DOCUMENT RESUME

ED 059 235 TM 001 044


| | |
|---|---|
| AUTHOR | Hall, Charles E. |
| TITLE | A Symbolic Logic for Representing Linear Models. |
| INSTITUTION | Educational Testing Service, Princeton, N.J. |
| REPORT NO | RB-71-60 |
| PUB DATE | Oct 71 |
| NOTE | 27p. |
| EDRS PRICE | MF-$0.65 HC-$3.29 |
| DESCRIPTORS | *Analysis of Variance; Computer Programs; Correlation; Factor Analysis; Factor Structure; Interaction; *Mathematical Logic; *Mathematical Models; Multiple Regression Analysis; *Statistical Analysis; *Symbols (Mathematics); Tests of Significance |



ABSTRACT

A set of symbols is presented along with logical operators which represent the possible manipulations of the linear model. The use of these symbols and operators is to simplify the representation of analysis of variance models, correlation models and factor analysis models. (Author)

RESEARCH BULLETIN

RB-71-60

# A SYMBOLIC LOGIC FOR REPRESENTING LINEAR MODELS

Charles E. Hall



This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the author. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
October 1971

A SYMBOLIC LOGIC FOR REPRESENTING LINEAR MODELS

Charles E. Hall

Abstract

A set of symbols is presented along with logical operators which represent the possible manipulations of the linear model. The use of these symbols and operators is to simplify the representation of analysis of variance models, correlation models and factor analysis models.

A SYMBOLIC LOGIC FOR REPRESENTING LINEAR MODELS

Charles E. Hall

The development of statistical inference from correlation theory to multivariate analysis of variance should be familiar to most readers, even though it may be beyond the interests of many. Most of this development has been concerned with the linear relationships among variables. Techniques of solving linear models of the relationships among variables for statistically relevant figures are not familiar to many because the general solutions for these models involve the techniques of matrix calculus. These techniques have become a little better known during recent years due to the increasing popularity of multivariate analysis and to the use of computers to generate solutions to large classes of statistical procedures.

The notational system presented here attempts to solve three problems: (1) a simplified representation of various linear model procedures, (2) an improved description of the matrix calculus procedures used to solve linear models, and (3) a simplified interface between the statistician and the computer program which does the calculations.

Following the presentation of the symbolism, there is a section showing how the symbolism can be used to expose the identity or similarity of various linear models.

## A Short History

Without belaboring the history of calculation procedures, it may be noted that the general solution for linear procedures seems to have been devised by R. A. Fisher, about 1930 or so, out of Gauss' method of reducing a general square matrix to a diagonal matrix. The first reasonable discussion of the calculation technique, called the "method of fitting constants," seems to have been undertaken by Wilks (1938) in a paper discussing the solution of non-orthogonal analysis of variance. An earlier paper by Wilks (1932) on the lambda criterion for multivariate distributions provides cues as to how this method can be applied to various kinds of problems in correlations and analysis of variance. A paper by Vartak (1955) showed how pseudovariables for factorial analysis of variance interactions could be generated from Kronecker products of main effect pseudovariables, thus simplifying the "method of fitting constants" as a calculation technique.

However, the history of calculation technique has ignored the "method of fitting constants" in the main. Instead, special calculation formulae were developed suitable to each of several analysis of variance designs. The bulk of practicing statisticians were unable to solve linear models unless their design conformed to those models for which special calculation formulae had been worked out.

With the advent of computers, special calculating formulae had to be abandoned and calculating technique reverted back to the "method of fitting constants" with some improvements. What had been too laborious a technique to do by hand became the most efficient method to program for a computer (see Bock, 1963). Currently, the major computer programs for computing analysis of variance and correlation analysis [for example, Cramer's MANOVA (Clyde, Cramer, & Sherin, 1966), Finn's Multivariance (1967) and Beaton's F4STAT (1971)] all use this form of calculation.

With the growth of generalized computer programs and the increasing scope of the problems they solve, it is becoming more and more obvious to all statisticians that all linear models are but variations on a central theme. There is no need for special calculation formulae for apparently diverse designs. With the unification of calculation technique and the attendant recognition of the unity of design types, there has also become a need for a notation system which will serve for all linear models.

To this point the following rules are proposed.

General Notation and Rules

It is general practice in texts on analysis of variance to refer to row and column effects of a factorial design as A or B and interaction effects as AxB. In the arithmetic of solving linear models, any effect, say A, has as many design parameters (pseudovariables, dummy parameters) as it has degrees of freedom. For convenience we will call these the A variables; likewise the design parameters for the B effect will be called the B variables. The interaction effect AxB is calculated using the Kronecker product of the A variables and the B variables; these product variables will be designated AB with no intervening symbol to designate interaction terms and Kronecker product of main effect design parameters.

Rule 1. Single upper-case Roman letters will designate main effect parameters of analysis of variance designs and solutions (excepting V and W).

Rule 2. Two or more upper-case Roman letters juxtaposed without intervening symbols will indicate the Kronecker product of main effect parameters and designate the interactions effect parameters in analysis of variance designs and solutions (excepting V and W).

In texts on correlation analysis, the measurements of variables are often designated x, y and z. It is also customary in analysis of variance to designate the variable which an experiment manipulates as x. In this system V will denote continuous variables.

Rule 3. All the continuous variables will be designated by the symbol V.

On occasion it is necessary to subdivide the design parameters for a main effect into subsets. Such occasions arise when the experimental treatments of an analysis of variance involve a control group and several treatments and one wishes to test control against the average of treatments and treatments among themselves. The subdivisions, partitions, of the main effect A will be labeled A1, A2, etc. The same holds true for subsets of the continuous variables.

Rule 4. Division of sets of parameters into subsets will be designated by an upper case Roman letter (not W) followed by a numeral 1, 2, 3, etc.

It is assumed that subsets will be enumerated ordinally from 1 to N and that the N subdivisions exhaust the parameters designated by the Roman letter. Examples: $\{A1,A2\} = A$, $\{V1,V2,V3\} = V$.

There are three basic types of linear model problems: correlation, analysis of variance and factor analysis. These three types of models can all be written as solutions to determinantal equations, regardless of the number of design parameters or continuous variables involved. For correlations we can have an equation of the type

$$| SS_H - \mu SS_T | = 0 \tag{1}$$

where $SS_H$ denotes sum of squares for hypothesis and $SS_T$ denotes sum of squares for total. For analysis of variance we can have an equation of the type

$$| SS_H - \lambda SS_E | = 0 \tag{2}$$

where $SS_E = SS_T - SS_H$ designates the sum of squares for error. For principal components analysis we have the equation

$$| \ SS_T - \nu \ I \ | = 0 \tag{3a}$$

where $SS_T$ is sometimes reduced to correlations.

Factor analysis may be distinguished from components analysis by altering (3a) to

$$| \ SS_T - \nu \ U^2 \ | = 0 \tag{3b}$$

where $U^2$ denotes uniqueness variance. In this case it is customary to reduce $SS_T$ to correlations and $U^2$ correspondingly. The various forms of factor and components analysis are usually capable of representation as eigenvalue problems and therefore have determinantal equations similar to (3a) or (3b) with only minor alterations.[1]

Rule 5. Every statement will begin with one of the acronyms ANOVA, CORREL or FACTOR to denote which linear model solution is required. A colon will follow the acronym.

The following simple rules are useful:

Rule 6. An equal sign will be used to separate hypothesis parameters from error parameters of an analysis. Hypothesis terms will be to the left, error terms to the right of the equal signs.

Rule 7. Commas will separate different tests in a model statement and a period will end the statement.

Rule 8. The letter W will designate the constant term of the model or grand mean of the data when used singly as a hypothesis term.

---

[1]Since the eigenvector's other basis of (3a) and (3b) form a basis for the vector spaces involved, any of the space must be a linear combination of the eigenvectors. Thus, any factor analytic solution must be expressible in terms of rotated eigenvectors of either (3a) or (3b).

Rule 9. The number 0 when used in conjunction with either a hypothesis term or an error term will designate a set of parameters included in a model but not tested.

Rule 10. The statistical estimate of errors is assumed to be the residual variance of the error variables after all hypothesis effects have been removed, unless otherwise indicated.

With these 10 rules it is possible to describe symbolically many of the common orthogonal designs.

Example (1) Factorial analysis of variance, multivariate or univariate, two factors.

ANOVA:W=0,A=V,B=V,AB=V.

W=0 indicates that the effect of the grand mean of the data is to be extracted but not tested. A=V indicates a test of the variables V over the main effect A. B=V indicates a test for the main effect B. AB=V indicates a test of the interaction effect.

Example (2) Canonical, multiple or product-moment correlation within a single sample.

CORREL:W=0,V1=V2.

The effect of the variable means is extracted but not tested. The variables in V1 are to be correlated with those in V2 and the correlation tested using V2 for errors.

Example (3) Factorial analysis of variance with partitioning in the A factor which is carried into the interaction.

ANOVA:W=0,A1=V,A2=V,B=V,A1B=V,A2B=V.

Example (4) Randomized Blocks (Snedecor, 1956, p. 300 <u>et seq</u>.).

ANOVA:W=0,B=0,T=V.

The effect B=0 is the effect of blocks (normally not tested) and T=V is the effect of treatments within blocks, pooled across blocks. Alternative representations of this design are available.

Example (5) Subjects by Treatments designs.

ANOVA:W=0,S=0,T=V.

S denotes the Subject effect, T the Treatment effect which is tested. Alternative representations of this design are available.

Example (6) Correlation from multiple samples.

CORREL:W=0,S=0,V1=V2.

The effect parameters S are included to eliminate the differences among sample means which would otherwise confound the test of the correlation between V1 and V2 (unfortunately this is not common practice).

Example (7) Factor analysis from a single sample.

FACTOR:W=0,0=V.

The variables V are the ones to be analyzed. Specification of type of factor analysis will not be discussed here.

Example (8) Factor analysis from multiple samples.

FACTOR:W=0,A=0,0=V.

The effect A is included to eliminate the differences among sample means which might otherwise give rise to spurious and unidentifiable factors (unfortunately this is not common practice).

<u>Rule 11</u>. Analysis of covariance is denoted in the error term by a slash, / , followed by an upper case Roman letter to denote which parameters are to be used as covariates. When the covariates are the same for all tests in the model the designation may follow the acronym and precede the colon.

Example (9) Analysis of covariance in a factorial design.

ANOVA:W=0,A=V1/V2,B=V1/V2,AB=V1/V2, or

ANOVA/V2:W=0,A=V1,B=V1,AB=V1.

The parameters V2 are the covariates. The attendant regression analysis may be described as a correlation model.

CORREL:W=0,A=0,B=0,AB=0,V2=V1.

Example (10) Partial correlation (product moment, multiple or canonical correlation).

CORREL:W=0,V3=V1/V2.

V2 are the covariates for the test of the correlation between V1 and V3.

Example (11) Partial correlation from multiple samples.

CORREL:W=0,A=0,V3=V1/V2.

Removal of the sample means from criteria (hypothesis variates), predictors (error variates) and covariates is accomplished by including dummy parameters A for the sample means as in examples 6 and 8 above.

Example (12) Homogeneity of regression in a factorial design ignoring other tests.

ANOVA:W=0,A=0,B=0,AB=0,V2=0,AV2=V1,BV2=V1,ABV2=V1.

The parameters AV2 are a Kronecker product of V2 and the parameters of A and test the homogeneity of regression over the A effect, say rows, of the design; BV2 tests homogeneity over the B effect, say columns; and ABV2 tests homogeneity in the AB interaction. Example 13 is more detailed.

Example (13) A complete study of analysis of covariance (Snedecor, 1956, p. 394 *et seq.*).

Let A be the main effect (states), V1 the variate and V2 the covariate.

ANOVA:W=0,A=0,V2=0,AV2=V1.

The error term for this homogeneity of regression problem is the residual after fitting individual regressions for each sample of A. W=0 and A=0 indicate parameters of an analysis of variance which are to be extracted independently of tests of the covariate regressions. V2=0 indicates that the "common" (customarily the "error") regression effect is to be eliminated before testing for homogeneity of regression. AV2=V1 test the homogeneity of regression effects over the factor A by obtaining the differences between regressions for each sample after the common regression has been removed.

This example shows the similarity of covariance analysis to analysis of variance by way of treating V2 as a factor of the design and AV2 as an interaction effect.

ANOVA:W=0,A=V1/V2,A=V2.

The test A=V1/V2 is the standard analysis of covariance. This model includes A=V2 as a check of the homogeneity of covariate means over the effect A. (Snedecor does not include this but it is suggested by many authors.) The test A=V1/V2 is the standard analysis of covariance. The error term here is V1 residual to A and the common regression of V2 on V1 assuming the test AV2=V1 to be null.

The "common" regression analysis, assuming homogeneity of regression, is formed as in example 10 and pools the individual regression coefficients of the samples.

## Method of Solution and Nonorthogonal Data

Two types of solutions exist for linear models. These might be termed the part solution and the partial solution as in correlation theory. The distinction between the two models arises when there are two sets of hypothesis parameters in a nonorthogonal analysis of variance to be tested separately against the same set of error parameters.

To distinguish between the two solutions let us consider the model: A=V, B=V. The partial solution tests B against V with the effects of A removed from both B and V, then tests A against V with the effects of B removed from both A and V. The part solution considers the order of the hypotheses and first tests B against V with the effects of A removed from both B and V, then tests A against V with B removed from V only. If A and B are uncorrelated, as the effect parameters of analysis of variance are with orthogonal data, the two models are identical because A and B are uncorrelated. When A and B are correlated, as in correlation or nonorthogonal analysis of variance, the test of B is the same for both solutions but the test of A is confounded with the test of B in the part solution since the variance of B which is correlated with the variance of A is pooled with the variance of A. Without describing the pros and cons of the two solutions it must be noted that if the test of B is significant in the part solution, no "exact" test of A can be made since A is confounded with B.

In computer programs it is desirable to use the part solution rather than the partial solution because the partial solution can be obtained from the part solution by simply reordering the variables as follows. Suppose we are forced to use the part solution for example 1.

ANOVA:W=O,A=V,B=V,AB=V.

The only test which has the desired partial is AB=V. To obtain all the tests as partials we should use all the models

ANOVA:W=O,A=O,B=O,AB=V,

ANOVA:W=O,A=O,AB=O,B=V,

and ANOVA:W=O,B=O,AB=O,A=V.

For a more formal discussion of the part method see Bock (1963). The partial method is discussed by Yates (1933).

When computer programs use the partial solution, it is impossible to obtain a part solution. Therefore, in this logic the part solution will be the rule.

Rule 12. The solution of the model is by the part method. The hypothesis of any test will be taken residual to all hypotheses to the left of it and will possibly be confounded with all hypotheses to the right of it. All errors are residual to all hypotheses and confounded with each other. The order of testing proceeds from left to right.

This rule allows some alternative statements for various models. For instance, a one-way analysis of covariance can be stated in various ways. Proof of the identity is presented in a later section.

ANOVA/V2:W=0,A=V1.

ANOVA:W=0,A=V1/V2.

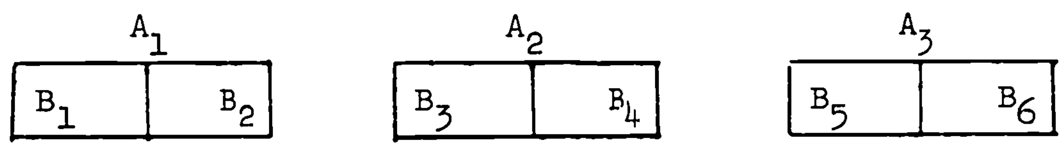
ANOVA:W=0,V2=0,A=V1.

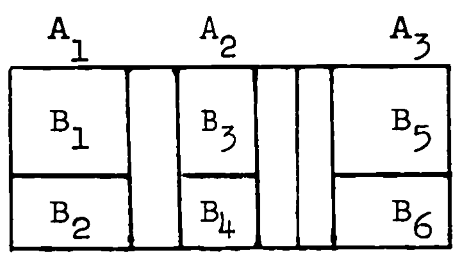
ANOVA:V2=0,W=0,A=V1.

With this rule, the usual orthogonality restrictions on examples 1 to 13 can be lifted without loss of meaning.

## Nested Designs

A nested design can also be viewed as a special kind of factorial. Scheffé (1959, p. 179) does this with some ease. If one views the primary sampling as the rows of a factorial design, then subsampling of the $i^{th}$ primary sample j times produces j subsamples as the column entries of the $i^{th}$ row. For the simple case of three primary samples with two subsamples each, the following two diagrams may be drawn to illustrate nesting. The first is a common method of diagramming nested designs, the second a variation which shows the similarity to factorials more clearly.




Statistical testing involves testing the differences of means between $B_1$ and $B_2$, the differences in means between $B_3$ and $B_4$, the differences in means between $B_5$ and $B_6$ and the differences among the means for $A_1$,$A_2$, and $A_3$ primary samples. Using the common notation of W to denote "within" we can promulgate the following rule.

Rule 13. Nesting of several subsamples, say B, within a larger sample A shall be denoted with an upper case Roman W as in BWA. When several samples are involved in A the notation for the B effect within sample $A_3$ will be denoted BWA3.

Example (14) The nesting design cited above.

ANOVA:W=0,BWA1=V,BWA2=V,BWA3=V,A=V.

Pooling of Parameters

Pooling effects is a convenience to have in some analytic problems. For instance, in some complex factorial designs, some statisticians pool high order interactions and test them collectively rather than singly. Such procedures may be allowed for here as follows.

Rule 14. The parameters for two effects may be pooled and tested collectively by placing a + (plus sign) between them. In a given level of nesting, when all nested effects are to be pooled before testing, the notation may be shortened by dropping the integers on the levels of the effect in which the nesting occurs.

Example (15) A four factor analysis of variance pooling third and fourth order interactions.

ANOVA:W=0,A=V,B=V,C=V,D=V,AB=V,AC=V,AD=V,BC=V,BD=V,CD=V,
ABC+ABD+ACD+BCD+ABCD=V.

Example (16) A nested design with two primary samples and subsamples nested in each (Winer, 1962, pp. 184-185) (fixed effects).

ANOVA:W=0,BWA1=V,BWA2=V,A=V (B effects not pooled),

or ANOVA:W=0,BWA1+BWA2=V,A=V (B effects pooled),

or ANOVA:W=0,BWA=V,A=V (short version equivalent to B effects pooled).

Example (17) A complicated nesting design, partially factorial.

This is the same design as Example 16 with each subsample divided into two categories, C, which are consistent over samples and subsamples (fixed effects) (Winer, 1962, p. 186).

ANOVA:W=0,BWA=V,C=V,A=V,BCWA=V.

Example (18) Pooling of error variables.

ANOVA:W=0,A=V1,B=V1+V2,AB=V1.

Components of Variance Models

There are two assumptions implicitly in the above rules which restrict the application of this logic to factorial designs. They are: (a) the error term for a test is the residual sum of squares remaining after all hypothesis effects have been removed from the data, and (b) the statement of the hypothesis in say A=V assumes that a hypothesis sum of squares is generated for the regression of V onto A.

These assumptions make it difficult to express a components of variance model in the logic. To include components of variance models in the logic it is necessary to invent a notation for generating a regression sum of squares as the error term of an analysis of variance.

Rule 15. When an error term is to be obtained as a sum of squares due to regression, an asterisk, *, will be used to separate the dummy parameters and the continuous variables which denote the regression.

Example (19) A subjects by treatments analysis.

ANOVA:W=0,S=0,T=ST*V.

This is an alternative representation of Example 5. In Example 5 the error term for the treatment effect was obtained as a residual after eliminating the grand mean, subjects and treatments effect. Here the error term is generated directly as an interaction sum of squares. This interaction sum of squares as a hypothesis would have been generated from a statement ST=V.

Example (20) A complete three factor components of variance model (Kempthorne, 1952, p. 110 et seq.). Only the tests of second order interactions are shown assuming the third order interaction to be the appropriate error term.

ANOVA:AB=ABC*V,AC=ABC*V,BC=ABC*V.

Note that it is not necessary to eliminate the grand mean or the main effects from the model since the error term is generated directly and not obtained as a residual.

Assuming that the second order interactions test null we can use other choice of error terms for various main effects.

ANOVA:A=ABC+AB*V,B=ABC+AB+BC*V,C=ABC*V.

Whether these choices are rational or not is irrelevant to the notation.

Example (21) A nested design with four levels of nesting and replication (Kempthorne, 1952, p. 107). Equal numbers of subsamples are assumed within each sample.

ANOVA:W=0,A=BWA*V,B=CWAB*V,C=DWABC*V,D=V.

Example (22) A nested design with two levels of nesting and unequal numbers of subsamples. Let A be two primary samples; A1 having two subsamples, A2 having four subsamples. To do this it will be necessary to treat the design as a full 2 x 4 factorial with some missing cells $A_1$ $B_3$ and $A_1$ $B_4$. The B effect will be partitioned: B1 will have one parameter and B2 will have two.

ANOVA:W=0,A=B1WA1+B1WA2+B2WA2*V,B1WA1=V,B1+B2WA2=V.

Shortened notation can be

ANOVA:W=0,A=B1WA1+BWA2*V,B1WA1=V,BWA2=V.

Extension

The notion of extension in factor analysis has been generalized to other forms of the linear model by Hall (1969). Briefly, the technique explores the relationship of an external variable to the canonical variables of a significance test.

Rule 16. When a significance test is to be extended to variables not involved in the test, an ampersand, &, will be used to separate the error variables from the variables of the extension.

Example (23) A multiple group discriminant analysis extended to a set of variables, V2.

ANOVA:W=0,A=V1&V2.

Some Miscellaneous Examples

These examples are included for whatever interest the reader may have in them.

Example (24) Two factor experiment with repeated measures on one factor (Winer, 1962, p. 302 et seq.). Two groups of subjects, S, are subjected to two treatments $A_1$ and $A_2$. Each group is given three successive treatments $B_1$, $B_2$ and $B_3$.

ANOVA:W=0,A=SWA,B=SWAB,AB=SWAB.

Example (25) Three schools are participating in an experiment on teaching fourth grade arithmetic. Each school has four classes, two teachers in method 1 and two teachers in method 2. Differences concerned method, teacher and school. All effects are fixed; all students and teachers assigned at random within schools. Third grade arithmetic scores are assigned as covariates.

(a) Check of whether covariate scores (V2) are randomly distributed among teacher (T), method (M) and school (S).

ANOVA:W=0,S=V2,TWMS=V2,MWS=V2.

The effect S=V2 would likely be significant and will be taken to be so here.

(b) Check the homogeneity of regression of third grade scores against final fourth grade scores within teacher and method for each school.

ANOVA:W=0,S=0,TWMS=0,MWS=0,V2=0,TV2WMS1=V1,TV2WMS2=V1,TV2WMS3=V1,MV2WS1=V1, MV2WS2=V1,MV2WS3=V1.

It will be assumed that all these tests will be null.

(c) Check for homogeneity of regression among schools.

ANOVA:W=0,S=0,TWS=0,MWS=0,TMWS=0,V2=0,SV2=V1.

If this test is significant, the experiment must go to an analysis of gains of the form V1-βV2 where the β is different for each school. The model follows from example of analysis of covariance. Where V2WS removes the separate school regressions, the appropriate statement is:

ANOVA:W=0,S=0,V2WS=V1,TWMS=V1,MWS=V1.

MWS has three degrees of freedom; if this is not significant there is no effect due to method. If it is significant it is entirely possible that the significant differences occurred in only one school, but this has not been examined. The differences in separate schools can be discovered as follows.

ANOVA:W+S+V2WS+TWMS=0,MWSL=VL,MWS2=V1,MWS3=V1.

## The Rules as an Operator Calculus on a Set of Vectors

Consider a set, $\theta$, of K vectors (the variables) in an N dimensional (the observations) vector space; $K \leq N$.

Notation. Subspaces of $\theta$ may be designated by letters of the alphabet, not W. Subsets of subsets are designated by a letter followed by an indexing digit (examples: A, B, A1, A2, etc.). The null space will be noted as $\Phi$ (zero); W used alone designates the unit vector.

Operator 1. Direct Product. The direct product of two vectors is a vector whose elements are the pairwise products of corresponding elements of the two vectors. The direct product of two subsets of vectors consists of all possible direct products of two vectors one from each of the subsets. This will be indicated by juxtaposing the two subspace notations without intervening symbols, i.e., AB, A1B, ABC (generating interaction parameters from main effects. This is an alternative to the Kronecker product).

## Some Algebraic Properties of the Notation

One of the basic properties of mathematics is that it provides a simple notation for representing a complex event. One of the advantages of mathematics is that the simplified notation can be manipulated to express relationships among complex events when these relationships are too complicated to be observed directly. For example, a sheepherder can merge two counted flocks of sheep and know the size of the resulting flock without counting it: the mathematics of addition relieves him of the necessity of counting the merged flock.

If this use of mathematics is accepted, the present notation scheme can be manipulated to show certain properties of linear models. For example, it is possible to show that some of the model statements made above are identical without going through the labor of calculating the model to find identical results. Consider the following two theorems and a lemma about identities.

Lemma: A given model statement produces a unique ordered series of hypotheses as subspaces of a vector space, when the hypotheses are correlated.

Proof: [The proof here is very simple and depends on the method of solution stated in Rule 12.]

Suppose we are given a linear model statement with hypotheses parameters $\alpha,\beta,\ldots\gamma$ and error parameters $\rho,\sigma,\ldots\tau$. Rule 12 says that the hypothesis of any test will be taken residual to all hypotheses to the left of it and (possibly) confounded with any hypothesis to the right of it. This makes it possible to write the hypotheses of a model statement as an ordered list of subspaces of a vector space: $\alpha,\beta,\ldots\gamma$ .

It is apparent on the face of it that $\beta,\alpha,\ldots\gamma$ is a different statement from $\alpha,\beta,\ldots\gamma$ when $\alpha$ and $\beta$ are correlated since in $\alpha,\beta,\ldots\gamma$ the hypothesis $\alpha$ is confounded with $\beta$ in the solution, but in $\beta,\alpha,\ldots\gamma$ the hypothesis $\alpha$ will be residual to $\beta$ in the solution, Q.E.D.

Example (26) Consider the alternative models

(a) ANOVA:W=0,A=V,B=V,AB=V

and

(b) ANOVA:W=0,B=V,A=V,AB=V.

If the design is not orthogonal (i.e., all cell frequencies are not the same), the A pseudovariables are correlated with the B pseudovariables.

If we follow the method of the lemma, these give rise to two ordered series of hypothesis subspaces (a) W, A, B, AB and (b) W, B, A, AB and are obviously not the same model.

Theorem: The following statements of a one-way analysis of covariance produce identical analyses:

(a) ANOVA:W=0,A=V1/V2.

(b) ANOVA:W=0,V2=0,A=V1.

(c) ANOVA:V2=0,W=0,A=V1.

(d) ANOVA/V2:W=0,V1=V2.

Proof: Statement (d) is equivalent to (a) by definition in Rule 11.

Statement (a) gives rise to an ordered series of hypotheses[2]

(a') W, V2, A; V1

since the analysis of covariance is a partial correlation technique with both A and V1 residual to the covariate.

Statement (b) gives rise to the ordered series of hypotheses

(b') W, V2, A; V1

which is the same as (a') above and therefore produces an analysis identical to (a) and (d).

Statement (c) gives rise to the ordered series of hypotheses

(c') V2, W, A; V1.

Since W and V2 are hypothesis statements to the left of A and V1, A and V1 are residual to them regardless of what order they have between them.

Since A=V1 is the only test being made (c) is equivalent to (b) and also (a) and (d). Q.E.D.

Theorem: The following statements about a subjects by treatments design are equivalent:

(a) ANOVA:W=0,S=0,T=V.

(b) ANOVA:T=ST*V.

Proof: [Subjects by treatments designs are customarily designs where a number of subjects are given a fixed number of treatments, all in the same sequence.]

Consider V to be univariate. If the number of subjects is s and the number of treatments is t, then there are $s \cdot t = 1 + (s-1) + (t-1) + (s-1)\cdot(t-1)$ observations on subjects and degrees of freedom in the design.

---

[2]Using a semicolon to separate hypotheses from errors.

Statement (a) gives rise to the ordered series of parameters W, S, T;V where V is residual to S, T/ and W. Since W has one parameter, S has s-1 and T has t-1 parameters, (s-1)·(t-1) degrees of freedom are residual in V.

By Rule 15 one can see that ST of statement (b) is a set of (s-1)·(t-1) dummy parameters which can be handled as if they were a hypothesis set. ST*V will generate a sum of squares with (s-1)·(t-1) degrees of freedom.

It is customary to use contrasts for S and T which are orthogonal to W and to each other. Thus (s-1)·(t-1) parameters of ST are also independent of W, S and T and, trivially, residual to them. From this we conclude that the sum of squares for ST*V has (s-1)·(t-1) parameters. Therefore we claim that V of statement (a) is identical to ST*V of statement (b).

Now: in statement A, since T is independent of W and S the parameters T in (a) are identical to those of T in (b).

Therefore both hypothesis tests give the same results. Q.E.D.

In addition to identity relationships between model statements there are other types of simple relationships among models. Consider the following definition.

Definition: Two models are said to be concordant if there exists an isomorphism between the subspaces of each.

This requires that there be a one to one correspondence between: (1) the ordered series of hypothesis subspaces of the two model statements and (2) the errors subspaces of the two model statements, in such a way as to preserve the mathematical relations between hypotheses and errors within both statements. This does not require that the vectors of the subspaces be related in any way.

For example, the two model statements

ANOVA:W=0,A=V,B=V,AB=V. (a)

and

CORREL:W=0,V2=V1,V3=V1,V3=V1,V2V3=V1 (b)

are concordant if we set up the correspondence

model (a) $W \leftrightarrow W$ model (b)
$A \leftrightarrow V2$
$B \leftrightarrow V3$
$V \leftrightarrow V1$

This correspondence must also extend to the mathematical operations of testing and Kronecker product. These models satisfy the definition of concordance because

model (a) $AB \leftrightarrow V2V3$ model (b)
$W{=}C \leftrightarrow W{=}0$
$A{=}V \leftrightarrow V2{=}V1$
$B{=}V \leftrightarrow V3{=}V1$
and $AB{=}V \leftrightarrow V2V3{=}V1$

In this example the ANOVA model is a two-factor factorial analysis of variance, the CORREL model is a study of response surfaces.

The model

ANOVA:W=0, A=V1, V2=V1, AV2=V1 (c)

is also in this concordance class since (c) we can set up the correspondence

model (a) $W \leftrightarrow W$ model (c)
$A \leftrightarrow A$
$B \leftrightarrow V2$
$V \leftrightarrow V1$

Discussion

It should be noted that, throughout the description of the notational scheme and its logical properties, there is no discussion of either the distributional form of the data or whether the models have any exact test in any distribution. Thus, it is up to the user to state the nature of data distributions and to determine whether an exact test is available. For instance, while components of variance models can be stated for nonorthogonal models there is no assurance that a solution exists which can be tested as an exact test.

A similar but less complex scheme was produced by Cramer (see Clyde *et al.*, 1966) for use with a program to compute multivariate analysis of variance. The instigation to develop this system came from Cramer's work and the need for a control mechanism for a similar type of computer program for linear models.

References

Beaton, A. E. (1971) F4STAT Manual (in process). Princeton, N. J.: Educational Testing Service.

Bock, R. D. (1963) Programming univariate and multivariate analysis of variance. Technometrics, 5: 95-117.

Clyde, O. J., Cramer, E. M., & Sherin, R. J. (1966) Multivariate Statistical Programs. Coral Gables, Florida: The University of Miami Biometric Laboratory.

Finn, J. D. (1967) Multivariance: Fortran Program for Univariate and Multivariate Analysis of Variance. Buffalo, New York: State University of New York at Buffalo, Department of Educational Psychology.

Hall, C. E. (1969) Extension of Discriminant Analysis and Multivariate Analysis of Variance. In Proceedings of the 76th Annual Convention of the American Psychological Association. Washington, D. C.: The American Psychological Association.

Kempthorne, O. (1952) The Design and Analysis of Experiments. New York: Wiley.

Scheffé, H. (1959) The Analysis of Variance. New York: Wiley.

Snedecor, G. W. (1956) Statistical Methods, 5th edition. Ames, Iowa: The Iowa State College Press.

Vartak, M. N. (1955) On an application of Kronecker product of matrices to statistical designs. Annals of Mathematical Statistics, 26: 420-438.

Wilks, S. S. (1932) Certain generalizations in the analysis of variance. Biometrika, 24: 471-494.

Wilks, S. S. (1938) The analysis of variance and covariance in nonorthogonal data. Metron, 13: 141-154.

Winer, B. J. (1962) Statistical Principles in Experimental Design. New York: McGraw-Hill.

Yates, F. (1933) Confounding in replicated experiments the principles of orthogonality. Journal of Agricultural Science, 23: 108-145.