3.7) \qquad V = \perp(V_L|L\in J(\mathcal{L})),$$

where $V_L = L\cap J(L)^{\perp}$, $L \in J(\mathcal{L})$. Next for $L \in J(\mathcal{L})$, let $r_L:V \to V$ denote the orthogonal projection onto V_L, $SSD_L(x) := \delta(r_L(x))$, $x \in V$, and $f_L := \dim(V_L)$. The <u>contrast (vector) estimates</u> and the <u>analysis of variance</u> (≡<u>ANOVA</u>) <u>table</u> corresponding to an observation $x \in V$ are then defined as the families

$$(3.8) \qquad (r_L(x)|L\in J(\mathcal{L}))$$

and

$$(3.9) \qquad ((SSD_L(x), f_L)|L\in J(\mathcal{L}))$$

respectively. Often, the quantities $s_L^2(x)$ also appear in the table, where

$s_L^2 := SSD_L/f_L$, $L \in J(\mathscr{L})$. The mean value ML estimators $\hat{\xi}_L$, $L\in\mathscr{L}$, (see (3.10)) can readily be obtained from the contrast estimates (see (3.12)), and the variance estimators s_{OL}^2, $L\in\mathscr{L}$, (see (3.10)) and LR test statistics Q_{MN}, $M,N\in\mathscr{L}$, (see (3.11)) can readily be obtained from the ANOVA table, together with their distributions, (see Remark 3.2 below).

3.5. Estimation and testing in a geometrically orthogonal design.

For $L \in \mathscr{L}$, let $p_L : V \to P_L$ denote the orthogonal projection onto P_L, $q_L := u_L \circ p_L$, $SS_L(x) = \delta(q_L(x))$, $x\in V$, $n_L := \dim(L)$ $(= \mathrm{tr}(q_L))$, and $n := n_V$. For $N \in \mathscr{L}$, $N \subset V$, the statistics $\hat{\xi}_N$ and s_{ON}^2 given by

$$(3.10) \qquad u_N(\hat{\xi}_N) = q_N \text{ and } s_{ON}^2 := \frac{SS_V - SS_N}{n_V - n_N}$$

are the mean value ML estimator for $\xi \in P_N$ and the unbiased estimator for $\sigma^2 \in \mathbb{R}_+$ under the model corresponding to N; $\hat{\xi}_N$ is normally distributed on P_L with mean ξ_N and with precision $\sigma^{-2}\delta\circ(u_N \times u_N)$ (or equivalently with variance $\sigma^2\delta^{-1}\circ(p_N^t \times p_N^t)$, where $p_N^t : P_N^* \to V^*$ is the dual mapping to p_N), and s_{ON}^2 is χ^2 distributed with $n_V - n_N$ degrees of freedom and scale $\sigma^2/(n_V - n_N)$.

For $M \in \mathscr{L}$ and $M \subset N \subset V$, the LR statistic Q_{MN} for testing the linear model corresponding to M against that corresponding to N is given by

$$(3.11) \qquad Q_{MN}^{2/n} = \frac{SS_V - SS_N}{SS_V - SS_M};$$

under the null hypothesis, $Q_{MN}^{2/n}$ has the beta distribution with $n_V - n_N$ and $n_N - n_M$ degrees of freedom.

<u>Remark 3.2</u>. In order to calculate the contrast estimates and the ANOVA table proceed as follows. For an observation $x \in V$ begin by calculating $q_M(x)$, $SS_M(x)$, and n_M <u>for all</u> $M \in J(\mathcal{L})$. The following formulas are immediate from (3.6):

$$q_M = \Sigma(r_L \,|\, L\in J(\mathcal{L}), L\sqsubseteq M), \quad M \in \mathcal{L}, \tag{3.12}$$

$$SS_M = \Sigma(SSD_L \,|\, L\in J(\mathcal{L}), L\sqsubseteq M), \quad M \in \mathcal{L}, \tag{3.13}$$

$$n_M = \Sigma(f_L \,|\, L\in J(\mathcal{L}), L\sqsubseteq M), \quad M \in \mathcal{L}. \tag{3.14}$$

The equations (3.12-14) <u>for</u> $M \in J(\mathcal{L})$ are then solved using the Möbius function μ for $J(\mathcal{L})$ to obtain:

$$r_M = \Sigma(\mu(L,M)q_L \,|\, L\in J(\mathcal{L})), \quad M \in J(\mathcal{L}), \tag{3.15}$$

$$SSD_M = \Sigma(\mu(L,M)SS_L \,|\, L\in J(\mathcal{L})), \quad M \in J(\mathcal{L}), \tag{3.16}$$

$$f_M = \Sigma(\mu(L,M)n_L \,|\, L\in J(\mathcal{L})), \quad M \in J(\mathcal{L}). \tag{3.17}$$

From these equations the contrast estimates (3.8) and the ANOVA table (3.9) are obtained. Finally one uses (3.12-14) again to calculate $q_M(x)$, $SS_M(x)$ and n_M, for $M \in \mathcal{L}\setminus J(\mathcal{L})$.□

<u>Remark 3.3</u>. If $L \in J(\mathcal{L})$, then we refer to L as a subspace of <u>main effect</u> and V_L as the subspace of the associated contrast (= contrast vectors). This agrees with the terminology used in classical ANOVA examples.□

If the family $\mathcal{M}$ of subspaces of interest is not g.o. wrt δ, but the smallest lattice $\mathcal{L}$ containing $\mathcal{M}$ is still distributive (hence finite), then the decompositions in (3.6-7) are not orthogonal (eg. Example 3.4). This implies that (3.12), (3.13), (3.15) and (3.16) are not valid. In this situation estimators and LR statistics cannot be determined from the contrasts and the ANOVA table. Furthermore, the SSD statistics defined by (3.16) are not independent. (Note that by Proposition 2.2, in this case there does exist some other inner product $\tilde{\delta}$, i.e., another precision structure, such that $\mathcal{L}$ becomes g.o. wrt $\tilde{\delta}$. However, $\tilde{\delta}$ may not be statistical meaningful.) If $\mathcal{L}$ is not distributive but still finite the decomposition of V cannot be defined. These considerations show that theonly designs of linear models for which meaningful and useable decompositions and ANOVA tables can be defined are the g.o. designs.

3.6. Orthogonal factor-generated designs.

Tjur (1984) has studied a subclass of g.o. designs of linear models, namely, the class of analysis of variance designs with orthogonal factors. His treatment and formulation of such designs is essentially the same as ours, with one important difference. The cardinality of Tjur's set $\mathcal{D}$ (= the set of all factors in the design - see below) used by him to index his orthogonal decomposition of the observation space (Tjur, p. 42) and to index his ANOVA table (Tjur, Section 5) may be strictly greater than the cardinality of our set $J(\mathcal{L}_{\mathcal{D}})$ (= the set of all join-irreducible elements of the lattice $\mathcal{L}_{\mathcal{D}}$ determined by $\mathcal{D}$ - see below). Thus his decomposition and ANOVA table may contain trivial components and entries, whereas ours does not (cf. (3.6-9)). This reflects the fact that, in the general g.o. design, the index set $J(\mathcal{L})$ has a fundamental connection with the lattice structure $\mathcal{L}$ of the design of subspaces.

We briefly review the formulation and notation in Tjur (1984) and compare them to ours. In Tjur (1984) the observation space is $V = \mathbb{R}^I$, with I a finite index set, δ is the usual inner product, and the design of linear models is determined by a finite set $\mathcal{D}$ of orthogonal factors. A factor is a mapping $\bar{F}: I \to F$, where F is a finite set. We shall in the present work suppose that factors are surjective. Usually one simply refers to "the factor F", subsuming the mapping $\bar{F}: I \to F$ from the context. Two factors are said to be equivalent if they induce the same partition of I. The set $\mathcal{F}(I)$ of equivalence classes is finite and is equipped as a poset by the ordering $\leq$ given by: $F_1 \leq F_2$ if the partition induced by F_2 is finer than that induced by F_1. By convention we do not distinguish between equivalent factors, hence speak of a factor as an element of $\mathcal{F}(I)$. In fact, $\mathcal{F}(I)$ becomes a finite lattice with the one point set 0 as the minimal element and I as the maximal element.

A factor F defines a subspace (P_L, u_L) of $\mathbb{R}^I$ in the following way:

$$
\begin{aligned}
P_L &:= \mathbb{R}^F, \\
L := L_F &:= \{(x_{\bar{F}(i)} | i \in I) \in \mathbb{R}^I | (x_f | f \in F) \in \mathbb{R}^F\}, \\
u_L((x_f | f \in F)) &= (x_{\bar{F}(i)} | i \in I).
\end{aligned}
\tag{3.18}
$$

The matrix X_F for u_L is given by $(X_F)_{if} = 1$ if $\bar{F}(i) = f$ and 0 otherwise. Thus the matrices for the orthogonal projections $p_L: \mathbb{R}^I \to \mathbb{R}^F$ and $q_L = u_L \circ p_L: \mathbb{R}^I \to \mathbb{R}^I$ become $(X_F^t X_F)^{-1} X_F^t$ and $X_F (X_F^t X_F)^{-1} X_F^t$ respectively. The matrix $(X_F^t X_F)$ is diagonal with diagonal $(n_f | f \in F)$, where $n_f = |\bar{F}^{-1}(f)|$, $f \in F$. For $x = (x_i | i \in I) \in \mathbb{R}^I$ it follows that

(3.19) $$q_L(x) = (\frac{1}{n_{\overline{F}(i)}} S_{\overline{F}(i)}(x) \mid i \in I),$$

where

(3.20) $$S_f(x) = \Sigma(x_i \mid i \in \overline{F}^{-1}(f)), \; f \in F.$$

Futhermore,

(3.21) $$SS_L(x) := \delta(q_L(x)) = \Sigma(S_f(x)^2/n_f \mid f \in F)$$

and

(3.22) $$n_L := \dim(L) = |F| = \mathrm{tr}(q_L).$$

It is obvious that equivalent factors correspond to equivalent subspaces and that the mapping $F \to L_F$ of $\mathcal{F}(I)$ into $\mathcal{L}(\mathbb{R}^I)$ is an injective homomorphism of posets with the property

(3.23) $$L_{F \wedge G} = L_F \cap L_G, \; F, G \in \mathcal{F}(I).$$

In the sequel, we shall frequently denote L_F simply by F.

It is seen from the definition of orthogonal factors in Tjur (1984) that two factors are orthogonal if and only if the corresponding subspaces are g.o. Since $\mathcal{D}$ is assumed to consist of orthogonal factors, the lattice $\mathcal{L} = \mathcal{L}_{\mathcal{D}}$ generated by $\mathcal{M} = \mathcal{M}_{\mathcal{D}} := \{L_F \mid F \in \mathcal{D}\}$ is g.o., therefore distributive and finite. Note that $\mathcal{D}$ and $\mathcal{M}_{\mathcal{D}}$ are isomorphic posets. It is seen from (3.23) that there is no loss of generality if one supposes that $\mathcal{D}$ is closed under $\wedge$; hence, as in Tjur (1984), we subsequently assume this.

In order to complete the comparison of our treatment with that of Tjur (1984) for this special case of orthogonal factor-generated designs, it is seen that only one question remains: what is the relation between the poset $J(\mathcal{L}_{\mathcal{D}})$ of join-irreducible elements and the poset $\mathcal{M}_{\mathcal{D}}$ of factor subspaces? It is routine to check that in general one has $\mathcal{M}_{\mathcal{D}} \supseteq J(\mathcal{L}_{\mathcal{D}})$; furthermore, this inclusion may be strict (cf. our Example 3.11 and subsequent Remark 3.4). When $\mathcal{M}_{\mathcal{D}} \supset J(\mathcal{L}_{\mathcal{D}})$ all terms in Tjur's decomposition (p. 42) and ANOVA table (Section 5) that are indexed by $F \in \mathcal{M}_{\mathcal{D}} \setminus J(\mathcal{L}_{\mathcal{D}})$ will be trivially 0. By contrast, no term in our decompositions (3.6-7) and ANOVA table (3.9) is trivial.

When $\mathcal{L} = \mathcal{L}_{\mathcal{D}}$ is generated by a set of orthogonal factors, the quantities q_M, SS_M, and n_M, $M \in J(\mathcal{L}_{\mathcal{D}})$ $(\subseteq \mathcal{M}_{\mathcal{D}})$ needed in Remark 3.2 are immediately obtained from (3.19-22). Thus to complete the determination of the contrast estimators and ANOVA table, as outlined in Remark 3.2, it remains only to find the subposet $J(\mathcal{L}_{\mathcal{D}})$ of $\mathcal{M}_{\mathcal{D}}$ and to find the Möbius function for $J(\mathcal{L}_{\mathcal{D}})$.

3.7. Examples.

In the remainder of this section we present a series of examples to illustrate the specification and analysis of g.o. designs of linear models. In each example, $V = \mathbb{R}^I$ and δ is the usual inner product. Many of our examples treats factor-generated designs, in which (following (Tjur (1984)) we identify the isomorphic posets $\mathcal{D}$ and $\mathcal{M}_{\mathcal{D}}$. Also in each figure below, the elements of $J(\mathcal{L})$ are circled on the graph of the poset $\mathcal{L}$.

Example 3.1. (Homogeneous observations $\equiv$ the i.i.d. case). Let $\mathcal{D} = \{I,0\}$. Then $\mathcal{L}_{\mathcal{D}} = J(\mathcal{L}_{\mathcal{D}}) = \mathcal{D}$ (see Figure 3.1) and the Möbius function for $J(\mathcal{L}_{\mathcal{D}})$ is specified by $\mu(0,I) = -1$.□

Example 3.2. (One-way analysis of variance). Let G be a finite set with

$|G|\geq 2$ and let I_g be finite sets with $|I_g|\geq 1$, $g \in G$. The set G indexes the distinct groups of observations, while the set I_g indexes the observations within the group with index $g \in G$. Suppose that there exists at least one $g \in G$ such that $|I_g|\geq 2$. Set $I = \dot{U}(I_g|g\in G)$ and let $\mathcal{D} = \{I,G,0\}$, where the factor $\bar{G}:I \rightarrow G$ is given by $\bar{G}(gi) = g$, $gi\in I_g$, $g\in G$. (The notation gi is the usual subscript notation indicating an observation within group g.) Here $\mathcal{L}_{\mathcal{D}} = J(\mathcal{L}_{\mathcal{D}}) = \mathcal{D}$ is a chain of three elements (see Figure 3.2) and the Möbius function for $J(\mathcal{L}_{\mathcal{D}})$ is specified in Remark 1.1.□

Example 3.3. (Comparison of several one-way analyses of variance). Let H be a finite set with $|H|\geq 2$, let G_h be finite sets with $|G_h|\geq 2$, $h\in H$ and let I_{hg} be finite sets $hg\in G_h$, $h\in H$. Suppose that there exist at least one $h \in H$ and one $hg\in G_h$ such that $|I_{hg}|\geq 2$. Set $I = \dot{U}(\dot{U}(I_{hg}|g\in G_h)|h\in H)$ and set $\mathcal{D} = \{I,G,H,0\}$, where $G = \dot{U}(G_h|g\in H)$, $\bar{H}(hgi)=h$, $\bar{G}(hgi)=hg$, $hgi\in I_{gh}$, $hg\in G_h$, $h\in H$. As in the previous examples $\mathcal{L}_{\mathcal{D}} = J(\mathcal{L}_{\mathcal{D}}) = \mathcal{D}$ is a chain (see Figure 3.3) and the Möbius function is specified in Remark 1.1.□

Example 3.4. (Two-way analysis of variance). Let $G = R\times C$, where R (for rows) and C (for columns) are finite sets with $|R|\geq 2$, $|C|\geq 2$ and let $I_{(r,c)}$ be a finite set with $|I_{(r,c)}|\geq 1$, $(r,c)\in R\times C$. Suppose that there exists at least one $(r,c)\in R\times C$ such that $|I_{(r,c)}|\geq 2$. Set $I=\dot{U}(I_{(r,c)}|(r,c)\in R\times C)$ and let $\mathcal{D} = \{I,R\times C,R,C,0\}$, where $\overline{R\times C}((r,c)i) = (r,c)$, $\bar{R}((r,c)i) = r$, $\bar{C}((r,c)i) = c$, $(r,c)i\in I_{rc}$, $(r,c)\in R\times C$. Then $\mathcal{D}$ is closed under $\wedge$ and the corresponding subspaces are g.o. if and only if $|I_{(r,c)}||I| = |I_{r\cdot}||I_{\cdot c}|$, $(r,c)\in R\times C$, where $|I_{r\cdot}|=\Sigma(|I_{(r,c)}||c\in C)$, $r\in R$, and $|I_{\cdot c}|=\Sigma(|I_{(r,c)}||r\in R)$, $c\in C$. The lattice $\mathcal{L}_{\mathcal{D}}$ consists of the five factor subspaces and the subspace L_R+L_C. Thus $\mathcal{L}_{\mathcal{D}} \supset J(\mathcal{L}_{\mathcal{D}}) = \mathcal{D}$ (see Figure 3.4). The Möbius function is specified by $\mu(R\times C,I) = \mu(R,R\times C) = \mu(C,R\times C) = \mu(0,R) = \mu(0,C) = -1$ and $\mu(0,R\times C) = 1$.□

Example 3.5. (Multi-way analysis of variance with one observation per cell). Let J be a finite set with $|J|\geq 2$ and let K_j be finite sets with $|K_j|\geq 2$, $j\in J$. Set $I = X(K_j|j\in J)$. For any $B\subseteq J$ define $F_B = X(K_j|j\in B)$ and let $\overline{F}_B: I\to F_B$ be the projection mapping i.e., $\overline{F}_B((k_j|j\in J))=(k_j|j\in B)$, $(k_j|j\in J)\in I$. Set $\mathcal{D} = \{F_B|B\in\mathcal{D}(J)\}$ (see Remark 1.1). Then $\mathcal{D}$ is closed under $\wedge$ since $F_B\wedge F_C = F_{B\cap C}$, $B,C\in\mathcal{D}(J)$. Furthermore $F_B\vee F_C = F_{B\cup C}$ shows that $\mathcal{D}$ is also closed under $\vee$. Also, all factor subspaces are g.o. It is seen that $\mathcal{L}_{\mathcal{D}} \supset J(\mathcal{L}_{\mathcal{D}}) = \mathcal{D}$ (see Figure 3.5 for $J = \{a,b,c\}$) and, since $\mathcal{D}$ is isomorphic to $\mathcal{D}(J)$ as a poset, the Möbius function is specified by $\mu(B,C) = (-1)^{|C|-|B|}$, $C\supseteq B$ (cf. Remark 1.1). Note that the ANOVA table is indexed by $\mathcal{D}(J)$.□

Example 3.6. (Comparison of several two-way analyses of variance with one observation per cell). Let G be a finite set with $|G|\geq 2$ and let R_g and C_g be finite sets with $|R_g|\geq 2$ and $|C_g|\geq 2$, $g\in G$. Set $I = \dot{\cup}(R_g\times C_g|g\in G)$ and $\mathcal{D} = \{I,R_G,C_G,G,0\}$, where $R_G = \dot{\cup}(R_g|g\in G)$, $C_G = \dot{\cup}(C_g|g\in G)$, $\overline{R}_G(g(r,c))=gr$, $\overline{C}_G(g(r,c))=gc$ and $\overline{G}(g(r,c))=g$, $g(r,c)\in R_g\times C_g$, $g\in G$. The factors are g.o. and $\mathcal{L}_{\mathcal{D}} \supset J(\mathcal{L}_{\mathcal{D}}) = \mathcal{D}$ (see Figure 3.6). The Möbius function is specified by $\mu(R_G,I) = \mu(C_G,I) = \mu(G,R_G) = \mu(G,C_G) = \mu(0,G) = -1$ and $\mu(G,I) = 1$.□

Example 3.7. (Split-plot design). Let G, R_g, C_g, $g \in G$ be as in Example 3.6 and suppose that R_g=R is independent of $g \in G$. Under this assumption we can add the factor R given by $\overline{R}(g(r,c))=r$ to the factors in Example 3.6 to obtain the split-plot design $\mathcal{D} = \{I, R_G, C_G, G, R, 0\}$. As in the previous examples every factor subspace is join-irreducible, i.e., $J(\mathcal{L}_{\mathcal{D}}) = \mathcal{D}$ (see Figure 3.7). The Möbius function is specified by $\mu(R_G, I) = \mu(C_G, I) = \mu(G, C_G) = \mu(G, R_G) = \mu(R, R_G) = \mu(0, R) = \mu(0, G) = -1$ and $\mu(G, I) = \mu(0, R_G) = 1$.□

Example 3.8. (Two-way comparison of several two-way analyses of variance with one observation per cell). In Example 3.6 replace G by G×H, where H is a finite set with $|H| \geq 2$. Then we can add to the factors in Example 3.6 the factors G and H given by $\overline{G}((g,h)(r,c))=g$ and $\overline{H}((g,h)(r,c))=h$, $(g,h)(r,c) \in R_{gh} \times C_{gh}$, $(g,h) \in G \times H$, to obtain $\mathcal{D} = \{I, R_{G\times H}, C_{G\times H}, G\times H, G, H, 0\}$. Here again $\mathcal{L}_{\mathcal{D}} \supset J(\mathcal{L}_{\mathcal{D}}) = \mathcal{D}$ (see Figure 3.8). The Möbius function is specified by $\mu(R_{G\times H}, I) = \mu(C_{G\times H}, I) = \mu(G\times H, R_{G\times H}) = \mu(G\times H, C_{G\times H}) = \mu(G, G\times H) = \mu(H, G\times H) = \mu(0, G) = \mu(0, H) = -1$ and $\mu(G\times H, I) = \mu(0, G\times H) = 1$.□

Example 3.9. (Two-way split-plot design). In Example 3.8, suppose that $R_{(g,h)} = R_g$, $g \in G$ is independent of $h \in H$ (cf. Example 3.7). Then we can add the factor $R_G = \dot{\cup}(R_g | g \in G)$ given by $\overline{R}_G((g,h)(r,c))=gr$, $(g,h)(r,c) \in R_g \times C_{gh}$, $(g,h) \in G \times H$, to the factors in Example 3.8 to obtain $\mathcal{D} = \{I, R_{G\times H}, R_G, C_{G\times H}, G\times H, G, H, 0\}$. Again $\mathcal{L}_{\mathcal{D}} \supset J(\mathcal{L}_{\mathcal{D}}) = \mathcal{D}$ (see Figure 3.9). The Möbius function is specified by $\mu(R_{G\times H}, I) = \mu(C_{G\times H}, I) = \mu(G\times H, R_{G\times H}) = \mu(G\times H, C_{G\times H}) = \mu(G, R_G) = \mu(G, G\times H) = \mu(H, G\times H) = \mu(0, G) = \mu(0, H) = -1$ and $\mu(G\times H, I) = \mu(G, R_{G\times H}) = \mu(0, G\times H) = 1$.□

<u>Example 3.10</u>. (Two-way analysis of variance with one observation per cell and subdivision of rows and colums) In Example 3.9 suppose furthermore that $C_{(g,h)}=C_h$, $h\in H$, is independent of $g\in G$. Then we can add the factor $C_H = \dot{\cup}(C_h|h\in H)$ given by $\overline{C}_H((g,h)(r,c))=hc$, $ghrc\in R_g\times C_h$, $(g,h)\in G\times H$ to the factors in Example 3.9, obtaining $\mathcal{D} = \{I, R_{G\times H}, R_G, C_{G\times H}, C_H, G\times H, G, H, 0\}$. Again $\mathcal{L}_{\mathcal{D}} \supset J(\mathcal{L}_{\mathcal{D}}) = \mathcal{D}$ (see Figure 3.10) and the Möbius function is specified by $\mu(R_{G\times H}, I) = \mu(C_{G\times H}, I) = \mu(R_G, R_{G\times H}) = \mu(C_H, C_{G\times H}) = \mu(G\times H, R_{G\times H}) = \mu(G\times H, C_{G\times H}) = \mu(G, R_G) = \mu(H, C_H) = \mu(G, G\times H) = \mu(H, G\times H) = \mu(0,G) = \mu(0,H) = -1$ and $\mu(G\times H, I) = \mu(G, R_{G\times H}) = \mu(H, C_{G\times H}) = \mu(0, G\times H) = 1$.□

In each of the preceding examples, $\mathcal{D} = J(\mathcal{L}_{\mathcal{D}})$, the poset $J(\mathcal{L}_{\mathcal{D}})$ is a lattice (i.e., $\mathcal{D}$ is closed under maximum ($\vee$) in $\mathcal{F}(I)$) and, in fact, even a distributive lattice. Thus the Möbius function for $J(\mathcal{L}_{\mathcal{D}})$ readily can be determined from Remark 1.3. The next example presents a case where again $\mathcal{D} = J(\mathcal{L}_{\mathcal{D}})$, $J(\mathcal{L}_{\mathcal{D}})$ is a lattice, but not a distributive lattice.

<u>Example 3.11</u>. (Latin square). Let R (for "Rows"),C (for " Columns") and G (for the third index) be finite sets with $|R| = |C| = |G| \geq 3$ and let $I \subseteq R\times C\times G$ be a subset such that the factors $\overline{R\times C}$, $\overline{C\times G}$ and $\overline{R\times G}$ on $R\times C\times G$, when restricted to I, are equivalent to $\overline{I}$. Set $\mathcal{D} = \{I, R, C, G, 0\}$. Then $\mathcal{L}_{\mathcal{D}} \supset J(\mathcal{L}_{\mathcal{D}}) = \mathcal{D}$ (see Figure 3.11) and $\mathcal{D}$ is not distributive. The Möbius function is specified by $\mu(R,I) = \mu(C,I) = \mu(G,I) = \mu(0,R) = \mu(0,C) = \mu(0,G) = -1$ and $\mu(0,I) = 2$.□

<u>Remark 3.4</u>. If we take $|R|=|C|=|G|=2$ in Example 3.11 we obtain an example where $\mathcal{D} \supset J(\mathcal{L}_{\mathcal{D}})$ (see Figure 3.12) and $J(\mathcal{L}_{\mathcal{D}})$ is <u>not</u> a lattice. (Note that

$\mathcal{D}$ is still closed under V in $\mathcal{F}(I)$.) In this example our ANOVA table (3.9) is different from Tjur (1984) (though not essentially different) since ours does not include term indexed by the factor I. Nevertheless in Tjur's ANOVA table the term indexed by I will be trivially 0. Thus, factor subspaces are not always join-irreducible.□

In the following three examples $\mathcal{L}$ is not a factor-generated analysis of variance design, hence (strictly speaking) these examples fall outside realm of Tjur's paper, but the sample space decompositions (3.6-7), the contrast estimates (3.8), and the ANOVA table (3.9) are still well-defined. Of course, the formulas (3.19-22) are not applicable for those subspaces $L \in J(\mathcal{L})$ that are not factor subspaces. Therefore in order to analyze these designs, we must find $J(\mathcal{L})$, calculate q_L, SS_L, and n_L for the remaining subspaces, and determine the Möbius function on $J(\mathcal{L})$.

Example 3.12. (Regression analysis). In the one-way analysis of variance (Example 3.2), assume that $|G|\geq 3$ and suppose that the qualitative index $g\in G$ is quantified by the family $(t_g \in \mathbb{R}|g\in G)$, where $|\{t_g|g\in G\}| \geq 2$. Set $\mathcal{L}$ = $\{\mathbb{R}^I, L_G, T, L_0\}$, where T is the subspace given by $P_T = \mathbb{R}^2$ and $u_T(\alpha,\beta)$ = $(\alpha+\beta t_g|gi\in I) \in \mathbb{R}^I$. Since $\mathcal{L}$ is a chain, $J(\mathcal{L}) = \mathcal{L}$ and the Möbius function is trivial. Set $S_t = \Sigma(|I_g|t_g \; |g\in G)$, $\bar{t}_{\bullet} = S_t/|I|$, $SS_t = \Sigma(|I_g|t_g^2 \; |g\in G)$, $SSD_t = SS_t - S_t^2/|I|$, $SP_t(x) = \Sigma(S_g(x)t_g \; |g\in G)$, and $SPD_t(x) = SP_t(x) - S_tS_0(x)/|I|$, where $S_g(x)$ and $S_0(x)$ are defined as in (3.20), $x\in\mathbb{R}^I$. Then

$$q_T(x) = (\hat{\alpha}(x)+\hat{\beta}(x)t_g|gi\in I),$$
$$SS_T(x) = (\hat{\alpha}(x)+\hat{\beta}(x)\bar{t}_{\bullet})^2+\hat{\beta}(x)^2 SSD_t,$$
$$n_T = 2,$$

where $\hat{\beta}(x) = SPD_t(x)/SSD_t$ and $\hat{\alpha}(x) = S_0(x)/|I|-\hat{\beta}(x)\bar{t}_{\bullet}$, $x\in\mathbb{R}^I$.□

Example 3.13. (Comparison of several regression lines). In Example 3.3, suppose that the qualitative index $hg \in G = \dot{\cup}(G_h|h\in H)$ is quantified by the family $(t_{hg}\in\mathbb{R}|hg\in G_h, h\in H)$, where $|\{t_{hg}|hg\in G_h\}| \geq 2$ for at least one $h \in H$. Set $\mathcal{L} = \{\mathbb{R}^I, L_G, L, T+L_H, T, L_H, L_0\}$, where T and L are subspaces given by $P_T = \mathbb{R}^2$, $u_T(\alpha,\beta) = (\alpha+\beta t_{hg}|hgi\in I) \in \mathbb{R}^I$, $P_L = (\mathbb{R}^2)^H$, $u_L(((\alpha_h,\beta_h)|h\in H)) = (\alpha_h+\beta_h t_{hg}|hgi\in I) \in \mathbb{R}^I$, respectively. Here $\mathcal{L}$ is a distributive (not necessarily g.o.) lattice with $J(\mathcal{L}) = \mathcal{L}\backslash\{T+L_H\}$ (see Figure 3.13). The Möbius function on $J(\mathcal{L})$ is specified by $\mu(L_G,\mathbb{R}^I) = \mu(L,L_G) = \mu(T,L) = \mu(L_H,L) = \mu(L_0,T) = \mu(L_0,L_H) = -1$ and $\mu(L_0,L) = 1$. Set $S_t^h = \Sigma(|I_{hg}|t_{hg}|g\in G_h)$, $|I_{h\bullet}| = \Sigma(|I_{hg}||g\in G_h)$, $\bar{t}_{h\bullet} = S_t^h/|I_{h\bullet}|$, $SS_t^h = \Sigma(|I_{hg}|t_{hg}^2|g\in G_h)$, $SSD_t^h = SS_t^h-(S_t^h)^2/|I_{h\bullet}|$, $h\in H$. Also set $S_t = \Sigma(S_t^h|h\in H)$, $\bar{t}_{\bullet\bullet} = S_t/|I|$, $SSD_t^{\bullet} = \Sigma(SSD_t^h|h\in H)$, $SS_t^{\bullet} = \Sigma(SS_t^h|h\in H)$, $SSD_t = SS_t^{\bullet}-(S_t^{\bullet})^2/|I|$. Furthermore, set $SP_t^h(x) = \Sigma(t_{hg}S_{hg}(x)|g\in G_h)$, $SPD_t^h(x) = SP_t^h(x)-S_t^hS_h(x)/|I_{h\bullet}|$, $h\in H$, and $SPD_t^{\bullet}(x) = \Sigma(SPD_t^h(x)|h\in H)$, $SP_t^{\bullet}(x) = \Sigma(SP_t^h(x)|h\in H)$, $SPD_t(x) = SP_t^{\bullet}(x)-S_t^{\bullet}S_0(x)/|I|$, where $S_{hg}(x)$, $S_h(x)$ and $S_0(x)$ are defined as in (3.20), $x \in \mathbb{R}^I$. Then

$$q_L(x) = (\hat{\alpha}_h(x)+\hat{\beta}_h(x)t_{hg}|hgi\in I),$$
$$SS_L(x) = \Sigma((\hat{\alpha}_h(x)+\hat{\beta}_h(x)\bar{t}_{h\bullet})^2+\hat{\beta}_h(x)^2SSD_t^h|h\in H),$$
$$n_L = 2|H|,$$

and

$$q_T(x) = (\hat{\alpha}(x)+\hat{\beta}(x)t_{hg}|hgi\in I),$$
$$SS_T(x) = (\hat{\alpha}(x)+\hat{\beta}(x)t_{\bullet\bullet})^2|I|+\hat{\beta}(x)^2SSD_t,$$
$$n_T = 2,$$

where $\hat{\beta}_h(x) = SPD^h_t(x)/SSD^h_t$, $\hat{\alpha}_h(x) = S_h(x)/|I_{h\cdot}|-\hat{\beta}_h(x)\bar{t}_{h\cdot}$, $h\in H$, $\hat{\beta}(x) = SPD_t(x)/SSD_t$, and $\hat{\alpha}(x) = S_0(x)/|I|-\hat{\beta}(x)\bar{t}_{\cdot\cdot}$. The subspaces L_H and T are not in general g.o.; nevertheless the projection q_M on $M := T+L_H$ is easely found to be $q_M(x) = (\tilde{\alpha}_h(x)+\tilde{\beta}(x)t_{hg}|hgi\in I) \in \mathbb{R}^I$, where $\tilde{\beta}_h(x) = SPD^{\bullet}_t(x)/SSD^{\bullet}_t$ $(= \Sigma(SSD^h_t\hat{\beta}_h(x)|h\in H)/SSD^{\bullet}_t)$ and $\tilde{\alpha}_h(x) = S_h(x)/|I_{h\cdot}|-\tilde{\beta}(x)\bar{t}_{h\cdot}$, $h\in H$, $x\in\mathbb{R}^I$. Also $SS_M(x) = \Sigma((\tilde{\alpha}_h(x)+\tilde{\beta}(x)\bar{t}_{h\cdot})^2|I_{h\cdot}||h\in H)+\tilde{\beta}(x)^2SSD^{\bullet}_t$, and $\dim(M) = |H|+1$. The subspaces T and L_H are g.o. if and only if $\bar{t}_{h\cdot} = \bar{t}_{\cdot\cdot}$, $h\in H$, and only in this case may the ANOVA table be defined. Furthermore, in this case $\tilde{\beta}(x) = \hat{\beta}(x)$, $x\in\mathbb{R}^I$.□

From a distributive lattice point of view, the next example is identical to the Latin square design in Example 3.11.

<u>Example 3.14</u>. (Two-way analysis of variance with one observation per cell and regression in the rows). In Example 3.4 suppose that $|I_{(r,c)}| = 1$, $(r,c)\in R\times C$, i.e., $I = R\times C$. Also suppose that the family $(t_r\in\mathbb{R}|r\in R)$ is a quantification of the qualitative index $r\in R$ such that $|\{t_r|r\in R\}| \geq 2$ and $\Sigma(t_r|r\in R) = 0$. Let $\mathcal{L} = \{\mathbb{R}^I, L_R+U+V, U+V, L_R+V, L_R+U, V, U, L_R, T\}$, where the subspaces T, U and V are given by $P_T = \mathbb{R}^2$, $u_T(\alpha,\gamma) = (\alpha+\gamma t_r|(r,c)\in I) \in \mathbb{R}^I$, $P_U = \mathbb{R}\times\mathbb{R}^C$, $u_U(\alpha,(\gamma_c|c\in C)) = (\alpha+\gamma_c t_r|(r,c)\in I) \in \mathbb{R}^I$, $P_V = \mathbb{R}^C\times\mathbb{R}$, $u_V((\alpha_c|c\in C),\gamma) = (\alpha_c+\gamma t_r|(r,c)\in I) \in \mathbb{R}^I$, respectively. Then $J(\mathcal{L}) = \{\mathbb{R}^I, V, U, L_R, T\}$ (see Figure 3.14) and V, U, L_R are g.o., hence $\mathcal{L}$ is g.o. Clearly the lattice $\mathcal{L}$ and the poset $J(\mathcal{L})$ are isomorphic to the lattice $\mathcal{L}_{\mathcal{D}}$ and the poset $J(\mathcal{L}_{\mathcal{D}})$ in Example 3.11, hence the Möbius functions are the same. Note that $L_R+L_C = L_R+V$. Set $SS_t = \Sigma(t_r^2|r\in R)$, $SP^c_t(x) = \Sigma(t_r x_{(r,c)}|r\in R)$, and $SP^{\bullet}_t(x) = \Sigma(SP^c_t(x)|c\in C)$, $x = (x_{(r,c)}|(r,c)\in R\times C) \in \mathbb{R}^I$.

From Examples 3.12 and 3.13 it follows that

$$q_U(x) = (\hat{\alpha}(x)+\hat{\gamma}_c(x)t_r \,|\, (r,c)\in I)\in\mathbb{R}^I$$
$$SS_U(x) = |I|\hat{\alpha}(x)^2+SS_t\Sigma(\hat{\gamma}_c(x)^2\,|\,c\in C),$$
$$n_U = |C|+1,$$

$$q_V(x) = (\hat{\alpha}_c(x)+\hat{\gamma}(x)t_r \,|\, (r,c)\in I),$$
$$SS_V(x) = |R|\Sigma(\hat{\alpha}_c(x)^2\,|\,c\in C)+|C|\hat{\gamma}(x)^2SS_t,$$
$$n_V = |C|+1,$$

and

$$q_T(x) = (\hat{\alpha}(x)+\hat{\gamma}(x)t_r \,|\, (r,c)\in I),$$
$$SS_T(x) = |I|\hat{\alpha}(x)^2+|C|\hat{\gamma}(x)^2SS_t,$$
$$n_T = 2,$$

where $\hat{\alpha}_c(x) = S_c(x)$, $\hat{\alpha}(x) = S_0(x)$, $\hat{\gamma}_c(x) = SP_t^c(x)/SS_t$ and $\hat{\gamma}(x) = SP_t^{\bullet}(x)/(SS_t|C|)$, $c\in C$, and where $S_c(x)$ and $S_0(x)$ are defined as in (3.20), $x = (x_{(r,c)}\,|\,(r,c)\in R\times C) \in \mathbb{R}^I$.□

In each of Examples (3.1-14) the poset $J(\mathcal{L}) \subseteq \mathcal{L}$ is closed under minimum ("$\wedge$") in $\mathcal{L}$. For a general distributive lattice $\mathcal{L}$ of subspaces, however, this property of $J(\mathcal{L})$ need not be valid, e.g., the lattice $\mathcal{L}$ in Figure 3.15, which, by Remark 2.1, can be represented as a g.o., hence distributive, lattice of subspaces.

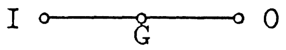
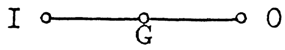
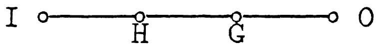
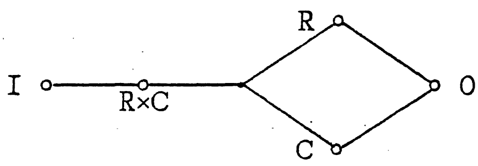
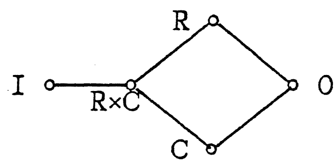
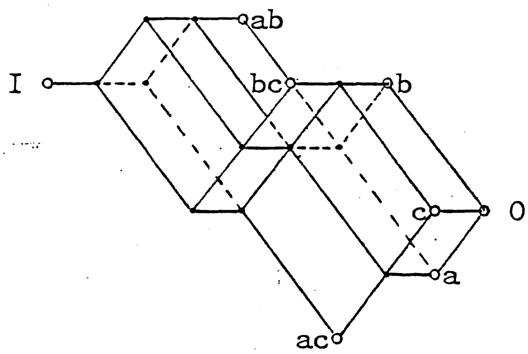
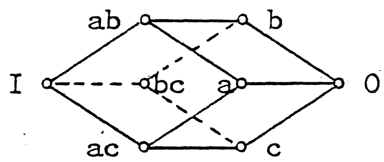
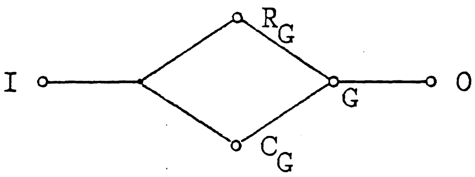
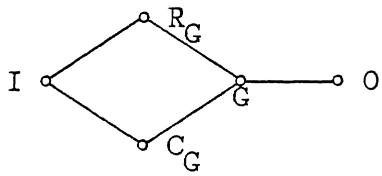
FIGURE	$\mathcal{L}$	$J(\mathcal{L})$
(3.1)		
(3.2)		
(3.3)		
(3.4)		
(3.5)		
(3.6)		

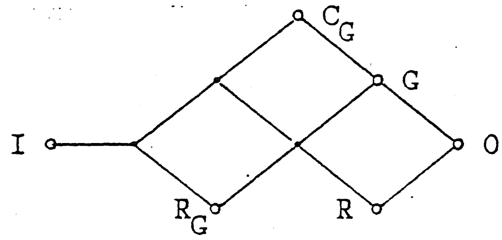
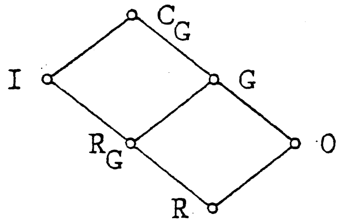
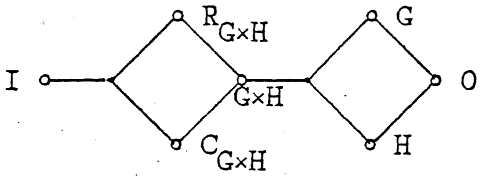
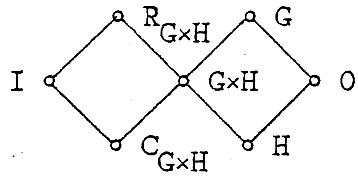
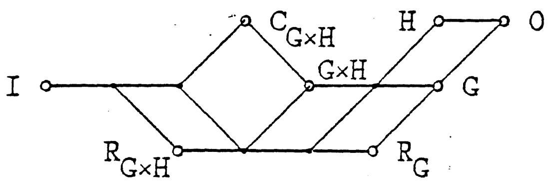
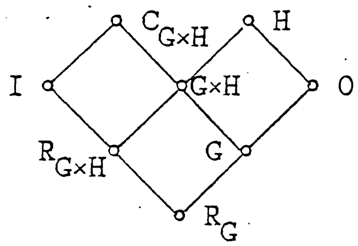
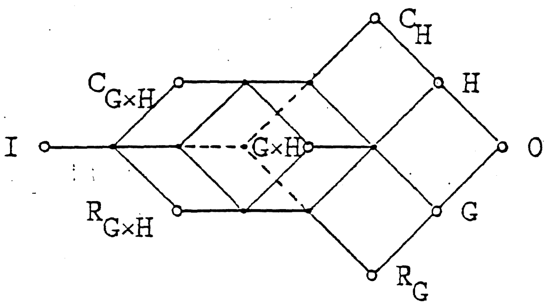
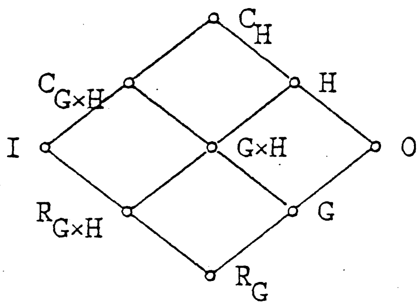
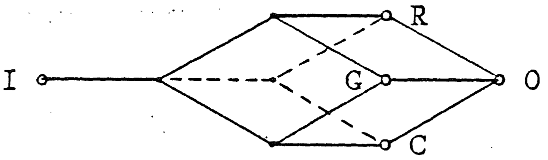
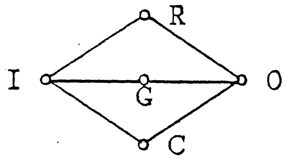
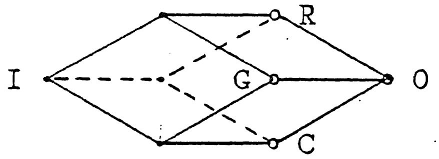
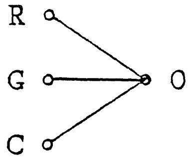
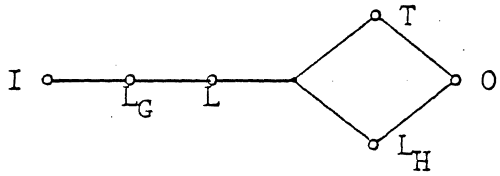
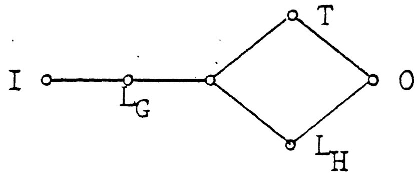
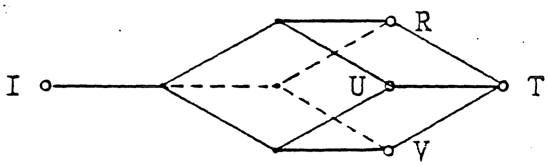
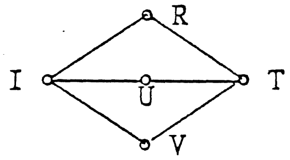
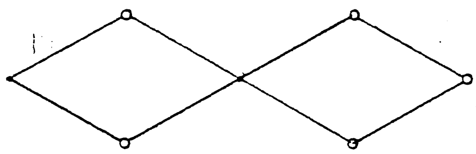
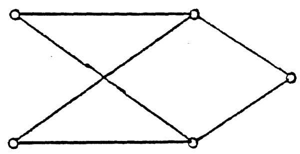
FIGURE	$\mathcal{L}$	$J(\mathcal{L})$
(3.7)		
(3.8)		
(3.9)		
(3.10)		

FIGURE	$\mathscr{L}$	$J(\mathscr{L})$
(3.11)		
(3.12)		
(3.13)		
(3.14)		
(3.15)		

4. VARIANCE COMPONENT DESIGNS.

4.1. Extending a design of linear models.

When analyzing an experiment by means of a linear model, or rather a design of linear models, it may occur that the smallest mean value subspace that fits the data is too complex to answer questions of fundamental interest for the experiment. For example in a two-way layout with significant interaction between treatments (rows) and individuals (columns), one is unable to estimate the effects of the treatments. One possible explanation for this occurence might be that the requirement that the covariance structure is known up to a positive constant is too restrictive. One way out of this difficulty is to allow a more complex covariance structure which may then permit a simpler mean value subspace to fit the data.

The extension of the covariance structure is usually limited by the requirement that the the ANOVA table still shall provide the necessary information for the analysis of the design. This approach has led to the extension of common analysis of variance designs to so-called variance component models (actually, designs), also called random effect models. The specific form of this extension is usually determined by declaring certain systematic (≡fixed) effects in the original design to be random effects. For example the variation between individuals may be declared to be random while the variation between treatments remains systematic.

Except for Tolver Jensen (1979) and Tjur (1984) the literature contains no clear definition of a variance component model or design. The main aim of this section is to define these concepts in a precise and general way and to show that the "canonical" variance component models that we define are exactly those that allow a complete solution of the likelihood inference problem using the associated ANOVA table. Furthermore our treatment of these designs will shed light on some of the well-known difficulties associated with so-called variance component models. It will be seen that these difficulties do not arise within the class of canonical variance component models.

4.2. Geometrically orthogonal variance component designs.

We shall define a canonical extension of a given class of a given g.o. design of linear models. We shall call this extension the geometrically orthogonal variance component design (g.o.v.c. design) associated with the given design of linear models.

For V and δ as in Section 3, let $\mathcal{L} \subseteq \mathcal{L}(V)$ be a g.o. sublattice of embedded subspaces with $V \in \mathcal{L}$ and $((P_L, u_L)|L\in\mathcal{L})$ an associated g.o. design as defined in Section 3. Let $\mathrm{Sub}^*(\mathcal{L})$ denote the set of all sublattices $\mathcal{M} \subseteq \mathcal{L}$ such that $V \in \mathcal{M}$. For $\mathcal{M} \in \mathrm{Sub}^*(\mathcal{L})$, let

$$V = \perp(W_M | M\in J(\mathcal{M})) \tag{4.1}$$

be the orthogonal decomposition of V wrt $\mathcal{M}$ (see Theorems 2.1 and 2.2), where $W_M = M \cap J(M)^{\perp}$, $M\in J(\mathcal{M})\setminus\{0\}$, $W_0 = 0$ (here $0 = 0_{\mathcal{M}}$). For $L \in \mathcal{L}$ and $\mathcal{M} \in \mathrm{Sub}^*(\mathcal{L})$, the canonical variance component model (c.v.c. model) determined by the pair $((P_L, u_L), \mathcal{M})$ (or simply $(L, \mathcal{M})$) is defined as follows: the observation space is V, the parameter space is $P_L \times \mathbb{R}_+^{J(\mathcal{M})}$, and the set of unknown probability measures on V consists of the normal distributions with mean $u_L(\xi)$ and precision

$$(4.2) \qquad (x,y) \to \Sigma(\sigma_M^{-2}\delta_M(x_M,y_M)\,|M \in J(\mathcal{M})),$$

where $x = (x_M|M\in J(\mathcal{M})) \in V$ and $y = (y_M|M\in J(\mathcal{M})) \in V$ (see (4.1)), δ_M is the restriction of δ to W_M, and $(\xi,(\sigma_M^2|M\in J(\mathcal{M}))) \in P_L\times \mathbb{R}_+^{J(\mathcal{M})}$. Clearly the c.v.c. model determined by $((P_L,u_L),\mathcal{M})$ is an extension of the original linear model determined by (P_L,u_L) (take $\sigma_M^2 = \sigma^2$, $M \in J(\mathcal{M})$).

Remark 4.1. If $\mathcal{M} \neq \mathcal{M}'$, then the c.v.c. models determined by $((P_L,u_L),\mathcal{M})$ and $((P_L,u_L),\mathcal{M}')$ may be identical in the sense that the sets of unknown probability measures are identical.□

Definition 4.1. The geometrically orthogonal variance component design (g.o.v.c. design) determined by the g.o. design $((P_L,u_L)|L\in\mathcal{L})$ is the set consisting of all canonical variance component models determined by $((P_L,u_L),\mathcal{M})$, $L \in \mathcal{L}$, $M \in \mathrm{Sub}^*(\mathcal{L})$.□

The subset obtained by fixing $\mathcal{M} = \{V\}$ is the original g.o. design. When there is no danger of confusion we shall refer simply to the g.o.v.c. design determined by $\mathcal{L}$.

Example 4.1. In the two-way analysis of variance design of Example 3.4, $\mathcal{L} = \mathcal{L}_{\mathcal{D}}$, $|\mathcal{L}_{\mathcal{D}}| = 6$ and $|\mathrm{Sub}^*(\mathcal{L}_{\mathcal{D}})| = 26$.□

4.3. ML estimation in a canonical variance component model.

Let $u:\mathcal{M} \to \mathcal{L}$ denote the embedding mapping $(u(M) = M)$. By Proposition 1.3, the mapping $\psi = J(u):J(\mathcal{L}) \to J(\mathcal{M})$ is a surjective poset homomorphism.

Lemma 4.1. Let $\perp(W_{M'}\,|M'\in J(\mathcal{M}))$ and $\perp(V_{L'}\,|L'\in J(\mathcal{L}))$ be the orthogonal decompositions of V with respect to $\mathcal{M}$ and $\mathcal{L}$ respectively. Then for $L \in \mathcal{L}$ and $M \in J(\mathcal{M})$,

(4.3) $$L = \perp(\perp(V_{L'}\,|L'\in J(\mathcal{L}),L'\subseteq L,\psi(L')=M')\,|M'\in J(\mathcal{M})),$$

(4.4) $$W_M = \perp(V_{L'}\,|L'\in J(\mathcal{L}),\psi(L')=M).$$

Proof: Since ψ is surjective it follows that $L = \perp(V_{L'}\,|L'\in J(\mathcal{L}),L'\subseteq L) = \perp(\perp(V_{L'}\,|L'\in J(\mathcal{L}),L'\subseteq L,\psi(L')=M')\,|M'\in J(\mathcal{M}))$. Since $[L'\subseteq M,\psi(L')=M'] \Leftrightarrow [M'\subseteq M,\psi(L')=M']$ whenever $M \in \mathcal{M}$, $M' \in J(\mathcal{M})$, and $L' \in J(\mathcal{L})$, in particular we have that

$$\begin{aligned} M &= \perp(\perp(V_{L'}\,|L'\in J(\mathcal{L}),L'\subseteq M,\psi(L')=M')\,|M'\in J(\mathcal{M})) \\ &= \perp(\perp(V_{L'}\,|L'\in J(\mathcal{L}),\psi(L')=M')\,|M'\in J(\mathcal{M}),M'\subseteq M). \end{aligned}$$

Then (4.4) follows from the uniqueness of the orthogonal decomposition wrt $\mathcal{M}$.□

From the lemma it follows that

(4.5) $$L = \perp(L^M|M\in J(\mathcal{M})),$$

where

(4.6) $$L^M := \perp(V_{L'}\,|L'\in J(\mathcal{L}),L'\subseteq L,\psi(L')=M) \subseteq W_M,\ M \in J(\mathcal{M})$$

Thus the c.v.c. model determined by $((P_L,u_L),\mathcal{M})$ is a product of linear models indexed by $J(\mathcal{M})$. The linear model corresponding to $M\in J(\mathcal{M})$ has observation space W_M, precision given by $\sigma_M^{-2}\delta_M$, and mean value subspace

L^M having the parametrization $u_L^M : P_L^M \to W_M$, where $P_L^M = \{\xi \in P_L | u_L(\xi) \in L^M\}$ and u_L^M is the restriction of u_L to P_L^M.

For any c.v.c. model in the g.o.v.c. design determined by $\mathcal{L}$, the ML-estimator and its distribution are thus readily obtained from this decomposition into a product of linear models. Furthermore, all of the quantities needed for the variance estimators σ_M^2, $M \in J(\mathcal{M})$, and their simultaneous distribution are obtained from the ANOVA table $((SSD_L(x), f_L) | L \in J(\mathcal{L}))$, $x \in V$, associated with the g.o. design determined by $\mathcal{L}$, cf. (3.8) and (3.9). Also, the mean value estimator $\hat{\xi}$ is the same as in the linear model given by L (and δ). The results are as follows. For the c.v.c. model determined by $((P_L, u_L), \mathcal{M})$, the ML estimator $(\hat{\xi}, (\hat{\sigma}_M^2 | M \in J(\mathcal{M})))$ of $(\xi, (\sigma_M^2 | M \in J(\mathcal{M}))) \in P_L \times \mathbb{R}_+^{J(\mathcal{M})}$ exists if and only if

$$(4.7) \qquad L^M \subset W_M, \quad M \in J(\mathcal{M}).$$

In this case it is unique and given by

$$(4.8) \qquad u_L \circ \hat{\xi} = q_L = \Sigma(r_{L'} | L' \in J(\mathcal{L}), L' \subseteq L),$$

$$(4.9) \qquad \hat{\sigma}_M^2 = m_M^{-1}(\Sigma(SSD_{L'} | L' \in J(\mathcal{L}), L' \not\subseteq L, \psi(L') = M)),$$

where

$$(4.10) \qquad m_M := \dim(W_M) = \Sigma(f_{L'} | L' \in J(\mathcal{L}), \psi(L') = M), \quad M \in J(\mathcal{M}).$$

The unbiased variance estimators

$$(4.11) \qquad s_M^2 = \frac{m_M}{m_M - l_M} \hat{\sigma}_M^2,$$

where

$$(4.12) \qquad l_M := \dim(L^M) = \Sigma(f_{L'}\,|L'\in J(\mathcal{L}),L'\subseteq L,\psi(L')=M),\ M\in J(\mathcal{M}).$$

To describe the distribution of the ML estimator, decompose $\xi \in P_L$ into $\xi = (\xi_M|M\in J(\mathcal{M})) \in X(P_L^M|M\in J(\mathcal{M}))$ and let $q_L^M:V\to V$ denote the orthogonal projection on L^M, $M \in J(\mathcal{M})$. Then the ML-estimator $\hat{\xi}_M$ for the component $\xi_M \in P_L^M$ is given by $u_L^M \circ \hat{\xi}_M = q_L^M = \Sigma(r_{L'}\,|L'\in J(\mathcal{L}),L'\subseteq L,\psi(L')=M)$, $M\in J(\mathcal{M})$. The distribution of the family $((\hat{\xi}_M,\hat{\sigma}_M^2)\,|M\in J(\mathcal{M}))$ is described as follows: the $2\times|J(\mathcal{M})|$ components are independent, $\hat{\xi}_M$ is normally distributed on P_L^M with mean ξ_M and precision $\sigma_M^{-2}\delta_M\circ(u_L^M\times u_L^M)$, and $\hat{\sigma}_M^2$ is χ^2-distributed with m_M-l_M degrees of freedom and scale σ_M^2/m_M, $M \in J(\mathcal{M})$.

<u>Remark 4.2</u>. Consider $M \in J(\mathcal{M})$ such that $M \subseteq L$. If $\psi(L') = M$ for some $L' \in J(\mathcal{L})$ then $L' \subseteq M$ and thus also $L' \subseteq L$. This shows that $L^M = W_M$. It is therefore seen from (4.7) that the condition

$$(4.13) \qquad M \not\subseteq L,\ \forall\, M \in J(\mathcal{M})$$

on $L \in \mathcal{L}$ is necessary for the existence of the ML estimator. In particular if $0_{\mathcal{L}} \in J(\mathcal{M})$ then (4.13) cannot hold, so the ML estimator cannot exist. If $J(\mathcal{M}) \subseteq J(\mathcal{L})$ then the condition (4.13) also becomes sufficient, since $[L^M = W_M,\ \psi(M) = M] \Rightarrow M \subseteq L$. The condition (4.13) can be interpreted as follows: if there is stochastic variation in the subspae $M \in J(\mathcal{M})$ then there cannot be a systematic effect associated with any $L \in \mathcal{L}$, that is "higher" than M, i.e., $L \supseteq M$.□

Example 4.1 (continued). Consider $L = L_C$ ($\in \mathscr{L}_{\mathscr{D}}$) and $\mathscr{M} = \{\mathbb{R}^I, L_{R\times C}, L_R\}$:= $\{I, R\times C, R\}$ ($\in \mathrm{Sub}^*(\mathscr{L}_{\mathscr{D}})$). Then $\mathscr{M} = J(\mathscr{M}) \subseteq J(\mathscr{L}_{\mathscr{D}})$. From (4.4), $W_I = V_I$, $W_{R\times C} = V_{R\times C} \perp V_C$, $W_R = V_R \perp V_0$, while $L_C^I = \{\underset{\sim}{0}\}$, $L_C^{R\times C} = V_C$, and $L_C^R = V_0$ from (4.6). (In the case where all $|I_{(r,c)}|$ are the same, we recognize the c.v.c. model determined by this $(L,\mathscr{M})$ as the extension to the maximal parameter domain of the two-way layout with random interaction effect, random row effect, and systematic column effect – see § 4.6 and Tjur (1984), § 7.6.) It is easy to verify (4.7) (and (4.13)), hence the ML estimator exists.

Next consider $L = L_R+L_C$ and $\mathscr{M}$ as before. Since $L_R+L_C \supseteq L_R \in J(\mathscr{M})$, (4.13) fails and the ML estimator does not exist. In fact, from (4.6) $(L_R+L_C)^I = \{\underset{\sim}{0}\}$, $(L_R+L_C)^{R\times C} = V_C$, and $(L_R+L_C)^R = V_R \perp V_0 = W_R(!)$.

Altogether there are 12 sublattices $\mathscr{M} \in \mathrm{Sub}^*(\mathscr{L}_{\mathscr{D}})$ such that $0 \notin \mathscr{M}$. For each such $\mathscr{M}$, the ML estimator exists for the c.v.c. model determined by $(L,\mathscr{M})$, $L \in \mathscr{L}_{\mathscr{D}}$, if and only if (4.13) holds. In this example it can be seen that 24 such c.v.c. models exist.□

4.4. Variance component models in a multi-way layout (Jensen (1979)).

Jensen (1979) studied the broad class of so-called variance component models (= random effect models) associated with the multi-way layout with one observation per cell (see our Example 3.5). One of his main results, Lemma 4.2 below, characterizes the subclass of such models which, in this example, are also c.v.c. models in our sense. Our treatment of so-called random effect models in this case (see § 4.6) was strongly influenced by his results, which we now review.

In the context of our Example 3.5, for every $B \in \mathscr{D}(J)$ and $k(B)$:=

$(k_j|j\in B) \in F_B$ let $Y^B_{k(B)}$ be a normally distributed stochastic variable with values in $\mathbb{R}$, mean value $\xi^B_{k(B)} \in \mathbb{R}$, and variance $\omega_B \geq 0$. Suppose, furthermore, that all the $Y^B_{k(B)}$, $k(B) \in F_B$, $B \in \mathcal{D}(J)$, are independent and set

$$(4.14) \qquad X_{k(J)} = \sum\left[Y^B_{k(B)} \middle| B \in \mathcal{D}(J)\right],$$

where $k(B) = \bar{F}_B(k(J))$, $B \in \mathcal{D}(J)$, and $X = (X_{k(J)}|k(J)\in F_J)$. Define

$$(4.15) \qquad Y^B := (Y^B_{k(B)}|k(B)\in F_B);$$

note that

$$(4.16) \qquad E(Y^B) = \xi^B := (\xi^B_{k(B)}|k(B)\in F_B)$$

and

$$(4.17) \qquad X = \Sigma(Y^B|B\in\mathcal{D}(J)).$$

The family of normal distributions of X parametrized by $(\xi^B|B\in\mathcal{D}(J)) \in \times(\mathbb{R}^{F_B}|B\in\mathcal{D}(J))$ and $(\omega_B|B\in\mathcal{D}(J)) \in [0,\infty[^{\mathcal{D}(J)}$ is thus a statistical model.

Consider the two hypotheses $H_{\mathcal{T}}$ and $H_{\mathcal{B}}$, where $\mathcal{T} \subseteq \mathcal{D}(J)$ and $\mathcal{B} \subseteq \mathcal{D}(J)$, defined as follows:

$$(4.18) \qquad H_{\mathcal{T}}\colon\ \xi^T = 0 \text{ if and only if } T \notin \mathcal{T},$$

$$(4.19) \qquad H_{\mathcal{B}}\colon\ \omega_B > 0 \text{ if and only if } B \in \mathcal{B}.$$

One usually thinks of Y^B, $B \in \mathcal{B}$, as the random effect and ξ^T, $T \in \mathcal{T}$, as the systematic (= fixed) effect. The natural requirements $J \in \mathcal{B}$ and $\emptyset \in \mathcal{T}$

are also imposed. In multi-way analysis of variance, all hypotheses of the form $H_{\mathcal{T}} \cap H_{\mathcal{B}}$ are referred to as variance component models. The hypothesis $H_{\mathcal{T}}$ is equivalent to the hypothesis $E(X) \in L_{\mathcal{T}} := \Sigma(L_T | T \in \mathcal{T}) \in \mathcal{L}_{\mathcal{D}}$; however, two distinct $\mathcal{T}$'s may give rise to the same $L_{\mathcal{T}}$ (cf. Tjur (1984), § 6.2).

Consider the c.v.c. model determined by $L \in \mathcal{L}_{\mathcal{D}}$ and $\mathcal{M} = \mathcal{L}_{\mathcal{D}}$. Let $(\sigma_B^2 | B \in \mathcal{D}(J)) \in \mathbb{R}_+^{\mathcal{D}(J)}$ be the variance parameters for the model. Jensen noted that any submodel defined by a set of equalities among the σ_B^2 is also a product of linear models, hence may be solved explicitly. He posed the question: which variance component models of the form $H_{\mathcal{T}} \cap H_{\mathcal{B}}$ are submodels of this type?

<u>Lemma 4.2 (Jensen (1979))</u>. The hypothesis $H_{\mathcal{T}} \cap H_{\mathcal{B}}$ is defined by a set of equalities among the σ_B^2, $B \in \mathcal{D}(J)$, if and only if $\mathcal{B}$ is closed under intersection.

Jensen noted that this result is not quite true as stated - one must ignore the restrictions imposed on the σ_B^2, $B \in \mathcal{D}(J)$, by the requirements $\omega_B > 0$, $B \in \mathcal{B}$ - cf. Tjur (1984), § 7.6 and our Remark 4.3.

In Theorem 4.1, we shall extend Jensen's lemma to the general case considered in this paper.

<u>4.5. Variance component models in factor-generated designs (Tjur (1984)).</u>

Tjur (1984, Section 7) defines a family of variance component models arising from a design generated by a set $\mathcal{D}$ of orthogonal factors with $I \in \mathcal{D}$ (see also our § 3.6). Each of his variance component models is determined by two subsets $\mathcal{T},\mathcal{B} \subseteq \mathcal{D}$ and by a formula similar to (4.14). He assume that $\mathcal{B}$ satisfies four conditions:

(C1) $\quad I \in \mathcal{B}.$

(C2) $\quad$ All factors in $\mathcal{B}$ are balanced.

(C3) $\quad \mathcal{B}$ is closed under formation of minima (in $\mathcal{D}$).

(C4) $\quad$ The matrices $X_B X_B^t$ are linearly independent, $B \in \mathcal{B}$.

The variance component model determined by $\mathcal{T}$ and $\mathcal{B}$ is equivalent (except for restriction on the parameter domain - cf. Tjur (1984, § 7.6) and our Remark 4.3) to the c.v.c. model given by $(L, \mathcal{L}_{\mathcal{B}})$, where $L = \Sigma(L_T | T \in \mathcal{T})$ and $\mathcal{L}_{\mathcal{B}}$ is the smallest lattice containing $\{L_B | B \in \mathcal{B}\}$, and therefore has an explicit solution. (Strictly speaking, it should also be assumed that $\mathcal{T} \neq \emptyset$.) In the following subsection we shall return to this case and discuss the relation between Tjur's conditions (C1)-(C4) and our specification of a c.v.c. model via the sublattice $\mathcal{L}_{\mathcal{B}}$.

4.6. Canonical variance component models and random effect models.

After the digression of the preceding two subsections, we now return to the general case condered in §4.2 and 4.3. We shall propose a general definition of a random effect model and obtain a necessary and sufficient condition for such a model to be a c.v.c. model (cf. Theorem 4.1).

Let $\mathcal{T}$ and $\mathcal{B}$ be two nonempty subsets of $\mathcal{L}$. Set

$$L_{\mathcal{T}} = \Sigma(T | T \in \mathcal{T}) = \Sigma(u_T(P_T) | T \in \mathcal{T}) \in \mathcal{L}. \tag{4.20}$$

Let $(Y^B|B\in\mathcal{B})$ be a family of independent stochastic variables, where Y^B is normally distributed on P_B with mean $\underset{\sim}{0}$ and precision $\omega_B^{-1}\delta\circ(u_B\times u_B)$. (Our ω_B correspond to Tjur's σ_B^2.) Set

$$(4.21) \qquad X = \xi + \Sigma(u_B(Y^B)|B\in\mathcal{B}),$$

where $\xi \in L_{\mathcal{T}}$ and $(\omega_B|B\in\mathcal{B}) \in \mathbb{R}_+^{\mathcal{B}}$. (Alternatively, in place of ξ in (4.21) we can write $\Sigma(u_T(\xi_T)|T\in\mathcal{T})$, where the parameter $(\xi_T|T\in\mathcal{T}) \in X(P_T|T\in\mathcal{T})$, but this parametrization is not in general one-to-one.) If the two conditions

$$(R1) \qquad \Sigma\mathcal{B} = V,$$

$$(R2) \qquad \mathcal{B} \subseteq J(\mathcal{L}_{\mathcal{B}})$$

are satisfied, where $\mathcal{L}_{\mathcal{B}}$ is the smallest lattice containing $\mathcal{B}$, we refer to the family of distributions of X in (4.21) as the <u>random effect model determined by</u> $(\mathcal{T},\mathcal{B})$.

From (4.21) it follows that the covariance V(X) is given by

$$\begin{aligned}
V(X) &= \Sigma(V(u_B(Y^B))|B\in\mathcal{B}) \\
&= \Sigma(V(Y^B)\circ(u_B^t\times u_B^t)|B\in\mathcal{B}) \\
&= \Sigma((\omega_B^{-1}\delta\circ(u_B\times u_B))^{-1}\circ(u_B^t\times u_B^t)|B\in\mathcal{B}) \\
&= \Sigma(\omega_B\delta^{-1}\circ((u_Bp_B)^t\times(u_Bp_B)^t)|B\in\mathcal{B}) \\
&= \Sigma(\omega_B\Sigma(\delta^{-1}\circ(r_L^t\times r_L^t)|L\in J(\mathcal{L}_{\mathcal{B}}),L\sqsubseteq B)|B\in\mathcal{B}) \\
&= \Sigma(\Sigma(\omega_B\delta^{-1}\circ(r_L^t\times r_L^t)|B\in\mathcal{B},B\sqsupseteq L)|L\in J(\mathcal{L}_{\mathcal{B}})) \\
&= \Sigma(\Sigma(\omega_B|B\in\mathcal{B},B\sqsupseteq L))\delta^{-1}\circ(r_L^t\times r_L^t)|L\in J(\mathcal{L}_{\mathcal{B}})),
\end{aligned}$$

where $u_B^t : P_B^* \to V^*$ and $r_L^t : V^* \to V^*$ are the dual mappings. The fifth equality follows from the relation (cf. (3.12)) $q_B = u_B p_B = \Sigma(r_L | L \in J(\mathcal{L}_{\mathcal{B}}), L \subseteq B)$. This shows that the random effect model (4.21) determined by $(\mathcal{T}, \mathcal{B})$ is a submodel (= hypothesis) $H'_{\mathcal{B}}$ ($\equiv H'_{\mathcal{T},\mathcal{B}}$) within the c.v.c. model determined by $(L_{\mathcal{T}}, \mathcal{L}_{\mathcal{B}})$. Specifically, $H'_{\mathcal{B}}$ is the submodel specified by the restrictions

$$(4.22) \quad (\sigma_L^2 | L \in J(\mathcal{L}_{\mathcal{B}})) \in \{(\sigma_L^2(\underline{\omega}) | L \in J(\mathcal{L}_{\mathcal{B}})) \in \mathbb{R}_+^{J(\mathcal{L}_{\mathcal{B}})} | \underline{\omega} = (\omega_B | B \in \mathcal{B}) \in \mathbb{R}_+^{\mathcal{B}}\},$$

where

$$(4.23) \qquad \sigma_L^2(\underline{\omega}) := \Sigma(\omega_B | B \in \mathcal{B}, B \supseteq L), \quad L \in J(\mathcal{L}_{\mathcal{B}}).$$

(We suppress the $\mathcal{T}$ in $H'_{\mathcal{B}}$ since the discussion and results in the remainder of § 4.6 depend only upon the covariance structure, not the mean value structure.)

<u>Remark 4.3</u>. Since $\mathcal{B} \subseteq J(\mathcal{L}_{\mathcal{B}})$ the formula (4.23) determines a linear isomorphism of $\mathbb{R}^{\mathcal{B}}$ given by

$$(4.24) \qquad \lambda_B = \Sigma(\omega_{B'} | B' \in \mathcal{B}, B' \supseteq B), \qquad B \in \mathcal{B}, \ (\omega_{B'} | B' \in \mathcal{B}) \in \mathbb{R}^{\mathcal{B}}$$

$$(4.25) \qquad \omega_B = \Sigma(\tilde{\mu}(B,B')\lambda_{B'} | B' \in \mathcal{B}), \quad B \in \mathcal{B}, \ (\lambda_{B'} | B' \in \mathcal{B}) \in \mathbb{R}^{\mathcal{B}}$$

where $\tilde{\mu}(B,B') = \mu(B',B)$, B, $B' \in \mathcal{B}$ and μ is the Möbius function for the poset $\mathcal{B}$. For $(\lambda_{B'} | B' \in \mathcal{B}) \in \mathbb{R}_+^{\mathcal{B}}$ in (4.24), $(\omega_B | B \in \mathcal{B})$ lies in a corresponding cone $\Omega \supseteq \mathbb{R}_+^{\mathcal{B}}$. It will be necessary to replace $\mathbb{R}_+^{\mathcal{B}}$ by Ω in the definition of $H'_{\mathcal{B}}$ (thereby obtaining an extended hypothesis $H_{\mathcal{B}}$) in order to establish the precise connection between random effect models and c.v.c. models (cf. Theorem 4.1). By allowing this extension, we are ignoring the restrictions on $(\sigma_L^2 | L \in J(\mathcal{L}))$ imposed by the original assumption that $\omega_B > 0$, $B \in \mathcal{B}$. This extension to Ω appears in the literature as the well known questions of allowing negative variance components in the model and of interpreting negative estimates when they occur (cf. Tjur (1984), §7.6).□

In general, even the extended model $H_{\mathcal{B}}$ cannot be solved explicitly. Further conditions must be imposed on $\mathcal{B}$ in order that $H_{\mathcal{B}}$ becomes a c.v.c.model.

Theorem 4.1. Under the conditions (R1) and (R2), the hypothesis $H_{\mathcal{B}}$ specified by replacing $\mathbb{R}_+^{\mathcal{B}}$ by Ω in (4.22) is a canonical variance component model if and only if the following condition holds:

$$(R2)^* \qquad \mathcal{B} = J(\mathcal{L}_{\mathcal{B}}).$$

In this case $H_{\mathcal{B}}$ is the canonical variance component model determined by $(L_{\mathcal{T}}, \mathcal{L}_{\mathcal{B}})$.

Proof: If $\mathcal{B} = J(\mathcal{L}_{\mathcal{B}})$ then obviously $|\mathcal{B}| = |J(\mathcal{L}_{\mathcal{B}})|$, hence it follows immediately from (4.22) that $H_{\mathcal{B}}$ is the c.v.c. model determined by $(L_{\mathcal{T}}, \mathcal{L}_{\mathcal{B}})$. Conversely, suppose that $H_{\mathcal{B}}$ is a c.v.c. model determined by $(L, \mathcal{M})$ for some $L \in \mathcal{L}$ and $\mathcal{M} \in \mathrm{Sub}^*(\mathcal{L})$. Then for every $K, M \in J(\mathcal{L}_{\mathcal{B}})$, $[\forall \underline{\omega} \in \Omega,\ \sigma_K^2(\underline{\omega}) = \sigma_M^2(\underline{\omega})]$ $\Leftrightarrow$ $[\{B \in \mathcal{B} | B \supseteq K\} = \{B \in \mathcal{B} | B \supseteq M\}]$ $\Leftrightarrow$ $[\{B \in J(\mathcal{L}_{\mathcal{B}}) | B \supseteq K\} = \{B \in J(\mathcal{L}_{\mathcal{B}}) | B \supseteq M\}]$ $\Leftrightarrow$ $K = M$. The second $\Leftrightarrow$ follows from the fact that every element in $J(\mathcal{L}_{\mathcal{B}})$ is an intersection of elements from $\mathcal{B}$. It follows that $|J(\mathcal{M})| \geq |J(\mathcal{L}_{\mathcal{B}})|$. Since also $|J(\mathcal{M})| = |\mathcal{B}|$ we conclude that $\mathcal{B} = J(\mathcal{L}_{\mathcal{B}})$ and and that $H_{\mathcal{B}}$ can be determined by $(L_{\mathcal{T}}, \mathcal{L}_{\mathcal{B}})$.□

Corollary 4.1. Assume that (R1) and (R2) hold.

(i): If $\mathcal{B}$ is closed under $\cap$, then $H_{\mathcal{B}}$ is a canonical variance component model.

(ii): If $\mathcal{B} \subseteq J(\mathcal{L})$, $J(\mathcal{L})$ is closed under $\cap$, and $H_{\mathcal{B}}$ is a canonical variance component model, then $\mathcal{B}$ is closed under $\cap$.

Proof: (i) Since every subspace in $\mathcal{L}_{\mathcal{B}}$ can be expressed as a sum of subspaces in $\mathcal{B}$, $J(\mathcal{L}_{\mathcal{B}}) \subseteq \mathcal{B}$, hence $J(\mathcal{L}_{\mathcal{B}}) = \mathcal{B}$ and the theorem applies. (ii) Since $\mathcal{B} \subseteq J(\mathcal{L})\cap\mathcal{L}_{\mathcal{B}} \subseteq J(\mathcal{L}_{\mathcal{B}})$, the theorem implies that $\mathcal{B} = J(\mathcal{L})\cap\mathcal{L}_{\mathcal{B}}$, thus is closed under $\cap$.□

Remark 4.4. If the condition $\mathcal{B} = J(\mathcal{L}_{\mathcal{B}})$ is not satisfied in Theorem 4.1 one can still estimate the unknown parameters $(\omega_B | B\in\mathcal{B})$ in the model as follows. Since $\mathcal{B} \subseteq J(\mathcal{L}_{\mathcal{B}})$ we have that

$$\sigma_L^2 = \Sigma(\omega_B | B\epsilon B, B\supseteq L), \quad L\epsilon J(\mathcal{L}_{\mathcal{B}}) \tag{4.26}$$

and

$$\omega_B^2 = \Sigma(\tilde{\mu}(B,B')\sigma_{B'}^2 | B'\epsilon\mathcal{B}), \quad B\epsilon\mathcal{B}, \tag{4.27}$$

First estimate σ_B^2, $B \in \mathcal{B}$, in the c.v.c. model determined by $(L_{\mathcal{T}}, \mathcal{L}_{\mathcal{B}})$ by the ML estimator $\hat{\sigma}_B^2$, $B \in \mathcal{B}$. Then the estimates $\hat{\omega}_B$ (possible negative) for ω_B, $B \in \mathcal{B}$, are obtained from (4.27).□.

Remark 4.5. By Theorem 4.1, any c.v.c. model given by $(L,\mathcal{M})$ can be interpreted in terms of a random effect model in the following way. Let $\mathcal{T}$ be any non-empty subset of $\mathcal{L}$ such that $L_{\mathcal{T}} = L$, take $\mathcal{B} = J(\mathcal{M})$, and consider the random effect model determined by $(\mathcal{T},\mathcal{B})$. Ignoring the restrictions on

$(\sigma^2_M | M \in J(\mathcal{M}))$ imposed by the restrictions $\omega_M > 0$, $M \in J(\mathcal{M})$, this random effect model is just a reparametrization of the original c.v.c. model. The correspondence between the parameters is given by

(4.28) $$\sigma^2_M = \Sigma(\omega_{M'} | M' \in J(\mathcal{M}), M' \supseteq M), \quad M \in J(\mathcal{M}),$$

and

(4.29) $$\omega_M = \Sigma(\tilde{\mu}(M,M')\sigma^2_{M'} | M' \in J(\mathcal{M})), \quad M \in J(\mathcal{M}),$$

where $\tilde{\mu}(M,M') = \mu(M',M)$ and μ is the Möbius function for $J(\mathcal{M})$. Thus every c.v.c. model can be represented as a random effect model.□

Remark 4.6. Consider the special case of the multi-way layout treated by Jensen (1979) (cf. §4.4). Recall that, as in Example 3.5, we may identify the isomorphic posets $J(\mathcal{L}_{\mathcal{D}(J)})$ and $\mathcal{D}(J)$. Since $\mathcal{D}(J)$ is closed under $\cap$ and $\mathcal{B} \subseteq \mathcal{D}(J)$, Jensen's Lemma 4.2 follows from Corollary 4.1.□

Remark 4.7. In the more general case of an orthogonal factor-generated design treated by Tjur (1984) (cf. our §3.6), let $\mathcal{T}$, $\mathcal{B}$ be nonempty subsets of $\mathcal{D}$. Let $(Y^B | B \in \mathcal{B})$ be a family of independent stochastic variables with Y^B normally distributed on $\mathbb{R}^B$ with mean $\underset{\sim}{0}$ and precision $\omega_B^{-1}\delta \circ (u_B \times u_B)$, where $(\omega_B | B \in \mathcal{B}) \in \mathbb{R}^{\mathcal{B}}_+$, δ is the usual inner product on $\mathbb{R}^I$, and $u_B : \mathbb{R}^B \to \mathbb{R}^I$ is the embedding determined by the factor B, i.e., $u_B((x_b | b \in B)) = (x_{\overline{B}(i)} | i \in I)$ (cf. (3.8)). The covariance matrix for Y^B then becomes $\omega_B(X_B'X_B)^{-1}$, where X_B is the matrix for u_B. As in (4.21), define

(4.30) $$X = \xi + \Sigma(X_B Y^B | B \in \mathcal{B}),$$

where $\xi \in L := \Sigma\mathcal{T}$. Note that in this case of a factor-generated design, the statistical model given by this family of distributions for X is of course the same as that given by the family of distributions for X in (4.21). This model does not coincide with the model considered by Tjur (see the first display in Section 7 of Tjur (1984), p.51) unless Tjur's condition (C2) that all factors $B \in \mathcal{B}$ are balanced is imposed, in which case the two models are identical, with our ω_B and Tjur's σ_B^2 related by $\omega_B = \sigma_B^2 |I| / |B|$. When the condition (C2) holds, we can make the following comparisons between Tjur's conditions and ours: (i) (C1) => (R1); (ii) under (R1), (R2) <=> (C4); (iii) under (R1) and (R2), (C3) => $(R2)^*$. (cf. Theorem 4.1 and Corollary 4.1.) Thus, when (C2) holds, Tjur's conditions (C1), (C3), and (C4) are (strictly) more restrictive than our conditions (R1) and $(R2)^*$ for a random effect model given by $H_{\mathcal{B}}$ to be a c.v.c. model.□

Example 4.2. (A Latin square of Latin squares.) Let $\{J, R_0, C_0, G_0, 0\}$ be a Latin square design and for each $j \in J$ let $\{I_j, R_j, C_j, G_j, 0_j\}$ be a Latin square design (cf. Example 3.11). Define $I := \dot{U}(I_j | j \in J)$, $R := \dot{U}(R_j | j \in J)$, $C := \dot{U}(C_j | j \in J)$, $G := \dot{U}(G_j | j \in J)$ and note that $J := \dot{U}(0_j | j \in J)$. Consider the set of factors $\mathcal{D} = \{I, R, C, G, J, \underline{R}, \underline{C}, \underline{G}, 0\}$ in I defined in the obvious way, e.g., $\overline{R}(ji) = j\overline{R}_j(i)$ and $\underline{\overline{R}}(ji) = \overline{R}_0(j)$, $ji \in I$. Suppose that $|R_j| = |C_j| = |G_j| = f$ is independent of $j \in J$. Then (C2) holds, i.e., all factors are balanced. (i) Set $\mathcal{B} = \{R, C, G, J\}$, let $\mathcal{T} \subseteq \mathcal{D}$ be arbitrary, and suppose that $f = 2$, $|R_0| = |C_0| = |G_0| = g > 2$. Then (C1) fails but (R1) and $(R2)^*$ are satisfied, hence the random effect model $H_{\mathcal{B}}$ is a c.v.c. model that is not contained in the class of variance component models described by Tjur (1984), Section 7. (ii) Set $\mathcal{B} = \mathcal{D}$, let $\mathcal{T} \subseteq \mathcal{D}$ be arbitrary, and suppose that $f > 2$ and $g = 2$. Then (C1) holds but (C3) fails. Nevertheless, (R1) and (R2)* hold so again $H_{\mathcal{B}}$ is a c.v.c. model not contained in Tjur's class.□

4.7. The covariance structure in a canonical variance component model.

In Remark 4.5 a second representation of a c.v.c. model were given. A third representation of a general c.v.c. model is obtained through its covariance structure Γ. The covariance Γ can be expressed in terms of the parameter $(\sigma_M^2 | M \in J(\mathcal{M}))$ or the parameter $(\omega_M | M \in J(\mathcal{M}))$. From (4.1) and from (4.21) with $\mathcal{B} = J(\mathcal{M})$ it is readily obtained that

$$\Gamma = \Sigma(\sigma_M^2 \delta^{-1} \circ (r_M^t \times r_M^t) | M \in J(\mathcal{M})) \tag{4.31}$$

and

$$\Gamma = \Sigma(\omega_M \delta^{-1} \circ (q_M^t \times q_M^t) | M \in J(\mathcal{M})) \tag{4.32}$$

If $V = \mathbb{R}^I$ and δ is the usual inner product, the matrix formulations of (4.31) and (4.32) become

$$\Gamma = \Sigma(\sigma_M^2 R_M | M \in J(\mathcal{M})) \tag{4.33}$$

and

$$\Gamma = \Sigma(\omega_M Q_M | M \in J(\mathcal{M})), \tag{4.34}$$

where R_M and Q_M are the $I \times I$ matrices for r_M and q_M respectively.

In (4.32) and (4.34) the parameter space for $(\omega_M | M \in J(\mathcal{M}))$ is $\Omega \supseteq \mathbb{R}_+^{J(\mathcal{M})}$ as defined in Remark 4.3 with $\mathcal{B} = J(\mathcal{M})$. Equivalently, Ω can be defined by requiring that Γ in (4.32) or (4.34) is positive definite. (In (4.31) and (4.33) Γ is positive definite if and only if $\sigma_M^2 > 0$, $M \in J(\mathcal{M})$).

In our opinion, this representations of a c.v.c. model in terms of its covariance structure are more appropriate for determining its statistical interpretation than its representation in terms of a random effect model.

Example 4.3. Consider the one-way analysis of variance layout in Example 3.2 with $\mathcal{D} = \{I,G,0\}$. If $L = L_0$ and $\mathcal{M} = \{I,G\}$ where G is not balanced, it follows from (4.33) that the entries of the covariance matrix Γ are given by

$$(4.35) \qquad \Gamma_{gi,\gamma j} = \begin{cases} \sigma_I^2(1-|I_g|^{-1}) + \sigma_G^2|I_g|^{-1}, & \text{for } g=\gamma,\ i=j \\ (\sigma_G^2-\sigma_I^2)|I_g|^{-1} & \text{, for } g=\gamma,\ i\neq j \end{cases}$$

and 0 for $g \neq \gamma$. This determines a c.v.c. model, which may be represented as a random effect model (cf. Remark 4.5) given by the stochastic variable X with coordinates

$$(4.36) \qquad X_{gi} = \alpha + Y_g + Y_{gi}, \quad g \in G,\ gi \in I,$$

where $\alpha \in \mathbb{R}$ and where the Y_g and Y_{gi} are all mutually independent and normally distributed on $\mathbb{R}$ with mean 0 but with unequal variances given by $V(Y_g) = \omega_G|I_g|^{-1}$, $V(Y_{gi}) = \omega_I$. Equivalently, in terms of $\underline{\omega} = (\omega_I,\omega_G)$ we have

$$(4.37) \qquad \Gamma_{gi,\gamma j} = \begin{cases} \omega_I + \omega_G|I_g|^{-1} & , \text{ for } g=\gamma,\ i=j \\ \omega_G|I_g|^{-1} & , \text{ for } g=\gamma,\ i\neq j \end{cases}$$

and 0 for $g \neq \gamma$. Although the c.v.c. model represented by (4.35) or (4.36) has a simple mathematical solution, it is likely that neither representation (4.35) nor (4.36) will make it palatable to an applied statistician, who usually would prefer the model (4.36) defined with equal variances $V(Y_g)$. Then, however, it is not a c.v.c. model and has no simple solution. Thus there is a trade-off between statistical appropriateness and mathematical tractability in random effect models. Of course, when G is balanced, i.e., $|I_g|$ is independent of $g \in G$, then the model reduces to the usual one-way layout with random group (=treatment) effect and this conflict disappears.□

Example 4.1 (continued). Consider the two-way analysis of variance layout with $|I_{(r,c)}| = n > 1$, $(r,c) \in R{\times}C$. If $L = L_C$ and $\mathcal{M} = \{I, R{\times}C, R\}$ it follows from (4.33) that the entries of the covariance matrix Γ are given by

$$(4.38) \quad n\Gamma_{(r,c)i,(\rho,\gamma)j} = \begin{cases} \sigma_I^2(n-1)+(\sigma_{R\times C}^2(|C|-1)+\sigma_R^2)/|C|, & \text{for } (r,c)i=(\rho,\gamma)j \\ \sigma_I^2+\sigma_{R\times C}^2(|C|-1)/|C|+\sigma_R^2/|C|, & \text{for } (r,c)=(\rho,\gamma),\ i\neq j \\ (\sigma_R^2-\sigma_{R\times C}^2)/|C|, & \text{for } r=\rho,\ c\neq\gamma \end{cases}$$

and 0 for $r\neq\rho$. This determines a c.v.c. model, which may be represented as a random effect model (cf. Remark 4.5) given by the stochastic variable X with coordinates

$$(4.39)\quad X_{(r,c)i} = \xi_c + Y_r + Y_{(r,c)} + Y_{(r,c)i}, \quad (r,c) \in R\times C,\ (r,c)i \in I,$$

where $\xi_c \in \mathbb{R}$, $c \in C$ and where the Y_r, $Y_{(r,c)}$ and $Y_{(r,c)i}$ are all mutually independent and normally distributed on $\mathbb{R}$ with mean 0 and variances given by $nV(Y_r) = \omega_R/|C|$, $nV(Y_{(r,c)}) = \omega_{R\times C}$ and $V(Y_{(r,c)i}) = \omega_I$. Equivalently, in terms of $\underline{\omega} = (\omega_I, \omega_{R\times C}, \omega_R)$ we have

$$(4.40)\qquad n\Gamma_{(r,c)i,(\rho,\gamma)j} = \begin{cases} \omega_I n + \omega_{R\times C} + \omega_R/|C|, & \text{for } (r,c)i=(\rho,\gamma)j \\ \omega_{R\times C} + \omega_R/|C|, & \text{for } (r,c)=(\rho,\gamma), i\neq j \\ \omega_R/|C|, & \text{for } r=\rho, c\neq\gamma \end{cases}$$

and 0 for $r \neq \rho$. Both representations (4.38) and (4.39) for this model can be easily interpreted statistically. This model may be called the two-way layout with random interaction effect, random row effect, and systematic column effect.□

<u>Example 4.1 (continued)</u>. In the preceding example let instead $\mathcal{M} = \{\mathbb{R}^I, L_{R\times C}, L_R + L_C\}$. Then

$$n\Gamma_{(r,c)i,(\rho,\gamma)j} = \begin{cases} \sigma_I^2(n-1) + (\sigma^2_{R\times C}(|R|-1)(|C|-1) + \sigma^2_{R+C}(|R|+|C|-1))/|R||C|, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{for } (r,c)i=(\rho,\gamma)j \\ \sigma^2_{R\times C}(|R|-1)(|C|-1) + \sigma^2_{R+C}(|R|+|C|-1))/|R||C|, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{for } (r,c)=(\rho,\gamma), i\neq j \\ \sigma^2_{R+C}(|R|-1)/|R||C|, & \text{for } r=\rho, c\neq\gamma \\ \sigma^2_{R+C}(|C|-1)/|R||C|, & \text{for } r\neq\rho, c=\gamma \\ -\sigma^2_{R+C}/|R||C|, & \text{for } r\neq\rho, \gamma\neq c, \end{cases}$$

which is difficult to interpret statistically even when $|R| = |C|$. Thus, even in the balanced case (i.e., when (C2) holds), our class of c.v.c. models may include models which are not readily interpretable.□

4.8. Testing canonical variance component models.

Lastly, we consider the problem of testing c.v.c. models within a g.o.v.c. design determined by a lattice $\mathcal{L}$. We shall see that the ANOVA table $((SSD_L(x), f_L) | L \in J(\mathcal{L}))$ (cf. (3.9)) contains all information needed for calculation of the LR test statistics and their distributions.

Consider the c.v.c. model determined by $(L, \mathcal{M})$. The parameter space for this model is $L \times \mathbb{R}_+^{J(\mathcal{M})}$. For $L_0 \in \mathcal{L}$ and $\mathcal{M}_0 \in \mathrm{Sub}^*(\mathcal{L})$ such that $L_0 \subseteq L$ and $\mathcal{M}_0 \subseteq \mathcal{M}$, the parameter space for the c.v.c. model determined by $(L_0, \mathcal{M}_0)$ is $L_0 \times \mathbb{R}_+^{J(\mathcal{M}_0)}$. This model is a submodel of the c.v.c. model $(L, \mathcal{M})$: simply note that the surjective poset homomorphism $\psi = J(u): J(\mathcal{M}) \to J(\mathcal{M}_0)$, where $u: \mathcal{M}_0 \to \mathcal{M}$ is the embedding, defines the injective mapping

(4.41)
$$\begin{aligned} \eta: L_0 \times \mathbb{R}_+^{J(\mathcal{M}_0)} &\to L \times \mathbb{R}_+^{J(\mathcal{M})} \\ (\xi_0, (\sigma_K^2 | K \in J(\mathcal{M}_0))) &\to (\xi_0, (\sigma_{\psi(M)}^2 | M \in J(\mathcal{M}))) \end{aligned}$$

between their two parameter spaces.

Remark 4.8. It may be seen from (4.41) that the c.v.c. model determined by $(V, \mathcal{L})$ contains all c.v.c. models in the design as submodels, while the c.v.c. model determined by $(0_{\mathcal{L}}, \{V\})$ is a submodel of every c.v.c. model in the design.□

Let H (respectively H_0) denote the hypothesis given by the c.v.c. model $(L,\mathcal{M})$ (respectively $(L_0,\mathcal{M}_0)$) and consider the problem of testing H_0 vs. H. The subhypothesis H_0 may be expressed in terms of the mean and variances as

$$H_0 = H_{0m} \cap H_{0v},$$

where

$$H_{0m}: \xi \in L_0$$
$$H_{0v}: \forall M,M' \in J(\mathcal{M}),\ \psi(M) = \psi(M') \Rightarrow \sigma_M^2 = \sigma_{M'}^2.$$

Equivalently, H_{0v} may be written as

$$H_{0v}: \ \forall M \in J(\mathcal{M}),\ \forall K \in J(\mathcal{M}_0),\ \psi(M) = K \Rightarrow \sigma_M^2 = \sigma_K^2.$$

When $L_0 \subset L$ and $\mathcal{M}_0 = \mathcal{M}$, we may refer to the testing problem H_0 vs. H as <u>testing systematic effects</u>, while when $L_0 = L$ and $\mathcal{M}_0 \subset \mathcal{M}$ we are <u>testing random effects</u>. Tjur (1984), §7.9 and 7.10, treats special cases of these two testing problems in the context of factor-generated designs (cf. our §4.5).

The LR statistic for testing H_0 vs. H is now obtained. Let $V = \perp(U_K | K \in J(\mathcal{M}_0))$ be the orthogonal decomposition wrt. $\mathcal{M}_0$ (cf. (4.1)). Then

$$U_K = \perp(W_M | M \in J(\mathcal{M}), \psi(M)=K)$$

(cf. (4.4)) and

$$\begin{aligned} L_0^K &:= \perp(W_M | M \in J(\mathcal{M}), M \sqsubseteq L_0, \psi(M)=K) \\ &\subseteq \perp(L^M | M \in J(\mathcal{M}), \psi(M)=K), \end{aligned}$$

where L^M is defined in (4.6). It thus follows that the existence of the ML estimator of $(\xi,(\sigma_M^2|M\in J(\mathcal{M}))) \in L \times \mathbb{R}_+^{J(\mathcal{M})}$ under H implies the existence of the ML estimator of $(\xi_0,(\sigma_K^2|K\in J(\mathcal{M}_0))) \in L_0 \times \mathbb{R}_+^{J(\mathcal{M}_0)}$ under H_0. From the general expression (4.9) for the ML estimator in a c.v.c. model, we readily obtain the LR statistic

$$\text{(4.42)} \qquad Q = \frac{\Pi([\hat{\sigma}_M^2]^{m_M/2}|M\in J(\mathcal{M}))}{\Pi([\tilde{\sigma}_K^2]^{k_K/2}|K\in J(\mathcal{M}_0))}$$

where for $M \in J(\mathcal{M})$ and $K \in J(\mathcal{M}_0)$, $m_M = \dim(W_M)$, $k_K = \dim(U_K)$, $\hat{\sigma}_M^2$ is the ML estimator for $\sigma_M^2 \in \mathbb{R}_+$ under H, and $\tilde{\sigma}_K^2$ is the ML estimator for $\sigma_K^2 \in \mathbb{R}_+$ under H_0.

To find the null distribution of Q, divide the numerator and the denominator in (4.42) by $\Pi([\sigma_K^2]^{k_K/2}|K\in J(\mathcal{M}_0))$. Since $k_K = \Sigma(m_M|M\in J(\mathcal{M}),\psi(M)=K)$, $K \in J(\mathcal{M}_0)$, this shows that the distribution of Q is independent of the unknown parameters under H_0. Using Basu's Lemma, it is seen the ML estimator under H_0 is independent of Q when H_0 holds. This implies that for $\alpha \in [0,\infty[$,

$$EQ^\alpha = \frac{E(\Pi([\hat{\sigma}_M^2]^{\alpha m_M/2}|M\in J(\mathcal{M})))}{E(\Pi([\tilde{\sigma}_K^2]^{\alpha k_K/2}|K\in J(\mathcal{M}_0)))}$$

Furthermore, from the distribution of the ML estimators as described following (4.12), it follows that

$$(4.43)\qquad EQ^{\alpha} = \frac{\Pi\left[m_M^{-\alpha m_M/2}\ \Gamma((l_M+\alpha m_M)/2)\ /\Gamma(l_M/2)\ |M\in J(\mathcal{M})\right]}{\Pi\left[k_K^{-\alpha k_K/2}\ \Gamma((j_K+\alpha k_K)/2)/\Gamma(j_K/2)\ |K\in J(\mathcal{M}_0)\right]},$$

where $l_M = \dim(L^M)$ and $j_K = \dim(L_0^K)$. It is obvious that the usual Box approximation for the distribution of Q is valid. Also, it is clear from (4.9), (4.10) and (4.12) that all quantities needed in (4.42) and (4.43) can be read out directly from the ANOVA table.

Usually it is more convenient to carry out the test for H_0 vs. H in several simpler steps. As always one should try to simplify the variance structure before the mean value structure. Thus H_{ov} vs. H should be tested first and then $H_{om}\cap H_{0v}$ vs. H_{ov}.

Testing H_{ov} vs. H is the problem of testing equality of the σ_M^2's within the families $(\sigma_M^2|M\in J(\mathcal{M}),\psi(M)=K)$, $K \in J(\mathcal{M}_0)$. The LR statistic Q_v for this problem is then a product, $Q_v = \Pi(Q_v^K|K\in J(\mathcal{M}_0))$, of the LR statistics Q_v^K, $K \in J(\mathcal{M}_0)$ for testing equality within each family. Under H_{0v}, the Q_v^K are independent. Usually, Q_v^K is replaced by the (unbiased) Bartlett test statistic. Thus the test for H_{ov} vs. H can be carried out in $|J(\mathcal{M}_0)|$ steps, each step a Bartlett test.

The testing problem $H_{0m}\cap H_{0v}$ vs. H_{0v} can be seen to be an independent product of testing problems indexed again by $K \in J(\mathcal{M}_0)$. The problem indexed by K is a testing problem in the linear model with observation space U_K and precision $\sigma_K^{-2}\delta_K$, where δ_K is the restriction of δ to U_K. The two subspaces of U_K that determine the hypotheses to be tested are L_0^K and $\perp(L^M|M\in J(\mathcal{M}),\psi(M)=K)$ $(\supseteq L_0^K)$. The LR statistic Q_m for testing $H_{0m}\cap H_{0v}$ vs. H_{0v} is thus a product $Q_m = \Pi(Q_m^K|K\in J(\mathcal{M}_0))$, where the Q_m^K, $K \in J(\mathcal{M}_0)$, are independent, Q_m^K being the LR statistic for the testing problem indexed by

K. Thus the test for $H_{0m} \cap H_{0v}$ vs. H_{0v} also can be carried out in $|J(\mathcal{M}_0)|$ steps.

Finally we again emphasize that all quantities needed to determine the above test statistics and their distributions can be obtained directly from the ANOVA table associated with the underlying g.o. design of linear models determined by $\mathcal{L}$. By considering the larger class of c.v.c. models in the g.o.c.v. design determined by $\mathcal{L}$, we only provide ourselves with an argument for comparing other pairs of sums of SSD's than those pairs allowed to be compared in the underlying design.

Acknowledgements. I am grateful to my colleagues Hans Brøns, Tue Tjur, and Søren Tolver Jensen for valuable discussions and comments. Their contributions to this area of statistics have been of invaluable help and inspiration for the present paper. Special thanks are due to Michael D. Perlman whose advice, corrections, and comments helped to improve the presentation of the paper.

References.

[1] Andersson, S.A. and Perlman, M.D. (1988). Conditional indepedence models determined by finite distributive lattices. In preparation.

[2] Grätzer, G. (1978). General Lattice Theory. Basel, Stuttgart: Birkhauser.

[3] Jensen, S.T. (1979). Varianskomponentmodeller i fuldstændige balancerede forsøg. Inst. Math. Statist., University of Copenhagen, manuscript in Danish.

[4] Tjur, T. (1984). Analysis of variance models in orthogonal designs (with discussion by R.A. Bailey, T.P. Speed, and H.P. Wynn). Int. Statist. Rev. 52, 33-81.

[5] Speed, T.P. and Bailey, R.A. (1982). On a class of association schemes derived from lattices of equivalence relations. In Algebraic Structures and Applications, Ed. P. Schultz, C.E. Praeger and R.P. Sullivan. New York: Marcel Dekker.

PREPRINTS 1986

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK, TELEPHONE + 45 1 35 31 33.

No. 1 Jespersen, N.C.B.: On the Structure of Simple Transformation Models.

No. 2 Dalgaard, Peter and Johansen, Søren: The Asymptotic Properties of the Cornish-Bowden-Eisenthal Median Estimator.

No. 3 Jespersen, N.C.B.: Dichotomizing a Continuous Covariate in the Cox Regression Model.

No. 4 Asmussen, Søren: On Ruin Problems and Queues of Markov-Modulated M/G/1 Type.

No. 5 Asmussen, Søren: The Heavy Traffic Limit of a Class of Markovian Queueing Models.

No. 6 Jacobsen, Martin: Right Censoring and the Kaplan - Meier and Nelson - Aalen Estimators.

PREPRINTS 1987

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK, TELEPHONE + 45 1 35 31 33.

No. 1 Jensen, Søren Tolver and Johansen, Søren: Estimation of Proportional Covariances.

No. 2 Rootzén, Holger: Extremes, Loads, and Strengths.

No. 3 Bertelsen, Aksel: On the Problem of Testing Reality of a Complex Multivariate Normal Distribution.

No. 4 Gill, Richard D. and Johansen, Søren: Product-Integrals and Counting Processes.

No. 5 Leadbetter, M.R. and Rootzén, Holger: Extremal Theory for Stochastic Processes.

No. 6 Tjur, Tue: Block Designs and Electrical Networks.

No. 7 Johansen, Søren: Statistical Analysis of Cointegration Vectors.

No. 8 Bertelsen, Aksel: On the Problem of Testing Reality of a Complex Multivariate Normal Distribution, II.

No. 9 Andersson, S.A. and Perlman, M.D.: Group-Invariant Analogues of Hadamard's Inequality.

No. 10 Hald, Anders: Two Generalizations of the Problem of Points by Bernoulli, de Moivre and Montmort.

No. 11 Andersson, Steen Arne: The Lattice Structure of Orthogonal Linear Models and Orthogonal Variance Component Models.